Multiple-shot Human Re-Identification by Mean Riemannian Covariance Grid

Sławomir Bąk, Etienne Corvee, Francois Brémond, Monique Thonnat
INRIA Sophia Antipolis, PULSAR group
2004, route des Lucioles, BP93
06902 Sophia Antipolis Cedex - France
firstname.surname@inria.fr

Abstract

Human re-identification is defined as a requirement to determine whether a given individual has already appeared over a network of cameras. This problem is particularly hard by significant appearance changes across different camera views. In order to re-identify people a human signature should handle difference in illumination, pose and camera parameters. We propose a new appearance model combining information from multiple images to obtain highly discriminative human signature, called Mean Riemannian Covariance Grid (MRCG). The method is evaluated and compared with the state of the art using benchmark video sequences from the ETHZ and the i-LIDS datasets. We demonstrate that the proposed approach outperforms state of the art methods. Finally, the results of our approach are shown on two other more pertinent datasets.

1. Introduction

Human re-identification is one of the most challenging and important problems in computer vision and pattern recognition. Only knowledge about identities of tracked persons can allow a system to fully understand the scene. The human re-identification problem can be defined as a determination whether a given person of interest has already been observed over a network of cameras. This issue is also called *the person re-identification problem*.

Person re-identification can be considered on different levels depending on information cues which are currently available in the system. Biometrics such as face, iris or gait can be used to recognize identities. Nevertheless, in most video surveillance scenarios such detailed information is not available due to video low-resolution or difficult segmentation (crowded environments, *e.g.* airports, metro stations). Therefore a robust modeling of a global appearance of an individual is necessary to re-identify a given person of interest. In these identification techniques (named *appearance-based approaches*) clothing is the most reliable information about an identity of an individual (there is an assumption that individuals wear the same clothes between different sightings). The model of an appearance has to handle differences in illumination, pose and camera parameters to allow matching appearances of the same individual observed in different cameras.

The main topic of this paper is a novel appearance-based approach which builds a specific human signature model to re-identify a given individual. In our approach a human detection algorithm is used to find out people in video sequences. Then, the detected individual is tracked to gather as many frames as possible. Our method belongs to the group of *multiple-shot* approaches where multiple images of a person are used to extract discriminative signature.

This paper makes the following contributions:

- We propose to use *the Mean Riemannian Covariance* (MRC) matrices blending the appearance information from multiple images. This mean covariance matrix keeps not only information about feature distribution but also carries out essential cues about temporal changes of an appearance (Section 3.2).
- We offer a novel kind of feature, *i.e.*, *the Mean Riemannian Covariance Grid* (MRCG) (Section 3.3). Our idea is to combine efficiency of the mean riemannian covariance descriptor with a spacial information carried out by a dense grid structure.
- We present an efficient method to enhance discriminative features to improve matching accuracy (Section 3.4). The experimental results show that we outperform existing methods without adopting any of complex machine learning schemes such as boosting [1, 11] or RankSVM [17].
- We introduce a new dissimilarity measure between signatures which is able to hold discriminative power coming from the informative dense grid structure of the MRC-s (Section 3.5).

We evaluate our approach in Section 4 before discussing related work and concluding.

Figure 1. The results of the query. The first image on the left is the query image. The true match is on the first position in the list.

2. Problem Definition

We lay the problem as the following. We generate human signature for each person detected and tracked in our video surveillance system. Let us denote a signature as $\mathfrak{s}_i^c$, where i encodes the person identity and c denotes the camera. The task is to find for each signature its corresponding signature in another camera. It is realized by querying the database of signatures $\mathfrak{s}_j^{c'}$, where $c \neq c'$ with signature of interest $\mathfrak{s}_i^c$. The results of the query is the list of the most similar signatures ordered by increasing dissimilarity (see Fig. 1). The position in the list of the true match is called the rank score.

