Depth-based Person Re-identification

Ancong Wu , Wei-Shi Zheng* , and Jian-Huang Lai
School of Information Science and Technology, Sun Yat-sen University, China
wuancong@mail2.sysu.edu.cn, wszheng@ieee.org, stsljh@mail.sysu.edu.cn

Abstract

Person re-identification aims to match people across non-overlapping camera views. For this purpose, most works exploit appearance cues, assuming that the color of clothes is discriminative in short term. However, when people appear in extreme illumination or change clothes, appearance-based methods tend to fail. Fortunately, depth images provide more invariant body shape and skeleton information regardless of illumination and color, but only a few depth-based methods have been developed so far. In this paper, we propose a covariance-based rotation invariant 3D descriptor called Eigen-depth to describe pedestrian body shape and the property of rotation invariance is proven in theory. It is also insensitive to slight shape change and invariant to color change and background. We combine our descriptor with skeleton-based feature to get a complete representation of human body. The effectiveness is validated on RGBD-ID and BIWI RGBD-ID datasets.

1. Introduction

The task of person re-identification is to match people in a distributed multi-camera surveillance system across different cameras. Most existing methods exploit appearance cues, *e.g.*, color features [7, 18, 16], texture-based features [29, 11, 33, 3], interest points [14], fisher vectors [20] and covariance descriptors [2]. A large amount of metric/subspace models [15, 31, 39, 21, 30, 38] have also been developed. These methods assume color is reliable and discriminative in short term.

However, in some certain situations, appearance cues are not reliable or not discriminative. For example, in short-term applications, lighting condition changes extremely or different people wear the same clothes. In long-term applications, people change clothes in different days. In these cases, appearance-based techniques tend to fail.

Is there anything that stays more invariant even when suffering from clothes change or extreme illumination? The answer is yes. In an indoor environment, Microsoft Kinect can easily obtain depth images and estimate body skeletons regardless of color and illumination, providing shape and physical information. Another advantage is that background of a depth image can be easily segmented, avoiding background influences in appearance-based methods.

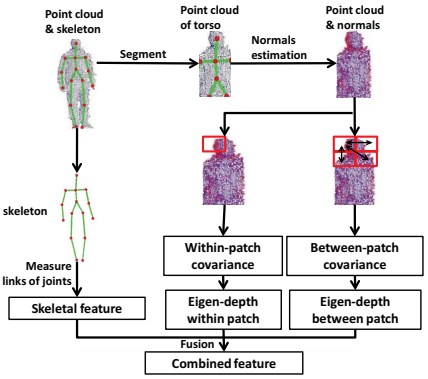


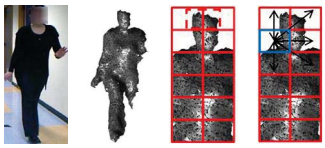
Figure 1. Pipeline of our depth-based re-identification feature extraction.

Recently, some re-identification methods based on RGB-D information have been developed. Existing works can be categorized into two groups. The first kind of methods combine appearance cues and depth information to get more discriminative feature representation [22, 23, 13, 32, 4, 1, 28, 35]. The second kind of methods exploit only depth information without using RGB images[5, 25, 26, 27, 19, 6]. So far, very few methods based on only depth information have been studied. Barbosa *et al*. exploited skeleton-based feature [5], but with the constraint of skeleton noise, the matching accuracy was limited. Munaro *et al*. applied Point Cloud Matching (PCM) [26], and combined it with skeleton-based feature in [25, 27]. PCM exploits raw data of 3D point coordinates without extracting robust features, so it tends to be sensitive to shape change of non-rigid human body.

