Human Gait Recognition Using Template Features with Opponent-Motion Model

Hossein.SamadiTazehKand
Isfahan University of Technology
Tehran Iran
Samadi_hosein@yahoo.com

Mozaffar Alizadeh
Department of Electrical Engineering Technical and Vocational University Tehran Iran
m-alizadeh@tvu.ac.ir

***Abstract*—In this paper, we apply a filtering scheme to encode motional information from style of human walking into single template. This filtering process utilizes bio-inspired basis kernels which are constructed from two separable spatial and temporal basis functions. To extract opponent motions from gait in a walking cycle, wefusethese kernels together to have spatiotemporal kernels and then convolve itwith gait stances. This paradigm has been repeatedineach frame of gait in a cycle and motional information is highlighted as salient regions inthe sequence of images. To make gait features (salient regions)distinctive enough within different individuals, the responses of gait filtering are aggregated together to get final Gait Salient Image (GSI). Our proposed model provides more representative gait feature since conventional gait templates such as Gait Energy Image (GEI) does not rely on any motion model. Extensive experiments on popular gait databases reveal that, our GSI-based features provides more competitive performance compared with recent published gait recognition approaches with efficiency and accuracy.**

***Index Terms*—Biometric identification, Gait recognition,Orientedmotion filters, Template-based feature.**

I. INTRODUCTION

GAITof people, i.e. manner of walking, has been known as an effective type of biometric for human identification in surveillance and security applications [1, 2]. This is mainly due to eases of capture of gait in unobtrusive way. However, any biometric system based on gait recognition suffers from some exterior factors. The source of variants could be of clothing, carriage conditions, type of shows, surface, angle of view and age. Although these effects (besides interior problems, like psychological mood), are challenging, gait is irreplaceable biometrics [1, 3] which can solve human authentication problems in surveillance situations [2, 3].

To handle the mentioned problems, we must describe the gait by distinctive motion model. To represent the human's motion by conventional gait features, two models named model-based and appearance-based are introduced [4, 5]. Meta-analysis of gait [3, 6, 8, 9] reveals that appearance is more consistent with variations of gait of individuals. Considering this hypothesis, it is reasonable to compute the motion model from the appearance of gait. Moreover for better gait recognition, sequence of moving shapes is aggregated and converted to single form called template [3]. Common appearance-based gait templates such as Gait Energy Image (GEI) [3], Gait Flow Image (GFI) [10] and Gait Entropy Image (GEnI) [11] have alternatively simple structure with low computational costs. Since the nature of gait is a spatiotemporal process, this conversion losses temporal ordering of gait sequence. Here, the proposed templates are spatial-based features including motion model implicitly.

Recently, some temporal templates are proposed to preserve temporal ordering of gait along time. For example, C. Wang et al. [12] inform Chrono-Gait Image (CGI) which time ordering of gait presents by multichannel colored image. T. Lam et al. [10] introduce Gait Flow Image (GFI) which template of gait generated using optical flow estimation. The timing issues in GFI include in gait template as optical flow model. Y. Makihara et al. [13] construct Gait Silhouette Volume (GSV) as a spatiotemporal model of walking person. Then, they apply Fourier Description Feature (FDF) to build different templates of gait across different view directions. Considering recent time-based templates, CGI template proposes robust gait feature for recognition. However,these time-based templates do not rely on a convenient human's motion model. Furthermore, the mentioned above models offer complex scheme of template generation due to recognition tasks. In this paper, we apply novel time-based filtering to extract the motional information from gait and encode it in a single template.

Since spatial-based templates such as GEI provides simple and effective structure, we pay more attention on developing a spatiotemporal motion model to include it in a single template.Our new template model is Gait Salient Image (GSI) where the motional information saves as salient regions in final map (template). The main part of GSI is utilizing a bio-inspired spatiotemporal filtering process [14]for gait. This model has been applied successfully in the action recognition task [15] while we propose it for gait recognition in this paper.The proposed filtering is a motion-energy map consists of four main steps: First,two spatial and two temporal basis functions have been set with predefined scales. Then, a pair of these bases is chosento form quadratic pairs of basis kernels. The quadratic pairs afterwards mixed together making quadratic pairs of oriented (spatiotemporal) kernels. The proposed oriented kernels are able to detect four opponent motions within human's movement. Finally, the sequence of input gait imageshas been convolved with these oriented kernels and responses competewith together resulting

