Video Object Segmentation based on Adaptive Background and Wronskian Change Detection Model

Deepak Kumar Panda
Dept. of Electronics and Communication Engg.
National Institute of Technology Rourkela
Rourkela 769008, India
Email: deepakkumar.panda@gmail.com

Sukadev Meher
Dept. of Electronics and Communication Engg.
National Institute of Technology Rourkela
Rourkela 769008, India
Email: smeher@nitrkl.ac.in

***Abstract*—In computer vision, detection of moving objects from a complex video scene is an important and challenging problem. It finds application in many computer vision and artificial intelligent systems. Background subtraction is a very popular and powerful technique in computer vision for moving object detection in the presence of stationary camera. In the proposed scheme, Wronskian change detection model (WM) is used to find out the change between the constructed background and the incoming video frame. In this paper we have used WM in the Gaussian distribution for video object segmentation. We have presented a new equation for variance updation in the neighbourhood. The parameters of Gaussian (i.e., the mean and the variance) are updated for linearly dependent pixels using a Gaussian weight learning rate in the neigbourhood. The result of the proposed scheme is found to provide accurate silhouette of moving objects in presence of illumination variation and un-stationary backgrounds like fountain, ocean, curtain and Train. We compare our method with other modelling techniques and report experimental results.**

***Index Terms*—Motion detection, background subtraction, Wronskian change detection, single Gaussian, illumination invariant.**

I. Introduction

Visual surveillance is an important application area of research in computer vision, which includes object detection, classification, person identification, tracking, and activity recognition steps for careful monitoring. All of these procedure starts with moving object detection as the initial step. Hence moving object detection algorithm should produce fewer false alarms and be real-time with less memory usage. A very popular and powerful method for moving object detection in presence of static camera is background subtraction (BGS). In comparison to other approaches such as optical flow, and frame differencing, BGS is computationally efficient and gives less false alarms for real-time surveillance. A basic requirement of BGS is modelling background and then comparing each incoming frame with the constructed background so as to detect changes. A threshold is selected for classifying pixel belonging to background or foreground. The threshold is judiciously chosen so as to minimize false positive as well as false negatives. The performance of BGS depends mainly on background modelling. In presence of complex dynamic scene BGS techniques have to deal with practical problems such as illumination variation, un-stationary background, noise, camera jitter, camouflage, shadows, relocation of background objects, stopping of moving object, slow and fast object moving together, initialization with moving objects, etc. In the due course of time, the background needs to be periodically updated to keep a check on the various problems engulfed with BGS.

The paper is organized as follows. In Section II, we present an extensive overview of the existing approaches adopted for background subtraction. Here in this we have presented the pros and cons of each of the algorithm. We have implemented some of the algorithms in order to compare them with our method. Section III gives a detailed description of Wronskian change detection model (WM). Our proposed method is well illustrated in Section IV. In Section V, we present a comparison of our proposed scheme with that of the state-of-the-art techniques in terms of visual as well as quantitative measures. Section VI concludes the work done and give future research directions.