As for challenges, noise of depth information and shape change of body bring us some difficulties. To address these problems, we propose a rotation invariant 3D descriptor Eigen-depth designed for describing body shape, which is covariance-based and insensitive to slight shape change. Here we explain the differences between Eigen-depth and related 3D descriptors. The descriptor in [8] for RGB-D images is covariance-based but not rotation invari-

*Corresponding author

978-1-4799-6100-9/15/$31.00 ©2015 IEEE



(a) (b) (c) (d)

Figure 2. Illustration of feature extraction region. (a) is sample RGB image, (b) is corresponding point cloud, (c) is within-patch feature extraction region, (d) is between-patch feature extraction region.

ant. RIFT2M [34] and Fehr's covariance descriptor [9] are based on 3D interest points and their rotation invariance relies on local point cloud orientation normalization. In comparison, our Eigen-depth feature is dense, containing more information of body shape. Moreover, rotation invariance is achieved in eigen analysis level. So normalization of point cloud is not needed and eigen analysis helps to get more robust and compact representation.

The pipeline illustrated in Figure 1 takes these steps: 1) Segmenting point cloud of torso and head and estimating normals, 2) extracting Eigen-depth and skeleton features and 3) combining Eigen-depth with skeleton-based feature.

Our contributions are twofold: 1) We propose Eigen-depth feature for depth-based re-identification and prove the rotation invariance in theory. 2) We combine Eigen-depth and skeleton for complete representation and it is effective.

The experiments on datasets RGBD-ID [5] and BIWI RGBD-ID [26] show the effectiveness of our approach for overcoming clothes change. Our method outperforms other existing depth-based re-id methods PCM [26], skeleton-based feature [26] and rotation invariant descriptors RIFT2M [34] and Fehr's descriptor [9].

2. Covariance and Eigen-depth Feature

This section will detail the extraction of rotation invariant[1] descriptor describing frontal body shape.

2.1. Basic Feature Extraction

After converting depth image to point cloud, we segment head and torso by skeleton joints for more reliable body parts, and estimate surface normals following [12]. The feature vector of a point is $F(x, y, z) = [x, y, z, n_x, n_y, n_z]^T$, composed by the coordinate (x, y, z) (unit: mm) and unit normal vector (n_x, n_y, n_z).

2.2. Local Depth Covariance

In order to depict the variation of local feature vectors, within-patch covariance and between-patch covariance are exploited. To some extent, slight shape change and noise can be absorbed by covariance matrices.

2.2.1 Within-patch Covariance

Point cloud is divided into 6×2 rectangular patches with 50% overlap, as shown in Figure 2 (c). In each patch, within-patch covariance is computed to describe shape.

For a patch $R_1 \subset F$, let $\{f_{1i}\}_{i=1..m}$ be the 6-dimensional feature vectors inside R_1. Within-patch covariance matrix C_1 is defined as follows:

$$C_1 = \frac{1}{m-1}\sum_{i=1}^{m}(f_{1i} - \mu_1)(f_{1i} - \mu_1)^T, \quad (1)$$

where μ_1 is the mean of the feature vectors of R_1.

2.2.2 Between-patch Covariance

While within-patch covariance describes shape in a region, the differences of shape between regions also contain discriminative information. Similar to the idea of standard covariance, we define a novel between-patch covariance to represent the relation between different patches.

As shown in Figure 2 (d), the point cloud is divided into 6×2 patches without overlap. For two adjacent patches $R_1 \subset F$ and $R_2 \subset F$, let $\{f_{1i}\}_{i=1..m}$ and $\{f_{2j}\}_{j=1..n}$ be the 6-dimensional feature vectors inside R_1 and R_2 respectively. We define the between-patch covariance matrix C_{12} as follows:

$$C_{12} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(f_{1i} - f_{2j})(f_{1i} - f_{2j})^T. \quad (2)$$

2.3. Extracting Rotation Invariant Eigen-features

Assume p, q are patches of a person in depth camera K. Let $\{f_{pi}^K\}_{i=1...n}$ denote feature vectors, C_p^K denote the within-patch covariance matrix for patch p and C_{pq}^K denote between-patch covariance matrix between patches p and q. We assume only view angle and position change in different cameras A and B. To express the transformation from camera A to camera B, let R_{AB1} denote the rotation transformation matrix of point coordinate (x, y, z), R_{AB2} denote the rotation transformation matrix of unit normal vector (n_x, n_y, n_z), $f_{AB} = [x_s, y_s, z_s, 0, 0, 0]^T$ denote the shift of pedestrian position. Then the transformations for feature vectors from camera A to camera B are

$$f_{1i}^B = R_{AB}(f_{1i}^A + f_{AB}), \quad (3)$$

$$f_{2i}^B = R_{AB}(f_{2i}^A + f_{AB}), \quad (4)$$

where $R_{AB} = \begin{pmatrix} R_{AB1} & O \\ O & R_{AB2} \end{pmatrix}$.

