Markerless Extraction of Gait Features using Haar-like Template for View-Invariant Biometrics

Imed Bouchrika [1,2]
Dept of Electronics & Computer Science[1]
University of Souk-Ahras
Souk-Ahras, Algeria, 41000
imed@imed.ws

Abdelhani Boukrouche [2]
PIMIS Research Laboratory[2]
University of Guelma
Guelma, Algeria, 24000
boukrouche.abdelhani@univ-guelma.dz

***Abstract*—Many research studies have recently shown the possibility of recognizing people by the way they walk i.e. gait. This research is mainly fuelled by the wide range of potential applications where gait biometrics could be useful as the case of visual smart surveillance and forensic systems. In this research paper, we present a Haar-like template for the temporal markerless extraction of gait features under various camera viewpoints. A markerless model-based method whereby angular model templates describing the human motion are employed to guide the extraction process. Gait features consist of the angular measurements for the lower legs in addition to the spatial displacement of the human body. To further refine gait features based on their discriminatory potency, a feature selection algorithm is applied using a newly proposed validation-criterion based on the proximity of neighbours belonging to the same class. Experimental results revealed that gait angular measurements derived from the joint motions can achieve a correct classification rate of 73.6% after applying a rectification process back into the sagittal plane.**

***Keywords*-Gait; Biometrics; Gait Recognition**

I. Introduction

Much research within the computer vision community is directed into the analysis of articulated objects with immerse attention on the analysis of human motion. Such research is fuelled by the wide range of applications where human motion analysis can be deployed such as biometrics, smart surveillance, human computer interaction and athletic performance analysis. A vision-based system for human motion analysis consists of three main consecutive stages: Detection, Tracking and Perception of either activity or identity. The majority of systems being used for motion analysis are commercially available with the main disadvantage of being marker-based. For marker-based solutions, reflective markers or sensors are attached at a number of key positions on the human body to capture their movement. However, most applications including automated surveillance require the use of fully automated markerless vision system in order to derive the joints positions. On the other hand, automated extraction of the joints coordinates is proven to be a difficult process as the human body encompasses a wide range of possible articulation due to its highly flexible structure and to self occlusion. Clothing type, segmentation errors and different viewpoints exacerbate further challenges for accurate localization of the joints.

The use of walking pattern for people recognition in surveillance systems has attracted recently considerable interest. The suitability of gait identification for surveillance applications emerge from the fact that gait i.e. the way of walking, can be perceived from a distance as well as its non-invasive and less-intrusive nature [1]. In fact, early studies by Johansson [2] which is conducted on analysing the human motion using Moving Light Displays have shown that an observer can recognise various types of activities performed by an actor based on seeing only the joint motions. Moreover, the observer can infer the gender of the actor and even further recognise the persons identity if they are familiar with their walking pattern. Early research work by Murray [3] had also confirmed that gait could be a potential biometrics for people identification. In their study, a total of 20 gait features which includes ankle rotation, trunk horizontal and vertical displacement are found to render uniquely the gait biometric signature for every person.

As a number of research studies confirmed the potentiality of using gait biometrics for real surveillance and forensic applications [1] based on silhouette-based methods, we explore in this research a model-based markerless approach for extracting gait features from walking subjects under different camera viewpoints. The method derives the joints positions of walking subject from uncalibrated single cameras through the use of a haar-like template matching approach. The gait signature is being constructed based on people being recorded from a sagittal view using the floating feature selection algorithm [4], [5]. For different cases of viewpoint, the gait biometric signature is reconstructed onto the sagittal plane through a rectification process [6] which transforms the gait angular data from a particular viewpoint back to the normal plane. This way, it is possible to recognize walking people seen in one camera view from data derived from a different viewpoint.

This paper is organized as follows. The next section outlines the previous approaches for markerless extraction of gait biometric features including mainly model-based and silhouette-based methods. The theoretical description of the presented markerless method for extracting and reconstructing a gait-based biometric signature is detailed in Section 3. The following section introduces the experimental results.

