# Real Time Efficient Foreground Extraction with Video Processing

Mr. Pawan Arun Gawade
*Department of Electronics & Telecommunication*
S. V. C. E. T. Rajuri, Pune
pawangawade1008@gmail.com

Mr. Manoj Kumar
*Department of Electronics & Telecommunication*
S. V. C. E. T. Rajuri, Pune
manoj1985.111@rediffmail.com

Mr. P. Balaramudu
*Department of Electronics & Telecommunication*
S. V. C. E. T. Rajuri, Pune
balram9hossanna@gmail.com

***Abstract*— Video analysis is large source of unstructured information today. Introduction to HEVC standard has tremendous potential to lower the cost of data transfer and storage. HEVC provides forty percent or more than that of reduction in bit rate at the same visual quality. Looking over problem of limited bandwidth it became necessary to compress video data. To compress the data spatial and de correlated data is removed. This causes compacted information to fill in as a key asset for machine learning with altogether less examples for preparing. In this venture, a great way to deal with closer view extraction/division is expected abuse novel spatio-transient de-related piece alternatives taken straightforwardly through HEVC compressed video. Various techniques, in distinction, work on pictures which are uncompressed providing vital storage and machine resources not just for the decryption method before data formatting. The planned approach has been quantitatively and qualitatively evaluated against Gaussian mixture model.**

***Index Terms* —CTU, *Background subtraction, HEVC, Dirichlet Process, Video surveillance etc.***

## I. INTRODUCTION

Background subtraction may be a method within an image process. An Associate degree of image's foreground is extracted for further additional process. An associate degree image's area of interest square measure objects including text, humans, cars etc. in its foreground. This system is used at the time of image preprocessing object where localization is needed. Background subtraction helps for detecting the moving objects in videos taken from static cameras [1].

The method of detecting the moving items between the modern-day body and a reference frame, are often referred as "Background image", or "Background model". Background subtraction is broadly carried out if the photo in query is part of a video circulation. Historical past subtraction provides important cues for numerous applications in pc imaginative and prescient, for example surveillance tracking or human poses estimation [2].

December 2012, Gary J. Sullivan and Thomas Wieg abstracted that "Overview of the high efficiency video coding (HEVC) standard" [3]. In April 2014, Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, Larry Davis published "Real-time foreground–background segmentation using codebook model [4]". In May 2014, T. Bouwmans abstracted that "Traditional and recent approaches in background modeling for foreground detection: An overview [5]". Due to the challenges of video surveillance, evaluations of recent background subtraction methods suffer from various shortcomings. They abstracted that "Background subtraction approach is employed to sight the moving object from background completely different strategies are planned to sight object motion by exploitation different background subtraction techniques over recent years [6].

In November 2015, B. Dey used the rising High-Efficiency Video cryptography for compression video. It guarantees that up to 50%-bit rate savings compared to the compression schemes obtainable now days. Also, the compression potential of HEVC offers a chance to use the information measures, new techniques of feature extraction are factored using quicker algorithms [7]. The state of the art (SoA) algorithms use model for associate degree freelance background which care for pictures which are uncompressed. Therefore, compressed videos are pre-processed which should be utterly de-compressed to claim about important time and memory [8].

Moving objects are segmented from the background by using motion detection algorithm. To implement this, take a picture as background and frames obtained at the time t, denoted by I(t). That will check with the background image denoted by B. By victimization image subtraction technique for every pixel in I(t), we can phase out the objects. And the pixel price is denoted by P[I(t)] and the corresponding pixels of the background image is denoted as P[B].

It is written as:

$$P[F(t)] = P[I(t)] – P[B] \quad (1)$$

At time t, the background is considered as frame. This distinct image shows the intensity of element locations that are modified within the 2 frames. Although we have removed the background, all foreground pixel area per unit are moving and every background pixels area per unit are static [3]. To boost the subtraction, threshold is placed on distinct image.

$$|P[F(t)] - P[F(t+1)]| > \text{Threshold} \quad (2)$$

The value of threshold, the difference image's pixels' intensities are filtered [4]. The accuracy depends on movement speed of movement in the scene. Higher threshold is required for Faster movements.

In this, a proficient way to deal with frontal area extraction is anticipated exploitation novel spatio-transient de-corresponded piece alternatives taken specifically from the HEVC packed video [5]. A large portion of strategies, in qualification, take a shot at uncompressed pictures giving essential stockpiling and machine assets for the decipherment strategy before arrangement or to highlight determination/extraction and foundation displaying stage tailing it.

