Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 8, August -2017

A PERSON DETECTION AND MULTI-VIEW VIDEO TRACKING K MEANS CLUSTERING OF FACES

[1]M.Senbagapriya, [2] Dr.P.Sumitra

[1]M.Phil Full time Rresearch Scholar, PG and Research Department of Computer Science
Vivekananda College of Arts and Sciences for Women (Autonomous),Elayamapalayam.

[2]Assistant Professor, PG and Research Department of Computer Science
Vivekananda College of Arts and Sciences for Women (Autonomous),Elayamapalayam.

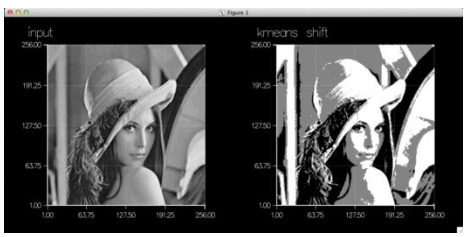
ABSTRACT:- *A generic methodology for the semi-automatic generation of reliable position annotations for evaluating multi-camera people-trackers on large video data sets. Most of the annotation data are automatically computed, by estimating a consensus tracking result from multiple existing trackers and people detectors and classifying it as either reliable or not.A small subset of the data, composed of tracks with insufficient reliability, is verified by a human using a simple binary decision task, a process faster than marking the correct person position. The proposed framework is generic and can handle additional trackers. In this thesis studied the most commonly used face edge detection techniques of K means clustering of faces. Higher-level edge detection techniques and appropriate programming tools only facilitate the process but do not make it a simple task.*

KEYWORDS: *Image processing, Digital Image Processing, Analog Image Processing Two dimensional signals.*

1.0 INTRODUCTION

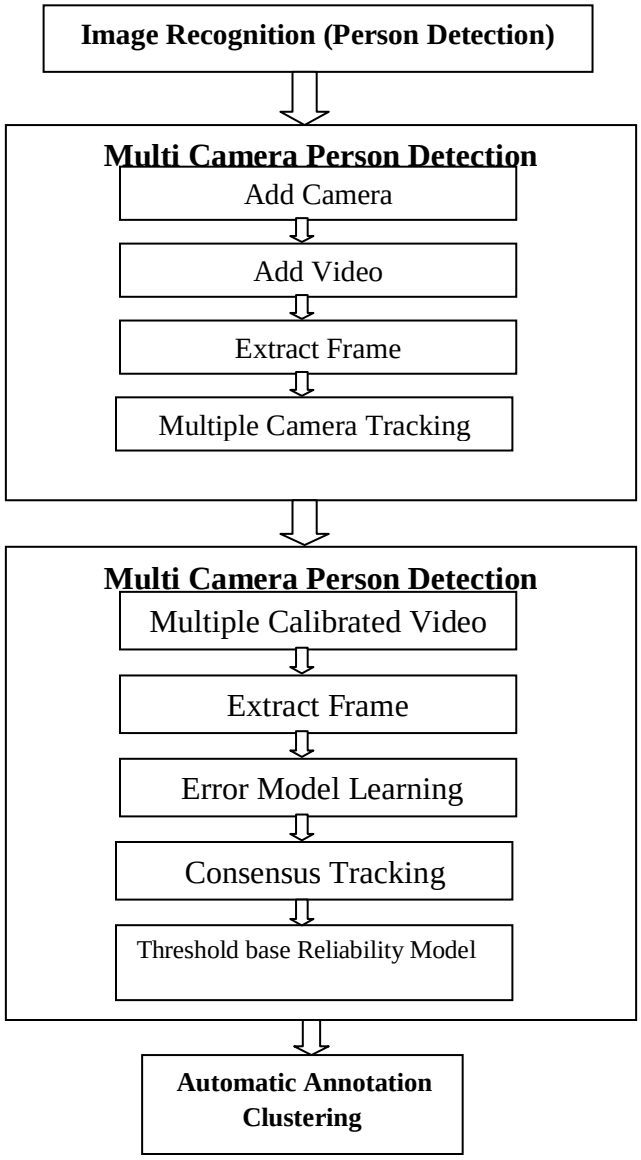
Identifying human actions in video is a challenging computer vision problem and the key technology for many potential video mining applications. Such applications become increasingly important with the rapid growth of personal, educational, and professional video data. The problem of automatic temporal annotation of realistic human actions in video using minimal manual supervision. To this end consider two associated problems weakly-supervised learning of action models from readily available annotations and temporal localization of human actions in test videos.To avoid the prohibitive cost of manual annotation for training, use movie track as a means of weak supervision. Scripts, however, provide only implicit, noisy, and imprecise information about the type and location of actions in video. They are addressing this problem with a kernel-based discriminative clustering algorithm that locates actions in the weakly-labeled training data. Using the obtained action samples, to train temporal action detectors and apply them to locate actions in the raw video data. The proposed system demonstrates that the weakly-supervised learning of action models leads to significant improvement in action detection.

