Performance Analysis of Multi-level HAAR in Background Removal for Object Detection

Meena Deshpande* and M. B. Veena

Electronics and Communication Engineering, BMS College of Engineering, Bengaluru – 560019, Karnataka, India; meenapandit26@gmail.com, veenamb.ece@bmsce.ac.in

Abstract

Objective: This study proposes performance improvement in speed of multi-level HAAR processed images for object detection. Method: Background subtraction algorithm is implemented using phase as a feature to reduce illumination variation. The algorithm is implemented on level 2 and level 3 HAAR compressed images. Simulation results are obtained on kit ware database. Findings: Simulation results show that object detection is faster in level 3 HAAR compressed images as compared to level 2 and level 1 HAAR compressed images. Average time required for processing single frame is in range of 6.53 to 29.22 ms in level 3 while that in level 2 is 6.65 to 36.46 ms. Improvement: Using this approach saving of 5% to 22% of processing time is observed at level 2 of HAAR while a saving of 9% to 48% of time is observed at level 3 of HAAR.

Keywords: Background Subtraction, Illumination Variation, Multi-Level HAAR, Phase as a Feature

1. Introduction

Over past, numerous approaches have been used to detect the object using compressed videos due to increasing demand of various video processing algorithms implemented on FPGA hardware¹. This study proposes the object detection on HAAR compressed videos taking in to consideration the memory and the power requirements for implementation. Many features are used for object detection like histogram, edge and corner. Statistical features include intensity, SVM, texture, etc.

One of the intricate task in moving object detection is detection of objects in presence of changing illumination conditions^{2.3}. Different approaches of object detection proposed include feature based object detection, template based object detection, etc. Recently many feature extraction methods are developed to overcome the challenges like changing illumination, occlusion, camouflage image etc. faced by background subtraction methods⁴. Proposed approach uses phase as a feature as it is invariant to illumination changes. To extract phase, Gabor filter is used as it is related to invariance to illumination. Feature extraction is performed on multi-level HAAR compressed images, since it is very effective in detecting the exact instants when a signal changes. HAAR wavelets are easy to implement and fast in computation. Results are generated for level1, level 2 and level 3 HAAR compressed images. Speed of operation is compared for three sample files for 45 frames processing showing the enhanced speed in higher levels of HAAR usage in pre-processing.

Organization of the paper is as follows: Related work is surveyed in Section 2, Section 3 briefs the methodology used. Experimental findings and observations are discussed in Section 4. Work is concluded in 5 and Section 6 discusses future work.

2. Literature Survey

In the study, "Background subtraction based on phase feature and distance transform" suggested detection of moving objects in changing illumination conditions. In^{5.6} proposes an algorithm for moving object detection in

video which is compressed using wavelet transform and frame difference method in "Moving object detection in wavelet compressed video". In² in the study "A Proposed Approach for Image Compression based on Wavelet Transform and Neural Network" proposes the selection of a suitable wavelet function to compress the image efficiently by training back propagation neural network for selection of suitable wavelet function. In but the proposed work did not produce a very high recognition t have reviewed different object tracking methods under different illumination conditions in the review paper "Illumination Condition Effect on Object Tracking". In the study "2-D Object Recognition Approach using Wavelet Transform" used wavelet transform for object detection suitable for greyscale images, however it does not worked well for color images. In the study "Moving Object Detection under Various Illumination Conditions for PTZ Cameras" has tried to overcome the problem in background modelling by adapting a texture based method XCS-LBP and GMM for segmenting the foreground. In⁸ study, "Multilevel decomposition Discrete Wavelet Transform for hardware image compression architectures applications" proposes a Discrete Wavelet Transform (DWT) for image compression applications to eliminate redundant information from the transmitted images or video frames over the wireless channel. New Image Compression Algorithm Using HAAR Wavelet Transformation is proposed in "A New Image Compression Algorithm using HAAR Wavelet" by Compression Ratio, PSNR, Threshold Value and Reconstructed Normalization were calculated⁹. Computational time and computational complexity is reduced in Fast HAAR wavelet transform. In¹⁰ suggested a modified HAAR for image compression in "Image Compression using HAAR and Modified HAAR Wavelet Transform". The advantage of HAAR and modified HAAR wavelet transform is the infrequent representation, fast transform and low memory space requirements¹¹.

