IOSR Journal of Engineering (IOSR JEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
PP 31-37

Efficient Foreground Extraction with Video Processing Technique

Jadhav Rahul Ramesh[1], Dr. Govind. U. Kharat[2]
[1] *ME Student, E & TC Engineering (VLSI & Embedded System), SPCOE, Otur, Pune, Maharashtra, India.*
[2]*Principal, SPCOE, Dumbarwadi (Otur), Pune, Maharashtra, India.*

Abstract: *Video analysis is big source of information today. Introduction to HEVC that is high-efficiency video coding 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 compressed data to serve as a vital resource for machine learning with significantly fewer samples for training. In this concept, an efficacious foreground extraction is obtain using novel spatio-temporal de-correlated block features taken directly from the HEVC compressed video [1]. Most of techniques, in contrast, work on uncompressed images providing significant storage and computational resources not only for the decoding process prior to initialization but also in order to feature selection/extraction and background modeling stage following it. The proposed approach has been qualitatively and quantitatively evaluated against Gaussian mixture model.*

Index Terms: *Background subtraction, HEVC, transform coding, statistical signal processing, video surveillance*

I. Introduction

Framework seduction is a method in the area of image processing. In that an image's foreground is extracted for further processing (object recognition etc.). Generally, an image's regions of interest are objects such as humans, cars, text etc. in its foreground. After the phase of picture preprocessing object restriction is required which may make use of this technique Background subtraction is a widely-used approach for detecting moving objects in videos from static cameras [1].

The rationale in the method is that of detecting the moving items from the distinction between the modern-day body and a reference frame, often referred to as "Background image", or "Background model". Foundation subtraction is extensively talking done if the photograph in question is a piece of a video dissemination. Old past subtraction gives essential for innumerable applications in inventive and insightful, for example surveillance tracking or human poses estimation. but, background seduction is mainly based on a static background speculation which is not applicable in real environments [2]. Reflections or animated images on screen lead to background changes with indoor scenes. Similarly, due to radiance, breeze, rainfall difference brought by weather, static backgrounds concept have difficulties with external scenes.

In Dec.2012, Gary J. Sullivan and Thomas Wieg abstracted that "High Efficiency Video Coding (HEVC) is currently being prepared for newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group [3].

In April 2014, Thanarat H. Chalidabhongse, David Harwood, Kyungnam Kim, they abstracted that "We present a real-time algorithm for foreground–background segmentation [4]. Sebastian Brutzer, Benjamin H¨oferlin, Gunther Heinemann Intelligent Systems Group, Universit¨at Stuttgart, Germany they abstracted that "Background subtraction is one of the key techniques for automatic video analysis, especially in the domain of video surveillance [5]. Even there is importance, assessment of present background subtraction concept with releted to the challenges of video surveillance caused from various shortcomings.

"Framework subduction approach is used to identify the working object from background. Different methods have been proposed to detect object motion by using different background subtraction techniques over recent years [6]. Olivier Barnich and Marc Van Droogenbroeck, Member, IEEE they abstracted a technique for motion detection that incorporates several innovative mechanisms [7].

In November 2015, BhaskarDey and Malay K. Kundu used the emerging High-Efficiency Video Coding (HEVC) [3] standard for video compression, promises up to 50%-bit rate savings compared against the best of compression schemes available today. While the compression efficiency of HEVC offers a unique opportunity to alleviate the bandwidth crunch, new methods of feature extraction directly from compressed video must be factored into the design of faster algorithms. However, most state-of-the-art (SoA) algorithms operate on uncompressed images with an independent background model for each pixel (pixel-based methods).

Therefore, compressed videos must undergo computationally intensive pre-processing to be completely de-compressed, claiming significant time and memory prior to the application of such algorithms. Real-time performances with accuracy comparable to those of pixel-based methods are targeted.

Finally, pixel value is propagated into the background model of a neighboring pixel when the pixel is found to be part of the background. By using pseudo code and the parameter values, we describe our method in full details and compare it to other background subtraction techniques [8].

