

HUMAN RE-IDENTIFICATION USING COLOUR AND TEXTURE FEATURES

R. Newlin Shebiah, S. Arivazhagan

Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi

ABSTRACT

Human Re-Identification turns out to be the most fascinating and perplexing tasks in the domain of smart video surveillance in recent time. Re-Identification is crucial in establishing reliable tagging of person across multiple cameras or even within the same camera to re-establish detached or lost tracks. The challenge in Person Re-Identification lies due to the visual and spatio-temporal uncertainty in the appearance of the person's across multiple cameras. Here, a viewpoint invariant human re-identification framework with colour and texture features is proposed. The descriptor encodes the visual features of the human from the chromatic content described by colour statistical features, Local Binary histogram and Histogram of Gradients derived from wavelet transformed input image. Similarity measure is found by the Euclidean distance measured between the extracted feature descriptors from the non-overlapping cameras. From the experimental results it is observed that the proposed system significantly out performs state-of-the-art algorithms on the VIPeR Human Re-identification dataset.

Key words: Re-Identification, Colour Features, Histogram of Gradients, Texture Features, Local Binary Patte.

I. INTRODUCTION

In recent times, there is an increase usage of data from video cameras to fight against accidents, offenses, mistrustful activities, terrorism and sabotage. The existing digital video surveillance system processes voluminous data but leaves the task of threat detection to the human operators. The manual analysis of video is subjective to error and it is labor intensive. Re-identification is an interesting computer vision task provides tools security applications like tracking of suspected person across non-overlapping cameras, and retrieval of the frames with the person of interest [Gianfranco Doretto, et al., 2011].

There are varieties of hurdles in developing algorithms for re-identification. The primary challenges are intra class variation and inter class variation when analyzed from different views of camera. The underlying challenge of the human re-identification problem arises from momentous changes in the appearance due to the variations in viewing angle, illumination and pose. In order to recognize individuals in a real video surveillance system, human silhouettes have to be detected automatically. This brings a new challenge which comes from the preprocessing step and is not directly related to the re-identification problem but has a large influence on recognition accuracy when dealing with real time video surveillance.

Conventional biometrics such as face, palm print, iris, finger print, and gait have been widely used for human identity recognition; however, they are hard to apply to the re-id problem since clear image of the target cannot be captured by the surveillance cameras. Usually in person re-identification visual cues of the person signifies the low-level feature descriptors, identified by the photometric properties like color [Madden et al., 2007, Raflaa et al., 2017, Orwell, et al., 1999, Park et al., 2006] geometric

properties such as texture and spatial structure [Ma and Su et al., 2006, Farenzena, et al., 2010], or combinations [Farenzena et al., 2010, Gray et al., 2008, Gheissari, et al., 2001]. The principles behind using low level features are those of simplicity and computationally efficient and it provides efficient inter personal discrimination ability across different camera views. After extracting a suitable representation, distance based matching like Euclidean distance, Manhattan distance or nearest neighbor classifier [Farenzena et al., 2010] or model-based matching techniques like support vector machine can be engaged for matching and re-identification. For re-identification selection of representative features and a good classifier is equally important. The re-identification procedure is supported by a distance metric such as, Euclidean, L1-Norm or Bhattacharyya preferred to compute the correspondence between samples. On the other hand, the optimized distance metric may or fused multilayer metrics with supplementary information to enhance the correct matches or to reduce mismatches. Further, approaches for improving re-identification matching or representations may be grouped as (i) unsupervised [Madden et al., 2007, Ma, and Su, et al., 2012, Farenzena et al., 2010, Gheissari et al., 2006] or (ii) supervised and (iii) semi-supervised. However, it is also expected to depict re-identification scheme with respect to improving either feature representation or matching task.

The remaining section of this document is formatted as follows: The next section converses about the Proposed Method for human re-identification. The Section 3 gives the Experimental Results and Discussion. Finally, Section 4 gives the Concluding remarks of the proposed method.

II. PROPOSED METHODOLOGY

In this section, the proposed methodology of human re-identification using colour and texture features is explained. The block diagram of the proposed method for human re-identification is shown in Figure 1.

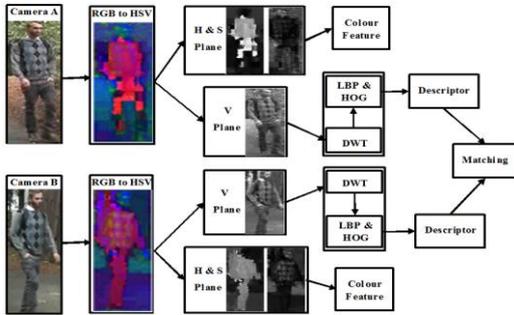


Fig 1: Block Diagram of Proposed method for Human Re-identification

Two Cameras from different locations were considered as input. From the input sequence, pedestrian is extracted by considering effective background subtraction algorithm and the detected person is tracked across the frames when considered a video. The HSV colour space representation is preferred for its unvarying nature. The hue is invariant to illumination and camera direction hence well suited for object retrieval. Here texture features from Local Binary pattern are computed from the luminance channel 'V', and statistical color features are computed from the chrominance channels 'H' and 'S' [Drimbarean, et al., 2001]. The component which corresponds to brightness of the color (V) is decomposed using Discrete Wavelet Transform [Daubechies et al., 1988, Strang. et al., 1996, Burrus, Goinath. et al., 1993] and LBP and HOG features were extracted from it. Feature matching is done by minimum distance method with Euclidean distance as a metric. Daubechies et al., [1988], Strang et al., [1996], Burrus et al., [1993] extracted LBP and HOG features from feature matching by doing they minimum distance method with Euclidean distance as a metric.