II. Review Of Background Subtraction Algorithms

In the last decade, a large number of researchers have proposed background subtraction techniques for handling illumination variation, un-stationary background, etc. Manzanera and Richefeu [1] proposed $\Sigma - \Delta$ estimation (SDE) based on non-linear recursive approximation for background modelling by incrementing the value by one if background pixel is smaller than image pixel, or decremented by one if background is greater than the current image pixel. The background estimated by this method is an approximation of the median of values from history of image frames. At every frame, the pixel of the background model is changed by one. So background constructed by SDE will not be adaptive for fast changing background scene. The W^4 system [2] proposed by Haritaoglu et al. models the background in gray scale by calculating minimum, maximum intensity values, and the maximum intensity difference between consecutive frames from the training frames. However, the background model fails in the presence of un-stationary background pixels, as threshold for changed detection is decided by the median of maximum absolute difference between pixels of consecutive frames. Wren et al. proposed Pfinder [3] for people segmentation, tracking and behavior understanding. It is based on the assumption that the intensity values of a pixel can be modelled by a single Gaussian (SG) distribution. Pfinder works well in indoor environment, can deal with small or gradual changes in the background and illumination variation but fails in the outdoor scene, when the background scene has multi-modal distributions. To overcome the problem of multi-modal background, Stauffer and Grimson [4] modeled each pixel intensity by a Gaussian Mixture Model (GMM) distributions. Durucan and Ebrahimi [5] exploited the concept of linear dependence and linear independence to find out the change. The Wronskian determinant is used to find out the change between the constructed background and the incoming video frame. The background image is chosen as frame which is devoid of any foreground object. The WM provides robust result in presence of noise and illumination variation. It can also detect moving object in presence of environmental condition such as changing cloud. However the method does not provide an adaptive threshold for WM. The algorithm relies on choosing the best reference image and threshold for every video sequence. Badri et al. [6] improved the algorithm [5] by integrating single Gaussian and WM. In the background construction, the authors have used linearly dependent pixel and during the updation of the background model they have considered linearly independent pixel. Thus there is a clear confusion regarding the choice of linear dependent or linear independent pixels for background modeling. The threshold for WM is made adaptive by using statistical standard deviation. However, the algorithm is not computationally efficient. The algorithm uses the total count of linearly dependent pixels at a location in calculating the background and the standard deviation. In calculating the standard deviation for adaptive threshold, it store the pixel values of current frame which are linearly dependent with the background model. Storing linearly dependent pixel till the current instant of time is not computationally justifiable for real-time surveillance.

The method proposed in this paper uses WM in the Gaussian distribution for video object segmentation. We have presented a new equation for variance updation in the neighbourhood. The parameters of Gaussian (i.e., the mean and the variance) are updated for linearly dependent pixels using a Gaussian weight learning rate in the neigbourhood. The steps for adaptive threshold is simple and yet effective as compared to the steps presented in [6] and the confusion regarding the updation of background model is avoided by updating the background pixels only for linearly dependent pixel. The result of the proposed scheme is found to provide accurate silhouette of moving objects in presence of illumination variation and un-stationary backgrounds such as fountain, ocean, curtain and Train.

III. Wronskian Change Detection Model

A simple and exhaustive test for determining the linear dependence or independence of vectors is the Wronskian function.

In general, for a real or complex valued functions $f_1, f_2, ..., f_n$ which are $n-1$ times differentiable on an interval $I_\dagger$, the Wronskian $W(f_1, f_2, ..., f_n)$ is defined by:

$$W = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix},\ x \ \epsilon\ I_\dagger \tag{1}$$

if $W(f_1, f_2, ..., f_n) \neq 0$ then $f_1, f_2, ..., f_n$ are independent, else $f_1, f_2, ..., f_n$ are dependent.

A set of functions $f_1(x), f_2(x), ..., f_n(x)$ is linearly dependent on an interval $I_\dagger$, if there exists constant $c_1, c_2, ..., c_n$ not all zeros such that

$$c_1 f_1(x) + c_2 f_2(x) + ... + c_n f_n(x) = 0 \tag{2}$$

Wronskian change detection model (WM) was initially proposed by Durucan and Ebrahimi [5]. WM is based on a vector model of images to find out the change between current frame and the reference background. Let the two components $\mu(E)$ and $I(E)$ represent the reference background and the current frame of the video. $\mu(E)$ and $I(E)$ are function of illuminance E and their derivative is given as $\mu' = d\mu/dE$ and $I' = dI/dE$.

The linear combination of these function can be represented in the from of eq (2)

$$c_1\mu + c_2 I = 0 \tag{3}$$

Solving eq (2) for c_2 yields

$$\frac{\mu}{I}c_1 + 1 \cdot c_2 = 0 \tag{4}$$

and the Wronskian for two linearly dependent function is

$$W = \begin{vmatrix} \frac{\mu}{I} & 1 \\ (\frac{\mu}{I})' & 0 \end{vmatrix} = -\left(\frac{\mu}{I}\right)' = -\frac{\mu'}{I} - \mu\left(\frac{1}{I}\right)' \tag{5}$$

