Improved Human Gait Recognition

Imad Rida[1(✉)], Ahmed Bouridane[2], Gian Luca Marcialis[3], and Pierluigi Tuveri[3]

[1] LITIS EA 4108 - INSA de Rouen, Saint Etienne du Rouvray, Rouen, France
imad.rida@insa-rouen.fr

[2] Department of Computer Science and Digital Technologies, Northumbria University, Newcastle, UK
ahmed.bouridane@northumbria.ac.uk

[3] Department of Electrical and Electronic Engineering, University of Cagliari, Cagliari, Italy
marcialis@diee.unica.it, tuveri.pierluigi@gmail.com

Abstract. Gait recognition is an emerging biometric technology which aims to identify people purely through the analysis of the way they walk. The technology has attracted interest as a method of identification because of its non-invasiveness, since it does not require the subject's cooperation. However, "covariates" which include clothing, carrying conditions, and other intra-class variations affect the recognition performances. This paper proposes a feature selection mask which is able to select most relevant discriminative features for human recognition to alleviate the impact of covariates so as to improve the recognition performances. The proposed method has been evaluated using CASIA Gait Database (Dataset B) and the experimental results demonstrate that the proposed technique yields 77.38 % of correct recognition.

Keywords: Biometrics · Gait · Model free · Feature selection

1 Introduction

Technology has invaded our lives as never before and the effectiveness of current security systems has become increasingly important. Biometric recognition aims to identify individuals using unique, reliable and stable physiological and/or behavioral characteristics such as fingerprint, palmprint, face, gait, etc. Gait recognition consists on discriminating among people by the way or manner they walk.

Gait recognition techniques can be classified into two main categories: model-based and model-free approach. Model based approach [1,2] models the person body structure, it uses the estimation over time of static body parameters for recognition (i.e. trajectory, limb lengths etc). This process is usually computationally intensive since we need to model and track the subjects body. On the other hand, the model free approach does not recover a structural model of human motion, instead it uses the features extracted from the motion or shape

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part II, LNCS 9280, pp. 119–129, 2015.
DOI: 10.1007/978-3-319-23234-8_12

for recognition. Compared to a model based approach, the model free approach requires much less computation cost, furthermore dynamic information results in improved recognition performance than static counterpart [3]. These reasons have motivated the researchers to introduce new feature representations in model free approach context. The major challenges of methods belong the model free gait recognition are due to the effect of various covariates as the presence of shadows, clothing variations and carrying conditions (backpack, briefcase, handbag, etc). Moreover, segmentation and the view dependency are further causes of gait recognition errors. This has motivated the work presented in this paper which aims to mitigate the effect of the covariates and improve the recognition performance.

The rest of this paper is organized as follows: Sect. 2 summarizes the previous works. Sect. 3 gives the theoretical description of the proposed method. Sect. 4 presents the experimental results. Sect. 5 offers our conclusion.

2 Related Works

There exists a considerable amount of work in the context of model free approaches for gait recognition. Benabdelkader et al. [4] introduced a self similarity representation to measure the similarity between pairs of silhouettes. Collins et al. [5] proposed a template based silhouette matching in some key frames. Recent trends seem to favor Gait Energy Image (GEI) representation suggested by Han and Bhanu [6]. GEI is a spatio-temporal representation of the gait obtained by averaging the silhouettes over a gait cycle. This representation has already been used in several state of the art works [7–10]. Yu et al. introduced a simple template matching technique based on the euclidian distance without data reduction and feature selection [11]. It has been found that the different clothing and carrying conditions between the gallery and probe sequences influence the recognition performances [6,11]. To overcome the limitations of the GEI presentation, Bashir et al. introduced a novel gait feature selection method named Gait Entropy Image (GEnI) [12]. It consists of computing Shannon entropy for each pixel over a gait cycle; in other terms it aims to distinguish static and dynamic pixels of the GEI. In this case GEnI represents a measure of feature significance (pixels with high entropy correspond to dynamic parts which are robust against appearance changes). In the same context Bashir et al. suggested a new gait representation called flow field [13] in order to represent a weighted sum of the optical flow corresponding to each coordinate direction of human motion. Rida et al. [19,20] proposed a supervised feature extraction method based on Modified Phase-Only Correlation (MPOC) algorithm which is an improved version of the Phase Only Correlation (POC). Recently Random Subspace Method (RSM) has been been used to reduce the effect of the covariates, the results showed very good performances in the USF database [14,15].

