ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Egocentric Activity Recognition using Combined SVMkNN Classifier

K.P.Sanal Kumar[1], R.Bhavani[2]

Asst. Professor/Programmer, Dept. of ECE, Annamalai University, Tamilnadu, India[1]

Professor, Dept. of CSE, Annamalai University, Tamilnadu, India[2]

ABSTRACT: Egocentric vision is a unique perspective in computer vision which is human centric. The recognition of egocentric actions is a challenging task which helps in assisting elderly people, disabled patients and so on. In this work, life logging activity videos are taken as input. There are 2 categories, first one is the top level and second one is second level. Here, the recognition is done using the features like Histogram of Oriented Gradients (HOG), Motion Boundary Histogram (MBH) and Trajectory. The extracted features are reduced using Principal Component Analysis (PCA). The reduced features are provided as input to the combined classifiers Support Vector Machine (SVM) and k Nearest Neighbor (kNN) (Combined SVMkNN). This classifier provided better results than other classifiers in the literature.

KEYWORDS: Egocentric; Histogram of Oriented Gradients; Motion Boundary Histogram; Trajectory; Life Logging Activity;

I. INTRODUCTION

The conventional method of activity recognition involves studying the behaviour of a person from one or morecameras [1]. A significantimprovement over the recent yearsin identifying actions has been identified using first person camera. However, majorchallenges in this field need subtle movements/gestures. Thisis due to occlusions and distractions from image regions where theactual action is taking place. Another method is implemented by using egocentric cameras with which the actions are obtained and analysed from the user's perspective.

Egocentric video analysis for user activities has attracted attention in the recent development technology. These videos offer a helping hand for the disabled or elderly people. Lifelogging is an activity of recording some portions of a person's life. Here, the videos taken are for activities like drinking, eating, house work etc. The recording is done automatically using wearable devices. Many applications of Lifelogging activity are behaviour analysis, lifestyle analysis, health monitoring and so on.

In this research work, HOG (Histogram of oriented gradients), Motion Boundary Histogram (MBH) and Trajectory are extracted. Here, HOG gives the static information, MBH gives the motion information and Trajectory capture the local motion information of the video.

II. RELATED WORKS

[2] Summarizes the evolution of the state of the art in FPV video analysis highlighting, among others, the most commonly used features, methods, challenges and opportunities within the field. [3] Our solution provides the flexibilitytocapturetheinteractionsdisregardingthenumberofindividualsinvolvedandtheirlevelof acquaintanceincontextwithavariabledegreeofsocialinvolvement. [4] Proposed anovelmethod for solvingthebodypartandmotion identification problem.[5] designed a new algorithm, Ensemble Actions EMT (EA-EMT),utilises the initial environment model as a library of state transition functions and applies a variationof prediction with experts to assemble and calibrate a revised model. [6] Present a new approach that exploits the inherent contextual information from structured hand labelling for pixel level hand detection and hand part labelling. In [7] office environment is taken egocentric activity recognition. Here, motion descriptors are extracted which are combined with

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

user eye movements. In [8], First Person Videos are temporally segmented into twelve hierarchical classes. Here, the problem of FPV ADL analysis is implemented for multi-task learning framework. [9] implemented the HOG feature and support vector machine (SVM) was applied to train an action classifier using HOG features. [10] Implemented the Probabilistic Neural network (PNN) classifier for classifying the actions of supplied and PCA for dimensionality reduction.[11] used a Hessian detector and HOG and Histograms of Optical Flow (HOF) descriptors alongwith a Bag-of-Features (BoF) representation to implement action recognition on simple and realistic datasets. In [12], the Motion Boundary Histogram (MBH) is a recent appealing approach to suppress the constant motion by considering the flow gradient. It is robust due to the presence of camera motion. [13] Proposed the motion boundary histograms (MBH) descriptor for human detection by computing derivatives separately for the horizontal and vertical components of the optical flow. The descriptor encodes the relative motion between pixels. [14] Recognizing activities of people by SVM multi-class classifier whose structure is determined by a clustering process.