A. Histogram of Oriented Gradients (HOG)

The histogram of oriented gradients (HOG) is a feature descriptor for the purpose of human detection [Dalal et al., [2005]. The occurrences of gradient orientation in localized region of an image are counted for calculating HOG. On a dense grid of uniformly spaced cells, overlapping local contrast normalization is used for improved accuracy. Normalization of color and gamma values is a necessary pre-processing step. The most used method is to apply the one dimension centered, point discrete derivative mask in the horizontal or vertical directions or in both. Specifically, this method filters the image with the kernels $[-1,0,1]$ and $[-1,0,1]^T$.

The histogram which computes the pixel density of each pixel in an image, is computed from the outputs from the Gradient information. The contrast of the images are enhanced by Histogram normalization. In

this method, normalized cumulative sum is used to transform the histogram of the image. Then the intensity values of the image are mapped to new intensity to give a uniform histogram of intensity values. Window Descriptor is used to indicate the person detected from the previous block. Though the window may be of any size, the thinner window is preferable. Linear SVM is used for classification of positive and negative samples.

B. Local Binary Pattern (LBP)

Local Binary Patterns (LBP) proposed by Ren et al., [2012], Mäenpää, et al., [2003], Karthikeyan, et al., [2017] and Gray, et al., [2007], which is a scale invariant texture measure gives the effective representation of the local characteristics around each pixel. In a 3x3 neighborhood each surrounding pixel is compared with its eight neighbors and encoded with binary values. Clockwise concatenation of all these binary codes starting from the top-left one produces a binary number and its corresponding decimal value is referred to as Local Binary Patterns or LBP codes. Since the basic LBP operator of size 3x3 neighborhood is unable to capture features at large scale usage of different sizes of neighborhoods was adapted. A more generic local neighborhood of sampling a set of points that form a circle with the pixel to be labeled as its center is considered, and bilinear interpolation is used to intercept the sampling points that do not fall within the pixels, thus allowing for any radius and any number of sampling points in the neighborhood to deal with the texture at different scales. Variants in LBP are introduced by different shapes in neighborhood calculation and varying encoding techniques. In case of multiscale LBP, the radius of concentric neighbourhood is varied to capture LBP feature at different scales. When large radius is considered more LBP patterns is generated thus dimensionality of the derived features has to be taken in to consideration. Thus feature reduction techniques for effective representation of the features are preferred. Here, Principal Component Analysis (PCA) is used for feature reduction. PCA is an efficient and widely used linear technique in statistics that represents the data in a new coordinate system in which basis vectors has the greatest variance in the data.

C. Wavelet Transform

The Discrete Wavelet Transform (DWT) based on time-scale representation and affords multi-resolution analysis of images that finds application in Denoising [Daubechies et al., 1988, Strang. et al., 1996, Burrus, et al., 1993], Compression, Image fusion, pattern recognition, texture analysis etc. DWT make use of the wavelet function to detect variations and scaling function for approximating signals in both decomposition and reconstruction of signals or images. The 2D-DWT used is identical to Laplacian pyramid decomposition, where the sub-bands are logarithmically spaced in frequency. Two-dimensional decomposition by wavelet transform of an image results in four frequency sub-bands labeled as low-low (LL),

- Gray, D. and H. Tao, Viewpoint invariant pedestrian recognition with an ensemble of localized features. European Conference on Computer Vision, Marseille, France Pp. 262–275 2008.
- Gray, D., Brennan, S. and H. Tao, Evaluating appearance models for recognition, reacquisition and tracking. Proc. IEEE Int'l Workshop on Performance Evaluation of Tracking and Surveillance (2007).
- Karthikeyan, M.V., M.V.J. Martin and Leo Manickam, A 128-bit secret key generation using unique ecg bio-signal for medical data cryptography in lightweight wireless body area networks. Pak. J. Biotechnol. 14(2): 257–264 (2017).
- Ma, B., Y. Su, and F. Jurie, BiCov: a novel image representation for person re-identification and face verification. In British Machine Vision Conference, pages 57.1–57.11. British Machine Vision Association (2012).
- Madden, C., E.D. Cheng, and M. Piccardi, Tracking people across disjoint camera views by an illumination-tolerant appearance representation. Machine Vision and Applications 18(3-4): 233–247 (2007).
- Mäenpää, T. and M. Pietikäinen, Multi-scale Binary Patterns for Texture Analysis. In: Bigun, J., Gustavsson, T. (eds.) SCIA-03. LNCS, Springer, Heidelberg 2749: 885–892 (2003)
- Maenpaa, T., The Local Binary Pattern Approach to Texture Analysis—Extensions and Applications Pp. 42–47 (2003)
- Orwell, J., P. Remagnino and G.A. Jones, Multiple camera color tracking, IEEE Workshop on Visual Surveillance (1999).
- Park, U., A. Jain, I. Kitahara, and N. Hagita, ViSE: Visual Search Engine Using Multiple Networked Cameras. International Conference on Pattern Recognition Pp. 1204–1207 (2006).
- Rafila S. and H. Hussian, Molecular detection of entamoeba gingivalis using polymerase chain reaction, Pak. J. Biotechnol. 14 (3): 351-354 (2017)
- Ren X.M., Wang X.F. and Y. Zhao, (2012) An Efficient Multiscale Overlapped Block LBP Approach for Leaf Image Recognition. In: Huang D.S., Ma J., Jo K.H., Gromiha M.M. (eds) Intelligent Computing Theories and Applications. ICIC-12. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg Vol. 7390 (2012)
- Strang. G and T. Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA (1996).