3 Methodology

3.1 Motivations

In this paper among all available feature representations we have chosen the so-called Gait Energy Image [6]: it is an easy and simple representation to compute, thus making it an effective compromise between the computational cost and the recognition performance. Its main drawback is common to all model-free approaches: covariates makes it unreliable.

The aim of this work is to improve the GEI representation by determining a mask capable to select the robust features against the covariates. The notion of gait mask was introduced by Foster et al. [16] where several predefined masks were used to capture gait characteristics. Fig 1 shows some of such predefined masks, where the gray parts represent the features selected whereas the black parts represent the non selected ones.

(a) (b) (c)

Fig. 1. Predefined masks introduced by Foster et al. to capture gait features [16].

In the current work we propose to estimate a mask instead of using a predefined one as suggested in the previous work by Foster et al. [16]. The calculation of the mask on all data will bias the results, furthermore the selection method (mask) should not be overspecialized for a particular and specific training set [17], all that has motivated to current work which has as particularity to estimate a fixed unique mask on a small feature selection set independently from training and testing sets (all selected sequences from the feature selection set were removed from the training and testing sets) capable to select relevant features from all GEIs under both carrying and clothing conditions (see Fig. 2). To calculate our mask we estimate the normal walk GEI called $\mathbf{M}$ by taking the mean of all normal walk GEIs within the feature selection set, after that we calculate the variation matrix $\mathbf{D}_k$ between each couple (estimated normal walk $\mathbf{M}$ and carrying-bag/wearing-coat GEIs) by taking the difference pixel by pixel. A mask $\{\mathbf{S}_k\}_{k=1}^{K}$ is defined for each variation matrix $\{\mathbf{D}_k\}_{k=1}^{K}$ which aims to select pixels with low variation value by assigning 1 for pixels with variation $\mathbf{D}_k(i,j)$ less than a threshold T and 0 otherwise. The masks $\{\mathbf{S}_k\}_{k=1}^{K}$ are combined together using a simple 'AND' operator to obtain our final mask $\mathbf{S}$ (see Sect. 3.3).

Our framework is divided into two main modules: the first one consists of calculating the mask on the feature selection set. The second module estimates

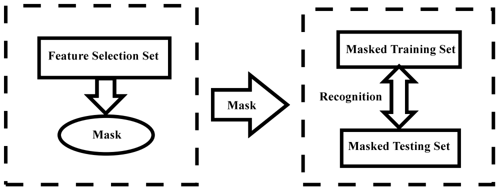


Fig. 2. Scheme of our framework.

the performance of our method (Correct Classification Rate) using GEI features selected with the resulting mask in the first module and Component Discriminant Analysis (CDA) [18] (see Sec. 3.4).

3.2 Gait Energy Image

GEI is a spatio-temporal representation of the gait patterns. It consists of representing the gait cycle using a single grayscale image obtained by averaging the silhouettes extracted over a complete gait cycle [6]. GEI is computed using the following equation:

$$\mathbf{G}(x,y) = \frac{1}{N}\sum_{t=1}^{N}\mathbf{B}(x,y,t) \tag{1}$$

where N is the number of the frames within a complete gait cycle, $\mathbf{B}$ is a silhouette image, x and y are the coordinates of the image and t is frame number in the cycle. Low and high intensity pixels of the GEI correspond to the dynamic and static parts of the body, respectively. Dynamic parts are most informative since they contain the information of the gait while static parts are sensitive since they contain the shape and contour information which can easily be influenced by the covariates [12].

(a) Normal Walk (b) Carrying Bag (c) Wearing Coat

Fig. 3. Gait energy image of an individual under different conditions.