III. PROPOSED WORK

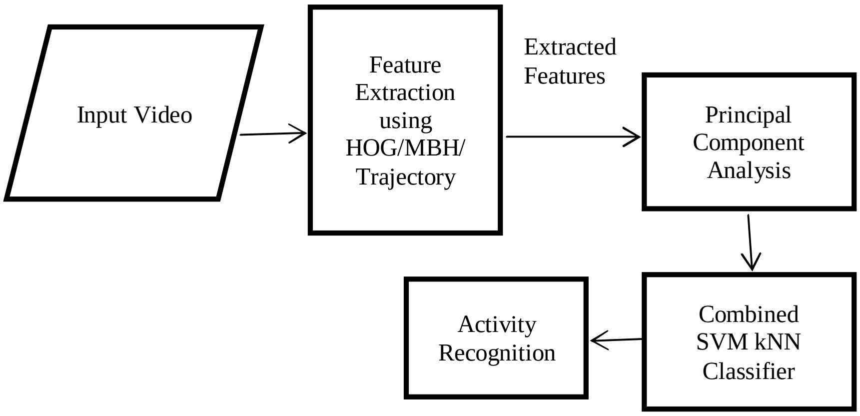

Figure 1. The Proposed model

The input video is fed where the feature extraction is applied. The features are Histogram of Oriented Gradients (HOG), Motion Boundary Histogram (MBH) and Trajectory. The extracted features are reduced using Principal Component Analysis (PCA). The reduced features are provided as input to the combined classifier (Combined SVMKNN) Support Vector Machine (SVM) and k Nearest Neighbor (kNN). The Figure 1 shows the proposed model.

IV. FEATURE EXTRACTION

Histogram of Oriented Gradients (HOG)

This feature is local shape information described by the distribution of gradients or edge directions. This generally focuses on static appearance information [15].

The input frame is chosen where the normalization is applied on both color values and gamma correction is applied. The next step is to apply gradient filter for finding the gradient values. The window is partitioned into adjacent, non-overlapping cells of size C×C pixels (C = 8). In each cell, a histogram of the gradient orientations is calculated which is binned into B bins.

Calculate the weights by using bilinear interpolation. These are subjected to normalization by concatenating overlapping blocks. The normalized block features are combined into a single feature vector which is the HOG. Initially the videos are divided into frames and then the feature extraction technique is applied. Here, 8102 features are extracted.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Motion Boundary Histogram (MBH)

The optical flow's horizontal and vertical components are separately represented using two scalar maps, of the motion components [16]. Histograms of oriented gradients are calculated for each of the two optical low component images, similar to HOG in still images. Due to low differences, the information about changes in motion boundaries is retained whereas the constant motion information is rejected. This leads to the cancelation of most of the effects of camera motion. Spatial derivatives are calculated for each and orientation is quantized into histograms. The magnitude used here is for weighting. The outcome of this process is a pair of horizontal (MBHx) and vertical (MBHy) descriptors. Here, 2050 features are extracted.

Trajectory

Here, feature points are tracked on each spatial scale separately. For each frameI_t, its dense optical flow field$\omega_t = (u_t, v_t)$ is calculated with respect to the next frameI_{t+1}, whereu_t and v_t are the horizontal and verticalcomponents of the optical flow. Given a point $P_t = (x_t, y_t)$ in frameI_t, its tracked position in frame $I_{t+1}$is smoothed by applying a median filter onω_t.

$$P_{t+1} = (x_{t+1}, y_{t+1}) = (x_t, y_t) + (M * \omega_t)|_{(x_t, y_t)} \qquad (1)$$

Where M is the median filtering kernel [17]. The size of the median filter kernel M is 3 × 3 pixels. As the median filter is more robust to outliers, it improves trajectoriesfor points at motion boundaries that would otherwise be smoothed out. To extract dense optical flow fields, the algorithm embeds a translation motion model between neighbourhoods of two consecutive frames [18]. Polynomial expansion is incorporated to approximate intensities of pixel in the neighbourhood. Points of subsequent frames are concatenated to form trajectories: $(P_t, P_{t+1}, P_{t+2}, \ldots)$.

For each frame, if no tracked point is found in its neighbourhood, a new point is sampled and added to the tracking process so that a dense coverage of trajectories is identified. The shape of a trajectory encodes local motion patterns. Given a trajectory of length L, its shape can be described by a sequence $(\Delta P_t, \ldots \Delta P_{t+L-1}, \ldots)$of displacement vectors$(\Delta P_t = ((P_{t+1} - P_t) = (x_{t+1} - x_t, y_{t+1} - y_t)$. The resulting vector is normalized by the sum of displacement vector magnitudes:

$$T = \frac{(\Delta P_t, \ldots \Delta P_{t+L-1})}{\sum_{j=t}^{t+L-1} \|(\Delta P_j\|} \qquad (2)$$

In this research work, 902 features are extracted.

