Human Body Posture Recognition Using Artificial Neural Networks

Manu Bali, Devendran V

Abstract-Activity acknowledgment has turned into a critical theme in recognizing the passionate action with numerous crucial applications. To abstract the concept, human detection is the very initial step to be performed in each field like surveillances, abnormal behaviour detection, crowd analysis and etc. In human association, understanding human practices is a testing issue in this day and age. Various poses of the body will be the challenging always for activity recognition. Hence, orientation / pose of the body detection become a vital one. In this work, Head is used as the key for body-pose recognition. Here, we present a framework, which is capable of finding the posture orientation of a human body in context of head position. Extracted features are used to train Artificial Neural Networks to obtain the Head Position. Our work consists of following steps: (i) Human Detection (ii) Upper Body / Lower Body Detection (iii) Head Detection (iv) Human Pose Recognition. This work is implemented in Matlab Platform. Datasets are taken randomly from Google Images.

Keywords –boundary extraction, upper and lower body segmentation, body-pose recognition.

I. INTRODUCTION

Activity recognition is an extraordinarily powerful research point in PC vision with various basic applications, including human PC interfaces, content-based video requesting, video perception, additionally, apply self-sufficiency, among others. Genuinely, visual action acknowledgment has been apportioned into sub-subjects, for instance, signal perceived for human PC interfaces, outward appearance acknowledgment, and development conduct acknowledgment for video perception. Anyway full-body exercises generally consolidate unmistakable developments what's more, require a united approach for acknowledgment, joining facial exercises, hand exercises and feet exercises [1].

Action acknowledgment is the path toward naming exercises, for the most part in the direct kind of a movement verb, using unmistakable discernments. Indeed, an action is a progression of advancements made by a human pro amid the execution of an undertaking. Along these lines, it is a four-dimensional protest, which may be additionally deteriorated into spatial and fleeting parts. Highlight extraction is the standard vision in real life acknowledgment and comprise in removing stance and movement prompts from the video that are discriminative regarding human activities [1]. Visual reconnaissance framework is utilized to screen a situation (or scene) with different camcorders. To comprehend the protest conduct continuously condition requires the improvement of PC vision calculations to distinguish, find and track the intrigued questions as it travels through nature. Astute visual observation framework is coordinated on consequently distinguishing occasion of intrigue, ready police when any risk found and make whole reconnaissance framework naturally. For instance, reconnaissance frameworks assume a noteworthy part in wellbeing observing for elderly maturing individuals at home. Likewise it's useful for watching human conduct and distinguishes illicit activities and occasions in broad daylight and private spots. In mind boggling and jumbled scene with direct number of moving articles (e.g. 10-15) the issue of identifying and following the articles is essentially muddled by impediment, where a question might be in part blocked or thoroughly vanish from camera see. The essential rationality of question division is to isolate the intrigued protests in unlabeled recordings from the foundation scene naturally [2].

II. RELATED WORK

Shirbhate et al. grouped feelings utilizing semantics-based approach. Encourage the activities are delegated sitting stance and standing stance. Here, sitting stance finishes up action to be perceived as either loose or hands on temple (strained). While standing stance finishes up movement perceived as lingering or fidgetting. Creators have decided on semantics-based approach rather than machine learning empowers us to identify the activities without requiring to prepare the framework. This additionally improves the framework execution shrewd; and empowers activity discovery progressively [1]. Vaniya et al. examined the different techniques and additionally calculations of question division (human). Creators will likewise examine the quality and shortcoming of calculations, complexities in movement understanding and recognize the conceivable future research challenges [2]. Jeroudi et al. goes for ordering human exercises in view of the inertial information gathered from a client's cell phone. An Online Sequential Extreme Learning Machine (OSELM) technique is executed to prepare a solitary concealed layer feed-forward system (SLFN). Test results with a normal exactness of 82.05% are accomplished [3].