opponent motion of gait in sequence.This spatiotemporal filtering [14] is a motion-energy map which is tuned for detection of opposite motions (e.g. leftward vs. rightward). Considering the proposed motion model, the main contribution of paper is utilizing it for gait recognition. To this end, we first divide one cycle of gait into two equal halves. This division is mainly due to remove similar patterns of motion (e.g. one stride has two consecutive analogous steps). Then we compute the responses of spatiotemporal filtering separately for each half. Finally, the GSI is constructed by aggregation of these responses over each half and averaging the results in one period of gait. The whole procedure of our GSI template generation is shown in Figure 1. In comparison with recently published gait recognition method (e.g. [3], [6], [10], [12]), our major contributions are:

- The GSI templates have simple structure and easy to compute.
- Our proposed model is almost parameter-free since there is needed to tune just few parameters. Meanwhile, the performance of recognition is somehow robust against the variation of parameters.
- In our GSI template, both of the spatial and temporal information of gaithave been suitably embedded. Previous works, e.g. [12], have been devoted additional process to extract contours from gait, while there are no more steps to extract contour in this paper.
- Finally, the proposed GSI model is robust against hard challenging gait conditions according to our experiments.

In the rest of the paper we will explain completely the GSI temporal template model and discuss on how to recognize the gait of human based on GSI. The paper is organized as follows: We review a brief survey in Section II, then details of proposed scheme including filtering process, computing opponent images and generating final GSI template is discussed in Section III. Section IV presents the human gait recognition based on GSI and Section V proposes the experimental results of the model. Finally, Section VI concludes the paper.

II. Related Works

There are two general approaches for gait feature extraction: model-based and appearance-based [4, 5]. In this section we will review these methods and discussonadvantages of each method. Also, we will explain thebackgrounds of spatiotemporal filtering.

Model-based approaches consider a predefined structure for motion model [16, 37]. In model-based approaches, the articulated parts of body are characterized by predefinedparameters which are then fitted based on gait features. Zhou et al. [17] propose two types of parameters: time-invariant (static) and time-variant (dynamic). Static parameters are length of limbs or size of body while dynamic parameters are the angle between limbsthat varying during the movement. To recognize the gait, the statistics of parameters in Bayesian framework have been learnedand fitted to prior knowledge. Although gait recognition based on Bayesian framework has high performance in indoor gait data, the approach is time-consuming and vulnerable to noise. An alternative method has been developed by Wang et al. [18]who employa condensation framework to track the walker and recover dynamic information. Moreover, parameters of model could be represented effectively by Hidden Markov Model (HMM) [19] or Fourier descriptors [20, 22].Also to extract robust gait features, statistical information of limbsmovement (like mean and variances of angles) could be used as parameters of model[4], [21], [22].

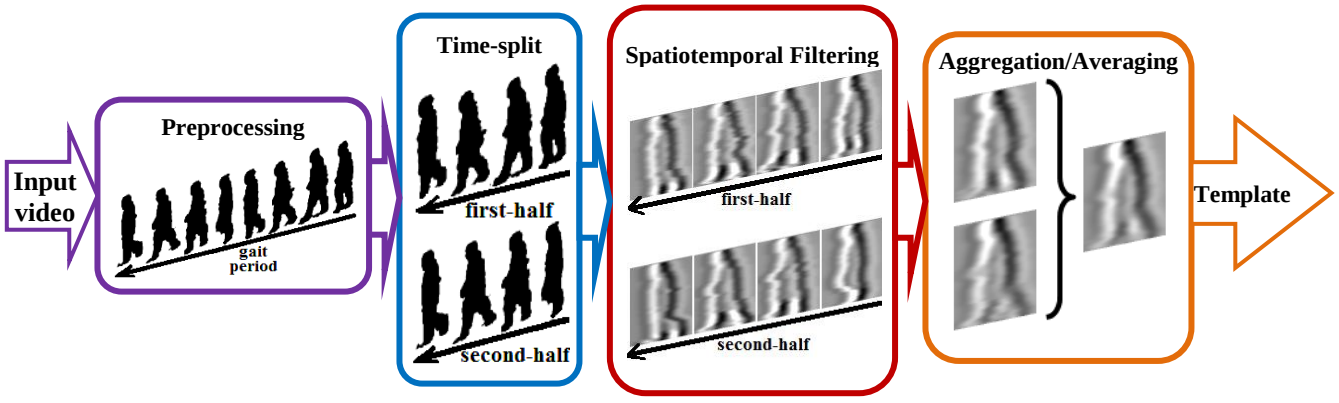


Figure 1.The proposed framework for salient time-based gait template. Here, the motional information of gait in one period of walking is preserved in the final map.

