Fast Appearance-based Person Re-identification and Retrieval Using Dissimilarity Representations

Riccardo Satta, Giorgio Fumera, and Fabio Roli

Dept. of Electrical and Electronic Engineering, University of Cagliari
Piazza d'Armi, 09123 Cagliari, Italy
e-mail: {satta, fumera, roli}@diee.unica.it
WWW: http://prag.diee.unica.it

Abstract. Person re-identification consists of recognizing an individual who has previously been observed over a camera network. It is a recently introduced computer vision task that can provide useful tools for many applications of video-surveillance. Person re-identification exhibits several challenging issues. Most notable ones are pose variations, partial occlusions, and changing lighting conditions. Another relevant issue is computational complexity. In this contribution, we present the results attained by the Pattern Recognition and Applications (PRA) Group in this field: 1) a framework for person re-identification methods; 2) a dissimilarity-based approach for speeding up existing methods; 3) a dissimilarity-based approach for retrieving images of individuals based on a textual query describing clothing appearance, instead of an image.

Keywords: person re-identification, video-surveillance systems, people retrieval, dissimilarity-based representations

1 Introduction

Person re-identification consists of recognizing an individual who has previously been observed over a camera network. It is a recently introduced computer vision task that can provide useful tools for many applications of video-surveillance, e.g., on-line tracking of individuals over different, non-overlapping cameras, and off-line retrieval of the video sequences containing an individual of interest. Formally, person re-identification can be modelled as an object *matching* problem. Each individual is represented using a *descriptor*, which is built from one or more video frames. The task is to rank the individuals in a "template" gallery (e.g., all the ones observed over a camera network in a given period of time) according to their similarity to a given "query" individual.

Due to unconstrained poses and low size of the region of video frames containing a person, face recognition techniques are usually ineffective for this task. For this reason, a widely used approach is to build descriptors based on clothing appearance ("appearance-based person re-identification") [2]. Other works exploit different "soft" biometric traits, like gait. Appearance-based person re-identification exhibits several challenging issues. Most notable ones are pose variations, partial occlusions, and changing lighting conditions. Another relevant issue is computational complexity.

In this contribution, we present the results attained by the Pattern Recognition and Applications (PRA) Group in this field: 1) a framework for appearance-based re-identification methods; 2) a dissimilarity-based approach for speeding up existing methods; 3) a dissimilarity-based approach for retrieving images of individuals based on a *textual* query describing clothing appearance.

In [8] we proposed a framework, called Multiple Component Matching (MCM), for appearance-based re-identification methods. MCM gives a unifying view of current works, and provides a common foundation to analyse and improve them, as well as to define new methods. MCM is based on a multiple component representation (e.g., patches, or interest points) and a body part subdivision (e.g., upper/lower body), that are common to most works.

In [6, 7] we addressed the issue of computational complexity, which is crucial for real-time applications, but has been overlooked in the literature so far. To this aim, we exploited dissimilarity-based representations [5], in which dissimilarities are computed to prototypes representative of *local* components. We developed the Multiple Component Dissimilarity (MCD) framework, that allows one to obtain a dissimilarity-based descriptor (basically, a vector of dissimilarity values) from any appearance-based descriptor. Once descriptors are turned to vectors of real numbers, their matching time becomes very fast.

While person re-identification methods can be useful in forensic investigations, they need an image of the individual of interest to be given as query. However, in many cases a *textual* description of clothing appearance may only be available (e.g., a description provided by a witness, such as "a person wearing a black t-shirt and white trousers"). Retrieving video sequences containing individuals that match a textual query would be very useful in such scenarios. This problem has been addressed by very few authors so far [9]. In our ongoing work we are investigating the use of our MCD framework to perform this task. While previous methods rely on ad hoc descriptors, our MCD-based approach can be applied to *any* appearance-based descriptor.

We present the MCM framework of [8] in Sect. 2. The MCD dissimilarity framework of ([6, 7]) is presented in Sect. 3. Finally, we illustrate our ongoing work on retrieving images of individual through a textual query, in Sect. 4. Conclusions and future research directions are drawn in Sect. 5.