Kwon et al. display a human movement acknowledgment mindful system utilizing multi-modular sensors associated with IoT gadgets and shopper gadgets in home. So as to comprehend a few sorts of client's circumstance inside a domain with IoT gadgets, a novel human action acknowledgment plot is required to oversee bunches of information for ensuring precise circumstance continuously [4]. Abdelhedi et al. propose a novel idea of mid-level portrayals to develop a discriminative and enlightening

Revised Manuscript Received on April 12, 2019.
Manu Bali, Lovely Professional University,Punjab, India
Devendran V, Lovely Professional University,Punjab, India

semantic idea for human activity acknowledgment. This work presents a mid-level portrayal in light of the Optical Flow (OF) strategy, Hu and Zernike minute together. First creators remove from every video, Uh and Uv movement vectors by framing movements ebbs and flows. Second, they decide the Hu minute and Zernike that fill in as the element vector of an activity. This strategy was tried and assessed through a characterization of the KTH and Weizmann datasets, with an Artificial Neural Network classifier (ANN). The outcomes demonstrate the exactness of the recommended approach [5]. Vivek et al. utilizes the picture handling procedures and designed learning frameworks to get the goal of sign video affirmation. To finish the proposed task execution it uses picture preparing methods like blueprint researching in light of edge acknowledgment, wavelet change, breaking down, broadening, cloud transfer, upheaval end, on preparing dataset. It in like manner uses Histogram Orientation Gradient called HOG for shape feature extraction and most essential part examination for rundown of capacities streamlining and diminishment. For result examination, this paper uses differing grouping accounts, for example, day circumstance, family data, official data, relations et cetera. Database of evacuated results are stood out from the video supported with the system as a commitment of the guarantor by an arranged unclear inference structure [6].

Sun et al. proposed enhanced fake state calculation to advance the parameters of SVM and connected to human action acknowledgment. Contrasted and other improvement calculations including essential fake province calculation, hereditary calculation and molecule swarm calculation on standard datasets, the proposed calculation can obtain higher grouping accuracy. Contrasted and fake settlement calculation in light of all dimensional pursuit, the enhanced calculation costs less running time. The proposed strategy is utilized as the classifier of human action and a high arrangement exactness is obtained [7]. Hanyuan et al. expects to decrease computational many-sided quality by diminishing the component measurement through dissecting the direct relationship between the highlights. In view of the help vector machine model of single-layer completely associated organizes, the preparation and acknowledgment time are fundamentally lessened while the acknowledgment precision is still guaranteed. The trial depends on general society dataset in the UCI Machine Learning Repository, and it utilizes Caffe, a profound learning system, to structure the help vector machine show. In the investigation, when the component measurement is decreased from 561 to 130, the preparation time can be diminished by 70% while the acknowledgment precision is kept at a promising 91% [8]. Yuan et al. center around geometric places of key parts of the face. Right off the bat, the face territory is distinguished in a photograph or video, at that point the key parts of the face is extricated and the position adjustment is performed. An arrangement of key focuses is found utilizing the relative position of the face. This procedure cannot just successfully maintain a strategic distance from the effect of the earth and the light on the example, yet additionally enormously enhance the acknowledgment of outward appearances. With the advancement of human-PC collaboration, outward appearance acknowledgment has turned into a hotly debated issue in the field of example acknowledgment. Following quite a while of advancement, outward appearance acknowledgment has made some progress, for example, HOG. In this paper, HOG highlight extraction strategy for outward appearance is utilized as a complexity. Test results demonstrate that the proposed calculation can extricate key data and accomplishes higher acknowledgment exactness [9].