However, model-based approach limits in some directions: 1) it is vulnerable to noise and occlusions, 2) there needs to define a robust initial model, 3) complexity of model raises exponentially by increasing number of parameters, and 4) it is not applicable for large population of databases. To alleviate these problems, appearance-based (or model-free) approach has been developed [3, 6, 12, 23]. In this approach, there is no need to define prior model of human movement because features of gait areextracted during recognition process. In the gait recognition studies, appearance-based features are roughly divided into two categories: Temporal-based features and Template-based features [12]. In formal approach, all frames of gait sequence are being kept in database for matching. In contrast,in template-based approach,the sequence of walking is represented by single gait template.

Sarkar et al. [6] propose baseline algorithm which the correlation between all silhouettes of each sequence are used as benchmark of recognition. In this method, the correlation of gait between gallery and probe sets is computed in three levels for matching: frame-to-frame, period-to-period and sequence-to-sequence matching. As an extension to temporal strategy, some approaches have been developed. Wang et al. [23] apply principle component analysis (PCA) to extract statistical non-model features from the gait frames of sequence. Liuand Sarkar [8] utilize population HMM (pHMM) and introduce Dynamics-Normalized Gait Representation Algorithm (DNGRA) to get a dynamics-normalize stance-framefor recognition of each individual.Although direct sequence matching is very simple and accurate, however it is time consuming and needs large volume of memory.

In the template-based models, the whole sequence of gait is convertedinto single template. Han et al. [3] introduce the original idea of gait template model by averaging the gait stances in one period. Xu et al. [24] utilizeDiscriminant Analysis with Tensor Representation (DATER) which the gait template is represented as 2D tensor in image. Moreover, General Tensor Discriminant Analysis (GTDA) is developedby Tao et al. [25] thatrobustgait template is achieved by Gabor features. Xu et al. [26] generalize GTDA and propose Locality-constrained Group Sparse Representation (LGSR) for gait recognition. Also, Chen et al. [27] proposeMultilinear Tensor-based non-Parametric dimension reduction (MTP)for gait recognition and Zhang et al. [28] apply a generalized MTP for gait recognition in low-resolution videos. In addition, Guoand Nixon [29] utilizeMutual Information (MI) to select important features for gait recognition. The above mentioned template models offer an effective gait representation in videos. However, such a conversion may lose the temporal information of human's motion.

To solve this problem, Lam et al. [10] recently proposeGait Flow Image (GFI) and Wang et al. [12] develop a colored gait image named Chrono-Gait Image(CGI), where inboth models the temporal information in single gait templatehas been preserved well. However, these models do not exploit an optimized gait motion model and have difficulties bothin computation of model and in matching templates between gallery and probe sets.

We want to address this problem by a simple opponent-based motion model. The basic idea behind this model is using spatiotemporal filtering to encode the motion data into single template. The response of filtering is a salient map where informative regions are marked in the final map. This opponent-based filtering scheme is inspired from human's visual cortex. Adelson et al. [14] propose earlier works on motion-energy map and utilize an opponent-based motion model consistent with visual cell. However, this model is non-casual and hence not biologically plausible [14, 30]. To make the model casual, Watson et al. [30] develop Adelson's motion model and suggest a temporally casual filter. The Watson's model is consistent with the electrophysiologicalstudies and psychophysical data of human's visualcontrast-sensitivity measurements. This model is capable of handling various types of motion. Recently, Shabani et al. [15] investigate this motion model as an effective model for action recognition.In this paper, we utilize this motion model to extract human's walking knowledge for gait recognition purposes. Our model is Gait Salient Image (GSI) where motional data would be encoded in a single template. Our experimental results on publicly gait databases (USF [6] and CASIA [7]) reveal that this spatiotemporal model could enhance recognition rate and is competitive to most popular template models. In the following section we will explain thoroughly the GSI template model.