2 The Multiple Component Matching framework

Despite the existing methods for person re-identification exhibit many differences, it can be noted that all of them are based on some part-based body representation (e.g. torso-legs), and/or exploit more or less implicitly the concept of multiple instances (patches, points of interest). Our MCM framework provides an unifying view of appearance-based person re-identification methods, embed-

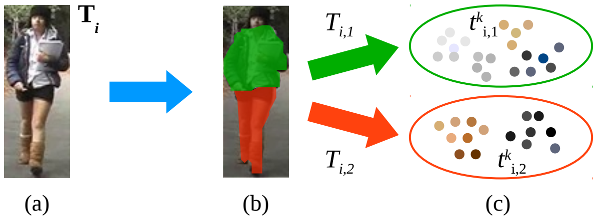


Fig. 1. An example of the MCM representation. (a) The image of an individual. (b) The body is subdivided into parts (in this case, upper and lower body). (c) Each part is a set of components (e.g., image patches), represented here as coloured dots.

ding the common concepts of multiple instance representation and body part subdivision. In MCM, individuals are represented as bags of instances, named "sets of components". Such components can be any kind of local features, like patches or interest points, extracted from the image of an individual. If a body part subdivision is used, a different set of components is extracted from each part. The rationale behind this representation is to gain robustness to partial occlusions and pose variations.

Formally, let $\mathcal{T} = \{\mathbf{T}_1, \dots, \mathbf{T}_N\}$ be a *template gallery* of N individuals. Each template $\mathbf{T}_i$ is represented via an ordered sequence of M sets, corresponding to M body parts ($M \geq 1$):

$$\mathbf{T}_i = \{T_{i,1}, \dots, T_{i,M}\}. \tag{1}$$

According to the multiple instance representation, each part $T_{i,m}$ is a set of $n_{i,m}$ of components described by feature vectors $\mathbf{t}^k_{i,m}$:

$$T_{i,m} = \{\mathbf{t}^1_{i,m}, \dots, \mathbf{t}^{n_{i,m}}_{i,m}\},\ \mathbf{t}^k_{i,m} \in \mathbb{X}, \tag{2}$$

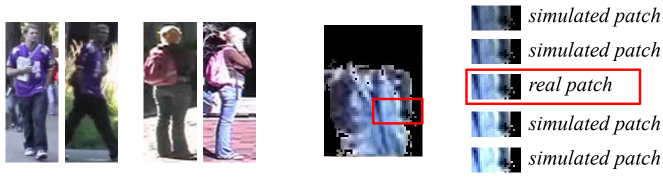
where $\mathbb{X}$ denotes the feature space (without losing generality, we assume that all sets are represented with the same features). An example of the MCM representation is shown in Fig. 1.

Given a probe $\mathbf{Q}$, represented as a sequence of parts as well, the task is to find the most similar template $\mathbf{T}^* \in \mathcal{T}$, according to a similarity measure $D(\cdot, \cdot)$:

$$\mathbf{T}^* = \arg\min_{\mathbf{T}_i} D(\mathbf{T}_i, \mathbf{Q}). \tag{3}$$

Given the above representation, D must be defined as a similarity measure between ordered sequences of sets; for instance, a linear combination of the M distances $d(T_{i,1}, Q_1), \dots, d(T_{i,M}, Q_M)$. The measure d can be in turn any similarity measure between sets. A suitable one is the *Hausdorff Distance* d_{H} [11].

Different implementations of MCM can be defined by a specific choice of its parameters, namely, the part subdivision, the components extracted from each part and their representation, and the similarity measures d and D. In [8], a simple, direct implementation of MCM, named MCMimpl, was proposed, that attained state-of-the-art performance. It adopts a two-part subdivision of the body



(a) (b)

Fig. 2. (a) Two examples of different images for the same individual: note the difference both in contrast and brightness. (b) Examples of four artificial patches simulating changing illumination (right), corresponding to the patch highlighted on the left.