3.3 Feature Selection Mask

Let consider L Gait Energy Image templates $\{\mathbf{G}_l\}_{l=1}^{L}$ characterizing normal gait walking, we calculate the mean GEI normal walk as follows:

$$\mathbf{M} = \frac{1}{L}\sum_{l=1}^{L} \mathbf{G}_l \tag{2}$$

The variation $\{\mathbf{D}_k\}_{k=1}^{K}$ for a given GEI template $\{\mathbf{G'}_k\}_{k=1}^{K}$ characterizing carrying bag or wearing coat walk is given by:

$$\mathbf{D}_k = \mathbf{G'}_k - \mathbf{M} \tag{3}$$

$\mathbf{D}_k$ represents a measure of feature significance (i.e. discriminative power) since pixels with large variation are more suspected to be affected by the covariates (it can be seen as an inverse relationship between variation value and importance). Someone can say that $\mathbf{D}_k$ can contain negative values and we should take the square when we calculate $\mathbf{D}_k$, this is not possible for the simple reason that two pixels with same position (i, j) from two different GEI templates $\mathbf{G}_1(i, j)$ and $\mathbf{G}_2(i, j)$ with the corresponding variations $\mathbf{D}_1(i, j) < \mathbf{D}_2(i, j)$ and $|\mathbf{D}_1(i, j)| = |\mathbf{D}_2(i, j)|$ don't have the same importance because $\mathbf{G}_2(i, j)$ has more intensity value than $\mathbf{G}_1(i, j)$, as consequence it is much more suspected to be affected by the covariates. To facilitate our calculations we normalize the matrix $\mathbf{D}_k$ values between 0 and 1.

A mask defines if a feature is selected therefore a binary representation is useful: assigning a value 1 or 0 corresponding to selected or unselected features respectively. The mask based GEI template is given by:

$$\mathbf{S}_k(i, j) = \begin{cases} 1, & if\; \mathbf{D}_k(i, j) \leq T \mid T \in [0, 1] \\ 0, & otherwise \end{cases} \tag{4}$$

Where T represents the threshold. The masks $\{\mathbf{S}_k\}_{k=1}^{K}$ are combined together using a simple binary 'AND' to obtain the final mask $\mathbf{S}$ which is given by:

$$\mathbf{S} = \mathbf{S}_1 \;\&\&, \cdots,\; \&\&\; \mathbf{S}_k, \cdots, \&\&\; \mathbf{S}_K \tag{5}$$

Algorithm 1. Mask calculaion algorithm.

1: **Input:** $\{\mathbf{G}_l\}_{l=1}^{L}$ (normal walk GEI templates)
   $\{\mathbf{G'}_k\}_{k=1}^{K}$ (carrying bag and wearing coat walk GEI templates)
   T: treshold
   Output: $\mathbf{S}$ (mask)
2: Calculate $\mathbf{M}$ using formula (2);
3: **for** $k = 1$ to K **do**
4:     Compute $\mathbf{D}_k$ using formula (3);
5:     Compute $\mathbf{S}_k$ using formula (4);
       $k = k + 1$;
6: **end for**
7: Compute the mask $\mathbf{S}$ using formula (5);

Where && is the binary operator. The whole process of mask calculation is summarized step by step in Alg. 1.

3.4 Canonical Discriminat Analysis

Canonical Discriminant Analysis (CDA) corresponds to Principal Component Analysis (PCA) followed by a Multiple Discriminant Analysis (MDA). The aim of the PCA is to be able to represent most of the variations of the original data using only a few principal components which are orthogonal to each others. MDA tries to maximize the distance between classes and preserve the distance inside the classes (the full explantation is found in [18]). The performance of our method is measured with the correct classification rate (CCR) which corresponds to the ratio of the number of well classified samples over the total number of samples.

Let n d-dimensional training GEI templates $\{\mathbf{g}_1, \cdots, \mathbf{g}_n\}$, where each template is a column vector obtained by concatenating the rows of the corresponding GEI. The feature selection is applied to these templates using the mask to obtain n d'-dimensional GEI templates $\{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$ where $d' < d$. PCA aims to minimize the following objective function:

$$J_{d''} = \sum_{k=1}^{n} \left\| \left(\mathbf{m} + \sum_{i=1}^{d''} a_{ki}\mathbf{u}_i \right) - \mathbf{x}_k \right\|^2 \tag{6}$$

where $d'' < d' < d$, $\mathbf{m} = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k$, $\{\mathbf{u}_1, \cdots, \mathbf{u}_{d''}\}$ set of orthogonal unit vectors representing new coordinate system of the subspace and a_{ki} is the projection of the k-th data over $\mathbf{u}_i$.$J_{d''}$ is minimized when $\mathbf{u}_1, \cdots, \mathbf{u}_{d''}$ are eigenvectors of the largest eigenvalues of the covariance matrix $\mathbf{C}$ given by:

$$\mathbf{C} = \sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m})(\mathbf{x}_k - \mathbf{m})^T \tag{7}$$

