2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

View-Invariant Gait Recognition via Deterministic Learning

Wei Zeng and Cong Wang

***Abstract*—In this paper, we present a new method to eliminate the effect of view angle for efficient gait recognition via deterministic learning theory. The width of the binarized silhouette models the periodic deformation of human gait silhouettes. It captures the spatio-temporal characteristics of each individual, represents the dynamics of gait motion, and can sensitively reflect the variance between gait patterns across various views. The gait recognition approach consists of two phases: a training phase and a recognition phase. In the training phase, the gait dynamics underlying different individuals' gaits from different view angles are locally accurately approximated by radial basis function (RBF) neural networks. The obtained knowledge of approximated gait dynamics is stored in constant RBF networks. In order to address the problem of view change no matter the variation is small or significantly large, the training patters from different views constitute a uniform training dataset containing all kinds of gait dynamics of each individual observed across various views. In the recognition phase, a bank of dynamical estimators is constructed for all the training gait patterns. Prior knowledge of human gait dynamics represented by the constant RBF networks is embedded in the estimators. By comparing the set of estimators with a test gait pattern whose view pattern contained in the prior training dataset, a set of recognition errors are generated. The average L_1 norms of the errors are taken as the similarity measure between the dynamics of the training gait patterns and the dynamics of the test gait pattern. Finally, comprehensive experiments are carried out on the CASIA-B and CMU gait databases to demonstrate the effectiveness of the proposed approach.**

I. Introduction

Gait as a biometric has recently gained considerable attention because of its unobtrusiveness and gait information can be captured at a distance from a camera. It has the potential to be applied in various areas of real-world applications such as surveillance, health care, entertainment, access control and border control. However, there are a number of covariate factors that affect gait recognition performance, such as lighting condition, clothes, carrying status, shoe type, walking speed, view angle and so on [1]. Hence, in gait recognition, one important requirement is of robustness to these variations. Among these factors, view angle is one of the most important factors which heavily affects gait recognition performance [2]. The difficulties lie in that gait appearance changes due to the variation of views or walking directions, and it is impossible to expect all the subjects to walk in a particular direction.

To handle the variation of gait sequences caused by different view angles, many researchers have proposed several methods to address this problem. They can be roughly divided into the following categories: (1) extracting view-invariant gait features; (2) projecting gait feature from one view angle to the other by using view transformation; and (3) synthesizing view angle based on a three-dimensional (3D) model.

In the first category, [3] synthesized a side view from any other arbitrary view using a single camera. Two methods, namely the perspective projection model and the optical flow based structure from motion equations, respectively, were adopted by working with calibrated single-camera system. However, this method assumed the subjects to be far away from the camera and could not cope with large variations in view angle.

In the second category, approaches aim to learn a mapping relationship between gait features of the same subject observed across views [4, 5]. When matching gait sequences from different views, the gait features are mapped into the same view before a distance measure is computed for matching.

In the third category, a 3D model or visual hull of the walking body is usually generated for recognition under multi-camera system. For example, in [6], a 3D human model is set up from video sequences captured by multiple cameras. The motion trajectories of lower limbs extracted from the 3D model are used as dynamic features, and linear time normalization is exploited for matching and recognition.

Gait is a dynamic shape model such that it varies in pose and size throughout a walking cycle [7]. In practice, the gait shape of an individual can be easily altered by many factors, particularly by the change of views. Individuals can walk at various directions to the cameras in any real-world situations, and gait shape changes nonlinearly according to views. This will lead to significant changes to walking patterns and generate difficulties for gait recognition.

In our previous studies, gait dynamics represented by suitable periodic features, such as lower limb joint angles and angular velocities, will be approximated by radial basis function (RBF) neural networks [8, 9]. The difference of gait dynamics between different individuals during walking can be used for gait recognition. Following this idea, we continue to search for periodic gait features that constitute

Wei Zeng is with the School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou 510640, China, and is with the School of Physics and Mechanical & Electrical Engineering, Longyan University, Longyan 364000, China (email: zw0597@126.com). Cong Wang is with the School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China (email: wangcong@scut.edu.cn).

This work was supported by the National Natural Science Foundation of China (Grant No. 61304084), by China Postdoctoral Science Foundation (Grant No. 2013M531851), by Natural Science Foundation of Guangdong Province (Grant No. S2013040016859), by the Educational and Scientific Research Project for Middle-aged and Young Teachers of Fujian Province (Grant No. JA13304), and by the Fundamental Research Funds for the Central Universities, SCUT (Grant No. 2013ZM0089).

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

the gait dynamics and can reflect the variance between gait patterns across various views. Recent gait research revealed that silhouette cues, such as the width of the outer contour and silhouette area, play a primary role in gait recognition. This finding enlightens us to extract silhouette features for gait identification.

II. PRELIMINARIES AND PROBLEM FORMULATION

Consider a general nonlinear human gait dynamical system in the following form:

$$\dot{x} = F(x;p), \quad x(t_0) = x_0 \tag{1}$$

where $x = [x_1, \ldots, x_n]^T \in R^n$ is the state of the system representing the gait features, p is a constant vector of system parameters (different p will in general generate different gait patterns under different view angles). $F(x;p) = [f_1(x;p), \ldots, f_n(x;p)]^T$ is a smooth but unknown nonlinear vector field.

