



Facial Re-identification on Non-overlapping Cameras and in Uncontrolled Environments

Everardo Santiago Ramírez¹(✉) , J. C. Acosta-Guadarrama¹ ,
Jose Manuel Mejía Muñoz¹ , Josue Dominguez Guerrero¹ ,
and J. A. Gonzalez-Fraga²

¹ Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez,
Ciudad Juárez, Mexico

{`everardo.santiago`, `juan.acosta`, `jose.mejia`, `josue.dominguez`}@uacj.mx

² Facultad de Ciencias, Universidad Autónoma de Baja California,
Ensenada, Mexico

`angel.fraga@uabc.edu.mx`

Abstract. Face re-identification is an essential task in automatic video surveillance where the identity of the person is known previously. It aims to verify if other cameras have observed a specific face detected by a camera. However, this is a challenging task because of the reduced resolution, and changes in lighting and background available in surveillance video sequences. Furthermore, the face to get re-identified suffers changes in appearance due to expression, pose, and scale. Algorithms need robust descriptors to perform re-identification under these challenging conditions. Among various types of approaches available, correlation filters have properties that can be exploited to achieve a successful re-identification. Our proposal makes use of this approach to exploit both the shape and content of more representative facial images captured by a camera in a field of view. The resulting correlation filters can characterize the face of a person in a field of view; they are good at discriminating faces of different people, tolerant to variable illumination and slight variations in the rotation (in/out of plane) and scale. Further, they allow identifying a person from the first time that has appeared in the camera network. Matching the correlation filters generated in the field of views allows establishing a correspondence between the faces of the same person viewed by different cameras. These results show that facial re-identification under real-world surveillance conditions and biometric context can be successfully performed using correlation filters adequately designed.

Keywords: Face re-identification and recognition · Biometrics · Correlation filters

1 Introduction

The goal of person re-identification is to determine if disjoint field-of-views have observed a specific person already detected in another field-of-view (FoV) [9, 28]. It can get performed after the events by processing stored videos and (or) images. That is useful, for example, to know where a person has been in previous time. On the other hand, the re-identification can be performed in real time, i.e., while the person crosses a camera network that belongs to a video surveillance system.

Person re-identification (Re-Id) is a primordial task for many critical applications, such as public safety, multi-camera tracking, and legal forensic search [9, 28]. However, it is an inherently challenging task because of the visual appearance of a person may change dramatically in camera views from different locations due to unknown changes in human pose, viewpoint, illumination, occlusion, cloth, and background clutter [10, 28].

Re-identification task requires person detection, tracking, synthesis, and matching task on digital images. In this sense, computer vision is a feasible technology as it can provide the image of a person in a discrete, remote, and non-intrusive way. The synthesis and matching are the most crucial task for a successful re-identification. Different views of the person in an FoV must be synthesized to generate the descriptor while matching task must compare the descriptors to determinate correspondences.

Descriptors can be generated using cues such as face [19, 23]; the full image of the person’s body [13, 20]; walking pattern (gait) [11]; height and build [5]; and head, torso, and limbs of a person [25]. Although the visual appearance of the whole body is the most exploited in the person re-identification, the face of the person allows a discreet, remote, and non-intrusive re-identification. Furthermore, the face allows a stable re-identification by a long time slot. For this reason, in this work, we propose to perform the re-identification by using correlation filters that synthesize face images detected in the FoVs.

The rest of the paper is organized as follows. Section 2 describes some related works; Sect. 3 presents the basics of correlation filters; Sect. 4 describes the proposed algorithm; Sect. 5 presents the results of experimental evaluation; and finally, the main conclusions of this work.

2 Related Work

Existing approaches for re-identification mainly focus on developing robust descriptors to capture the invariant appearance of a person’s face in different camera views. For this purpose, it can be exploited from the face either the full image or a set of features extracted from it.

A significant challenge in re-identification is class imbalance. In some FoVs, the number of faces captured from a person-of-interest gets outnumbered by the face images of other people. Two-class classification systems designed using imbalanced data tend to recognize better the class with the most significant number of samples. A learning algorithm named Progressive Boosting (PBoost)

can address this problem, which progressively inserts uncorrelated samples into a Boosting procedure to avoid losing information while generating a diverse pool of classifiers [17].