The d''-dimensional feature vector $\mathbf{y}_k$ obtained from $\mathbf{x}_k$ is given by:

$$\mathbf{y}_k = [a_1, \cdots, a_{d''}]^T = [\mathbf{u}_1, \cdots, \mathbf{u}_{d''}]^T \, \mathbf{x}_k, \ k = 1, \cdots, n \tag{8}$$

As suggestion in [6] we retain $d'' = 2c$ eigenvectors after applying PCA. Suppose that the n d''-dimensional principal vectors $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$ belong c classes, MDA is a supervised learning method which seeks a transformation matrix $\mathbf{W}$ that maximizes the ratio of the between-class scatter matrix S_B to the within-class scatter matrix S_W given by:

$$J(\mathbf{W}) = \frac{|\mathbf{W}^T S_B \mathbf{W}|}{|\mathbf{W}^T S_W \mathbf{W}|} \tag{9}$$

The within-class scatter matrix in the PCA subspace S_W is defined as $S_W = \sum_{i=1}^{c} S_i$ where:

$$\begin{cases} S_i = \sum_{\mathbf{y}\in\mathcal{D}_i}(\mathbf{y}-\mathbf{m}_i)(\mathbf{y}-\mathbf{m}_i)^T \\ \mathbf{m}_i = \frac{1}{n\,i}\sum_{\mathbf{y}\in\mathcal{D}_i}\mathbf{y} \\ \{\mathcal{D}_i\}_{i=1}^{c} \; training\ data\ of\ class\ i\ of\ size\ n_i \end{cases} \tag{10}$$

The between-class scatter in the PCA subspace S_B is given by:

$$S_B = \sum_{i=1}^{c} n_i(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \tag{11}$$

where $\mathbf{m} = \frac{1}{n}\sum_{\mathbf{y}\in\mathcal{D}}\mathbf{y}$. $J(\mathbf{W})$ is maximized when the columns of $\mathbf{W}$ are the generalized eigenvectors that correspond to $c-1$ nonzero eigenvalues in:

$$S_B\mathbf{w}_i = \lambda_i S_W \mathbf{w}_i \tag{12}$$

where $\mathbf{w}_i$ is the i-th column of the matrix $\mathbf{W}$. The corresponding generalized eigenvectors are denoted by $\mathbf{v}_1, \cdots, \mathbf{v}_{c-1}$. The $(c-1)$-dimensional feature vector $\mathbf{z}_k$ in the MDA subspace is obtained from the d''-dimensional principal component vector $\mathbf{y}_k$:

$$\mathbf{z}_k = [\mathbf{v}_1, \cdots, \mathbf{v}_{c-1}]^T \mathbf{y}_k, \quad k = 1, \cdots, n \tag{13}$$

4 Experiments

We have used CASIA database (dataset B) [11] to evaluate our method. It is a multiview gait database containing 124 subjects captured from 11 different angles. Each subject has six normal walking sequences (SetA), two carrying-bag sequences (SetB) and two wearing-coat sequences (SetC). The first four sequences of setA noted as (SetA1) are used for training. The two remaining noted as (SetA2), (SetB) and (SetC) are used for testing the effect of view angle variations, clothing and carrying conditions respectively. In our work we focus on the effect of clothing, carrying conditions and experiments are carried out under 90° view using 64×64 GEI resolution. We determine our mask from a feature selection set independent from training and testing sets (all selected sequences from the feature selection set were removed from the training and testing sets). To create our feature selection set we randomly select 24 subjects without replacement as follows: for each subject 3 sequences are randomly chosen corresponding to the three situations (normal, carrying bag, wearing coat) so that 72 sequences are obtained. To make our method robust and avoid the

Algorithm 2. The Evaluation Method

Input: feature selection set
for $p = 1$ to P **do**
  Randomly select without replacement of 15 subjects from feature selection set;
  Select related GEI templates corresponding to the three variants (normal, carrying bag, wearing coat);
  Calculate the mask based on Alg. 1 ;
  Estimate the best threshold value T using 3-folds Cross-Validation;
  Select the mask corresponding to the best threshold performance;
end for

overspecialization we have applied the evaluation strategy described in Alg. 2 on feature selection set for $P = 5$ (The threshold T is estimated using a 3-folds Coss-Validation for $T \in [0, 1]$ with a step of 0.1).