Assumption 1: The gait system state x remains uniformly bounded. Moreover, the system trajectory starting from x_0, denoted as $\varphi_\varsigma(x_0)$, is in either a periodic or periodic-like (recurrent) motion regardless of the view angle variation.

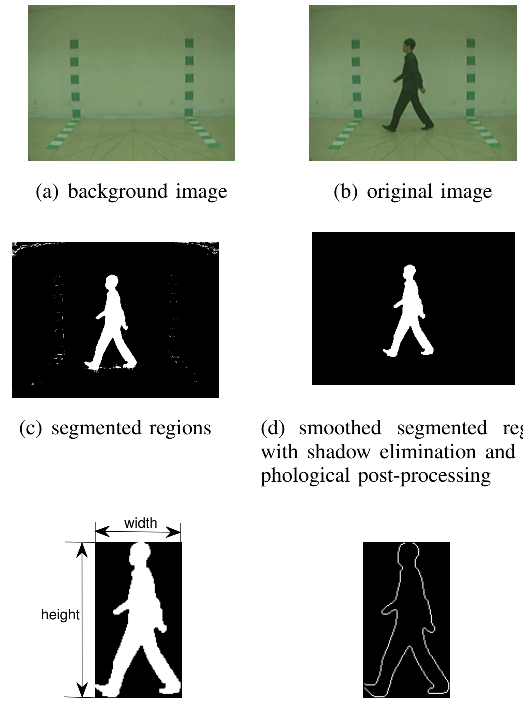
For the human gait system, Ref. [10] suggested that human gait is a form of periodic or quasi-periodic motion no matter the view angle variation is large or small . Our objective is to choose suitable human gait features satisfying Assumption 1 and design a dynamic RBF network to identify and approximate the unknown vector $F(x;p)$ under different view angles. The approximation result can be used to represent the gait dynamics which will be stored and used for view-invariant gait recognition.

III. GAIT FEATURE EXTRACTION AND GAIT SIGNATURE DERIVATION

In this section, we investigate four kinds of periodic silhouette width features. They seem suitable for capturing discriminatory gait information from sequences of extracted silhouettes and for representing gait dynamics under different view angles. An important issue in gait recognition is the extraction of salient features that will effectively capture gait characteristics. The features must reflect the view variation and should yield good discriminability across individuals. For each image sequence, we first introduce the silhouette segmentation and preprocessing methods, then present the proposed silhouette representation method. After that, we describe how to obtain silhouette features of each gait sequence. Finally, the gait signature consisting of gait dynamics under different view angles is derived.

A. Silhouette Extraction

Silhouettes in a walking image sequence can be extracted using the method proposed in [11]. Note that some extracted silhouettes are incomplete. To solve this problem , we use mathematical morphology method to fill in holes and remove noise. Edge images are produced by applying the Canny operator with hysteresis thresholding. Finally, the body silhouette is determined followed by dilation and erosion. A bounding box is then placed around the part of the motion image that contains the moving person. These boxed binarized silhouettes can be used directly as image features or further processed to derive the width vector and the ratio vector of the silhouette's height and width as in the next item. Fig. 1 shows an example of human silhouette extraction of a gait image.



(a) background image

(b) original image

(c) segmented regions

(d) smoothed segmented regions with shadow elimination and morphological post-processing

(e) binary silhouette with a bounding box

(f) silhouette contour

Fig. 1. Extraction of a moving silhouette

B. Width of the Outer Contour of the Binary Silhouette

Width contains structural as well as dynamical information of gait. Owing to influenced by the traction force produced by muscle and skeleton, athletics apparatus (foot, hip and hand etc.) of an individual are always in periodic diversification state, and the width of silhouette produced from the projection in photography plane also changes in a periodic state. Between the width feature of different individuals' gaits, there is a lot of dissimilarity in period and distribution due to the diversity in individuals' geometrical configuration, state of healthiness, psychology and view angle.

Distance between left and right extremities of the silhouette gives the width vector. From the binarized silhouette, the left and right boundaries are traced. In order to reflect the local shape and its change of human gait influenced by view angle, the gait silhouette has been divided into four equal subregions from top to bottom, namely subregion 1, subregion 2, subregion 3 and subregion 4, as shown in

Fig. 2.

Width calculation is shown in Fig. 2. Here, X axis denotes the row index and Y axis denotes the width associated with that row. H is the height of the silhouette. The width along a given row is simply the difference between leftmost and rightmost boundary pixels (1-valued) in that row. For a binary gait image $b(X,Y)$ indexed spatially by pixel location (X,Y), Y_X^1 represents the Y-coordinate of the leftmost boundary pixel in the Xth row, and Y_X^2 represents the Y-coordinate of the rightmost boundary pixel in the same row.