V. DIMENSIONALITY REDUCTION USING PRINCIPAL COMPONENT ANALYSIS(PCA)

The obtained feature vectors are of large size. Thislarge dimension results in a very slow computation. As a result, the dimensions are reduced using PrincipleComponent Analysis.

In statistics, Principal Components Analysis (PCA) [19] is a technique which is implementedto reduce the dimensionality of a high dimensional data set. It is a linear transformationthat chooses a new coordinate system for the data set. The new coordinate system is a representation of the directions along which the variance of the data is highest. ThePCA can be used for reducing dimensionality in a data set whileretaining the characteristics of the data set that contribute most to its variance, bykeeping lower-order principal components and ignoring higher-order ones.

Applying Principal Component Analysis

Algorithm for PCA

Step 1: The column vectors are generated from the input feature vectors.

Step 2. The covariance matrix of the two column vectors formed in step 1 is calculated.

Step 3. The diagonal elements of the 2×2 covariance vector would contain the variance of each column vector with itself, respectively.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

Step 4. The Eigen values and the Eigenvectors of the covariance matrix are computed [20].

Initially the covariance matrix is calculated and then the Eigen vectors are derived. Once the Eigenvectors are found from the covariance matrix, the next stepis to rank them by their eigenvalue, in the descending order that results in arranging thecomponents in the order of significance. To reduce the dimensionality of the data-set, ignore thecomponents of lower significance [21]. Here, each feature vector is reduced to 800.

VI. COMBINED SVM KNNCLASSIFIER (COMBINED SVMKNN)

Support Vector Machine (SVM) is a supervised learning method for classification. Here, a hyperplane is created through which data is classified into groups Support vector machine(SVM) basically uses support vectors to create a hyperplane.k Nearest Neighbour algorithm is basically a machine learning algorithm. kNN algorithm is used to find new data which is added to training data set. Support vector machine (SVM) and k nearest neighbour (kNN) algorithms are combined together to evaluate false positive rate is known as Combined Support Vector Machine k Nearest Neighbour(Combined SVMkNN) algorithm. Here, in SVM classifier, the Guassian kernel is defined as

$$K(x, \acute{x}) = exp\left(\frac{-\|x-\acute{x}\|^2}{2\sigma^2}\right) \quad (3)$$

$\|x - \acute{x}\|^2$is squared euclidean distance and σ is a measure of expansion.

Algorithm for Combined SVMkNN

Step 1: Select data from different class;

Step 2: Separate data by using SVM classifier and kNN classifier.

Step 3: Calculate False Positive Rate for both the classifiers

Step 4: If FPR(SVM) is greater than FPR(kNN)

then apply kNN for clustering the data.

Step 5: If new data added to data set then update dataset;

Repeat the above steps for all data in the data set.

These two algorithm works together in Combined SVMkNN algorithm in which, support vector machine (SVM) uses training data set to learn something from data set. If any new is added to its dataset. It is updated by k nearest neighbour (kNN) algorithm.

VII. PERFORMANCE METRICS

Accuracy is defined as the ratio between the summation of true positive rate and true negative rate and the total population.

$$\text{Accuracy} = Tp+Tn / (Tp+Tn+Fp+Fn) \quad (4)$$

Where,

Tp is the number of items correctly classified as positive class

Tn is the number of items correctly classified as negative class

Fp is the number of items wrongly classified as positive class

Fn is the number of items wrongly classified as negative class

VIII. EXPERIMENTAL RESULTS

In our work, the videos are classified into two levels. They are 5 top level categories and 13 second level categories. The top level categories are motion, social interaction, office work, food and house work. The second level categories are walk straight, walk back and forth, walk up and down, running, talk on the phone, talk to people, watch videos, use

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

internet, write, read, eat, drink and housework. Table 1 shows the accuracy of both classifiers for top level. Table 2 depicts the accuracy of both classifiers for second level.