978-1-4799-5906-8/14/$31.00 ©2014 IEEE

II. Related Work

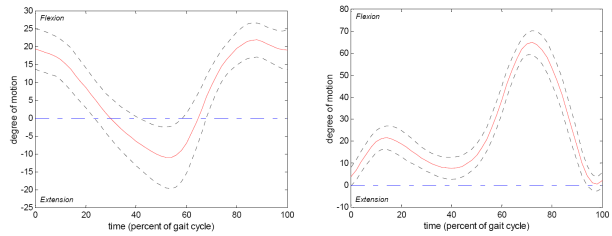
Since human motion analysis is one of the most active and challenging research topics in computer vision, many research studies have aimed to build a robust vision system capable of overcoming the difficulties imposed during the extraction and tracking of gait-based features. Various methods are surveyed by [7]. Two approaches are being used for human motion anaylsis: model-based and non-model based methods. For the first category, a priori model is initially established in order to match real images to this predefined template, and thereafter extracting the corresponding features once the best match is deduced. The stick and volumetric models [8] were the most commonly used methods for human motion analysis. Guo *et al* [9] represented the human body structure in the silhouette by a stick figure model which is composed of ten articulated sticks with six joints. Rohr [10] proposed a volumetric model through the use 14 elliptical cylinders in order to model the human body. Bouchrika *et al.* [11] proposed a model-based approach for extracting the positions of the joints through the use of Elliptic Fourier Descriptors and evidence gathering algorithm. A parameterized representation of the spatial templates of the joints is presented accounting for the different shape transformation including scaling, rotation and translation. The model-based approach is the most popular method being used for human motion analysis due to its advantages [12]. It can extract detailed and accurate motion data, as well as having the capability to cope well with occlusion and self-occlusion.

For the non-model based method, feature correspondence between successive frames is based upon prediction, velocity, shape, texture and colour. For related work concerning the use of model-free methods for gait biometrics, Little *et al.* [13] constructed gait signature using dense optical flow achieving a correct classification rate of 95% tested on a database of 6 subjects. Huang *et al.* [14] introduced a new approach for gait identification based on statistical analysis. The biometric signature for every person is derived by combining the eigen-space transformation (EST) with the canonical space transformation (CST). These projections are being applied for the sake to reduce the dimensionality of the silhouette data and increase the inter-class separability respectively. Recently, Bashir *et al* [15] proposed the Gait Entropy Image which encodes into a single image the randomness of pixel values in the silhouette images over a complete gait cycle. Kusakunniran [16], [17] proposed a silhouette-based approach for feature extraction to view-invariant gait biometrics.

III. Methods

A. Markerless Feature Extraction

For recovering the lower limbs positions, a *Haar-like* feature template matching process is proposed for the localization of the legs using motion information for a single walking subject. The approach does not depend on background subtraction for the derivation of gait features. This is because it is computationally expensive and complex to deploy background subtraction for real-time surveillance applications due the process of updating the background model which is influenced by a number of factors such as background clutter weather, conditions and other outdoor environmental factors.



(a) (b)

Figure 1. Gait Angular Motion: (a) Hip. (b) Knee.

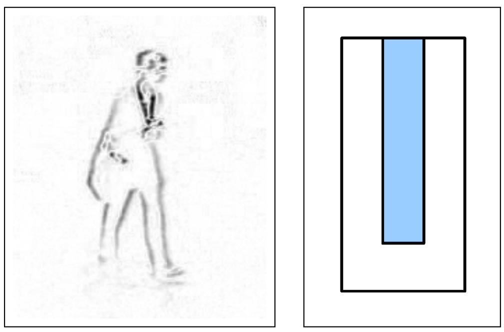
For the marker-less extraction of human gait feature, motion models are derived based on medical data that describes the angular motion for both the knee and hip at different phases of the gait cycle as shown in Figure (1). The dashed lines represent the maximal and minimal points for the angular data of the human gait based on medical studies. A full gait cycle is defined as the interval between two successive instances of initial foot-to-ground contact of the same foot [18]. The hip initially bends by approximately 20° throughout the terminal stance phase and it extends afterwards until it reaches approximately 10° during the stance phase. Throughout the pre-swing and most of swing stage, the hip flexes to nearly 20 degrees , and then it starts to extend just before the next initial contact with the ground. As shown in Figure (1), the knee is almost fully extended during the first part of the mid-stance, it gradually begins to flex down to reach about 20 degrees. The knee extends almost fully and then flexes to reach 40 degrees through the pre-swing stage. After the foot leaves the ground, the knee flexes to a maximum of 60 to 70 degrees during mid-swing, then it extends again to prepare for the next initial contact to the ground.

