A Review on Different Methods of Person Re-identification

[1]JASHER NISA A J, [2] SUMITHRA M D
[1]Mtech Student, [2]Assistant Professor
[1]Computer Science & Engineering
[1]L B S Institute of Technology for Women, Trivandrum, India

Abstract— Person re-identification is a fundamental task in intelligence video surveillance and has taken as an area of intense research in the past few years. Given an image or video of one person taken from one camera, person re-identification is the task of identifying the same person from a video or image taken from another camera. Apart from surveillance, it has applications in robotics, forensics and multimedia. This paper will summarizes several developments in the recent literature and discuss various available techniques used in person re-identification. Especially, their pros and cons are mentioned and are compared.

Index Terms— Silhouette segmentation, Structural Element (STEL), Histogram of Oriented Gradients (HoG), Color Context People Descriptor (CCPD).

I. INTRODUCTION

One of the fundamental tasks of a distributed multi-camera surveillance system is to associate people across camera views at different locations and at different time. This is known as person re- identification problem. Person re-identification finds its application in forensic and in long term multi camera tracking. For surveillance applications, an individual disappearing from one view need to be matched in one or more views at different locations at different time and have to differentiate a number of visually similar but different candidates in that view. Person re-identification consists of matching the observations of individuals across different views in a network of cameras [1]. This is a non-trivial problem, since the appearance of individual varies with scene, due to different acquisition devices and different illuminations shadows, occlusion, different pose of the same person that has searched for.

II. RE-IDENTIFICATION PIPELINE

There are many challenges faced by forensic analysis of video from multi camera CCTV networks, including data overload from different cameras, limited amount of absorption leading to important events and targets being missed. An automated re- identification system takes as input either tracks or bounding boxes containing segmented images of individual person. For automatically matching person at different locations over time captured by different camera views, re-identification process contains following steps [1].

- **a)** Extraction of imagery features such as color, texture, spatial structure or its combinations. These features are easily and reliably measurable. They also have a capable of providing inter-personal discriminations and intra-camera invariance.
- **b)** Constructing a representation or descriptor for example, a histogram of features.
- **c)** Matching specified images or bounding boxes against a gallery of persons in another camera view using some model-based matching procedure or by using a neural network.

III. METHODS EMPLOYED IN PERSON RE-IDENTIFICATION

Re-identification can be done using either appearance based or gait based methods. In appearance based method, generate invariant signature from color, texture and other appearance based properties. In gait based method, extract features from gait and motion of the person of interest [2]. Person re-identification problem will be more convenient if the region of interest (ROI) can be segmented for further processing. Segmentation will reduce the number of pixels to be processed.

Segmentation

It is one of the pre-processing step including Human detection, Background elimination and Shadow suppression. If this preprocessing is carried out, it will definitely improve the time complexity and accuracy of the result. First step after data acquisition is background elimination which is needed to detect person or region-of - interest [2].

1. BACKGROUND ELIMINATION:

In some specific applications, the background information's are not always relevant since the images are taken from different cameras with different backgrounds. But in all other applications, it is necessary to eliminate this background information, thus we can use one of these techniques such as silhouette segmentation, which are more accurate than any other methods [2].

Silhouette segmentation:

According to Mark Barnard, Matti Matilainen and Janne Heikkila [3], it is the most naive approach for background elimination. Here, classify each pixel on the silhouette outline into one of the 4 body parts, head, body, arm and legs. In order to cope with uncertainty in the recognition process we use a ratio of the likelihood between the most likely class and the second most likely class for each pixel on the silhouette outline. Any pixel where the likelihood falls below a certain threshold is discarded as the uncertainty on the pixel is considered too high. A significant problem in dealing with silhouette data is that it can be very noisy. This noise is in the form of shadows, changes in lighting conditions or the background matching the foreground we wish to segment. This noise can cause dramatic changes in the shape of the silhouette and so seriously degrade the performance of any shape recognition method. Another source of error in silhouette segmentation is occlusion. This is particularly true in the case of surveillance and monitoring.

2. HUMAN DETECTION:

There are several approaches like HOG, pixel wise method to detect human from the image and extract features from the bounding boxes that surround the human body.

