www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

Vol. 03, Issue 05, May 2023, pp : 1169-1173

e-ISSN : 2583-1062

Impact Factor : 5.725

ROBUST FEATURE-BASED AUTOMATED MULTI-VIEW HUMAN ACTION RECOGNITION SYSTEM

Dr. M. Praneesh[1], B. Udayakumar[2], M. Selvakumar[3]

[1]Assistant Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamilndu, India

[2,3]PG Scholar, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, Tamilndu, India

DOI: https://www.doi.org/10.58257/IJPREMS31409

ABSTRACT

Automated human action recognition has the potential to play an important role in public security, for example, in relation to the Multiview surveillance videos taken in public places, such as train stations or airports. This paper compares three practical, reliable, and generic systems for Multiview video-based human action recognition, namely, the nearest neighbor classifier, Gaussian mixture model classifier (GMM), and the nearest mean classifier. To describe the different actions performed in different views, view-invariant features are proposed to address Multiview action recognition. These features are obtained by extracting the holistic features from different temporal scales which are modeled as points of interest which represent the global spatial-temporal distribution. Experiments and cross-data testing are conducted on the KTH, WEIZMANN, and MuHAVi datasets. The system does not need to be retrained when scenarios are changed which means the trained database can be applied in a wide variety of environments, such as view angle or background changes. The experiment results show that the proposed approach outperforms the existing methods on the KTH and WEIZMANN datasets.

Keywords: Gaussian mixture model classifier, human action recognition, temporal domain, Features.

1. INTRODUCTION

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually, Image Processing system includes treating images as two-dimensional signals while applying already set signal processing methods to them. Recognition on the basis of how an individual walks has attracted interest from defense and other agencies for remote surveillance or infrared recordings of movement in an area under covert surveillance. The technology essentially involves dynamic mapping of the changing relationships of points on a body as that person moves. It centered on the 'stride pattern' of a sideways silhouette, with a few measurement points from the hip to feet. More recent research appears to be encompassing people in the round and seeking to address the challenge of identification in adverse conditions.

2. EXISTING APPROACH

Willems *et al.* proposed the spatio-temporal domain which is an extension of the SURF descriptor. Schuldt *et al.* and Dollar *et al.* described sparse spatio-temporal features to deal with the complexity of human action recognition. Schuldt *et al.* proposed the representation of action using 3D spatio-temporal interest points captured from video frames. Schuldt also produced a histogram of informative words for each action adopting the codebook and bag-of-words (BOW) approach. Most of the early work assumes that the action is captured from a static viewpoint without any camera movement. These approaches only extend the system from a single viewpoint to a multi-view dataset. Existing approaches have difficulty ensuring the performance of the classifier when the viewpoint or environment changes.

3. MODELING AND ANALYSIS

3.1 Video Acquisition

Testing videos are collecting from KTH database. After the video acquisition, frames extraction is performed. for this method frames were extracted based on pixel wise.

3.2 Moving Object Localization

In this phase, moving objects are detected and localized based on Gaussian mixture model. In action recognition, detecting and segmenting the foreground object without the noise produced by camera movements, zoom, shadows etc. is difficult. To do this, the model can be divided into the following steps. Firstly, the Gaussian mixture model (GMM) is used to construct the background and obtain the silhouette by background subtraction. Secondly, the

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 05, May 2023, pp : 1169-1173

e-ISSN : 2583-1062

Impact Factor : 5.725

Prewitt edge detector can be used to segment the objects from the foreground. The GMM is a common and robust method in background construction. For the purpose of action recognition in a complex scene condition, the GMM is used to build the background image. Finally, blob analysis based segmented moving objects are localized

3.3 Feature Extraction & Reduction

In this phase, localized moving objects features are extracted and reduced. Features mean detection of interest points in localized moving object and find strongest features in object via Harris Spatio temporal corner detector.

3.4 Action Recognition

The human action recognition process is done over the extracted features. The main novelty here is the adoption of Nearest Mean Classifier (NMC). NMC is applied over the features and the action recognition is done.