Substitute equations (3) and (4) into (1) and (2), then we have

$$C_1^B = R_{AB}C_1^A(R_{AB})^T, C_{12}^B = R_{AB}C_{12}^A(R_{AB})^T. \quad (5)$$

Since R_{AB1} and R_{AB2} satisfy $R_{AB1}(R_{AB1})^T = I$ and $R_{AB2}(R_{AB2})^T = I$. So $R_{AB}(R_{AB})^T = I$, and R_{AB} is an orthogonal transformation. Hence, within-patch covariance matrices C_1^B and C_1^A have the same eigenvalues, between-patch covariance matrices C_{12}^B and C_{12}^A have the same eigenvalues.

It can be seen that eigenvalues of within-patch covariance and between-patch covariance are invariant to rotation and position change.

[1] The rotation invariance is invalid when pedestrian turns the back to camera, since the descriptor describes frontal body shape.

2.3.1 Analysis

Here we provide more in-depth analysis about the role of those eigenvalues.

Let $C_1, C_2 \in Sym^+(6, \mathbb{R})$ denote two covariance matrices. The eigenvalue decomposition of C_1 and C_2 are $C_1 = U_1 DIAG(\lambda_{1i}) U_1^T$ and $C_2 = U_2 DIAG(\lambda_{2i}) U_2^T$, where $\lambda_{11}, \lambda_{12}, ..., \lambda_{16}$ and $\lambda_{21}, \lambda_{22}, ..., \lambda_{26}$ are eigenvalues in descending order, U_1 and U_2 are orthogonal matrices containing eigenvectors.

We note that rotation of point cloud and normal vectors can be normalized by matching the principal axis of C_2 and C_1 according to the descending order of eigenvalues. One can find an orthogonal transformation matrix Q such that $QU_2 = U_1$, where Q is the rotation transformation we want to eliminate. Hence, we construct a normalized covariance matrix $C_2^N = U_1 DIAG(\lambda_{2i}) U_1^T$, where λ_{2i} are eigenvalues of C_2 and U_1 contains eigenvectors of C_1. This is the rotation normalization from C_2 to C_1.

We use eigenvalues to construct feature vectors, let $X_1 = [ln\lambda_{11}\ ln\lambda_{p2}\ ...\ ln\lambda_{16}]^T, X_2 = [ln\lambda_{21}\ ln\lambda_{22}\ ...\ ln\lambda_{26}]^T$, then we can have the following theorem.

Theorem 1 *Computing Euclidean distance between X_1 and X_2 is equivalent to computing the geodesic distance between covariance matrix C_1 and rotation normalized covariance matrix C_2^N on the Riemannian manifold.*

Proof. The Euclidean distance between X_1 and X_2 is

$$\|X_2 - X_1\|_2 = \sqrt{\sum_{i=1}^{6} (ln\lambda_{2i} - ln\lambda_{1i})^2} = \sqrt{\sum_{i=1}^{6} ln^2 \frac{\lambda_{2i}}{\lambda_{1i}}}. \tag{6}$$

The geodesic distance between C_1 and C_2^N on Riemannian manifold [10] can be determined as follows:

$$dist(C_1, C_2^N) = \sqrt{\sum_{k=1}^{6} ln^2 \lambda_k(C_1, C_2^N)}, \tag{7}$$

where $\lambda_k(C_1, C_2^N)_{k=1...6}$ are the generalized eigenvalues of C_1 and C_2^N, determined by $\lambda_k C_1 x_k - C_2^N x_k = 0$, *i.e.*, eigenvalues of $C_1^{-1} C_2^N$.

