

A NOVEL PROBABILISTIC BASED IMAGE SEGMENTATION MODEL FOR REALTIME HUMAN ACTIVITY DETECTION

D.Ratna kishore¹, Dr. M. Chandra Mohan² and Dr.Akepogu. Ananda Rao³

¹Asst.Prof, Department of Computer Science & Engineering, JNTUA, Andhrapradesh, India

²Professor, Department of Computer Science & Engineering, JNTUH, Hyderabad, Telangana State, India

³Professor in CSE, Director, Academic & Planning, JNTUA, Anantapuramu, Andhrapradesh, India

ABSTRACT

Automatic human activity detection is one of the difficult tasks in image segmentation application due to variations in size, type, shape and location of objects. In the traditional probabilistic graphical segmentation models, intra and inter region segments may affect the overall segmentation accuracy. Also, both directed and undirected graphical models such as Markov model, conditional random field have limitations towards the human activity prediction and heterogeneous relationships. In this paper, we have studied and proposed a natural solution for automatic human activity segmentation using the enhanced probabilistic chain graphical model. This system has three main phases, namely activity pre-processing, iterative threshold based image enhancement and chain graph segmentation algorithm. Experimental results show that proposed system efficiently detects the human activities at different levels of the action datasets.

KEYWORDS

Human activity detection, Graphical models, Markov model, probability density distribution

1. INTRODUCTION

Human activity segmentation not only based on the human activity trained data, but also based on their domain knowledge such as the connectivity, contextual information, shape and the local homogeneity. Probabilistic graphical models in the machine learning are powerful statistical methods, that are used to find the image entities, their prior knowledge along with the uncertainties. Basically probabilistic graphical models are two types : the directed acyclic graph based model and the undirected graph based model. The undirected graph based model is used to find the entity feature extraction among the stratified random variables. The hidden markov model is the well known undirected graph based model for image segmentation. But, the markov random field(MRF)[1] model incorporates the spatial relationships among the graph nodes. So, an improvement to this MRF model is a conditional random field (CRF) is also an undirected graph based model applicable for image segmentation. Traditionally, a human activity can be detected

by using color based segmentation such as k-means, agglomerative and histogram approaches with edge detection algorithms. Digital image segmentation plays a significant role by partitioning an image into similar objects based on domain knowledge. Feature selection is an important requirement in most of the digital image segmentation techniques. Depending on the extracted features, the segmentation approaches are classified into 3 categories, namely, edge-based, region based and thresholding based segmentations.

Fig 1, represents the basic human activity recognition process. Here, image filtering methods are applied to remove noise or to optimize the brightness of the high resolution images. In the segmentation step traditional techniques like Graph based segmentation, K-means, Fuzzy-Means, etc. are used. Finally, the number of human activities are detection in each regions.

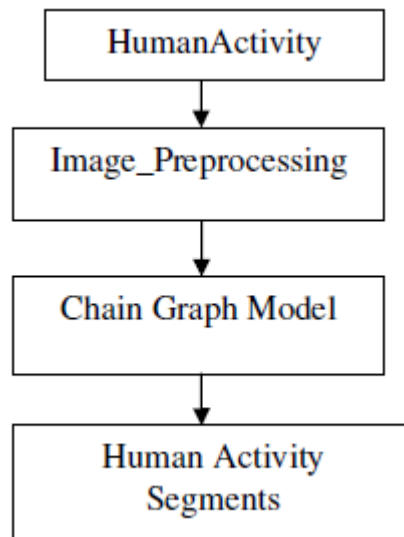


Fig.1 Traditional Human activity Segmentation Process

This paper provides the study of segmentation approaches for human activity recognition. Traditional study has limitations in feature selection and over segmentation.

Image thresholding: Thresholding based mechanism is based on choice of threshold and determines an intensity value to classify the image into desired values called a threshold. Failing to find such a threshold may lead to poor classification or segmentation[2].

Activity Segmentation: Segmentation is dividing the image data into groups of similar object classes. Each group indicates the cluster, which may contain the objects of similar type or features. The performance of the traditional segmentation algorithms is based on initial random assignments and similarity metrics.

Level Set Segmentation: Level set models use dynamic boundary variations for image segmentation. In these models, initial segmentation is performed using one of the traditional algorithms such as K means, FCM, KFCM etc. with trail and error basis. The parameters of the level set segmentation control the initial gradient level set function and spread of Gaussian smoothing function[2-3].

