A REVIEW ON PERSON RE-IDENTIFICATION TECHNIQUES

[1]RESHMA O S, [2]RESHMA SHEIK
[1]Mtech Student, [2]Assistant Professor
[1]Computer Science and Engineering,
[1]T K M College of Engineering, Kollam, India

Abstract: ***In investigation, person re-identification is also a tough task of matching persons determined from utterly completely different camera views. It is necessary applications in AI, threat detection, human trailing and activity analysis. Person re-identification is also a tough analysis topic as a result of partial occlusions, low resolution images and massive illumination changes. Also, person determined from utterly completely different camera views has very important variations on poses and viewpoints. This paper summarizes the challenges related to the person re-identification jointly discuss varied techniques utilized person re-identification.***

Index Terms - Gaussian mixture model (GMM), Video-surveillance, Histogram of Oriented Gradients (HOG), Bidirectional ranking.

I. INTRODUCTION

In recent years, closed-circuit television plays a vital role publically, and has become a lot of necessary in crime investigation. During this state of affairs, investigators wish to find and track the person in vicinity lined by multiple cameras. Here, manual browsing is time overwhelming and is not economical for the crime investigation. To unravel this downside, person re-identification that is matching persons across totally different camera views, attracts a lot of and a lot of analysis interests.

In video surveillance system, a sequence of video frames is obtained from their source mainly from closed circuit television (CCTV) and processing of these frames help to extract relevant information. Surveillance in public places is mainly for monitoring various locations and people in that locations and observing their behaviors. Nowadays, events like terrorist attacks have occurred more frequently in different public places. So, there is a growing need for video network systems to guarantee the safety of people in those areas.

In addition, intelligent surveillance has proven to be a useful tool for detecting and preventing potentially violent situations in public transport such as airports, train stations or even inside trains and air planes. The growth of computational capabilities in intelligent systems provided more opportunities in video surveillance system. This includes segmentation, object detection, tracking, classification etc. Person re-identification is one of the most important topic in this area.

In 1961, Alvin Plantinga [1] provided one of the first definitions of re -identification: "To re-identify a particular, then, is to identify it as (numerically) the same particular as one encountered on a previous occasion". According to Paul McFedries [2], re-identification is the process of matching unknown huge amount of data with the individuals who provided the data. In generally, person re-identification can be defined as the process of recognizing individuals over different camera views in various locations under the condition of large illumination variations. In other words, it is the process of finding a person of interest from large number of images and video frames that track the persons in a network of cameras. The important applications of person re-identification are in video surveillance such as human tracking, human retrieval, and activity analysis. Searching a person from a large number of images and videos are time consuming. In this situations person re-identification technique saves a lot of human efforts. At the same time, it is a challenging research topic in computer vision due to large illumination variations, low resolution images, pose variation, background noises and occlusions.

When a person stays within a single camera view, the system has knowledge about his location, position, background and lighting condition. When the person moves from one camera view to another camera view the important question is, how does the system know that person observed from the camera was same as the person observed in another camera. This issue is called re-identification problem. It is the technique of recognizing persons separated in time and location. Person re-identification is a complex problem due to the lack of spatial continuity for the information received from different camera observations.

Traditionally, person re-identification can be considered as a matching problem. There is a gallery set and probe set. Gallery set contains huge amount of candidate person images and probe set contains query person images. For each test image or group of test images of an unknown person, the goal of person re-identification is to return a ranked list of individuals from the gallery set as shown in figure 1. It is a key point of many applications, at the same time a most challenging research topic in Pattern Recognition.

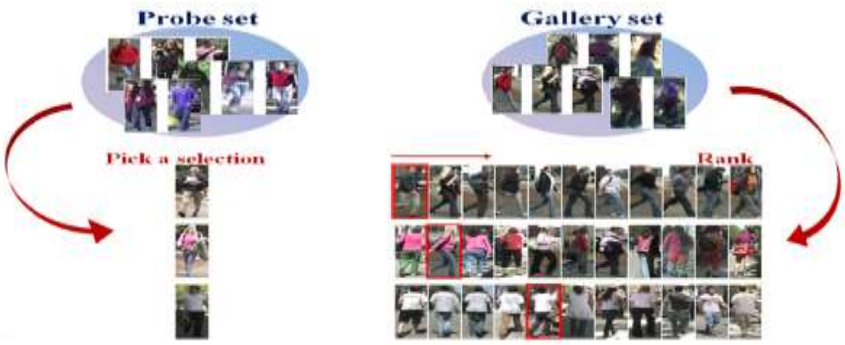

Figure1.Person Re-identification

II. CHALLENGES IN PERSON RE-IDENTIFICATION

The issues in person re-identification can be classified into two categories i.e. Inter camera and Intra camera issues. Inter-camera issues can cause problems in identifying people in a network of cameras. The different illumination conditions that exist at different camera views should be consider here. The cameras have differences in illumination, because they have different features even if the cameras are from the same manufacturer. Another important issue is the different poses of persons in different camera views. These problems decrease the detection rate (Figure 2).

