Person Re-identification in frontal gait sequences via Histogram of Optic flow Energy Image

Athira Nambiar[1], Jacinto C. Nascimento[1], Alexandre Bernardino[1], and José Santos-Victor[1]

Institute for Systems and Robotics, Instituto Superior Técnico,
Universidade de Lisboa, 1049-001 Lisboa, Portugal

Abstract. In this work, we propose a novel methodology of re-identifying people in frontal video sequences, based on a spatio-temporal representation of the gait based on optic flow features, which we call Histogram Of Flow Energy Image (**HOFEI**). Optic Flow based methods do not require the silhouette computation thus avoiding image segmentation issues and enabling online re-identification (Re-ID) tasks. Not many works addressed Re-ID with optic flow features in frontal gait. Here, we conduct an extensive study on CASIA dataset, as well as its application in a realistic surveillance scenario- HDA Person dataset. Results show, for the first time, the feasibility of gait re-identification in frontal sequences, without the need for image segmentation.

Keywords: Gait analysis, Optic Flow, Histogram of Flow, Gait Energy Image.

1 Introduction

Over the past few years, human gait, has been receiving unprecedented attention from pattern recognition and computer vision communities, as a rich behavioural soft-biometric cue. A plethora of studies have been conducted on the visual analysis of human motion and automated person Re-identification/ Recognition. The cognitive and psychological studies have proclaimed that humans are able to identify his peers by their distinct gait signature [1]. Human gait, which includes both the body appearance and the dynamics of walking [2], is considered to be quite pertinent in visual surveillance scenario. This is because gait analysis is unobtrusive, does not require explicit user cooperation, is perceivable from a distance, and is unique for each individual.

During the last decade, a number of gait analysis techniques have been proposed towards Person Re-identification/ Recognition. Re-identification (Re-ID) is the process of identifying the same individual in different time instances either in the same camera or in different cameras. Re-identification is associated with change in appearance (carrying bags and different clothings etc.) and uncontrolled conditions (changes in illumination, pose and background). Recognition is a special case of Re-ID, where there is no apparent change in the appearance of the subject and the operator has much control on the conditions (same camera, no change in pose/ background/ illumination etc.).

The gait analysis techniques are broadly categorised as model based and model free approaches. The former leverage explicit either structural or motion gait models, whose parameters are estimated using the underlying kinematics of human motion in a sequence of images [3], whereas the latter category operate directly on the gait image sequences without fitting any underlying structure [4,5,6]. Although the model based approach is less susceptible to the changes in viewing angle or clothing, the inaccurate model fitting (mainly due to noisy data) may lead to the poor recognition performance. Instead, the model free approaches provide better performance by using information directly from the temporal evolution of the gait image sequences.

In classical gait analysis, the most commonly used views are lateral and frontal views. Most of the state-of-the-art techniques address the lateral case, in which the gait can be better observed. Lateral views have the advantage of minimizing perspective distortion and the amount of self occlusion, however, they cannot be applied in narrow passages, since very few gait cycles are observed in those conditions. Hence, in many real world scenarios like indoor narrow corridors and confined spaces, systems that rely on frontal gait analysis are preferred due to the convenience to be installed in confined spaces, as well as the capability to capture longer video sequences, at the same time impose more challenges in terms of perspective and occlusions.

In this paper, we propose a novel framework for model-free and frontal gait analysis for person Re-ID, by amalgamating the Histogram of Flow (HOF) [7] into the the framework of Gait Energy Image (GEI) [4], and building upon the advantages of both representations. First, the HOF represents the dynamic gait characteristics by encoding the pattern of apparent motion of the subject in a visual scene. Second, GEI enables to average the energy information over a gait cycle to obtain the spatio-temporal gait signature. Our major contributions are:

- A new technique (termed as **HOFEI**) for person Re-ID in frontal videos leveraging optic flow features.
- Our proposal does not require binary silhouettes, instead computes global dense motion descriptors directly from raw images. This not only bypasses the segmentation and binarization phases, but also facilitates online Re-ID.
- Proposal of a new gait period estimation directly from the temporal evolution of the HOF computed at the lower limbs.
- The demonstration of an optic flow-based method to frontal gait recognition, which, to the best of our knowledge, is absent in the literature.