III. GAIT SALIENT IMAGE (GSI)

Generally speaking, each individual has periodic gait with fix period [3, 6, 12]. Moreover, it is found that there are sub-cycles with homogenous motions within each period of gait.To compute efficient spatiotemporal model, cycles with similar patterns should be identified before filtering. For example we divide each period into two equal cycles (Figure 1) to compute the responses in each halfseparately. For this purpose, we need to compute the gait period toextract and align spatiotemporal features of gait efficiently.In the next subsection we review the period detection step. Then, we will focus on the main parts of our model on the following subsections.

A. Preprocessing and Gait Period Estimation

To achieve best performance in gait recognition system, some preprocessesare required to calculate from gait video sequences. The major steps are background subtraction, foreground alignment, gait period detection and time-splitprocedure. First, a background subtraction technique like Gaussian Mixture Model (GMM) is needed to perform on the original gait sequence [6].By modeling background, gait of individual could be derived from foreground model. Then, foreground images align according to center of image and resize to obtain normalized silhouettes with fix size and aspect ratio. Once the foreground is modeled, the template features of gait could be calculated from silhouettes.

Considering that regular human walking is a periodical motion, it is essential to calculate the period of gaitbefore

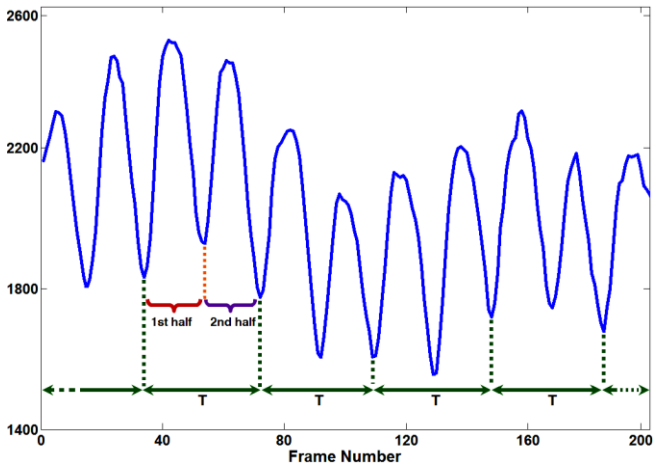


Figure 2.Variation of lower-half of foreground pixels in silhouette over time. The horizontal axis indicates frame number and vertical axis shows $p_f(t)$. Here T is the gait period containing 1st and 2nd similar halves.

filtering. Taking into account the period of gait, we are able to synchronize the starting phase of filtering in whole sequence.

This period could be easily computed from the aligned silhouettes by counting the foreground pixels over time, $p_f(t)$.To increase the sensitivity, it is convenient to limit the counting area to legs region. In baseline algorithm [6], the counting area limits to bottom half of silhouette while in other models like CGI [12], it is narrowed to legs from above ankle to knee.The reason is that in CGI, for keeping time-ordering well, it is required to compute the period more accurately. Since our GSI model in this paper utilizes a motion model to preserve spatiotemporal information, we just compute the period simply like the way in Baseline algorithm [6]:

$$p_f(t) = \sum_{i=0.5h}^{h} M_{LRi}(t), \qquad (1)$$

where h is the height of silhouette andSOP_{LRi}(t) is sum of pixels in the i^{th} line of silhouette image from leftmost to rightmost foreground pixels at time $t.p_f(t)$ has a local maximum when two legs are farthest apart from each other and reaches a local minimum when the two legs entirely overlap. By computing the extremums,the period is simply measured by median of distances between two consecutive minimums. Figure 2 shows an instance of $p_f(t)$ over time where the period is indicated as T under the curve. Here, for better visualization the $p_f(t)$ data has been smoothed.