For all rows, the holistic width feature W_d is generated by

$$W_d = \max_{X\in[0,H]} (Y_X^2 - Y_X^1) \tag{2}$$

For specific rows, the desired width features W_d^1 and W_d^2 of subregions 3 and 4 are generated by the following equations, respectively.

$$W_d^1 = \max_{X\in[\frac{2}{4}H,\frac{3}{4}H]} (Y_X^2 - Y_X^1) \tag{3}$$

$$W_d^2 = \max_{X\in[\frac{3}{4}H,H]} (Y_X^2 - Y_X^1) \tag{4}$$

where d denotes the dth frame binary silhouette. During a walking period, the width of subregions 1 and 2 changes slightly with the view angle variation while the width of subregions 3 and 4 changes significantly. In accordance with life experience, the upper limbs swing lightly while the lower limbs swing more sharply when the view changes.

The mean and median width W_d^3 and W_d^4 of the silhouette will reflect the variation of the whole silhouette when the view changes, and will be obtained as:

$$W_d^3 = \underset{X\in[0,H]}{mean} (Y_X^2 - Y_X^1) \tag{5}$$

$$W_d^4 = \underset{X\in[0,H]}{median} (Y_X^2 - Y_X^1) \tag{6}$$

Hence, the four vectors W_d^1, W_d^2, W_d^3 and W_d^4 thoroughly reflect the influence of view angle variation on silhouette shapes. Fig. 3 and Fig. 4 show the width feature curves of one person under different view angles. It is seen that the four width vectors are periodic or quasi-periodic.

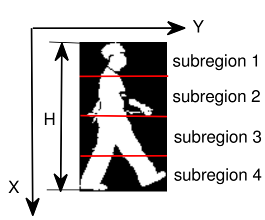


Fig. 2. Width feature extraction.

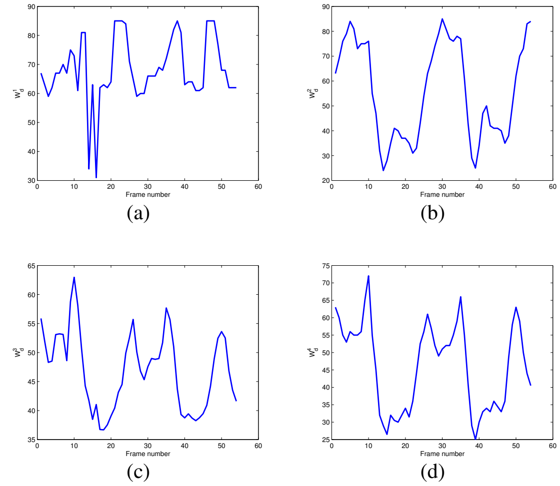


Fig. 3. The four width vectors for one person under 36° view angle in a gait sequence.

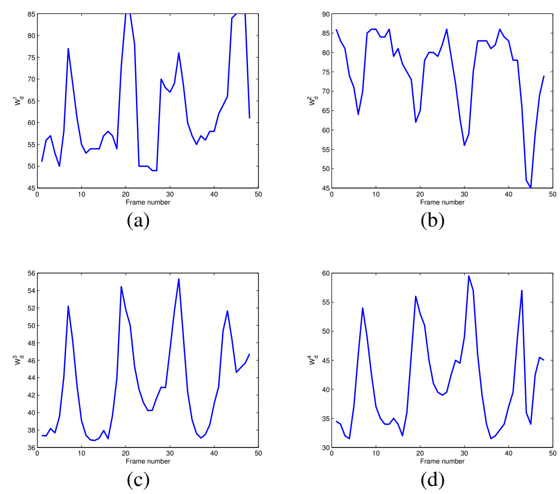


Fig. 4. The four width vectors for one person under 108° view angle in a gait sequence.

C. Derivation of Gait Signature

Gait signatures are the most effective and well-defined representation method for dynamic gait analysis. They can be extracted by motion information from human gait. To recognize individuals walking in different view angles by their gait easily, we need to firstly select the most efficient gait features which can best represent the gait characteristics and reflect the view variation.

The width vectors W_d^1, W_d^2, W_d^3 and W_d^4 are combined as silhouette features for gait recognition. They capture the spatio-temporal characteristics of each individual, represent the dynamics of gait motion, and can sensitively reflect the variance between different gait patterns generated by different view angles. This derives the gait signature consisting of gait dynamics under different view angles.

IV. TRAINING AND LEARNING MECHANISM BASED ON SILHOUETTE FEATURES

In this section, based on deterministic learning theory we present a scheme for identification of gait system dynamics under different view angles.

In order to more accurately describe the human walking, the gait dynamics can be modeled as the following form:

$$\dot{x} = F(x;p) + v(x;p) \tag{7}$$

where $x = [x_1, \ldots, x_n]^T \in R^n$ are the states of system (7) which represent the combined silhouette features of the human body under different view angles, p is a constant vector of system parameters. $F(x;p) = [f_1(x;p), \ldots, f_n(x;p)]^T$ is a smooth but unknown nonlinear vector representing the gait system dynamics under different view angles, $v(x;p)$ is the modeling uncertainty. The system trajectory starting from initial condition x_0, is denoted as $\varphi_\zeta(x_0)$.

Since the modeling uncertainty $v(x;p)$ and the gait system dynamics $F(x;p)$ cannot be decoupled from each other, we consider the two terms together as an undivided term, and define $\phi(x;p) := F(x;p)+v(x;p)$ as the general gait system dynamics. The objective of the training or learning phase is to identify or approximate the general gait system dynamics $\phi(x;p)$ under different view angles to a desired accuracy via deterministic learning.