$$\begin{aligned} C_1^{-1} C_2^N &= (U_1 DIAG(\lambda_{1i}) U_1^T)^{-1} (U_1 DIAG(\lambda_{2i}) U_1^T) \\ &= U_1 DIAG(\frac{\lambda_{2i}}{\lambda_{1i}}) U_1^T. \end{aligned} \tag{8}$$

Hence, the generalized eigenvalue of C_1 and C_2^N is

$$\lambda_i(C_1, C_2^N) = \frac{\lambda_{2i}}{\lambda_{1i}}. \tag{9}$$

Substituting equation (9) into (6) and (7), we get

$$\|X_2 - X_1\|_2 = dist(C_1, C_2^N). \tag{10}$$

It can be seen that the geodesic distance on Riemannian manifold $dist(C_1, C_2^N)$ is equivalent to the Euclidean distance between feature vectors X_1 and X_2.

Eigen-depth Features. The theorem tells if there is only rotation variation, the logarithm eigenvalue vector can convert distance between covariance matrices on Riemannian manifold to Euclidean distance between two vectors. It makes large amount of existing feature learning and quantization methods applicable. Since there may be bias when estimating the covariance matrix C, empirically, we compute eigenvalues of $C + \alpha I$ to obtain more robust result. So that the Eigen-depth feature of patch p is $X_p = [ln(\alpha + \lambda_{p1})\ ln(\alpha + \lambda_{p2})\ ...\ ln(\alpha + \lambda_{p6})]^T$. Here we set $\alpha = 1$. Using eigenvalues makes the feature representation compact.

3. Experiments

We evaluated our method on two RGB-D person re-identification datasets RGBD-ID [5] and BIWI RGBD-ID [26] collected by Kinect. The results are presented in Cumulated Matching Characteristics (CMC) curve [24].

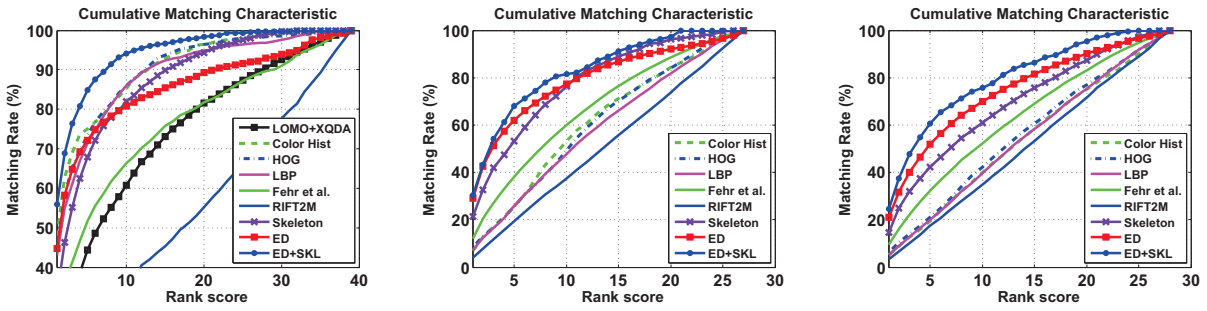
Feature Fusion and Matching. Skeleton-based feature [26] is constructed by 13 distances between skeleton joints (unit: cm). Eigen-depth feature (ED) and skeleton-based feature (SKL) are concatenated directly without setting weights. Feature dimension is reduced to 100 by PCA and reduced to $c - 1$ by LDA [36], where c is the number of classes. Then Euclidean distance is computed to match.

Compared Methods. We compared with appearance-based features including color histogram (RGB, HSV and YCbCr space) [7], HOG [29] and LBP [37], rotation invariant 3D descriptors including RIFT2M [34] and Fehr's descriptor [9], and depth-based re-identification descriptor skeleton-based feature [26]. Here RIFT2M and Fehr's were densely extracted as Eigen-depth. In the matching stage, LDA [36] was applied to all features, except that skeleton-based feature used Euclidean distance to match. The state-of-the-art appearance-based method LOMO+XQDA [17] was compared on RGBD-ID, but it cannot be applied on BIWI RGBD-ID for there is only one camera view in training set.