3. Human Appearance Model

In this section we propose a new appearance model based on the MRC matrices extracted from tracks of a specific individual. The input of our approach is a set of cropped images corresponding to human detection and tracking results. We handle color dissimilarities caused by camera illumination difference by applying the histogram equalization [12] to each of the color channels (RGB). Then, such normalized image is divided into a grid structure of overlapping cells. Nevertheless, before explaining details concerning cells of the grid, we present a brief overview of *the covariance descriptor*.

3.1. Covariance descriptor

In [19] the covariance of d-features has been proposed to characterize a region of interest. The descriptor encodes information of the variances of the defined features inside the region, their correlations with each other and a spatial layout. It is shown that the performance of the covariance features is superior to other methods as rotations and illuminations changes are absorbed by the covariance matrix.

Covariance matrix as a positive definite and symmetric matrix can be seen as a tensor. The main problem is that such defined tensor space is a manifold that is not a vector space with the usual additive structure (do not lie on Euclidean space). Hence, many usual operations (like the mean or the distance) need a special treatment. Therefore, our covariance manifold is specified as Riemannian to determine a powerful framework using tools from differential geometry [16]. We use the distance definition proposed by [7] to compute the dissimilarity between two covariance matrices C_i and C_j

$$\rho(C_i, C_j) = \sqrt{\sum_{k=1}^{d} \ln^2 \lambda_k(C_i, C_j)} \tag{1}$$

where $\lambda_k(C_i, C_j)_{k=1\ldots d}$ are the generalized eigenvalues of C_i and C_j, determined by

$$\lambda_k C_i x_k - C_j x_k = 0, \quad k = 1 \ldots d \tag{2}$$

and $x_k \neq 0$ are the generalized eigenvectors.

3.2. Mean Riemannian Covariance (MRC)

Let $C_1, \ldots, C_N$ be a set of covariance matrices. The Karcher or Fréchet mean is the set of tensors minimizing the sum of squared distances. In the case of tensors, the manifold has a non-positive curvature, so there is a unique mean value μ

$$\mu = arg \min_{C \in \mathcal{M}} \sum_{i=1}^{N} \rho^2(C, C_i). \tag{3}$$

where ρ is the covariance matrix distance (Eq. 1).

Since covariance matrices lay in a Riemannian manifold we use the intrinsic Newton gradient descent algorithm to compute the approximation mean covariance at step $t+1$

$$\mu_{t+1} = exp_{\mu_t}\left[\frac{1}{N}\sum_{i=1}^{N} log_{\mu_t}(C_i)\right] \tag{4}$$

where exp_{μ_t} and log_{μ_t} are specific operators uniquely defined on the Riemannian manifold. This iterative gradient descent algorithm usually converges very fast (in experiments 5 iterations were enough, which is similar to [16]).

3.3. Mean Riemannian Covariance Grid (MRCG)

In this section we define the novel Mean Riemannian Covariance Grid (MRCG) and explain its merits. The proposed human signature has been designed to deal with low resolutions images and crowded environments where more specialized techniques (*e.g.* based on body parts detectors) might fail. We combine dense descriptors philosophy [4] with extremely effectiveness of the MRC descriptor.

Once color has been normalized, we scale every human image into a fixed size $W \times H$ pixels. Then, an image is divided into a dense grid structure with overlapping spatial square regions (*cells*). First, such dense representation makes the signature robust to partial occlusions. Second, as the grid structure, it contains a relevant information about spatial correlations between the MRC *cells* which is essential to carry out discriminative power of the signature. Moreover, as we use covariance matrices to describe characteristic of the cells, it is an efficient fusion of different types of features and their modalities.

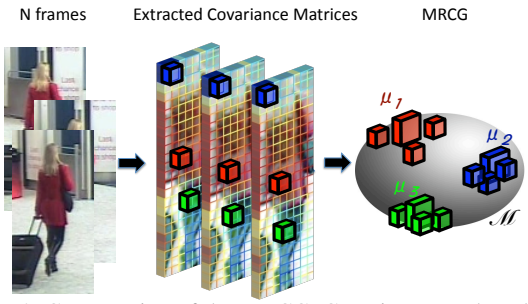


Figure 2. Computation of the MRCG. Covariances gathered from tracking results are used to compute the MRC using Riemannian manifold space (depicted with the surface of the sphere).