Zhao et al. proposed a strategy in view of a tri-pivotal gyrator for fall occasions acknowledgment. A tri-pivotal whirligig is put at the client's abdomen to gather tri-hub precise speed data. Keeping in mind the end goal to encourage information preparing and extricate highlights, ongoing information are isolated into an arrangement of continuous and mostly covering windows. Three time-space includes that mirror the contrasts between the falls and different developments in our everyday lives are extricated from these backs to back information windows. At that point, every one of these windows is delegated speaking to either a fall or a non– fall occasion by utilizing a prepared machine learning classifier. Choice Tree is picked as the classifier in light of its low calculation multifaceted nature and simple usage on inserted frameworks [10]. Sainath et al. proposed an approach to extract the multilevel segmentation of a standing human body. In the strategy, firstly face was detected then skin detected using color base classification then followed by upper body and lower body detection. The main step is the detection of face in this approach [11]. Vette et al. A model is proposed that uses UBSP set of all vertebral trunk segments that is accurate andcomplete. The author in this paper takes the advantage of Visible Human Project to identify a complete set of upper body segment parameters as needed for state-of-the-art, 3D dynamic modeling [12]. Eslami et al. In this paper, the author extended the existing ShapeBoltzmannMachine(SBM) for the extraction of foreground object part. Proposed model is named as Multinomial SBM i.e. MSBM extract both local and global statistics of part shapes accurately.Object segmentation are done based on parts by performing probabilistic inference [13]. Shinde et al. proposed an automatic extraction of human bodies from single images. For the Localization of human body, position, dimensions, and color of the face are used and the upper and lower body models are constructed using anthropometric constraints, and estimation of the skin color. The performance of the algorithms was tested and achieved the accuracy of 89.53% and 97.68% [14]. Dharmateja et al. proposed a strategy for extraction of human bodies from single images automatically, in the case of almost upright poses in cluttered environments. Different levels of segmentation granularity are combined to extract the pose with highest potential. The segments that belong to the human body arise through the joint estimation of the foreground and background during the body part search phases, which alleviates the need for exact shape matching. Experimental results show the proposed method performs better in terms of accuracy [15]. Deboeverieet al. author has discussed the

idea that human body parts are approximated by nearly cylindrical surfaces, of which the axes of minimum curvature accurately reconstruct the human body skeleton. Then the segmentation is evaluated with a line segment distance method between reconstructed human body skeletons and ground truth skeletons. Experimental results are then compared with human body parts segmentations on the basis of mean shift, normalized cuts and watersheds and showed the better accuracies in case of proposed model [16]. Reddy et al. extraction of standing human bodies from single images strategy is discussed in the paper where the highly dimensional pose space, scene density, and various human appearances are handled in better way compared to conventional state of art methods. A bottom-up technique for human bodies extraction form the single images, in the case of almost upright poses in cluttered environments. Face recognition is done as a major part of the methodology [17]. Yadav et al. proposed a technique for the extraction of standing human bodies where the highly dimensional pose space, scene density, and various other human appearances are handled in better way than the existing state of arts techniques. Simulation results achieved the better performance [18].

Buehler at al. the author proposed a model that detect and track the articulated pose of a human in signing videos of more than one hour in length. The model accurately localizes hands and arms, despite fast motion and a cluttered and changing background. A complete model that accounts for self-occlusion of the arms is proposed and simulated results shows the better performance [19]. Magarat al. a method is proposed that automatically recover human bodies from photos. The haar cascades are used todetect human body that is haarcascade_upperbody and haarcascade_lowerbodyhelps to segment the upper and lower body of the human. MCTD algorithm is used to perform CT (coarse torso) detection for extracting the upper body segmentation. MOH based graph-cut algorithm is used to extract the Lower body. Experimental results showed the method achieved high performance than the conventional techniques [20].

III. PROPOSED METHODOLOGY

In the proposed technique, object detection is done from the video by applying the filters to enhance the quality of frames extracted from the video, then extracting the boundary of the frame to detect the object. Contour selection is done for dividing the image into different sections like upper half and lower half. Now the concept is how upper and lower half detection of the object proves itself good. To answer this query we have take two subplots in which the one is showing the upper half of the human body and another one is showing the lower half of the body. After comparing these two subplots we are able to generate the particular activity performed by the object in that video sequencing.