Problem Statement

The segmentation of human emotions or expressions are challenging factor in many realtime human activity identification. It is common in many domains such as biometric identification and medical fields which exhibit irregular boundaries and discontinuities. Also, segmented sub-regions can only be partitioned when several modalities are merged, which requires an efficient registration process in the preprocessing step. In the traditional human activity detection models, intra and inter region resolutions may affect the segmentation accuracy. Due to the partial volume effect, inherent noise, wide range of image features and spatial constraints, conventional models fail to detect the exact location and segmentation of the human activities in large training datasets[3,4].

The rest of this paper is organized as follows, Section II describes the literature study of segmentation models, threshold based models, classification models, Section III describe the chain based graph models, Section IV describes the experimental results and in Section V, we conclude with the model.

2. RELATED WORK

The traditional graph based models for image segmentation are based on either directed graphical or undirected graphical models. But, they often fail to capture the complex region relationship of different activities. In [5], color texture analysis to extract spectral correlation features is proposed. Both approaches doesn't handle texture boundary extraction due to noise in the image. A more improved approach is proposed in the Texem model , which consists of a conditional dependency between neighbor regions and it is totally based on Gaussian mixture model. Digital image thresholding is one the most popular method due to its simple implementation and intuitive features. Consequently, the threshold measures for multimodal histograms must be minima among the two maxima. A few techniques enlighten the histogram peaks in image feature extraction stage so as to provide the threshold detection. The main drawback of this method is to segregate the object from background if the object and background regions are identical pixel distribution. Edge based segmentation works well against irregularities in image attributes such as texture, gray level, color etc. These irregularities are known as edges and are noticed using edge detection operations, some of the generally used operations are prewitt, laplace , sobel etc. Segmentation using edge-based method could be used as incomplete due to the occurrence of stray, broken or noisy edges. Enhanced image processing is essential to extract the edges corresponding to rational objects. Several edge based methods have been proposed in the literature, but the frequently accepted segmentation systems are edge based thresholding, which is used to clear noisy edges in bright conditions. Edge image thresholding directs to stray edges in the presence of noise where the actual edges are frequently missing[6][7]. Stray edge problems can be cleared if the edge properties are decided with respect to the mutual neighbor, while the presence of edge based on the strength of edges in the near neighborhood. Region based segmentation approach which depends on the homogeneity measure to divide and merge regions in an image so as to broaden semantic or useful division in the processed image.

Conventional Human Activity Detection Methods

The region growing model is proposed in C-V segmentation. For a given frame I on the domain, C represents the closed curve region and i_1 an i_2 represents the average inside and outside

intensities of the closed curve C . The basic assumption of this segmentation by optimizing the energy function as:

$$E(i_1, i_2, C) = \mu_x \cdot len(C) + \omega \cdot area(C) + \theta_1 \int_{ins(C)} |I - v_1|^2 dx dy + \theta_2 \int_{ins(C)} |I - v_2|^2 dx dy$$

Where $\mu_x > 0$, $\omega > 0$, $\theta_1, \theta_2 > 0$ are fixed values.

The partial differentiation equation corresponding to the given C-V model is evaluated as

$$\partial \varphi / \partial t = \frac{df(x)}{dx} [\mu_x \text{div}(\nabla \varphi / |\nabla \varphi|) - \omega - \theta_1 (I - v_1)^2 + \theta_2 (I - v_2)^2]$$

Where

$$v_1 = \int_{\Omega} I \cdot f(x, y) dx \cdot dy / \int_{\Omega} f(x, y) dx \cdot dy$$

$$v_2 = \int_{\Omega} I \cdot (1 - f(x, y)) dx \cdot dy / \int_{\Omega} (1 - f(x, y)) dx \cdot dy$$

$$f(x) = 1 \text{ if } x \geq 0$$

$$= 0 \text{ if } x < 0$$

This model assumes that the image intensity values in each adjacent block always uniform. Self Organizing Map (SOM), as part of competitive learning neural network technique has been used to develop the vector quantization process. The role of SOM for vector quantization process is mainly due to the pixel similarity between the region learning applied in the self organizing map. Neural units in the competitive layer need to be nearly equal to the number of regions specified in the segmented image. This is the main disadvantage of traditional SOM for image segmentation.