3.1. Evaluation on RGBD-ID

RGBD-ID [5] contains 79 people, in which groups "Walking1" and "Walking2" are with frontal view and some people change clothes. In each group, a person has five point clouds and skeletons. Following the open-set evaluation protocol of re-identification, we randomly sampled half of people in "Walking1" for training and the remaining for testing. Images of testing persons in "Walking1" were randomly selected as gallery and all images of testing persons in "Walking2" were for probe. In one-shot experiments, one image of a person was randomly selected as gallery while in multi-shot setting was five. The procedure was repeated 10 times, and average CMC curves were computed. The results are reported in Figure 3 (a) and Table 1.

It is observed that Eigen-depth feature is more effective than RIFT2M and Fehr's, indicating that it describes body shape better. As shape and skeleton are complementary, the combination of Eigen-depth and skeleton-based feature achieves encouraging performance of Rank-1 accu-



(a) RGBD-ID (one-shot) (b) BIWI "Still" (one-shot) (c) BIWI "Walking" (one-shot)

Figure 3. Performance on RGBD-ID and BIWI RGBD-ID. Our approach: ED, ED+SKL.

Table 1. RGBD-ID dataset: Rank-1 and Rank-5 accuracy.

Setting	One-shot		Multi-shot	
Method	Rank-1	Rank-5	Rank-1	Rank-5
LOMO+XQDA	16.10%	44.31%	18.00%	47.74%
Color Hist	47.90%	74.97%	48.92%	74.82%
HOG	45.03%	73.49%	45.33%	73.95%
LBP	42.92%	71.33%	45.64%	72.36%
RIFT2M	7.13%	22.77%	8.77%	27.69%
Fehr's	24.26%	51.64%	30.56%	58.67%
Skeleton	33.13%	67.85%	37.33%	71.13%
ED	44.67%	72.10%	51.59%	76.15%
ED+SKL	**55.95%**	**84.77%**	**61.23%**	**87.64%**

Table 2. BIWI RGBD-ID "Still" and "Walking": Rank-1 and Rank-5 accuracy.

Probe	Still				Walking			
Setting	One-shot		Multi-shot		One-shot		Multi-shot	
Method	Rank-1	Rank-5	Rank-1	Rank-5	Rank-1	Rank-5	Rank-1	Rank-5
Color Hist	7.02%	25.47%	10.61%	31.92%	5.43%	19.56%	5.86%	21.70%
HOG	8.42%	25.69%	12.35%	30.39%	6.38%	21.00%	6.94%	23.29%
LBP	7.37%	26.04%	10.87%	33.57%	4.87%	20.04%	5.34%	23.31%
RIFT2M	4.04%	19.52%	4.34%	20.78%	3.25%	17.46%	3.75%	18.31%
Fehr's	12.08%	38.17%	14.06%	43.78%	9.33%	32.39%	12.09%	39.60%
Skeleton	21.33%	53.32%	26.55%	62.73%	14.52%	42.36%	16.94%	47.18%
ED	28.98%	61.85%	36.22%	**73.11%**	20.90%	51.98%	28.71%	63.85%
ED+SKL	**30.52%**	**67.86%**	**39.38%**	72.13%	**24.47%**	**60.63%**	**29.96%**	**65.18%**

racy 55.95%. Appearance-based methods do not completely fail, because not all people change clothes.

3.2. Evaluation on BIWI RGBD-ID

BIWI RGBD-ID [26] contains three groups data of 50 people. In each group, a person has about 300 frames of depth images and skeletons. In "Training", people rotated and walked. In "Still", people slightly moved in place with frontal view. In "Walking", people walked in different views. Only 28 People in "Training" were recorded in "Still" and "Walking" and they dressed differently.

We selected frames with frontal and side view by face detection for evaluation as in [26]. We selected people recorded only in "Training" for training and the remaining for testing. We selected testing images in "Training" as gallery, "Still" and "Walking" as probe. The other experiment settings followed the tests on RGBD-ID. The performance is reported in Figure 3 (b), (c) and Table 2.

