A LIGHTWEIGHT ENCRYPTION METHOD FOR PRIVACY PROTECTION IN SURVEILLANCE VIDEOS

V.K.J. Sri Pujythaa E.Suba Lakshmi Mrs.P.Brundha.
Deptof CSEDept of CSE HOD/ Dept of CSE
Francis Xavier Engineering College

ABSTRACT:Video surveillance systems are becoming important for crime investigation and by increasing the number of cameras installed in public space. Many cameras are installed at fixed positions to observe a wide and complex area. Detection of suspicious human behavior is important. Due to random nature of human movements, detecting the suspicious human behavior can be very difficult. An approach to the problem of automatically tracking people and detecting unusual or suspicious movements we use Closed Circuit TV (CCTV) videos. We are introducing a system that works for surveillance systems installed in indoor environments like entrances/exits of buildings, corridors, etc. This work presents a framework that processes video data obtained from a CCTV camera fixed at a particular location.

INTRODUCTION:

Investigation and so the number of surveillance cameras installed in public space is increasing. Many cameras installed at fixed positions are required to observe a wide and complex area, so observation of the video pictures by human becomes difficult. So there is a need for automation and dynamism in such surveillance systems. In order to allow the different users (operators and administrators) to monitor the system selecting different Quality of Service (QoS) are required depending on the system status and to access live and recorded video from different localizations i.e. from their mobile devices. More concretely, in Internet Protocol (IP) surveillance systems some resources involved are limited or expensive. So a technology using automatic detection of intruders (using image processing systems) and automatic alert systems will provide competitive advantage for surveillance systems.

Extended version of programming paradigms have allowed the dynamism and flexibility of distributed environments. Thus Service-Oriented approaches is providing lots of developing decoupled applications in heterogeneous networks by defining their concept of service.

SOA context has a service that receives and sends messages through well-defined interfaces, allowing building more complex applications . This concept can be applied to QoS-aware (Quality of Service) systems, to ease the configuration and reconfiguration of applications. This paper addresses recognition of natural human actions in diverse and realistic video settings. This is challenging but the important subject has mostly been ignored in the past due to several problems. It is the lack of realistic and annotated video data sets. Its contribution is to address the limitation and to investigate the use of movie scripts. We evaluate alternative methods from scripts and show benefits of a text-based classifier .Using the retrieved action samples for visual learning ,we next turn to the problem of action classification in video. They present a new method for video classification that builds upon and extends several recent ideas including local space-time features, space-time pyramids and multichannel nonlinear SVMs. Their experimental results demonstrate that post-processing techniques can significantly improve the

foregroundsegmentation masks produced by a BGS algorithm. They provided recommendations for achieving robust foreground segmentation based on the lessons learned performing this comparative study. We address the problem of learning view-invariant 3Dmodels of human motion from motion capture data, in or-der to recognize human actions from a monocular video sequence with arbitrary viewpoint.

EXISTING WORK:

Laptev et al [1] proposed the Learning realistic human actions from movies.They addressed recognition of natural human actions in detail and realistic video settings. This is challenging but important subject has been mostly ignored in the past due to several problems.

Address the limitation and investigate the use of movie scripts is necessary. They evaluated alternative methods for action retrieval from scripts and show benefits of a text-based classifier.They presented a new method for video classification that builds upon and extends several recent ideas including local space-time features, space-time pyramids and multichannel nonlinear SVMs.

Their experimental results demonstrate that post-processing techniques can significantly improve the foreground segmentation masks produced by a BGS algorithm. They provided recommendations for achieving robust foreground segmentation based on the lessons learned performing this comparative study.
This method shows how to improve state of the art results on the standard KTH action data set by achieving 91.8%. Given the inherent problem of noisy labels in automatic annotation, Here it is particularly investigated and shown high tolerance of this method to annotation errors in the training set. Finally this method is applied to learn and to classify challenging action classes in movies and show promising results.

J.Nielbes et al [2] proposed Unsupervised learning of humanaction categories using spatial-temporal words.They addressed the problem of learning view-invariant 3Dmodels of human motion from motion capture data, in or-der to recognize human actions from a monocular video sequence with arbitrary viewpoint.

Now propose a Spatio-TemporalManifold (STM) model to analyze non-linear multi variate time series with latent spatial structure and we need to apply it . A novel alignment algorithm Dynamic Manifold Warping (DMW) and a robust motion similarity metric are proposed both in 2D and3D. DMW extends previous works on spatiotemporal alignment by incorporating manifold learning.