3.5 Proposed System

The proposed system has following steps,

1. First, videos are acquiring from database.
2. Second, frames extraction is performing to further process.
3. Third, implementing of detection and localization of moving objects in each frame based on gaussian mixture model, prewitt filter and blob analysis.
4. Fourth, interest feature points are extracted from localized moving object and strongest key points are extracted based on Harris Spatio temporal corner detector.
5. Finally, Nearest Mean Classifier is proposed to recognize the human actions based on interest features.

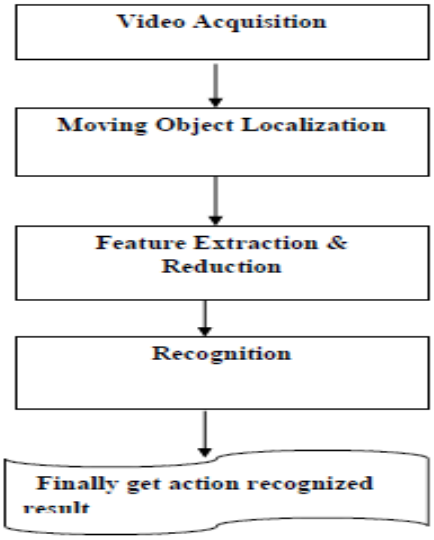

Figure1: Proposed System Architecture.

4. RESULTS AND DISCUSSION

In this Section results and discussion of the study is carried out Nearest neighbor algorithm were implemented the results were presented in figure 2 to

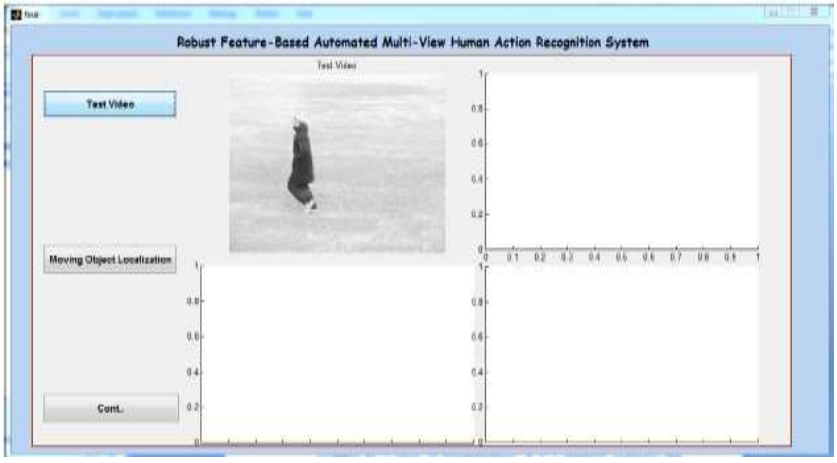

Figure2: Test Input video

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 05, May 2023, pp : 1169-1173