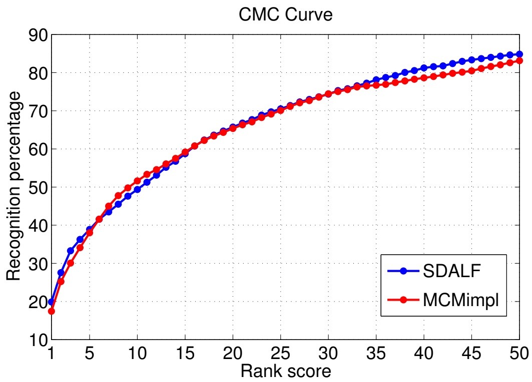


Fig. 3. *MCMimpl* compared with *SDALF* on the VIPeR dataset.

into torso and legs as in [3], and extracts multiple, partly overlapping, rectangular patches from each part. Every patch is described by an HSV histogram and its vertical position. MCMimpl inherits robustness to partial occlusions and pose variations from the multiple component/multiple parts representation of MCM. To obtain robustness also to changing lighting conditions, *synthetic* components corresponding to different lighting conditions are added together with the original ones. Since light variations usually result in a change of both brightness and contrast of the image (see for example Fig. 2-a), we developed a *simulation* algorithm that changes both of them simultaneously Fig. 2-b). Further details on the algorithm can be found in [8].

The performance of MCMimpl was assessed in [8] on the VIPeR dataset [4], a challenging corpus composed by two non overlapping views of 632 different pedestrians, which show varying changing conditions and pose variations. Ten random subsets of 316 pedestrians were drawn from the original dataset. The gallery set is composed by the first image of each person; the probe set, by the second one. Images of the probe set are compared to the images of the gallery set to find the best match. The corresponding cumulative matching characteristics (CMC) curve (that is, the probability of finding the correct match over the first n ranks) is shown in Fig. 3, and compared with the SDALF method [3].

3 The Multiple Component Dissimilarity framework

In [6] we showed that the MCM framework suggests a general method to reduce the computational complexity of appearance-based re-identification methods. We observed that, regardless of the specific descriptor (i.e., the body part model and the kind of local components), if the clothing of different individuals share similar local characteristics, their set-of-components representation exhibits some redundancy, due to similar local components. Such redundancy can be reduced by turning any set-of-components representation into a dissimilarity-based one [5], which consists of representing each individual as a vector of dissimilarity values to pre-defined *visual prototypes*. We proposed to define also prototypes as sets of components, each corresponding to a local characteristic shared by different individuals.

This allowed us to define the Multiple Component Dissimilarity (MCD) framework in [6, 7]. MCD provides a dissimilarity-based version of any re-identification method that can be framed in MCM, adopting a very compact representation of individuals (basically, a small set of dissimilarity vectors). As such, matching two descriptors becomes very fast, as it reduces to compare two vectors of real numbers. Dissimilarity-based descriptors of MCD try to keep the same discriminative capability and robustness of the original re-identification method, by exploiting the same part subdivision and local features, but grouped in homogeneous clusters. The MCD framework is summarised in the rest of this section.

The first step of MCD is prototype construction. Prototypes are constructed from a given gallery of, say, N individuals, starting from their MCM descriptors, denoted as:

$$\mathcal{I} = \{\mathbf{I}_1, \ldots, \mathbf{I}_N\} . \tag{4}$$

According to MCM (Sect. 2), each descriptor $\mathbf{I}_i \in \mathcal{I}$ is an ordered sequence of M elements, one for each body part. The m-th element is a set of n_m feature vectors:

$$\mathbf{I}_i = \{I_{i,1}, \ldots, I_{i,M}\}, \quad I_{i,m} = \{\mathbf{x}_m^1, \ldots, \mathbf{x}_m^{n_m}\}, m = 1, \ldots, M . \tag{5}$$

The procedure for prototype construction is the following. For each body part $m = 1, \ldots, M$:

1. Merge the feature vectors of the m-th part of each $\mathbf{I} \in \mathcal{I}$ into a set $X_m = \bigcup_{j=1}^{N} I_{j,m}$;
2. Cluster the set X_m into a set $\mathbf{P}_m$ of N_m clusters, $\mathbf{P}_m = \{P_{m,1}, \ldots, P_{m,N_m}\}$. Take each cluster as a prototype for the m-th body part.