The pipeline of our proposed algorithm is shown in Fig. 1, which will be detailed in the forthcoming sections. The rest of the paper is structured as follows. Section 2 discusses related works in the area. Section 3 presents our proposed method, with a detailed description of HOF computation and the gait feature extraction. Section 4 presents the experimental results to demonstrate the effectiveness of the proposed algorithm. Finally, Section 5 concludes this work.

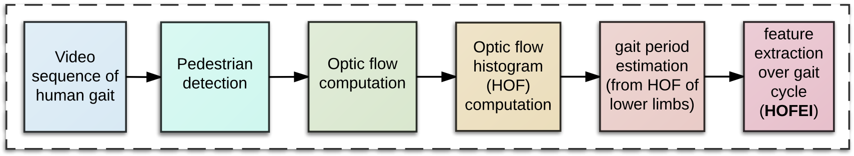


Fig. 1. Proposed pipeline of the gait analysis.

2 Related work

One of the most acclaimed research in model free gait recognition viz., Gait Energy Image (GEI) by Han et. al [4], presented the idea of generating spatio-temporal description by averaging the normalized binary silhouette over gait cycle. Inspired by the same, many other energy image improvements were carried out e.g., active energy image [10], gait entropy image [12], gradient histogram energy image [11] to quote a few.

The idea of HOF was adopted from Histogram of Oriented Gradients (HOG) [13], which divide the image into cells and compile a histogram of gradient directions, weighed by its magnitude for the pixels within each cell. The same approach has been extended to the optic flow and the spatial derivatives of its components [7,15]. Optic flow and their histograms have also been proposed for gait analysis such as in [16] by using motion intensity and direction from optical flow field, while in [8] a silhouette based gait representation has been used to generate gait flow image. In the field of optic flow based gait recognition also some energy image concepts were proposed in the recent works by [9] and [8]. However, both of those works were reasonably insufficient to convey the motion information of the whole human body since their optic flow measurements are on the binary silhouette edges. Much inspired from the aforementioned literature studies, here we propose a novel spatio-temporal gait representation termed as Histogram Of Flow Energy Image (**HOFEI**), which is a dense descriptor computed over the entire body parts.

Different from the aforementioned literature on optic flow based gait recognition conducted in the lateral view, we demonstrate the potential of our proposal in the front view (in HDA and CASIA dataset), for which no similar state-of-the-art using optic flow has been reported. However, there have been some works in CASIA dataset frontal sequences, leveraging the binary silhouettes for gait recognition. Chen et al [19] demonstrated the performance of various gait features including Gait Energy Image (GEI), Gait History Image (GHI), Gait Moment Image (GMI), Frame Difference Energy Image (FDEI) etc. in each view angle, from frontal to rear view. We will show that our proposed method is competitive with this state-of-the-art, while using an optic-flow method, which does not require silhouette segmentation.

3 Methodology

In this section, the target representation strategy via HOF, gait period estimation and the generation of gait signature **HOFEI** are explained in detail.

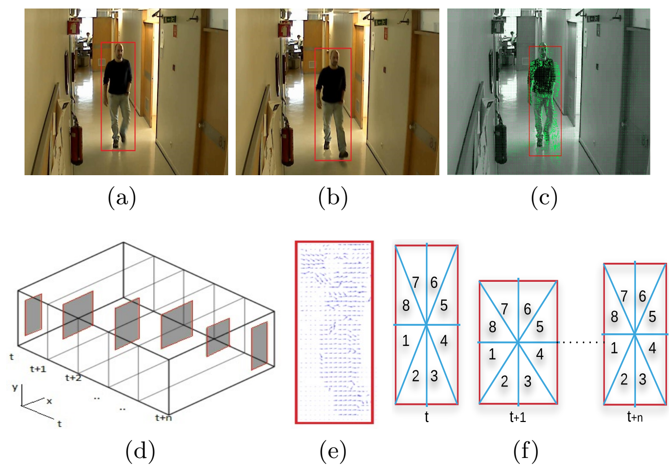


Fig. 2. Optic flow computation (*top*) and polar sampling scheme for the computation of Histogram of Optic Flow (HOF) (*bottom*); (a) and (b) show the adjacent video frames of gait. (c) shows the optic flow of the person computed. (d) The cuboid represents a slice of the video sequence spanning a gait cycle (n frames). The shaded regions are the Bounding Box (BB) of the person detected in each frame. (e) A sample of person's optic flow inside the BB. (f) Polar sampling of histogram of flow HOF in each of the images during a gait cycle, whose average results in the **HOFEI** gait signature.