An optimization to C-V model was implemented by Zhang et al[8] which consider intensity data to find the refined edges. Thus, using gradient information , the gradient based C-V model is proposed. The basic optimization in GCV model is:

$$E(v_1, v_2, ds_1, ds_2, \eta) = \mu_x \int_{\Omega} 0.5(|\nabla \eta| - 1)^2 dx dy +$$

$$+ \theta_1 \int_{ins(C)} |I(x, y) - v_1|^2 dx \cdot dy + \theta_2 \int_{ins(C)} |I(x, y) - v_2|^2 dx \cdot dy +$$

$$\int_{\Omega} \int_{ins(C)} |I(x, y) - v_1|^2 dx \cdot dy + \theta_2 \int_{ins(C)} |I(x, y) - v_2|^2 dx \cdot dy$$

The HSOM straightly addresses the drawbacks of the traditional SOM. HSOM is the mixture of self organizing and topographic mapping technique. HSOM combines the idea of data abstraction and independent of image features.[2] implemented an automatic image segmentation technique using thresholding approach. This is based on the pre-assumption that neighbor pixels whose

value (color, value, gray level, texture, etc.) lies within a certain range belongs to the identical class and good segmentation results include only two opposite components can be achieved. [9] implemented the edge detection and thresholding as one of the main aspects of image segmentation comes prior to image recognition system and feature extraction for analyzing human activities. In this approach, an edge detection and thresholding techniques are implemented on different medical images, geo images to quantify the stability of error rates. This thresholding algorithm is used to detect edges and removing noisy regions using histogram peaks. [6] Implemented the image analysis and comparison of these methods are executed based on the assessment criteria and classification method to analyze performance metrics. [7] Implemented the thresholding methods on scalar images using binary partitioning and image intensities. Segmentation is the process of grouping all related intensity pixels more than the user specified threshold into one class, and all other pixels into second class is called multi thresholding. In the paper [8], a combination of wavelet co-occurrence texture and wavelet statistical features have been used to segment the abnormal human activities. They proposed the framework in four phases: Segmentation, Discrete wavelet decomposition, feature selection and classification. However, the main problem of this system is that, it need new training for SVM classifier whenever there is change in image activity. As the distribution of the intensity of each activity is complex, traditional methods become difficult to process.

In most of the human activity images, multiplicative noise is involved with unknown noise variance. Most of the traditional denoising models work for the additive noise, so in order to convert the multiplicative noise to additive logarithmic function is applied to the image as follows.

$$\log(\text{mul}(x,y))=\log I(x,y)+ \log(\text{addnoise}(x,y))$$

Problems in Traditional Human activity segmentation Algorithms:

- Conventional Segmentation algorithms fail to differentiate the human activities with the corner edges.
- Each activity depends on the histogram and user specified threshold. Need to specify a dynamic or global and local threshold.
- Fails to identify the symmetrical activities due to noisy pixels in the central region.
- Fails to detect new types of activities in the test data.
- Difficult to segment the human expressions using the activity list.

Probabilistic Chain Graphical Model:

Each probabilistic graph model consists of two phases: quantitative and qualitative parts. Each probabilistic graph model is denoted by a graph $G=(V,E)$; where V is the set of vertices and E is the edge set. Each node in the vertices set represent the random variable and the link or edge represents the prior probabilistic relationships among the node random variables. In the novel chain graphical model , graph contains both directed and undirected edges[9]. Each edge in the novel chain graphical model represents the joint probability distribution among the random variables. Let x_1,x_2,x_3,x_4 and x_5 represents the random variables in the novel graph model, then the probability computation can be derived using the following equations[9][10] as shown in Fig 2.

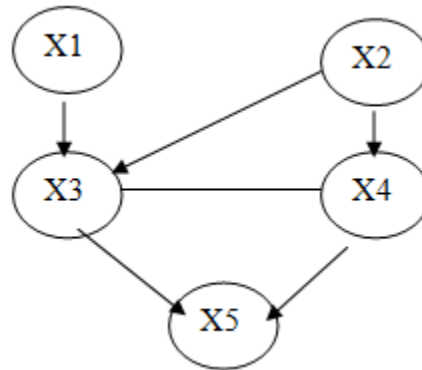


Fig 2: Directed Graphical Model

$$\phi(X 3) = Prob(x 3 / x 1, x 2)$$

$$\phi(X 4) = Prob(x 4 / x 2)$$

$$\phi(X 5) = Prob(x 5 / x 3, x 4)$$

$$Prob(x) = Prob(x 1, x 2). Prob(x 3, x 4 / x 1, x 2)$$

$$* Prob(x 5 / x 3, x 4)$$

3. PROBABILISTIC BASED HUMAN ACTIVITY DETECTION MODEL

The conditional dependence represents the relationship of the occurrence of one object being dependent on the occurrence of the other object and asymmetric in nature. Also, the homogeneous association usually computes the similarity among the objects.

The main aim of the segmentation algorithm is to partition an image into homogeneous regions (classes) with respect to one or more characteristics. Proposed segmentation algorithm is based on probabilistic estimation of the human action and non-actions regions. In this human activity detection algorithm, the basic chain graph algorithm is applied on the preprocessed image for initial segments. Initial segments have a large number of over segmented regions which are very difficult to detect or process the activities in the image. In order to overcome this issue, the probabilistic estimation method is applied in the over-segmented regions to extract the activities by merging the estimated regions[9].