The performance of proposed methodology, inspected the consensus tracking results for the full dataset frame-by-frame, correcting the generated annotations where needed. For solely the automatic part, we found that for the whole dataset, considering an accuracy of 60cm, 97.8% of the dataset (420000 frames) was correctly annotated. Using the proposed semi-automatic method, we show that it is possible to isolate most of the wrong annotations and to obtain up to 99% correct annotations. In this thesis present additional results for several performance specifications. The proposed approach allows comparing the results of a given tracker to the consensus of other trackers. The ultimate goal is to provide a graphical user interface which allows easy visualization of those portions of a data set with large discrepancies between tracking results. In some cases these discrepancies will be due to errors in the reference data. Based on our current internal graphical user interface, it would be easy to allow researchers to indicate segments of video data in which their new tracker outperforms the consensus tracker.



The performance of our method with experiments on a multi-camera video dataset of about 6 hours duration, showing scalability in semi-automatic annotation for long multi-camera sequences for the first time. This leads us to the final contribution of this paper: the annotated data itself, which is available in two versions: the first version results from the proposed procedure; the second from the exhaustive inspection to create ground truth. Since annotated multi-camera data sets of such length are not yet reported in literature, we hope these annotated sequences will be useful to the research community.

In this addition auto face identification of person tracking in video has drawn most research interests and led to many interesting applications. Since huge variation in the appearance of each character is found, it is a challenging problem. Existing methods evaluates promising results in clean environment, the performances are limited in complex movie scenes due to the noises generated during the face tracking and face clustering process. This study presents two schemes of global face- Person tracking based framework for robust person annotation and identification. In this thesis main contributions of this study include the first noise insensitive character relationship representation is incorporate, next study introduces an edit operation based graph matching algorithm, next complex person tracking are handled by simultaneously graph partition and graph matching and beyond existing person annotation approaches. The proposed schemes demonstrate state-of-the-art performance on video person annotation in various video frames.



2.0 K MEANS CLUSTERING

In addition to existing system implementation, the proposed system includes identifying face/names for different movies with different character names also. Enhanced Sobel Edge Annotation Algorithm (Enhnaced K-Means clustering) is applied for clustering process. And also, different person names for given face is tracked even if movies vary. Median filter concept is applied to remove the noise either before/after the clustering process.

2.1 PROPOSED ANNOTATION ALGORITHMS

- Frames from multiple movie files are consolidated (grouped) as if they are taken from single movie.
- For the given face (in the selected frames), the names appeared are grouped. For example if Name Jack from Frame 1 of movie 1 and Name George from frame 1 of Movie 2, occurred for the same face, then both are considered as same actor name.
- Likewise all the faces in all the frames are checked for combined name appearances in both movies taken for frame selection.
- Common names for same face data in two movie frames are compared for occurrences in multiple places and treated as same character.