Figure 2. Differences in poses and lighting conditions in four different cameras

Some of the problems in person re-identification are due to varying light condition at different times of the day (Figure 3). In addition, low resolution cameras reduce the quality of frames. So, the methods depend on quality of frames decrease the detection rate. Another important issue is occlusions in camera frames create difficulty in image segmentation (Figure 4). This is called Intra-camera issues.

Figure 3. Lighting conditions at different times of day

Figure 4. Sample of occlusion in scene

The system must be powerful against illumination variations, background noises and occlusions in order to identify the same person in different camera views.

III. METHODS USED FOR PERSON RE-IDENTIFICATION

Many methods are based on gait or shape of the moving person, different camera views can also cause problems in these methods. Occlusion in public places is another issue that should be addressed. Some methods are based on appearance of people. In that situations clothing of the object shown from one camera to another should not be changed; otherwise the system would fail. Persons observed from different camera views has significant variations on their poses, these changing poses will create difficulties for approaches that try to extract a model based on the movement of the person. Hence, more flexible methods are needed for preventing the failure. Generally, person re-identification can be done using two methods either appearance based or gait (motion) based. The appearance based methods try to extract signature from colour, texture and other appearance properties. In other ways, gait based methods try to extract features from the motion of persons. This gait based methods are restricted due to view point variations in which person's motion is not clearly shown.

Whether the using method is appearance based or gait based, re-identification consist of three main stages. The first stage is to extract imagery features. The second stage is to constructing a representation. Final stage is to compare the features between query and the gallery set to calculate the similarities.

Before going through feature extraction, some of the pre-processing techniques are performed for improving person re-identification. These pre-processing stages include background elimination and human detection. Manual silhouette segmentation [3] used for background elimination. This method suffers from noise and occlusion. Another method for background removal is using Gaussian mixture models (GMM) [4]. However, GMM is sensitive to illumination changes. Histogram of oriented gradients (HOG) [5] is one of the methods used for human detection. HOG is a dense feature extraction method, in which extracts features for all locations in the images. However, varying poses and viewpoints reduce the detection rate.

There are various works related to person re-identification. These works can be classified into three categories: Supervised methods, unsupervised methods, and other methods. Supervised methods focuses on learning feature representation, and unsupervised methods focuses on feature extraction.

3.1 Unsupervised Methods

Unsupervised methods focus on feature design of given images. Farenzena et al. [6] used the STEL model to separate foreground and background. Then, head, torso and legs are separated by an asymmetry human partition. A symmetry based partition was proposed to divide body and legs into right and left parts. Finally, weighted HSV histograms and high structured patches are used to extract features. The combination of the distance computed on each features used as the result matching distance.

Chang et al. [7] introduced an improved random walk algorithm into person re-identification. In which foreground and background separation are achieved by combining the shape information and color seed into the Random Walk formulation. Finally, for feature matching 1-D RGB signal, HSV color histograms, local binary pattern (LBP) and scale invariant local ternary pattern (SILTP) were used.

Zhao et al. [8] have proposed unsupervised salience learning for person re-identification. In which a combination of color histograms and SIFT features are used as the feature. Here LAB color histogram used as the color feature. For obtaining color feature, the image is divided into several blocks, and from each block a LAB color histogram is extracted. SIFT descriptor is useful for handling illumination variations and pose variations. So, SIFT descriptor can be used as the complementary feature to color histograms. Finally, for selecting possible candidates they have used adjacency search and k-nearest neighbour algorithm.

Ma et al. [9] introduced a BiCov descriptor for handling illumination changes and background variations. Actually, BiCov descriptor trusted on the combination of Biologically Inspired Features (BIF) and covariance descriptor for computing the similarity of the BIF features at neighboring scales.

Wu et al. [10] build a scalable face image retrieval system. In which, both global and local features are used for representing face. They have utilized special properties of faces for designing new component based local features in indexing stage. Then a novel identity based quantization schema is used for quantizing these local features into visual words. To refine candidate person images, they have constructed a reference images set. However, a new feature extraction process is needed for re-rank the initial ranking list.