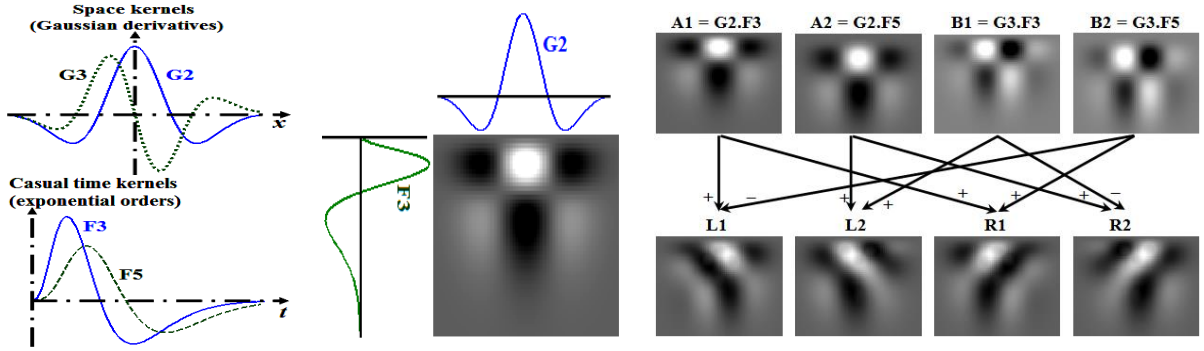
As it is shown in Figure 2, there are homogenous patterns in each period of gait. The reason is that there are similar motions during walking. For example in two halves of one period (as indicated in Figure 2), foots have identical motion which differ in direction. The spatiotemporal filtering in this paper is capable of detecting direction of motion, but in our final template both leftwards and rightwards motions must have the same cost. To reduce the dependencies of model to the direction of motion, we divide each period into two halves and compute the responses of filtering for each half separately. Theprocedure of the period intervalsdivision into multiple cycles (e.g. two equal halves) is named as time-split procedure in this paper as shown in Figure 1. Now within these sub-periods we can compute the responses of filtering as is explained in the next subsection.

B. Salient Map Generation

The main contribution of this paper is using a bio-inspired time-casual filtering to generate a salient gait model. In this model both spatial and temporal information of gait would be preserved well.To import motional information into a single gait template, some strategies exist.An efficient model is proposed by Wang's multichannel color coding as CGI [12]. In CGI, spatial and temporal information are preserved as contour of gait and its color, respectively.This model, as mentioned before, does not represent any explicit model of the gait motion. But in our GSI, a spatiotemporal filtering is used to encode the motion.The filtering scheme in this paper is derived from Adelson's motion model [14] which is easily built from two linearly-separable spatial and temporal signals. However, the primary model is not time-casual where its usage in real world scenarios will be a challenging task. For this purpose and based on Adelson's model [14], the Watson's time-casual filtering scheme [30] has been developed by Shabani et al. [15]. Then, the modified motion model has been applied for action recognition tasks successfully.The time-casual model describes motion perception model of human's action which is consistent with real human's visual system. The process of basiskernels generation for time-casual filtering in this model is explained graphically in Figure 3.

As it is showninFigure 3 (a), the motion perception model consists of four basic signals: (1) two spatial signals which is derivative of Gaussian kernels, and (2) two time-casual temporal functions which are Watson's temporally kernels [30]. These basis signalsare combined together formingquadratic pairs of1D+t spatiotemporal kernels, named R_1, R_2, L_1 and L_2in Figure 3 (c). Each of these four kernels is sensitive to one direction of motion. In other words,R_1 and R_2 could detect rightward motions while L_1 and L_2reveal leftward motions in a video sequence.

After computing these kernels, steps of proposed filtering for gait recognition is as follows: First, each gait silhouette has been convolved with spatiotemporal kernels to make four spatiotemporal maps. Then, responses from similar direction are normalized and addedtogether. By aggregating theseresponses, two motion-energy maps (one for leftward and one for rightward) would be derived. Finally, the salient map of gait silhouette is generated by subtraction of two energy maps. These steps have been shown in Figure 4.



(a) (b) (c)

Figure 3. (a) Spatial and temporal basis functions, (b) Combination of the functions to create basis kernels, (c) Process of spatiotemporal kernels generation. Here,A_1, A_2, B_1, B_2 quadratic pairs are basis kernels and L_1, L_2, R_1, R_2 quadratic pairs are oriented kernel (or spatiotemporal kernel) [13, 14].