2.2 ADVANTAGES

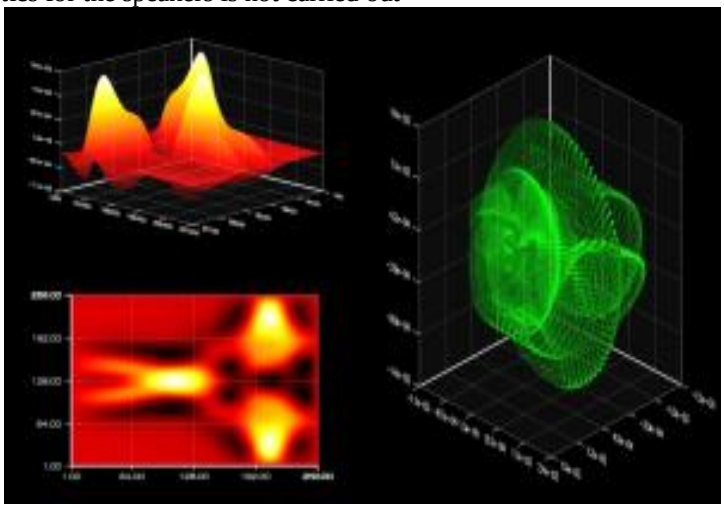
- Not sensitive to noises, since noise removal can be applied either before or after clustering process.
- Different character name for given face is tracked even if movies vary.
- Noise removal process is carried out.
- The sequential statistics for the speakers is also carried out.

2.3 OBJECTIVE OF THE PROPOSED SYSTEM

- Different type of the noise removal process is to be carried out for identifying the noise in the selected video for that clustering process is carried out in this noise removal.
- The previous systems are not studied such type of the noise removal and identifying images like process.
- The proposed system is carried out with different clustering for the noise removal.
- The proposed system also concentrates in detecting the face of the person who is found in the particular video.
- Total number of times the person image occur in that is been clearly calculated in this proposed system.

2.4 NEED FOR PROPOSED SYSTEM

- While continuous measures of the strength of relationship hold complete information, but it is highly sensitive to noises.
- Same person name for given face is tracked even if movies vary.
- Noise removal process is not discussed.
  The sequential statistics for the speakers is not carried out



3.0 MODULE DESCRIPTION

1. Recognizing Human face
2. Add pattern and Select snapshot
3. Video dataset
4. Split into frames
5. Add Annotation
6. Add and view person name
7. Persons face reorganization
8. Persons face in movies recognition
9. Person name wise clustering
10. K Means clustering of person name

3.1 RECOGNIZING HUMAN FACE

In this module the human face that is to be found out in the video should be noticed previously into the database that should be saved in the database. For storing that the camera should be added for taking the snaps of the human in particular position or the place where the camera located. The previously added camera details are also view in the grid view control that is done in this module. The different electrical consumables that is in the camera are also noticed in this module to locate the human image taken in clear manner this process is been called.

With out having searching image pattern image in the video is not possible to found the occurrence of the face in the video that user needs to found.

3.2 ADD PATTERN AND SELECT SNAPSHOT

In this module human face is chosen and then it is been added to the database for that the specific area of the human face is selected and stored as the pattern. With that selected pattern image the searching is carried out. For the selection of the snapshot that is taken in the camera will be noticed. The date and time the snap are taken is stored in the database.

The view is also provided for the pattern that is already stored that details are stored in the database and will be generated as the view using the grid view control. The selected pattern is done using the command click the specific camera is chosen and with that camera id the snap is been waited. View is generated using the grid view control and the snapshots stored will be shown in this grid view.

3.3 VIDEO DATASET

In this module the video is added that is in that video the person or the suspects face is been recorded that video will be stored in the database for the search. The total video that is stored in the database can be viewed in this view dataset using the grid view control.

3.4 SPLIT INTO FRAMES

This module the video or the image is applied into frames and then each image is split into frames and stored to the database for the future reference. For that process command click is used to split the image into frames is been done in this module. For doing this picture box control is used and explore control is to explore the split image to the user.

3.5 ADD ANNOTATION

In this module the movie id and the frame id is been selected using the listbox the previously added movie and the split frames are stored in this annotation module. The selected details are stored to the database. This module also has view that is designed using the grid view control. The added annotation is also having in this module.

3.6 ADD AND VIEW PERSON NAME

This module have specific segment that is person details are added to the database that is person to search for is added in this module. The view will be generated for the stored person details in this module. Select the movie id that is been choose from the list box the person who is appeared in the video is stored in this module. And additionally the character he got in that video is displayed in this module.

This also includes the annotation titles that are video added for the annotation and the frames they split for the annotation are shown in this module. This view is created using the grid view controls. This module also includes the separate view for displaying the person character name and the movie id are shown in this module.