Pedronette et al. [11] have proposed new re-ranking algorithm that exploit contextual information. In the content based image retrieval there is a query image, the approach aims to retrieve more similar images from a collection. Here, contextual information improving the efficiency of CBIR systems for generating ranking list. Context images are created by using the ranked list and distance measures, which brought computation complexity.

Huang et al. [12] have studied about two mechanisms, visual consistency and visual latency. In current web image search engines, search query and images that are closely related to that search query are visually similar. Also, user's eyes can capture salient images than ones in low level vision. Here, results from the search engines are re-ranked based on the visual saliency and consistency. So, re-ranked images consider both vision and content. However, the characteristic of saliency is not suitable for person re-identification.

Zhu et al. [13] have proposed a clustering algorithm in re-ranking. Rank-order distance, the core of the algorithm, uses neighbouring information in the dataset for measuring the dissimilarity between pair of faces. Here, faces of the same person will have similar neighbours. First a ranking order list is created for every faces by sorting all the faces in the dataset. Then ranking order is used for calculating the rank order distance. Finally, a rank-order distance based clustering algorithm is designed based on the new distance, which iteratively groups all faces into small number of groups. Here, only context distance is used for re-ranking which ignores content distance. But content distance is essential in some situation, so ignoring content distance reduces detection rate.

One disadvantage of unsupervised methods is that, information represented by their feature designs is not fully utilized. Also, it takes too much time for feature extraction and matching, which reduces its application in real time systems.

3.2 Supervised Methods

Generally, supervised methods require training samples with identity labels. These methods have lower flexibility for large gallery objects. Du et al. [14] have proposed a new approach called, random ensemble of color features (RECF). In which color features are obtained from six kinds of popular color spaces such as RGB, normalized RGB, HSV, YCbCr, CIE XYZ and CIE Lab. This information is used by a random forest to learn the similarity function between pair of person images.

Gray and Tao [15] have used ensemble of localized features (ELF), an efficiently designed object representation for viewpoint invariant pedestrian recognition. In computer vision, recognition of viewpoint invariant pedestrian is one of the most important problems. If there are two objects with unknown viewpoint and pose, it is very difficult to matching them. Here, designing a feature space instead of a specific feature to solve the problem and a machine learning algorithm is used for getting the best representation. Also, this approach generates a single similarity function by combing different kinds of single features.

Bak et al. [16] have used haar-like features and dominant color descriptor (DCD) for person re-identification problem. Generally, persons observed from one camera view is different from those observed in another camera. Human signatures are used for person re-identification in order to handle difference in illumination, pose and camera parameters. For obtaining most discriminative haar-like feature set for each individual, the AdaBoost schema is used in haar-based approach. Finally, human signature is obtained by combining the set of haar-like features through a cascade. The dominant colors of upper and lower parts are extracted for obtaining DCD signature. Haar-like features are suitable for handling view point and pose changes, however foreground extraction seems to be a difficult task in person re-identification.

Prosser et al. [17] have considered person re-identification as a relative ranking problem. Previous approaches tried to extract discriminative features then perform a matching based on distance measures. In this approach, by focusing on a highest ranked person, team find out most similar persons for a given query image. So, person re-identification problem is considered as a relative ranking problem rather than an absolute scoring problem. Computational scalability limitations of existing RankSVM models are resolved by using an Ensemble RankSVM.

Schwartz et al. [18] have proposed a feature selection approach, called partial least square (PLS) for learning discriminative appearance models. The applications like tracking and people recognition uses appearance information. In the case of appearance based models when the number of persons being considered increases, there is a chance for ambiguities among classes. Also for each appearance limited numbers of training samples are available. Here, a rich set of feature descriptors based on color, textures and edges are used for reducing the ambiguity. Partial least square (PLS) is a powerful statistical tool used for weighting the features based on their discriminative power for different appearance. This approach can reduce the background influence, but in the case of variable number of persons it is not flexible.

3.2 Other Methods

This category includes those methods other than supervised and unsupervised for handling pose variations, light conditions and occlusions. Wang et al. [19] have computed the similarity between image regions by using appearance models. An appearance model can be constructed by computing local descriptors in an image, and then different strategies are used for aggregating them. In this work multi-layer appearance modeling framework was introduced for extracting discriminative features for computing the similarity. The local description of the image is computed by the first layer, and second layer computing spatial relations between appearance labels. This work is mainly focus on computational complexity.

Bak et al. [20] introduced a spatial covariance descriptor for human signature computation. Human signatures should handle the variations in illumination, pose and camera parameters in order to re-identify persons. First Histogram Oriented Gradients (HOG) is used for detecting human body parts, and then spatial covariance regions are extracted from these body parts. Then a covariance descriptor is used for computing the similarity between corresponding body parts. Finally, spatial pyramid matching is used for computing a new dissimilarity measure between human signatures.