The determinant of the Wronskian matrix yields zero for linearly dependent function

$$0 = -\frac{\mu'}{I} - \mu\left(\frac{1}{I}\right)' = \frac{\mu}{I^2} - \left(\frac{1}{I}\right) \tag{6}$$

$$= \frac{\mu^2}{I^2} - \left(\frac{\mu}{I}\right) \tag{7}$$

The change between the reference background μ and the incoming frames I for every pixel of a video can be calculated as:

$$W(x,y) = \left(\left(\frac{\mu(x,y)}{I_t(x,y)}\right)^2 - \frac{\mu(x,y)}{I_t(x,y)}\right) \tag{8}$$

WM exploits the spatial redundancy in a $m \times m$ neighbourhood and it represents each pixel in the frame in terms of a vector $\vec{\mu}$ and $\vec{I}$ of size $[m \times m, 1]$. The vector model for a pixel in the frame consists of a central pixel and its neighbouring pixel. In order to detect the change between the current frame and the reference background, a linear independence test is carried out between the two. A pixel is said to be changed pixel, if it satisfies linear independency test. A pixel at position (x,y) in the t th frame can be considered as a changed pixel, if W deviates from zero. In WM, each pixel in the center pixel is associated with its neighbouring pixel called as region of support for the corresponding vectors. The region of support can have various sizes, eg., (a) 3×3, (b) 5×5, (c) 7×7. The pixels are scanned from left to right in every row of $m \times m$ neighbourhood and are placed in the column vector of size $[m \times m, 1]$. The eq (8) can be changed to include the region of support for every pixel in the frame and is given in eq (9).

$$W(x,y) = \frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{\vec{\mu}^{i}_{t-1}(x,y)}{\vec{I}^{i}_{t}(x,y)}\right)^2 - \frac{\vec{\mu}^{i}_{t-1}(x,y)}{\vec{I}^{i}_{t}(x,y)}\right) \quad (9)$$

where n represents the number of pixel in the vector image. The factor $1/n$ is added to normalize the results to the vector dimensions so that the same threshold can be applied for different vector dimension.

Algorithm 1 Proposed background subtraction using Wronskian change detection model

$\alpha = c \cdot Gaussianweights$ \\ size of α is $n \times n$
$W(x,y) = \frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{\vec{\mu}^{i}_{t-1}(x,y)}{\vec{I}^{i}_{t}(x,y)}\right)^2 - \frac{\vec{\mu}^{i}_{t-1}(x,y)}{\vec{I}^{i}_{t}(x,y)}\right)$
if $W(x,y) \leq K\sigma_{t-1}(x,y)$ **then**
  $p_t(x,y) = 0$
  $j = 1$
  for $u = x - \lfloor m/2 \rfloor, \ldots, x + \lfloor m/2 \rfloor$ **do**
    $k = 1$
    for $v = y - \lfloor m/2 \rfloor, \ldots, y + \lfloor m/2 \rfloor$ **do**
      $\mu_t(u,v) = \alpha(j,k) I_t(u,v) + (1 - \alpha(j,k))\mu_{t-1}(u,v)$
      $\sigma^2_t(u,v) = \alpha(j,k)\left(\frac{1}{n}\sum_{i=1}^{n}\left(\vec{I} - \vec{\mu}\right)^2\right) + (1 - \alpha(j,k))\,\sigma_{t-1}$
      $k = k + 1$
    end for
    $j = j + 1$
  end for
else
  $p_t(x,y) = 1$
end if

IV. Proposed Method

The proposed scheme uses WM in Gaussian distribution for video object segmentation. The change between the constructed background and the incoming video frame is calculated using WM. A pixel is said to be changed pixel, if it satisfies linear independence test. A simple and exhaustive test for determining linear dependence/linear independence of two vectors (i.e., the background and the current image) is done using Wronskian function. If the value of Wronskian matrix W does not deviate much from zero, the pixel is considered as a background else if the deviation is larger than $K\sigma$, the pixel is considered as a foreground. The reference model of the background and parameter for the threshold are updated only when the pixel of incoming current frame are found to be linearly dependent with the background model. In our proposed scheme, we update the parameters (i.e., the mean and the variance) over a neighbourhood $m \times m$ using a Gaussian weight only for every linearly dependent pixels. This allows to take into account spatial relationships in a weighted manner among incoming pixel with its surrounding.