Histogram of flow: We leverage the HOG encoding scheme mentioned in [13] on the human detection bounding boxes (BB). We provide 2 choices for the human detection BB: either by using the *'Ground truth'* annotations provided, or by using the *'Optic flow'* features to detect the moving section in the image. In this work, we use the default *'Ground truth'* BB. Then, the relative motion distributions of the peripheral human body parts - heads, arms and legs - are described within this BB. In contrast to the original HOG encoding scheme using grid of rectangular cells which overlap, here we use polar cells which better represent the spatial locations of limbs and head along time. Fig. 2(a)-(c) show the optic flow computation over a continuous walking sequence of frontal gait and Fig. 2(d)-(f) illustrate the sampling scheme of HOF.

When an optic flow image is provided, the first step is to divide it into cells according to the polar sampling strategy mentioned above, followed by the computation of histogram of flow orientation weighed by its magnitude. Let nR be the number of angular regions (i.e. cells) and nB be the number of bins that define each cell. Hence, the HOF features are parameterised as follows:

$$\mathbf{HOF^t} = \left[HOF^t{}_1 \cdots HOF^t{}_i \cdots HOF^t{}_{nR}\right] \in \mathbb{R}^{nR\times nB} \qquad (1)$$

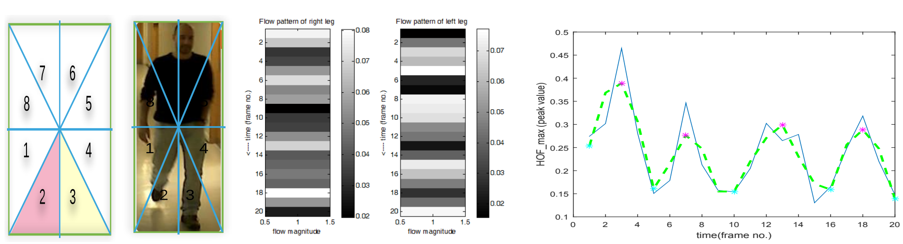
where $HOF^{\mathbf{t}}{}_i$ denotes the normalized HOF computed at cell i at frame t. Fig. 2 illustrates this, where a polar sampling scheme with 8 angular cells is shown (see Fig. 2 (f)). $\mathbf{HOF^t}$ is of dimension 64 with nR=*8* and nB=*8*. We compute the $\mathbf{HOF^t}$ for each frame throughout the video sequence S and the representation for the $\mathbf{HOF_S}$ is expressed as follows:

$$\mathbf{HOF_S} = \left[\mathbf{HOF^t} \mid \cdots \mid \mathbf{HOF^{t+\tau}}\right]^\top \in \mathbb{R}^{\tau \times nR \times nB.} \tag{2}$$

where τ denotes the number of frames in the video sequence.

Gait cycle estimation: Humans walk in a periodic fashion. In order to have coherent and reliable gait signature, it is necessary to estimate the gait features over a gait cycle, which acts as the functional unit of gait. A gait cycle is the time period or sequence of events/ movements during locomotion in which one foot contacts the ground to when that same foot again contacts the ground. In our proposal, the estimation of gait period is computed directly from the optic flow measured within the subjects' BB in raw images. This bypasses the computational load related to the traditional image segmentation and other image pre-processing steps in gait period computation.

We extract the periodicity encoded in the HOF sampling cells corresponding to the lower limbs. This choice is motivated by the fact that, the periodic information can reliably be obtained using the dynamic motion cues from the legs. The periodicity of right and left legs induces a similar periodic pattern in its corresponding optic flow. For instance, in a frontal gait sequence, as shown in Fig. 3(a) and Fig. 3(b), the polar sampling of cells 2 and 3, correspond to the location of legs in the image. More specifically, cell 2 and cell 3 correspond to the right and left leg, respectively.