IV. Conclusion

In this paper, we briefly discuss the application and challenges of person re-identification. In addition, we also discuss the general methods used for person re-identification. Existing methods are mainly based on single directional matching. Single directional matching is not efficient in the case of variation of illumination and poses. Based on the drawbacks introduce a bi-directional re-ranking method for getting better results. To provide associate insight for the long run analysis direction, we tend to highlight the strategies that are capable of being helpful within the forthcoming analysis.

V. Acknowledgment

We would like to express our sincere gratitude to Dr.S.Ayoob, Principal of T.K.M College of Engineering, Dr.Ansamma John, Associate Professor and Head of the Department, CSE, TKMCE, for constant support and encouragement throughout the successful completion of this paper. I also express my thanks to all my friends and well-wishers, who greatly helped me in my endeavour.

References

[1] A. Plantinga. 1961. Things and persons, The Review of Metaphysics, pp.493–519.
[2] wordspy.Availableathttp://www.wordspy.com/words/reidenti_cation.asp.
[3] Annesley, J., Orwell, J., Renno. 2005. Evaluation of MPEG7 color descriptors for visual surveillance retrieval, Proc.Second Joint IEEE Int. Workshop on VS-PETS, Beijing, China, pp. 105-112.
[4] Stauffer, C., Grimson, W.E.L. 1999. Adaptive background mixture models for real-time tracking, Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, vol. 2, no. C, pp. 246–252.
[5] Dalal N., Triggs B.2005.Histograms of oriented gradients for human detection, Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CPVR), CA., USA, pp. 886 to 893.
[6] Farenzena, M., Bazzani, L., Perina, A., Murino, V., Cristani, M. 2010, Person re-identification by symmetry-driven accumulation of local features, In computer vision and pattern recognition(CVPR), pp. 2360-2367.
[7] Chang, Y.-C., Chiang, C.-K., Lai, S.-H.2012, Single-shot person re-identification based on improved random-walk pedestrian segmentation, In intelligent signal processing and communication systems (ISPACS).
[8] Zhao, R., Ouyang, W., Wang, X. 2013, Unsupervised salience learning for person re-identification, In Computer Vision and Pattern Recognition (CVPR).
[9] Ma, B., Su, Y., Jurie, F. 2012, Bicov: A novel image representation for person re-identification and face verification, British Machive Vision Conference, Guildford, United Kingdom, p. 11.
[10] Wu Z, Ke Q, Sun J, Shum HY. 2011, Scalable face image retrieval with identity-based quantization and multireference reranking, IEEE Trans Pattern Anal Machine Intell 33(10):1991–2001.
[11] Pedronette D, Guimares C, da S Torres R. 2011, Exploiting contextual spaces for image re-ranking and rank aggregation, In: Proceedings of the 1st ACM international conference on multimedia retrieval.
[12] Huang J, Yang X, Fang X, Lin W, Zhang R. 2011, Integrating visual saliency and consistency for re-ranking image search results, IEEE Trans Multimedia, 13(4):653–66.
[13] Zhu C, Wen F, Sun J. 2011, A rank-order distance based clustering algorithm for face tagging, In: IEEE conference on computer vision and pattern recognition (CVPR), pp 481–488.
[14] Du, Y., Ai, H., Lao, S. 2012, Evaluation of color spaces for person re-identification, In International Conference on pattern recognition (ICPR), pp. 1371–1374.
[15] Gray, D., & Tao, H. 2008, Viewpoint invariant pedestrian recognition with an ensamble of localized features, In European Conference on computer vision (ECCV), pp. 262-275.
[16] Bak, S., Corvee, E., Bremond, F., Thonnat, M. 2010, Person re- identification using Haar-based and DCD-based signature, In Workshop on Activity Monitoring by Multi-Camera Surveillance Systems.

[17] Prosser, B., Zheng, W.-S., Gong, S., Xiang, T.2010,Person re-identification by support vector rankin, In British Machine Vision Conference (BMVC).

[18] Schwartz, W., & Davis, L. 2009, Learning discriminative appearance-based models using partial least squares, In XXII Brazilian symposium on computer graphics and image processing (SIBGRAPI).

[19] Wang, X., Doretto, G., Sebastian, T., Rittscher, J., Tu, P.2007, Shape and appearance context modeling, In International Conference on computer vision (ICCV).

[20] Bak, S., Corvee, E., Bremond, F., Thonnat, M. 2010, Perso re-identification using spatial covariance regions of human body parts, In Advanced Video and Signal-Based Surveillance.