The proposed approach constructs initially the motion map image based on the change detection for the inter-frame difference. The only constraint of using frame differencing is that the camera must be in a stationary position. Moving pixels of a walking person across consecutive frames are detected with the momentum to derive better edge features. The motion map M_t at frame t is approximated as the difference of two consecutive frames I_t and I_{t+1} as shown in the following equation:

$$M_t = \|I_t - I_{t+1}\| \tag{1}$$

Thresholding is applied on the resulting image M_t in order to reduce the artifacts. A sample motion image is illustrated in Figure (2) for a walking subject.

A *Haar-like* template [19] is being utilized for the localization of the gait features due to their robust and fast performance in real-time systems from object recognition to pedestrian detection. The template is shown in Figure (2) which is based on the outlier of the lower part of the leg. Let p_t^{ankle} is the candidate position of the ankle at t^{th} frame. To localize the ankle position, different templates are derived accounting for the different rotation and translation parameters. The templates are superimposed against the motion map at the candidate point p estimating the match score S as given in Equation (2). The position of the ankle is obtained once a maximum value for S is achieved. The similarity score signifies how well the matching template for the lower leg is superimposed on top of the motion map frame. It is estimated as the total sum of larger values inside the haar template divided by the accumulated lower values within the region that are less than a certain threshold which is experimentally set to $\tau = 20$.



(a) (b)

Figure 2. Markerless Gait Feature Extration: (a) Motion Image. (b) Haar-based Matching Template.

$$S(x, y, \alpha) = \frac{\sum_{i \in P_{x,y,\alpha}} P_{x,y,\alpha}(i) \times Z(P_{x,y,\alpha}(i))}{\sum_{i \in P_{x,y,\alpha}} P_{x,y,\alpha}(i) \times |1 - Z(P_{x,y,\alpha}(i))|} \tag{2}$$

such that α is the rotation angle and Z is given as :

$$Z(i) = \begin{cases} 1 & \text{if } i > \tau \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In contrast to using a per-frame algorithm for pose recovery [20], a frame-to-frame approach is being adopted for the markerless extraction procedure whereby results from the previous frame are exploited to guide the matching process in subsequent frames. Initially, the search process is performed over the whole motion region to find the best match for the leg. In order to limit the search space for a candidate point and refine further the extraction accuracy, kinematical and anthropometric constraints [21] including spatial as well as angular data derived from the gait motion models described in earlier section are imposed during the extraction process. For example, during the striking phase, one of the legs will be almost stabilised at the same position and therefore the ankle spatial movement is enclosed within a smaller region whilst the rotation parameter α is usually constrained within a specific range depending on the phase of the gait cycle. The pose estimation algorithm for the lower legs partially depends on the anatomical proportions of the human body segments, based on medical results of anatomical studies [22] for a person of height H:

$$\begin{array}{l} y'_{hip} = \min(\mathbf{y}_{sil}) + 0.5 \cdot H \\ y'_{knee} = \min(\mathbf{y}_{sil}) + 0.75 \cdot H \\ y'_{ankle} = \min(\mathbf{y}_{sil}) + 0.90 \cdot H \end{array} \tag{4}$$

During the double-support phase of the gait cycle where the legs overlap, it is difficult to extract the lower limbs accurately because of the self-occlusion due to the overlap. Therefore, the matching process is applied for the striking leg using kinematic gait constraints that can assist with the localization. The swinging leg is not dealt with during the overlap phase due to self-occlusion. The overlapping starts when the Euclidean distance between the two ankles is less than a certain threshold which is related to the person height. The extraction of the swinging leg during the overlap is resumed after a certain number of frames which is defined from the average gait cycle model. Experimentally, the number of frames is set to 6 for a video recorded with a frame rate of 25 frames/seconds. In order to extract the joints positions as well the angular measurements α when the legs overlap, a 3^{rd} order polynomial interpolation process is performed.

