www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 4 Issue 3 March 2015, Page No. 10908-10913

A Proficient Extreme Learning Machine Approach For Tracking And Estimating Human Poses.

Dr.P. Tamije Selvy[1], T. Renuka Devi[2], R.Siva Keerthini[3], S.Umamaheswari[4]

[1] Associate Professor, Sri Krishna College of Technology, Coimbatore – 42, Tamilnadu, India
[1]tamijeselvy@gmail.com
[2,3,4] UG Student , Sri Krishna College of Technology, Coimbatore -42, Tamilnadu, India
[2]sakthirenu23@gmail.com,
[3]keerthini.chandraj@hotmail.com,
[4]umasivakumaar@gmail.com

Abstract: ***Tracking and Estimation of human pose in real time is a difficult problem with many interesting applications. Automated tracking is useful in variety of domains including human computer interaction, gait analysis, the film industry and entertainment. The existing system uses different algorithm to estimate and track human poses but the limitation is due to the error rate which is above 10%. In the proposed system effective filtering technique and background subtraction technique is used in order to remove clutters and noises. The objects in each frame are tracked and the corner points are identified. The corner points are used to identify objects that are human. Finally, Extreme Learning Machine classifier is used to identify and estimate the exact human pose from video.***

Keywords: Harris Corner Detection, Gaussian Mixture Model, Extreme Learning Machine, Grey Level Difference Method.

1. Introduction

Image processing is a field of study which deals with several applications such as human pose estimation, surveillance, etc. In Computer Vision, Human pose estimation is the key problem that has been in research for more than 15 years. Human pose estimation has its wide range of applications in Gaming, Human Computer Interaction, and Biometrics. Determining the exact human pose is a difficult task due to the cluttered background, high computational cost. This paper deals with identifying the human objects in a video sequence and to tracking and estimate the exact human pose using robust and powerful Extreme Learning Machine (ELM) Algorithm.

2. Literature Survey

The related work can be generally divided into the following four categories.

2.1 Exemplar-Based Approach

Labeled exemplars are stored and then matched to the test image. The pose of the test image is assumed to be the same as that of the most similar exemplar. Each exemplar is matched to a test shape in a 2-D view, using the technique of shape context matching in conjunction with a kinematic chain-based deformation model [1].When the number of exemplars is large, there is a huge computational cost.

2.2 Human Body Pose Estimation Using Silhouette Shaping Analysis

Top-down approaches can include many parameters and consequently explain the image observation accurately. This is termed an analysis by- synthesis approach. Yang and Ramanan [2] incorporated the co-occurrence relations between body parts into the pictorial structure model, and each part is represented by a mixture of templates. To overcome this bottle-neck, some researchers resort to state-space decomposition methods. Decentralization strategy fails if the projection of the 3-D model is not initialized properly.

2.3 Real Time Pose Estimation of Articulated Objects Using Low Level Motion

Srinivasan et al. [3] proposed a bottom-up parsing of complete body masks guided by a parse tree. At each level of the parsing process, the partial body masks are found by shape matching with exemplars. Daubney et al. [4] used sparse motion features and determined the probability that a certain movement region belongs to a specific body part. False detection of limbs results in inaccurate pose estimation and increases the computation time to check them.

2.4 A Hybrid Approach

This combines the merits of top-down and bottom-up approaches for pose estimation. Zhang et al. conducted a hybrid strategy combining both deterministic and stochastic search, where visual cues such as edge and skin are used to facilitate the searching process.

Motivation

Tracking and estimating human poses ia challenging task due to its cluttered background, poor illumination or partial self-occlusion. The traditional systems remove the clutters, poor illumination. But the performance is not efficient enough when the background of the images have many clutters also these systems can estimate the human poses only from static images. To overcome these short comings, a system is proposed which is much more efficient in identifying and estimating the human poses.
Here background subtraction technique is used to remove the clutters from the image and the ELM algorithm is used to extract human object from the images and to estimate human poses accurately. Section 3 describes the modules in the proposed human pose estimation system. Section 4 details about Performance Evaluation In terms of Execution time, Specificity, Sensitivity, Accuracy. The last section concludes the paper with the scope of future Enhancement.