The face images captured in an FoV could appear in arbitrary poses, resolutions in different lighting conditions, noise and blurriness. That problem got addressed with a Dynamic Bayesian Network (DBN) that incorporates the information from different cameras, proposed in [1]. In [3], given a set of face images captured under realistic surveillance scenarios, they train a Support Vector Machine (SVM) as the descriptor of the person. In [19], the authors proposed to overcome the variability of facial images by embedding a distance metric learning into set-based image matching. The distance metric learning scheme learns a feature-space by mapping to a discriminative subspace. Person identification in video data is naturally performed on tracks, not on individual frames, to average out noise and errors under the assumption that all frames belong to the same person [4]. In [7] they propose a framework based on Convolutional Neural Network (CNN). First, a set of attributes is learned by a CNN to model attribute patterns related to different body parts, and whose output gets fused with low-level robust Local Maximal Occurrence (LOMO) features to address the problem of the large variety of visual appearance.

The person re-identification systems are computationally expensive. In order to obtain results in real-time, powerful computing equipment is a must in the implementation of existing approaches. That allows systems to re-identify a person from hundreds, even thousands, of potential candidates in a matter of seconds.

As can be seen in the related work, the approaches for re-identification focus their effort on characterizing the facial variations into a descriptor. However, no effective re-identification exists in uncontrollable environments. That raises the need to develop new strategies in order to obtain a robust re-identification.

3 Correlation Filters

Correlation pattern recognition is based on selecting or defining a reference signal $h(x, y)$, called a correlation filter, and then on determining the degree of similarity between the reference and test signals [18]. Correlation filters can be designed in either the spatial domain or frequency domain by Fourier Transform (FT). The correlation process using the FT is given by:

$$g(x, y) = \mathcal{F}^{-1}\{F(k, l) \cdot H^*(k, l)\}, \quad (1)$$

where $g(x, y)$ is the correlation output; \mathcal{F}^{-1} is the inverse of FT; $F(k, l)$ and $H(k, l)$ are the FT of the test signal $f(x, y)$ and reference signal $h(x, y)$, respectively; “ \cdot ” is an element-wise multiplication; and “ $*$ ” represents the complex conjugate operation. The correlation output $g(x, y)$ should exhibit a brightness, named correlation peak, when $f(x, y)$ is similar to $h(x, y)$. This process is depicted by Fig. 1a.

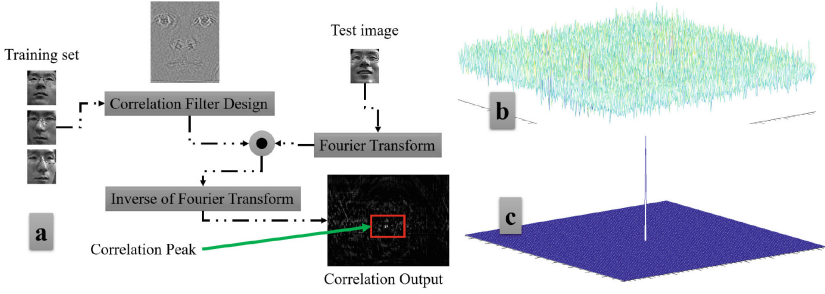


Fig. 1. Scheme of the correlation process in the Fourier domain.

The sharpness of the correlation peak indicates a similarity degree between the test and reference signals. A correlation output for signals of the same class could contain a sharp high peak, as in Fig. 1c; while in another case, the correlation output only contains noise and no prominent peak, as in Fig. 1b. The peak-to-sidelobe ratio (psr) can measure the sharpness of the peak:

$$psr = \frac{(peak - \mu_{area})}{\sigma_{area}}, \quad (2)$$

where μ_{area} and σ_{area} , respectively, are the mean and standard deviation of some area or neighborhood around, but not including, the correlation peak. Consider th as a recognition threshold; thus $psr \geq th$ means that $f(x, y)$ and $h(x, y)$ belong to the same class.