The orientation of the upper legs is extracted at each image $\mathbf{T} = [t_1, t_2, \ldots, t_\varphi, \ldots, t_F]$ with a coarse to fine estimation procedure where at first, the hips position is performed with

$$\begin{cases} x'_{hip\ell} = \frac{1}{P} \cdot \sum_{j=1}^{P} \widetilde{x}_j + (2\ell - 3) \cdot H \cdot \ \mu \cdot \ 10^{-3} \\ y'_{hip\ell} = y'_{hip} \cdot (2\ell - 3) \cdot \left(\frac{\widetilde{x}_P - \widetilde{x}_1}{2}\right) \cdot \sin(0.3 \cdot \mu) \end{cases} \tag{5}$$

such that $\widetilde{\mathbf{X}} = [\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_j, \ldots, \widetilde{x}_P]$ is the subset P ($P \leq MaxX$) horizontal coordinates from extracted motion region for the subject S [21].
The equation 5 puts in relationship the horizontal hip position and walking direction μ, computed with respect to the horizontal axes of the image reference. These relationships are obtained with regression analysis of the 3D Georgia Tech motion capture data. μ is approximated as the angle of inclination of the walking straight line which is deduced from the detected heelstrikes on the ground.

B. Biometric Gait Signature

Having extracted the joints' position of a walking subject, the subsequent phase is to rectify the extracted data back into the sagittal plane. The method proposed by the authors in [6] is based on four main assumptions which are: human gait pattern is periodic; people tend to walk along a straight line; the distances between the joints are constant; and the articulation of every joint is occurring approximately on the same plane. Thus, multiple periods of walking motion does appear analogous to a

single gait period seen from various cameras related by linear translation and the trajectories of the joints lie in an auto-epipolar configuration.
If $\mathbf{j}_i^{\ell}$ is the set of joints coordinates for each of the legs $\ell = \{1, 2\}$ at the i^{th} frame within the image reference system, the relationship between $\mathbf{j}_i^{\ell}$ and the corresponding trajectories in the worldspace is $\mathbf{j}_i^{\ell} \times \mathbf{P}_i \cdot \mathbf{J}^{\ell} = 0$, where $\mathbf{P}_i = \left[\mathbf{R}_{\mathbf{e}}^T, -i\mathbf{e}_0\right]$ and $\mathbf{R}_{\mathbf{e}}^T$ is the rotation matrix for aligning the epipolar vector $\mathbf{e}_0$ with the horizontal axis X. Then,

$$\mathbf{j}_i^{\ell} = \mathbf{P}_i \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_{\mathbf{V}}^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{H}_{\mathbf{V}} \end{pmatrix} = \mathbf{H} \cdot \mathbf{J}^{\ell} \qquad (6)$$

The limb plane transformation matrix is expressed with $\mathbf{H}_{\mathbf{V}}$ so that the two cross section plane lines are centred as well normalised with respect to the axes Y and Z and parallel with Y. As the lengths of the articulated limbs $\mathbf{D}_{\ell}^2 = \Delta\mathbf{j}_i^{\ell\mathbf{T}} \Delta\mathbf{j}_i^{\ell}$ are assumed constant over all consecutive frames, the pose difference vectors for the legs segments at two adjacent frames, $\Delta\mathbf{j}_i^{\ell}$ and $\Delta\mathbf{j}_{i+1}^{\ell}$, are related by

$$\Delta\mathbf{j}_i^{\ell\mathbf{T}} \cdot \mathbf{H}^{\mathbf{T}} \cdot \mathbf{H} \cdot \Delta\mathbf{j}_i^{\ell} = \Delta\mathbf{j}_{i+1}^{\ell\mathbf{T}} \cdot \mathbf{H}^{\mathbf{T}} \cdot \mathbf{H} \cdot \Delta\mathbf{j}_{i+1}^{\ell} \qquad (7)$$

