Person Re-Identification by Learning Invariant Color Features

Ms. Habeeba Rizwana
PG Scholar
Electronics and Communication Engineering
MET Engineering College

Abstract:

Person re-identification is a challenging problem due to the change in visual appearance caused by varying lighting conditions. Matching people across multiple camera views known as person re-identification. The perceived color of the subject appears to be different with respect to illumination. In this paper, a data driven approach is proposed for learning color patterns from pixels sampled from images across two camera views. It is perceived that, even though pixel values of same color would be different across views, they should be encoded with the same values. This framework is two fold, learning a linear transformation using color features and the other one is a dictionary to encode pixel values. The performance of the proposed work is superior when compared to the existing framework that uses color constancy model.

Keywords—re-identification; learning color patterns; linear transformation; color constancy model

I. INTRODUCTION

A. *Need for Person Reidentification*

Matching pedestrians across multiple CCTV cameras have gained a lot of interest in recent years. Despite several attempts to address these challenges, it largely remains challenging mainly due to the following reasons. First, the images are captured under different lighting conditions. Therefore the perceived color of the subject appears to be different with respect to the illumination. Second, from surveillance cameras, no biometric aspects are available. Third, most often, the surveillance cameras will be of lower resolution. Modern person re-identification systems primarily focus on two aspects. (1) A feature representation for the probe and gallery images and (2) a distance metric to rank the potential matches based on their relevance. In the first category, majority of the works has been done on designing low level features. Since each of the features capture different aspects of the images, usually a combination of these features are used to obtain a richer signature. In the second category, the person re-identification is formulated as a ranking or a metric learning problem.

Discontinuous tracking of people across large sites, that is the search of a person of interest in different non-overlapping locations over different camera views, is a crucial task known as people reidentification. A more formal definition of the reidentification paradigm can be summarized as follow: given a probe set acquired at location X at time T0, re-identification aims to match its items with the subjects in a gallery set, collected at a different location Y at time T1.

Biometrics seems a viable solution to solve this problem when human subjects are to be tracked. It is widely accepted that biometric recognition under controlled data acquisition conditions is a relatively mature technology, which has proved to be effective for the actors of security agencies, public transports, governments, and in independent technology evaluation initiatives. However, the feasibility of biometric techniques under uncontrolled data acquisition conditions still raises considerable and often well-motivated scepticism. Just for this reason, re-identifying people moving across different sites covered with nonoverlapping cameras, could be the ideal scenario for testing technology under uncontrolled data acquisition conditions. As the image data (video) is rather large, the challenge will be to develop fast strategies for human biometric tagging. The application scenarios provide an extremely challenging video analysis task. Video streams will be captured both indoor and outdoor, and the size of people in the images (pixels on target) will range from a few tens to several hundreds pixels. face and gait biometrics are particularly promising for reidentification, since they can operate at a distance and do not require a detailed and/or high resolution image of the subject and/or its biometric traits.

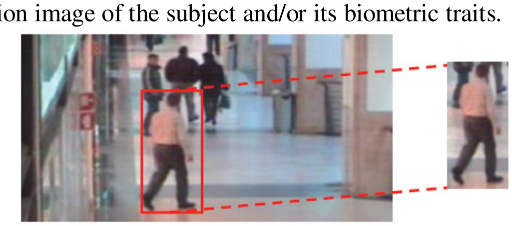


Figure 1.1 A typical person re-id scenario.

It is a highly challenging problem since human appearance usually exhibits large variations across different cameras. This variation is due to variability in backgrounds, sensor characteristics, lighting conditions, view-points, and human poses. Besides, distinct people may look similar if they wear clothes with the same color, which in turn increases the difficulty of finding correct associations.

B. *Problems associated with Person Reidentification*

One of the most critical challenges in facing the person re-identification problem is to recognize the same person viewed by disjoint, possibly nonoverlapping cameras, at different time instants and locations. Two fundamental problems are critical for person reidentification, feature representation and metric learning. An effective feature representation should be robust to illumination and viewpoint changes, and a discriminant metric should be learned to match various person images. Many efforts have been made along the two directions to tackle the challenge of person re-identification. Another aspect of person re-identification is how to learn a robust distance or similarity function to deal with the complex matching problem.