Table 1: Accuracy for top level categories

Features / Classifiers	Histogram of Oriented Gradients	Motion Boundary Histogram	Trajectory
CombSVMkNN	82.34	83.79	85.21
[21]	77.42	82.46	84.50

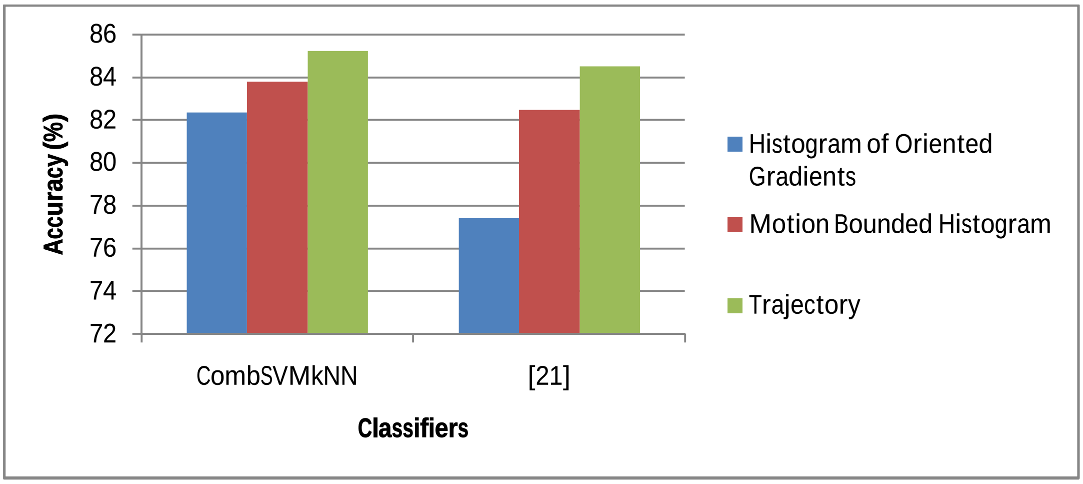

Figure 2: The chart to depict accuracy for top level categories

Table 2: Accuracy for second level categories

Features / Classifiers	Histogram of Oriented Gradients	Motion Boundary Histogram	Trajectory
CombSVMkNN	72.48	79.53	76.17
[21]	68.15	78.04	74.46

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

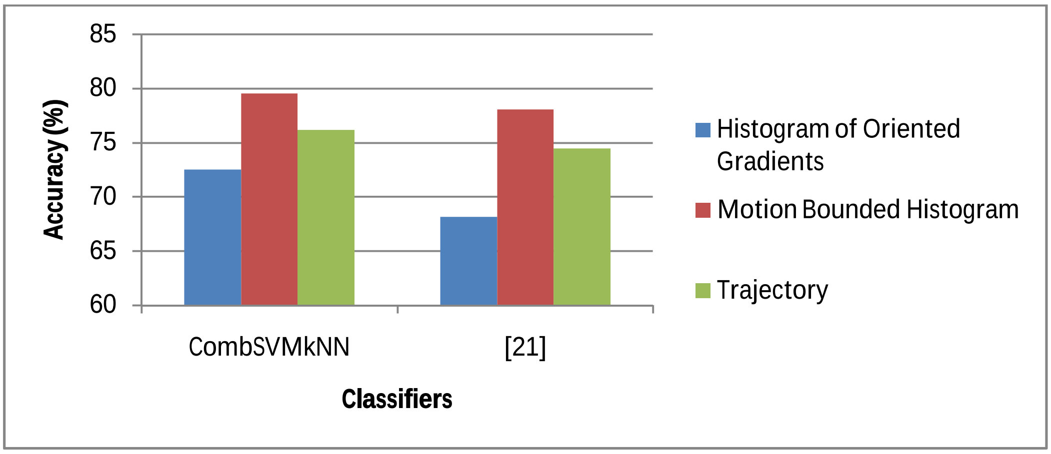

Figure 3: The chart to depict accuracy for second level categories

From the above tables 1 and 2 and figures 2 and 3, CombSVMkNNperformed the best for both categories than literature methods. ThereforeCombSVMkNNis better than other classifiers.

IX. Conclusion

In this research work, Combined SVMkNN classifier provided better results. SVM provides advantages like it is accurate and robust even when the training sample has some bias. It delivers unique solution since the optimality problem is convex. kNN classifier is a simple classifier. These advantages are combined together to form Combined SVMkNN classifier which provides better results than other classifiers in literature.