Let $C_1^p, \ldots, C_N^p$ be a set of covariance matrices extracted during tracking of N frames corresponding to image square regions at position of the cell p. We define the MRC as the mean covariance of these covariance matrices (see Section 3.2) computed using the Riemannian manifold space (see Fig. 2). The mean covariance matrix as an intrinsic average blends the appearance information from multiple images. This mean covariance matrix keeps not only information about features distribution but also carries out essential cues about temporal changes of the appearance related to the position of the cell p. All MRC *cells* compose a full grid, named as Mean Riemannian Covariance Grid (MRCG). We prove efficiency of MRCG in the experimental results in Section 4.

3.4. MRC Discriminants

The goal of using discriminants is to identify the relevance of MRC *cells*. We present an efficient way to enhance discriminative features to improve matching accuracy.

Given a set of signatures $\mathfrak{S}^c = \{\mathfrak{s}_i^c\}_{i=1}^n$ where s_i^c is a signature i from camera c, MRCG is represented by $\mathfrak{s}_i^c = \{\mu_{i,1}^c, \mu_{i,2}^c, \ldots, \mu_{i,m}^c\}$ and m is the number of *cells* in the grid. For each $\mu_{i,j}^c$ we compute the variance between the human signatures from camera c defined as

$$\sigma_{i,j}^c = \frac{1}{n-1} \sum_{k=1;k\neq i}^{n} \rho^2(\mu_{i,j}^c, \mu_{k,j}^c). \tag{5}$$

Hence for each human signature $\mathfrak{s}_i^c$ we obtain the vector of discriminants related to our MRC *cells*, $d_i^c = \{\sigma_{i,1}^c, \sigma_{i,2}^c, \ldots, \sigma_{i,m}^c\}$. Here the idea is similar to methods derived from text retrieval where the frequency of *terms* is used to weight relevance of a *word*. As we do not want to quantize covariance space, we use $\sigma_{i,j}^c$ of the MRC *cell* to extract its relevance. The MRC is assumed to be more significant when its variance is larger in the class of humans. Here, it is a kind of *"killing two birds with one stone"*: 1) it is obvious that the most common patterns belong to the background (the variance is small); 2) the patterns which are far from the rest are at the same time the most discriminative (the variance is large). We thought about normalizing the $\sigma_{i,j}^c$ by the variance *within the class* (similarly to Fisher's linear discriminants). Nevertheless, the results have shown that such a normalization does not improve matching accuracy. We think this is a consequence that the given number of frames per individual is not enough to obtain the reliable variance of MRC *within the class*.

Scalability: Discriminative approaches are often accused of non-scalability (like [14, 18]). It is true that in these approaches an extensive learning phase is necessary to extract discriminative signatures. This makes these approaches very difficult to apply in real scenario where in every new minute new people appear. Fortunately, our approach by using very simple discriminative method is able to perform in the real system. It is true that every time when a new signature is created we have to update all signatures in the database. Nevertheless, for $10,000$ signatures, the update takes less than 30 seconds. Moreover, we do not expect more than such amount of signatures into database as the re-identification approaches are constraint to *one day period* (the strong assumption about the same clothes).

3.5. Grid Matching

Given the extracted human signatures we introduce a way to effectively distinguish individuals. As already mentioned the human signatures consist of a set of the MRC *cells* structured into a dense grid. In general the matching of two signatures $\mathfrak{s}_A$ and $\mathfrak{s}_B$ is carried out by maximizing the similarity measure. We shift one signature over another in x and y-direction to reduce body alignment issues. When we shift signature we preserve relative position between MRC cells to avoid wasting of discriminative property. In our experiments the signature is shifted over another not more than width of a cell to maximize similarity. The similarity between two human signatures $\mathfrak{s}_A$ and $\mathfrak{s}_B$ is defined as

$$S(\mathfrak{s}_A, \mathfrak{s}_B) = \sum_{i \in K} \frac{\sigma_{A,i} + \sigma_{B,i}}{\rho(\mu_{A,i}, \mu_{B,i})} / \parallel K \parallel \tag{6}$$

where K stands for the set of *cells* in signature $\mathfrak{s}_A$ which have corresponding *cells* in signature $\mathfrak{s}_B$; ρ is the covariance distance; $\sigma_{A,i}$ and $\sigma_{B,i}$ are the discriminants of the corresponding MRC-s.