The Histogram of Oriented Gradients (HOG):

One of the most beneficial techniques for human detection, proposed by Dalal and Triggs [4] describes the distribution of intensity gradients or edge directions of a local object and shape within an image [8, 9]. It is useful in the case in which of video frames are not available, but needs training samples like structure element (STEL). HOG is a dense feature extraction method for images. Dense means that it extracts features for all locations in the image (or a region of interest in the image). The fundamental basic behind HOG is that we can characterize the objects appearance and shape by having either distribution of local intensity gradients or edge positions. The distribution of local intensity gradients can obtain by dividing the image into small cells. Each cell contains a local 1-D histogram of gradient directions or edge orientations over the pixels of the cell. The combined histogram entries form a representation. Before using them, it is also useful to contrast-normalize the local responses for better invariance to illumination, shadowing etc. This can be done by accumulating a measure of local histogram over larger spatial regions called blocks and using the results to normalize all of the cells in the block. This normalized descriptor blocks can be referred to as Histogram of Oriented Gradient (HOG) descriptors. Combing the detection window with a dense grid of HOG descriptors and using the combined feature vector in a conventional SVM based window classifier gives human detection chain.

3. SHADOW ELIMINATION:

The shadow which is present after the background subtraction can be removed using chromaticity method. Priya Garg and Kirtika Goyal proposed a method [5] in which Image used for shadow removal has been loaded as RGB image and three components of RGB are put in three columns of array. Red and blue intensities of all pixels are divided by geometric mean of Red, green and blue pixels and are stored in two separate arrays. Then natural log has been taken for both arrays in order to plot the one dimensional arrays obtained above in two dimensional orthogonal spaces. Let two arrays obtained at this step are array3 and array4. array3 = logarray1, array 4 = logarray2. A direction perpendicular to these two arrays has been obtained in order to remove shading or shadow effect. The direction has been found by equation,

$$\cos(angle)*array3+\sin(angle)*array4 \qquad \ldots\ldots (1)$$

Then, vary the angle and find the direction where there is maximum removal of shadow in the image and change the pixel value of original image, after changing the pixel value of original image apply some morphological operations and thresholding techniques to detect and remove the shadow.

Spatial features in re-identification

The holistic description or a part-based or region-based description of an individual can be utilized for modeling the appearance of the individual. In both appearance and motion-based approaches**,** the spatial information extraction helps to extract more robust features and finally obtain a better re-identification rate. Using the fixed proportions of the bounding box surrounding the person of interest, partitioning can be done but this cannot properly separate the regions and portions [3].

Symmetrical and asymmetrical axes:

The algorithm proposed by Farenzena et al [6] is an effective portioning algorithm and uses many approaches [11, 12, and 13]. In this algorithm three main body regions head, torso and leg are divided by two horizontal asymmetry axes. To divide head and torso uses the maximum difference between the number of pixels in two moving rectangles, which sweeps the image. The maximum color difference in these bounding boxes is estimated to divide the torso and legs. To differentiate two leg regions, a vertical axis of the appearance symmetry is estimated. To make this method pose independent, symmetrical axes are used.

Appearance-based models

Appearance-based methods are more suitable in re-identification because the entire process of re-identification can be completed within a short time period. The varying pose and illumination issues have a direct effect on the appearance features of different images, the availability and discrimination of appearance-based features are the main reason for using them in most of the works on re-identification [3].

Clor context people descriptor:

Belongie et al [7] proposed a shape context structure [11], which is used to implement the color context people descriptor (CCPD). In CCPD, on the center of the segmented object (which is torso or legs), the shape context structure is placed. Then, a color histogram is generated, based on the pixels radial and angular bins. The leg portion will provide more or less color information to the histogram based on its pose. Therefore the bottom histogram for the same person will be different from pose to pose. So, ignore the background pixels and only consider the legs pixels to make the descriptor more discriminative. A back projection algorithm is used to identify the pixels that represent the legs.

Depending on the pose, the legs may contribute more or less color information to the histogram. Therefore the bottom histogram for the same person will be different from pose to pose. Thus, it is important to ignore the background pixels and only consider the legs pixels to make the descriptor more accurate. To identify the pixels that represent the legs, we can use a back projection algorithm. To improve the problem of entering some unwanted pixels from the background in histogram computation when applying CCPD on the entire torso region, apply a background/foreground segmentation algorithm on detected person bounding box and then apply CCPD.