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

Fig. 3. (a) Cells 2 and 3 represent sampling cells corresponding to the lower limbs. (b) Person's BB under polar sampling scheme depicts that the major area of motion pattern is described by the lower limbs cells. (c) Magnitude of the highest peak of the histogram of the right and left legs (cell 2 and 3) during a walking sequence. It is worth mentioning that the minimum value corresponding to the *stance* phase in one leg is accompanied with the maximum value corresponding to the *swing* phase in the other leg. (d) Estimation of gait period. The frames within two adjacent peaks (in *Magenta* markers) denote a gait cycle. (see online version for colours).

In order to estimate the gait period, we leverage the subset of histogram bins corresponding to cells 2 and 3, i.e., $HOF^t{}_2$ and $HOF^t{}_3$, which represents the lower limbs motion patterns, whose amplitude provides a good signal-to-noise ratio for detection. Then, we compute $\mathbf{HOF^t}$ throughout the video sequence corresponding to either $HOF^t{}_2$ or $HOF^t{}_3$ (since both are complementary). We can notice that this evolution undergoes a periodic pattern as depicted in Fig. 3(c),(d). Fig. 3(d) shows a periodic sinusoidal curve generated by plotting the HOF peaks of a single leg against the frame (as a function of time). A moving average filter is employed to smooth the obtained curve measurements (see green dashed curve), and the peaks of the filtered gait waveform allow us to identify the gait cycles. The frames between two consecutive peak points represent a gait cycle. Fig. 3(c) visualizes the simultaneous evolution of the HOF pattern peaks of both legs i.e., the amplitude of the highest peak in the histogram of each corresponding leg over time, are complementary since stride phase in one leg is accompanied by the stance phase in the other and vice versa.

Histogram of flow Energy Image: Based on the gait period estimation as well as the HOF features over video sequences, we compute HOF Energy Image (**HOFEI**), which is used as the key descriptor of each person. Inspired by the GEI scheme, HOF energy image is obtained by averaging the $\mathbf{HOF^t}$ representations over a full gait cycle, as follows:

$$\mathbf{HOFEI} = \frac{1}{t_2 - t_1} \sum_{t=t_1}^{t_2} \mathbf{HOF^t} \tag{3}$$

where t_1 and t_2 are the beginning and ending frame indices of a gait cycle and $\mathbf{HOF^t}$ is the histogram of flow of the person at time instant t, as defined in Equation (1). More intuitively, the **HOFEI** gait signature provides the relative motion of each body part with respect to the other, over a complete gait cycle.

4 Experimental Results

Experiments are conducted in two scenarios: Re-ID in controlled scenario vs Re-ID in uncontrolled (busy office) scenario. For the former, we use CASIA dataset B [14] which contains multiple videos of subjects including normal and apparel change (*bag*, *overcoat*) conditions, which makes it suitable for Re-ID scenario. Nevertheless, there is much control over the pose, illumination and background. Hence, it is also suitable to study the recognition of the subject under similar conditions. Hence, we conduct an extensive study on both the re-identification as well as recognition analysis in CASIA dataset. After this feasibility analysis, we apply our algorithm on a more realistic dataset (HDA Person dataset [17]) which is used for benchmarking video surveillance algorithms. In contrast to the CASIA dataset, HDA provides uncontrolled environment conditions (change in illumination, pose changes and occlusions) as well as lower frame rate (5fps) similar to a real world video surveillance system, which enables to conduct a Re-ID task in realistic scenario.

(a) (b) (c) (d) (e) (f)

Fig. 4. Some sample images from CASIA database and HDA database. (a)-(c) show various appearance (*'normal walk', 'carrying bag', 'wearing coat'*) conditions of subjects in CASIA dataset B. (d)-(e) depict the position of subjects in HDA dataset, at various distances $\mathbf{D_{far}}$, $\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$ respectively.

4.1 Re-ID in controlled scenario : CASIA dataset

CASIA is one of the largest databases available for gait recognition and related research [1]. Among the available four different datasets, we used Dataset B for our experiments. Dataset B [14] is a large multiview gait dataset collected indoors with 124 subjects and 13640 samples from 11 different views ranging from 0 to 180 degrees. In our experiments, we consider only the frontal walks (0 degrees), i.e., walking towards the camera. Database B contains three variations, namely view angle, clothing and carrying condition changes, and also presents the human silhouettes for each case. For each person, it contains 10 different video sequences (6 '*normal*' walk, 2 '*bag carrying*' walk and 2 '*overcoat wearing*' walk). Please refer to Fig. 4(a)- (c) for samples from CASIA dataset.