4. Experimental results

In this section the extensive evaluation of our approach is presented. The performance is shown using the Cumulative Matching Characteristic (CMC) curve suggested in [10] as the validation method for the re-identification problem. The CMC curve represents the expectation of finding the correct match in the top n matches. In order to provide quantitative results for our MRCG, we consider ETHZ [18] and i-LIDS [13, 21] datasets.

Experimental setup: Every human image is scaled into a fixed size of 64×192 pixels (size of the grid). We extract the MRC *cells* of 16×16 pixels, on a fixed grid of 8 pixels step (it gives in total 161 *cells*). Feature vector consist of 11 features:

$$\left[x, y, R_{xy}, G_{xy}, B_{xy}, \nabla^R_{xy}, \theta^R_{xy}, \nabla^G_{xy}, \theta^G_{xy}, \nabla^B_{xy}, \theta^B_{xy}\right] \quad (7)$$

where x and y are pixel location, R_{xy}, G_{xy}, B_{xy} are RGB channel values and ∇ and θ corresponds to gradient magnitude and orientation in each channel, respectively.

4.1. ETHZ dataset

ETHZ dataset was originally used for human detection [5]. In [18] this data have been adjusted for re-identification purposes[1] The modified dataset consists of three sequences: SEQ. #1 contains 83 pedestrians, SEQ. #2 contains 35 pedestrians and SEQ. #3 contains 28 pedestrians. The main drawback of this dataset is that the re-identification is performed in the same camera. Since the human images are very similar we randomly pick up a set of $N = 10$ frames from the beginning and from the end of each sequence to maximize challenging aspects and to be comparable with [3, 6]. The evaluation was repeated 20 times to obtain reliable statistics. We compare our MRCG with HPE [3], PLS [18] and SDALF [6] (see Fig. 3). As we can see, our MRCG obtain the best results in all of the sequences. It shows how well the MRC *cells* can handle appearance variations. As MRCG consist of dense structured grid it is able to handle partial occlusions and small scale changes.

In our belief, despite such challenging aspects as illuminations changes and occlusions, the ETHZ dataset is not challenging enough to evaluate re-identification approaches. One of the most challenging issues in the re-identification problem is due to different camera settings, different color responses, different camera view points and different environments, which is not in this case. Hence, we have also evaluated our approach on images from more challenging i-LIDS (MCTS) dataset.

4.2. i-LIDS datasets

The experiments are performed on images from the 2008 i-LIDS Multiple-Camera Tracking Scenario (MCTS) dataset with multiple camera views. The evaluation dataset contains 476 images with 119 individuals automatically extracted by [21]. This dataset is very challenging since there are many occlusions and often only the top part of the person is visible.

We compared our approach with methods which obtained the best performance on this dataset: SCR [2], HPE [3], Appearance Context [21] and SDALF [6]. As SCR belongs to *single-shot* approaches, we extended SCR to *multiple-shot* approach by applying *"set matching"* (the minimal distance between pair of images). This makes our evaluation more fair to SCR method. The extended SCR method is noted as M-SCR. Unfortunately, i-LIDS [21] dataset does not fit very well for *multiple-shot* signature because the number of images per individual is very low (in average 4). Moreover, for 22 individuals there are only 2 images given (one from each camera). Hence, in evaluation we use maximally $N = 2$ images to create human signature (like in [6]). Then, we applied simple affine transformation on these images (coordination of transformation matrix were changed by 5% and rotation angle was in range of $[-6°; 6°]$) to obtain our MRCG (the power of our descriptor is obtained by an intrinsic average which blends the appearance information).

The results with the state of the art approaches are reported in Fig. 4 (a). Our MRCG outperforms the state of the art. It proves that MRCG is highly informative descriptor which can handle camera differences.

