



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 11, November 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

A Survey Paper on CNN Based Enhanced Review Authenticity Platform

Swapnil Dhamdhere, Aditi Kulkarni, Ajinkya Chawane, Tanmay Dangat, Prof. Pradnya Kasture

Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

Department of Computer Engineering, RMD Sinhgad School of Engineering, Pune, India

ABSTRACT: A recent quick advance in the quality of capturing devices with less expensive but superior technology has made it so simple to capture photographs with excellent quality and good size. Videos are a collection of time-stamped sequential images. interval. Video can therefore give us more details about our object as the circumstances change throughout time. Videomanipulation manually is therefore virtually impossible. Hence, in order to process these videos, we need an automated device. One such attempt to track objects in films has been undertaken in this thesis. To automatically monitor the object in a video recording, numerous algorithms and technological advancements have been made. Basic object detection compares the current frame of video at the pixel level to a static background frame. The current method in this field tries to identify the important object in video frames first. Among other things, selecting appropriate features and models for identifying and tracking the desired item from a video is one of the primary challenges in object tracking. Intensity, form, colour, and feature points are some popular options for selecting appropriate features to categories visual objects.

KEYWORDS: Object detection, Frame difference, Vision and scene understanding, Background subtraction, Scale invariant feature transform.

I. INTRODUCTION

An important and difficult task in the field of computer vision is the attempt to identify, track, and detect objects over a series of images known as video. Instead of having human operators watch over computers, it is helpful to comprehend and describe object behavior. It seeks to find moving things in video files or security camera footage. Finding an object or objects using a single camera, many cameras, or a particular video file is known as object tracking. The development of high-quality imaging sensors, improvements to picture quality and resolution, and the exponential growth in computing power are all necessary for the development of new, effective algorithms and their use in object tracking[1].

Object tracking fundamentally entails estimating the location of a particular region in successive frames in a video sequence[7]. Properly detecting objects can be a particularly challenging task, especially since objects can have rather complicated structures and may change in shape, size, location and orientation over subsequent video frames. Various algorithms and schemes have been introduced in the few decades, that can track objects in a particular video sequence, and each algorithm has their own advantages and drawbacks. Any object tracking algorithm will contain errors which will eventually cause a drift from the object of interest. The better algorithms should be able to minimize this drift such that the tracker is accurate over the time frame of the application. In object tracking the important challenge that has to consider while the operating a video tracker are when the background is appear which is similar to interested object or another object which are present in the scene.

Motivation

The rapid improvement in technology makes video acquisition sensor or devices better in compatible cost. This is the cause of increasing the applications in different areas that can more effectively utilize that digital video. Digital videos are a collection of sequential images with a constant time interval. So there is more information is present in the video about the object and background are changing with respect to time. After studying the literature, it is seen that detecting and tracking of objects in a particular video sequence or any surveillance camera is a really challenging task in computer vision application. Video processing is really time consuming due to a huge number of data is present in video sequence. The area of video tracking is currently immense interest due to its implication in video surveillance, security, medical equipments, robotic systems. Video tracking offers a context for extraction of significant information such as scene motion, background subtraction, object classification, interaction of object with background and other objects from a scene, human identification, behavior of human with object and background, etc. Therefore, it is seen that there is a wide range of research possibilities are open in relation to video tracking.

II. REVIEW OF LITERATURE

Z.A. Khan, W. Sohn, "Real Time Human Activity Recognition System based on Radon Transform", IJCA Special Issue on Artificial Intelligence Techniques - Novel Approaches Practical Applications, AIT – 2011.: A real time human activity recognition system based on Radon transform (RT), Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) is presented.

Artificial Neural Nets (ANN) is used to recognize different human activities.

agadeesh B, Chandrashekar M Patil, "Video Based Action Detection and Recognition Human using Optical Flow and SVM Classifier", IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2016, India: The data extracted using optical flow is converted to binary image. Then Histogram of Oriented Gradient (HOG) descriptor is used to extract feature vector from the binary images. These feature vectors are given as training features to Support Vector Machine (SVM) classifier to prepare a trained model.

Aishwarya Budhkar, Nikita Patil, "Video-Based Human Action Recognition: Comparative Analysis of Feature Descriptors and Classifiers", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 6, June 2017: In this paper, video based action recognition is performed on KTH dataset using four combinations of two feature descriptors and two classifiers. The feature descriptors used are Histogram of Oriented Gradient Descriptor (HOG) and 3- dimensional Scale Invariant Feature Transform (3D SIFT) and classifiers used are Support Vector Machine (SVM) and K Nearest Neighbour (KNN). Features are extracted from frames of training videos using descriptor and clustered to form Bag-of-words model.

Chengbin Jin, Shengzhe Li, Trung Dung Do, Hakil Kim, "Real-Time Human Action Recognition Using CNN Over Temporal Images for Static Video Surveillance Cameras", Information and Communication Engineering, Inha University, Incheon, Korea: This approach predicts human actions using temporal images and convolutional neural networks (CNN). CNN is a type of deep learning model that can automatically learn features from training videos. Although the state-of-the-art methods have shown high accuracy, they consume a lot of computational resources. Another problem is that many methods assume that exact knowledge of human positions.

olliffe, I.T., "Principal component analysis", Springer Series in Statistics, 2nd ed., Springer, 2002.: The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables.