Based on deterministic learning theory [12], the following dynamical RBF networks are employed to identify the gait system dynamics $\phi(x;p) = [\phi_1(x;p), \ldots, \phi_n(x;p)]^T$:

$$\dot{\hat{x}} = -A(\hat{x} - x) + \hat{W}^T S(x) \tag{8}$$

where $\hat{x} = [\hat{x}_1, \ldots, \hat{x}_n]$ is the state vector of the dynamical RBF networks, $A = diag[a_1, \ldots, a_n]$ is a diagonal matrix, with $a_i > 0$ being design constants, localized RBF networks $\hat{W}^T S(x) = [\hat{W}_1^T S_1(x), \ldots, \hat{W}_n^T S_n(x)]^T$ are used to approximate the unknown $\phi(x;p)$.

The NN weight updating law is given by:

$$\dot{\hat{W}}_i = \dot{\tilde{W}}_i = -\Gamma_i S(x)\tilde{x}_i - \sigma_i \Gamma_i \hat{W}_i \tag{9}$$

where $\tilde{x}_i = \hat{x}_i - x_i, \tilde{W}_i = \hat{W}_i - W_i^*$, W_i^* is the ideal constant weight vector such that $\phi_i(x;p) = W_i^{*T} S(x) + \epsilon_i(x)$, $\epsilon_i(x) < \epsilon^*$ is the NN approximation error, $\Gamma_i = \Gamma_i^T > 0$, and $\sigma_i > 0$ is a small value.

With Eqs. (7)-(8), the derivative of the state estimation error $\tilde{x}_i$ satisfies

$$\dot{\tilde{x}}_i = -a_i\tilde{x}_i + \hat{W}_i^T S(x) - \phi_i(x;p) = -a_i\tilde{x}_i + \tilde{W}_i^T S(x) - \epsilon_i \tag{10}$$

By using the local approximation property of RBF networks, the overall system consisting of dynamical model (10) and the NN weight updating law (9) can be summarized into the following form in the region Ω_ζ

$$\begin{bmatrix} \dot{\tilde{x}}_i \\ \dot{\tilde{W}}_{\zeta i} \end{bmatrix} = \begin{bmatrix} -a_i & S_{\zeta i}(x)^T \\ -\Gamma_{\zeta i} S_{\zeta i}(x) & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_i \\ \tilde{W}_{\zeta i} \end{bmatrix} + \begin{bmatrix} -\epsilon_{\zeta i} \\ -\sigma_i \Gamma_{\zeta i} \hat{W}_{\zeta i} \end{bmatrix} \tag{11}$$

and

$$\dot{\hat{W}}_{\bar{\zeta} i} = \dot{\tilde{W}}_{\bar{\zeta} i} = -\Gamma_{\bar{\zeta} i} S_{\bar{\zeta} i}(x)\tilde{x}_i - \sigma_i \Gamma_{\bar{\zeta} i} \hat{W}_{\bar{\zeta} i} \tag{12}$$

where $\epsilon_{\zeta i} = \epsilon_i - \tilde{W}_{\bar{\zeta} i}^T S_{\bar{\zeta}}(x)$. The subscripts $(\cdot)_\zeta$ and $(\cdot)_{\bar{\zeta}}$ are used to stand for terms related to the regions close to and far away from the trajectory $\varphi_\zeta(x_0)$. The region close to the trajectory is defined as $\Omega_\zeta := Z|dist(Z, \varphi_\zeta) \leq d_\iota$, where $Z = x, d_\iota > 0$ is a constant satisfying $s(d_\iota) > \iota$, $s(\cdot)$ is the RBF used in the network, ι is a small positive constant. The related subvectors are given as: $S_\zeta(x) = [s_{j1}(x), \ldots, s_{j\zeta}(x)]^T \in R^{N_\zeta}$, with the neurons centered in the local region Ω_ζ, and $W_\zeta^* = [w_{j1}^*, \ldots, w_{j\zeta}^*]^T \in R^{N_\zeta}$ is the corresponding weight subvector, with $N_\zeta < N$. For localized RBF networks, $|\tilde{W}_{\bar{\zeta} i}^T S_{\bar{\zeta}}(x)|$ is small, so $\epsilon_{\zeta i} = O(\epsilon_i)$.

The norminal part of system (11) is referred to as system (11) without the terms $-\epsilon_{\zeta i}$ and $-\sigma_i \Gamma_{\zeta i} \hat{W}_{\zeta i}$. For the human gait system, Ref. [10] suggested that human gait is a form of periodic or quasi-periodic motion. In Section III, we have shown that the silhouette features are quasi-periodic signals generated from the free human walking sequences. Hence, the NN input $x = [W_d^1, W_d^2, W_d^3, W_d^4]^T$ is quasi-periodic.