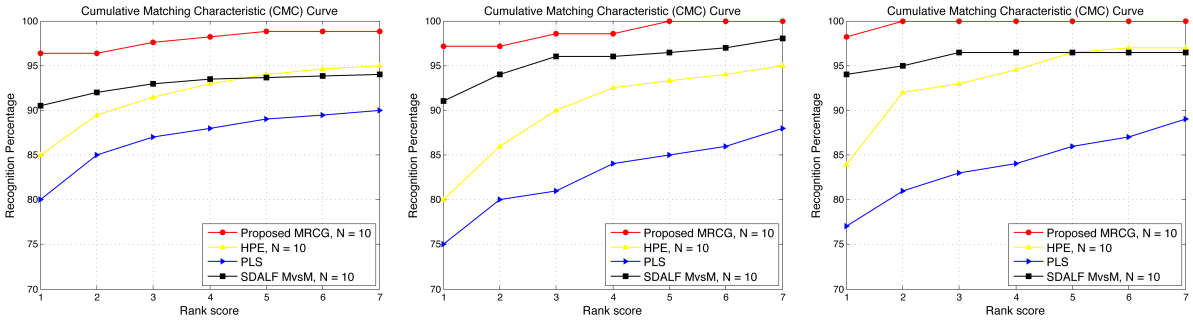
As iLIDS [21] does not fit well for *multiple-shot* signature we decide to evaluate our approach on two new sets of individuals from i-LIDS data [13]. These datasets finally satisfy all requirements of multiple-shot person reidentfication.

i-LIDS-MA [13]. This dataset contains 40 individuals extracted from two cameras. For each individual 46 frames are annotated manually from both cameras. Therefore we have $40 \times 2 \times 46 = 3680$ annotated images. For each pedestrian we create human signature using $N = 1$ (for SCR [2]) or $N = 10$ (for M-SCR and MRCG) randomly selected images. Then, every signature is used as a query to the gallery set of signatures from different camera. The procedures were repeated 10 times and average CMC curves together with our MRCG results are displayed in Fig. 4 (b). MRCG again proves its efficiency. Moreover, in comparison to M-SCR, our MRCG condenses information from the set of frames into compact and highly informative signature. The results show that the MRCG is extremely efficient gathering information using Riemannian manifold space.

As i-LIDS-MA is a manually annotated dataset, it still does not reflect real video surveillance scenario where humans are detected and tracked automatically. Consequently, we use second dataset (i-LIDS-AA [13]) where images of humans are extracted automatically using HOG-based detector. In this case, detection and tracking results are noisy which makes the dataset more challenging.

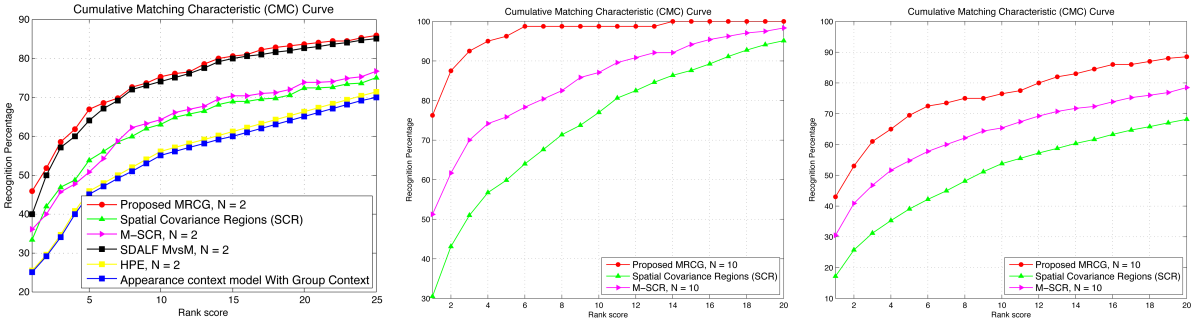
i-LIDS-AA [13]. This dataset contains 100 individuals on 10754 images. The evaluation scheme was the same as for i-LIDS-MA dataset. The performance on this dataset is shown in Fig. 4 (c). The results show again that our descriptors outperform significantly SCR and M-SCR. Nev-

[1]ETHZ Dataset for Appearance-Based Modeling: http://www.umiacs.umd.edu/~schwartz/datasets.html



(a) SEQ. #1 (83 pedestrians) (b) SEQ. #2 (35 pedestrians) (c) SEQ. #3 (28 pedestrians)

Figure 3. CMC curves obtained on ETHZ dataset. Our descriptor is noted as MRCG. We compare our method with the results of HPE [3], PLS [18] and SDALF [6].