Preprocessing

Preprocessing of human action images is the primary step in real time detection, to minimize noise and to optimize the image quality. An adaptive median filter is used to enhance the image quality as well as remove the poison noise from the images. In the median filter, a window moves along the image and the computed median value of the window pixels becomes the output. It preserves the edges and reduces the noise in the image. Each pixel is replaced with the median value of the neighborhood of the input pixels as shown in Fig 3.

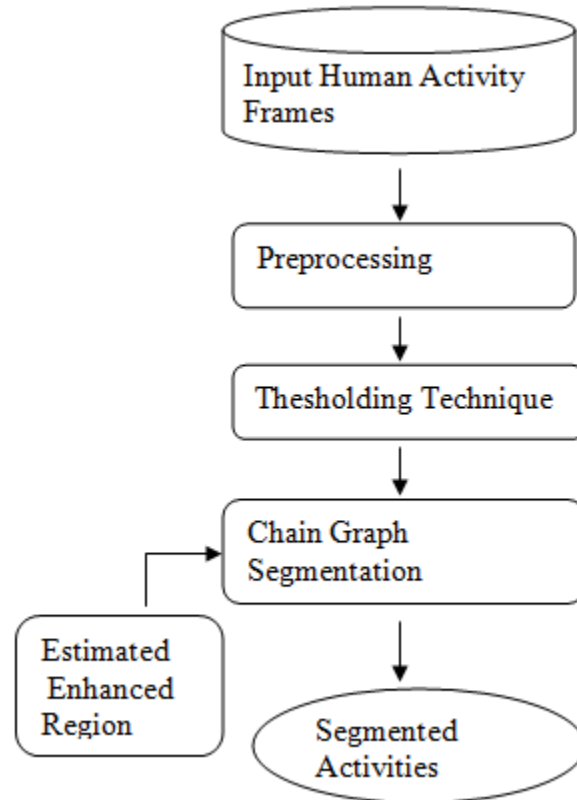


Fig 3: Enhanced Chain Graph Model

Iterative Threshold Algorithm

Step 1: Divide the human activity frame into two classes Cls_1 and Cls_2 with foreground level fl and background level bl with the region mean value Th such that

$$bl = \{Th+1, Th+2, Th+3 \dots N-1\}$$

where N is the total number of levels.

$$Th = \min\left\{\sum_{x=0}^{N-1} x^2 * prob(x)\right\}$$

$$\text{where } prob(x) = v_x / |v|$$

$$fl = \{0, 1, 2 \dots Th\} \text{ and}$$

$$prob(x) \geq 0,$$

$$\sum_{x=0}^{N-1} prob(x) = 1$$

Step 2: Determine the foreground level lower threshold using C-V model.

The variance of foreground level $fl = \{0, 1, 2 \dots Th\}$ is

$$\frac{\partial \phi}{\partial t} = \frac{df(x)}{dx} [\mu_x \text{div}(\nabla \phi / |\nabla \phi|) - \omega - \theta_1 (I - v1)^2 + \theta_2 (I - v2)^2]$$

$$v1 = \int_{\Omega} I \cdot f(x, y) \, dx \cdot dy / \int_{\Omega} f(x, y) \, dx \cdot dy$$

$$v2 = \int_{\Omega} I \cdot (1 - f(x, y)) \, dx \cdot dy / \int_{\Omega} (1 - f(x, y)) \, dx \cdot dy$$

Where

$$f(x) = 1 \text{ if } x \geq 0$$

$$= 0 \text{ if } x < 0$$

$$T_1 = \frac{\partial \phi}{\partial t} * \sum (x * prob^2(x) - x^2 * prob) / |N|$$

Step 3: Determine upper threshold .

$$T_2 = \sum_{x=Th+1}^{N-1} prob(x) * \sum_{x=Th+1}^{N-1} (x * prob^2(x) - x^2 * prob(x)) / |N|$$

Step 4: Compute the intra and inter activity variance of each block

$$\sigma_{intercls_1}^2 = \sum_{x=0}^{T_k} (x - (\sum_{x=0}^{T_k} x * n_x / |n_x|)) * (\sum_{x=0}^{T_k} n_x / |n_x|)$$

$$\sigma_{intercls_2}^2 = \sum_{x=T_k+1}^{N-1} (x - (\sum_{x=T_k+1}^{N-1} x * n_x / |n_x|)) * (\sum_{x=T_k+1}^{N-1} n_x / |n_x|)$$

Step 5: Intra-class variance can be calculated using

$$\sigma_{intra}^2 = \sigma_{intercls_1}^2 + \sigma_{intercls_2}^2$$

Step 6: Repeat steps 2 to 5 until the condition is satisfied.