3.1 Synthetic Discriminant Function

A Synthetic Discriminant Function (SDF) filter is a linear combination of MFs [6]. They design this filter by using a training set, T , composed of images with general distortions expected for the object of interest, and it is, therefore, robust at recognizing an object that presents distortions similar to those found in T .

Let $T = \{f_1(x, y), f_2(x, y), \dots, f_N(x, y)\}$ be the training set, and \mathbf{x}_j the column-vector form of $f_j(x, y)$. The vector \mathbf{x}_j is constructed by lexicographic scanning, in which each image is scanned from left to right and from top to bottom. Each vector is a column of the training data matrix, $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$. The SDF correlation filter is given by:

$$\mathbf{h} = X(X^+X)^{-1}\mathbf{u}, \quad (3)$$

where X^+ is the complex conjugate transpose, and $\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]^+$ is a vector of size N that contains the expected values at the origin of the correlation output for each training image. Typical values for \mathbf{u} are 1 for *true* class, and 0 for *false* class.

3.2 Unconstrained Optimal Tradeoff Synthetic Discriminant Function

The Unconstrained Optimal Trade-off Synthetic Discriminant Function (UOTSDF) filter was designed to produce sharp, high peaks in the presence of noise and low-light conditions. This filter is given by:

$$h = (\alpha D + \sqrt{-1\alpha})^{-1}, \quad (4)$$

where α is the normalizing factor, and $D = \frac{1}{N \cdot d} \sum_{i=1}^N (X_i X_i^*)$. The symbols d and X are the amount of pixels in $f_i(x, y)$ and FT of $f_i(x, y)$.

3.3 Maximum Average Correlation Height

Maximum Average Correlation Height (MACH) filter is designed to be tolerant to noise and distortion while addressing the issue of sensor noise and background clutter. MACH filter is given by:

$$h = \alpha(S + C)^{-1}m, \quad (5)$$

where α is a normalizing coefficient,

$$S = \frac{1}{N \cdot d} \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})^*, \quad (6)$$

where X_i is FT of the i -th training image, C is the covariance matrix of the input noise, and m is a vector containing the average of the training images.

In recent years, correlation filters have achieved impressive results in discrimination, efficiency, location accuracy, and robustness [22].

Furthermore, correlation filters could be designed for reliable recognition of partially occluded objects as is described in [14] and [15]. The SDF, UOTSDF, and MACH correlation filters described in this section were used to test the proposed algorithm for re-identifying persons by their facial images captured in disjoints FoVs.

4 Correlation Filters Based Algorithm for Facial Re-identification

This section describes the proposed algorithm for re-identifying a person by their face images by using correlation filters. Since it is all about a video-based person, re-identification must perform in a discreet remote non-intrusive way; the face recognition process makes that possible and allows the proposed algorithm to work in a biometric context.

4.1 Framework for Facial Re-identification

Here is described the proposed framework for facial re-identification in an unconstrained environment. Let $G(V, E)$ be a graph that represents a camera network, where V is the set of FoVs, and E is the set of edges between them. To re-identify a person by their face image on this camera network, necessarily different tasks must work in collaboration: the face detection, identification, tracking, synthesis, and matching.

Figure 2 shows the correlation-filters-based algorithm for re-identifying a person by their face images. A facial descriptor gets synthesized from a set of face images captured in the FoVs, where the person gets last seen. Then, the facial descriptors will correlate (match) for determining correspondence. As one can see in Fig. 2, the person gets identified by using the set of face images, recorded in the FoV, where seen for the first time.

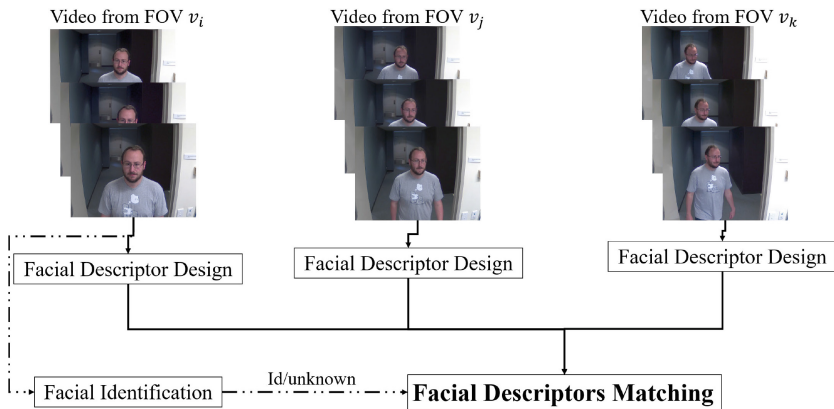


Fig. 2. Basic scheme of proposed algorithm for re-identifying persons by their face images.