It can be seen from Fig. 4 that the threshold value $T = 0.6$ is giving the best performance for the $P = 5$ experiments, we combine the resulting $P = 5$ masks of the experiments together with a simple 'AND' operator to obtain our mask used to select relevant features and remove irrelevant ones.

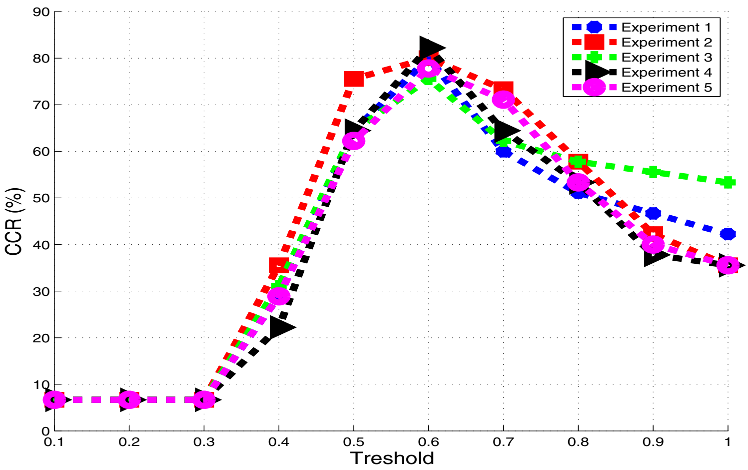
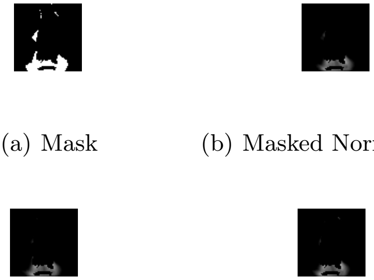


Fig. 4. Correct classification rate of the carried out experiments on the feature selection set using different threshold values.

Fig. 5 shows the calculated mask by our method as well as the masked GEI under different conditions (the white part from the mask represents the selected features). We can notice that our mask selects features from the bottom part of the GEI template which represent the dynamic movement of the legs, this part is robust against the covariates and is discriminative, our mask selects also some



(a) Mask (b) Masked Normal GEI

(c) Masked Carrying GEI (d) Masked Clothing GEI

Fig. 5. The mask obtained by our method and the corresponding masked GEI of an individual under different conditions.

features from the top part of the GEI corresponding the dynamic motion of the hands during the walk and head shape.

Tab. 1 represents the results obtained by our method compared to the reported results of four other state-of-the-art methods. It can be seen that our method loses slightly a bit performance in the case of normal walk condition and improves considerably the performance in the case of clothing conditions. Moreover, it makes the best compromise between gait walk conditions performance which can be seen by the mean and the standard deviation, which outperform the other ones. This can be explained by the fact that our method eliminates features from the top of the GEI template which, in turn, improve the recognition performance in the case of normal and carrying bag walks while these features are considerably affected in the presence of wearing coat covariates.

Table 1. Comparison of CCRs (In percent) from several different algorithms on CASIA database using 90° view.

Method	Normal	Carrying-Bag	Wearing-Coat	Mean	Std.Dev.
Han et al. [6]	99.60	57.20	23.80	60.20	37.99
Yu et al. [11]	97.60	32.70	52.00	60.77	33.33
Bashir et al. [12]	**100.00**	78.30	44.00	74.10	28.24
Bashir et al. [13]	97.50	**83.60**	48.80	76.63	25.09
Our Method	95.97	63.39	**72.77**	**77.38**	**16.77**

5 Conclusions

This paper has presented a feature selection mask for improved gait recognition. The proposed mask demonstrates attractive results in the presence of clothing covariates and makes the best compromise between different gait walk recognition performances.