e-ISSN : 2583-1062

Impact Factor : 5.725

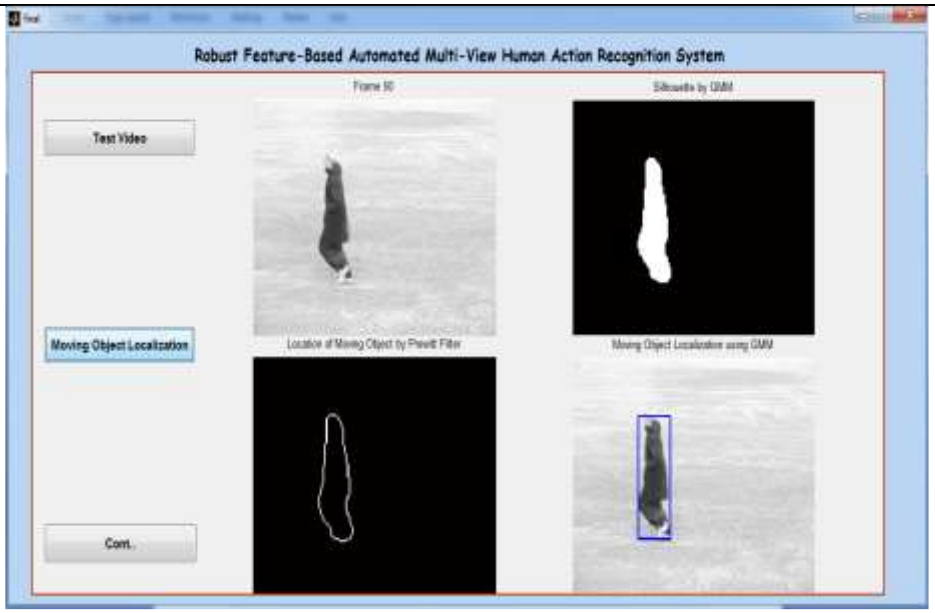

Figure 3: Moving object Localization

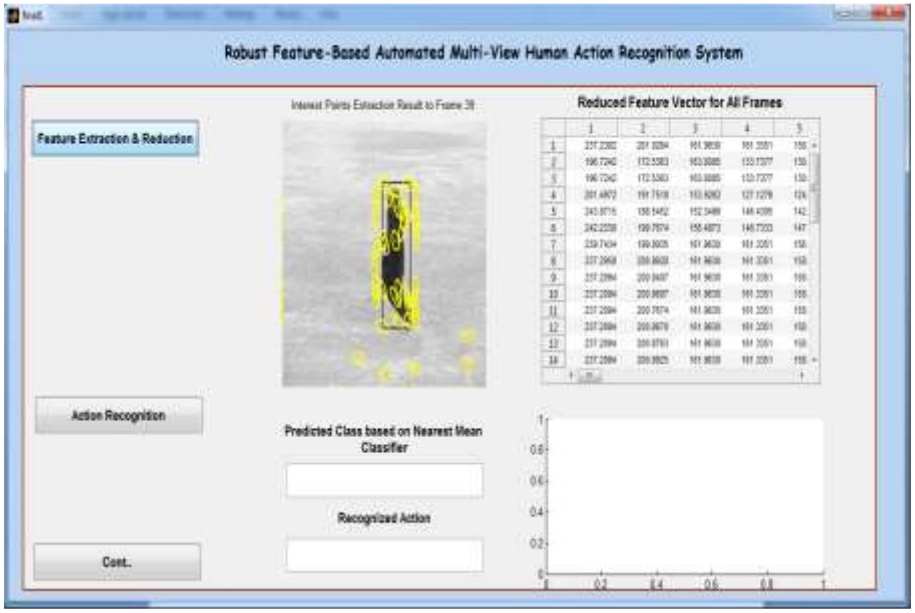

Figure 4: Moving object Localization with frame calculation

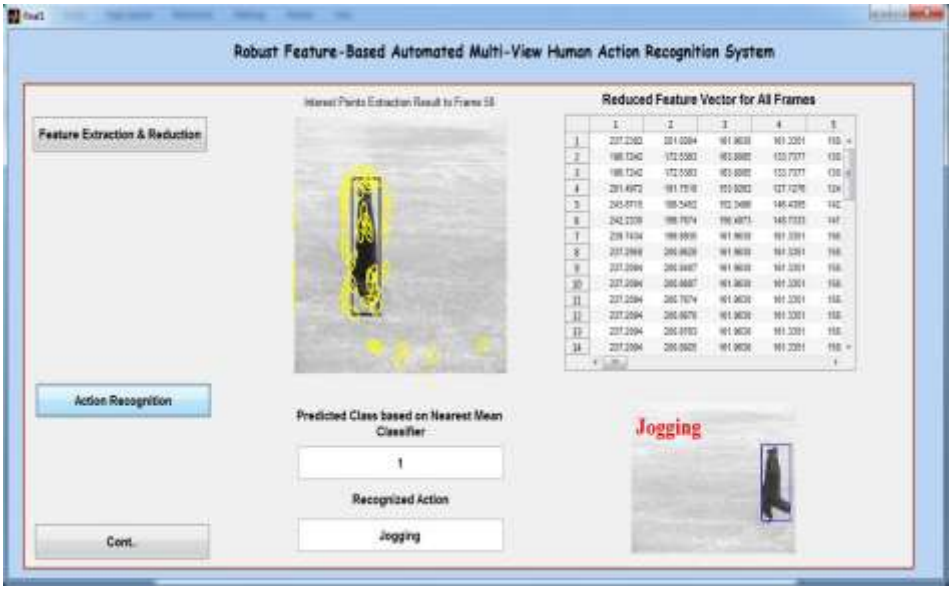

Figure 5: Recognition Action

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 05, May 2023, pp : 1169-1173