Upper Half	Lower half	Activity Generation
Straight	Straight	Standing
Straight	Bend	Sitting
Straight	Moving at uniform speed	Walking
Straight	Moving at non-uniform speed	Jogging

Fig 1: Shows the comparison of upper and lower half.

This research work follows following steps:

1. Get the video: the first and foremost step to be followed is to capture the video frame.
2. Divide it into frames: after capturing the video, it is partitioned into various frames for further processing.
3. Take the average of all the frames: all the frames are then evaluated further to calculate an average of all the frames.
4. Apply filters: after estimating average of the frames, filters are applied for further processing. Filtering is a technique for modifying or enhancing an image. For example, you can filter an image to emphasize certain features or remove other features. Image processing operations implemented with filtering include smoothing, sharpening, and edge enhancement.
5. Boundary extraction: after applying filtering on video frames, next step is to extract boundary of the detected object.
6. Human detection: after extracting boundary of an object, detection of object is performed.
7. Divide the image into upper half and lower half by using contour selection:

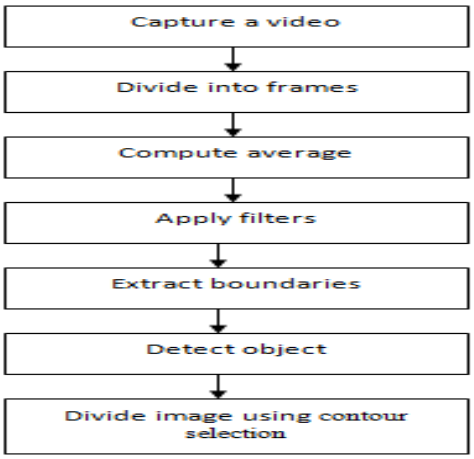

Fig 2: Flowchart of proposed technique

IV. EXPERIMENTAL RESULTS

The various challenges to be faced during the literature survey are as follows:

- The detection of objects in video sequencing rather than images.
- The performance gets reduced while dealing with larger datasets.
- The preprocessing of video processing is somehow difficult in case of picture division, extraction of features and distribution of various patterns.
- Less optimized video representation leads to low recognition of actions.

- Object tracking among low resolution videos needs an expensive computational cost.
- To reduce the identification time of an object.
- To avoid the impact of environment and light effectively.

This section presents results of the proposed technique.

(i) Extraction of frame from a video:

Fig 3: Original frame

Above figure presents original frame of video.

(ii) Extraction of background:

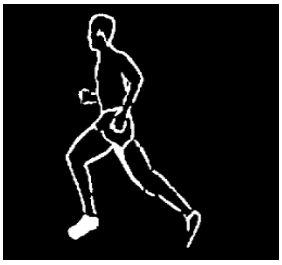

Fig 4: Background extraction of original frame'

After capturing original frame, background is extracted for further processing. Above figure shows extracted boundaries of the above captured frame.

(iii) Apply mask on a extracted frame:

Fig 5: Image mask

(iv) Apply edge detection on the object in an image:

Fig 6: Edge detection

Above figures show image mask and edge detection of a frame.

(v) Division of whole body into upper and lower half:

Fig 7: Upper bodyFig 8: Lower body

Above figures presents separation of image into upper and lower half by using contour selection.

(vi) Detection of upper portion features (Head/Face):

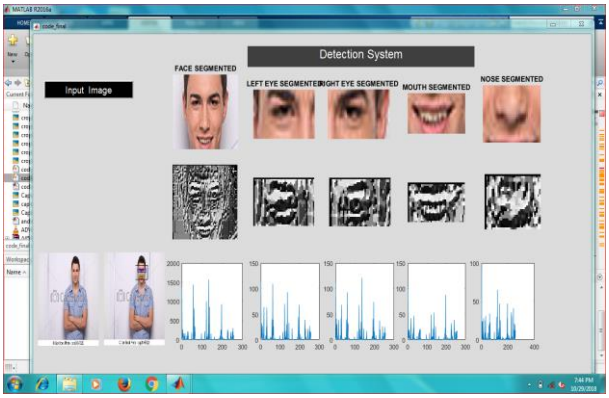

Fig 9. Head/Face Features Extraction

In this step we have to detect the facial features like eyes, mouth and nose. The reason behind this is to extract only the head portion from the upper half of the body so that on the basis of head orientation we can easily detect the appropriate posture of the human body.

