Human Activity Recognition and Classification Using Local Invariant Feature Extraction

G P Hegde
Department of Information Science and Engineering SDMIT Ujire

Nireeksha, Nithini B K
Department of Information Science and Engineering SDMIT Ujire

Soundarya M B, Supreetha B V
Department of Information Science and Engineering SDMIT Ujire

***Abstract*:- Human action recognition from realistic videos attracts more attention in many practical applications such as on-line video surveillance and content-based video management. Single action recognition always fails to distinguish similar action categories due to the complex background settings in realistic videos. This paper explains the various studies on human activity recognition and also focuses on steps needed for recognition of human activity using local invariant methods and classification of human activity frames.**

General Terms:- Your general terms must be any term which can be used for general classification of the submitted material such as Pattern Recognition, Security, Algorithms et. al.

Keywords:- Activity recognition, Local Invariant Method

1. INTRODUCTION

Applications such as surveillance, video retrieval and human-computer interaction require methods for recognizing human actions in various scenarios. Typical scenarios include scenes with cluttered, moving backgrounds, non-stationary camera, scale variations, individual variations in appearance and cloth of people, changes in light and view point. This work demonstrates how the action recognition can be achieved using local measurements in terms of spatio temporal interest points. Such features capture local motion events in video and can be adapted to the size, the frequency and the velocity of moving patterns, hence, resulting in video representations that are stable with respect to corresponding transformations. In this work systematic study on human activity recognition using local invariant methods are described. Comparison of various studies understands that local feature is preserved for accurate recognition of human activity. The redundant information is removed during activity detection and classification of human activity frames also has been carried out by support vector machine (SVM) classifier.

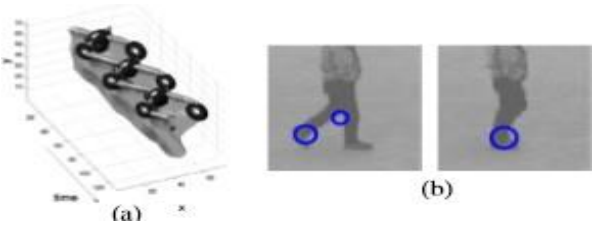


Figure 1: Local space-time features detected for a walking pattern

: (a)3-D plot of a spatio-temporal leg motion (up side down) and corresponding features (in black) (b) Features overlaid on selected frame sequence.

Figure2. Action and scenarios. Data base (available on request): examples of sequences corresponding to different types of action.

2. LITERATURE REVIEW

Ren and Xu (2002) [1] presented a new system for teachers' natural complex action recognition in the smart classroom in order to realize an intelligent cameraman and virtual mouse. First, the system proposes a hybrid human model and employs a second order B-spline function to detect the two shoulder joints in the silhouette image to obtain the basic motion features including the elbow angles, motion parameters of the face and two hands. Then, a primitive-based coupled hidden Markov model (PCHMM) is presented for natural context-dependent action recognition. Last, some comparison experiments show that PCHMM is better than the traditional HMM and coupled HMM.

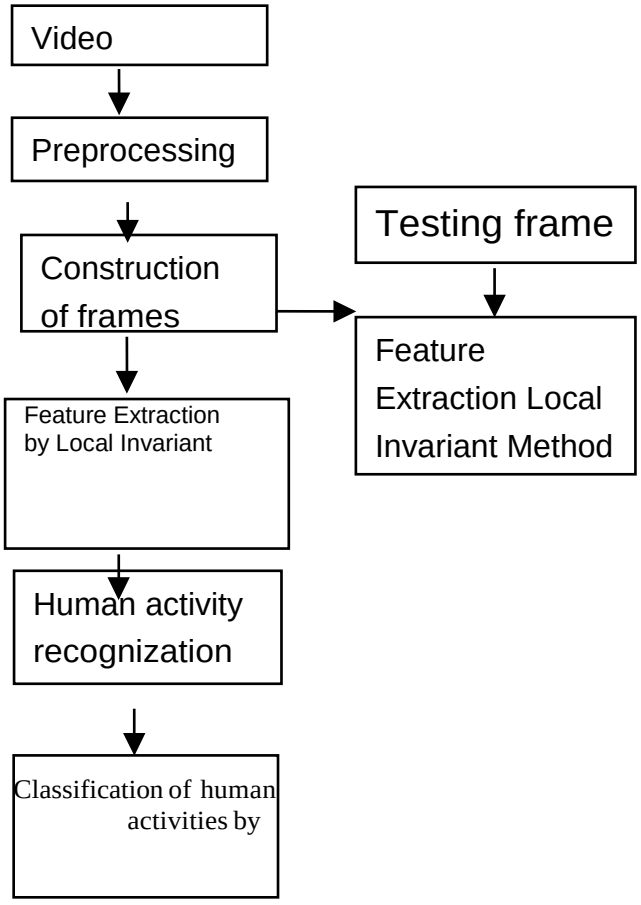
Akilandasowmya and P.Sathiya [2] describes that

Human activity recognition is an important research area of computer vision. There is an urgent mechanism to automatically detect and retrieve semantic events in videos based on video contents. Low-level video sequence contents is translated into high level video to sequence content is a interesting research topic in recent years. Its application include automated video surveillance schemes, intensive care system, airports, analysis of physical condition of people and variety of systems. Which include human-computer interfaces.