The facial identification, synthesis, and matching modules details are in the next subsections.

4.2 Facial Identification

In the traditional face recognition task, only a few samples per subject are necessary to identify face images [2, 8]. However, face images in uncontrollable environments are usually accompanied with appearance changes in lighting, pose, expression, resolution, and occlusion. As a result, the person gets identified by using only the most representative face images, i.e., using only the images that contribute to their recognition.

The proposed identification process is depicted in Fig. 3a. Before the identification, we must build a gallery of biometric templates G_{bt} .

Given a sequence S recorded in the i -th FoV, and where appears the person-of-interest, p ; each frame is searched for face images to generate the set $F_{p,i}$. This set could contain face images with different distortions, and not all are suitable for using them to generate the biometric template. That raises the need for selecting only the most representative face images from $F_{p,i}$. That is why we used the approach proposed in [16]. The selected set must contain face images that best describe the set $F_{p,i}$. Images with variable illumination and noise may not be considered in the training set since image processing algorithms can remove them successfully.

A UOTSDF filter is synthesized with the set selected from $F_{p,i}$ and used as the biometric template $H(k, l)$ for p . For each known person, a biometric template must get synthesized and stored in G_{bt} . Once the gallery G_{bt} is ready, the identification works as follows. Face images $F_{q,j}$ of person q , detected in the j -th FoV, are correlated with the biometric template $H_i(k, l) \in G_{bt}$ for calculating the average similarity score \overline{psr} . The person-of-interest's ID corresponds to the biometric template with the highest average similarity score.

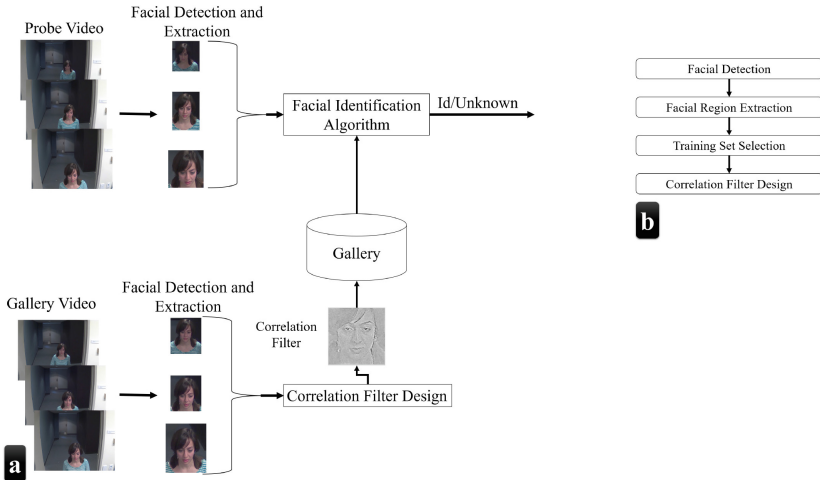


Fig. 3. Facial identification scheme for the proposed re-identification algorithm.

The gallery of biometric templates is generated once and updated eventually to adapt to changes of the face, while the identification performs each time that a person gets detected for the first time. The filters of the identification module are different from those used in the re-identification that we describe below.