978-1-5386-4008-1/17/$31.00 

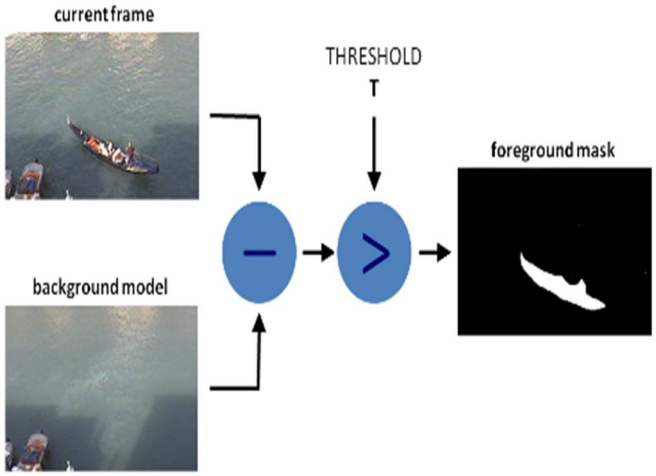



Fig.1 Foreground mask of image

High intensity Video cryptography is that the most up to date video cryptography ordinary of the ITU-T Video cryptography pros bunch [3] [8]. The most point of the HEVC institutionalization exertion is to adjust impressively enhanced pressure execution in respect to existing gauges inside the fluctuate of half piece rate diminishment for equivalent tangible action video quality. This paper gives a diagram of the specialized choices and qualities of the HEVC ordinary. The anticipated approach has been subjectively and quantitatively assessed against Gaussian mixture model.

Dirichlet handle Gaussian blend models is utilized to surmised per-pixel foundation circulations taken after by probabilistic regularization. In this, per pixel modes are consequently number by utilizing non-parametric Bayesian technique, keeping away from over-/under-fitting.

## II. BACKGROUND SUBTRACTION WITH HEVC VIDEO

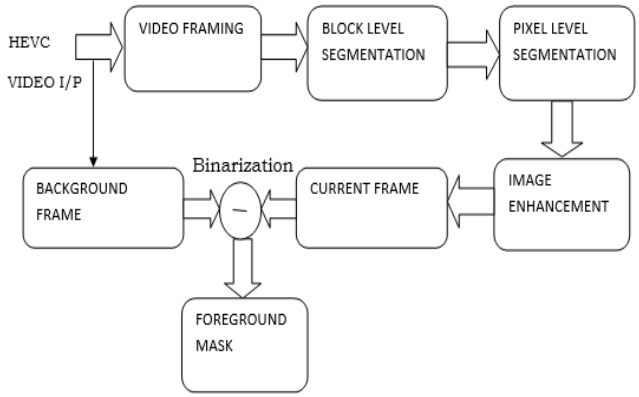



Fig. 2. Block diagram of background subtraction with HEVC video

### *A. HEVC Video I/P*

HEVC could be a video compression modern, a successor to H.264/MPEG-4 AVC. HEVC is communicated to twofold the information pressure quantitative connection contrasted with H.264/MPEG-4 AVC on the indistinguishable level of video high-quality. At the equivalent piece rate, it's going to rather be wonted to give significantly propelled video quality. It's going to encourage 8K UHD and resolutions the greatest sum as 8192×4320.

Now a day, the greatest supply of unstructured huge data is Surveillance video. HEVC has a huge position in bringing down the costs that are identified with transmission capacity. A few the advantages of HEVC over the inheritance MPEG-four propelled Video Coding (AVC), are that forty rates or additional piece rate decrease at the equivalent obvious quality. Considering data transfer capacity limits, video records are compacted basically by putting off spatial and fleeting relationships that exist in its uncompressed shape. This causes compacted records.

### *B. Video Framing*

A video records the incorporate frames. These frames appear before at a rate greater than our perception of vision. It gives a sensation of movement of an object. We can see by this movement by looking just at the screen where frames are appearing at high rate. Hence one can state that video records are nothing but casings of fundamental substance. Those fundamental elements are utilized for chronicled past subtraction.