After recovering the fronto-parallel structure of subject gait, the representation of the leg joints function $\left[\mathbf{J}_x^{\ell}(t), \mathbf{J}_y^{\ell}(t)\right]$ is found by fitting a modified Fourier series to the data with fixed fundamental frequency f_0 and period T:

$$\mathbf{J}_x^{\ell}(t) = v_x t + \sum_{k=1}^{n} A_k \cos\left(2\pi k f_0 \left(t + \frac{(\ell - 1)\,T}{2}\right) + \phi_k\right) + \mathbf{J}_{x0}^{\ell} \qquad (8)$$

analogously for $\mathbf{J}_y^{\ell}(t)$. Thus, the projection of the leg joints on the lateral plane is obtained with an optimized procedure in the following way

$$\breve{\mathbf{J}}^{\ell}(t) = \left[\begin{array}{ccc} h_1 & h_2 & h_3 \end{array}\right] g\left(t + \frac{(\ell - 1)\,T}{2} : f_0, \mathbf{D}_{\ell}, v_x, v_y, F\right) \qquad (9)$$

where $g(t)$ is the bilateral Fourier series function with coefficients F and h are the values of the inverse normalization transform matrix.
Therefore, starting from a video sequence from a single camera and without any calibration, the proposed markerless system, in junction with [6], estimates the gait parameters projected on the lateral plane.

Feature selection is a critical task for most of the pattern recognition problems. This process is aimed to derive as many discriminative characteristics as possible whilst removing the redundant and irrelevant features which may degrade the classification rate. It is practically infeasible to run an exhaustive search for all the possible combinations of features in order to derive the optimal subset for recognition due to the high dimensionality of the features space. For this reason, we utilised the Adaptive Sequential Forward Floating Selection (ASFFS) search algorithm.

The feature selection procedure is purely based on an evaluation function that determines the usefulness or discriminativeness of each feature in order to obtain the ideal subset of features for the classification process. A number of approaches [7] are based mainly on statistical metric measures which use the scatter or distribution of the training samples in the feature space such as the Bhattacharyya metric. Although, statistical methods have the benefit of low-cost implementation, they have been proved to offer poor estimate of the classification rate because of their independence from the final classifier. The proposed algorithm uses a validation-based evaluation criterion to derive the subset of features that minimises the classification errors as well as ensure good separation between the different clusters. As opposed to the voting procedure employed within the *KNN* classifier, the evaluation criterion uses different coefficients w to further weigh more importance of nearest neighbours. The probability value for a candidate s_c to belong to a given class c is given in the following equation (10):

$$f(s_c) = \frac{\sum_{i=1}^{N_c - 1} z_i w_i}{\sum_{i=1}^{N_c - 1} w_i} \qquad (10)$$

where N_c is the number of instances belonging to class c, and the coefficient w_i for the i^{th} nearest candidate point is inversely related to proximity as:

$$w_i = (N_c - i)^2 \qquad (11)$$

The value of z_i is given by as:

