IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

A Review on Human Activity Recognition System

N. Geetha[1*], E. S. Samundeeswari[2]

[1]Department of Computer Applications, Vellalar College for Women, Tamilnadu, India
[2]Department of Computer Science, Vellalar College for Women, Tamilnadu, India

[*]*Corresponding Author: geethuqyn@gmail.com*

Available online at: www.ijcseonline.org

Accepted: 19/Dec/2018, Published: 31/Dec/2018

***Abstract* –** Action recognition being one of the hot topics of many research efforts and being useful in so many commercial and scientific fields. Action recognition concerns the extraction ofactivity based knowledge, image data relationship or other patterns implicitly or explicitly stored in the images. Action in images is one of the powerful sources of high-level semantics. Recognition can be used for recognizing activities occurring in a particular scene. There is a need of effective and efficient methods to be encounter for recognizing the activities of human. The goal of this work is to study various recognition methods of common human actions represented in images. In this research, we present detailed insights on existing works and the methodologies used by researchers for recognizing the human activities. Comparison among different human activities by similarity systems is particularly challenging owing to the great variety of techniques implemented to represent likeness and the dependence that the results present of the used image dataset. This will be helpful to the researchers for their future research direction in this area.

***Keywords* –** Human Activities**,** Segmentation, Feature Extraction, Classification, Human Action Recognition

I. INTRODUCTION

Image processing isa growing technology and can be used in a various fields like optical imaging, medical imaging, computer vision, computer graphics etc.Image processing is a method that takes image as input and performs some operations on it and gives image as output. The input image may be a photograph or video frame (considered as two dimensional signals) and the output may be the parameters related to that input image. The output image is enhancedby taking the useful information from it.Digital image processing enhances the image by modifying its image qualities. Operations of image processing are divided into image enhancement, restoration, segmentation, extraction and recognition [9].

Recently still image-based human action recognition has become an active research topic in computer vision and pattern recognition. The result of various research studies indicates that the success of action recognition problems highly depends on the appropriate feature extraction process. It focuses on identifying a person's action or behavior from a single image. Unlike the traditional action recognition approaches where videos or image sequences are used, a still image contains no temporal information for action characterization. Thus the prevailing spatio-temporal features for video-based action analysis are not appropriate for still image-based action recognition. It is more challenging to perform still image-based action recognition than the video-based one, given the limited source of information's as well as the cluttered background of images. Therefore it is demanding to develop robust and efficient methods for still image-based action recognition to understand the web images better for image retrieval or search. Based on the emerging research in recent years, it is time to review the existing approaches to still image-based action recognition and inspire more efforts to advance the field of research [4].

Human behavior Analysis, using visual information in a givenimage or sequence of images, has been an active area of researchin computer vision community. The image captures byconventional camera does not provide the suitable information toperform comprehensive analysis. Most of the existingwork focuses on body part detection and pose estimation. Agrowing research area addresses the recognition of human actionsbased on depth images. Due to the large diversity of human body, size, appearance, posture, motion, clothing, view angleand illumination changes, besides the complexityof human actions, the task of automatically recognizinghuman actions is very challenging [5].

Section I contains the introduction of image processing and fundamental theories for human activity recognition system. Section II covers the literature survey of emerging

techniques for human activity recognition system as well as several extended work in these areas. Section III concludes the human activity recognition techniques and identifying the future directions in this field.

II. LITERATURE SURVEY

To better understand Human Action Recognition System, it is useful to review and examine the existing research works in literature**.** Therefore, recent approaches and methodologies used for recognizing actions from images and image sequences have been discussed.

Sumaira Ghazal, Umar S. Khan has proposed an algorithm for posture recognition in still images using 2D pose information from human skeleton is implemented. An approach bases on angles and distances between joints is adopted for classification between sitting and standing postures. The openpose library is used to obtain human skeleton information from the images. The system takes an image as an input and generates an output containing the locations of 18 key-points for each person in the image. Total 18 key-points are obtained out of which 6 joints i.e. Left hip, left knee, left ankle, right hip, right knee and right ankle joints' positions. They are used to extract features to classify between sitting and standing postures. A rule based decision-making approach is used to classify sitting and standing postures [10].