Foreground or moving objects are segmented from the background by using motion detection algorithm. To implement this, take an image as background and take the frames obtained at the time t, denoted by I(t) to compare with the background image denoted by B. In this, we can segment out the objects simply by using image subtraction technique for each pixel in I(t), take the pixel value denoted by P[I(t)] and subtract it with the corresponding pixels on the background image denoted as P[B]

In mathematical equation, it is written as:

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

The background is assumed to be a frame at time *t*. This differentiation picture would show strength for the pixel locations which have changed in the two frames. Though we have removed the background, this will work for cases where all foreground pixels are moving and all background pixels are static [3]. Threshold is put on this difference

image to improve the subtraction.

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

The differentiation of clarity of image pixels' power are 'maximum' or filtered depend on value of maximum [4]. The accuracy is dependent on speed of movement in the scene. Faster movements may require higher thresholds

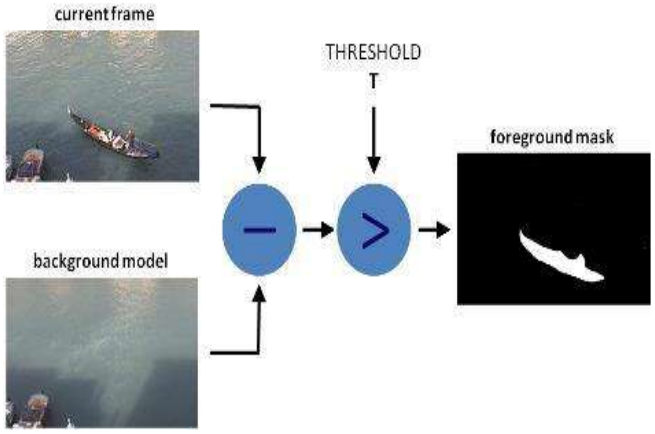


Fig.1 Foreground mask of image

Here, an efficacious approachto foreground subtraction is assumed using the concept of novel spatio-temporal block features which is taken directly from the HEVC technique compressed video [5]. Most of techniques, in contrast, work on uncompressed images providing significant storage and computational resources for the decoding process prior to initialization or to feature selection/extraction and background modeling stage following it.

In this concept HEVC is the latest video programming excellence of the ITU-T Video Coding Experts team and the ISO/IEC Moving Picture Experts Group (MPEG) [3]. The main aim of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standards in the range of half bit-rate reduction for equal perceptual video quality. This paper provides an overview of the technical features and characteristics of the HEVC standard. The proposed approach has been qualitatively and quantitatively evaluated against Gaussian mixture model

.

II. Background Subtraction With Hevc Video

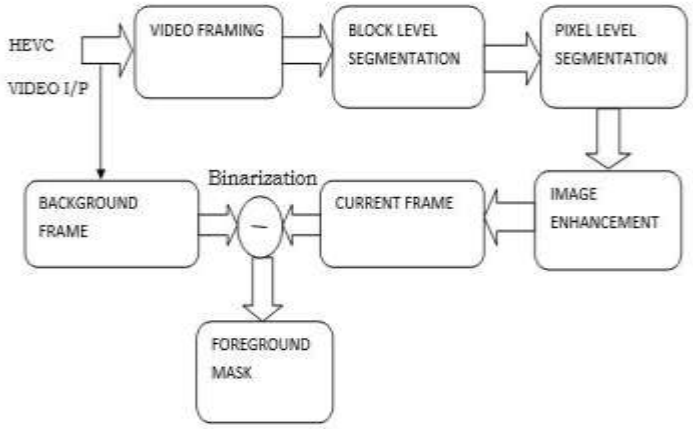


Fig. 2. Block diagram of background subtraction with HEVC video

A. HEVC Video I/P

HEVC is a video compression fashionable, a successor to H.264/MPEG-4 AVC. HEVC is stated to double the information compression ratio in comparison to H.264/MPEG-4 AVC on the identical level of video high-quality. at the equal bit rate, it may instead be used to offer significantly advanced video quality. It may help 8K UHD and resolutions as much as 8192×4320.