A significant challenge for person re-id is that people are often captured in different camera views at significantly different distances to the cameras, resulting in very different image resolutions. This difference in resolution between matching views, compounded by changes in lighting, pose and occlusion, makes re-identification extremely hard and unreliable. When it comes to the low-resolution (LR) person re-id problem, that is, matching LR person images to normal (higher) resolution ones, most (if not all) methods would simply normalize input images to a uniform normal scale. However LR person images contain much less information than those of normal resolution and many appearance details have been lost. A simple image magnification by interpolation thus would not recover the lost information in the LR person images.

II. IMAGE ACQUISITION AND PREPROCESSING

A. Color Correction

Regarding color correction, as depth increases, colors drop off one by one depending on their wavelength. First of all, red color disappears at the depth of 3 m approximately. At the depth of 5 m, the orange color is lost. Most of the yellow goes off at the depth of 10 m and finally the green and purple disappear at further depth. The blue color travels the longest in the water due to its shortest wavelength. The underwater images are therefore dominated by blue-green color. Also the light source variations will affect the color perception. As a consequence, a strong and non uniform color cast will characterize the typical underwater images.

Histogram equalization (HE) is popular method of contrast adjustment using the images histogram and also enhances a given image. In this method, transformation T is to be designed in such a way that the gray value in the output is equally distributed in [0, 1]. It is also called histogram flattening. Histogram equalization method in which histogram is modified by spreading the gray level areas. When an image's histogram is made equal, all pixel values of the image are redistributed so there are approximately an equal number of pixels to each of the user-specified output gray-scale classes e.g., 32, 64, and 256. Contrast is increased at the most populated range of brightness values of the histogram. For very bright or dark parts of the image, it automatically reduces the contrast associated with the ends of a normally distributed histogram. It can also divide pixels into different groups, if few output values are over a wide range. But this method is effective only when the original image has poor contrast to start with, otherwise it may degrade the image quality.

B. Dealing with Illumination Variation

Color is an important feature for describing person images. However, the illumination conditions across cameras can be very different, and the camera settings might also be different from camera to camera. Therefore, the perceived colors of the same person may vary largely from different camera views.

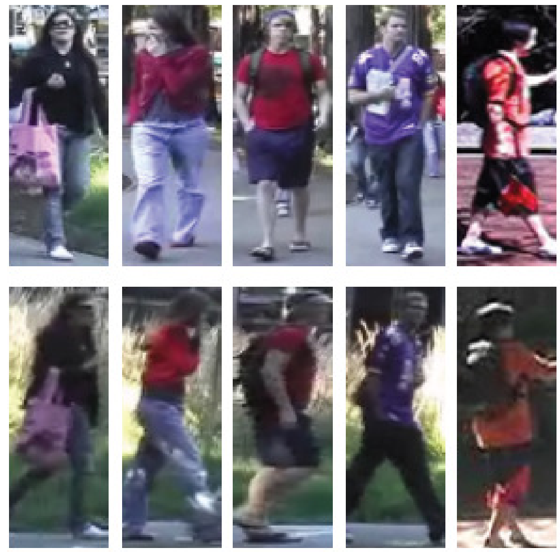


Figure 1.2 Example pairs of images from the VIPeR database

The above figure shows some sample images from the VIPeR database. It can be seen that images of the same person across the two camera views have a large difference in illumination and color appearance.

C. Dealing with Viewpoint Changes

Pedestrians under different cameras usually appear in different viewpoint. For example, a person with frontal view in a camera may appear in back view under another camera. Therefore, matching persons in different viewpoints is also difficult. To address viewpoint changes, it has been checked all subwindows at the same horizontal location, and maximize the local occurrence of each pattern (i.e. the same histogram bin) among these subwindows. The resulting histogram achieves some invariance to viewpoint changes, and at the same time captures local region characteristics of a person.

III. LITERATURE SURVEY

Discontinuous tracking[1] of people across large sites, that is the search of a person of interest in different non-overlapping locations over different camera views, is a crucial task known as people reidentification. A more formal definition of the reidentification paradigm can be summarized as follow: given a probe set acquired at location X at time T0, re-identification aims to match its items with the subjects in a gallery set, collected at a different location Y at time T1. Biometrics seems a viable solution to solve this problem when human subjects are to be tracked. It is widely accepted that biometric recognition under controlled data acquisition conditions is a relatively mature technology, which has proved to be effective for the actors of security agencies, public transports, governments, and in independent technology evaluation initiatives. However, the feasibility of biometric techniques under uncontrolled data acquisition conditions still raises considerable and often well-motivated scepticism. Just for this reason, re-identifying people moving across different sites covered with nonoverlapping cameras, could be the ideal scenario for testing technology under uncontrolled data acquisition conditions. As the image data (video) is rather large, the challenge will be to develop fast strategies for human biometric tagging. The application scenarios provide an extremely challenging video analysis task. Video streams will be captured both indoor and outdoor, and the size of people in the images (pixels on target) will range from a few tens to several hundreds pixels. face and gait biometrics are particularly promising for reidentification, since they can operate at a distance and do not require a detailed and/or high resolution image of the subject and/or its biometric traits.