(vii)To detect the posture of human body we have to follow the following work flow:

Using ANN Using contour

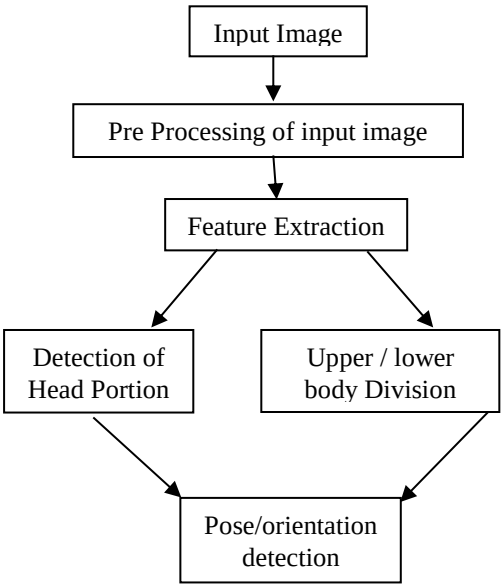

Fig 10: Flow Chart on Proposed Work

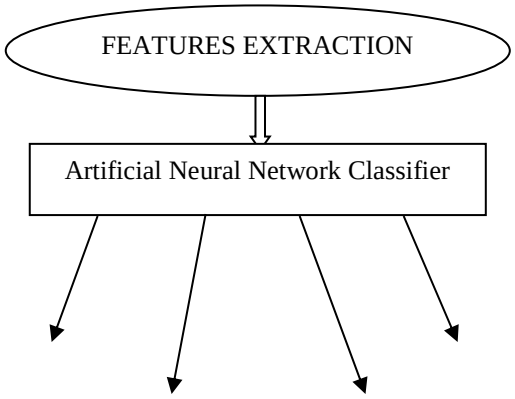

Fig. 11 Four Different Poses Recognition

V. CONCLUSION

In this work, we present an object detection framework using boundary extraction and contour selection. To conclude the whole work we can say that we cannot only recognize the numerous activities to be performed by a human body but we can also consider the different postures to be made by them. With the help of detection features of upper body(head) and lower body(legs). We can become able to find the particular direction of orientations of those features. Detection of various poses proves itself much easier if it can be done only on the basis of this division of upper and lower portions of the whole. This work does it's role clearly and able recognize more than 90% accuracy which combines the head detection and lower/upper body segmentation and recognition as a whole.