3. Proposed System

The goal of the proposed system is to provide the system with greater accuracy in identifying the human poses. One of the key issues in identification is when the images are with cluttered background and poor-illumination. To overcome this problem background subtraction method is used. The images for the proposed system are taken from Weizzman Dataset. The system has the following phases I. Preprocessing using Histogram Equalization II. Background Subtraction using Gaussian Mixture Model. III. Identifying the objects from the frames using Monte Carlo and Harris Corner Algorithm. IV. Identifying the human objects using the features derived. V. Human Pose Estimation using Extreme Learning Machine Algorithm. Figure 1 gives the complete overview of the system.

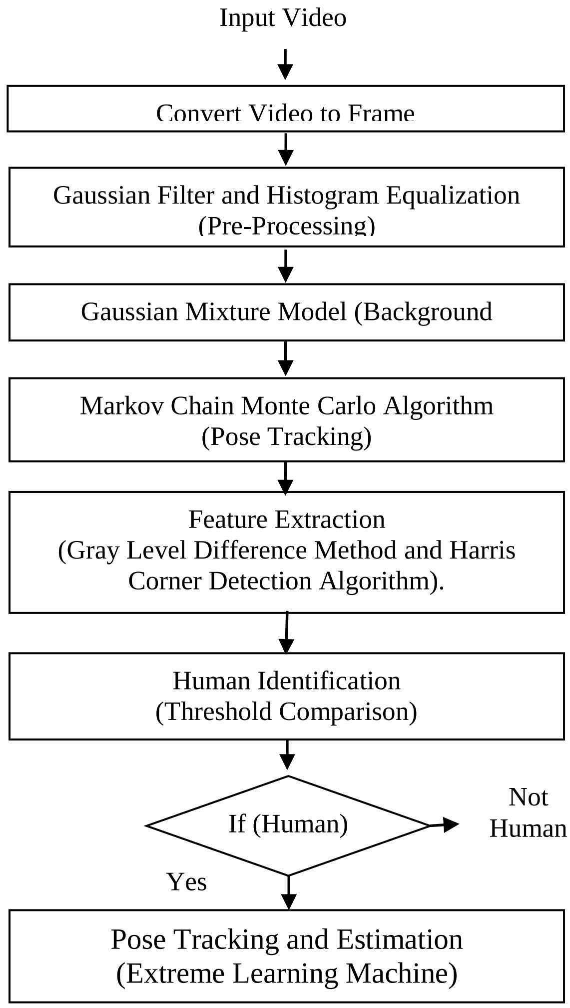

Figure 1: System Overview

3.1 Pre-Processing

Histogram Equalization is a method in image processing for contrast adjustment and for enhancement of images.The image is loaded into the dataset. Input image is resized and the color image is first transformed into a gray level image. Filtering technique such as mask filter and weighted average filter are applied for removing noise. Weighted average filter [R] is calculated using equation (1) gives more weight to the center value.

$$R=\begin{bmatrix}1 & 1 & 1\\ 1 & -8 & 1\\ 1 & 1 & 1\end{bmatrix} \quad \text{\%\% Filter mask} \qquad (1)$$

By using histogram equalization, the quality of the image is enhanced. The output of this module is shown in Figure 2.

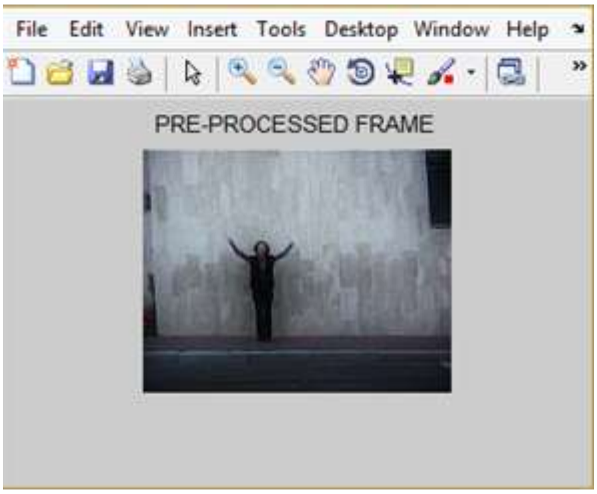

Figure 2: Pre-processing

3. 2 Background Subtraction

Background subtraction or Foreground detection is a technique in image processing where the unwanted background is subtracted and the region of interest is extracted. Background subtraction is a widely used technique for detecting moving objects from a static frame. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called "background image", or background model".