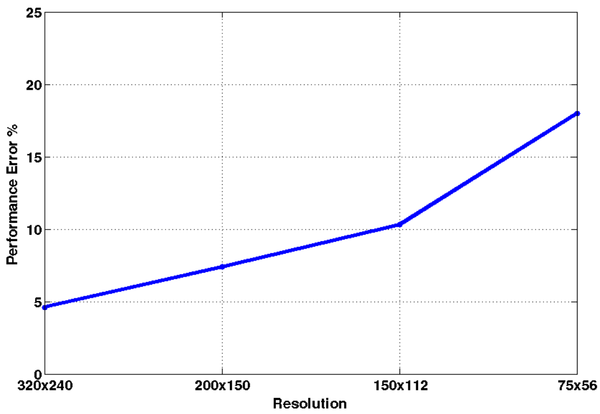
$$z_i = \begin{cases} 1 & \text{if } nearest(s_c, i) \in c \\ 0 & \text{otherwise} \end{cases} \qquad (12)$$

such that the $nearest(s_c, i)$ method returns the i^{th} nearest candidate point to the sample s_c. The Euclidean distance is used to find the nearest neighbours.

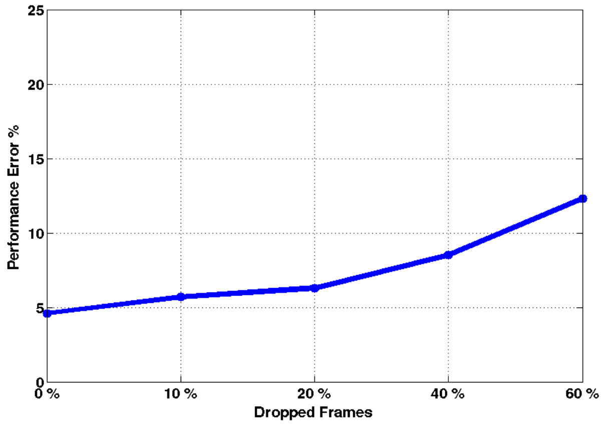
IV. Experimental Results

The performance of the model-based method proposed for recovering the joints' positions of walking subjects is assessed on videos with small resolution. A set of 18 sequences are taken from the CASIA-B gait database [23] and used for the evaluation of gait identification across different viewpoints. The collected dataset consists of 6 different viewpoints (36^o, 54^o, 72^o, 90^o, 108^o, 126^o) with 3 sequences for every camera view. Manual labelling of the videos is performed to collect ground truth data. Figure 3 presents the performance error of the algorithm for the localisation of the joints at various resolutions. The Euclidean distances between the extracted joints and manually labelled data (i.e., ground truth data) are used to approximate the performance error which is estimated as the average of the distances normalised to the subjects' heights. The resolution of the video sequences are reduced gradually from an original size of 320×240 pixels with the aspect ratio kept constant. The algorithm is still able to derive the joints of the legs with an acceptable accuracy level for a resolution of 144×180 with a performance error of 14.3%. However, the algorithm does not cope well when the video resolution is further reduced to 75×56 pixels.

Furthermore, a number of experiments are conducted using the same video dataset to investigate the algorithm potentials for handling occlusion. For the case of full occlusion, the performance error is simulated by dropping



(a)



(b)

Figure 3. Performance Analysis for The Joint Extraction.

a number of frames from every 25 frames (25 is the original frame rate) of the video sequences which is equivalent to changing the frame rates. Figure (3) depicts the performance error which is computing by dropping the frame rates for videos resized at fixed resolution of 320x240 pixels. From this analysis, the extraction performance is observed to have little effect even for the case when dropping the frame rate to reach 60% as the extraction method predicts the joint trajectories for the missing frames using the temporal and spatial templates. This is one of the merit for using model-based methods as to their potentials of handling self-occlusion robustly.

To assess the efficiency of the presented approach for tracking walking people in multi-camera surveillance scenarios, the system is evaluated on a variety of cases and conditions. The Casia-B database is set as the real test-bed of the view-invariant markerless method since every single persons walking pattern is simultaneously recorded from a large number of different views. This establishes confidence in assessing the viewpoint invariant approach. The markerless feature extraction method is applied on

