A FRAMEWORK FOR PEOPLE RE-IDENTIFICATION IN MULTI-CAMERA SURVEILLANCE SYSTEMS

Sirine Ammar, Nizar Zaghden and Mahmoud Neji
MIRACL: Multimedia InfoRmation systems and Advanced Computing Laboratory, Tunisia

ABSTRACT

People re-identification has been a very active research topic recently in computer vision. It is an important application in surveillance system with disjoint cameras. This paper is focused on the implementation of a human re-identification system. First the face of detected people is divided into three parts and some soft-biometric traits are extracted from each part. In second step, we can recognize people even if their faces are hidden or they are with back appearance. The features extraction will be carried out according to the overall characteristics of the complete images of different persons. An algorithm that identifies people from their body shape will be developed. A powerful representation of the person based on the characteristics of color, texture and shape as well as different soft-biometric features is suggested. The experiments are carried out on SAIVT-SoftBio database which consists of videos from disjoint surveillance cameras as well as some static image based datasets (MUCT, VIPeR, CVSRP).

KEYWORDS

People Re-Identification, Soft-Biometric, Surveillance System

1. INTRODUCTION

Today, video surveillance systems are widely used in our daily life. These surveillance systems are applied in many areas including home security, surveillance of public places, control some private access, resolution of criminal investigations and so on. In many surveillance purposes, it is desirable to determine whether a person has already been detected by a multi-camera system, known as people re-identification. Biometric information can be very effective in identifying people, but there are limits to use them, namely the need for high-resolution images and are captured in the closest distance. In fact, the principal problem in tracking or people re-identifying is to create a model that can represent people in a unique and accurate way. In this paper, a framework is suggested to construct human appearance model for people re-identification in disjointed camera surveillance system. This approach allows to resolve some problems (theft in public or private places, criminal investigations, etc) and their possible unfavorable results.

The rest of the paper is organized as follows: Section 2 gives a general view on re-identification approaches, followed by the choice of our method compared to the state of the art in section 3. Finally, conclusion is addressed in section 4.

2. GENERAL VIEW ON RE-IDENTIFICATION APPROACHES

2.1 Global Approaches vs. Local Approaches

2.1.1 Global Approaches

The global approaches allow the recognition of images based on visual similarities measured on the whole images. A globally described image is represented by a single attribute vector. Many global approaches exist to make re-identification of people. Among the best known are the works of (D. Gray / 2008) which consist of extracting texture features such as Schmid and Gabor, as well as eight color channels from the RGB, YCbCr, and HSV color spaces. In (A Derbel et al. / 2014), two new representations of the motion distribution

named gait frequency representation (RFD) and gait envelope representation (RED) are presesnted. These two representations are computed from the binary images of the silhouettes. In the work of (Nakayima et al. 2003) Two-dimensional normalized color histograms are computed; r = R/(R+G+B), g = G/(R+G+B). The characteristics of shapes were calculated by counting the pixels along the rows and columns of the extracted body images.

2.1.2 Local Approaches

This type of approaches consists of segmenting the image to divide it into local areas and then calculating the characteristics for each of the extracted regions. Most of these methods use points of interest (Nizar Zaghden et al/2013) that are based on characteristic structures (blobs or ridges) in the image. Many local approaches exist to make re-identification of person. Among the best known, we can cite the work (W. Schwartz et al/2010), in which characteristics of texture (Local Binary Pattern) and shape (Histogram of Oriented Gradient) and color (captured by averaging the intensities of pixels) are combined to represent each block of the cropped face. (O. Hamdoun / 2008) uses interest points descriptors to obtain the signature for a person. In (Vaquero, 2009) First, human body is segmented in face, torso and legs. Then, Normalized color histogram is extracted for each body part in HSL space. The region of the face is partitioned into three parts as presented in Figure 1 and some soft biometric features are extracted from each head area.

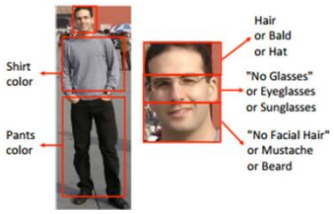

Figure 1. Body Parts and Attributes Considered in the Implementation

In (Lyes Hamoudi / 2011) the elements of the foreground are obtained. After, the entire blob is divided into three areas using height ratios equal to [1/5, 3/10, 1/2] of the total blob size as presented in Figure 2. The color characteristics used are the mean values of the colors. The extracted texture characteristics are based on the co-occurrence matrix.