The first frame is stored as a background image. The current frame with the pre stored background image is subtracted and the difference is estimated. If the difference is greater than the bound threshold, then it determines that the pixel to pixel on the moving target, or as the background pixel. Threshold value is estimated using the equation (2) and (3).

$$R_k(x, y) = f_x(x, y) – B(x, y) \tag{2}$$

$$D_x(x,y) = \begin{cases} 1\ background\ Rk(x,y) > T \\ 0\ \ target\ Rk(x,y) \leq T \end{cases} \tag{3}$$

Where,

B(x,y) – Co-ordinates to represent background image.

$f_x(x,y)$ - Co-ordinates to represent the size of the first frame.

$D_x(x,y)$ - Co-ordinates to represent the image based on the threshold value

The output from background subtraction module is shown in figure 3.

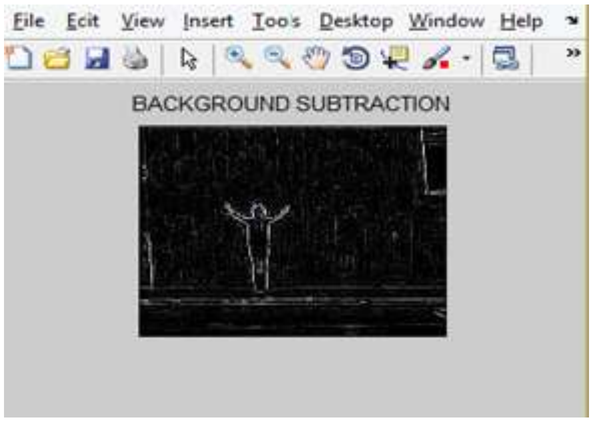

Figure 3. Background Subtraction

3.3 Object Tracking

Markov chain is a class for sampling from a probability distribution based on constructing a Markov chain that has equilibrium distribution.

The human body is modeled as a three level tree structure which consists of ten body parts: The root: torso, The second level: head, left/right upper arm, The third level: left/right lower arm, left/right lower leg. The pose of the human body is $X = \{X_1, X_2, \cdot\cdot\cdot, X_{10}\}$, where *Xi* represents an individual articulated part *i*. In detail, body part *i* is represented by a rectangle, and parametrically characterized as $Xi = \{x_i, y_i, \theta_i, l_i, w_i\}$ where x_i, y_i and θ_i are position and orientation relative to its parent, and l_i, w_i are length and width respectively. Object tracking is shown in Figure 4.

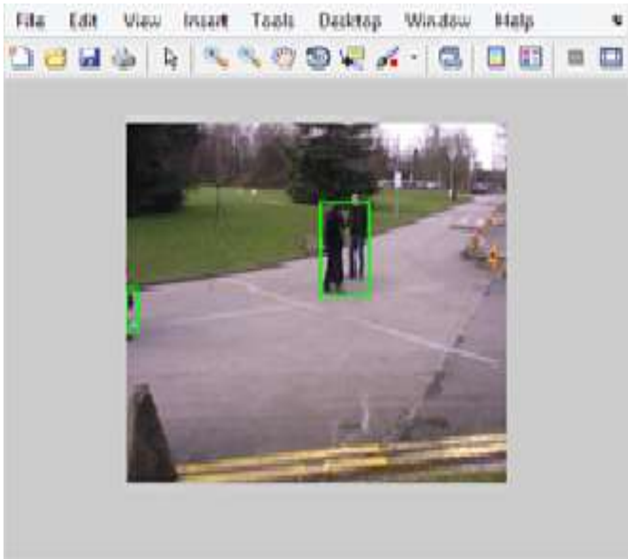

Figure 4: Object Tracking

3.4 Gray level difference method

In photography and computing, a gray scale digital image is an image in which the value of each pixel is a single sample that carries only intensity information. So, it is also called as black-and-white image. Grey level difference method is summarized as follows,

Step 1: For each pixel in image, get the different red, green and blue values of the pixel.