Each prototype is a set of visually similar image components, which can belong to different individuals. In turn, each original descriptor $\mathbf{I}$ consists of a set of components for each body part. Thus, to create a dissimilarity vector from $\mathbf{I}$, dissimilarities can be evaluated via a distance measure between sets. In [6] the *k-th Hausdorff Distance* was used, due to its robustness to outliers. The above procedure returns M sets of prototypes, one for each body part:

$$\mathcal{P} = \{\mathbf{P}_1, \ldots, \mathbf{P}_M\} . \tag{6}$$

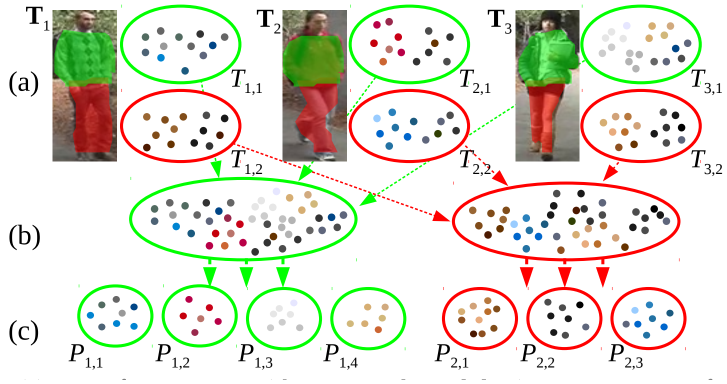


Fig. 4. Generation of the prototype gallery in MCD. In this example, the body is subdivided into two parts: upper (green) and lower (red) body. (a) A template gallery of three individuals, represented according to MCM. (b) All the components of the same part are merged. (c) A clustering algorithm is applied, and a number of prototypes (clusters) are generated for each part.

Fig. 4 summarises the process of prototypes generation, for $M = 2$ parts.

Once the prototypes have been constructed, given the MCM descriptor of any individual, $\mathbf{I} = \{I_1, \ldots, I_N\}$, its MCD descriptor is obtained as a sequence of M dissimilarity vectors $\mathbf{I}^{\mathrm{D}} = \{I_1^{\mathrm{D}}, \ldots, I_M^{\mathrm{D}}\}$, where:

$$I_m^{\mathrm{D}} = \big(d(I_m, P_{m,1}) \ldots d(I_m, P_{m,p_m})\big), m = 1, \ldots, M \ . \tag{7}$$

The MCD representation of above was exploited in [6, 7] for the task of person re-identification. In this context, it exhibits two clear advantages over complex descriptors used by other methods. One is a considerable reduction in storage requirement: only a small set of dissimilarity vectors (one for each body part) for each individual, and the set of prototypes, need to be stored. The other advantage is a reduction in processing time: comparing descriptors becomes as simple as computing distances between vectors of real numbers, which is very fast with modern CPUs. This can enable several useful applications, like real-time re-identification of an individual, among a huge number of candidates.

3.1 Trade-off between re-identification accuracy and matching time

The results reported in [7] showed that this method exhibits a much lower matching time and memory requirement than its non-dissimilarity-based version, as expected. However, this is sometimes attained at the expense of a lower re-identification accuracy. Trading a lower accuracy for a lower processing time can be nevertheless advantageous in certain scenarios. As an example, consider a real time application in which individuals observed by different, non-overlapping cameras are automatically tracked, and a human operator can select an individual of interest from one of the videos, and ask the system to re-identify it (again,

	MCMimpl	$\text{MCMimpl}^{\text{Dis}}$
Avg time for template descriptor creation	93.7 ms (1)	17.5 ms
Avg time for probe descriptor creation	6.8 ms (1)	17.5 ms
Avg time for a single match	28.6 ms	0.004 ms
Avg total time for a single run (119 templates)	2719.1 sec	63.5 sec
Avg total time for a single run (316 templates)	2887.6 sec	87.2 sec
Avg total time for a single run (474 templates)	6521.0 sec	134.5 sec
Avg total time for a single run (632 templates)	11550.6 sec	179.4 sec