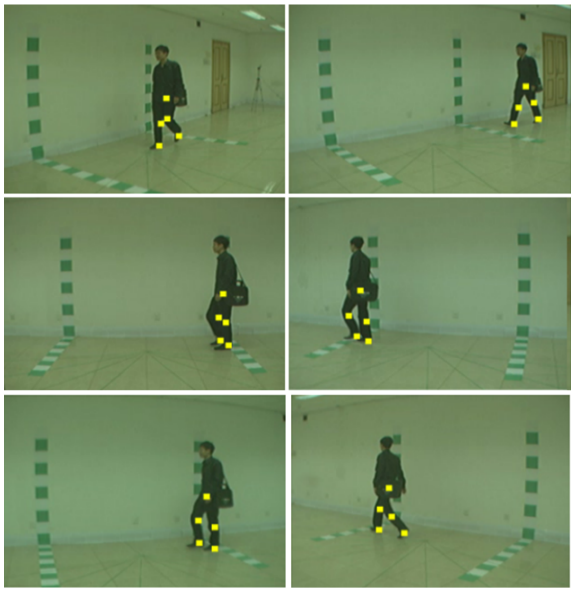


Figure 4. Markerless Extraction of Gait Features on CASIA-B dataset.

the CASIA-B gait database for a collection of videos consisting of 2270 sequences for 65 walking people with an average of 6 different sessions for every viewpoint per candidate. In this study, we have considered taking only 6 different camera orientations which are (36^o, 54^o, 72^o, 90^o, 108^o, 126^o). The pose for the legs has been estimated on a frame by frame basis. The hip and knee angles have been extracted for each viewpoint and for each person. Figure 4 illustrates an example of for the extraction process of gait features for the 6 different orientations.

To analyse the impact of viewpoint as a covariate factor for the performance of gait biometrics, we have measured initially the correct classification rates for the non-rectified gait features. Afterwards, the viewpoint rectification process is invoked to project the gait angular features back into the sagittal view (i.e. lateral view). The Cumulative Match Score (CMS) metric is used in this evaluation. The measure assesses the ranking potency of the biometric system by giving a list of scores that indicates the probabilities that the true identification trial for a given test sample is within the top n matched class labels. For this study, correct classification rates of 73.6% and 100% are attained for 1^{st} and 11^{th} rank respectively for the rectified gait features. However, lower recognition rates of 32.0% and 67.5% are achieved for raw un-rectified data for the same ranks. The CMS score at 1st rank is considered the correct classification rate. Figure 5 depicts the CMS plot of gait recognition for both rectified and unrectified gait features at different ranks.

V. Conclusion

We have taken an important step in deploying a markerless model-based extraction method for gait biometrics. We present a new approach for people identification from different uncalibrated camera viewpoints based on gait analysis. Identification signature is derived from gait kine-

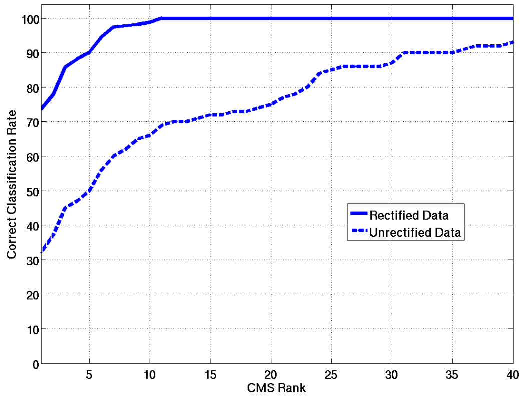


Figure 5. CMS curve of the gait identification for the rectified and unrectified data.

matical data that are obtained using a marker-less feature extraction algorithm which is based on Haar-like template matching. Experimental results revealed the potential of our method to recognize walking people over different views using the marker-less pose recovery with an attained correct classication rate of of 73.6%. This is to conclude that identity recognition can be achieved by gait analysis and we have encouraging results using the marker-less feature extraction and rectification approach. This is an crucial milestone in translating gait biometrics into real cases whereby calibration parameters are not an option to be obtained such as in forensic, security and surveillance applications.

REFERENCES

[1] M. S. Nixon, I. Bouchrika, B. Arbab-Zavar, and J. N. Carter, "On use of biometrics in forensics: gait and ear," in *European Signal Processing Conference*, 2010.