After enhancing the given image using a threshold based algorithm, the next step is to find the human activities in the enhance image using segmentation algorithm.

Human Activity Segmentation Algorithm

Optimized Region Merging Activity Segmentation Algorithm

Input: Image threshold image.

$e(x)$ =edgeset values along with probabilistic values of first region.

$e(y)$ = edgeset values along with probabilistic values of first region.

$p(x)$ =intensity values of first region.

$p(y)$ =intensity values of second region.

Output : Enhanced Human Activity Segments.

Procedure:

Segmented regions S_i ;

Human activities estimated over segmented Regions S'_i generated from chain graphical model.

Segmented_Regions[]=Apply Chain Graphical Model Algorithm

For each region reg in Segmented_Regions

Do

If (reg is oversegmented)

Then

S_i =reg;

End if

done

Set i=0.

For each pair of adjacent oversegmented activities in S_i

do

Calculate the similarity between the oversegmented activities (O_i, O_j) along with the area as

$$L(i, j) = (1 / 2\pi) * e^{-|O_i - O_j|^2 / 2 * \max\{\sigma_i^2, \sigma_j^2\}}$$

Done

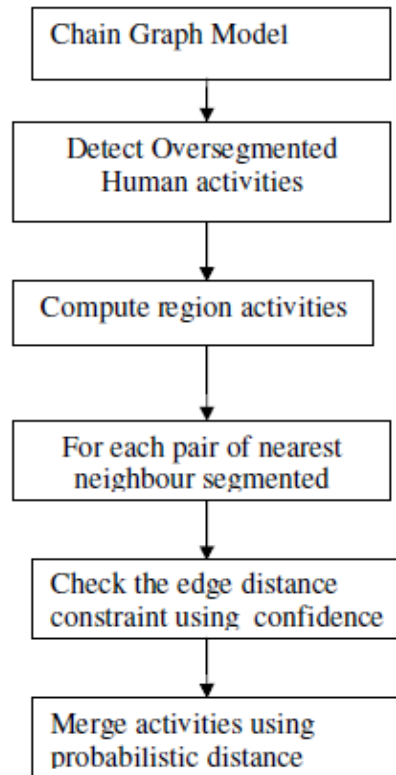


Fig 4: Chain Graph based Activity Segmentation Algorithm

For each pair of nearest neighbor over-segmented activities

do

$$R(A)=\text{RegionrArea}(A), R(B)=\text{RegionArea}(B).$$

Calculate mean similarity value of all the over-segmented activities in List L as θ

$$\theta = L(i, j) / N;$$

Where N is the number of oversegmented regions.

$$\text{If} ((R(A) \parallel R(B)) < \theta)$$

Then

$$\text{prob}(p(x)) =$$

$$\eta = \text{prob}(p(x) / p(y))$$

$$\eta = \frac{1}{(2\pi)^{N_1} * \sqrt{\text{cov}(p(x), p(y))}} * e^{-(1/2) * N_2}$$

The variance of edge sets x is evaluated as :

$$\partial e(x) / \partial t = \frac{df(x)}{dx} [\mu_{e(x)} \text{div}(\nabla \varphi / |\nabla \varphi|) - \omega - \theta_1 (I - v_1)^2 + \theta_2 (I - v_2)^2]$$

$$v_1 = \int_{\Omega} I \cdot f(x, y) \, dx \cdot dy / \int_{\Omega} f(x, y) \, dx \cdot dy$$

$$v_2 = \int_{\Omega} I \cdot (1 - f(x, y)) \, dx \cdot dy / \int_{\Omega} (1 - f(x, y)) \, dx \cdot dy$$

Where

$$f(x) = 1 \text{ if } e(x) \geq 0 \\ = 0 \text{ if } e(x) < 0$$

$$e'_x = \text{Max}\{\partial e(x) / \partial t, |E \cap \{p_x\}| / |\{p_x\}|\}$$

The variance of edge sets y is evaluated as :

$$\partial e(y) / \partial t = \frac{df(x)}{dx} [\mu_{e(y)} \text{div}(\nabla \varphi / |\nabla \varphi|) - \omega - \theta_1 (I - v_1)^2 + \theta_2 (I - v_2)^2]$$

$$v_1 = \int_{\Omega} I \cdot f(x, y) \, dx \cdot dy / \int_{\Omega} f(x, y) \, dx \cdot dy$$

$$v_2 = \int_{\Omega} I \cdot (1 - f(x, y)) \, dx \cdot dy / \int_{\Omega} (1 - f(x, y)) \, dx \cdot dy$$