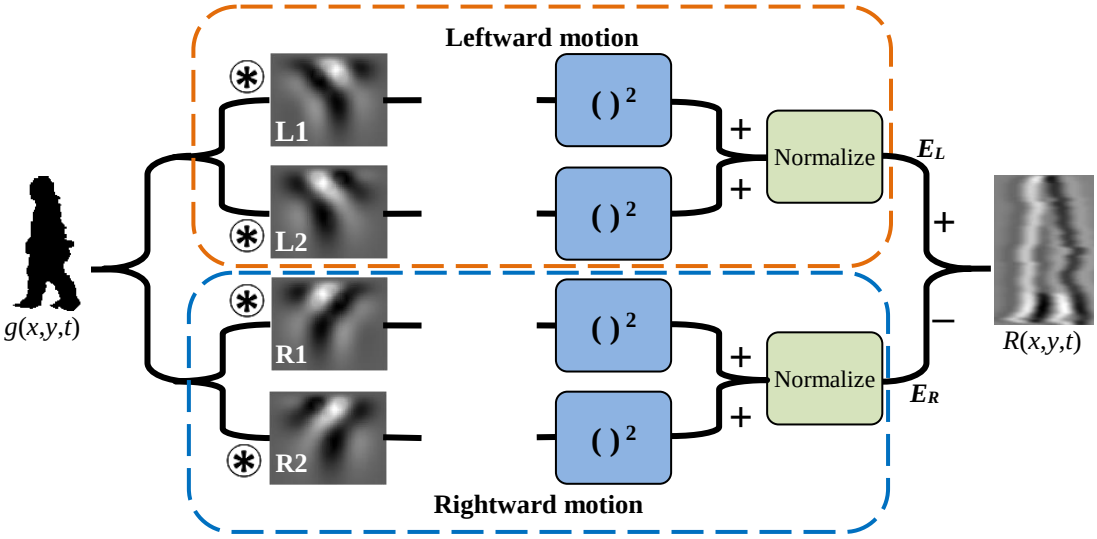


Figure 4.The proposed spatiotemporal filtering approach for gait recognition.

The spatial part of mentioned basic kernels in motion perception model defined as derivatives of Gaussian;

$$G_n(X,\sigma)=\frac{\partial^n}{\partial X^n}\left\{\frac{1}{2\pi\sigma^2}\exp\left(-\frac{|X|^2}{2\sigma^2}\right)\right\}, \qquad (2)$$

where$X$ and σ are spatial coordination and its scale respectively.Also the temporal part of kernels is Watson's exponential function [31];

$$F_n(t)=\left[\frac{1}{n!}-\frac{(kt)^2}{(n+2)!}\right](kt)^n\exp(-kt). \qquad (3)$$

Here,n and k=1/τ are temporal order and temporal scale respectively. Also t is frame indexvarying up to the number of silhouettes in one sequence (t_{max}). In this paper, t_{max} sets to one half of gait period (length of sequence). It can be easily proved that any orders of this temporal function is casual because it takes zero for values of$t\leq 0$.

According to Figure 3, the motion map consists of four 1D signals; two for spatial encoding and two for temporal measurements. For this purpose, the second and third derivatives of Gaussian kernels (Equation 2) are used as spatial measurements whilethird and fifth orders of Watson's kernel(Equation 3) are utilized for temporal model.

In the next step, the basis kernels will be generated simply by products of spatial and temporal 1-D signals. This process has been shown graphically in Figure 3 (b) for G_2 and F_3. By mixing these signals, we will have four oriented kernels named A_1, A_2, B_1 and B_2in Figure 3. These kernels could detect motions in different scales which is convenient forspatiotemporal kernels generation.

By takingbasis kernels as inputs of filtering, spatiotemporal kernels could beeasily computed by linear combination of inputs. As it is shown in Figure 3 (c),proposed kernels are calculated as;

$$\begin{cases} L_1 = A_1 - B_2; & L_2 = A_2 + B_1 \\ R_1 = A_1 + B_2; & R_2 = A_2 - B_1 \end{cases}. \qquad (4)$$

It is obvious from Figure 3 that set of $\{L_1,L_2\}$ and $\{R_1, R_2\}$are used to measureleftward and rightward motions in video respectively. To extract the motional information from the input sequence, each frame of the video should be

convolved with the four above kernels. Then, the responses for each leftwards and rightwards motions are added together to make normalized motions in each direction. Finally, to compute the opponent motion map of the input video, the final normalized responses of filtering are subtracted. The proposed spatiotemporal filtering (based on Adelson's model) for gait silhouette at time *t* has been shown in Figure 4.