According to Theorem 1 in [12], the regression subvector $S_{\zeta i}(x)$ satisfies PE condition almost always. This will lead to exponential stability of $(\tilde{x}_i, \tilde{W}_{\zeta i}) = 0$ of the nominal part of system (11) [13]. Based on the analysis results given in [12], the NN weight estimate error $\tilde{W}_{\zeta i}$ converges to small neighborhoods of zero, with the sizes of the neighborhoods being determined by $\epsilon_{\zeta i}$ and $\| \sigma_i \Gamma_{\zeta i} W_{\zeta i}^* \|$, both of which are small values. This means that the entire RBF network $\hat{W}_i^T S(x)$ can approximate the unknown $\phi_i(x;p)$ along the trajectory φ_ζ, and

$$\phi_i(x;p) = \hat{W}_i^T S(x) + \epsilon_{i1} \tag{13}$$

where $\epsilon_{i1} = O(\epsilon_{\zeta i})$.

By the convergence result, we can obtain a constant vector of neural weights according to

$$\bar{W}_i = mean_{t\in[t_a,t_b]} \hat{W}_i(t) \tag{14}$$

where $t_b > t_a > 0$ represent a time segment after the transient process. Therefore, we conclude that accurate identification of the function $\phi_i(x;p)$ is obtained along the trajectory $\varphi_\zeta(x_0)$ by using $\bar{W}_i^T S_i(x)$, i.e.,

$$\phi_i(x;p) = \bar{W}_i^T S(x) + \epsilon_{i2} \tag{15}$$

where $\epsilon_{i2} = O(\epsilon_{i1})$ and subsequently $\epsilon_{i2} = O(\epsilon^*)$.

Hence, locally-accurate identification of the gait system dynamics $\phi_i(x;p)$ under different view angles to the error level ϵ^* is achieved along the trajectory $\varphi_\zeta(x_0)$. Time-varying gait dynamical patterns can be effectively represented by the locally-accurate NN approximations of the gait system dynamics, and this representation is time-invariant. The gait system dynamics under different view angles constitute a uniform training gait patterns dataset which is used for the following view-invariant gait recognition.

V. GAIT RECOGNITION MECHANISM

In this section, we present a scheme for rapid recognition of human gait using the learned gait system dynamics under different view angles.

Consider a training dataset containing dynamical human gait patterns φ_ζ^k under different view angles, $k = 1, \ldots, M$, with the kth gait training pattern φ_ζ^k generated from

$$\dot{x} = F^k(x; p^k) + v^k(x; p^k), \quad x(t_0) = x_{\zeta 0} \tag{16}$$

where $F^k(x; p^k)$ denotes the gait system dynamics, $v^k(x; p^k)$ denotes the modeling uncertainty, p^k is the system parameter vector.

As shown in Section IV, the general gait system dynamics $\phi^k(x; p^k) := F^k(x; p^k) + v^k(x; p^k)$ can be accurately identified and stored in constant RBF networks $\bar{W}^{k^T} S(x)$. By utilizing the learned knowledge obtained in the training phase, a bank of M estimators is first constructed for the trained gait systems as follows:

$$\dot{\bar{\chi}}^k = -B(\bar{\chi}^k - x) + \bar{W}^{k^T} S(x) \tag{17}$$

where $k = 1, \ldots, M$ is used to stand for the kth estimator, $\bar{\chi}^k = [\bar{\chi}_1^k, \ldots, \bar{\chi}_n^k]^T$ is the state of the estimator, $B = diag[b_1, \ldots, b_n]$ is a diagonal matrix which is kept the same for all estimators, x is the state of an input test pattern generated from Eq. (7).

In the test phase, by comparing the test gait pattern generated from human gait system (7) with the set of M estimators (17), we obtain the following recognition error systems:

$$\begin{aligned} \dot{\tilde{\chi}}_i^k &= -b_i \tilde{\chi}_i^k + \bar{W}_i^{k^T} S_i(x) - \phi_i(x; p), \\ & \quad i = 1, \ldots, n, \quad k = 1, \ldots, M \end{aligned} \tag{18}$$

where $\tilde{\chi}_i^k = \bar{\chi}_i^k - x_i$ is the state estimation (or synchronization) error. We compute the average L_1 norm of the error $\tilde{\chi}_i^k(t)$

$$\|\tilde{\chi}_i^k(t)\|_1 = \frac{1}{\mathrm{T}_c} \int_{t-\mathrm{T}_c}^{t} |\tilde{\chi}_i^k(\tau)| d\tau, \quad t \geq \mathrm{T}_c \tag{19}$$

where T_c is the cycle of human gait.

The fundamental idea of human gait recognition is that if one person appearing whose view pattern contained in the prior training dataset similar to the trained gait pattern s ($s \in \{1, \ldots, k\}$), the constant RBF network $\bar{W}_i^{s^T} S_i(x)$ embedded in the matched estimator s will quickly recall the learned knowledge by providing accurate approximation to the human gait dynamics. Thus, the corresponding error $\| \tilde{\chi}_i^s(t) \|_1$ will become the smallest among all the errors $\| \tilde{\chi}_i^k(t) \|_1$. Based on the smallest error principle, the appearing person can be recognized. We have the following recognition scheme.

Human gait recognition scheme: If there exists some finite time t^s, $s \in \{1, \ldots, k\}$ and some $i \in \{1, \ldots, n\}$ such that $\| \tilde{\chi}_i^s(t) \|_1 < \| \tilde{\chi}_i^k(t) \|_1$ for all $t > t^s$, then the appearing person can be recognized.