REFERENCES

1. Neha Shirbhate, Kiran Talele, "Human Body Language Understanding for Action Detection usingGeometric Features", IEEE, 2016, pp. 603-607.
2. Sandipkumar M Vaniya, Dr. B. Bharathi, "Exploring Object Segmentation Methods in VisualSurveillance for Human Activity Recognition", IEEE, International Conference on Global Trends in Signal Processing, Information Computing and Communication, 2016, pp. 520-525.
3. Yazan Al Jeroudi, M. A. Ali, MarsadLatief, RiniAkmeliawati, "Online Sequential Extreme Learning Machine Algorithm Based Human Activity Recognition using Inertial Data", IEEE, 2015.
4. Eunjung Kwon, Hyunho Park,SungwonByon, Eui-Suk Jung, Yong-Tae Lee, "A Human Activity Recog□n ition-Aware FrameworkUsing Multi-modal Sensor Data Fusion", IEEE, International Conference on Consumer Electronics, 2018.
5. Slim ABDELHEDI, Ali WALI, Adel M. ALIMI, "Human activity recognition based on mid-levelrepresentations in video surveillance applications", IEEE, International Joint Conference on Neural Networks, 2016, pp. 3984-3989.
6. Vivek D. Lad, Ramesh M. Kagalkar, "Methodology for Real Time Hand GestureRecognition and Generating Text DescriptionUsing Histogram Techniques", IEEE, International Conference on Intelligent Computing and Control, 2017.
7. Xuekai Sun, Haiquan Wang, Fanbing Zhu, "Human Activity Recognition Based on ImprovedArtificial Bee Colony Algorithm", IEEE, 2017, pp. 381-385.
8. Hanyuan Xu, Zhibin Huang, Jue Wang, Zilu Kang, "Study on Fast Human Activity Recognition Based on Optimized FeatureSelection", IEEE, International Symposium on Distributed Computing and Applications to Business, Engineering and Science, 2017, pp. 109-112.
9. Caiyou Yuan, Qingxiang Wu, Caiyun Wu, Pengfei Li, Yanan Zhang, Yao Xiao, "Expression Recognition Algorithm Based on theRelative Relationship of the Facial Landmarks", "EEE, International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, 2017.
10. ShizhenZhao, Wenfeng Li, WenyuNiu, Raffaele Gravina, Giancarlo Fortino, "Recognition of Human Fall Events Based on SingleTri-axial Gyroscope", IEEE, 2018.
11. Munde Sainath S. ,Dr.H.S. Fadewar, Shivpuje Prakash R., MundeSaraswati and MundeMaroti, "A Mechanism For Extraction Of Standing Human Body From Photo Using Multilevel Segmentation", Advances in Computational Sciences and Technology, 2017, pp. 1675-1681.
12. Albert H. Vette1, Takashi Yoshida1, T. Adam Thrasher, Kei Masani, and Milos R. Popovic, "A Complete, Non-Lumped, and Verifiable Set of Upper Body Segment Parameters for Three-Dimensional Dynamic Modeling", Medical Engineering & Physics, 2010.
13. S.M.AliEslami, ChristopherK.I.Williams, "A Generative Model for Parts-based Object Segmentation ",Advances in Neural Information Processing Systems 25 (NIPS 2012), 2012.
14. Mrs.R. Shinde, Mrs.R.S.Jamdar, "A Methodology for Extracting Standing Human Bodies from Single Images", Journal Of Information, KnowledgeAnd ResearchIn Electronics And Communication Engineering, 2017.
15. PenugondaDharmateja, P. Lakshmi Devi, "A Novel Approach for Segmentation of Human Bodies in Complex Background Images", International journal for innovative technologies, 2016.
16. Francis Deboeverie, Roeland De Geest, TinneTuytelaars, Peter Veelaert and Wilfried Philips, "Curvature-basedHumanBodyPartsSegmentationinPhysiotherapy", 2015
17. P. Lokeshwara Reddy, R. V. Sree Hari, "Extraction of Standing Human Bodies from Images using Multi Level Segmentation", International Journal of Scientific Engineering and Technology Research, 2017 pp. 5208-5213.
18. Abhishek Yadav and Vijay Kumar Joshi, "Human Body Detection with Segmentation Using Spline Regression", International Journal of Informative & Futuristic Research, 2017, pp. 7440-7449.
19. Patrick Buehler, Mark Everingham, Daniel P. Huttenlocher, Andrew Zisserman, "Upper Body Detection and Tracking in Extended Signing Sequences",International Journal of Computer Vision Springer, 2011.
20. Ashwini Magar, Prof.J.V.Shinde,"A New Approach of Human Segmentation from Photo Images" ,International Journal of Scientific and Research Publications, 2015
21. Khamparia, A and Pandey, B," SVM and PCA based learning feature classification approaches for E-learning system" International Journal of Web-based Learning Performance in Defence-Based Enhanced Learning, Vol. 9, No. 1, pp.37-50,2018.