e-ISSN : 2583-1062

Impact Factor : 5.725

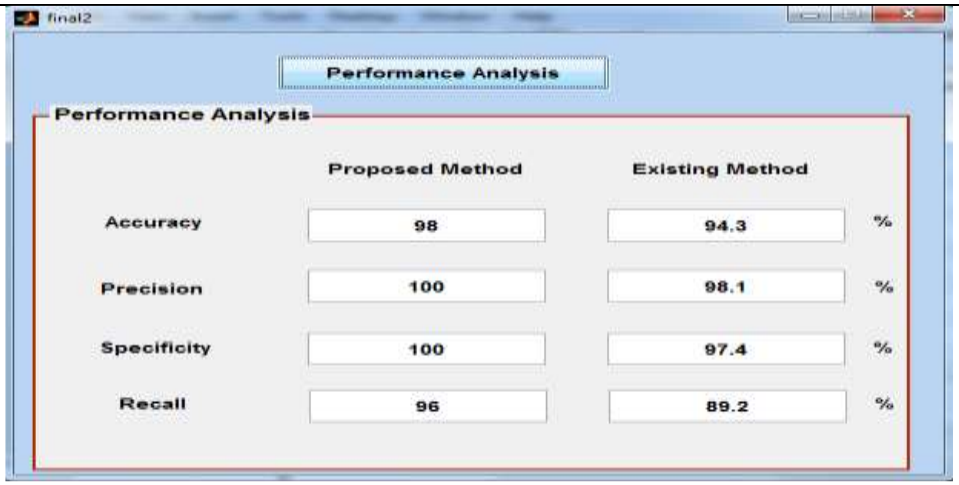

Figure 6: Performance Analysis

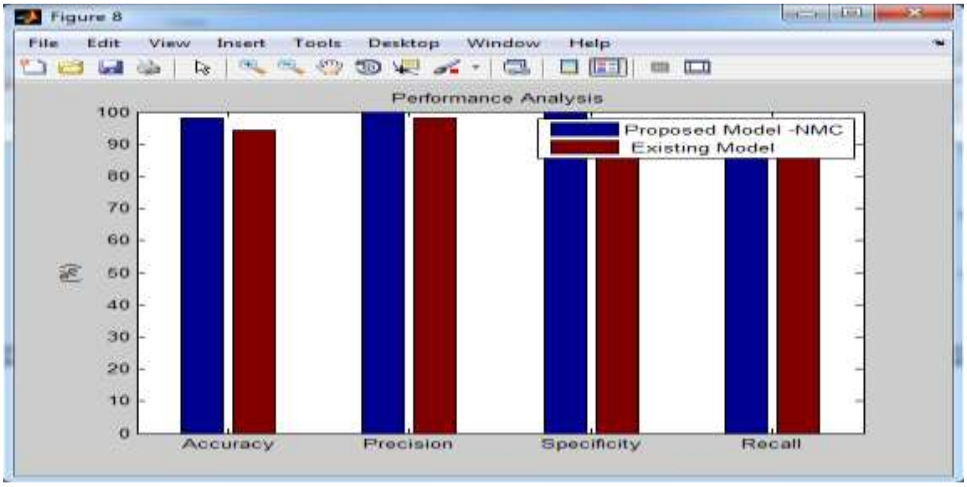

Figure 7: Performance Chart

5. CONCLUSION

This paper presents an approach for real-world applications which automatically labels the beginning and ending of an action sequence. The system uses the proposed view-invariant features to address multi-view action recognition from different perspectives for accurate and robust action recognition. The view-invariant features are obtained by extracting holistic features from different temporal scale clouds, which are modeled on the explicit global, spatial and temporal distribution of interest points. The experiments on the KTH and WEIZ- MANN datasets demonstrate that using view-invariant features obtained by extracting holistic features from clouds of interest points is highly discriminative and more robust for recognizing actions under different view changes. The experiments also show the proposed approach performs well with cross-tested datasets using previously trained data, which means there is no need to re-train the system if the scenario changes.

6. REFERENCES

[1] K. G. Derpanis, M. Sizintsev, K. J. Cannons, and R. P. Wildes, ``Action spotting and recognition based on a spatiotemporal orientation analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 3, pp. 527_540, Mar. 2013.

[2] A. Gilbert, J. Illingworth, and R. Bowden, ``Action recognition using mined hierarchical compound features," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 5, pp. 883_897, May 2011.

[3] L. Liu, L. Shao, X. Zhen, and X. Li, ``Learning discriminative key poses for action recognition," IEEE Trans. Cybern., vol. 43, no. 6, pp. 1860_1870, Dec. 2013.

[4] Y. Yang, I. Saleemi, and M. Shah, ``Discovering motion primitives for unsupervised grouping and one-shot learning of human actions, gestures, and expressions," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 7, pp. 1635_1648, Jul. 2013.

[5] Z. Jiang, Z. Lin, and L. S. Davis, ``Recognizing human actions by learning and matching shape-motion prototype trees," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 3, pp. 533_547, Mar. 2012.