$$\alpha = c \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (10)$$

In WM, we have taken the ratio of background image to current image, as compared to Badri et al. [6] which uses the ratio of current image to the background. WM provides accurate shape detection of moving objects for the ratio of background image to current image in comparison to the ratio of current image to the background. Badri et al. [6] have used N initial frames to calculate the initial background and the variance. In our proposed scheme, the initial background and variance is calculated in the neigbourhood from the first frame of the video. In this paper, we assume that the spatial neighbourhood variation is similar to the temporal variation.

The proposed background subtraction is well illustrated in algorithm 1.

V. Experimental Results

The performance of the proposed algorithm is evaluated using several publicly available video sequences. Both indoor and outdoor scenes were considered for the analysis. The effectiveness of the proposed scheme is demonstrated on video sequences such as “MSA”[1], “Intelligent Room”, “Curtain”, “Water Surface”, “Ocean Waves”[2], “Fountain”, and “Train”. The sequence “Intelligent Room” is from CVRR Laboratory ATON project[3]. The video sequence namely “Curtain”, “Water Surface”, and “Fountain” are complex video sequence with unstationary background in the scene, are taken from I2R dataset [4]. The frame number used for testing of the proposed algorithm for different video dataset is shown in Table I. The values of constant K for different video sequence is given in Table II. The constant c is fixed at 0.001 across all experiments. The region size for the WM calculation and parameter updation is set as 3×3.

[1]http://cvprlab.uniparthenope.it
[2]http://imp.iis.sinica.edu.tw/ytchen/testvideos.rar
[3]http://cvrr.ucsd.edu/ aton/ shadow/index.html
[4]http://perception.i2r.a-star.edu.sg/bk_model

TABLE I
FRAME NUMBER CONSIDERED FOR VIDEO SEQUENCE

Video	MSA	IR	Curtain	WS	OW	Fountain	Train
Frame	1100	95	22774	1515	510	1430	70065

TABLE II
THE VALUES OF K FOR DIFFERENT VIDEO SEQUENCE

Sequence	MSA	IR	Curtain	WS	OW	Fountain	Train
K	0.4	0.5	0.5	0.5	0.4	0.3	1.5

A. Quantitative Evaluations

To validate the proposed scheme, results obtained by it are compared with those of manual thresholding-based background subtraction (SBS), $\Sigma - \Delta$ background subtraction (SDE) [1], W^4 background subtraction [2], Wronskian change detection scheme (WM) [5], Badri et al. [6], single Gaussian (SG) [3], GMM [4]. No pre-processing or post-processing operations have been applied in any of these algorithms to maintain the fairness in comparison.

1) Accuracy Metrics: For measuring accuracy, different metrics such as Recall, Precision, F_1, Similarity and PCC is calculated.

Recall, also known as detection rate. It is the ratio of detected true positives to the total number of pixels present in the ground truth

$$Recall = \frac{t_p}{t_p + f_n} \tag{11}$$

Precision, also known as positive prediction which gives the percentage of detected true positives as compared to the total number of pixels detected by the method

$$Precision = \frac{t_p}{t_p + f_p} \tag{12}$$

F_1 metric, also known as Figure of Merit or F-measure. It is the weighted harmonic mean of precision and Recall

$$F_1 = \frac{2 * Recall * Precision}{Recall + Precision} \tag{13}$$

pixel-based Similarity measure is defined as

$$Similarity = \frac{t_p}{t_p + f_n + f_p} \tag{14}$$

and finally, we consider percentage of correct classification (PCC), given by

$$PCC = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \times 100 \tag{15}$$

Here true positive (t_p) represents the number of pixels classified correctly as belonging to the foreground and true negative (t_n), which counts the number of background pixel classified correctly. The false positive (f_p) is the number of pixels that are incorrectly classified as foreground and false negatives (f_n) represents the number of pixels which are wrongly labelled as background but should have been classified as foreground.