Table 1. Comparison of the computational requirements of MCMimpl and MCMimpl^{Dis}. Notes: (1) in MCMimpl, the construction of a template descriptor includes the generation of simulated patches, and thus requires a higher time than the construction of a probe descriptor.

in real time). In this case, the overall time t_{s} needed to find the person of interest among the ranked list of individuals returned by the system, is the sum of two quantities [7]:

$$t_{\text{r}} = t_{\text{p}} + t_{\text{s}} \tag{8}$$

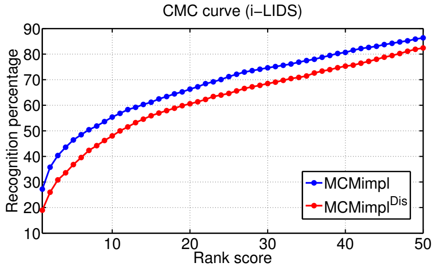
The first quantity is the average processing time $t_{\text{p}} = t_{\text{d}} + N t_{\text{m}}$, given by the time needed for constructing the probe descriptor t_{d} plus the time needed to match it to the N template descriptors (t_{m} denotes the time for performing a single match). The second quantity, t_{s}, is the average time spent by the operator to search the person of interest among the ranked results. It is given by $t_{\text{s}} = t_{\text{c}} E\{R\}$, where t_{c} is the time required by the operator to compare the probe image with one template image, and $E\{R\}$ is the expected rank of the probe individual, that can be computed from the Cumulative Matching Characteristics (CMC) curve.

3.2 Experimental results

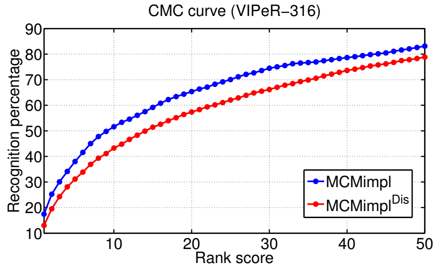
In [7] we applied the MCD framework to the MCMimpl method of Sect. 2, and conducted a thorough experimental evaluation of the resulting dissimilarity-based method (denoted in the following as $\text{MCMimpl}^{\text{Dis}}$). Experiments were carried out on two benchmark datasets used in many previous works: VIPeR [4], and a set of images taken from the i-LIDS MCTS video dataset [12]. We considered three subdivisions of VIPeR, with respectively 316, 474, and 632 pairs of individuals. We refer to the above three versions of VIPeR respectively as VIPeR-316, VIPeR-474, and VIPeR-632. The i-LIDS dataset contains 476 images of 119 different pedestrians. One image for each person was randomly selected to build the template gallery; the other images formed the probe gallery.

Here we sum up the results obtained. First, we evaluated the average processing time of $\text{MCMimpl}^{\text{Dis}}$, attained on a 2.4 GHz CPU, and compared it with MCMimpl. Results are shown in Table 1.

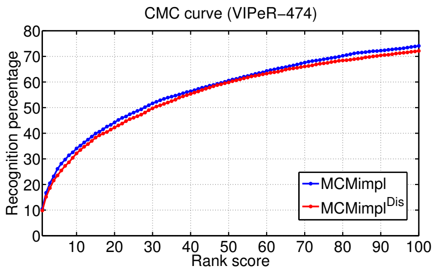
As expected, $\text{MCMimpl}^{\text{Dis}}$ clearly outperforms MCMimpl in terms of processing time. Regarding re-identification accuracy, the average CMC curves of MCMimpl and $\text{MCMimpl}^{\text{Dis}}$ on the four datasets are reported in Fig. 5(a)–(d).



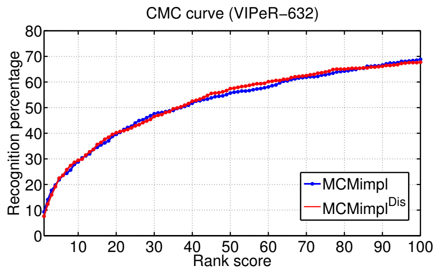
(a)



(b)



(c)



(d)

Fig. 5. CMC curves attained by MCMimpl and MCMimpl$^{\mathrm{Dis}}$ on the four datasets used in the experiments.