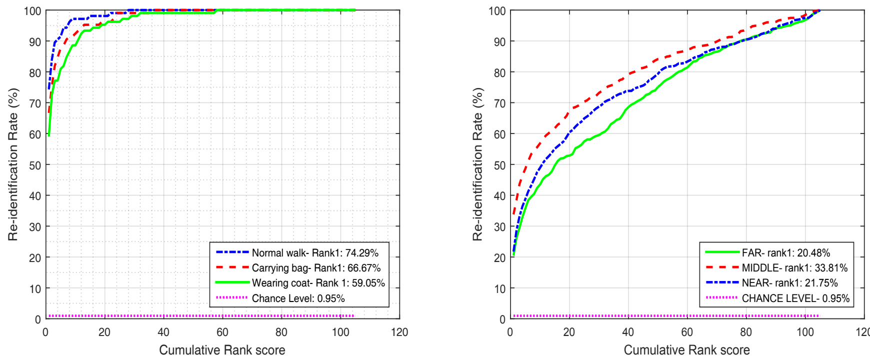
In order to evaluate the performance of our system towards long term re-identification, we conducted experiments not only under normal scenario, but also in the apparel change situations. For each of these experiments we considered 105 subjects, out of all the available 124 subjects. Videos in which the optical flow information can not be successfully extracted are excluded. For each of these available 1050 videos, we could get at least 3 gait cycles, in order to have enough data for training and testing. Then, for each gait cycle, the corresponding **HOFEI** is extracted. Regarding the dense optical flow computation, we use Stefan's implementation [2], which provides robust flow estimation by various methods of which, we select the Lucas- Kanade method [18].

Three main experiments are carried out in this dataset: The first is to verify the recognition performance under the same appearance and similar conditions. The second experiment is the Re-ID test conducted in order to verify performance under different appearance conditions. The third experiment is to test the influence of the distance of the subject in the performance of our system.

Experiment 1) Recognition in regular conditions: In this experiment, we only consider the '*normal*' type videos. The first four sequences are used for training and the last two are placed into the probe set. Then for each person's probe sequences, we compute the minimal Euclidean distance between any of the HOFEI descriptors in the probe and those of each person on the gallery. The minimal distance (most similar) gallery sequence is selected as the best matching

[1] `http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp`

[2] `http://www.mathworks.com/matlabcentral/fileexchange/44400-tutorial-and-toolbox-on\-real-time-optical-flow`



(a) (b)

Fig. 5. Re-ID results: (a) presents the CMC curves obtained for Experiment 1 and 2 for different probe cases viz., *normal* case, *bag* carrying case and *coat* wearing case. A chance level of 0.95% is also denoted in *magenta*. The Rank1 recognition achieved for normal, bag and overcoat are 74.29% (78 times the chance level), 66.67% (70 times the chance level), 59.05% (62 times the chance level) respectively. (b) depicts the CMC curves obtained for Experiment 3 at various distance probes viz., $\mathbf{D_{far}}$, $\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$. Middle case outperforms the others. (see online version for colours).

and sets the identity of the recognized person. The distances to the other persons in the gallery are used to provide a ranked list of identifications, for evaluation. Blue dotted curve in Fig. 5(a) shows the Correct Classification Rate (CCR) of this experiment, in terms of Cumulative Matching Characteristic (CMC) Curve. CMC curve shows, how often on average, the correct person ID is included in the best K matches against the training set for each probe. We could observe that a high CCR rate of 74.29% (78 times the chance level), has been achieved under the regular '*normal* walking' conditions.

A similar evaluation strategy, but using silhouette-based approach, had been carried out in [19] in all the view angles in CASIA dataset *'normal'* sequences. In order to conduct a reasonable comparison with our approach, we select the frontal view results they obtained by using various gait features (GEI, GHI, GMI, FDEI). Table 1 shows the comparison results of our strategy (HOFEI) against them in the ascending order. We can observe that CCR of our approach lies in between the others. The higher performance of FDEI and GEI could be attributed to their usage of segmented binary silhouettes and a more powerful classification method (HMM), whereas we use more flexible optic flow based features and a simpler classification method (Euclidean distance). Therefore, we consider the proposed feature competitive with the state-of-the-art, while more versatile. Since it does not require any pre-segmentation phase, it is easier to use in automatic RE-ID systems.