Surveillance video is the biggest supply of unstructured large information these days. HEVC popular is poised to have a large position in lowering the costs related to transmission and storage. Several the benefits of HEVC over the legacy MPEG-four advanced Video Coding (AVC), is a spectacular forty percentages or extra bit-rate reduction at the equal visible quality.

B. Video Framing

A video record includes frames. those frames whilst appear earlier than us in a charge more than our belief of imaginative and prescient, gives a sensation of an object moving earlier than us, through taking a gander at the screen on which outlines are performing at high charge. Henceforth one could state that outlines are the basic element of a video report. The basic substances are utilized for authentic past subtraction.

Picture division is the technique for parceling a virtual picture into more than one portion (sets of pixels, also known as super pixels) [6]. The point of division is to clear up and exchange the outline of an image into the idea that is significant and low hard to explore limits

C. Block Level Segmentation

Coding tree unit is the premier preparing unit of the High Effectiveness Video Coding (HEVC) video standard and conceptually corresponds in structure to macro block devices that had been used in numerous preceding video standards. CTU is likewise called largest coding unit (LCU). A CTU may be between 16×16 pixels and sixty 64x64four pixels in size with a bigger size usually growing coding efficiency. Square dimension division comprises of acting a harsh square stage division of each casing by way of choosing of a fixed of ability CTUs that are occupied absolutely or partly by using elements of moving objects [7].

D. Pixel Level Segmentation

Block level segmentation includes performing a finer pixel-degree segmentation by way of casting off pixels from the selected CTUs which might be similar (in depth) to the corresponding background model.

E. Image Enhancement

To get lovely outcomes in picture handling its miles continually necessary that input images must be of good quality. But almost this isn't that a great deal easy. In light of stand-out reasons like clamor, gentle impacts we get low or mid-range quality pictures. Therefore, it turns into important to enhance their exceptional [8]..

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 being idle. While it is looked at, and separated from current casing yield will distinguish moving article in a present casing.

H. Foreground Frame

Frontal area outline is last aftereffects of distinction between current body and history variant. It indicates the shifting item in a binary form.