VI. EXPERIMENTS

In our experiments, two widely adopted multiview gait databases are used to evaluate the performance of the proposed method, which include: CASIA-B gait database [14] and CMU MoBo gait database [15]. These databases directly support the study of gait recognition with respect to the variation of views. Moreover, from the research perspective, there are different advantages from the two databases: 1) the CASIA-B gait database contains a large number of subjects; 2) the CMU Mobo gait database has been widely used by a large number of papers. The experiments are implemented using matlab software and tested on an Intel Core i5 3.5GHz computer with 4GB RAM.

Different from experiments in other works, we construct a uniform training dataset consisting of gait patterns under different view angles. When a test pattern whose view angle cannot be known priorly but contained in the training dataset appears, it can be rapidly recognized. This makes it more applicable in real-world applications. Since the training dataset is different from that of other works, we do not compare the recognition performance of our method with other benchmark methods in [4–6].

A. CASIA-B Gait Database

The CASIA-B is a large dataset, including 124 different subjects (93 males and 31 females) with variations in view angle and walking status (normal, in a coat, or with a bag) [14]. The videos were synchronously captured from 11 different views (namely 0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162° and 180°) in a well controlled laboratorial environment. There were 11 USB cameras around the left side of the subject when he/she was walking, and the angle between two nearest views is 18°. The video sequences have spatial resolution and frame rate of 320×240 pixels and 25 frames per second, respectively. Fig. 5 shows sample images in this gait database. As we only focus on the view factor affecting the gait recognition performance, six normal walking sequences of each subject collected from all these 11 views are selected in the following experiments. The corresponding width features are extracted for each sequence and the silhouette image is resized to a fixed 128×88 image.

Fig. 5. Sample frames from 11 different views of one subject in the CASIA-B gait database.

Eleven experiments designed for this database are listed in Table I. We assign sequences to training set for all the 124 subjects to construct a uniform training database. Each subject contains three sequences for each view angle. That is, there are $124 \times 3 \times 11 = 4092$ patterns in the training dataset. Based on the method described in Section III, we extract all the 124 persons' silhouette features through walking image sequences, which means the input of the RBF networks $x = [W_d^1, W_d^2, W_d^3, W_d^4]^T$. In order to eliminate the data difference between different silhouette features, all the silhouette feature data is normalized to $[-1, 1]$. One example of silhouette feature normalization for a walking image sequence is shown in Fig. 6.

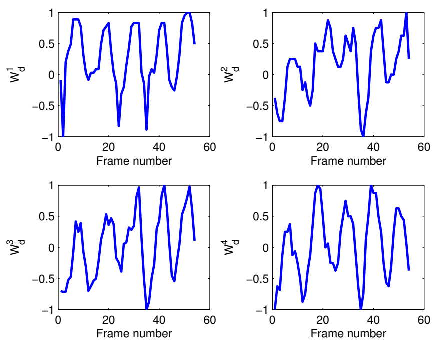


Fig. 6. Normalization of silhouette feature data.

Fig. 7 and Fig. 8 show an example of the training and recognition of the person numbered 001 under different view angles in the 124-person dataset. Consider recognition of the test person 001 under 54° view angle. There are three sequences to training set (labeled '001-nm-01-054', '001-nm-02-054' and '001-nm-03-054' in the database) and three sequences to test set (labeled '001-nm-04-054', '001-nm-05-054' and "001-nm-06-054' in the database). In the training phase, the RBF network $\hat{W}_i^T S_i(x)$ is constructed in a regular lattice, with nodes $N = 83521$, the centers μ_i evenly spaced on $[-1,1] \times [-1,1] \times [-1,1] \times [-1,1]$, and the widths $\eta = 0.15$. The weights of the RBF networks are updated according to Eq. (9). The initial weights $\hat{W}_i(0) = 0$. The design parameters for (8) and (9) are $a_i = 0.5, \Gamma = diag\{1.5, 1.5, 1.5, 1.5\}, \sigma_i = 10, (i = 1, \ldots, 4)$. The convergence of neural weights is shown in Fig. 7, which demonstrates partial parameter convergence, that is, only the weight estimates of some neurons whose centers close to the orbit are activated and updated. These weights converge to their optimal values W_i^*. Based on the deterministic learning theory, the gait dynamics $\phi_i^k(x; p^k)$ can be locally accurately approximated by $\bar{W}_i^{k^T} S_i(x), (k = 1, \ldots, 4092)$ along recurrent system trajectory, then these constant weights are stored for each training pattern.

In the test phase, by using the constant networks $\bar{W}_i^{k^T} S_i(x), (k = 1, \ldots, 4092)$, 4092 RBF network estimators are constructed based on (17). The parameters in (17) and (19) are $b_i = -25$ $(i = 1, \ldots, 4), T_c = 1.01s$. Consider recognition of the test person 001 under 54° view angle (represented by the test pattern '001-nm-04-054') by 4092 training patterns. The average L_1 norms of the synchronization errors, that is, $\| \tilde{x}_i^k(t) \|_1$ $(k = 1, \ldots, 4092)$ are shown

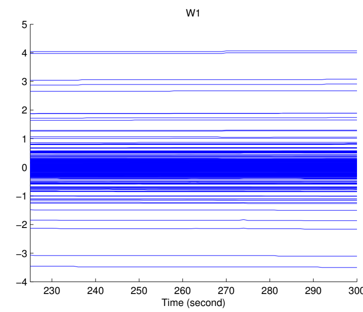


Fig. 7. Partial parameter convergence of $\hat{W}_1$ in one training pattern.

in Fig. 8. It is obvious that after certain time, the average L_1 norm generated by the training pattern '001-nm-01-054' becomes smaller than the others.