References

1. Poppe, R., " A survey on vision-based human action recognition", Image and VisionComputing, Elsevier, 2010.
2. Alejandro Betancourt, PietroMorerio, Carlo S. Regazzoni, and Matthias Rauterberg, "The Evolution of First Person Vision Methods:A Survey",IEEE Transactions On Circuits And Systems For Video Technology, Vol. 25, No. 5, May 2015.
3. Stefano Alletto, Giuseppe Serra, Simone Calderara, Rita Cucchiara, "Understandingsocialrelationshipsinegocentricvision", PatternRecognition, Elsevier, vol.48, pp.4082–4096, 2015.
4. Sihao Ding, QiangZhai , Ying Li , Junda Zhu , YuanF.Zheng , Dong Xuan, "Simultaneous body part and motion identification for human-following robots", Pattern Recognition, Elsevier,vol.50, pp.118–130, 2016.
5. ZinoviRabinovich, Nicholas R. Jennings Robotics and Autonomous Systems, "A hybrid controller based on the egocentric perceptual principle", Robotics and Autonomous Systems, Elsevier, vol.58,no.9, pp.1039-1048.
6. XiaolongZhu, XuhuiJia,Kwan-YeeK.Wong, "Structuredforestsforpixel-levelhanddetectionandhandpart labeling", Computer Vision and Image Understanding, Elsevier,vol.141,pp. 95–107,2016.
7. K. Ogaki, K. M. Kitani, Y. Sugano, and Y. Sato, "Coupling eye-motion and ego-motion features for first-person activity recognition," in CVPR Workshop on Egocentric Vision, 2012.
8. Kishor B. Bhangale and R. U. Shekokar, "Human Body Detection in static Images Using HOG & Piecewise Linear SVM", International Journal of Innovative Research & Development, vol.3 no.6, June 2014.
9. C. P. Huang, C. H. Hsieh, K. T. Lai and W. Y. Huang, "Human Action Recognition Using Histogram of Oriented Gradient of MotionHistory Image", 2011 First International Conference on Instrumentation, Measurement, Computer, Communication and Control, Beijing,2011,pp.353-356.
10. Human Action Recognition in Video using Histogram of Oriented Gradient (HOG)Features and Probabilistic Neural Network (PNN) A. Jeyanthi Suresh, P.Asha,International Journal of Innovative Research in Computer and CommunicationEngineering, Vol. 4, Issue 7, July 2016.
11. Improving Robustness and Precision in GEI + HOG Action Recognition Tenika P. Whytock , Alexander Belyaev , Neil M. Robertson 9th International Symposium, ISVC 2013, Rethymnon, Crete, Greece, July 29-31, 2013. Proceedings, Part I pp 119-128.
12. H.Wang, A. Klaser, C. Schmid, and C.-L. Liu. Action recognition by dense trajectories. In CVPR, Jun. 2011.
13. N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance, "in European Conference on Computer Vision, 2006.
14. HuiminQian, Yaobin Mao, Wenbo Xiang, Zhiquan Wang, "Recognition of human activities using SVM multi-class classifier", Pattern Recognition Letters, Elsevier, vol. 31, no.2, pp.100–111, 2010.
15. Kishor B. Bhangale and R. U. Shekokar, "Human Body Detection in static Images Using HOG & Piecewise Linear SVM", International Journal of Innovative Research & Development, vol.3 no.6, June 2014.
16. Heng Wang, Alexander Kaiser, CordeliaSchmid, Cheng-Lin Liu. Dense trajectories and motion boundary descriptors for action recognition. [Research Report] RR-8050, 2012.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 5, Issue 10, October 2017

17. N. Sundaram, T. Brox, and K. Keutzer, "Dense point trajectories by GPU-accelerated large displacement optical flow," in European Conference on Computer Vision, 2010.
18. AlirezaSarveniazi, "An Actual Survey of Dimensionality Reduction" American Journal of Computational Mathematics, vol.4, pp.55-72, 2014.
19. S.Anu H Nair, P.Aruna, "Image Fusion Techniques For Iris And Palmprint Biometric System", Journal of Theoretical and Applied Information Technology, Vol. 61 No.3, ,pp.476-485, March 2014.
20. ArunGarg, Richika, "A Novel Approach for Brain Tumor Detection Using Support Vector Machine, K-Means and PCA Algorithm", International Journal of Computer Science and Mobile Computing, vol.4,no.8, pp. 457-474, August- 2015,.
21. Sibo Song, Vijay Chandrasekhar, Ngai-Man Cheung, Sanath Narayan, Liyuan Li, Joo- Hwee Lim , "Activity Recognition in Egocentric Life-Logging Videos", vol.9010 Lecture Notes in Computer Sciencepp 445-458, Springer 2014.