Image segmentation is that the method of parceling a reflection into very one sections (sets of pixels, furthermore alluded to as super pixels) [6]. The aim of segmentation is to adjust and additionally exchange the representation of a photo into a certain something that is a great deal of substantive and less troublesome to investigate. Image segmentation is regularly wonted to discover questions and cutoff points (strains, bends, so forward.) in pics.

### *C. Block Level Segmentation*

CTU is the essential handling unit of the HEVC video standard what's more, adroitly compares in structure to large scale piece 2 devices that had been utilized as a part of various going before video measures. CTU is in like manner called largest coding unit (LCU). A CTU might be between 16×16 pixels and sixty 64x64four pixels in size with a greater size typically developing coding proficiency. Square level division comprises of acting an unpleasant square stage division of each casing by method for picking of a settled of capacity CTUs that are involved totally or somewhat by utilizing components of moving objects [7].

### *D. Pixel Level Segmentation*

Block level segmentation includes performing a better pixel-degree division by method for pushing off pixels from the chose CTUs. It is comparable (inside and out) to the comparing foundation display. This fine level course method protects the edges at closer view outline. Diminishing clamor arrange at edges of exchanging article.

### *E. Image Enhancement*

To get pleasant results in image processing it's miles continually necessary that input images must be of good quality. But almost this isn't that a great deal easy. Because of one-of-a-kind reasons like noise, mild influences we get low or medium quality images. Consequently, it turns into important to enhance their exceptional [8]. To enhance the pleasant of photo we're the usage of image enhancement set of rules. This algorithm complements the image by way of

specializing in parameters like contrast, brightness adjustment.

### F. Current Frame

The modern frame obtained from HEVC video is now de-noised, more desirable and processed with one of a kind segmentation. This frame is used to locate current moving object in a frame.

### G. Background Frame

It's far reference frame of predefined region. It is considering to be idle. While it is compared, and extracted from current frame output will detect shifting object in a current frame.

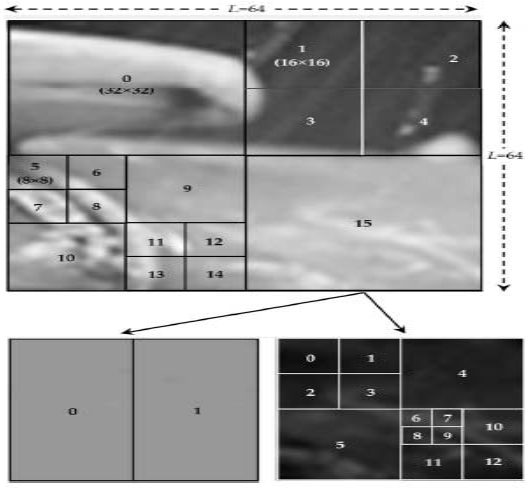
### H. Foreground Frame

Foreground frame is final results of difference between current body and history version. It indicates the shifting item in a binary form.

## III. CODING TREE UNIT

A compacted HEVC video comprises of an arrangement of pictures or casings. Each of that pictures or edges is part into non-covering squares known as coding tree unit (CTU) [8].


The luma Coding tree block of size 64x64 Coding tree unit which is at the location 14 is split into the CQT structure as shown below.




Predicted information of the CB number 15 is split into 2PBs

Residual information of 32x32 CB number 15 is split into recursively into a RQT consisting of 13 TBs

Fig.3. Top image is an example of 256x256 image

The CTU is that rudimentary unit of pressure. Coding Tree Square (CTB)keeps data for each shading component in structures. A luma CTB covers a rectangular picture zone of L × L trial of the luma portion also, the relating chroma CTBs cover L/2 × L/2 trial of each of the two chroma parts. Each CTB can be composed into a quadtree structure, the separation down to 8 × 8 (in units of luma tests) ranges. This quadtree like structure is called coding quad tree. An illustration shown in below Fig. 3 (top), the 64 ×64 luma CTB identifying with the CTU at range 14 is seemed to contain two 32 × 32, six 16×16, and eight 8×8 areas. They are called as coding unit (CBs). The spatial and the transient redundancies exclusively of a given CB are obstructed into parts that are expected from in advance coded hinders inside a comparative packaging (called intra-forecast), and from the neighboring frames (called Inter forecast). The pieces are called forecast squares (PBs). Between expectations of a PB is a short lived de-relationship technique by which perhaps a couple sensible reference pieces are picked.