3. Methodology Used

The work proposed in this study focuses on object detection. It is implemented on enhanced images such that the object detection shows remarkable reduction in execution time. Proposed work cites the reference of object detection in which object detection is done under complex environment. In^{12,13} Phase feature approach. which is suitable for background modelling is used. This is done by taking every pixel created as a group of adaptive phase features. Distance transform is used to get clear results by accumulating the neighboring foreground pixels. Basic methodology consists of following steps:

• Construct the Gabor filter, 2D Gabor filter is a Gaussian kernel function modulated by sinusoidal plane wave:

$$\begin{split} \Psi_{\varphi,v}(z) &= \frac{\|k|\Psi,v\|^2}{\sigma^2} e^{\left(-\|k_{\varphi,v}\|^2/2\sigma^2\right)} \left[e^{-ik_{\varphi,v}z} - e^{-\sigma^2/2} \right] \\ k_{\varphi,v} &= \left(\frac{k_{jx}}{k_{jy}} \right) = = \left(\frac{k_v \cos \phi_\varphi}{k_v \sin \phi_\varphi} \right) \\ k_v &= f_{max}/2^{v/2}, \quad \phi_\varphi = \varphi(\pi|\varphi_{max}) \\ B &= \sum_{i=1}^K \omega_i \sum_{i=1}^K \omega_i j = 1, \quad \Phi \text{ denotes feature value b} \end{split}$$

 φ is the orientation and varies from 0 to max1 v varies from 0 to max -1. σ is the Gaussian width to wavelength.

$$G_{\varphi,v}(z) = I(z)^{*\varphi_{\varphi,v}(z)}$$

* is a convolution operator and G(z) is a outcome of convolution expressed as amplitude part and phase part. Gabor filter is constructed with 4 frequencies and 6 orientations which means $v_{max} = 4$ and $\phi_{max=6}$.

 v_{max} and φ_{max} respectively depict the number of frequencies and orientations. Frequency of the Gabor wavelet is calculated as per the formula;

$$k_v = \frac{J_{max}}{2^{v/2}}$$

Based on Gabor wavelength, patch size is selected for convolution is $4 \ge 4$. This gives different values of phase for different frequencies and orientations. Number of effective Gabor phases for every pixel was done by analyzing the $4 \ge 6$ convolution results.

- Initialize the phase based background model. For each new frame do the patch based Gabor filter convolution.
- Find the phase feature Ø with existing model. The phase feature is defined as:

$$\phi(z) = \sum_{l=1}^{L} \theta_l(z)$$

Where z = Image position.

 \emptyset (z) = Feature value at position z.

 $\theta_{l}(z)$ = Gabor phase value fulfilling the condition that its corresponding amplitude is the lth highest

L= number of selected Gabor phases.

Based on the stability and distinguishability L = 5.

• Compare the pixel phase feature with existing models. If matched model is found update the

matched model parameters. Use the model with lowest weight is replaced. Sort the background models according to ω/σ . Select first B distributions as background.

- If feature Ø fits at least one background model then current pixel=background pixel else current pixel is equal to foreground model.Σ^L_{i=1}θ_i(z)
- Phase feature of a particular pixel is calculated
 (Ø) is the probability of observing the present phase feature which is calculated as

 $P(\emptyset) = \sum_{i=1}^{k} \omega_i N(\emptyset, \mu_i, \sigma_i)$

Where K = no. of Gaussians,

N = Gaussian distribution,

 μ i =mean value,

 σi = variance of i-th Gaussian,

 ω i = weight of ith Gaussian in the mixture ,

 $\sum_{i=1}^{K} \omega_i j = 1, \Phi$ denotes feature value b

Feature U is computed and compared with present set of k models to check for matching models for every new pixel. Matching condition is defined as the feature value within the standard deviation of the distribution.