2) Accuracy Results: Quantitative accuracy result is obtained from accuracy metrics, namely, $Recall$, $Precision$, $Similarity$, F_1, and PCC. All these metrics should give higher measure for accurate shape detection of moving objects. To measure the accuracy, we have considered manually segmented ground-truth images of moving objects. The results for the proposed method and other state-of-the-art BGS scheme are tested on wide variety of video sequences such as “MSA”, “Intelligent Room”, “Curtain”, “Water Surface”, “Ocean Waves”, “Fountain”, “Train”. The binary segmented output of the proposed scheme and other BGS technique are compared with their corresponding ground-truth images and the results are provided in Tables III, IV, V, VI, VII, VIII and IX. It is concluded from these tables that SBS and SDE works well in presence of stationary background and its performance degrades for un-stationary background. The performance of Badri et al. BGS [6] fails to detect when the foreground object is darker than the background. They have used the ratio of current image to the background, which gives a smaller value in the Wronskian calculation. The value of K reported in their paper is from 1 to 5. This value is not giving any detection and therefore we have used the value of K to be in range of 0.01 to 0.1. Our proposed algorithm works well in presence of both stationary and un-stationary background. It is observed from the complex dynamic scene of “Fountain” dataset that the PCC for the proposed scheme is increased to 98.46 (from 96.1 by SBS, from 95.77 by SDE, from 74.38 by W4, from 86.29 by WM, from 93.31 by SG and from 91.2 by GMM). Similarly, for “Train” dataset the PCC obtained from the proposed technique is increased to 98.75 (from 91.71 by SBS, from 87.7 by SDE, from 77.02 by W4, from 91.84 for WM, from 95 by SG and from 90.13 by GMM). The proposed research uses Wronskian framework for calculating the change between constructed background and the incoming video frame over a neighborhood. The parameters of Gaussian are updated in a weighted manner, which extracts accurate silhouette of moving objects as compared to other BGS scheme in “Fountain” and “Train” dataset.

B. Qualitative Evaluations

We observe that the proposed scheme provides robust performance for dynamic backgrounds than other considered BGS techniques. Experimental results demonstrate the effectiveness of the proposed method in providing a promising and accurate silhouette of moving object in presence of complex video sequence as shown in Fig. 1 .

VI. CONCLUSION

In this paper, we have presented a novel algorithm which is simple and efficient for moving object detection using Wronskian matrix in the Gaussian distribution. We have presented a new equation for variance updation in the neighbourhood. The parameters of Gaussian (i.e., the mean and the variance) are updated for linearly dependent pixels using a Gaussian weight learning rate in the neigbourhood. The algorithm takes the advantage of spatial redundancy in moving object detection.

TABLE III
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "MSA" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.7934	0.9329	0.8575	0.7506	99.58
SDE	0.9545	0.1421	0.2474	0.1411	90.82
W4	0.5628	0.6985	0.6234	0.4528	98.92
WM	0.8926	0.4834	0.6272	0.4569	98.32
SG	0.895	0.9872	0.9388	0.8848	99.82
Badri et al.	0.1843	0.0304	0.0522	0.0268	89.42
GMM	0.9198	0.7371	0.8184	0.6926	99.35
Proposed	0.7107	0.8465	0.7727	0.6296	99.34

TABLE IV
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "INTELLIGENT ROOM" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.8709	0.92	0.8948	0.8096	99.7
SDE	0.9884	0.163	0.2798	0.1627	92.54
W4	0.9038	0.0501	0.0949	0.0498	74.72
WM	0.9902	0.3259	0.4904	0.3249	96.98
SG	0.9679	0.8724	0.9177	0.8479	99.75
Badri et al.	0.0873	0.0091	0.0165	0.0083	84.83
GMM	0.9813	0.5686	0.72	0.5625	98.88
Proposed	0.9706	0.8847	0.9257	0.8617	99.77

TABLE V
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "CURTAIN" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.8323	0.9056	0.8674	0.7658	98.05
SDE	0.8897	0.3972	0.5492	0.3786	88.82
W4	0.9177	0.3715	0.5289	0.3596	87.49
WM	0.9662	0.7073	0.8167	0.6902	96.68
SG	0.7774	0.7539	0.7655	0.62	96.35
Badri et al.	0.3265	0.1426	0.1985	0.1102	79.81
GMM	0.9611	0.6282	0.7598	0.6126	95.35
Proposed	0.8769	0.9398	0.9073	0.8303	98.63