[6] K. Guo, P. Ishwar, and J. Konrad, ``Action recognition from video using feature covariance matrices," IEEE Trans. Image Process., vol. 22, no. 6, pp. 2479_2494, Jun. 2013.

[7] Y. Chen, Z. Li, X. Guo, Y. Zhao, and A. Cai, ``A spatio-temporal interest point detector based on vorticity for action recognition," in Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW), Jul. 2013, pp. 1_6.

[8] S. Samanta and B. Chanda, ``Space-time facet model for human activity classi_cation," IEEE Trans. Multimedia, vol. 16, no. 6, pp. 1525_1535, Oct. 2014.

[9] Z. Moghaddam and M. Piccardi, ``Histogram-based training initialization of hidden Markov models for human action recognition," in Proc. 17th IEEE Int. Nat. Conf. Adv. Video Signal Based Surveill. (AVSS), Sep. 2010, pp. 256_261.

[10] Y. Wang, L. Wu, and X. Huang, ``Action recognition using tri-view constraints," in Proc. 8th IEEE Int. Conf. Adv. Video Signal-Based Surveill. (AVSS), Aug. 2011, pp. 107_112.

[11] M Praneesh and Jaya R Kumar. Article: Novel Approach for Color based Comic Image Segmentation for Extraction of Text using Modify Fuzzy Possiblistic C-Means Clustering Algorithm. IJCA Special Issue on Information Processing and Remote Computing IPRC(1):16-18, August 2012. Published by Foundation of Computer Science, New York, USA.

[12] M. N. Kumar and D. Madhavi, ``Improved discriminative model for view- invariant human action recognition," Int. J. Comput. Sci. Eng. Technol., vol. 4, no. 3, pp. 1263_1270, 2013.

[13] D. Wu and L. Shao, ``Silhouette analysis-based action recognition via exploiting human poses," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 2, pp. 236_243, Feb. 2013.

[14] S. Singh, S. A. Velastin, and H. Ragheb, ``MuHAVi: A multicamera human action video dataset for the evaluation of action recognition methods," in Proc. 17th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS), Sep. 2010, pp. 48_55.

[15] G. Willems, T. Tuytelaars, and L. Van Gool, ``An efficient dense and scale- invariant spatio-temporal interest point detector," in Computer Vision_ ECCV. Berlin, Germany: Springer, 2008, pp. 650_663.