	MCMimpl	MCMimpl$^{\mathrm{Dis}}$
Re-identification time t_{r}, *i-LIDS* dataset ($t_{\mathrm{c}} = 0.5$ sec)	13.600 sec	12.331 sec
Re-identification time t_{r}, *VIPeR-316* dataset ($t_{\mathrm{c}} = 0.5$ sec)	21.268 sec	16.730 sec
Re-identification time t_{r}, *VIPeR-474* dataset ($t_{\mathrm{c}} = 0.5$ sec)	56.039 sec	43.213 sec
Re-identification time t_{r}, *VIPeR-632* dataset ($t_{\mathrm{c}} = 0.5$ sec)	74.023 sec	57.830 sec

Table 2. Comparison of the overall re-identification time of MCMimpl versus MCMimpl$^{\mathrm{Dis}}$.

MCMimpl$^{\mathrm{Dis}}$ attained a worse recognition performance than MCMimpl on the smallest template galleries (i-LIDS and VIPeR-316).The accuracy gap diminished as the template gallery grows. This suggests that, when the number of templates is very high, as in many practical applications, the dissimilarity-based version of a re-identification method can attain the same performance as the original method, while requiring much lower computational resources. Moreover, by using the model of Sect. 3.1, we proved that even if the performance are worse, the trade-off between accuracy and processing time can be advantageous in the application scenario described in Sect. 3.1 The overall re-identification time t_{r} of MCMimpl and MCMimpl$^{\mathrm{Dis}}$, evaluated through Eq. (8) is reported in Table 2.

The processing time of MCMimpl$^{\mathrm{Dis}}$, as well as of any dissimilarity-based method obtained via MCD, is affected by prototype construction. This can be a problem, especially in applications where new templates can be added on-line

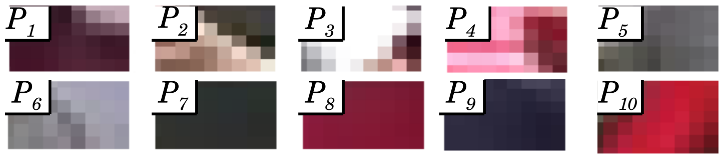


Fig. 6. Prototypes obtained from the upper body parts of a set of individuals.

during system operation, e.g. new individuals that are observed by a camera network. Nevertheless, in [7] we showed that the prototype gallery can be constructed using only a subset of the whole template gallery, or even using gallery of individuals *different* than the template gallery, without affecting the performance. This can avoid to re-build the prototype gallery (and thus, the dissimilarity representation of the existing templates) each time a new template is added to the system. In particular, in the latter case prototypes can be generated off-line, prior to system operation. To this aim, it would be desirable to use a dataset with a wide range of different clothing characteristics.

4 A general method for appearance-based people search using MCD

In our ongoing work, we are considering a task similar to person re-identification, that we call "appearance-based people search". It consists of finding, among a set of images of individuals, the ones relevant to a *textual* query describing clothing appearance of an individual of interest. This can be useful in applications like forensics video analysis, where a textual description given by a witness can be available, instead of an image of the individual of interest. To our knowledge, an analogous task ("person attribute search") was considered so far only in [10, 9]. We devised a general approach to extend appearance-based person re-identification systems, exploiting the same descriptors of clothing appearance to enable also the people search functionality based on a textual query. Our approach relies on dissimilarity-based descriptors, which can be obtained using the Multiple Component Dissimilarity (MCD) framework of [6] from *any* appearance descriptor that uses a body part subdivision and a multiple instance representation. The advantage of dissimilarity descriptors is that they allow one to build detectors of the attributes of interest (e.g., the presence of a colour in the torso) using supervised classifiers, without requiring techniques tailored to the specific, original descriptors, as in the approaches of [10, 9].

