RESEARCH ARTICLE OPEN ACCESS

Human Re-Identification In Surveillance Network

Shubham Lavania
Department of Electronics and Communication Enginnering VIT University Vellore, India

Abstract
This paper presents an appearance-based model to address the human re-identification problem. Human re-identification is an important and still unsolved task in computer vision. Re-identification refers to the problem of establishing correspondence among various observations of the same subject viewed at different time instances in different camera positions. It can also be defined as the process to match persons observed in non overlapping camera views with visual features for inter-camera tracking. The ambiguity increases with the number of candidates to be distinguished. Simple temporal reasoning can simplify the problem by pruning the candidate set to be matched. In many systems there is a requirement to identify individuals or determine whether a given individual has already appeared over a network of cameras. The human appearance obtained in one camera is usually different from the ones obtained in another camera. In order to re-identify people a human signature should handle difference in illumination pose and camera parameters. The paper focuses on a new appearance model based on Mean Riemannian Covariance (MRC) patches extracted from tracks of a particular individual to capturing from a distance even at low resolution without active co-operation of subjects, has motivated us to use it for re-identification in the computer surveillance network.
Keywords: Re-identification, Riemannian manifold, Human detection, Gait, Phase of motion

I. INTRODUCTION

Recently, cameras spread out across various domains that range from personal computers, video games, home surveillance applications,to large camera networks. These camera networks facility access to sports venue, monitored environments, such as airports, metro stations or car parks. A natural consequence of such situation is a need for an automated extraction of high-level semantic information from extremely large volumes of recorded video data. In many surveillance systems, the most desirable information is knowledge about an identity of a detected object. An appearance-based approach relies on modeling a human signature using tracking and detection results. Re-identification techniques, which exploit appearance information using only one image are referred to as single-shot approaches and until now they were the most popular methods. Currently, identification accuracy is improved by integrating information over many images. The methods, which employ multiple images of the same person as a training data, are called multiple-shot approaches. As re-identification concerns a large sets of individuals acquired from different cameras, we need to provide a distinctive and invariant signature. This signature has to be based on discriminative features to allow browsing the most similar signatures over a network of cameras. It can be achieved by signature matching, which has to handle differences in illumination, pose and camera parameters. Note that, intercamera variations in lighting conditions, differences in illumination, different camera parameters, changes in object orientation and object pose make this task extremely difficult. Besides, occlusions (caused by other people or objects of the scene) and self-occlusions (caused by body parts) make the re-identification problem one of the hardest tasks in the video surveillance domain. For high level analysis of surveillance video, the first step is to track the moving subjects in order to get their spatiotemporal information. Tracking subjects in a video captured from a single camera is usually done by the data association technique .However, effective video surveillance over a large area requires the use of multiple cameras. In case of overlapping cameras, joint camera calibration and spatiotemporal information help to establish the correspondence between subjects (Khan and Shah, 2006). But if there is no overlap between cameras, the problem of identity management of subjects viewed in different cameras at different time instants becomes challenging.

Identification is the process of recognizing a subject who has left one camera and then reappeared in another camera. The uncertain behavior of subjects across the 'blind gaps' makes the task harder.Re-identification techniques help to reassign the same identity to a subject entering a camera that was previously attached to him when exiting a previous camera. Thus, re-identification extends tracking beyond blind gaps. Spatiotemporal methods simplify the re-identification procedure. But, they fail to resolve identity when multiple persons, walking at normal velocity, exit and enter two consecutive cameras at the same time. To handle this situation phase of motion can be used. Phase of motion is defined as the physical shape of a person at a particular time instant while walking. Multiple persons exiting a camera at the same time are unlikely to have the same phase of motion. If the exit phase of motion and the transition time are known, the probable entry phase of motion of that person can be obtained using a phase of motion estimation model. Thus, phase of motion can resolve the ambiguity arising due to similar entry/exit time by multiplepersons.

The major drawback of spatiotemporal and phase of motion based models is in the assumption on the behavior of subjects in the blind gap. The subjects are assumed to walk at normal velocity without any hindrance in the blind gap which may not always be valid in reality. In order to handle the above-mentioned situations, biometric based features such as gait can be used for re-identification. Gait of a person is a periodic activity where each gait cycle covers two strides: the right foot forward and the left foot forward. Gait consists of both shape and dynamics where shape refers to physical build of a person seen in different gait phases and dynamics refers to motion dynamics of the person in a gait cycle.

II. USE OF TECHNIQUES

The most common way of extracting appearance models in automatic surveillance systems is by rst detection of objects in an image, and then by tracking them using different methods and strategies. We can detect and track humans using HOG-based detector and we handle color dissimilarities caused by camera illumination differences by applying histogram equalization.

The Histogram of Oriented Gradient (HOG) based technique to automatically detect humans in different scenes before their visual signatures are extracted for re-identication purposes and it also uses using a Sobel convolution kernel, in a multi-resolution framework.