First, the input image, $g(x,y,t)$,at time *t*is convolved with four spatiotemporal kernels;

$$\begin{cases} R_{L1} = L_1 * g(x,y,t); & R_{L2} = L_2 * g(x,y,t) \\ R_{R1} = R_1 * g(x,y,t); & R_{R2} = R_2 * g(x,y,t) \end{cases}. \quad (5)$$

In the filtering procedure, the dynamic ranges of pixels in responses would be changed.In the primary Adelson's model [14, 15]it is not mentioned to this issue. However, to extract normalized motion model from aligned gait silhouettes, we add normalization step in this paper. If we note $Rg_t(x,y)$ as generic term of responses at time *t*, it would be normalized by following formula;

$$Rg'_t(x,y) = \frac{Rg_t(x,y) - \min_{x,y} Rg_t(x,y)}{\max_{x,y} Rg_t(x,y) - \min_{x,y} Rg_t(x,y)}. \quad (6)$$

This normalization maps the range of pixels in four spatiotemporal responses to range [0, 1]. By summing these normalized responses, two energy models, E_L and E_R, for leftward and rightward motion (in Figure 4) will be generated respectively. Finally, the leftward motion energy is subtracted from the rightward motion to obtain an opponent-based model. The proposed salient map, $R(x,y,t)$, at time *t*has been shown at the right most of Figure 4. It could be seen from the final model that salient brighter and darker regions are more sensitive torightward and leftward motions.The mechanism of our salient map generation in Figure 4 is a nonlinear process.This issue is not too problematic since the input kernels are composed of linearly-separable spatiotemporal signals. However, the motion perception modelprovides useful toolsin human motion analysis such as [15]; insensitive to contrast, invariant to phase (or initial stance of the gait in a cycle) and consistent with the perceptual studies. In other words, with using this model, both spatial and temporal information of gait will be preserved well in the final template.

The proposed model of filtering in Figure 4 is computed for silhouettes in a cycleforming the sequence ofsalient maps. In the next subsection we will discuss on how to create the GSI template based on these salient maps.

C. ModelRepresentation

The salient map, mentioned in subsection III-B,will be calculated in each frame of silhouette in a cycle. The proposed filtering is capable of modeling a variety of human's motion from slow to fast movement. This is mainly due to high-pass filtering of oriented kernels to input video sequence. To keep this diversity of motionsin single template, it is required to sum all of the responses along time.

Given the response of gait silhouette in t^{th} frameR(x,y,t), or simply R_t, the sequence of salient maps states as;

$$\mathfrak{R}(\mathbf{x},\mathbf{y}) = \{R_0, R_1, \ldots, R_{t\max-1}\}, \quad (7)$$

wheret_{max} is the number of silhouettes in a cycle. Each element in $\mathfrak{R}(\mathbf{x},\mathbf{y})$ contains motional information in image with respect to specific time stamp. To preserve spatiotemporal information in one sequence, we sum all salient maps overthe cycle;

$$OGI(x,y) = \sum \mathfrak{R}(x,y) = \sum_{k=0}^{t_{\max}-1} R_k. \quad (8)$$

Here, OGI(x,y) (Oriented Gait Image) contains all motion maps in one cycle. Specifically, thetrajectory of human movement (and also temporal ordering)would be preserved in OGI since local motions are summed up in the final map. This trajectory could reflect the manner of walking (i.e. gait) in the sequence accurately. Note that OGIis a normalized map in range [0, 1] because it is the summation of normalized salient maps.

To compute the final GSI template we have to average the OGIs in a sequence. Since in one period of the gait, there are similar motions, we divide each period into multiple cycles and compute the OGI for each cycle separately. This division guarantees that in a period,each OGIdoes not have the similar motions. Generally, suppose each period of the gait includes*p* cycles, our GSI model is written as;

$$\mathrm{GSI}(x,y) = \frac{1}{p}\sum_{i=1}^{p} OGI_i(x,y). \quad (9)$$

Here, the values of GSI template are also normalized to range[0, 1] since the OGIs are the normalized templates. In our model, each period of gait is divided into two equal halves (Figure 1), and hence we set p=2. As an example, the process of GSI computing is shown in Figure 5. The first row of Figure 5 shows some silhouettes chosen from the first half of the gait period. The filtering responses, $R(x,y,t)$, for the gait sequence of first row are shown in second row respectively. Furthermore, sum of the responses, or OGI_1, along with OGI_2 (from second half of the gait) are shown in the rightmost of third row.Finally, our final GSI templates for this individual and others are shown in the rest of third row.