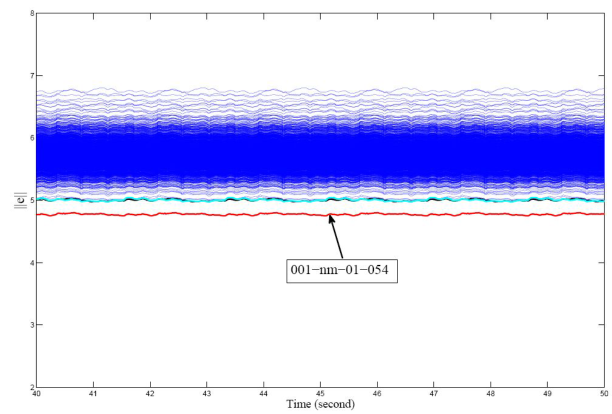


Fig. 8. Recognition of person 001 under 54° view angle using smallest error principle.

The recognition performance of the proposed methods is reported in terms of the correct classification rate (CCR). The Rank-1 recognition performance of our approach is presented in Table II.

TABLE II
GAIT RECOGNITION PERFORMANCE (%, RANK-1) ON THE CASIA-B GAIT DATABASE.

Experiment	CCR (%)
A	55.4%
B	44.2%
C	66.7%
D	77.8%
E	77.8%
F	87.9%
G	66.7%
H	77%
I	75.8%
J	76.7%
K	57.9%

B. CMU MoBo Gait Database

This database comprises gait sequences from 25 subjects and two different walking speeds, namely slow walking and fast walking. It contains walking subjects captured from six

TABLE I
ELEVEN EXPERIMENTS ON THE CASIA-B GAIT DATABASE FOR ROBUSTNESS TEST.

Experiment	Gallery Set (containing 11 view angles)	Probe Set	Gallery Size	Probe Size
A	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	0°	124×3×11	124×3
B	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	18°	124×3×11	124×3
C	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	36°	124×3×11	124×3
D	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	54°	124×3×11	124×3
E	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	72°	124×3×11	124×3
F	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	90°	124×3×11	124×3
G	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	108°	124×3×11	124×3
H	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	126°	124×3×11	124×3
I	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	144°	124×3×11	124×3
J	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	162°	124×3×11	124×3
K	0°, 18°, 36°, 54°, 72°, 90°, 108°, 126°, 144°, 162°, 180°	180°	124×3×11	124×3

cameras located in positions as shown in Fig. 9. We use five (see Fig. 10) out of the six available viewing directions, omitting the north view, since it is practically identical to the south view (i.e., the frontal view).

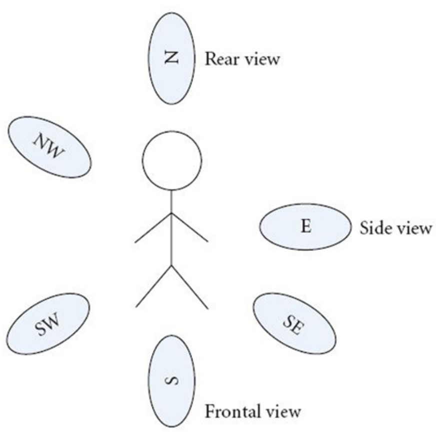
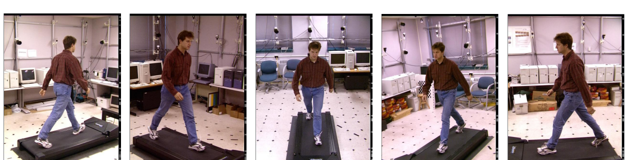


Fig. 9. Camera arrangement in the CMU MoBo gait database [16]. Six cameras are oriented clockwise in the east (E), southeast (SE), south (S), southwest (SW), northwest (NW), north (N), with the walking subject facing toward the south.



(a) NW (b) SW (c) S (d) SE (e) E

Fig. 10. Five different views in the CMU MoBo database.

Several experiments designed for this database are listed in Table III and Table V. The process of training and recognition is similar to the examples of CASIA-B database shown in Section VI-A and is omitted here for clarity and conciseness. The Rank-1 performance of our approach is presented in Table IV and Table VI.

It is seen from the experiments that the proposed method achieves very promising performance no matter the view change is small or significantly large.

VII. CONCLUSIONS

A new view-invariant gait recognition approach based on silhouette features via deterministic learning theory is presented in this paper. Based on the method for feature extraction, silhouette features representing the gait dynamics and reflecting the view variation can be extracted. The gait system dynamics under different view angles can be accurately approximated by RBF networks and the obtained knowledge will be stored in constant RBF networks. A uniform training dataset consisting of gait patterns under different view angles is constructed. Then, according to the dynamical estimators and the smallest error principle, the view-invariant gait recognition can be achieved.