4.3 Facial Synthesis

A robust facial descriptor must allow a re-identification either inter/intra FoVs and short/long intervals of time. The algorithm for synthesizing correlation filter

as a facial descriptor in the i -th FoV is in Fig. 3b. First, the most frontal faces in each video frame get detected. Second, the region of the detected faces must get extracted from the video frames. Let $F_{p,i} = \{f_1(x, y), \dots, f_N(x, y)\}$ be a set of those face images. Third, face images in $F_{p,v}$ contain variations concerning illumination conditions, pose, sharpness, as well as size because the person moves away or approaches to the camera. Using all these images may lower the performance of re-identification algorithms, and it is computationally expensive. Therefore, it is more appropriate to choose a subset $T_{p,i} \subset F_{p,i}$ that best describes the faces in i . Fourth, $T_{p,i} \subset F_{p,i}$ is used to train a UOTSDF filter $H_{p,i}(k, l)$, which in turn works as a descriptor of the person. Any correlation filter described in Sect. 3 or another reported in the literature could be suitable for this purpose. This module generates a filter in each FoV, where a person gets viewed and use them in the matching module.

4.4 Descriptors Matching

Descriptors matching could be performed in two ways: (a) once a person has left the camera network, and (b) while a person walks through the network of cameras. In this work, the descriptors matching task runs while the person walks through the camera network for real-time face Re-Id, as depicted by Algorithm 1.

Given the sequence S of the current FoV, a facial descriptor $H_{i-1}(k, l)$ generated in a previous FoV, a recognition threshold th , and a binary vector of matching results; the algorithm works as follows. First, build a facial descriptor $H_i(k, l)$ for faces in S , as described in Sect. 4.3. Later, calculate the psr of the correlation between $H_i(k, l)$ and $H_{i-1}(k, l)$, according to (2). Then, put 1 to vector \mathbf{c} if $psr_{i,j} \geq th$, which means that there is a correspondence. Otherwise, put 0 to \mathbf{c} . Finally, return $H_i(k, l)$ and \mathbf{c} .

Algorithm 1. Facial descriptors matching.

Data: $video_i, H_{i-1}(k, l), th, \mathbf{c}$
Result: $H_i(k, l), \mathbf{c}$

- 1 $H_i(k, l) = BuildFacialDescriptor(video_i)$
- 2 $psr_{i,j} = AverageSimilarity(H_i(k, l), H_{i-1}(k, l))$
- 3 **if** $psr_{i,j} \geq th$ **then**
- 4 $\mathbf{c} = [\mathbf{c}, 1]$
- 5 **else**
- 6 $\mathbf{c} = [\mathbf{c}, 0]$
- 7 **end**
- 8 **return** $H_i(k, l), \mathbf{c}$

Once the person leaves the camera network, the binary vector \mathbf{c} contains all the matches. If all values in \mathbf{c} are 1, then it means that the compared facial descriptor belongs to the same person identified in the first FoV. We performed experiments under real-world surveillance conditions to verify the proposed algorithm.

5 Experimental Evaluation

This section presents the performance of the proposed correlation-filter-based algorithm for re-identifying a person by their face images regarding Detection and Identification Rate (DIR), Verification Rate (VR), Mean Reciprocal Rank (MRR), and the correctness of the descriptors matching. In a first experiment, we tested the facial identification module in DIR, VR and MRR metrics by using UOTSDF filter. In a second experiment, the facial descriptors matching module got tested by using UOTSDF, SDF, and MACH correlation filters. In this section, we compare the performances for correlation filters and LBP images [27] at face re-identification. We used the *psr* measure as the similarity score in correlation filters, while for LBP images we used the Structural Similarity (SSIM) in [24]. Both in identification and correspondence modules, the subsets were selected by using the approach proposed in [16].

5.1 Data Set Configuration

The experimental evaluation of the proposed algorithm got conducted on the publicly available ChokePoint¹ video dataset. This dataset consists of 25 subjects (19 males and six females) in Portal 1 and 29 subjects (23 males and six females) in Portal 2 such as described in [26]. The recording of Portal 1 and Portal 2 are one month apart. The dataset has a frame rate of 30 fps, and the image resolution is 800×600 pixels.

Three case studies (see Table 1) were used to evaluate the performance of the proposed algorithm under different environments and time interval between recording. These case studies contain the most frontal faces, and they are from the ChokePoint website. Case Study 1 was recorded in a short time interval and indoor scene, while Case Study 2 got recorded both indoor and outdoor scenes and with short time interval as well. Similarly, Case Study 3 got recorded in both indoor and outdoor scenes although with a long time interval.