Figure 2. Body Detection and Split Procedure

In (Farenzena / 2010), a System Drive Accumulation of Local Features (SDALF) approach is presented. In order to individualize the human parts, the author used the bilateral chromatic operator and the spatial coverage operator. As for the representation, weighted color histogram, Maximally Stable Color Regions (MSCR) and Recurrent High-Structured Patches (RHSP) are described.

2.2 Intrusive Approaches vs. Non-Intrusive Approaches

Approaches can be categorized into two broad groups: intrusive or non-intrusive. Intrusive techniques require that the individual be in direct contact with the acquisition equipment or located at a small distance compared to the camera. The non-intrusive approaches consist of identifying persons without having touch with the acquisition equipment. Among these methods we can cite the works of (Vaquero/2009) and (Lyes Hamoudi/2011).

In the table below (table1), we present a classification of the different approaches of people re-identification by focusing on the human model.

Table 1. Summary Table of People Re-Identification Approaches

References	Features	Categories		representation			Categories	
		local	global	color	textu re	shap e	intru sive	Non-intrusive
(D. Gray/2008)	Schmid, Gabor, eight color channels		✓	✓	✓			✓
(Ahmed Derbel / 2014)	RFD, RED		✓			✓		✓
(C. Nakajima et al. 2003)	Color histogram, shape histogram		✓	✓		✓		✓
(W. Schwartz et al/2010)	LBP, HOG, average color values	✓		✓	✓	✓		✓
(O.Hamdoun /2008) (Nizar Zaghden /2013)	interest points descriptors	✓				✓		✓
(Vaquero/2009)	A normalized color histogram.	✓		✓				✓
(Lyes Hamoudi/2011)	Average color values, The co-occurrence matrix	✓		✓	✓			✓
(Farenzena/2010)	weighted color histogram, MSCR, RHSP	✓		✓	✓			✓

2.3 Supervised Approaches vs. Unsupervised Approaches

2.3.1 Supervised Approaches

In this classification approach, knowledge of the meaning of each class is required. The number of classes is fixed (in prior) and the available images are already allocated to the different classes of the database. The aim is thus to assign new images to the appropriate classes following two types of methods.

Discriminative methods - Discriminative models like SVM and boosting are widely used for feature learning in order to find the discriminant representations of the Region Of Interest. The matching phase, in (Truong Cong et al. 2010a) is based on the SVM method. In (Gray et al. / 2008). The proposed similarity function is a weighted ensemble of likelihood ratio tests, constructed with the AdaBoost algorithm.

Metric learning - Another direction is to learn task-specific distance functions with metric learning algorithms. (Weinberger and Saul/2009) briefly introduce the metric learning framework for large margin nearest neighbor (LMNN) classifier. The goal is to learn a linear transformation which minimizes the distance between each point of the small learning set and its K nearest neighbors similarly labeled, while maximizing the distance between all points labeled differently according to a constant margin.

2.3.2 Unsupervised Approaches

The unsupervised approach consists in determining the different classes without any previous knowledge. In (Farenzena *et al.* /2010) [6] As for matching, similarity between two images is defined as the weighted sum of the Bhattacharyya and the euclidean distances.

In the table2, we present a classification of different similarity calculation and machine learning methods.

Table 2. Summary Table of Similarity Measures and Matching Methods of People Re-Identification Approaches

references	method	classification	
		Supervised	Unsupervised
(Gray *et al.* / 2008)	Boosting	✓	
(Farenzena/2010)	Bhattacharyya distance, Euclidean distance		✓
(Truong Cong *et al.*/2010a)	Support Vector Machine	✓	
(Weinberger andSaul/2009)	LMNN	✓	