Our intuition is that the clothing characteristics that can be detected by a given appearance descriptor, according to its low-level features and part subdivision (e.g., "red shirt") may be encoded by one or more visual prototypes. For example, the rectangular image patches in Fig. 6 are sample components of 10 prototypes, extracted from the upper body parts of individuals taken from the VIPeR data set, using the MCD implementation of [6]. Intuitively, descriptors

Class (cardinality)	MCD_1	MCD_2	MCD_3
red shirt (51)	**0.845**	0.780	0.792
blue/light blue shirt (34)	**0.645**	0.523	0.494
pink shirt (35)	0.534	**0.578**	0.461
white/light gray shirt (140)	**0.771**	0.736	0.758
black shirt (156)	0.728	0.705	**0.736**
orange shirt (10)	**0.689**	0.580	0.463
violet shirt (18)	0.422	0.235	**0.433**
green shirt (34)	**0.687**	0.594	0.619
short sleeves (220)	0.631	0.608	**0.643**
red trousers/skirt (16)	0.713	0.638	**0.916**
black trousers/skirt (12)	0.683	0.607	**0.711**
white/light gray trousers/skirt (81)	**0.758**	0.639	0.635
blue/light blue trousers/skirt (175)	**0.641**	0.622	0.620
short trousers/skirt (82)	0.416	0.393	**0.557**

Table 3. Average break-even point attained using the considered descriptors.

of people wearing a red shirt should exhibit a high similarity to prototypes P_8 and P_{10}, while a high similarity to P_3 can be expected for a white shirt.

Following the above intuition, a possible approach to perform appearance-based people search through an existing appearance descriptor, is to: (i) identify a set $\mathcal{Q} = \{\mathbf{Q}_1, \mathbf{Q}_2, \ldots\}$ of clothing characteristics that can be detected by the given descriptor, named *basic queries* (for instance, if the descriptors separates lower and upper body parts, and uses colour features, one basic query can be "red trousers/skirt".); (ii) construct a detector for each basic query $\mathbf{Q}_i$, using dissimilarity values as *features* of a supervised classification problem.

Complex queries can be built by connecting basic ones through Boolean operators, e.g., "red shirt AND (blue trousers OR black trousers)". Given a set of images, those relevant to a complex query can simply be found by combining accordingly the subsets of images found by each basic detector.

We point out that the above approach for building detectors is independent of the original appearance descriptor.

4.1 Experimental results

We evaluated our people search approach using two different descriptors previously proposed for person re-identification. The first descriptor is MCMimpl [8]. The second is the SDALF descriptor proposed in [3]. We also used a third descriptor, that is a variation of the first one: it uses a pictorial structure [1] to subdivide body into nine parts: arms and legs (upper and lower, left and right), and torso. The corresponding implementations of our people search method are denoted respectively as MCD_1, MCD_2 and MCD_3.

Experiments were carried out using the VIPER data set [4]. We labelled a subset of 512 images with 14 different clothing characteristics. They are reported in Table 3, and the corresponding number of relevant images is shown between

Fig. 7. Top-ten images retrieved by MCD_1, for ”red shirt” (top) and ”short sleeves” queries, sorted from left to right for decreasing relevance score. A single non-relevant image is present (highlighted in red).

brackets. We evaluated the retrieval performance of our approach on each basic query, for each considered descriptor, using the precision-recall (P-R) curve. For each basic query we randomly subdivided the data set into a training and a testing sets of equal size, and trained an SVM classifier on training images to implement a detector. The P-R curve was evaluated on testing images by varying the SVM decision threshold. This procedure was repeated ten times, and the resulting P-R curves were averaged. The performance on each basic query is summarised in Table 3 as the corresponding average break-even point (BEP), which is the point of the P-R curve whose precision equals recall. The best performance for each basic query is highlighted in bold. An example of the ten top-ranked images for two basic queries is shown in Fig. 7. Our method attained a good performance with all descriptors.

We also developed a simple prototype that enables an operator to retrieve images from a given data set, building the query using a pre-defined set of basic queries. A screen-shot showing the GUI of this prototype is reported in Fig. 8. The lower-left box allows one to retrieve images from a given dataset, by building a complex query through a combination of the concepts in Table 3.