The problem of appearance-based person re-identification is considered in this paper that has been drawing an increasing amount of attention in computer vision. It is a very challenging task since the visual appearance of a person can change dramatically due to different backgrounds, camera characteristics, lighting conditions, view-points, and human poses. Among the recent studies on person re-id, color information plays a major role in terms of performance. Traditional color information like color histogram, however, still has much room to improve. It is proposed that to apply semantic color names to describe a person image, and compute probability distribution on those basic color terms as image descriptors. To be better combined with other features, we define our appearance affinity model as linear combination of similarity measurements of corresponding local descriptors, and apply the RankBoost [2] algorithm to find the optimal weights for the similarity measurements. This work is evaluated on the highly challenging VIPeR dataset, and shown improvements over the state-of-the- art methods in terms of widely used person re-id evaluation metrics.

A novel appearance based method is used for person re-identification, that condenses a set of frames of the same individual into a highly informative signature, called Histogram Plus Epitome, HPE[4]. It incorporates complementary global and local statistical descriptions of the human appearance, focusing on the overall chromatic content, via histograms representation, and on the presence of recurrent local patches, via epitome estimation. The matching of HPEs provides optimal performances against low resolution, occlusions, pose and illumination variations, defining novel state-of-the-art results on all the datasets considered

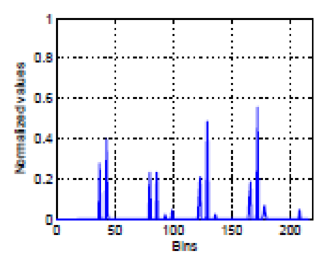
IV. METHODOLOGY

Person re-identification is a challenging problem due to the change in visual appearance caused by varying lighting conditions. Matching people across multiple camera views known as person re-identification. The perceived color of the subject appears to be different with respect to illumination. In this paper, a data driven approach is proposed for learning color patterns from pixels sampled from images across two camera views. It is perceived that, even though pixel values of same color would be different across views, they should be encoded with the same values. This framework is two fold, learning a linear transformation using color features and the other one is a dictionary to encode pixel values. The performance of the proposed work is superior when compared to the existing framework that uses color constancy model.

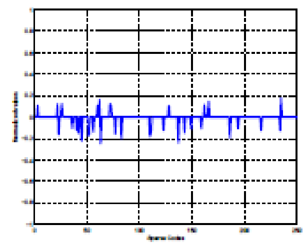
A. *Patch Formation*

The appearance of color changes across camera views due to a stark change in illumination. Fig.3.2a and Fig.3.2 d show an example of such changes in the appearance. It can be seen that the patches sampled, as shown in Fig.3a and Fig.3d, appears to be of different color. Histogram of the sampled patches appears to be as shown in the Fig.3.2b and Fig.3.2e respectively. And Fig.3.2c shows the encoding of the view of the image.

(a) View A (Image and the Extracted Patch)

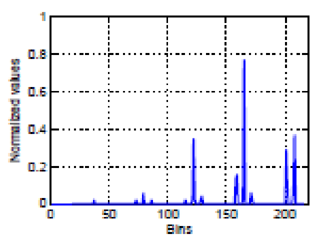


(b) RGB Histogram



(c) Encoding using our approach

(d) View B (Image and the Extracted Patch)



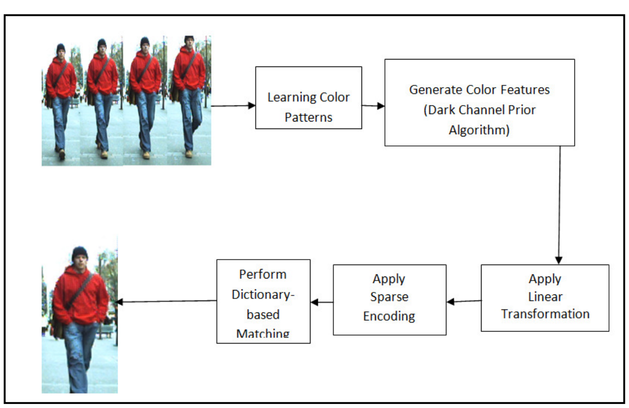
(e) RGB Histogram

Figure 3.2 Extracted Patches

B. Algorithm- Dark Channel Prior Algorithm

The algorithm works as follows,

i. Estimating the transmission
ii. Apply soft matting algorithm to refine the transmission
iii. Recovering the scene radiance
iv. Estimating the atmospheric light