3. CHOICE OF OUR METHOD COMPARED TO THE STATE OF THE ART

The characteristics presented in global approaches are calculated in an overall aspect on the image. Although the overall representations of the images are easy to construct and invariant to the position of the objects, they only provide a rough and not detailed aspect of the image, therefore, some localized information are lost. In order to overcome these problems the image can be represented as set of local image parts. On the other hand, dealing only with local characteristics risks losing the overall sense of the image, by submerging it in a stream of unnecessary little details. Furthermore, this type of approaches require good image resolution. In

our image search system, we combine global and local approaches in the measurement of the similarity between two images. To identify people by face, it requires videos or images with high resolution and being captured in nearest distance. Also, face identification is mainly focusing in the frontal face. That's why the intervention of people is required. We can recognize people even if their faces are hidden or they are with back appearance. An algorithm that identifies people from their body features will be developed. First of all, it is imperative to carry out a robust and adequate segmentation of human body, which is to isolate and identify each of the person's limbs. Then, several measurements may be taken. In matching phase, If an entire appearance of the person's body is available, it will be better to use this information in order to limit the number of searches in the database. This means that it is not necessary to browse the entire database, because many people will be eliminated only at the sight of the body.Thus, sufficient precision must be achieved for this be truly discriminatory and for candidates not to be falsely identified and / or excluded.

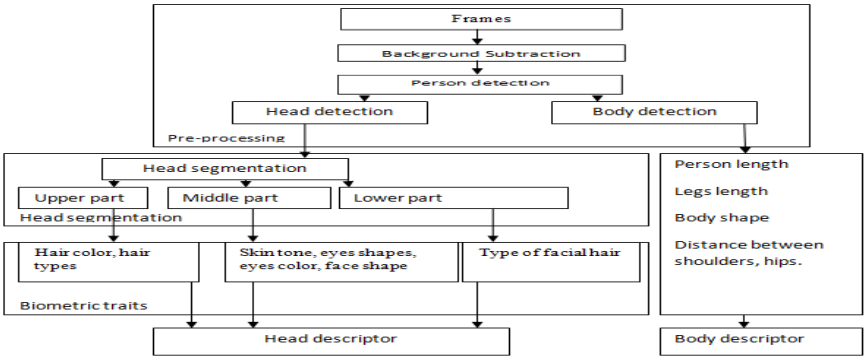

Figure 3. Overview of the Proposed Framework

4. CONCLUSION

In this paper, we present a person re-identification framework for multi-camera surveillance system based on both soft-biometric features and the overall appearance of person body. People detection, body segmentation and features extraction are included in this framework. In correspondence phase, first the comparison will be carried out according to the captures of faces of the different people. In second stage the comparison will be carried out according to the overall characteristics of the complete images of the different persons.

REFERENCES

Ahmed Derbel et al, Jun 2014 Human identification in frontal view using the silhouette motion distribution. *Shape Recognition and Artificial Intelligence (RFIA)* 2014, France.

C. Nakayima et al, September 2003. Full-body person recognition system. *Pattern Recognition*, 36(9) :1997–2006.

D. A. Vaquero et al, 2009. Attribute-based people search in surveillance environments, in: *Workshop on Applications of Computer Vision (WACV)*, pp. 1–8.

D. Gray, H. Tao, 2008 Viewpoint invariant pedestrian recognition with an ensemble of localized features, in: *Proc. European Conference on Computer Vision (ECCV)*, pp. 262–275.

Lyes Hamoudi, 2011 Learning techniques application to person detection and recognition, à Lille 1

M. Farenzena et al,2010 Person re-identification by symmetry-driven accumulation of local features, in: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2360–2367. pp. 381–390.

Nizar Zaghden, Remy Mullot, and Adel M.Alimi, « Categorizing ancient documents », IJCSI 2013, Vol. 10, Issue 2, No 2, March 2013.

O. Hamdoun et al, 2008 Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences, in: *Second ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC)*, pp. 1–6.

Truong Cong et al, (2010a). People re-identification by spectral classification of silhouettes. *Signal Processing*, 90(8): 2362–2374.

Weinberger, K., Saul, L(2009). Distance metric learning for large margin nearest neighbor classification. *The Journal of Machine Learning Research* 10, 207–244.

William Schwartz et al. 2010 A robust and scalable approach to face identification. *European Conference on Computer Vision (ECCV)*, pages 476–489.