In BIWI RGBD-ID, people appeared in different view angles, especially in "Walking", so the problem becomes more challenging. Appearance-based methods fail because most people change clothes so that color is not reliable. Our Eigen-depth descriptor outperforms SKL, RIFT2M and Fehr's and deals with rotation better. The combination of ED and SKL presents the best performance.

3.3. Close-set Evaluation

The experiment results of the existing depth-based methods PCM [26], PCM+SKL [25] on BIWI RGBD-ID were reported in [26, 25] using close-set settings. We follow their experiment settings and test ED and ED+SKL on BIWI RGBD-ID. Rank-1 accuracy and normalized area under CMC curve (nAUC) are reported in Table 3.

Table 3. BIWI RGBD-ID: close-set evaluation

Probe	Still		Walking	
Method	Rank-1	nAUC	Rank-1	nAUC
PCM	32.5%	89.0%	22.4%	81.6%
PCM+SKL	38.6%	**91.8%**	27.4%	87.4%
ED	48.7%	90.9%	38.3%	87.5%
ED+SKL	**54.3%**	90.5%	**42.4%**	**87.5%**

Our methods ED and ED+SKL outperform PCM and PCM+SKL. Since both ED and PCM are rotation invariant, Eigen-depth feature is more tolerant to body shape change than PCM and more suitable for person re-identification.

4. Conclusion

In this paper, we propose an effective rotation invariant 3D descriptor Eigen-depth for depth-based person re-identification. Exploiting depth information can deal with clothes change, extreme illumination and background influences in appearance-based methods. The rotation invariance is based on eigen analysis of covariance matrices and it is proven in theory. The experiment results on two RGB-D re-identification datasets show that, Eigen-depth feature describes body shape and deals with rotation better than existing methods, and appearance-based methods suffer from clothes change. Combining Eigen-depth feature with skeleton-based feature presents the best performance.

Acknowledgement

This work was supported partially by the National NSFC (No. 61472456, No. 61573387, No. 61522115, No. U1135001), Guangzhou Pearl River Science and Technology Rising Star Project (No. 2013J2200068), and in part by

Fundamental Research Funds for the Central Universities (No. 15lgjc).

References

[1] A. Albiol, J. Oliver, and J. Mossi. Who is who at different cameras: People re-identification using depth cameras. *IET computer vision*, 2012. 1

[2] S. Bak and F. Brémond. Re-identification by covariance descriptors. In *Person Re-Identification*. 2014. 1

[3] S. Bak, E. Corvee, F. Bremond, and M. Thonnat. Person re-identification using haar-based and dcd-based signature. In *AVSS*, 2010. 1

[4] D. Baltieri, R. Vezzani, and R. Cucchiara. Learning articulated body models for people re-identification. In *ACMMM*, 2013. 1

[5] B. I. Barbosa, M. Cristani, A. Del Bue, L. Bazzani, and V. Murino. Re-identification with rgb-d sensors. In *ECCV Workshop*, 2012. 1, 2, 3

[6] M. Castrillon-Santana, J. Lorenzo-Navarro, and D. Hernandez-Sosa. People semantic description and re-identification from point cloud geometry. In *ICPR*, 2014. 1

[7] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani. Person re-identification by symmetry-driven accumulation of local features. In *CVPR*, 2010. 1, 3

[8] D. Fehr, W. J. Beksi, D. Zermas, and N. Papanikolopoulos. Rgb-d object classification using covariance descriptors. In *ICRA*, 2014. 1

[9] D. Fehr, A. Cherian, R. Sivalingam, S. Nickolay, V. Morellas, and N. Papanikolopoulos. Compact covariance descriptors in 3d point clouds for object recognition. In *ICRA*, 2012. 2, 3

[10] W. Förstner and B. Moonen. A metric for covariance matrices. In *Geodesy-The Challenge of the 3rd Millennium*. 2003. 3

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

[12] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. *Surface Reconstruction from Unorganized Points*. 1992. 2

[13] V. John, G. Englebienne, and B. J. Kröse. Person re-identification using height-based gait in colour depth camera. In *ICIP*, 2013. 1