Arie et al. (2002) [3], develop a novel method for view-based recognition of human action/activity from videos. By observing just a few frames, we can identify the activity that takes place in a video sequence. The basic idea of multidimensional indexing method is that activities can be positively identified from a sparsely sampled sequence of a few body poses acquired from videos.

Davide Anguita et al.[4] presented an Activity-Based Computing aims to capture the state of the user and its environment by exploiting heterogeneous sensors in order to provide adaptation to exogenous computing resources. When heterogeneous sensors are attached to the subject's body, they permit continuous monitoring of numerous physiological signals.

3. DESIGN ANDIMPLEMENTATION



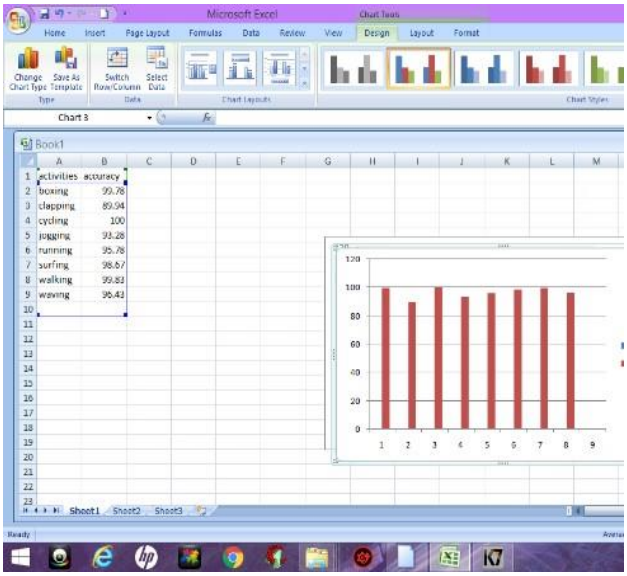
In Figure 3 video sequence is considered as a spatio-temporal intensity volume from which motion cues of human actions are firstly extracted through differencing adjacent frames .Backgrounds are simultaneously suppressed without suffering from expensive computations resulting from tracking or background subtraction. Then construct a spatio-temporal Laplacian pyramid Construction of spatio-temporal Gaussian pyramid and Laplacian pyramid.as follows. The obtained volumes with DOF are repeatedly filtered with Gaussian weighting functions and sub sampled to generate volumes with regularly reduced resolutions. These comprise a series of low-pass filtered copies of original volumes, namely a spatio-temporal Gaussian pyramid, in which the bandwidth decreases at one-octave per step. To directly represent the volumes in terms of voxel intensity values, however, is inefficient due to the high correlation among these voxels. Therefore, the smoothed 3D volumes are decomposed into a set of spatio-temporal band-pass filtered volumes called a spatio-temporal Laplacian pyramid by differencing adjacent levels of the Gaussian pyramid.

Features with different sizes are appropriately localized at each level of the pyramid, as the band-pass filtered volume represents a particular fineness of detail at each scale. Subsequently, we apply a feature extraction step. A bank of 3D Gabor filters is then applied to the original volume and each level of the Laplacian pyramid to enhance edge and orientation information. To extract invariant and discriminative features, a nonlinear max pooling technique is performed within bands of Gabor filters and over spatio-temporal neighborhoods, resulting in robustness to spatial and temporal shifts, partial occlusions and noise. Our feature extraction process from a raw video sequence is illustrated in SVM classification combined with motion descriptors in terms of local features (LF) and feature histograms (HistLF) define two novel methods for motion recognition. In this section we evaluate both methods on the problem of recognizing human action and

compare the performance to other approaches using alternative techniques for representation and classification. human activity is recognized using local invariant methods and classification of human activity frames also carried out by support vector machine (SVM) classifier.

2. RESULT

Figure 3: Architecure of Human Activity Recognition and Classification.



3. CONCLUSION

Multiple actions of human are recognized and also proposed method provide clear separation between leg actions and arm action. Confusion between different human activities like walking, jogging etc are classified and performance evaluation have been carried out. Based on activity accuracy is calculated.

Figure 4 (top) shows recognition rates for all of the methods. To analyze the influence of different scenarios we performed .

Based on activities accuracy is calculated that is shown in column chart.

4. REFERENCES

[1] Guangyou Xu, Human action recognition in smart classroom , IEEE International conference

[2] G.Akilandasowmya, P.Sathiya, P.AnandhaKumar Human action recognition in research area of computer vision , IEEE International Conference on automatically detect and retrieve semantic events in video, 2015 Seventh International Conference on,15-17 Dec.2015

[3] Arie et al. Human activity recognition using multidimensional indexing IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 24,Issue: 8, Aug 2002.

[4] Davide Anguita et al. Human Activity Recognition on Smartphone's Using a Multiclass Hardware-Friendly Support Vector Machine volume 7657.nterfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

[5] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[6] Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[7] Y.T. Yu, M.F. Lau, "A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions", Journal of Systems and Software, 2005, in press.

[8] Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S.Mullender