This model is applied in application domains like remote monitoring, control and surveillance. Classic approaches to real-time do not provide the flexibility and fault-tolerance required in new emerging environments. Their approach addresses these new challenges by combining concepts from the service oriented paradigm and distributed real-time systems.

Now evaluate and compare the approach to state-of-the-art methods on motion capture data and realistic videos. Experimental results demonstrate the effectiveness of our approach, which yields visually appealing alignment results, produces higher action recognition accuracy, and can recognize actions from arbitrary views with partial on conclusion.

P. Matikainen et al [3] proposed the Representing pairwise spatial and temporal relations for action recognition.They presented a model for quality-of-service (QoS) in distributed systems with real-time and fault-tolerance requirements. This model can be applied application domains like, for example, remote monitoring, control and surveillance.

Human action recognition in videos is a important problem with wide applications. State of the art approaches often to adopt the popular bag-of-features representation based on isolated local patches. Motion patterns such as object relationships are mostly discarded.

They Adopted global and local reference points. This approach operates on top of visual codewords derived from local patch trajectories, and therefore does not require accurate foreground-background separation, which is typically a necessary step to model object relationships. Through

an extensive experimental evaluation, They show that the proposed representation offers very competitive performance on challenging bench mark datasets, and combining it with the bag-of-features representation leads to substantial improvement. On Hollywood2, Olympic Sports, and HMDB51 datasets, they obtain 59.5%, 80.6% and 40.7% respectively, which are the best reported results to date.

The standard approaches to real-time systems do not provide flexibility and fault-tolerance required in emerging environments. It is needed to combine a high degree of dynamism with temporal predictability.

P. Natarajan et al [4] proposed the View and scale invariant action recognition uses multi view shape-flow models.Actions in real world applications typically take place in cluttered environments with large variations. They present an approach to simultaneously track and recognize known actions that is robust tovariations. In this approach first render synthetic poses from multiple view points using Mocap data. Represent them in a Conditional Random Field(CRF) using shape similarity and the transition potentials are computed using optical flow. They enhance these basic potentials with terms to represent spatial and temporal constraints and call our enhanced model the Shape,Flow,Duration-Conditional Random Field(SFD-CRF). .

This demonstrated approach on videos from multiple viewpoints and in the presence of background clutter. Then, based on the shadow model, an MRF model is constructed for shadow removal. First, although this method is a chroma-based method, they made the pre-classifier accurate and adaptive to the change of shadow by using the statistical features of shadow at the global level.This demonstrated approach on videos from multiple viewpoints and in the presence of background clutter.

Tracking information can make it global-level statistical information as robust. Second, they constructed an MRF model to represent the dependencies between the label of a pixel and the shadow models of its neighbors. After experiments the proposed method shows that it is efficient and robust

P. Yan et al [5] proposed Learning 4d action feature models for arbitary view action recognition.This is an approach to jointly learn a set of view specific dictionaries and a common dictionary for cross view action recognition. The set of view-specific dictionaries is used to learn specific views meanwhile the common dictionary is shared. This approach represents videos in each view using both the corresponding view-specific dictionary and the common dictionary. It encourages the set of videos taken from different views.

In this way, they can align view-specific features in the sparse feature spaces spanned by the view specific dictionary set and transfer the view-shared features in the sparse feature space spanned by the common dictionary. The incoherence between the common dictionary and the view-specific dictionary set and view-shared features separately. In addition, the learned common dictionary not only has the capability to represent actions from unseen views, but also makes this approach effective in a semi-supervised setting in the target view. Extensive experiments using the multi-view IXMAS dataset demonstrate that this approach out performs many recent approaches for cross-view action recognition.

First demonstrate a video surveillance system consisting of passive and active pan/tilt/zoom (PTZ) cameras that intelligently responds to scene complexity, capturing higher resolution video when there are fewer people and capturing lower resolution video as the number of pedestrians present in the scene increases automatically. Lastly they developed a behavior based-controllers for passive and active cameras, making these cameras to carry out multiple observation tasks in a parallel manner.

PROPOSED WORK.

This paper addresses recognition of natural human actions in diverse and realistic video settings. This is challenging but important subject has been mostly ignored in the past due to several problems. One of them is the lack of realistic and annotated video data sets. The first way is to address its limitation and to investigate the use of movie scripts . We evaluate alternative methods from scripts and show benefits of a text-based classifier .Using the retrieved action samples for visual

learning ,we next turn to the problem of action classification in video. We present a new method for video classification that builds upon and extends several recent ideas including local space-time features, space-time pyramids and multichannel nonlinear SVMs. Their experimental results demonstrate that post-processing techniques can significantly improve the foreground segmentation masks produced by a BGS algorithm. They provided recommendations for achieving robust foreground segmentation based on the lessons learned performing this comparative study. We address the problem of learning view-invariant 3Dmodels of human motion from motion capture data, in or-der to recognize human actions from a monocular video sequence with arbitrary viewpoint. We propose a Spatio -Temporal Manifold (STM) model to analyze non-linear multivariate time series with latent spatial structure and apply it to recognize actions in the joint-trajectories space. Based on STM, a novel alignment algorithm Dynamic Mani-fold Warping (DMW) and a robust motion similarity metric are proposed for human action sequences, both in 2D and3D.