Nabil Zerrouki, FouziHarrou, Ying Sun, AmraneHouacinehas presented a computer vision-based methodology for human action recognition using Adaboost algorithm. Vision-based method includesa four major phases namely: (1) data collection, (2) image segmentation, (3) feature extraction and (4) action classification. During data acquisition, video sequences correspondingto different human activities are recorded. The segmentation phase consists of extracting the body's silhouette from the input frame (image sequence). In this paper, human body segmentation is based on the background subtraction technique. A threshold is needed to decide whether a pixel belongs to the foreground or the background. After the background subtraction, some noise regions can be entered. To eliminate this noise, the morphological operator based on erosion and dilation operators with 3×3 structuring elements is applied. The feature extraction step determines discriminative information is needed to describe the human silhouette. The human body is divided into five portions. Partitioning is performed by tracing five segments from the body's center of gravity. First segment is vertical. Second and third segments are located at 45◦ on either side of the vertical segment. Fourth and the fifth segments are situated at 100◦ on either side of the third and fourth segments. Finally, the classification phase is required to distinguish between different human activities. The classification process consists of two steps: (i) assigned the video sequences as training samples (ii) classified the sequences according to their corresponding feature spaces via a trained classifier model. Adaboost classifier is to use a sequence of weak classifiers in order to form a powerful classifier.Experiments show that AdaBoost classifier has achieved the most accurate results with much less tuning of parameters or settings than other commonly used classifiers such as KNN, Naïve bayes, Neural Network and SVM [7].

Zhichen Zhao, Huimin Ma, ShaodiYouhas proposed the idea of semantic body part actions to improve single image action recognition. The proposed methodology combines body actions and part actions for action recognition. In this work, seven body parts and their semantic part actions (head, torso, arms, hands and lower body) are defined. Entire body actions are combination of semantic part actions. A deep neural network based system is proposed. First, multiple body parts are localized by a key-point prediction network. The key-point prediction task is also known as pose estimation. Second, for each body parts, a Part Action Network is used to predict semantic body part actions. The Part Affinity Fields Network (PAF) is a pose estimation method, which can be handled multi-person tasks, to predict key-points. The PAF network receives a person bounding box image as input and produces all possible landmarks locations and landmarks are grouped into multiple people. From that choose the largest one and generate part bounding boxes by post-processing. Bounding boxes are computed as the minimum bounding boxes enclosing the related key-points. Part Action Network (PAN) which receives both images and localized parts as input, and jointly learns body actions, part actions and fusion features for action prediction. In this paper,3 classification networks are trained namely the Baseline Network, the Part-based Network and the Part Action Network. The baseline network learns the mapping relationship from body appearance to body actions. The part-based network learns mapping relationship from part appearance to body actions. Part Action Network to be learned and predicted part actions combines global body actions and local part actions. Part Action Network method reaches the best results for 7 out of 10 activities [12].

GhazaliSulong, Ammar Mohammedali which includes the following basic steps: preprocessing (foreground extraction) to reduce noise from given image, segmentation (background subtraction) to extract useful features from object and classification to sortout these features by the classifier. It is highly promising compared to the existing one of 85% over last decade [3].

Chaitra B H, Anupama H S, Cauvery N K has presented a recognition system for human actions using a novel self-organizing map (SOM) based retrieval system. By using two-dimensional discrete cosine transforms (2D-DCT) to extract image vectors. These vectors are input to the

neural network classifier, which uses SOM algorithm to recognize elementary actions from the images (trained). SOM has three phases of life cycle: learning phase, training phase and testing phase.The system is evaluated in MATLAB using an image database of 30 action images, containing six subjects and each subject having 5 images with different actions. After training for approximately 3000 epochs the system has achieved a recognition rate of 98.16% for fastest network training time.Epochs are neural network training parameters. The SOM neural network is tested to determine the optimal number of epochs to be used for neural network training time to find the best possible recognition rate.The system having less computational requirement this make system well suited for low cost, real-time hardware implementation [1].

MoinNabi, Mohammad Rahmati which includes the following basic steps: First extract all poselets in the images for using as the descriptor of human's activity. These extracted poselets can be lead to build a Poselet Activation Vector (PAV) for every image. These PAVs has used to make a poselet by Image matrix. This matrix is similar to the matrix of word by document in the text space; so used to text-based approaches for extracting the basic concepts of human pose (Latent Poselets). Model the latent topics of human poses by using extracted vectors and P-LSA.Finally, the human's actionis recognized in a query image by using the trained SVM (Support Vector Machine) on the extracted bag of latent poselets. The result shows the significant improvements in some action classes such as Walking and Running [6].

Weilong Yang, Yang Wang, and Greg Mori has proposed a model as action recognition and poses estimation. The main novelty of this model is treat the human pose information as latent variables that will help with recognition. It uses the output of a pose estimation algorithm as the input of an action recognition system.The latent pose of this approach is not restricted to any specific type of pose representation.It uses an exemplar-based pose representation as "poselet" is proposed. "poselet" is to refer to a set of patches not only with similar pose configuration, but also from the same action class.The action-specific parts contain more discriminative information; itdecides to select the poselets per action.In this work manually annotate the pose with 14 joints on the human body on all the images in the dataset.For each poselet, trained an SVM classifier based on the HOG descriptors extract from image patches at the locations of the corresponding poselet in the training set.Localizing the body parts information is useful for improving the final action classification results [11].