Face images in ChokePoint video dataset have variations regarding illumination conditions, pose, sharpness, as well as scale since the person can approach or move away from the camera. For evaluation purpose, we re-sized the face images to 32 by 32 pixels, and we averaged ten executions to achieve the results.

5.2 Numerical Results

Suppose that an operator visually recognizes the person on the video and thus it is only necessary to verify the person's identity. Now suppose that the operator can not visually identify the person, so the algorithm has to search the gallery for the identity of such person. The performance in the first case gets measured by VR metric, while in the second case by the DIR metric.

Table 2 shows the performance of the facial recognition module. Regarding DIR and MRR metrics, the correlation filters obtained the best performance of

¹ <http://arma.sourceforge.net/chokepoint/>.

Table 1. Case studies used in the experimental evaluation.

Case study 1	Case study 2	Case study 3
P1E_S1_C1	P2E_S2_C2	P1E_S1_C1 P1E_S3_C3
P1L_S1_C1	P2L_S2_C2	P1E_S2_C2 P1E_S4_C1
P1E_S3_C3	P2E_S4_C2	P1L_S1_C1 P1L_S3_C3
P1L_S3_C3	P2L_S4_C2	P1L_S2_C2 P1L_S4_C1
P1E_S2_C2	P2E_S1_C3	P2E_S2_C2 P2E_S4_C2
P1L_S2_C2	P2L_S1_C1	P2E_S1_C3 P2E_S3_C1
P1E_S4_C1	P2E_S3_C1	P2L_S2_C2 P2L_S4_C2
P1L_S4_C1	P2L_S3_C3	P2L_S1_C1 P2L_S3_C3

Table 2. Performance of the facial identification module using correlation filters and LBP images.

Case study	Descriptor	DIR	VR	MRR
1	CF	0.85	0.99	0.91
	LBPi	0.89	1.00	0.93
2	CF	0.86	0.96	0.89
	LBPi	0.80	0.99	0.86
3	CF	0.81	0.97	0.91
	LBPi	0.79	1.00	0.85

both indoor/outdoor scenarios and short/long time interval (Case Studies 2 and 3). These results are evidence that the correlations filters are suitable for facial identification under re-identification conditions. The LBP images obtained the best performance under indoor scenarios and in short time interval–Case Study 1. Moreover, regarding VR metric, LBP images obtained the best performance in three case studies of person re-identification.

In the identification task, each element of the set $F_{p,i}$ gets compared against $N = |G_{bt}|$ biometric templates. This process could produce a list of possible matches. Thus, the biometric template must be able to produce a similarity score higher than other possible matches for the authentic person. The MRR metric measures the performance of the correlation filter to produce a similarity score higher than other possible matches.

The performance of facial descriptors matching module was calculated using the UOTSDF, SDF, and MACH filters. We measured the ability of the descriptor generated in the i -th FoV to match with the descriptor generated in the j -th FoV. As one can see in Table 1, each person appears in eight, eight, and 16 FoVs in the case studies 1, 2, and 3, respectively. Thus, $M = |V_p|$ becomes the number of FoVs, where the p -th person appears, so are performed $M - 1$ comparisons.

The performance of the descriptor matching module was calculated in terms of the number of successful matches between the total numbers of matches. We executed a total of seven matches for case studies 1 and 2, while 15 for case study 3. Both correlation filters and LBP images obtained 100% of effectiveness, which shows the competitiveness of the correlation filters against other approaches in face Re-Id.

5.3 Discussion

Although LBP image obtained similar performance to correlation filters in some cases, it does not allow to synthesize a set of images into a single signal. So, both recognition and matching are computationally expensive and thus not suitable for real-time systems. Furthermore, LBP image requires a big set of images for obtaining good performance. On the other hand, correlation filters can synthesize several images into a single signal. This synthesis improves the speed of the proposed algorithm for re-identifying a person by their facial images.

The results achieved by our proposed algorithm in the recognition module are competitive with those reported in the literature. In [26] they tested the LBP+ Multi-Region Histogram (MRH) and Average MRH, obtaining $VR = 86.7\%$ and $VR = 87.7\%$, respectively, on the ChokePoint dataset. Our proposed method outperforms these results in all case studies. In [12] they proposed a covariance descriptor based on bio-inspired features, getting $VR = 87\%$ in the best case; although that result is on a different dataset, our proposed algorithm achieved better results in the same metric. The descriptor matching module achieved results competitive with those approaches that use facial images for re-identification, reported in [17] and [21, 23].