## IV. DIRICHLET PROCESS GAUSSIAN MIXTURE MODEL (DP-GMM)

A Gaussian mixture version is a probabilistic version that assumes all the records factors are generated from an aggregate of a finite variety of Gaussian distributions with unknown parameters [9]. Exclusive classes to estimate Gaussian aggregate models, that correspond to exclusive estimation strategies, detailed under. One among them is DP-GMM.

### A. DP-GMM

A Dirichlet system is a chance distribution whose area is itself a set of possibility distributions. It is regularly utilized in Bayesian inference to describe the prior expertise approximately the distribution of random variables, this is, how probable it is that the random variables are dispensed in line with one or another distribution [8] [11].

In simple phrases, we can say that Dirichlet process works on probability distribution of stochastic processes. The term stochastic process describes events or structures which are unpredictable because of the have an impact on of a random variable.

### A. Flow Chart

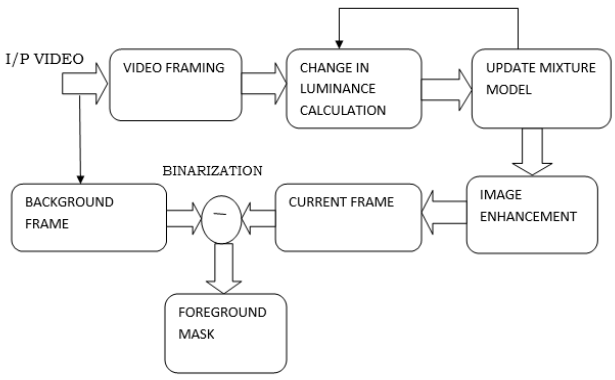



Fig.4. Block diagram of background subtraction with HEVC video by using DP-GMM.

This approach is a non-parametric Bayesian method that spontaneously estimates the wide variety of combination additives is automatically estimate via this approach to version the pixel's background colour distribution [12].

In DP-GMM as shown in figure due to light effect the colour distribution preserve on converting, this will cause misguided result right here we are able to use stochastic process if you want to don't forget the all chances of exchange in shade distribution. Those adjustments in colour distribution might be updated to combination model continuously. This non-stop ongoing procedure will give updated improved photograph frame. Gain of this method is that current frame is now adaptive to change in coloration distribution with the aid of considering probabilities over possibility.

## V. RESULTS

Result section consist of 3 parts. Firstly, Background frames and for ground mask of that frames shows in figure 5. Secondly, figure 6 shows a) Ground truth image b) Correct background subtracted image. Finally, comparison is done on qualitative as well as quantitative basis of the proposed method (figure 8) with Dirichlet method (figure 7) is shown. we conclude that proposed method is better than dirichlet method for background subtraction. Table shows the quality parameters used for simulation

TABLE I. QUALITY PARAMETERS USED FOR SIMULATION

| METHODS | MSE | PSNR(dB) | CORRELATION |
|---|---|---|---|
| DIRICHLET METHOD | 3.2289e+03 | 26.0806 | 0.8948 |
| PROPOSED METHOD | 718.1134 | 39.1377 | 0.8871 |

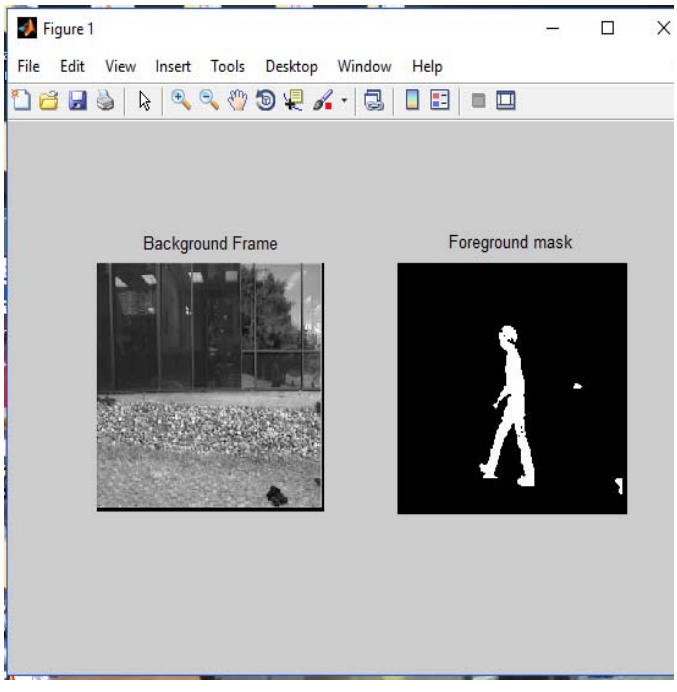



Fig. 5 Background frames and for ground mask of that frames.