E. Mohammadi, Q.M. Jonathan Wu, M. Saif, "Human Activity Recognition Using an Ensemble of Support Vector Machines", 2016 IEEE International Conference on High Performance Computing Simulation (HPCS), July 2016: Human Activity Recognition Using an Ensemble of Support Vector Machines is employed to improve the classification performance by fusing diverse features from different perspectives. The Dempster-Shafer fusion and product rule from the algebraic combiners have been utilized to combine the outputs of single classifiers.

T. B., Hilton, A., and Kruger, V., "A survey of advances in vision-based human motion capture and analysis", Computer Vision and Image Understanding, vol. 104, pp. 90-126, 2006 Human motion capture continues to be an increasingly active research area in computer vision with over 350 publications over this period. A number of significant research advances are identified together with novel methodologies for automatic initialization, tracking, pose estimation, and movement recognition. Recent research has addressed reliable tracking and pose estimation in natural scenes. Progress has also been made towards automatic understanding of human actions and behavior.

III. OPEN ISSUES

In this thesis our aim is to improve the performance of object detection and tracking by contributing originally to two components (a) motion segmentation (b) object tracking. Automated tracking of objects can be used by many interesting applications. An accurate and efficient tracking capability at the heart of such a system is essential for building higher level vision-based intelligence. Tracking is not a trivial task given the non-deterministic nature of the subjects, their motion, and the image capture process itself. The objective of video tracking is to associate target objects in consecutive video frames. We have to detect and track the object moving independently to the background. In this there are four situations to be considered in the account: Single camera and single object,

Single camera and multiple objects,

Multiple cameras and single object,

Multiple cameras and multiple objects. From the previous section it is found that there are many challenges in detecting of an object and tracking of objects and also recognition for fixed camera network. To set up a system for automatic segmentation and tracking of moving objects in stationary camera videoscenes, which may serve as a foundation for higher level reasoning tasks and applications, and make significant improvements in commonly used algorithms. Finally, the aim is to show how to perform detection and motion-based tracking of moving objects in a video from a stationary camera.

Therefore, the main objectives are:

To analyze segmentation algorithm to detect the objects.

To analyze some tracking method for tracking the single objects and multiple objects.

IV. CONCLUSION

Object detection and tracking is an important task in computer vision field. In object detection and tracking it consists of two major processes, object detection and object tracking. Object detection in video image obtained from single camera with static background that means fixing camera is achieved by background subtraction approach. In this thesis, we tried different videos with fixed camera with a single object and multiple objects to see it is able to detect objects. Motion based systems for detecting and tracking given moving object of interest can be created. Using SIFT feature extraction first feature of the object and the frame has detected to match the interested object. Since for feature extraction, SIFT algorithm has been used so tracker is invariant to representation of interested object.

REFERENCES

- [1] Z.A. Khan, W. Sohn, "Real Time Human Activity Recognition System based on Radon Transform", IJCA Special Issue on Artificial Intelligence Techniques - Novel Approaches Practical Applications, AIT – 2011.
- [2] Jagadeesh B, Chandrashekar M Patil, "Video Based Action Detection and Recognition Human using Optical Flow and SVM Classifier", IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2016, India
- [3] Aishwarya Budhkar, Nikita Patil, "Video-Based Human Action Recognition: Comparative Analysis of Feature

Descriptors and Classifiers”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 6, June 2017

[4] Chengbin Jin, Shengzhe Li, Trung Dung Do, Hakil Kim, "Real-Time Human Action Recognition Using CNN Over Temporal Images for Static Video Surveillance Cameras", Information and Communication Engineering, Inha University, Incheon, Korea

[5] Jolliffe, I.T., "Principal component analysis", Springer Series in Statistics, 2nd ed., Springer, 2002.

[6] E. Mohammadi, Q.M. Jonathan Wu, M. Saif, "Human Activity Recognition Using an Ensemble of Support Vector Machines", 2016 IEEE International Conference on High Performance Computing Simulation (HPCS), July 2016

[7] T. B., Hilton, A., and Kruger, V., "A survey of advances in vision-based human motion capture and analysis", Computer Vision and Image Understanding, vol. 104, pp. 90-126, 2006.

[8] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.

[9] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 1, pp. 39-51, 2002.

[10] C. Wojek, P. Dollar, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 4, p. 743, 2012.

[11] H. Kobatake and Y. Yoshinaga, "Detection of spicules on mammogram based on skeleton analysis." IEEE Trans. Med. Imag., vol. 15, no. 3, pp. 235-245, 1996.

[12] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in ACM MM, 2014.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in NIPS, 2012. [7] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime multi-person 2d pose estimation using part affinity fields," in CVPR, 2017.

[14] Z. Yang and R. Nevatia, "A multi-scale cascade fully convolutional network face detector," in ICPR, 2016.

[15] C. Chen, A. Seff, A. L. Kornhauser, and J. Xiao, "Deepdriving: Learning affordance for direct perception in autonomous driving," in ICCV, 2015.

[16] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3d object detection network for autonomous driving," in CVPR, 2017.

[17] A. Dunder, J. Jin, B. Martini, and E. Culurciello, "Embedded streaming deep neural networks accelerator with applications," IEEE Trans. Neural Netw. & Learning Syst., vol. 28, no. 7, pp. 1572-1583, 2017



Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details