Step 2: Implement math operation to turn those different pixel values into single gray value as in equation (4).

$$Gray = (Red + Green + Blue) / 3; \qquad (4)$$

Step 3: Replace RGB (red, green, blue) value with new gray value.

3.5 Harris Corner Detection

In Computer Vision, corner detection is an approach to extract certain kinds of features and infer the content of the Image. The resultant image is shown in figure 5.

The steps involved in Harris corner detection are,

Step 1: Use gradient formulation as per equation (5) to detect response at any shift (x, y).

$$E(u,v)= \sum_{(x,y)} w(x,y)[I(x+u,y+v)-I(x,y)]^2 \qquad (5)$$

Where,

W(x,y) – Window Function

I(x + u,y + v) - Shifted Intensity

I (x,y) - Intensity Instead of 0-1 weight use Gaussian function w(x , y).

Step 2: Reformulation is calculated by using the equation (6)

$$M=\sum_{(x,y)} w(x,y) \begin{bmatrix} Ix2 & Ix\,Iy \\ Ix\,Iy & Iy2 \end{bmatrix} \qquad (6)$$

Step 3: Eigen values of M correspond to principle curvature of the local autocorrelation function is calculated by using the following equation (7).

$$R=\det M - k\,(\mathrm{trace} M)^2 \qquad (7)$$

Where,

k – Empirical constant (k = 0.04-0.06).

Step 4:

if (R is large) return corner,

if (R is negative with large magnitude) return edge,

if (|R| is small) return (flat region).

The output from this module is shown in figure 5.

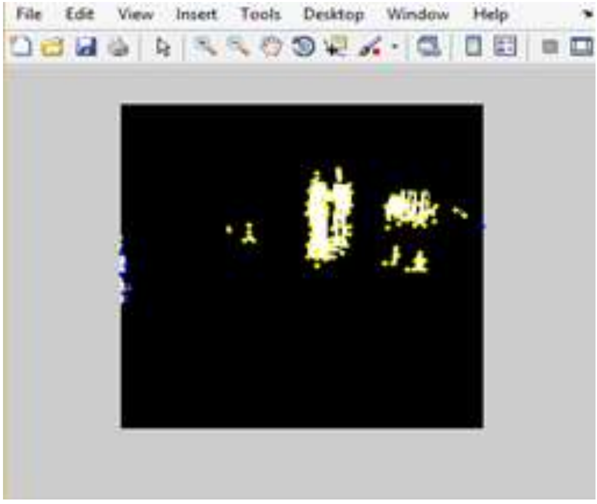

Figure 5: Feature Extraction

3.6 Human Identification

The tracked object can be any objects in the frame (say, car, man, animal etc,). In order to determine human follow the steps given below.

Input: Corner Points, Texture Points

Output: Human Identification

Step1: Find the mean value of the corner points and texture values in the frame.

Step 2: Compare mean with the threshold value.

Step 3: If the mean value is equal to threshold value, the tracked object is human otherwise it is not human. Figure 6 displays the object whether it is human or not.

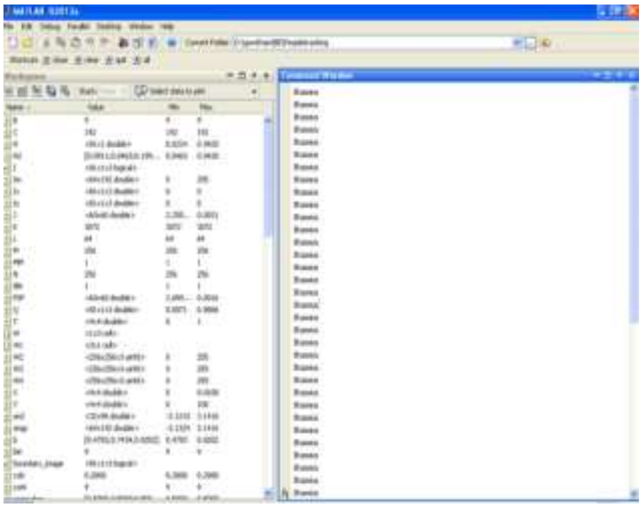

Figure 6: Human Identification.