TABLE VI
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "WATER SURFACE" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.6842	0.9294	0.7882	0.6504	97.13
SDE	0.8659	0.6983	0.7731	0.6302	96.04
W4	0.7657	0.4508	0.5674	0.3961	90.9
WM	0.9424	0.6101	0.7407	0.5882	94.86
SG	0.7287	0.9596	0.8283	0.707	97.65
Badri et al.	0.2412	0.1174	0.1579	0.0858	79.96
GMM	0.8772	0.814	0.8444	0.7307	97.48
Proposed	0.9198	0.8886	0.9039	0.8247	98.48

TABLE VII
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "OCEAN WAVES" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.7623	0.7538	0.758	0.6103	97.95
SDE	0.8533	0.3456	0.492	0.3262	92.59
W4	0.905	0.4008	0.5556	0.3846	93.91
WM	0.7852	0.7829	0.784	0.6448	98.18
SG	0.8501	0.6428	0.7321	0.5774	97.38
Badri et al.	0.3993	0.1616	0.2301	0.13	88.75
GMM	0.8777	0.6103	0.72	0.5625	97.13
Proposed	0.7676	0.8955	0.8266	0.7045	98.65

TABLE VIII
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "FOUNTAIN" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.4639	0.3748	0.4147	0.2616	96.1
SDE	0.7016	0.3845	0.4968	0.3305	95.77
W4	0.7754	0.0847	0.1528	0.0827	74.38
WM	0.8213	0.1566	0.263	0.1514	86.29
SG	0.5836	0.2418	0.342	0.2063	93.31
Badri et al.	0.7574	0.0383	0.0729	0.0378	42.64
GMM	0.7902	0.2235	0.3484	0.2109	91.2
Proposed	0.6869	0.7716	0.7268	0.5708	98.46

TABLE IX
RECALL, PRECISION, F-MEASURE, SIMILARITY, AND PCC FOR "TRAIN" VIDEO SEQUENCE

Approach	**Recall**	**Precision**	**F-measure**	**Similarity**	**PCC**
SBS	0.8753	0.314	0.4622	0.3006	91.71
SDE	0.9601	0.2437	0.3887	0.2413	87.7
W4	0.8258	0.1312	0.2264	0.1277	77.02
WM	0.9095	0.3223	0.4759	0.3123	91.84
SG	0.8836	0.4429	0.59	0.4185	95
Badri et al.	0.2775	0.0625	0.102	0.0537	80.09
GMM	0.8985	0.2791	0.4259	0.2706	90.13
Proposed	0.8687	0.8313	0.8496	0.7386	98.75

The algorithm does not use a series of background images to construct the initial background rather the initial background and variance is calculated in the neigbourhood from the first frame of the video sequence. The results obtained by the proposed scheme are found to provide accurate shape detection of moving objects in complex video sequence. The proposed scheme for moving object detection is illumination invariant, removes shadows and reflection, and works well in presence of un-stationary background like fountain, ocean, curtain and Train. However, the algorithm fails to detect in presence of large swaying of trees. In our future research, this problem will be addressed.

REFERENCES

[1] A. Manzanera and J. C. Richefeu, "A new motion detection algorithm based on $\Sigma - \Delta$ background estimation," *Pattern Recognition Letters*, vol. 28, no. 3, pp. 320–328, Feb. 2007.

[2] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 809–830, 2000.

[3] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, Jul. 1997.

[4] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 747–757, 2000.

[5] E. Durucan and T. Ebrahimi, "Change detection and background extraction by linear algebra," *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1368–1381, Oct. 2001.

[6] B. N. Subudhi, S. Ghosh, and A. Ghosh, "Change detection for moving object segmentation with robust background construction under wronskian framework," *Machine Vision and Applications*, vol. 24, no. 4, pp. 795–809, 2013.