[2] G. Johansson, "Visual Perception of Biological Motion and a Model for its Analysis," *Perception and Psychophysics*, vol. 14, pp. 201–211, 1973.

[3] M. P. Murray, "Gait as a Total Pattern of Movement." *American Journal of Physical Medicine*, vol. 46, no. 1, pp. 290–333, 1967.

[4] I. Bouchrika, "Gait analysis and recognition for automated visual surveillance," Ph.D. dissertation, University of Southampton, 2008.

[5] I. Bouchrika and M. S. Nixon, "Gait-based pedestrian detection for automated surveillance," in *International conference on computer vision systems*, 2007.

[6] N. Spencer and J. N. Carter, "Towards Pose Invariant Gait Reconstruction," *In Proceedings of IEEE International Conference on Image Processing*, vol. 3, 2005.

[7] M. S. Nixon, T. N. Tan, and R. Chellappa, *Human Identification Based on Gait*. Springer-Verlag New York, Inc. Secaucus, NJ, USA, 2005.

[8] J. H. Yoo, M. S. Nixon, and C. J. Harris, "Extraction and description of moving human body by periodic motion analysis," *Proceedings of Proc. ISCA 17th International Conference on Computers and Their Applications*, pp. pp. 110–113, 2002.

[9] Y. Guo, G. Xu, and S. Tsuji, "Understanding human motion patterns," in *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, vol. 2, 1994, pp. pp. 325–329.

[10] K. Rohr, "Towards model-based recognition of human movements in image sequences," *CVGIP: Image Understanding*, vol. 74, no. 1, pp. pp. 94–115, 1994.

[11] I. Bouchrika and M. S. Nixon, "Markerless Feature Extraction for Gait Analysis," *In Proceedings of IEEE SMC Chapter Conference on Advanced in Cybernetic Systems*, pp. 55–60, 2006.

[12] N. Huazhong, T. Tan, L. Wang, and W. Hu, "People tracking based on motion model and motion constraints with automatic initialization," *Pattern Recognition*, vol. 37, no. 7, pp. pp. 1423–1440, 2004.

[13] J. Little and J. Boyd, "Recognizing People by their Gait: the Shape of Motion," *Videre: Journal of Computer Vision Research*, vol. 1, no. 2, pp. 2–32, 1998.

[14] P. S. Huang, C. J. Harris, and M. S. Nixon, "Human Gait Recognition in Canonical Space using Temporal Templates," *In IEE Proceedings of Vision, Image and Signal Processing*, vol. 146, no. 2, pp. 93–100, 1999.

[15] K. Bashir, T. Xiang, and S. Gong, "Gait recognition without subject cooperation," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 2052–2060, 2010.

[16] W. Kusakunniran, Q. Wu, J. Zhang, Y. Ma, and H. Li, "A new view-invariant feature for cross-view gait recognition," *Information Forensics and Security, IEEE Transactions on*, vol. 8, no. 10, pp. 1642–1653, 2013.

[17] W. Kusakunniran, "Recognizing gaits on spatio-temporal feature domain," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 9, pp. 1416–1423, Sept 2014.

[18] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic Extraction and Description of Human Gait Models for Recognition Purposes," *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1–41, 2003.

[19] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, "Pedestrian detection using wavelet templates," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 193–199.

[20] G. Mori and J. Malik, "Recovering 3d human body configurations using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, pp. 1052–1062, 2006.

[21] M. Goffredo, I. Bouchrika, J. N. Carter, and M. S. Nixon, "Performance analysis for automated gait extraction and recognition in multi-camera surveillance," *Multimedia Tools and Applications*, vol. 50, no. 1, pp. 75–94, 2010.

[22] W. T. Dempster and G. R. Gaughran, "Properties of body segments based on size and weight," *American Journal of Anatomy*, vol. 120, no. 1, pp. 33–54, 1967.

[23] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," in *Proc. of Int. Conference on Pattern Recognition*, vol. 4, 2006, pp. 441–444.