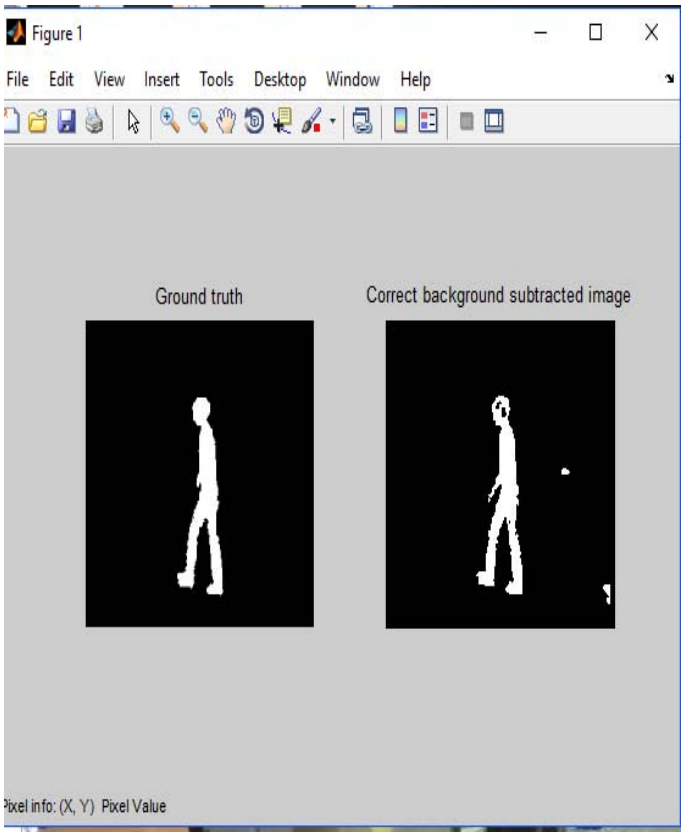



Fig. 6. a) Ground truth image b) Correct background subtracted image.