This technique was originally designed to detect faces using the assumption that facial features remain approximatively at the same location. However, locations of human silhouette features do not remain constant in templates with varying poses (e.g. knees are constantly changing position when walking).Hence the algorithm modification was proposed in the paper[3].According to it at specific locations around the human silhouette as shown in figure 2. The most dominant cells used to characterize human shapes are the 15 most dominant cells selected among 252 cells covering the human sample area as discussed in paper[8].These most dominant cells are the cells having the closest HOG vector to the mean HOG vector calculated over the vectors (of the corresponding cell) from a human database. The system was trained using 10; 000 positive and 20; 000 negative image samples from the NICTA database as mentioned in paper[4].

The next step basically involves the Color normalization One of the most challenging problem using the color as a feature is that images of the same object acquired under different cameras show color dissimilarities.Histogram equalization stretches a range of histogram to be as close as possible to an uniform histogram.The approach is based on the idea that among all possible histograms,an uniformly distributed histogram has maximum entropy. Maximizing the entropy of a distribution, we maximize its information and thus histogram equalization maximizes the information content of the output image. We apply the histogram equalization to each of the color channels (RGB) to maximize the entropy in each of those channels and obtain the invariant image

The other technique includes the interfacing of human reidentification in video by integrating the face profile and gait the technical approach for the same is as follows like the gait recognition is an emerging biometric technology which involves people being identified purely through the analysis of the way they walk. While research is still underway, it has attracted interest as a method of identification because it is non-invasive and does not require the subject’s cooperation. Gait recognition could also be used from a distance, making it well-suited to identifying perpetrators at a crime scene. But gait recognition technology is not limited to security applications – researchers also envision medical applications for the technology. For example, recognizing changes in walking patterns early on can help to identify conditions such as Parkinson’s disease and multiple sclerosis in their earliest stages.

There are two main types of gait recognition techniques currently in development. The first, gait recognition based on the automatic analysis of video imagery, is the more widely studied and attempted of the two. Video samples of the subject’s walk are taken and the trajectories of the joints and angles over time are analyzed. A mathematical model of the motion is created, and is subsequently compared against any other samples in order to determine their identity.

The second method uses a radar system much like that used by police officers to identify speeding cars. The radar records the gait cycle that the various body parts of the subject create as he or she walks. This data is then compared to other samples to identify them.

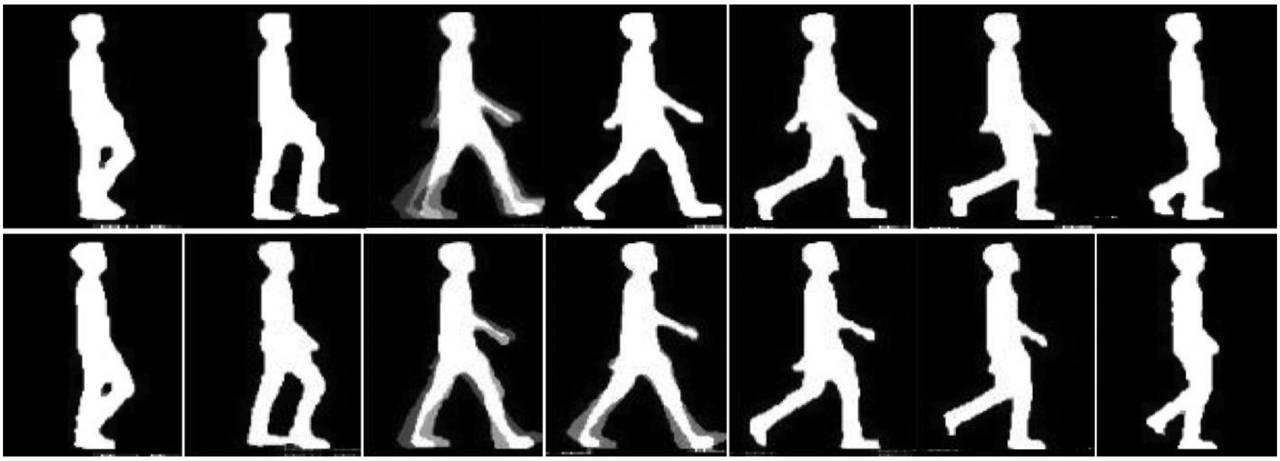

Fig. 1. PEI images of a subject having 14 key poses. This figure basically lead the emphasis on the gait motion tracking of the person while walking in different poses and its analysis with the help phase of motion techniques.