As future work we will investigate the robustness of the mask in case of view angle variation between training and testing data and extend the results to USF database to compare our method with others using this dataset [15].

References

1. Yam, C., Nixon, M.S., Carter, J.N.: Automated person recognition by walking and running via model-based approaches. Pattern Recognition **37**(5), 1057–1072 (2004)
2. Niyogi, S.A., Adelson, E.H.: Analyzing and recognizing walking figures in XYT. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 1994, pp. 469–474. IEEE (1994)
3. Wang, L., Ning, H., Tan, T., Hu, W.: Fusion of static and dynamic body biometrics for gait recognition. IEEE Transactions on Circuits and Systems for Video Technology **14**(2), 149–158 (2004)
4. Benabdelkader, C., Cutler, R.G., Davis, L.S.: Gait recognition using image self-similarity. EURASIP Journal on Advances in Signal Processing **2004**(721765), 572–585 (2004). doi:10.1155/S1110865704309236
5. Collins, R.T., Gross, R., Shi, J.: Silhouette-based human identification from body shape and gait. In: Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 366–371. IEEE (2002)
6. Han, J., Bhanu, B.: Individual recognition using gait energy image. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(2), 316–322 (2006)
7. Tao, D., Li, X., Wu, X., Maybank, S.J.: General tensor discriminant analysis and gabor features for gait recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(10), 1700–1715 (2007)
8. Xu, D., Yan, S., Tao, D., Zhang, L.: Human gait recognition with matrix representation. IEEE Transactions on Circuits and Systems for Video Technology **16**(7), 896–903 (2006)
9. Xu, D., Yan, S., Tao, D., Lin, S., Zhang, H.-J.: Marginal fisher analysis and its variants for human gait recognition and content-based image retrieval. IEEE Transactions on Image Processing **16**(11), 2811–2821 (2007)
10. Zhang, E., Zhao, Y., Xiong, W.: Active energy image plus 2DLPP for gait recognition. Signal Processing **90**(7), 2295–2302 (2010)
11. Yu, S., Tan, D., Tan, T.: A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: 18th International Conference on Pattern Recognition, ICPR 2006, pp. 441–444. IEEE
12. Bashir, K., Xiang, T., Gong, S.: Gait recognition without subject cooperation. Pattern Recognition Letters **31**(13), 2052–2060 (2010)
13. Bashir, K., Xiang, T., Gong, S.: Gait representation using flow fields. In: BMVC, pp. 1–11 (2009)
14. Guan, Y., Wei, X., Li, C.-T., Marcialis, G.L., Roli, F., Tistarelli, M.: Combining gait and face for tackling the elapsed time challenges. In: 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–8. IEEE
15. Guan, Y., Li, C.-T., Roli, F.: On reducing the effect of covariate factors in gait recognition: a classifier ensemble method. IEEE Trans. on Pattern Analysis and Machine Intelligence (2014) doi:10.1109/TPAMI.2014.2366766

16. Foster, J.P., Nixon, M.S., Prgel-Bennett, A.: Automatic gait recognition using area-based metrics. Pattern Recognition Letters **24**(14), 2489–2497 (2003)
17. Zheng, S., Zhang, J., Huang, K., He, R., Tan, T.: Robust view transformation model for gait recognition. In: 2011 18th IEEE International Conference on Image Processing (ICIP), pp. 2073–2076. IEEE
18. Huang, P.S., Harris, C.J., Nixon, M.S.: Recognising humans by gait via parametric canonical space. Artificial Intelligence in Engineering **13**(4), 359–366 (1999)
19. Rida, I., Bouridane, A., Al Kork, S., Bremond, F.: Gait recognition based on modified phase only correlation. In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D. (eds.) ICISP 2014. LNCS, vol. 8509, pp. 417–424. Springer, Heidelberg (2014)
20. Rida, I.: Al Maadeed, S., Bouridane, A. Gait recognition based on modified phase-only correlation. Signal, Image and Video Processing (2015). doi:10.1007/s11760-015-0766-4