- When the new phase feature and the mean value of the background model both fall into this singular area, the distance between them is computed as $2 \text{ L}\pi$ Dist, where Dist is the absolute value between the two values.
- If such models exist, the first matching one is selected for updating its parameters and the model weights are updated as:

 $\omega_i = (1 - \alpha_1)\omega_i + \alpha_1 M_k \alpha_1 M_k$

Where $\alpha_1 \in (0,1)$ is the 1st learning rate and for the first matching Gaussian $M_{k}=1$ and is zero for others⁴.

- Blob aggregation is performed using distance transform and morphological operations is used¹.
- In order to make reduction in the processing time of the frames we involve the HAAR transform for multiple levels (1, 2 and 3). Each frame is transformed using HAAR operator before applying the above mentioned Gabor filter based phase computation.
- The HAAR transform reduce the area of the frame by one fourth in every level, thereby permitting a

very easy mapping of threshold for background foreground discrimination of the objects.

4. Experimental Results

The entire process is implemented using MATLAB running on a 64 bit windows 10 system.. The code records the processing time for each of the frames at multiple HAAR transform level using the tic and toc commands of MATLAB on a frame by frame basis.

The process is designed with the intuitive approach that as the HAAR transform performs the compression of low-spatial frequency information into the top left corner quadrant of the frame and only high spatial frequency information into the rest 3 quadrants, the top left image provides a compressed frame with similar information as the original one. As the level of HAAR transform is increased the frame keeps on reducing by one fourth in area reducing the pixel count to be processed by the phase calculation module and then the actual foreground background identification. This leads to a reduced process time if a higher level of HAAR is used before the phase processing.

We observe that the process shows some jumps in the timing for the frame processing which is the result of the impact of the memory allocation done by the operating system for all the programs as per the program demands as well as background process. It is observed that the initial frames are more susceptible to this kind of noise and hence we had used the values after first 80 frames for comparative study. Starting from this frame 80 a span of 45 frames till frame 125 was used for multiple files for the study.

The graph now is seen to be less noisy as compared to the previous ones and also clearly indicates that even though the process itself has interference from OS and other programs still the higher level provides better processing time for the frame. Table 1 shows the timings obtained for the three files used for evaluation of the scheme.

The Average time comparison is shown in Table 2, it is observed that as the level is increased the average processing time for the frame reduces.

Figure 1 presents the comparative graphs for the average time, it is interesting to note that there is a very wide variation in the average time for each of the files, but as the HAAR level is increased there is reduction. This indicates that the video contents do dominate the process in terms of actual time Figures 2 to 6. The time saved is presented in the Table 3.

File 1:	VIRAT_1.mp4			
	Level 1	Level 2	Level 3	
Total				
Time (ms)	1373.5777	1296.8135	1247.9823	
Avg (ms)	29.86038478	28.19159783	27.13005	
Tot. Saved				
%	0	5.58863179	9.684923854	
Per frame Saved				
(ms)	0	1.668786957	2.730334783	
File 2:	VIRAT_S.mp4			
	Level 1	Level 2	Level 3	
Total				
Time (ms)	2489.1681	1932.7617	1548.7271	
Avg (ms)	46.96543585	36.46720189	29.22126604	
Tot. Saved				
%	0	22.35310665	48.65788679	
Per frame Saved				
(ms)	0	10.49823396	17.74416981	
File 3:	visiontraffic.avi			
	Level 1	Level 2	Level 3	
Total				
Time (ms)	341.0392	306.1063	300.3951	
Avg (ms)	7.413895652	6.654484783	6.530328261	
Tot. Saved %	0	10.2430747	13.27777311	
Per frame Saved				
(me)		0 759/1087	0 883567391	

Table 1. Processing time using HAAR transform formultiple files



Figure 1. Comparison of the timing for multiple HAAR levels.