6 Conclusion

In this paper, we propose a correlation-filter-based algorithm for re-identifying people by their face images. We tested the proposed algorithm with the UOTSDF filter in the facial identification module, achieving the best performance in the DIR metric for indoor/outdoor scenes. We assessed the matching module by using UOTSDF, SDF, and MATCH filters. Although they all scored perfect matches, we observed that the SDF filter does not affect the speed of the proposed algorithm. On the other hand, the UOTSDF and MATCH filters require a matrix inversion, which slows down the re-identification process.

We can highlight at least four strengths of the proposed algorithm. First, correlation filters can be successfully applied to facial re-identification as shown by the results previously presented. Second, since correlation filters synthesize the training sets into a single signal, it helps to maintain or even to improve the speed of re-identification algorithms. Third, the identification task in the re-identification algorithm avoids the need for using an operator and allows the development of intelligent re-identification systems. Fourth, the selection of a subset of the most representative face images, in the identification and matching

modules, improves their precision. Due to these four strengths, the proposed algorithm is suitable for re-identifying a person by their face images in a camera network, and it also promises further results in future work for both intelligent and big-data processes.

References

1. An, L., Kafai, M., Bhanu, B.: Dynamic Bayesian network for unconstrained face recognition in surveillance camera networks. *IEEE J. Emerg. Sel. Top. Circ. Syst.* **3**(2), 155–164 (2013). <https://doi.org/10.1109/JETCAS.2013.2256752>
2. Apicella, A., Isgro, F., Riccio, D.: Improving face recognition in low quality video sequences: single frame vs multi-frame super-resolution. In: Battiato, S., Gallo, G., Schettini, R., Stanco, F. (eds.) *ICIAP 2017*. LNCS, vol. 10484, pp. 637–647. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-68560-1_57
3. Bäuml, M., Bernardin, K., Fischer, M., Ekenel, H.K., Stiefelhagen, R.: Multi-pose face recognition for person retrieval in camera networks. In: *Proceedings—IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2010* (i), pp. 441–447 (2010). <https://doi.org/10.1109/AVSS.2010.42>
4. Bäuml, M., Tapaswi, M., Stiefelhagen, R.: A time pooled track kernel for person identification. In: *Proceedings of the 11th International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*. IEEE, 26–28 August 2014
5. Bedagkar-Gala, A., Shah, S.K.: A survey of approaches and trends in person re-identification. *Image Vis. Comput.* **32**(4), 270–286 (2014). <https://doi.org/10.1016/j.imavis.2014.02.001>
6. Casasent, D., Chang, W.T.: Correlation synthetic discriminant functions. *Appl. Opt.* **25**, 2343–2350 (1986)
7. Chen, Y., Duffner, S., Stoian, A., Dufour, J.Y., Baskurt, A.: Deep and low-level feature based attribute learning for person re-identification. *Image Vis. Comput.* **79**, 25–34 (2018). <https://doi.org/10.1016/j.imavis.2018.09.001>
8. Cui, Z., Chang, H., Shan, S., Ma, B., Chen, X.: Joint sparse representation for video-based face recognition. *Neurocomputing* **135**, 306–312 (2014). <https://doi.org/10.1016/j.neucom.2013.12.004>
9. Gong, S., Cristani, M., Loy, C.C., Hospedales, T.M.: The re-identification challenge. In: Gong, S., Cristani, M., Yan, S., Loy, C.C. (eds.) *Person Re-Identification*. *ACVPR*, pp. 1–20. Springer, London (2014). https://doi.org/10.1007/978-1-4471-6296-4_1
10. Li, W., Zhu, X., Gong, S.: Person re-identification by deep joint learning of multi-loss classification. *CoRR abs/1705.04724* (2017). <http://arxiv.org/abs/1705.04724>
11. Liu, Z., Zhang, Z., Wu, Q., Wang, Y.: Enhancing person re-identification by integrating gait biometric. In: Jawahar, C.V., Shan, S. (eds.) *ACCV 2014*. LNCS, vol. 9008, pp. 35–45. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16628-5_3
12. Ma, B., Su, Y., Jurie, F.: Covariance descriptor based on bio-inspired features for person re-identification and face verification. *Image Vis. Comput.* **32**(6), 379–390 (2014). <https://doi.org/10.1016/j.imavis.2014.04.002>
13. Ren, Y., Li, X., Lu, X.: Feedback mechanism based iterative metric learning for person re-identification. *Pattern Recognit.* **75**, 1339–1351 (2018). <https://doi.org/10.1016/j.patcog.2017.04.012>