[14] K. Jungling, C. Bodensteiner, and M. Arens. Person re-identification in multi-camera networks. In *CVPR*, 2011. 1

[15] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof. Large scale metric learning from equivalence constraints. In *CVPR*, 2012. 1

[16] I. Kviatkovsky, A. Adam, and E. Rivlin. Color invariants for person reidentification. *PAMI*, 2013. 1

[17] S. Liao, Y. Hu, X. Zhu, and S. Z. Li. Person re-identification by local maximal occurrence representation and metric learning. In *CVPR*, 2015. 3

[18] C. Liu, S. Gong, C. C. Loy, and X. Lin. Person re-identification: what features are important? In *ECCV Workshop*, 2012. 1

[19] J. Lorenzo-Navarro, M. Castrillón-Santana, and D. Hernández-Sosa. An study on re-identification in rgb-d imagery. In *Ambient Assisted Living and Home Care*. 2012. 1

[20] B. Ma, Y. Su, and F. Jurie. Local descriptors encoded by fisher vectors for person re-identification. In *ECCV Workshop*, 2012. 1

[21] A. Mignon and F. Jurie. Pcca: A new approach for distance learning from sparse pairwise constraints. In *CVPR*, 2012. 1

[22] A. Mogelmose, C. Bahnsen, T. B. Moeslund, A. Clapés, and S. Escalera. Tri-modal person re-identification with rgb, depth and thermal features. In *CVPR Workshop*, 2013. 1

[23] A. Mogelmose, T. B. Moeslund, and K. Nasrollahi. Multimodal person re-identification using rgb-d sensors and a transient identification database. In *IWBF*, 2013. 1

[24] H. Moon and P. J. Phillips. Computational and performance aspects of pca-based face-recognition algorithms. *Perception-London*, 2001. 3

[25] M. Munaro, A. Basso, A. Fossati, L. V. Gool, and E. Menegatti. 3d reconstruction of freely moving persons for re-identification with a depth sensor. In *ICRA*, 2014. 1, 4

[26] M. Munaro, A. Fossati, A. Basso, E. Menegatti, and L. V. Gool. *One-Shot Person Re-Identification with a Consumer Depth Camera*. 2014. 1, 2, 3, 4

[27] M. Munaro, S. Ghidoni, D. T. Dizmen, and E. Menegatti. A feature-based approach to people re-identification using skeleton keypoints. In *ICRA*, 2014. 1

[28] J. Oliver and A. Albiol. 3d descriptor for people re-identification. In *ICPR*, 2012. 1

[29] O. Oreifej, R. Mehran, and M. Shah. Human identity iecognition in aerial images. In *CVPR*, 2010. 1, 3

[30] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian. Local fisher discriminant analysis for pedestrian re-identification. In *CVPR*, 2013. 1

[31] B. Prosser, W.-S. Zheng, S. Gong, T. Xiang, and Q. Mary. Person re-identification by support vector ranking. In *BMVC*, 2010. 1

[32] R. Satta, F. Pala, G. Fumera, and F. Roli. Real-time appearance-based person re-identification over multiple kinecttm cameras. In *VISAPP*, 2013. 1

[33] W. R. Schwartz and L. S. Davis. Learning discriminative appearance-based models using partial least squares. In *SIBGRAPI*, 2009. 1

[34] L. J. Skelly and S. Sclaroff. Improved feature descriptors for 3d surface matching. In *Optics East*, 2007. 2, 3

[35] B. Takac, A. Catala, M. Rauterberg, and W. Chen. People identification for domestic non-overlapping rgb-d camera networks. In *SSD*, 2014. 1

[36] A. R. Webb. *Statistical pattern recognition*. 2003. 3

[37] Y. Zhang and S. Li. Gabor-lbp based region covariance descriptor for person re-identification. In *ICIG*, 2011. 3

[38] R. Zhao, W. Ouyang, and X. Wang. Person re-identification by salience matching. In *ICCV*, 2013. 1

[39] W.-S. Zheng, S. Gong, and T. Xiang. Reidentification by relative distance comparison. *PAMI*, 2013. 1