Experiment 2) Re-identification under change in appearance: In this experiment, we use all the 6 *'normal'* type videos for training, and *'wearing*

Table 1. Comparative analysis of our method against silhouette-based approaches in [19], for the frontal gait sequences of CASIA dataset B. The proposed method (HOFEI) is shown in bold letters.

GMI	GHI	**HOFEI**	GEI	FDEI
68.5%	71.8%	**74.3%**	91.1%	95.2%/100%

coat'/ *'carrying bag'* type videos for testing. In the '*bag*' case, we keep both the bag carrying sequences as the probe whereas all the 6 *'normal'* video sequences as the training set. A similar method is employed for *overcoat* scenario as well. Classification is similar to Experiment 1 (NN classifier+ Euclidean distance). The recognition results obtained are presented in Fig. 5(a). The apparel change recognition rates for bag (*red curve*) and coat (*green curve*) scenarios are 66.67% (70 times the chance level) and 59.05% (62 times the chance level) respectively. The lower CCR of *overcoat* condition could be ascribable to the global change in the flow features, whereas the *bag* either influence only a local flow change, or being occluded in some cases (occluded by hand, as in Fig. 4(b) or occluded while wearing as a backpack). No similar results in the appearance change conditions have been encountered in [19] for comparative evaluation.

Experiment 3) Variable distance to camera: Here, we are testing the robustness of the system when the subject is at different distance to the camera. In frontal sequences, the variability of the gait features with distance may have a significant impact on performance. Here we study the ability of the method in recognizing persons at a distance for which there are no gallery examples. We consider the '*normal'* type of videos for this experiment. In order to verify the impact of different distances, we conduct 3 case studies. In contrast to the previous experiments carried out on sets of videos, here we are conducting the analysis on each gait cycle instance. Performance will be lower than in the previous experiments, that used all gait cycles in the sequence for the classification. However, in this experiment we are not comparing absolute performance, but relative performance according to camera distance.

There are minimum of 3 gait cycles in each video sequence. In the first case study we keep all the *'normal'* gait cycle snippets seen at far distance $\mathbf{D_{far}}$ as the probe. The training set in this case is the *'normal'* $\mathbf{D_{middle}}$ and *'normal'* $\mathbf{D_{near}}$. Hence per person, we have 6 $\mathbf{D_{far}}$ probe and 12 training set ($\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$). Then, in the second case study, the $\mathbf{D_{middle}}$ is considered as the probe and $\mathbf{D_{far}}$ and $\mathbf{D_{near}}$ are kept as the training sets. Similarly, in the third case study, $\mathbf{D_{near}}$ videos are the probe and the others are kept as the training set. The Re-ID results are shown in Fig. 5(b). We can observe an expected drop in the CCR rate while conducting Re-ID with each gait cycle as the probe in this Experiment 3, rather than sets of videos as the probe in Experiment 1 & 2. $\mathbf{D_{middle}}$ case outperforms the other two cases, with 33.81% rate (35 times the chance level) whereas the far and near cases have recognition rates 20.48% (21 times the chance level) and 21.75% (22 times the chance level) respectively. In the case of $\mathbf{D_{middle}}$ as the probe, higher recognition rate could be attributed to the fact that, trained on the extreme ranges the classifier performs an interpolation when predicting values for the middle range, whereas in the other two $\mathbf{D_{far}}$ and $\mathbf{D_{near}}$ cases it has to extrapolate to one of the extremes, which is often an ill-posed operation.