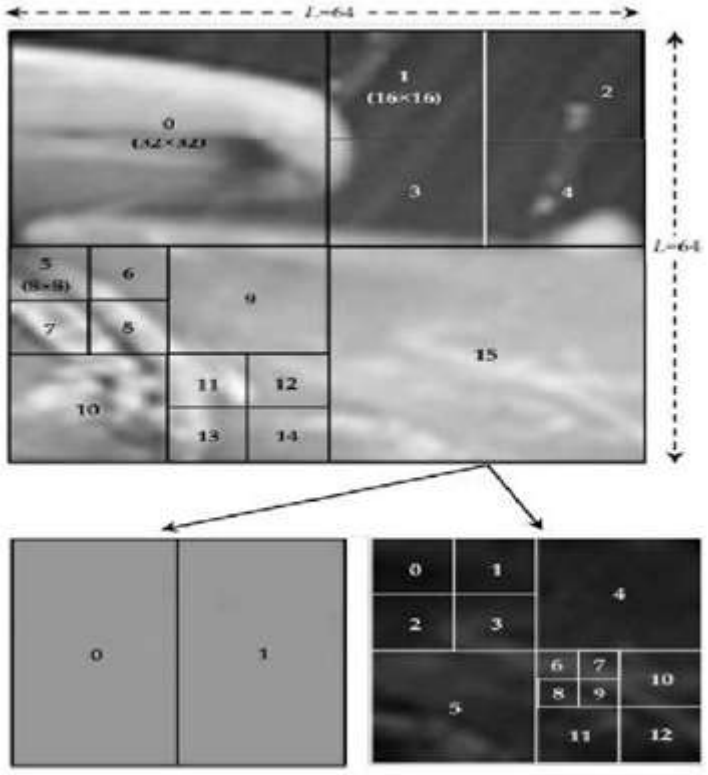
III. Coding Tree Unit

A compressed HEVC video consists of a sequence of pictures or frames, each of which is split into non-overlapping blocks called coding tree unit (CTU). The CTU is the fundamental unit of compression, which maintains information for each colour component in structures known as the coded tree block (CTB). A luma CTB covers a rectangular picture area of $L \times L$ samples of the luma component and the corresponding chroma CTBs cover $L/2 \times L/2$ samples of each of the two Chroma components. Thevalue of L (heretofore used to denote the size of luma CTBs) for a given sequence is fixed by the encoder and signalled by a sequence parameter which may be either equal to 16, 32, or 64. Each CTB can be split recursively into a quad tree structure, all the way down to 8 × 8 (in units of luma samples) regions. The quad tree structure is known as the coding quad tree (CQT). So, for the example shown in Fig. 1 (top), the 64 × 64 luma CTB corresponding to the CTU at location 14 is shown to consist of two 32 × 32, six 16×16, and eight 8×8 regions. These regions are called coding blocks (CBs). The spatial and the temporal redundancies respectively of a given CB are reduced by splitting it into blocks that were predicted from previously coded blocks within the same frame (called intra-prediction), as well as from the neighboring frames (called inter-prediction). The blocks are called prediction blocks (PBs). Inter-prediction of a PB is a temporal de-correlation technique by which one or two suitable reference blocks are selected

Above figure shows an example of 256 x 256 image partitioned into sixteen 64 × 64 CTUs (top image). The luma CTB corresponding to CTU at location 14 is split into a CQT structure (middle image). The CB number 15 of the CQT is shown to be coded using a predicted information consisting of 2 PBs (shown just above, to the left) and the prediction residual, which is split into an RQT consisting of 13 TBs (shown just above, to the right). i.e., the residual, still contains correlation between spatially neighbouring pixels. Therefore, the residual is further split into a quad tree structure, called the residual quad tree (RQT), into transform blocks of suitable sizes ranging from 4 × 4 to 32 × 32 samples [1] [28].

The TBs are subjected to transform coding and quantization. The purpose of transform coding to decompose a batch of correlated signal samples into a set of uncorrelated spectral coefficients, with energy concentrated in as few coefficients as possible. This compaction of energy permits a prioritization of the coefficients, with the more energetic ones receiving a greater allocation of encoding bits. The resulting transform coefficients are quantized. The purpose of quantization is to map a large set of input coefficients to a smaller set of representative quantization levels. This results in rounding of coefficient values to some acceptable unit of precision specified by a quantization parameter (*QP*). The output of a quantizer is typically a sparse array of quantized coefficient levels, mainly containing zeros. Finally, the non-zero coefficient levels are entropy coded. This process eliminates *coding* redundancy by assigning fewer bits to more frequently-occurring symbols

The luma CTB corresponding to the 64x64 CTU at location 14 is split into the CQT structure shown in below



Predicted information of the Residual information of 32x32 CB number 15 is split into 2PBs CB number 15 is split into recursively into a RQT consisting of 13 TBs

Fig.3. An example of 256x256 image divide into 16 64x64 CTUs (top image), the luma CTB corresponding to CTU at location 14 is split CQT structure (middle image). The CB number 15 of the CQT is shown to be coded using predicted information (last image)

IV. Dirichlet Process Gaussian Mixture Model (Dp-Gmm)

A Gaussian mixture category is a measure category 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 [11].

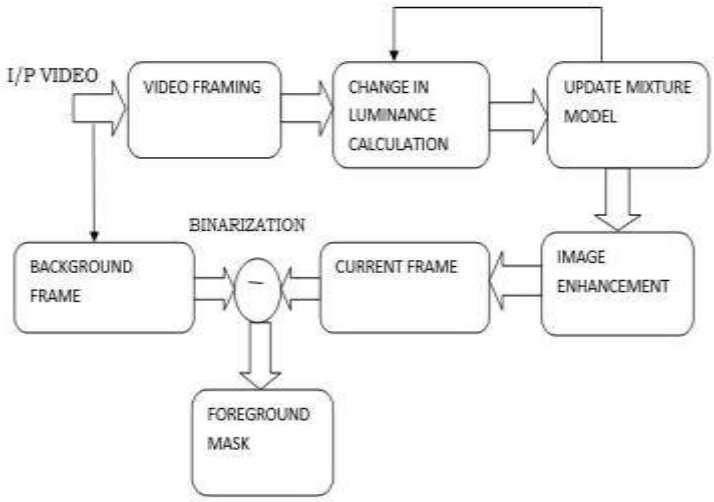


Fig.4. Block diagram of background subtraction with HEVC video by using DP-GMM.