Where

$$f(x) = 1 \text{ if } e(y) \geq 0 \\ = 0 \text{ if } e(y) < 0$$

$$\text{and } e'_y = \text{Max}\{\partial e(y) / \partial t, |E \cap \{p_y\}| / |\{p_y\}|\}$$

$$\lambda = |E \cap \{e_x \cup e_y\}| / (\min(e'_x, e'_y) * |\{e_x \cup e_y\}|)$$

Where

End if

If ($\lambda \leq \eta$)

Then

Merge($p(x)$, $p(y)$)

```
Else  
Break.  
End if  
done
```

4. PERFORMANCE ANALYSIS

The evaluation of the proposed system is carried out using standard performance measures which are of falsepositiverate(fpr), truepositiverate(tp), accuracy and errorrate, true positives(tp), false positives(fp), true negatives(tn), false negatives(fn). (fp+tn) indicates the total number of pixels represented foreground and background. Thus,in this paper the video sequences human activity 1 and 2, visionTraffic are adopted as the training sequences. Data sets like CAVIAR (2004), CASIA ACTION(2007) and ETISEO(2005) are used in our experimental study for human action recognition.

$$fpr = \frac{fp}{fp + tn} ; tpr = \frac{tp}{tp + fn} ;$$

$$accuracy = \frac{tp + tn}{tp + fn + fp + tn} ;$$

$$errorrate = 1 - accuracy$$

Existing Hybrid Probabilistic Model



Fig 5: Activity frame in atrium video



Fig 6: Detection of human activity using traditional model

```
=====
FPR 0.00225
TPR 0.90102
Accuracy 0.92688
Error Rate (%)2.45250
```

Proposed Model Results:

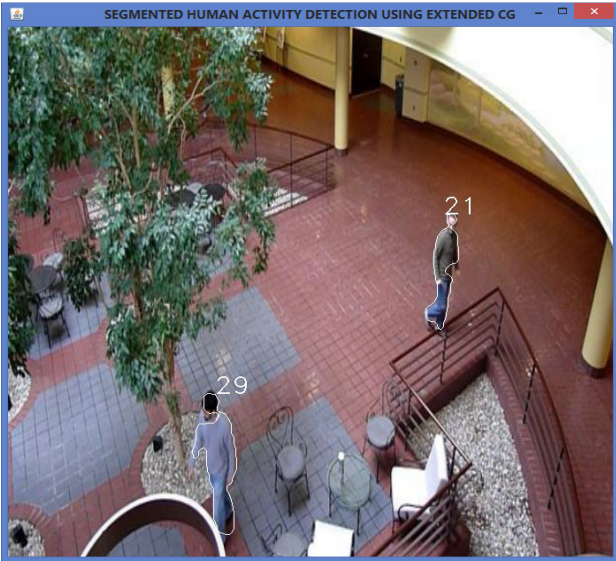


Fig 7: Activity frame in atrium video

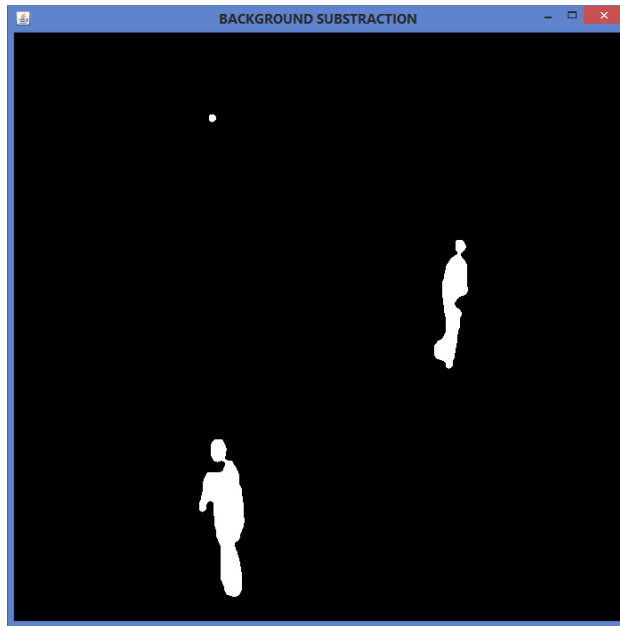


Fig 8: Proposed model human activity detection

Table 1: Avg False Positive Rate of the existing and proposed model

Dataset	Avg False Positive Rate(HPM)	Proposed Avg FPR
Human Activity 1	0.02254	0.0142
Atrium	0.03203	0.01153
Human Activity 2	0.025474	0.01253

Table 1 shows the average false positive rate of the hybrid probabilistic model and the proposed model. by observing the table it is clear that the proposed technique can perform better false positive rate compare to the existing technique.