3.7 PERSONS FACE REORGANIZATION IN SINGLE IMAGE

In this module the image is selected and using the red green and blue check boxes the image is selected to find the faces and that will be stored to the database. The size of the filter is also choosing using the list box control. Before that to load the image into the frame command click is made to add the image in the frame for that single image is been used.

3.8 PERSONS FACE RECOGNITION VIDEO FRAMES

In this module the video is been choose to in from that video the frames are selected to recognize the face in that video. Once the video is selected and the added frame in the image is been processed after the process command click is done. The face recognized in the single image is selected for that searching is done in this module and frames are identified and shown the occurrence of that frames that video. This also includes the filter size and the red green blue check boxes to choose the video frames. The selected image are stored in the data base after the save image click is processed.

3.9 PERSON NAME WISE CLUSTERING

This module movie id is selected and person name is also given for the searching of the person that is the person who is in the video. That is the name fetched for the finding of occurrence and the persons occurred as the group or cluster is also fetched from that video.

The person who is acted in the movie and how many times his image is displayed in that video is shown in this module. And this also includes the person name and name in each movie is also shown in this module. This also includes the person whose image occurred as the group and who is not with that group is listed in the message box. The frame selection details are also shown in this module. If the cluster forming details and the other details can be updated using update command click.

3.10 K MEANS CLUSTERING OF PERSON NAME

k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters. K Means clustering is applied for clustering process. And also, different character names for given face is tracked even if movies vary. Median filter concept is applied to remove the noise either before/after the clustering process. After the movie id selection, faces are clustered such that K Means clustering is applied with 'N' clusters is given as input.

Based on the color difference in the bitmap pixels, the face similarity is calculated. This form, the results of the K-Means clustering of the face image results are generated and displayed in Panel control. The clustered data is displayed in the Panel control with different color pixels. Then to display the cluster information, Tooltip option is implemented in this module.

4.0 CONCLUSION

A methodology for generating reference tracking data in long multi-camera videos, based on the consensus of a detector and several trackers. For multi-camera annotation, our methodology is the first to estimate the reliability of annotations, and to offer the possibility of balance accuracy and human-effort in the final annotation result. A novel probabilistic framework is to learn the error models of trackers, and how to apply them to estimate target position. Previous methods did not model the tracking error statistics of multiple trackers to increase the reliability of the estimated position.
The valuation of the accuracy and reliability of the proposed methodology 6 hour dataset, by a comprehensive visual inspection. Scalability of a semi-automatic methodology for annotation in multi-camera data sets of such length is addressed for the first time.

5.0 REFERENCES

1) A.Milan, K. chindler, and S. Roth, "Challenges of ground truth evalua-tion of multi-target tracking," in *Proc. CVPRW*, Jun. 2013, pp. 735–742.
2) H. Wu, A. C. Sankaranarayanan, and R. Chellappa, "Online empirical evaluation of tracking algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, pp. 1443–1458, Aug. 2010.
3) M. Kristan *et al.*, "The visual object tracking VOT2013 challenge results," in *Proc. ICCVW*, Dec. 2013, pp. 98–111.
4) M. Kristan *et al.*, "The visual object tracking VOT2014 challenge results," in *Proc. ECCV Workshops*, 2014, pp. 191–217.
5) J. Yuen, B. Russell, C. Liu, and A. Torralba, "LabelMe video: Building a video database with human annotations," in *Proc. ICCV*, Sep./Oct. 2009, pp. 1451–1458.

6) B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: A database and Web-based tool for image annotation," *Int. J. Comput. Vis.*, vol. 77, nos. 1–3, pp. 157–173, 2008.
7) A. Dorado, J. Calic, and E. Izquierdo, "A rule-based video annotation system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 5, pp. 622–633, May 2004.
8) M. Bertozzi, A. Broggi, R. Grisleri, A. Tibaldi, and M. D. Rose, "A tool for vision based pedestrian detection performance evaluation," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2004, pp. 784–789.
9) C. Spampinato, B. J. Boom, and J. He, "MTAP special issue on meth-ods and tools for ground truth collection in multimedia applications," *Multimedia Tools Appl.*, vol. 70, no. 1, pp. 409–412, May 2014.
10) S. Vijayanarasimhan and K. Grauman, "Active frame selection for label propagation in videos," in *Proc. ECCV*, 2012, pp. 496–509.