Christian Thurau, Vaclav Hlavachas presented a method for recognizing human actions based on pose primitives. In learning mode, the parameters representing poses and activities are estimated from videos. In run mode, the method can be used both for videos or still images. Action classes are represented by histograms of poses primitives. Action recognition is based on a simple histogram comparison.The contribution of this paper is threefold: (i) Presented a complete approach for recognizing activities from single images and image sequences (ii) Extended a Histogram of Oriented Gradient (HOG) based descriptor for recognizing pose primitives (iii) Developed a histogram based action recognition approach that assign the pose to a suitable action category. The task for action recognition is to an action class based on a single recognized pose primitive, or a sequence of recognized poses [2].

NazliIkizler, R. GokberkCinbins, SelenPehlivan and Pinar Duyguluhasproposed a concept for representing the pose with a spatial and orientationalhistogramming of rectangular regions on a parse probability map by using circular Histogram of Rectangles. This approach starts with employing a pose extractor, and then representing the pose via distribution of its rectangular regions. So it is needed to estimate the place and pose of the person. The approach uses edge and region features, and constructs two deformable models using Conditional Random Fields (CRF). The edge-based deformable model is used to estimate the initial body part positions. Using the previously obtained estimate, the method creates a region based model (parse) that represents an image for each one of the body parts. Histogram of Oriented Rectangles is used for representing the pose. Then computed the histogram of extracted rectangular regions based on their orientations. The rectangles are histogrammed over 15◦ orientations. For still images, doing this histogramming over the spatial circular grids and defined circular HORs (CHORs), as opposed to original N ×N grid form. This is mostly because cannot bedefined the explicit height of the human figure due to the discrepancies of the parse. Using circular grid helps to capture the angular positions of the parts more reliable. Linear Discriminant Analysis (LDA) is used to reduce the feature dimension from 144 to 50 and obtain a more compact and discriminative feature representation. Binary SVMs (classification) is used to obtain the highest probable class label [8].

The focus of this review is to summarize the various existing human activity recognition techniques are shown in Table 1.

Table 1. Summarization of various human activity recognition system methods

Authors	Proposed Techniques	Dataset Used	Activities Recognized	Accuracy (%)

Sumaira Ghazal, Umar S. Khan [2018]	Openpose library, Rule based decision making approach	INRIA Dataset, Freiburg Dataset (238 images)	sitting and standing postures	97
		MPII Human Pose Dataset (229 images)		93
		Image Database (images collected from Internet)		95.2
Nabil Zerrrouki, FouziHarrou, Ying Sun, AmraneHouacine [2017]	Background subtraction technique, Thresholding, Morphological operator, Adaboost classifier method	URFD fall detection Dataset	walking, standing, bending, lying, squatting, and sitting	96.6
Zhichen Zhao, Huimin Ma, Shaodi You [2017]	Deep neural network based system, 3 classification networks: the Baseline Network, the Part-based Network and the Part Action Network	PASCAL VOC 2012 dataset(500 images)	jumping, phoning, play instruments, reading, riding bike, riding horse, running, taking photo, using computer and walking	91.2
		Stanford-40 dataset (100 images)		
GhazaliSulong, Ammar Mohammedali [2015]	Preprocessing, Segmentation and Classification	INRIA and KTH Dataset (780 images - 64*128 pixels format)	running, walking, jumping, standing and sitting	86.2
Chaitra B H, Anupama H S, Cauvery N K [2013]	Two-dimensional Discrete Cosine Transform (2D-DCT), Neural network classifier - SOM algorithm	Action Image Database (30images - 3000 epochs)	walking, running, jumping, falling, bending, sitting and standing	98.2
MoinNabi, Mohammad Rahmati [2012]	Poselet Activation Vector (PAV), Text-based approaches (Latent Poselets), P-LSA (Bag of Latent Poselets), SVM (Support Vector Machine)	PASCAL VOC 2010 action classification dataset	phoning, play instruments, reading, riding bike, riding horse, running, taking photo, using computer and walking	58.6
Weilong Yang, Yang Wang, and Greg Mori [2010]	Pose estimation algorithm, Exemplar-based pose representation, SVM classifier	Image Database (2458 images collected from Internet)	running, walking, playing golf, sitting, and dancing	61.1
Christian Thurau, Vaclav Hlavac [2008]	Histogram of Oriented Gradient (HOG) based descriptor (poses primitive), Histogram based action recognition approach	Weizmann action-recognition data set (130 pose primitives)	run, wave1, wave2, jump, pjump, skip, side, jack, bend and walk	83.3
NazliIkizler, R. GokberkCinbis, SelenPehlivan and Pinar Duygulu [2007]	circular Histogram of Rectangles, Conditional Random Fields (CRF) - edge-based deformable model and region based model, Linear Discriminant Analysis (LDA), binary SVM classification	Image Database (467 - images collected from various sources like Google Image Search, Flickr, BBC Motion database)	running, walking, catching, throwing, crouching and kicking	85.1