14. Ruchay, A., Kober, V., Gonzalez-Fraga, J.A.: Reliable recognition of partially occluded objects with correlation filters. *Math. Probl. Eng.* **2018**, 8 p. (2018). <https://doi.org/10.1155/2018/8284123>. Article ID 8284123
15. Santiago-Ramirez, E., Gonzalez-Fraga, J.A., Lazaro-Martinez, S.: Face recognition and tracking using unconstrained non-linear correlation filters. In: *International Meeting of Electrical Engineering Research ENIINVIE* (2012)
16. Santiago-Ramirez, E., Gonzalez-Fraga, J.A., Gutierrez, E., Alvarez-Xochihua, O.: Optimization-based methodology for training set selection to synthesize composite correlation filters for face recognition. *Signal Process.: Image Commun.* **43**, 54–67 (2016). <https://doi.org/10.1016/j.image.2016.02.002>
17. Soleymani, R., Granger, E., Fumera, G.: Progressive boosting for class imbalance and its application to face re-identification. *Expert. Syst. Appl.* **101**, 271–291 (2018). <https://doi.org/10.1016/j.eswa.2018.01.023>
18. Vijaya-Kumar, B.V.K., Mahalanobis, A., Juday, R.: *Correlation Pattern Recognition*. Cambridge University Press, Cambridge (2005)
19. Wang, G., Zheng, F., Shi, C., Xue, J.H., Liu, C., He, L.: Embedding metric learning into set-based face recognition for video surveillance. *Neurocomputing* **151**(P3), 1500–1506 (2015). <https://doi.org/10.1016/j.neucom.2014.10.032>
20. Wang, J., Zhou, S., Wang, J., Hou, Q.: Deep ranking model by large adaptive margin learning for person re-identification. *Pattern Recognit.* **74**, 241–252 (2017). <https://doi.org/10.1016/j.patcog.2017.09.024>
21. Wang, J., Wang, Z., Liang, C., Gao, C., Sang, N.: Equidistance constrained metric learning for person re-identification. *Pattern Recognit.* **74**, 38–51 (2018). <https://doi.org/10.1016/j.patcog.2017.09.014>
22. Wang, Q., Alfalou, A., Brosseau, C.: New perspectives in face correlation research: a tutorial. *Adv. Opt. Photon.* **9**(1), 1–78 (2017). <https://doi.org/10.1364/AOP.9.000001>
23. Wang, Y., Shen, J., Petridis, S., Pantic, M.: A real-time and unsupervised face re-identification system for human-robot interaction. *Pattern Recognit. Lett.* (2018). <https://doi.org/10.1016/j.patrec.2018.04.009>
24. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004). <https://doi.org/10.1109/TIP.2003.819861>
25. Watson, G., Bhalerao, A.: Person re-identification using deep foreground appearance modeling. *J. Electron. Imaging* **27** (2018). <https://doi.org/10.1117/1.JEI.27.5.051215>
26. Wong, Y., Chen, S., Mau, S., Sanderson, C., Lovell, B.C.: Patch-based probabilistic image quality assessment for face selection and improved video-based face recognition. In: *IEEE Biometrics Workshop, Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 81–88. IEEE, June 2011
27. Yang, B., Chen, S.: A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image. *Neurocomputing* **120**, 365–379 (2013). <https://doi.org/10.1016/j.neucom.2012.10.032>. *Image Feature Detection and Description*
28. Zheng, L., Yang, Y., Hauptmann, A.G.: Person re-identification: past, present and future. *CoRR* **14**(8), 1–20 (2016)