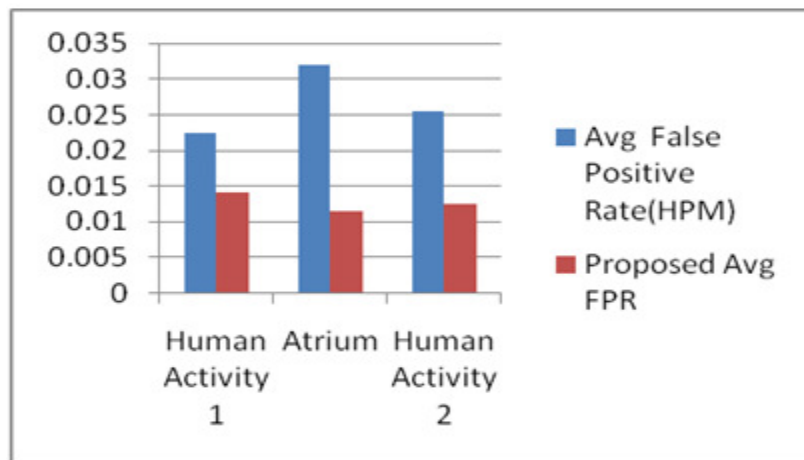


Fig 9: Avg False Positive Rate of the existing and proposed model

Fig 9 shows the average false positive rate of the hybrid probabilistic model and the proposed model. by observing the table it is clear that the proposed technique can perform better false positive rate compare to the existing technique.

Table 2: Avg true postive rate of existing and proposed models

Dataset	Avg TPR(HPM)	Proposed TPR
Human Activity 1	0.913	0.974343
Atrium	0.9434	0.98232
Human Activity 2	0.953344	0.9798

Table 2, shows the average true positive rate of the hybrid probabilistic model and the proposed model. by observing the table it is clear that the proposed technique can perform better false positive rate compare to the existing technique.

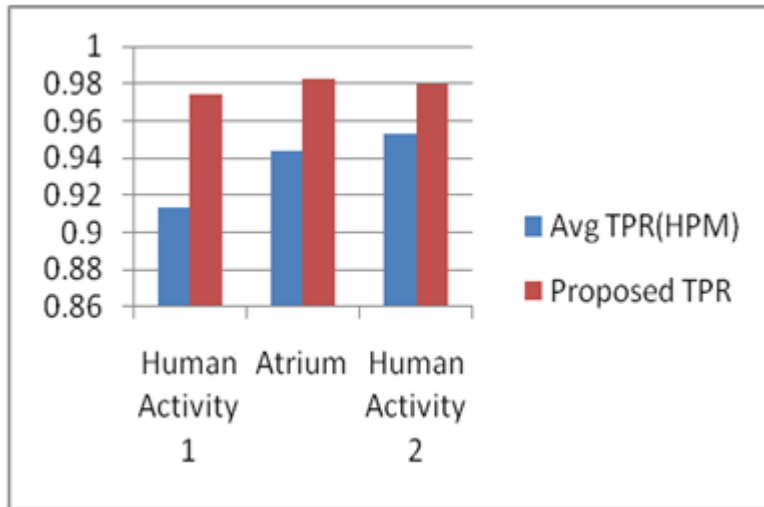


Fig 10. Avg true positive rate of existing and proposed model

Fig 10 shows the average true positive rate of the hybrid probabilistic model and the proposed model. by observing the table it is clear that the proposed technique can perform better false positive rate compare to the existing technique.

Table 3: Average Runtime comparison of existing and proposed models

Dataset	Execution Runtime (Existing)ms	Execution Runtime (Proposed)ms
Human Activity 1	4266	3855
Atrium	6756	4645
Human Activity 2	5756	3979

Table 3 shows the runtime comparison of existing and proposed model using the sample frames in the video. From the table, it is observed that proposed model has low runtime comparing to existing hybrid probabilistic model.