3.7 Pose Tracking and Estimation

Extreme Learning Machine (ELM) is a supervised learning for a class of feed forward neural networks with random weights that has recently been used with success for the classification of hyper spectral images. It has its applications in character recognition, pose estimation, spam filtering, search engine and computer vision.,

Using classification technique, different poses are trained and tested using ELM algorithm [1].

Input: Corner points, texture values, foreground area values.

Output: Human Pose Estimation.

The following are the steps involved in identifying the human poses, and the resultant is shown in figure 7.

Step 1: Extract the human, corner points, texture points and foreground values.

Step 2: Train the values to estimate the poses using the dataset.

Step 3: For the subsequent frames,

- Test for human candidate using the trained human dataset.
- Recursively obtain the human for each of the frames.
- Determine pose of human in different frames.

The pose estimation result is shown in Figure 7.

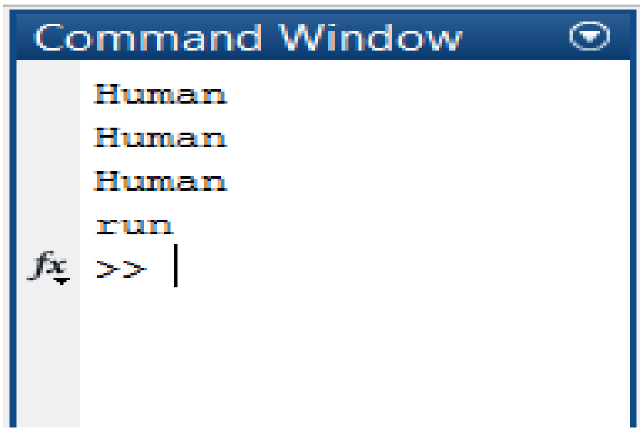

Figure 7: Pose Tracking and Estimation

4. Performance Evaluation

4.1 Execution Time

The duration of the project execution is evaluated and compared with the existing systems as in Table. 1 and Figure. 7 respectively.

Table 1: Execution time Comparison

S.No	Methods	No. of Iteration	Time (seconds)
1	Proposed algorithm	55	45
2	Markov Chain Monte Carlo algorithm	50	60
3	Exemplar Based algorithm	6	120

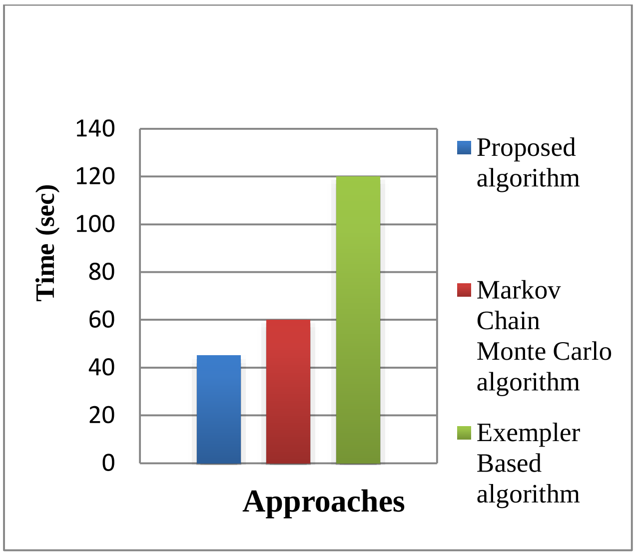

Figure 8: Execution Time

4.2 Accuracy Estimation

The performance measure of the algorithm is calculated by comparing the performance of SVM [5] and ELM algorithm. It is evaluated based on the three validation measures (i.e.) Sensitivity, Specificity and Accuracy. They are based on True Positive (TP), False Positive (FP), True negative (TN) and False Negative (FN). Sensitivity measures the ability of the proposed method to identify the abnormal cases. Sensitivity is calculated as per the formula (8). Specificity measures the ability of the method to identify normal cases; it is calculated from the following formula (9). The accuracy of the system is identified using (10).The performance of our system using ELM is compared with SVM and the comparison table and comparison chart is shown in the Table 2 and Figure 8 respectively.