V. EXPERIMENTAL SETUP

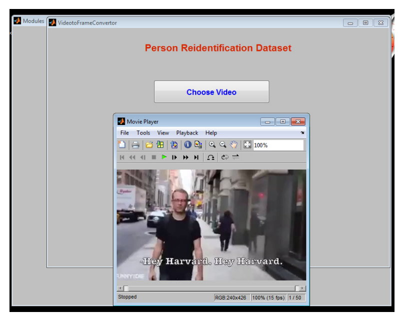
The proposed work uses illumination correction and find an optimal transformation to encode the pixel values in such a way that pixels corresponding to similar colors are close enough. The linear transformation and dictionary for sparse coding must be learned jointly. The joint learning improves the performance significantly for all the datasets. This is due to the fact that, for the encoding of each pixel, an optimal dictionary which can give same representation for pixels of same color has to be learned together with the linear transformation.

i. Loading a Video

Using the open dialog box, the user can select a video file of .avi format.

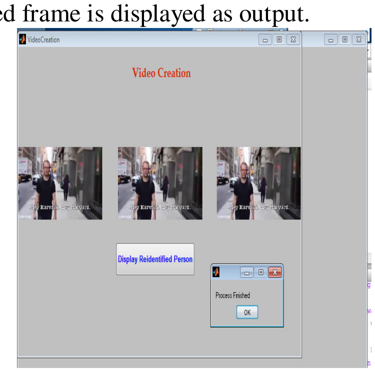
ii. Video to Frame Conversion

To perform the deblur and deskew operation fastly, the video need to be transformed to frames. And hence read the .av file and get the video information and write it as an image.



iii. Person Reidentification

Restored frame is displayed as output.



VI. CONCLUSION

A novel method based on data driven framework is proposed in this work for learning color features to handle illumination and other lighting condition changes across two camera views. In contrast to the previous works based on auto-encoders and sparse coding, in this method it combines them to learn a robust encoding jointly by forcing similar colors to be close to each other. The results are evaluated several baseline methods for achieving color constancy and have shown superior performance over all of them. By combining with other types of learned low-level and high-level features, it achieves promising results in several benchmark datasets.

References

1. Kviatkovsky, A. Adam, and E. Rivlin, "Color invariants for person reidentification," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, no. 7, pp. 1622–1634, July 2013.

2. D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, "Custom pictorial structures for re-identification," in Proceedings of the British Machine Vision Conference. BMVA Press, 2011, pp. 68.1–68.11.

3. B. Ma, Y. Su, and F. Jurie, "BiCov: a novel image representation for person re-identification and face verification," in British Machive Vision Conference, Guildford, United Kingdom, 2012, p. 11 pages.

4. D. Gray and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in Computer Vision ECCV 2008, ser. Lecture Notes in Computer Science, D. Forsyth, P. Torr, and A. Zisserman, Eds. Springer Berlin Heidelberg, 2008, vol. 5302, pp. 262–275.

5. D. Gray, S. Brennan, and H. Tao, "Evaluating appearance models for recognition, reacquisition, and tracking," in 10th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), 09/2007 2007.

6. M. Hirzer, C. Beleznai, P. M. Roth, and H. Bischof, "Person reidentification by descriptive and discriminative classification," in Proc. Scandinavian Conf. on Image Analysis, 2011, the original publication is available at www.springerlink.com.

7. J.-M. Morel, A. B. Petro, and C. Sbert, "Fast implementation of color constancy algorithms," Proceedings of the SPIE, vol. 7241, 2009.

8. E. H. LAND and J. J. McCANN, "Lightness and retinex theory," J. Opt. Soc. Am., vol. 61, no. 1, pp. 1–11, Jan 1971.

9. D. Forsyth, "A novel algorithm for color constancy," International Journal of Computer Vision, vol. 5, no. 1, pp. 5–35, 1990.

10. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," Image Processing, IEEE Transactions on, vol. 20, no. 9, pp. 2475–2489, Sept 2011.