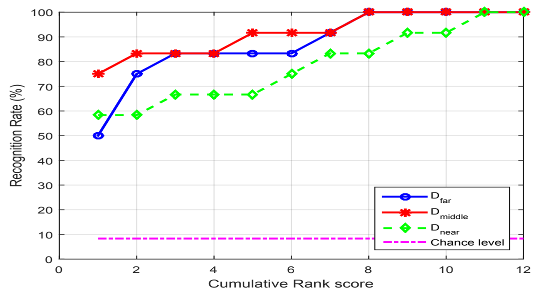
4.2 Re-ID in uncontrolled scenario: HDA Person Dataset

HDA dataset [17] [3], is a labelled image sequence data set for research on high-definition surveillance. The dataset was acquired from 13 indoor cameras distributed over three floors of one building, recording simultaneously for 30 minutes during a busy noon hour inside a University building. Among the 13, we select only a single camera recording (Camera19), containing frontal gait sequences. The camera has the VGA resolution of 640×480, with a frame rate of $5fps$. In this experiment we considered 12 people that crossed the whole corridor, and for which we could get at least 3 gait cycles in order to have enough data for training and testing. We collect each subject's walking frames, and from them we extract minimum three gait cycles and their corresponding **HOFEI**. Unlike the CASIA dataset, HDA is uncontrolled scenario since it contains varying illumination conditions during the walk, changing backgrounds, break points in between the walks (entry/ exit in the room along the way), occlusions by other person/ wall/ image boundary as well as self occlusions, also slight changes in the pose and limb movements during the walks.

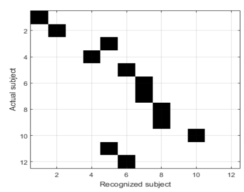
Due to the limitation of larger video sequences as well as varying appearance conditions per person, we exclude the CASIA counterpart Experiment 1 and Experiment 2 in HDA dataset. Here we only conduct Experiment 3, quite similar to the one carried out in CASIA dataset. We consider three cases in which we compute the **HOFEI** descriptor: far ($\mathbf{D_{far}}$), middle ($\mathbf{D_{middle}}$) and near ($\mathbf{D_{near}}$) sequences, as depicted in Fig. 4(d)-(f). Under this set of descriptors, we perform a leave-one-out evaluation where one set is kept as the probe and the other two sets as the gallery (i.e, a total of three trials). Thus, in each trial we have 24 training descriptors in the Gallery and we test against 12 test probes. Then, each test sample will search for the minimal Euclidean distance between itself and the gallery descriptors, under the nearest neighbor classification method. Fig. 6 demonstrates the recognition results in terms of Cumulative Matching Characteristic (CMC) curve as well as confusion matrix. The highest Rank-1 recognition rate of 75% (9 times the chance level) is achieved while using $\mathbf{D_{middle}}$ as the testing data. At the same time, the Rank-1 accuracy achieved by the test sets $\mathbf{D_{far}}$ and $\mathbf{D_{near}}$ are 50% and 58.33%, respectively.

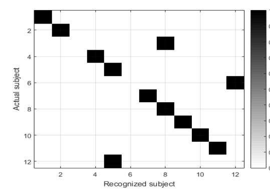
Referring to the CMC curve, another interesting observation is that the cumulative recognition rate improves drastically for both $\mathbf{D_{middle}}$ as well as $\mathbf{D_{far}}$ cases in comparison with $\mathbf{D_{near}}$, with the number of trials. This accentuates that gait sequences are better observed in far sequences than the closer ones since video frames close to the camera may undergo occlusions and thus result in poor encoding of the body flow features.

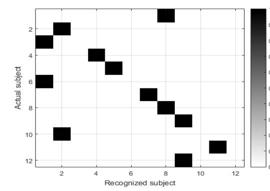
[3] `http://vislab.isr.ist.utl.pt/hda-dataset/`



(a)







(b) (c) (d)

Fig. 6. Recognition results: (a) presents the CMC curves obtained on 3 different probe cases viz., $\mathbf{D_{far}}$, $\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$. A chance level of 8.333% is also denoted in *magenta*. The Rank1 recognition achieved for $\mathbf{D_{far}}$, $\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$ are 50% (6 times the chance level), 75% (9 times the chance level), 58.33% (7 times the chance level) respectively. (b)-(d) show the confusion matrices for the 3 probe cases $\mathbf{D_{far}}$, $\mathbf{D_{middle}}$ and $\mathbf{D_{near}}$ respectively.