$$\text{Sensitivity} = \frac{TP}{(TP+FN)} \quad (8)$$

$$\text{Specificity} = \frac{TN}{(TN+FP)} \quad (9)$$

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (10)$$

Table 2: Comparison Table for ELM and SVM

S.No	Method	Specificity (%)	Sensitivity (%)	Accuracy (%)
1	SVM	97.98	94.17	92
2	ELM	98.45	94.77	92.8

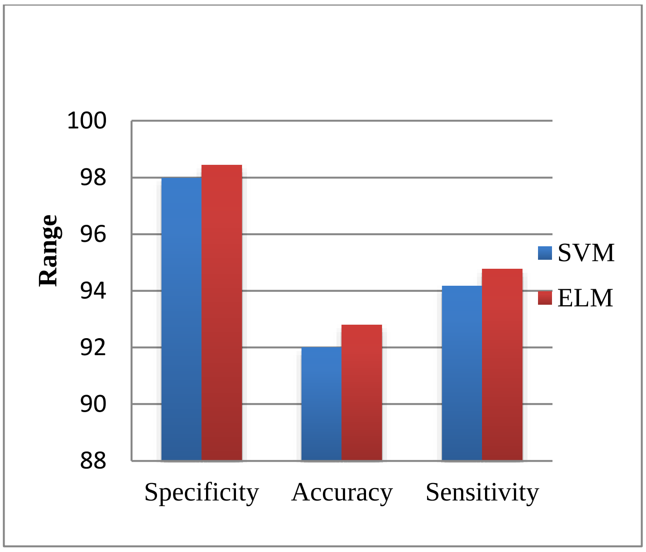

Figure 9: Comparison chart for ELM and SVM

5. Conclusion

The proposed system estimates the human pose effectively compared to earlier systems. The efficiency of the system is achieved through Gaussian filter and histogram equalization. The corner points of the objects identified by object tracking are used to identify the human. By using Extreme Learning machine different poses are estimated efficiently and effectively. The results from experiments carried out shows that the implemented human pose estimation using ELM approach is better than the earlier systems in terms of sensitivity, specificity and prediction accuracy. The system can be enhanced using optimization techniques such as Particle swarm optimization (PSO) and Ant Colony Optimization Technique (ACO).

References

[1] A. Faith and G. Mori, “Human pose estimation using motion exemplars,”in*Proc. Int. Conf. Comput. Vision*, 2007, pp. 1–8.

[2] A. Mittal, L. Zhao, and L. Davis, “Human body pose estimation using silhouette shape analysis,” in *Proc. Int. Conf. Adv. Video Signal Based Surveill.* 2003, pp. 263–270.

[3] B. Daubney, D. Gibson, and N. Campbell, “Real-time pose estimation of articulated objects using low-level motion,” in *Proc. Int. Conf. Comput. Vision Pattern Recognit.*, 2008, pp. 1–8.

[4] P. Srinivasan and J. Shi, “Bottom-up recognition and parsing of the human body,” in *Proc. Int. Conf. Comput. Vision Pattern Recognit.*, 2007, pp. 1–8.

[5] P.Tamije Selvy-(2012), An Improved GA-MILSVM Classification approach for diagnosis of breast lesions from stain images, International Journal of Advances in Engg& Tech., Sep. 2012,ISSN:2231-1963

[6] C. Stauffer and W. Grimson, “Adaptive background mixture models for real-time tracking,” in Proc. Int. Conf. Comput. Vision Pattern Recognit., vol. 2, 1999, pp. 246–252.

Author Profile

Dr.P. Tamije Selvy received B .Tech (CSE), M.Tech (CSE), and PhD (CSE) in 1996, 1998 and 2013respectively. She has been working as faculty in reputed Engineering Colleges. At present, she is working as Associate Professor in the department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore. Her Research interest includes Image Processing, Data Mining and Pattern Recognition.

Ms. T. Renuka Devi is currently pursuing Bachelor’s degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Image Processing and Data Mining.

Ms .R. Siva Keerthini is currently pursuing Bachelor’s degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Image Processing and Data Mining

Ms. S. Umamaheswari is currently pursuing Bachelor’s degree in Computer Science and Engineering at Sri Krishna College of Technology, Coimbatore. Her area of interest includes Image Processing and Data Mining