



7th International Conference on Advances in Computing & Communications, ICACC-2017, 22-24 August 2017, Cochin, India

## Robust Action Recognition framework using Segmented Block and Distance Mean Histogram of Gradients Approach

Vikas Tripathi<sup>a,b,\*</sup>, Durgaprasad Gangodkar<sup>b</sup>, Ankush Mittal<sup>b</sup>, Vishnu Kanth<sup>b</sup>

<sup>a</sup>Uttarakhand Technikal University, Dehradun, India-248007

<sup>b</sup>Graphic Era University, Dehradun, India-248002

---

### Abstract

This paper propose two novel algorithms, segmented block of mean image with normalization and distance mean histogram of gradients for generating descriptor. Feature analysis and classification done with the help of Random forest. Our approach performs better than benchmark, gradient based approaches with average accuracy 56.59% on HMDB dataset. We have also tested our approach on ATM video dataset. Video sequences have been analyzed by varying block size of mean image and Number of frames for mean image. Average accuracy 94.5% has been achieved during testing on ATM dataset, where proposed framework has been able to recognize activities efficiently.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 7th International Conference on Advances in Computing & Communications.

*Keywords:* HOG; Activity recognition; HMDB; Random forest;

---

### 1. Introduction

In recent years, computer vision society has explored a variety of application in human computer interaction, sport analysis, medical expert systems and content based video/ image retrieval etc. The development of these applications is based on standard computer vision approaches, such as action recognition, object identification and localization, image classification and application oriented activity analysis. However, it still remains a challenging task to develop frameworks which suits to the requirements in real-world scenarios [1-5]. Human activity analysis has been attracting a comprehensive range of consciousness over the last few years owing to its extensive range of implementation in event detection.

\* Corresponding author : VIKAS TRIPATHI. Tel.: 9634596042

*E-mail address:* vikastripathi.be@gmail.com

Even though copious theories have emanated, human activity analysis is still a challenging problem by considering the inconsistency in appearance, posture, vesture and background clutter. Significant research has been devoted to tracking human movement in images and videos as discrepancies in illumination, shadows as well as inter and intra person occlusion renders this a difficult task. Applications such as surroundings surveillance, activity recognition etc. necessitates methods for analyzing human actions in diverse scenarios. Such conditions acquaint challenging problems that have been observed in computer vision in the past. Strength in providing personalized support for different applications and fields, this research field has caught the interest of several computer science communities. Analyzing human activities on the basis of videos has attained a lot of attention amongst the research workers [5].

The current approaches for action recognition primarily focus on encoding motion as a feature [6]. Other recognition frameworks that search for a semantically efficient and more accurate representation have focused on individual pose, face, specific object and scene identification[7][8], which itself is difficult procedure. In this paper we propose two new approaches segmented block of mean image with normalization and distance mean histogram of gradients (DMH). Fusions of these two individual approaches provide great results in the most efficient and resource effective manner. The Video is divided into frames and two individual descriptors are applied to obtain pure motion into a single frame from a series of three images, the resultant motion image is very effective which generate the activity in a dense representation. We have demonstrated our results on the well versed dataset such as ATM, HMDB and achieved great results in comparison to the well-known approaches. Random forest is used for classification as it is has a distinct tree approach to precisely eliminate noise in dataset and provide better results than other classifiers. The paper further consist four sections review of Literature, applied Methodology, Result analysis and Conclusion drawn from paper.

## 2. Literature Survey

Lot of research has been done in the field of computer vision and various approaches have been proposed to detect and recognize action. Human-action recognition is more complicated as it involves very slight difference in motion but varies in the meaning of the action [8][9]. Bobick and Davis [10] have proposed a recognition approach that uses motion semantics based recognition by first representing the motion in form of spatial pattern and then describing motion of the object. Scale invariant feature transform is extensively applied descriptors by several researchers. In this every video volume is segmented into several blocks. For each block by applying either oriented gradients or optical flow, descriptor values are generated and finally all values are concatenated of various nearby blocks to get the final descriptor. Laptev et al. [11] proposed the grouping of two descriptors, histogram of oriented Gradient (HOG) and histogram of flow (HOF) where flow is optical flow, for effective recognition. Dollar et al.[12][13] performed comparative analysis of various local descriptors based on normalization of pixel intensities, gradient information and flow analysis. Scovanner et al. [14] and Klaser et al. [15] presented feature descriptor based on three dimensional gradients where oriented gradients of temporal analysis are also designed. Everts et al. [16] proposed to analysis of colour channels for effective identification of action. Further Wang et al. [17] [18] and Jiang et al. [19] effectively used information of flow and shown improvement in accuracy in comparison to fusion of HOG and HOF approach. Use of this approach with dense trajectory provides better outcomes as compared to dense trajectories alone [20]. Further divergence, curl and shear features based descriptor [21] used to recognize activity in video. S. Sadanand and Jason J. Corso [22] proposed, action bank using object bank method to generate bank of activities which is a upper level depiction of features in spatiotemporal environment.

Several researchers have used tree based ensemble learning. Bag of words [24] (BOF) and boosting [25] are two widely used approaches in current research. space time interest points [26] based methods, such as local ternary patterns [27], dense points [28] and discriminative category specific features [29] all have used BOF on features. Breiman [30] proposed random forests (RF) where each node is separated by using the most excellent among a subset of predictors randomly chosen at that node. This approach performs efficiently compared to many other classifiers like SVM and neural networks etc. It is also robust against over fitting. Bosch et al. [31] used RF for classification of feature vectors which get extracted by using decomposition scheme based on spatial pyramid. As [30] has proved that performance of Random forest using random number of features is better than bagging approach. Hence we have used RF for classification.

Download English Version:

<https://daneshyari.com/en/article/6902311>

Download Persian Version:

<https://daneshyari.com/article/6902311>

[Daneshyari.com](https://daneshyari.com)