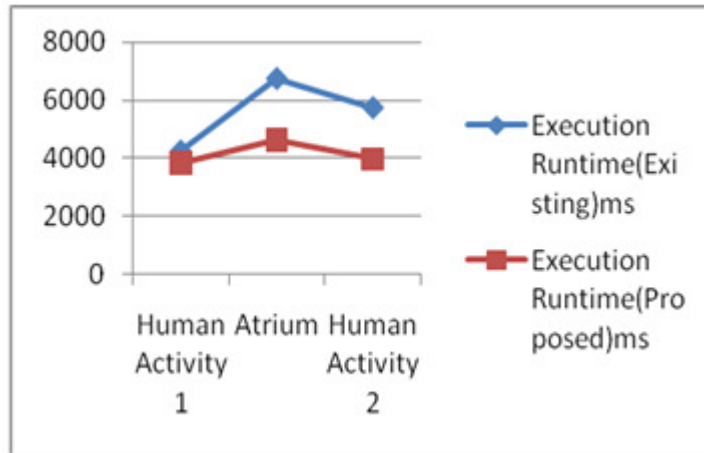


Fig 11: Average Runtime comparison of existing and proposed models

Fig 11, shows the runtime comparison of existing and proposed model using the sample frames in the video. From the table, it is observed that proposed model has low runtime comparing to existing hybrid probabilistic model.

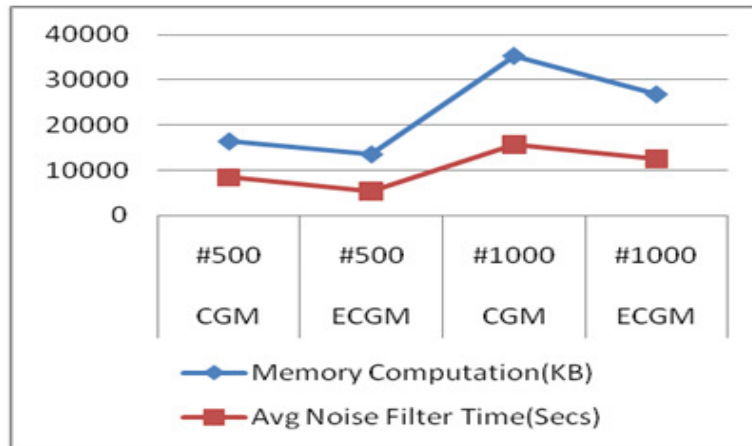


Fig 12: Comparison of storage and runtime in different sample video sequences

5. CONCLUSION

In this paper, we have summarized the various image segmentation approaches and its limitations on the human activity detection process. The study also reflects the various region growing models used for threshold based segmentation. In this paper, we have studied and proposed a natural solution for automatic human activity segmentation using the enhanced probabilistic chain graphical model. This system has three main phases, namely activity pre-processing, iterative threshold based image enhancement and chain graph segmentation algorithm. Experimental results show that proposed system efficiently detects the human activities at different levels of the action datasets. These models fail to differentiate the facial human activities due to noisy corner edges. In future, this work can be extended to detect various facial segmented expressions in 3D images using a novel probabilistic graphical segmentation model.

REFERENCENCES

- [1] C. Dharmagunawardhana, S. Mahmoodi, M. Bennett and M. Niranjana, "Gaussian Markov random field based improved texture descriptor for image segmentation", *Image and Vision Computing*, vol. 32, no. 11, pp. 884-895, 2014..
- [2] M. Zeng, T. Han, Q. Meng, Z. Bai and Z. Liu, "Image thresholding based on edge information analysis", 2012 5th International Congress on Image and Signal Processing, 2012.
- [3] S. Gupta and K. Ramakrishnan, "Learning Feature Trajectories Using Gabor Filter Bank for Human Activity Segmentation and Recognition", 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, 2008.
- [4] H. Zhang and L. Parker, "Fuzzy Temporal Segmentation and Probabilistic Recognition of Continuous Human Daily Activities", *IEEE Transactions on Human-Machine Systems*, vol. 45, no. 5, pp. 598-611, 2015.
- [5] B. Sun and J. He, "Discriminative dictionary based representation and classification of image texture", Sixth International Conference on Digital Image Processing (ICDIP 2014), 2014.
- [6] N. Wang, "Color Image Edge Detection Based on Cube Similarity", 2015 IEEE International Conference on Computational Intelligence & Communication Technology, 2015.
- [7] X. Zhang and C. Liu, "An ideal image edge detection scheme", *Multidimensional Systems and Signal Processing*, vol. 25, no. 4, pp. 659-681, 2013.
- [8] K. Zhang and K. Lam, "A Level Set Approach to Image Segmentation With Intensity Inhomogeneity", *IEEE Trans. Cybern.*, vol. 46, no. 2, pp. 546-557, 2016.
- [9] Lei Zhang and Qiang Ji, "Probabilistic Image Modeling With an Extended Chain Graph for Human Activity Recognition and Image Segmentation", *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2401-2413, 2011.
- [10] L. Herranz and Shuqiang Jiang, "A probabilistic model for food image recognition in restaurants", 2015 IEEE International Conference on Multimedia and Expo (ICME), 2015.