IV. CONCLUSION

Person re-identification (re-id) problem has many applications such as long-term multi- camera tracking and forensic search. In this paper, we provide a review of the existing research on re-identification which contains both the appearance and gait/motion descriptors and point out the limitations and advantages of the various available methods for re-identification. To provide an insight for the future research direction, we highlight the methods that are capable of being useful in the forthcoming research.

ACKNOWLEDGMENT

We are greatly indebted to our principal Dr. K. C. RAVEENDRANATHAN, Dr. SHREELEKSHMI R, Professor, Head of the Department of Computer Science and Engineering, Mr. MANOJ KUMAR G., Associate Professor, Department of Computer Science and Engineering, LBS Institute of Technology for Women who has been instrumental in keeping my confidence level high and for being supportive in the successful completion of this paper. We would also extend our gratefulness to all the staff members in the Department;also thank all my friends and well-wishers who greatly helped me in my endeavor. Above all,we thank the Almighty God for the support, guidance and blessings bestowed on us, which made it a success.

REFERENCES

[1] Shaogang Gong, Cristani,shuicheng Yan, Chen Change Loy Editors, “Person Re-Identification”.

[2] Mohammad Ali Saghafi1, Aini Hussain1, Halimah Badioze Zaman2, Mohamad Hanif Md. Saad, “Review of person re-identification techniques”, IET Computer Vision, IET Computer Vision, February 2013.

[3] Mark Barnard, Matti Matilainen and Janne Heikkil , “A body part segmentation of noisy human silhouette images”, , Proc. Second Joint IEEE Int. Workshop on VS-PETS, Beijing, China, October 2005, pp. 105 to 112.

[4] Dalal N., Triggs B., “Histograms of oriented gradients for human detec tion”,Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CPVR), CA., USA, June 2005, pp. 886 to 893.

[5] Priya Garg, Kirtika Goyal, “Detection and Removal of Shadow using Chromaticity”,Pattern Recognit. Lett., Elsevier, 2012, 33, (14), pp. 1908 to 1915.

[6] Farenzena, M., Bazzani, L., Perina, A., Murino, V., Cristani, M, “Person re- identification by symmetry-driven accumulation of local features”,Proc. Int. Conf. Computer Vision and Pattern Recognition (CVPR), San Francisco, USA, June 2010, pp. 2360 to 2367.

[7] Belongie, S., Malik, J., Puzicha, J, “Shape matching and object recognition using shape contexts”,IEEE Trans. Pattern Anal. Mach. Intell., 2002, 4, (4), pp. 509 to 522.

[8] Xiang, J.: ‘Active learning for person re-identification’. Proc. Int. Conf. Machine Learning and Cybernetics (ICMLC), Xian, China, July 2012, pp. 336–340.

[9] Bak, S., Corvee, E., Brémond, F., Thonnat, M.: ‘Person re- identification using haar-based and DCD-based signature’. Proc.Seventh IEEE Int. Conf. Advanced Video and Signal Based Surveillance, Boston, USA, September 2010, pp. 1–8.

[10] Belongie, S., Malik, J., Puzicha, J.: ‘Shape matching and object recognition using shape contexts’, IEEE Trans. Pattern Anal.Mach. Intell., 2002, 4, (4), pp. 509–522

[11] Satta, R., Fumera, G., Roli, F., Cristani, M., Murino, V.: ‘A multiple component matching framework for person re-identification’. Proc. 16th Int. Conf. Image Analysis and Processing (ICIAP), Ravenna, Italy, September 2011, pp. 140–149

[12] Eisenbach, M., Kolarow, A.: ‘View invariant appearance-based person re-identification using fast online feature selection and score level fusion’. Proc. IEEE Ninth Int. Conf. Advanced Video and Signal-based Surveillance (AVSS), Beijing, China, September 2012, pp. 184–190

[13] Martinel, N., Micheloni, C.: ‘Re-identify people in wide area camera network’. Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops, RI, USA, June 2012, pp. 31–36.