ALGORITHM USED

Absolute Differential Estimation.

This model is used to identify the absolute differential image by comparing the background model and incoming video frame at each frame. If any changes will be detected that fame is taken as motion detected image.

Cauchy Distribution Model.

After calculating Absolute differential image the Cauchy distribution model is used to detect the differentiation between the moving objects and the background region can be achieved using the Cauchy distribution model at each pixel.

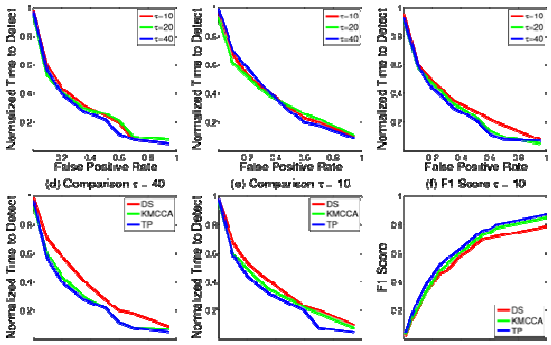

FIGURE 1: MVHAUS-PI-AMOC

MECHANISM:

User Authentication For Application.

User authentication is a way of identifying the user and verifying the user to see wether they are allowed to access some restricted service. The main aim of this modules is to authenticate the user to application to view the motion detected image. This modules include username and password for authentication to application Detecting Image using Cauchy distribution Model

Detecting Image using Cauchy distribution Model.

The important aim of this module is to detect the motion or movement in the particular area. The motion detection is done using two models they are the Cauchy distribution model and Absolute Differential Estimation. Absolute Differential Estimation is used to compare the background frame and incoming video frame to see if any changes occur in incoming video frame .And the

Cauchy distribution Model is used to mainly detect the pixel of moving object in the detected incoming video frame.

VIEWINGTHE DETECTED IMAGE

Android application will receive a notification(GCM) in the project id which is registered in Google account. Application id will unique for each application. After receiving GCM alert from the server and the user needs to authenticate the application . This image can be viewed using the url link which is received from the gcm alert.

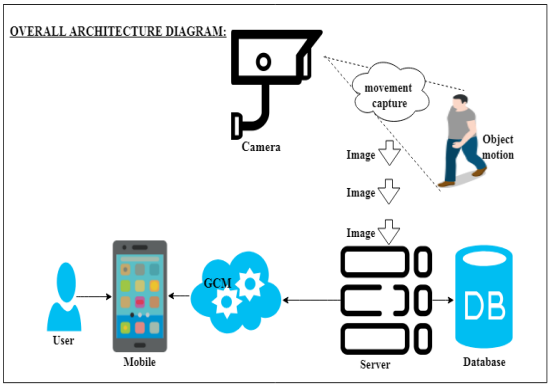

FIGURE2:SYSTEMARCHITECTURE

CONCLUSION:

By using this application deaf person can easily interact with normal person anywhere, and he can also use this application for mobile sign translation using VSR and by using UTF-7 he can communicate in daily activates without dialling number. We can use this application for mobile sign translation using VRS, and with UTF-7 communication can be made without dialling number Though this project has many added advantage, in future we like to upgrade this into the next level that is not only just viewing the captured image, we can also entire clip of what happened and what has been captured.

REFERENCES:

[1] Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, “Learning realistic human actions from movies,” in Proc.

[2] J. Nielbes, H. Wang, and L. Fei-Fei, “Unsupervised learning of human action categories using spatial-temporal words,” IJCV, vol. 79,pp. 299–318.

[3] P.Matikainen, M. Hebert, and R. Sukthankar, “Representing pairwise spatial and temporal relations for action recognition,” in Proc. ECCV, 2010, vol. 6311, pp. 508–521.

[4] P. Natarajan and R. Nevatia, “View and scale invariant action recognition using multiview shape-flow models,” in Proc. CVPR, 2008.

[5] P. Yan, S. M. Khan, and M. Shah, “Learning 4d action feature models for arbitary view action recognition,” in Proc. CVPR.