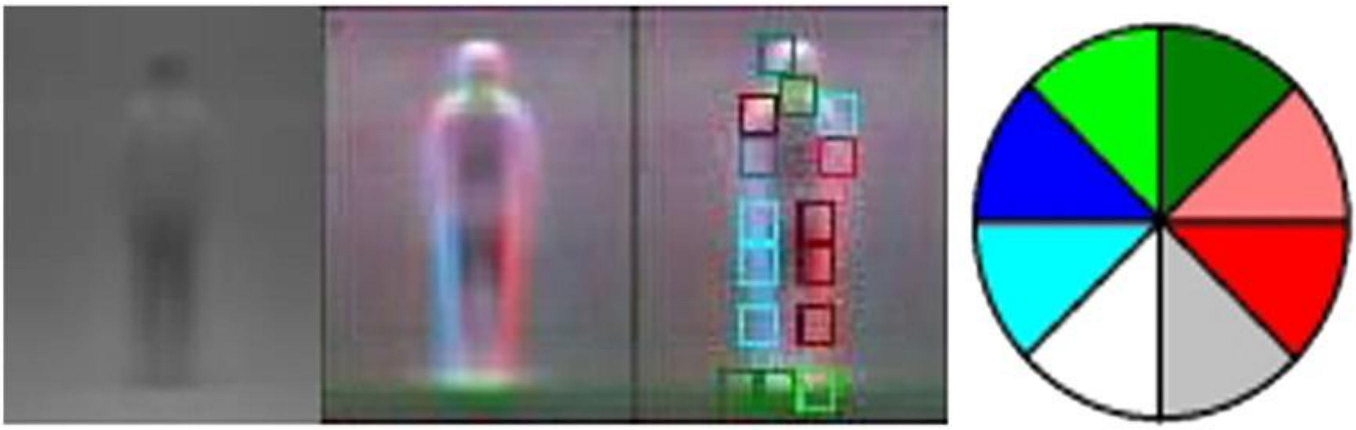

Fig. 2. Mean human image with corresponding edge magnitudes and the 15 most dominant cells. From the left: the first image shows the mean human image calculated over all positive samples in the database; the second image shows the corresponding mean edge magnitude response; the third image shows this later image superposed with the 15 most dominant cells of size 8×8 pixels. The cell bounding boxes are drawn with a color set by their most dominant edge orientation with the scheme defined in the last image.

III. CONCLUSION

In this paper, we have presented a re-identification method thatis able to track moving subjects in different locations while observingthem through multiple non-overlapping cameras. The proposed approach is based on a combination of gait, phase of motion, and the MRC model. The spatiotemporal model islearnt in the training phase based on parameters like entry/exitlocations, velocities, and transition time. Gait based feature helps insuch scenarios to re-identify him. The proposed method combining these three features is able to track subjects even if they change Speed, or stop for some time in the blind gap. The response time of the combined approach is also quite promising.. Another point of consideration is occlusion, which affects the overall re-identification accuracy. Previously it has been shown that the proposed MRC patchescomputed using a Riemannian manifold theory can extract essential informationon an appearance of a human and its variability. The experiments show that a joint combination of these distinctive patchesconstructs a robust invariant human signature, which can handle differences in camera parameters in paper[11]. Therefore the three techniques are combined together for the re-identification of the human being in the surveillance network and the future work will include to implement the same techniques in the real world applications like terrorist and mob identification and other security hazards which may occur in the public places like airports, stations and malls.

REFERENCES

[1] Rabiner, L.R., 1989. A tutorial on hidden markov models and selected applications inspeech recognition. Proc. IEEE 77 (2), 257–286.

[2] Rahimi, A., Darrell, T., 2004. Simultaneous calibration and tracking with a network of non-overlapping sensors. Proc. CVPR 1, 187–194.

[3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, M. Cristani, Person reidentification by symmetry-driven accumulation of local features, in: CVPR, 2010.

[4] A. Yilmaz, O. Javed, M. Shah, Object tracking: A survey, ACM Computer Survey 38. doi: http://doi.acm.org/10.1145/ 1177 352.1177355

[5] Veeraraghavan, A., Srivastava, A., Roy-Chowdhury, A.K., Chellappa, R., 2009. Rate- invariant recognition of humans and their activities.IEEE Trans. Image Process. 18 (6), 1326–1339.

[6] Wechsler, H., 2010. Intelligent biometric information management. Intell.Info. Manage. 2, 499–511.

[7] Huang, X., Boulgouris, N.V., 2009. Model-based human gait recognition using fusion of features. In: Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing. Pp.1469–1472.

[8] Slawomirbak,Etineecorvee,Francois Bremon Boosted human reidentification techniques using Riemannian manifolds

[9] Javed, O., Shafique, K., Rasheed, Z., Shah, M., 2008. Modeling inter-camera spacetimeand appearance relationships for tracking across non-overlapping views. Comput.Vision and Image Understanding 109, 146–162.

[10] Kale, A., Sundaresan, A., Rajagopalan, A.N., Cuntoor, N.P., Roy-Chowdhury, A.K., Kruger, V., Chellappa, R., 2004.Identification of humans using gait. IEEE Trans. Image Proc. 13, 1163–1173.

[11] Kettnaker, V., Zabih, R., 1999. Bayesian multi-camera surveillance. Proc. CVPR, 117–123.

[12] Aditi roy, Shamik sural,Jayanta Mukherjee A hierarchical approach combing gait and phase of motion with spatiotemporal model with person re-identification

[13] Khan, S.M., Shah, M., 2006. A multiview approach to tracking people in crowded scenes using a planar homography constraint. Proc. ECCV, 133–146.