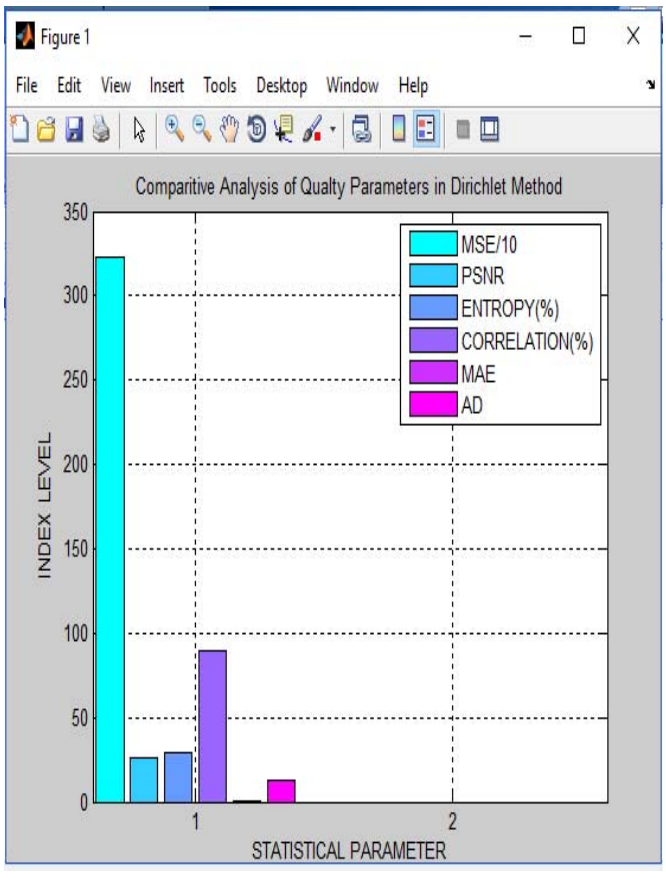



Fig.7. The statistical parameters of Dirichlet process

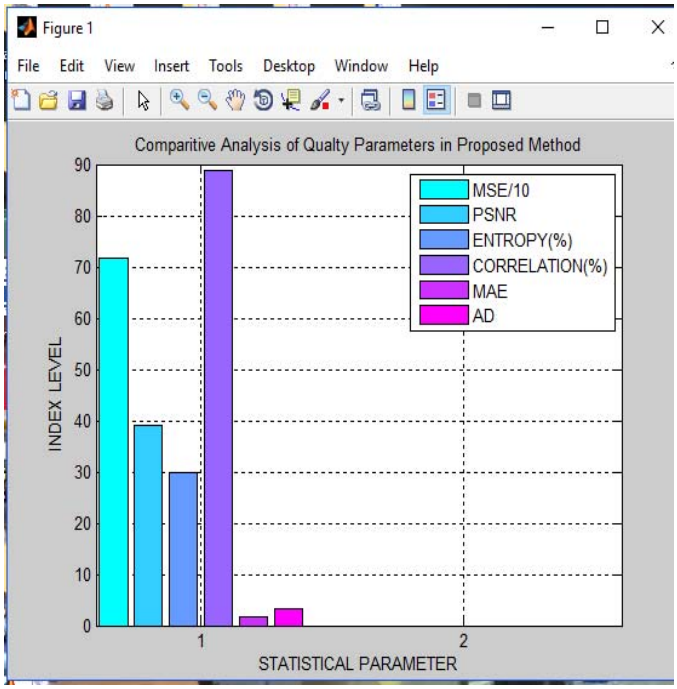



Fig. 8. The statistical parameter of Proposed method

## VI. CONCLUSION

In this paper, we proposed a method by using CTU features of HEVC compressed video for extracting foreground objects. We conclude that Background subtraction methods with HEVC algorithm and Gaussian mixture model algorithm have wide scope for analysis purpose. Hence, I have selected these techniques for comparative analysis purpose.


## ACKNOWLEDGMENT

Author Mr. Pawan Arun Gawade thanks Prof. Manoj Kumar, H.O.D. (E&TC) in S.V.C.E.T., Rajuri, Pune for guiding throughout her work. Also thanks to Mr. P. Balaramudu, Vice- Principal of S.V.C.E.T., Rajuri, Pune for guiding throughout work.


## *References*


[1] O. Al-Kofahi, R. Radke, S. Andra and B. Roysam, "Image change detection algorithms: A survey," *IEEE Trans. Image Process*., vol. 14, no. 3, Mar. 2005.

[2] "Surveillance video: The biggest big data," by T. Huang, Comput. Now, vol. 7, no. 2, Feb. 2014.

[3] "Overview of the high efficiency video coding (HEVC) standard," by G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, Dec. 2012.

[4] "Real-time foreground–background segmentation using codebook model" by Kyungnam Kim, David Harwood, Larry Davis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol*. 36, NO. 4, April 2014

[5] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," Comput. Sci. Rev., vols. 11–12, May 2014.

[6] N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," *Mach. Vis. Appl.,* vol. 8, no. 3, 1995.

[7] "Pfinder: Real-time tracking of the human body," by C. R. Wren, T. Darrell, and A. P. Pentland. *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 19, no. 7, Jul. 1997.

[8] " Efficient Foreground Extraction from HEVC Compressed Video for Application to Real-Time Analysis of Surveillance 'Big' Data" by Bhaskar Dey and Malay K. Kundu. Senior Member *IEEE Transactions on Image Processing, Vol. 24, No. 11,* November *2015*

[9] "Image segmentation in video sequences: A probabilistic approach," by N. Friedman and S. Russell in *Proc. 13th Conf. Uncertainty Artif. Intell.,* 1997.

[10] W. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proc. *IEEE Conf. Comput. Vis. Pattern Recognit*., Jun. 1999.

[11] Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in Proc. *IEEE Int. Conf. Pattern Recognit*., Aug. 2004.

[12] T. Ellis, "Self-adaptive Gaussian mixture model for urban traffic monitoring system," in Proc. *IEEE Int. Conf. Comput. Vis*., Nov. 2011.

[13] Wang, and J.-K. Wu, "A dynamic conditional random field model for foreground and shadow segmentation," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 28, no. 2, Feb. 2006.