III. CONCLUSION

Automatic human action/activity recognition has been one of the challenging issues in computer vision in recent years. It is of great importance in various applications in artificial intelligence like video surveillance, computer games, robotic and human computer interactions. When studying the related work on this issue, it has been found that the human action recognition system can be designed by the following steps of processes such as feature extraction, action representation, and classification.

The review aims at stressing out the need of automating their processing and classification with the purpose of recognizing human activities from an image dataset. This paper presents a survey on various human activity recognition techniques that was proposed earlier by researchers for the better development in the field of human activity recognition system. All these techniques have their own advantages; In other words, there is not a single technique that fits best in all category of human activities. We attempt to sum up the methods and techniques used to recognize the human activities. It will be helpful to the

researchers for understand and compare the related advancements in this area.

REFERENCES

[1] Chaitra B H, Anupama H S, Cauvery N K, *"Human Action Recognition using Image Processing and Artificial Neural Networks*", International Journal of Computer Applications (0975 – 8887), Volume 80 No. 9, October **2013**.

[2] Christian Thurau, Vaclav Hlavac, *"Pose Primitive based Human Action Recognition in Videos or Still Images"*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **2008**.

[3] GhazaliSulong, Ammar Mohammedali, *"Recognition of Human Activities from Still Image using Novel Classifier"*, Journal of Theoretical and Applied Information Technology, Vol. 71 No. 1, 10th January **2015**.

[4] GuodongGuo, Alice Lai, *"A Survey on Still Image based Human Action Recognition"*, Journal of Pattern Recognition, Vol. 47, pp-3343 to 3361, 9th May **2014**.

[5] Md. Atiqur Rahman Ahad, J.K. Tan, H.S. Kim and S. Ishikawa, *"Human Activity Recognition: Various Paradigms"*, International Conference on Control, Automation and Systems, pp-1896 to 1901, Oct. 14-17, **2008** in COEX, Seoul, Korea.

[6] MoinNabi, Mohammad Rahmati, *"Human Action Recognition in Still Images using Bag of Latent Poselets"*, CVMP2012 (Conference on Visual Media Production), **2012**.

[7] Nabil Zerrrouki, FouziHarrou, Ying Sun, AmraneHouacine, *"Adaboost-based Algorithm for Human Action Recognition"*, **2017**.

[8] NazliIkizler, R. GokberkCinbins, SelenPehlivan and Pinar Duygulu, *"Recognizing Actions from Still Images"*, Indian Council of Philosophical Research, ICPR, New Delhi,**2007**.

[9] Suganya V, *"A Survey on Image Processing and Human Action Recognition"*, International Journal of Latest Trends in Engineering and Technology (IJLTET), ISSN: 2278-621X, Vol. 6, Issue 1, 1st September **2015**.

[10] Sumaira Ghazal, Umar S. Khan, *"Human Posture Classification using Skeleton Information"*, International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), **2018**.

[11] Weilong Yang, Yang Wang, and Greg Mori, *"Recognizing Human Actions from Still Images with Latent Poses"*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **2010**.

[12] Zhichen Zhao, Huimin Ma, Shaodi You, *"Single Image Action Recognition using Semantic Body Part Actions"*, IEEE International Conference on Computer Vision (ICCV), **2017**.

Authors Profile

Ms. N. Geetha is currently pursuing Ph.D. and working as Assistant Professor in Department of Computer Applications, Vellalar College for Women, Erode, Tamilnadu, India since 2010. She has published more than 4 research papers in reputed international online journals. Her main research work focuses on Image Processing Techniques, Data Mining Algorithms, Network Security, Cloud Security and Privacy. She has 9 years of teaching experience and 5 years of Research Experience.

Dr E. S. Samundeeswari is currently working as Associate Professor in Department of Computer Science, Vellalar College for Women, Erode, Tamilnadu, India since 1989. She is currently working as Controller of Examinations in her institution since 2018 and a life member of the ISCA (Indian Science Congress Association) and ISTE (Indian Society for Technical Education). She has published more than 16 research papers in reputed international online journals including SCI and Scopus. Her main research work focuses on Image Processing Techniques, Network Security, Cloud Security and Privacy, Big Data Analytics, Data Mining, IoT and Computational Intelligence based education. She has 28 years of teaching experience and 22 years of Research Experience.