5 Conclusions and future works

In this paper, we presented the results attained by the Pattern Recognition and Applications (PRA) Group in the field of people re-identification. Concerning future developments, we are mainly interested on two research directions. First, to use the dissimilarity-based representation of our MCD framework as a general way to accumulate information from video sequences using any existing appearance-based descriptor. In this case, different dissimilarity vectors coming from different frames is available for each person. These vectors could be used as a training set to train a statistical classifier able to recognise that person. Second, to further develop our work on appearance-based people search, eventually realising a deployable software that can be used by forensic investigators in real-world applications.

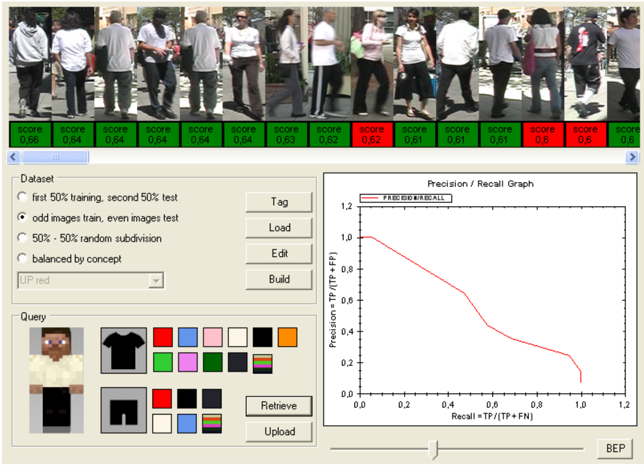


Fig. 8. A screenshot of the demo application for appearance-based people search.

References

1. Andriluka, M., Roth, S., Schiele, B.: Pictorial structures revisited: People detection and articulated pose estimation. In: Proc. IEEE CVPR. pp. 1014–1021 (2009)
2. Doretto, G., Sebastian, T., Tu, P., Rittscher, J.: Appearance-based person reidentification in camera networks: problem overview and current approaches. Journal of Ambient Intelligence and Humanized Computing 2, 127–151 (2011)
3. Farenzena, M., Bazzani, L., Perina, A., Murino, V., Cristani, M.: Person re-identification by symmetry-driven accumulation of local features. In: Proc. IEEE CVPR. pp. 2360–2367 (2010)
4. Gray, D., Brennan, S., Tao, H.: Evaluating appearance models for recognition, reacquisition, and tracking. In: Proc. IEEE Int. Workshop PETS. pp. 41–47 (2007)
5. Pekalska, E., Duin, R.P.W.: The Dissimilarity Representation for Pattern Recognition: Foundations And Applications World Scientific Publishing Co., Inc., (2005)
6. Satta, R., Fumera, G., Roli, F.: Exploiting dissimilarity representations for person re-identification. In: Proc. Int. Workshop SIMBAD. pp. 275–289 (2011)
7. Satta, R., Fumera, G., Roli, F.: Fast person re-identification based on dissimilarity representations. Pattern Recognition Letters, Special Issue on Novel Patt. Rec.-Based Methods for Reidentification in Biometric Context (2012, in press)
8. Satta, R., Fumera, G., Roli, F., Cristani, M., Murino, V.: A multiple component matching framework for person re-identification. In: Proc. ICIAP. vol. 2, pp. 140–149 (2011)
9. Thornton, J., Baran-Gale, J., Butler, D., Chan, M., Zwahlen, H.: Person attribute search for large-area video surveillance. In: Proc. IEEE Int. Conf. HST. pp. 55–61 (2011)
10. Vaquero, D., Feris, R., Tran, D., Brown, L., Hampapur, A., Turk, M.: Attribute-based people search in surveillance environments. In: Proc. IEEE WACV (2009)
11. Wang, J., Zucker, J.D.: Solving the multiple-instance problem: A lazy learning approach. In: Proc. ICML (2000)
12. Zheng, W.S., Gong, S., Xiang, T.: Associating groups of people. In: Proc. BMVC (2009)