5 Conclusions & Future work

We analysed the potential of exploiting histogram of optic flow for frontal human gait analysis for Person re-identification. The main advantage of such a methodology is that no silhouette segmentation is required and thus can be facilitated towards online Re-ID system. A novel idea of flow based gait period estimation as well as a novel Histogram of Optic flow Energy Image (**HOFEI**) over the entire body are proposed in this work. We experimented the proposed framework upon a controlled benchmarking gait dataset (CASIA dataset) and a more unconstrained, thus harder, benchmarking video surveillance dataset (HDA Person dataset). We verified the effectiveness of the proposed method in both cases, under very different background clutter and sampling rates (25Hz in CASIA vs 5Hz in HDA). Extensive studies were conducted in CASIA dataset, i.e., regular case, change in appearance and influence of variable distance. Promising results were reported in each experiment, showing a Re-ID rate of 74.29% (78 times the chance level) in the *normal* scenario. In HDA dataset person Re-ID also a good performance rate of 75% (9 times the chance level) was reported, under different camera distance conditions. In future work, we plan to extrapolate this work towards pose invariant person re-identification scenario.

Acknowledgements. This work was supported by the FCT projects [UID/EEA/50009/2013], AHA CMUP-ERI/HCI/0046/2013 and FCT doctoral grant [SFRH/BD/97258/2013].

References

1. Stevenage, S.V., Nixon, M.S., Vince, K.: Visual analysis of gait as a cue to identity. Applied Cognitive Psychology,13, 513–526 (1999)
2. Lee, L., Grimson, W., Gait Analysis for Recognition and Classification. Proc. IEEE International Conference on Automatic Face and Gesture Recognition, (2002)
3. Cunado, D., Nixon, M., Carter, J.: Automatic Extraction and Description of Human Gait Models for Recognition Purposes. Computer Vision and Image Understanding, 90,1-41 (2003)
4. Han, J., Bhanu, B.: Individual recognition using gait energy image. Pattern Analysis and Machine Intelligence, 28, 316-322 (2006)
5. Chen, S., Gao, Y.: An Invariant Appearance Model for Gait Recognition. Proc. IEEE Conference on Multimedia and Expo, pp. 1375 - 1378 (2007)
6. Goffredo, M., Carter, J., Nixon, M.: Front-view Gait Recognition. International Conference on Biometrics: Theory, Applications and Systems, pp. 1-6 (2008)
7. Dalal, N., Triggs, B., Schmid, C., Human Detection Using Oriented Histograms of Flow and Appearance. ECCV, pp. 428-441 (2006)
8. Lam, T.H.W., Cheung, K.H., Liu, J.N.K.: Gait flow image: A silhouette-based gait representation for human identification. Pattern Recognition, 44, 973987 (2011)
9. Yang, Y., Tu, D., Li, G.: Gait Recognition Using Flow Histogram Energy Image. International Conference on Pattern Recognition, pp. 444 - 449 (2014)
10. Zhang, E., Zhao, Y., Xiong, W., Active energy image plus 2DLPP for gait recognition. Signal Processing, 90, 2295-2302 (2010)
11. Hofmann, M., Rigoll, G.: Improved gait recognition using gradient histogram energy image, ICIP, pp. 1389 - 1392 (2012)
12. Bashir, K., Xiang, T., Gong, S.: Gait recognition without subject cooperation. Pattern Recognition Letters, 31, 2052-2060 (2010)
13. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. Computer Vision and Pattern Recognition, pp. 886 - 893 (2005)
14. Yu, S., Tan, D., Tan, T.: A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. International Conference on Pattern Recognition, pp. 441 - 444 (2006)
15. Moreno, P., Figueira, D., Bernardino, A., Victor, J.S.: People and Mobile Robot Classification Through Spatio-Temporal Analysis of Optical Flow. International Journal of Pattern Recognition and Artificial Intelligence, 29 (2015)
16. Bashir, K., Xiang, T., Gong, S.: Gait Representation Using Flow Fields. British Machine Vision Conference (BMVC) (2009)
17. Nambiar, A., Taiana, M., Figueira, D., Nascimento, J.C., Bernardino, A.: A multi-camera video dataset for research on high-definition surveillance. International Journal of Machine Intelligence and Sensory Signal Processing, 1, 267-286 (2014)
18. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. Proceedings of Imaging Understanding Workshop, pp. 674-679 (1981)
19. Chen, C., Liang, J., Zhao, H., Hu, H., Tian, J.: Frame difference energy image for gait recognition with incomplete silhouettes. Pattern Recognition Letters, 30, 997-984, (2009)