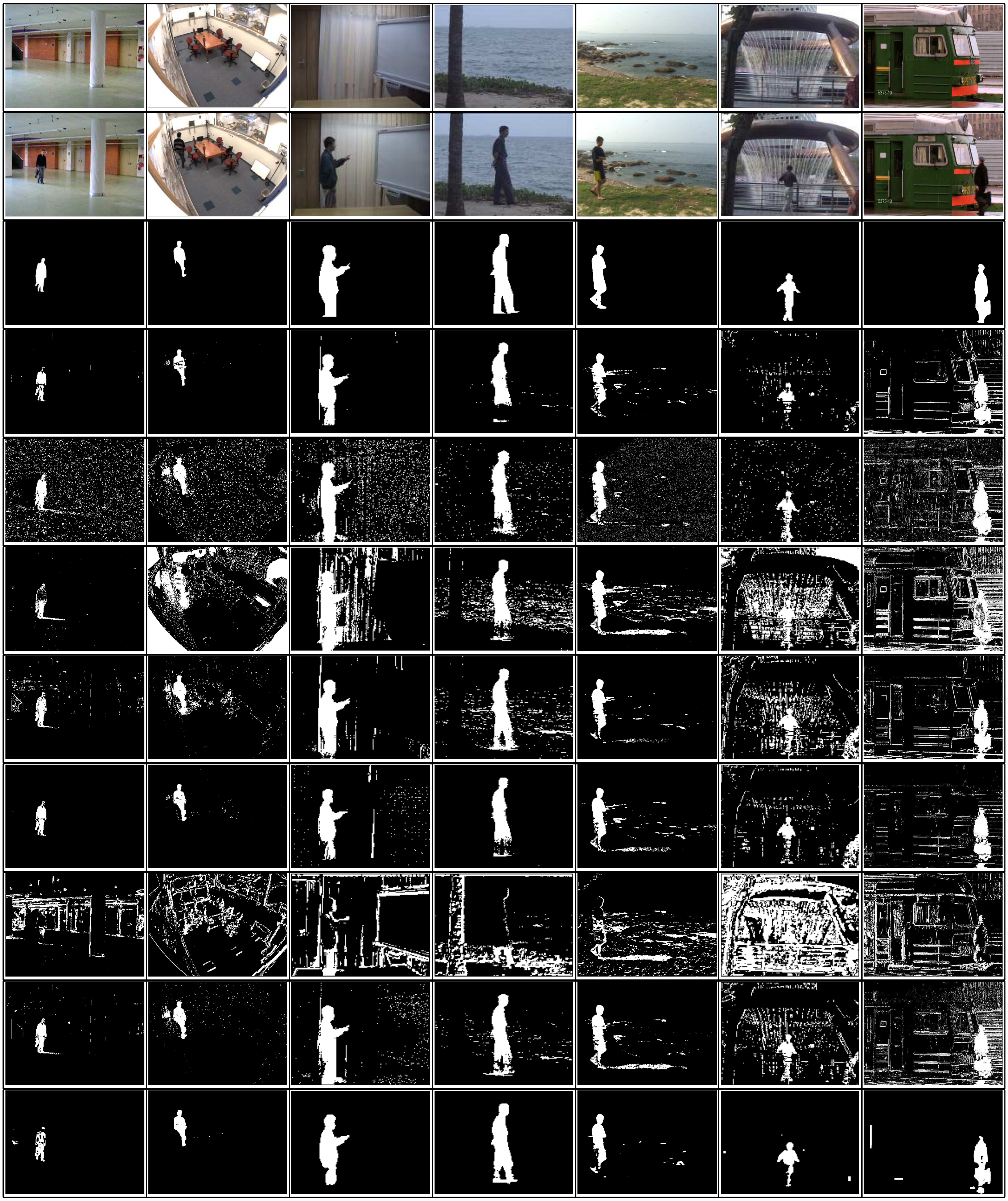

Fig. 1. Left to right: MSA, Intelligent Room, Curtain, Water Surface, Ocean Waves, Fountain, Train. Top to bottom: Original Image, Test Image, Ground truth, Moving object detection for SBS, SDE, W4, WM, SG, Badri et al., GMM, Proposed scheme