REFERENCES

[1] D.S. Matovski, M.S. Nixon, S. Mahmoodi and J.N. Carter, "The effect of time on gait recognition performance," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 543-552, 2012.

[2] R. Cilla, M.A. Patricio, A. Berlanga and J.M. Molina, "A probabilistic, discriminative and distributed system for the recognition of human actions from multiple views," *Neurocomputing*, vol. 75, no. 1, pp. 78-87, 2012.

[3] A. Kale, K.R. Chowdhury and R. Chellappa, "Towards a view invariant gait recognition algorithm," in: *Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance*, pp. 143-150, 2003.

[4] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo and Y. Yagi, "Gait recognition using a view transformation model in the frequency domain," in: *Proceedings of the 9th European Conference on Computer Vision*, pp. 151-163, 2006.

[5] W. Kusakunniran, Q. Wu, J. Zhang and H. Li, "Cross-view and multi-view gait recognitions based on view transformation model using multi-layer perceptron," *Pattern Recognition Letters*, vol. 33, no. 7, pp. 882-889, 2012.

[6] G. Zhao, G. Liu, H. Li and M. Pietikainen, "3D gait recognition using multiple cameras," in: *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pp. 529-534, 2006.

[7] A. Veeraraghavan, A.K. Roy-Chowdhury and R. Chellappa, "Matching shape sequences in video with ap-

TABLE III
FIVE EXPERIMENTS ON THE CMU MOBO GAIT DATABASE FOR ROBUSTNESS TEST. THE IMAGES OF FAST WALKING ARE SELECTED AS THE GALLERY AND PROBE SETS.

Experiment	Gallery Set (fast walking)	Probe Set (fast walking)	Gallery Size	Probe Size
A	NW+SW+S+SE+E	NW	$25\times3\times5$	25×3
B	NW+SW+S+SE+E	SW	$25\times3\times5$	25×3
C	NW+SW+S+SE+E	S	$25\times3\times5$	25×3
D	NW+SW+S+SE+E	SE	$25\times3\times5$	25×3
E	NW+SW+S+SE+E	E	$25\times3\times5$	25×3

TABLE IV
GAIT RECOGNITION PERFORMANCE (%, RANK-1) ON THE CMU MOBO GAIT DATABASE.

Experiment	CCR (%)
A	92%
B	92%
C	52%
D	80%
E	96%

TABLE V
FIVE EXPERIMENTS ON THE CMU MOBO GAIT DATABASE FOR ROBUSTNESS TEST. THE IMAGES OF SLOW WALKING ARE SELECTED AS THE GALLERY AND PROBE SETS.

Experiment	Gallery Set (slow walking)	Probe Set (slow walking)	Gallery Size	Probe Size
F	NW+SW+S+SE+E	NW	$25\times3\times5$	25×3
G	NW+SW+S+SE+E	SW	$25\times3\times5$	25×3
H	NW+SW+S+SE+E	S	$25\times3\times5$	25×3
I	NW+SW+S+SE+E	SE	$25\times3\times5$	25×3
J	NW+SW+S+SE+E	E	$25\times3\times5$	25×3

TABLE VI
GAIT RECOGNITION PERFORMANCE (%, RANK-1) ON THE CMU MOBO GAIT DATABASE.

Experiment	CCR (%)
F	88%
G	72%
H	56%
I	76%
J	92%

plications in human movement analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 12, pp. 1896-1909, 2005.

[8] W. Zeng and C. Wang, "Human gait recognition via deterministic learning," *Neural Networks*, vol. 35, pp. 92-102, 2012.

[9] W. Zeng and C. Wang, "Model-based human gait recognition via deterministic learning," *Cognitive Computation*, doi: 10.1007/s12559-013-9221-4.

[10] M.P. Murary, A.B. Drought and R.C. Kory, "Walking pattern of movement," *American Journal of Medicine*, vol. 16, no. 1, pp. 290-332, 1967.

[11] S. Sarkar, P.J. Phillips, Z. Liu, I.R. Vega, P. Grother, K.W. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 162-177, 2005.

[12] C. Wang, T. Chen, G. Chen and D.J. Hill, "Deterministic learning of nonlinear dynamical systems," *International Journal of Bifurcation Chaos*, vol. 19, no. 4, pp. 1307-1328, 2009.

[13] J. Farrell, "Stability and approximator convergence in nonparametric nonlinear adaptive control," *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 1008-1020, 1998.

[14] S. Yu, D. Tan and T. Tan, "Modelling the effect of view angle variation on appearance-based gait recognition," in: *Proceedings of the 7th Asian Conference on Computer Vision*, pp. 807-816, 2006.

[15] R. Gross and J. Shi, "The CMU motion of body (mobo) database technical report," Robotic Inst., Carnegie Mellon Univ., Pittsburgh, PA, CMU-RITR-01-18, 2001.

[16] X. Huang and N.V. Boulgouris, "Human gait recognition based on multiview gait sequences," *EURASIP Journal on Advances in Signal Processing*, 2008.