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

➢ Qualitative Evaluation of proposed method by using following parameters

The important parameters of an image are Mean square error (MSE), Peak signal to noise ratio (PSNR) and Correlation. In that method, we calculate the parameters of image and represent it in graphical form. The evaluation is based on parameters like MSE, PSNR Correlation.

MSE= sum ((a1-b1) ^2))/ (M*N)

PSNR= 20*log10 (255*255/MSE)

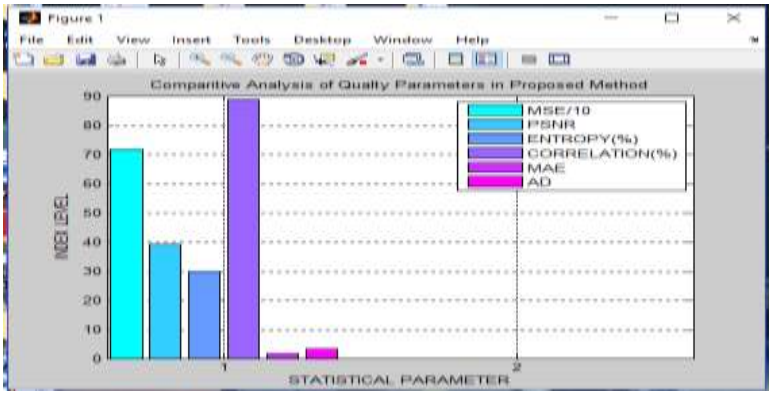


Fig. 5. Comparison of H.264 (CBR) & DP-GMM

VI. Conclusion

In this paper, we proposed a method for extracting foreground objects using CTU features of HEVC compressed video. 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.jadhav Rahul Ramesh thanks to Dr.G.U.kharat Guide & Principal guiding throughout work also Prof. Chinchole sir, H.O.D. (E&TC) in S.P.C.O.E.Otur, Pune for guiding throughout work. Also thanks to Prof.Bansode Rahul, ME Coordinator of S.P.C.O.E.Otur, and Pune for guiding throughout work.

References

[1]. R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. ImageProcess.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[2]. T. Huang, "Surveillance video: The biggest big data," Comput. Now, vol. 7, no. 2, Feb. 2014. [Online]. Available: http://www.computer. Org/portal/web/computing now/archive/february2014

[3]. G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. CircuitsSyst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[4]. Kyungnam Kim, Thanarat H. Chalidabhongse, David Harwood, Larry Davis."Real-time foreground–background segmentation using codebook model". *IEEE Transactions on Pattern Analysis andMachine 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, pp. 31–66, May 2014.

[6]. N. J. B. McFarlane and C. P. Schofield, "Segmentation and tracking of piglets in images," *Mach. Vis. Appl.*, vol. 8, no. 3, pp. 187–193, 1995.

[7]. C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pent land, "Pfinder: Real-time tracking of the human body," *IEEE Trans. Pattern Anal.Mach. Intell*., vol. 19, no. 7, pp. 780–785, Jul. 1997.

[8]. BhaskarDey and Malay K. Kundu, Senior Member*," "* Efficient Foreground Extraction From HEVC Compressed Video for Application to Real-Time Analysis of Surveillance „Big‟ Data*" IEEE*

[9]. *Transactions On Image Processing, Vol. 24, No. 11,* November *2015*

[10]. N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach," in *Proc. 13th Conf. Uncertainty Artif.Intell.*,1997, pp. 175–181.

[11]. C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proc. *IEEE Conf. Comput. Vis.Pattern Recognit*., Jun. 1999, pp. 246–252.

[12]. Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in Proc. *IEEE Int. Conf. Pattern Recognit*., Aug. 2004, pp. 28–31.