(a) i-LIDS [21] (119 pedestrians) (b) i-LIDS-MA (40 pedestrians) (c) i-LIDS-AA (100 pedestrians)

Figure 4. CMC curves obtained on i-LIDS datasets; (a) We compare our methods with the results of SCR [2] and M-SCR, HPE [3], SDALF [6] and Appearance context model with Group Context [21]; (b) and (c) We compare our methods with the results of SCR [2] and M-SCR.

ertheless the performance is not very high in comparison with the results obtained on i-LIDS-MA. It shows one of the main limitations that our approach performance directly depends on human detection results (*e.g.* detected bounding boxes not accurately centered around the people, only part of the people are detected due to occlusion). However, the results show that despite this limitation our descriptor still performs better than the state of the art approaches.

5. Related Work

Recently, the person re-identification problem became one of the most important tasks in video surveillance. There is a natural consequence of an invention of robust human detection algorithms to extend approaches for recognition purposes. The appearance-based re-identification techniques were focused on associating pairs of images, each containing one instance of individual. These methods are named *single-shot* approaches [2, 15, 20] and until now they were the most popular techniques. Currently researches try to improve identification accuracy by integrating information over many images. The group of methods which employs multiple images of the same person as training data is called *multiple-shot* approaches.

As to *single-shot* approaches, in [15] the clothing color histograms taken over the head, shirt and pants regions together with the approximated height of the person were used as the discriminative feature. Similarly, clothing segmentation together with facial features [8] were employed to recognize individuals. Shape and appearance context model is proposed in [20]. A pedestrian image is segmented into regions and their color spatial information is registered into a co-occurance matrix. This method works well if the system considers only a frontal viewpoint. For more challenging cases, where viewpoint invariance is necessary, the ensemble of localized features (ELF) [11] has been proposed. Instead of designing a specific feature for characterizing people appearance, a machine learning algorithm constructs a model that provides maximum discriminability by filtering a set of simple features. Enhancement of discriminative power of each individual signature with respect to the others was also the main issue in [14]. Pairwise dissimilarity profiles between individuals have been learned and adapted into nearest neighbor classification. Similarly, in [18], a rich set of feature descriptors based on color, textures and edges has been used to reduce the amount of ambiguity among human class. The high-dimensional sig-

nature was transformed into a low-dimensional discriminant latent space using a statistical tool called Partial Least Squares (PLS) in one-against-all scheme. Nevertheless in both methods, an extensive learning phase based on the pedestrians to re-identify is necessary to extract discriminative profiles what makes the approaches non-scalable. The person re-identification problem has been reformulated as a ranking problem in [17]. The authors presented extensive evaluation of learning approaches and show that a ranking relevance based model can improve the reliability and accuracy.

Concerning *multiple-shot* approaches, in [9] the spatiotemporal graph was generated for ten consecutive frames for grouping spatiotemporally similar regions. Then, clustering method is applied to capture the local descriptions over time and improve matching accuracy. In [1], the AdaBoost was applied to extract the most discriminative and invariant haar-like features. Here, again one-against-all learning scheme was used to catch human dissimilarities. In [6], the authors proposed to combine three features: 1) chromatic content (HSV histogram); 2) maximally stable colour regions (MSCR) and 3) recurrent highly structured patches (RHSP). The extracted features were weighted by the distance with respect to the vertical axis to minimize effects of pose variations. Recurrent patches were also proposed in [3]. Epitome analysis was used to extract highly informative patches form the set of images.