The output for various cases are presented below for a sample video is presented.



Figure 2. Ten frame average comparison.







Figure 4. Output for frame no. 82 with no object detected.

Table 2. Average	time using	HAAR	transform
(multiple files)			

Average Time (ms)	File 1	File 2	File 3
Level 1	29.86038	46.96544	7.413896
Level 2	28.1916	36.4672	6.654485
Level 3	27.13005	29.22127	6.530328

Table 3. Time saved percentage over level 1 usingHAAR transform (multiple files)

Saved time % over Level 1	File 1	File 2	File 3
Level 2	5.58863179	22.35310665	10.2430747
Level 3	9.684923854	48.65788679	13.27777311



Figure 5. Frame no. 101 results with one object detected.



Figure 6. Output for frame no. 118 with one object detected.

5. Conclusion

Despite the fact that there is no uniformity in the processing times of the various videos it is observed that, a pre-processing step applied to the phase based foreground background separation process helps in reducing time. A reduction of 5% to 2 % of processing time is observed at level 2 of HAAR while a saving of 9 to 48% of time is observed for level 3 HAAR. This enhances the speed of object detection in changing illumination conditions.

6. Future Work

Apart from involving other transforms in the spatial domain like wavelets, the work can be expanded by involving use of HOG features as a step for pre-processing prior to the phase based foreground background detection. The design can be accommodated on any targeting FPGA device to design a reconfigurable processor.

7. References

- Xue G, Sun J, Song L. Background subtraction based on phase feature and distance transform. PatternRecognition Letters. 2012; 33(12):1601–13. https://doi.org/10.1016/j. patrec.2012.05.009
- Ugur Toreyin B, Enis Cetin A, Aksaya A, Bilgay Akhan M. Moving object detection in waveletcompressed video. Signal Processing: Image Communication. 2005; 20(3):255–64. https://doi.org/10.1016/j.image.2004.12.002
- A proposed approach for image compression based on wavelet transform and neural network. 2017.https://thesai. org/Publications/ViewPaper?Volume=8&Issue=9&Code=I JACSA&SerialNo=59
- Singh N, Sabnis M. Object detection and tracking in moving background under different environmentalconditions. IOSR Journal of Computer Engineering (IOSR-JCE). 2016; 18(3):116–9.
- Komagal E, Sarah Sweetlyn Christella B. Moving object detection under various illuminationconditions for PTZ cameras. Indian Journal of Science and Technology. 2016; 9(45):1–7.
- 6. Object tracking under illumination variations using 2D-cepstrum characteristics of the target. 2010.https://ieeexplore.ieee.org/document/5662076
- A comparison of PCA and ICA for object recognition under varying illumination. 2002. https://ieeexplore.ieee. org/document/1047783
- Illumination condition effect on object tracking: A review. 2014. file:///C:/Users/a/Downloads/ 1097-1-1095-1-10-20150217.pdf
- Ugur Toreyin B, EnisCetin A, Aksay A, Bilgay Akhan M. Moving object detection in wavelet compressed video. Signal Processing: Image Communication. 2005; 20(3):255–64. https://doi.org/10.1016/j.image.2004.12.002
- 10. HAAR 2010. https://arxiv.org/abs/1010.4084
- A new image compression algorithm using HAAR wavelet transformation. 2013. https:// www.semanticscholar.org/ paper/A-New- Image- Compression- Algorithm- using- Haar-Mehala/ b84059909b99bf1 636f601af 77f2a22a6a 7136a0
- 12. FPGA implementation of image block generation and color space conversion for the Gaussian mixture model. 2017. https://ieeexplore.ieee.org/document/8081895
- 13. Ghorpade A, Katkar P. Image compression using HAAR transform and modified fast HAAR wavelet transform. International Journal of Scientific and Technology Research. 2014; 3(7):302–5