6. Conclusions

We have proposed a new approach for the human reidentification problem. The extensive evaluation has been performed on the ETHZ and the i-LIDS datasets. It has been shown that the MRCG computed using a Riemannian manifold theory can extract an essential information about appearance of the human and its variability. The experiments prove efficiency of the approach outperforming state of the art accuracy. In the future work we will investigate how to minimize the influence of noisy human detection and tracking on our human signature. Also we are planing to consider 2D/3D body parts modeling to improve matching of different poses of individuals.

Acknowledgements

This work has been supported by Agence National de la Recherche (ANR), VIDEO-ID and VANAHEIM project.

References

[1] S. Bak, E. Corvee, F. Bremond, and M. Thonnat. Person re-identification using haar-based and dcd-based signature. In *2nd Workshop on AMMCSS*, 2010. 1, 6

[2] S. Bak, E. Corvee, F. Bremond, and M. Thonnat. Person re-identification using spatial covariance regions of human body parts. In *AVSS*, 2010. 4, 5

[3] L. Bazzani, M. Cristani, A. Perina, M. Farenzena, and V. Murino. Multiple-shot person re-identification by hpe signature. In *ICPR*, pages 1413–1416, 2010. 4, 5, 6

[4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR*, pages 886–893, 2005. 2

[5] A. Ess, B. Leibe, and L. Van Gool. Depth and appearance for mobile scene analysis. In *ICCV*, pages 1–8, Oct. 2007. 4

[6] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. In *CVPR*, pages 2360–2367, 2010. 4, 5, 6

[7] W. Förstner and B. Moonen. A metric for covariance matrices. In *Quo vadis geodesia ...?, Festschrift for Erik W. Grafarend on the occasion of his 60th birthday, TR Dept. of Geodesy and Geoinformatics, Stuttgart University*, 1999. 2

[8] A. C. Gallagher and T. Chen. Clothing cosegmentation for recognizing people. In *CVPR*, pages 1–8, 2008. 5

[9] N. Gheissari, T. B. Sebastian, and R. Hartley. Person reidentification using spatiotemporal appearance. In *CVPR*, pages 1528–1535, 2006. 6

[10] D. Gray, S. Brennan, and H. Tao. Evaluating Appearance Models for Recognition, Reacquisition, and Tracking. *PETS*, 2007. 3

[11] D. Gray and H. Tao. Viewpoint invariant pedestrian recognition with an ensemble of localized features. In *ECCV*, pages 262–275, 2008. 1, 5

[12] S. D. Hordley, G. D. Finlayson, G. Schaefer, and G. Y. Tian. Illuminant and device invariant colour using histogram equalisation. *Pattern Recognition*, 38, 2005. 2

[13] http://www-sop.inria.fr/members/Francois.Bremond/topicsText/reidentification.html. September 2011. 3, 4

[14] Z. Lin and L. S. Davis. Learning pairwise dissimilarity profiles for appearance recognition in visual surveillance. In *ISVC*, pages 23–34, 2008. 3, 5

[15] U. Park, A. Jain, I. Kitahara, K. Kogure, and N. Hagita. Vise: Visual search engine using multiple networked cameras. In *ICPR*, pages 1204–1207, 2006. 5

[16] X. Pennec, P. Fillard, and N. Ayache. A riemannian framework for tensor computing. *Int. J. Comput. Vision*, 66(1):41–66, 2006. 2

[17] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang. Person re-identification by support vector ranking. In *BMVC*, pages 21.1–21.11, 2010. 1, 6

[18] W. R. Schwartz and L. S. Davis. Learning discriminative appearance-based models using partial least squares. In *SIBGRAPI*, pages 322–329, 2009. 3, 4, 5

[19] O. Tuzel, F. Porikli, and P. Meer. Region covariance: A fast descriptor for detection and classification. In *ECCV*, pages 589–600, 2006. 2

[20] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu. Shape and appearance context modeling. In *ICCV*, pages 1–8, 2007. 5

[21] W.-S. Zheng, S. Gong, and T. Xiang. Associating groups of people. In *BMVC*, London, 2009. 3, 4, 5