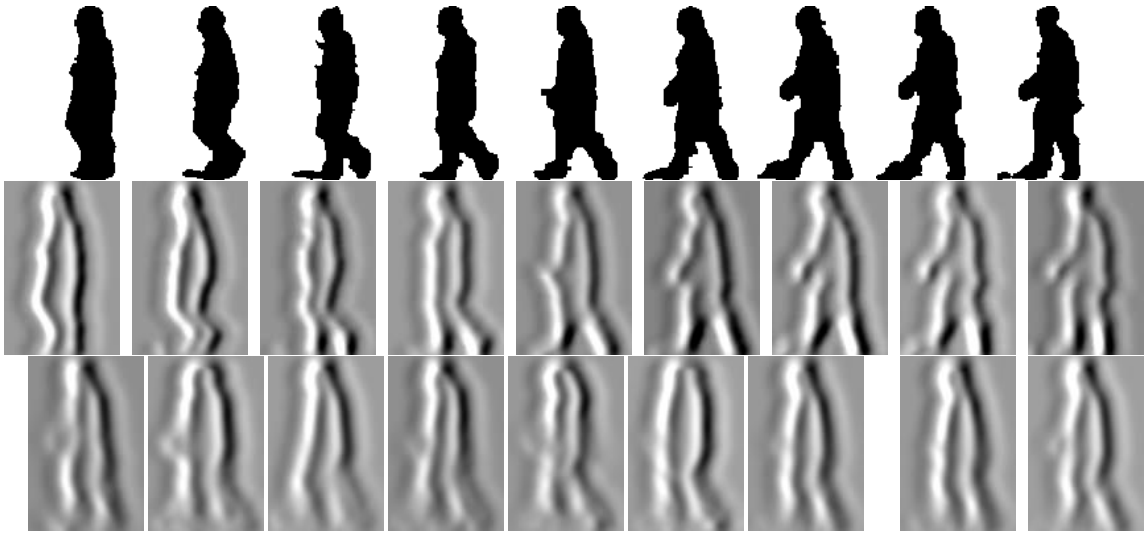


Figure 5. The flow of generating GSI templates.The first row shows silhouettes from the first half of period. The second row shows responses offiltering. Two rightmosttemplatesin the third row are OGI_1, OGI_2while the other templates are GSI of this individual and othersaccordingly.

IV. Gait Recognition Using GSI

The GSI template, explained in Section III, is computed for all individuals from gallery and probe sets. To recognize the gait, these GSI templates between the gallery and probe sets should be compared. However, direct matching of GSI from probe sets to gallery is not practical. This is due tosome reasons: 1) The gait samples are captured from similar conditions so the gait features may be overfitted. 2) The number of templates in training set is small and limited, so the gait features cannot characterize the topology of fundamental gait space. 3) Since each pixel of GSI template is considered as one dimension, the dimensionality of feature space is much higher than the training samples. This limitation of the training samples is known as undersample problem (USP) [12, 25]. To addressthe limitation of training samples, several methods have been proposed. The simple strategy is to increase the size of training set by generating synthetic templates. This ideahas been applied first for GEI [3] template and further for CGI temporal template [12]. For fair comparisons, in both GEI and CGI templates,the real and synthetic sets are compared separatelyand the final accuracy isobtained by fusion of the results.However, the result from comparisons of synthetic templates is somehow artificial. In this paper, we compute the real templates of gallery and probe sets and compare them accordingly.

To address the USP without losing the computational efficiency, we employ a two-staged dimensionality reduction technique. Here, simple Principle Component Analysis (PCA) [31, 32] followed by Linear Discriminant Analysis (LDA) [33, 34] technique is applied to GSI template. This approach,known as PCA+LDA,is found as an effective approach in face recognition [35]. For this purpose, the real templates in the gallery set are projected into a low-dimensional subspace by PCA first and then by LDA.By computingthe PCA+LDA projection matrix from real templates in gallery set, real templates in the probe set will also be projected into a low-dimensional subspace [3].

Suppose there is c classes in the gallery set where each one hasn_i(i=1,…,c) feature vectors in low-dimensional subspace. Also, let r_i be the feature vector of the i^{th} class (individual) in gallery set and let $\hat{r}_p$ (p=1,…,n_R) be the feature vector of individuals in probe sets, both in feature space. The distance of real template R_p(in probe) to real template R_i(in gallery) could be measured according to Euclidean distance between the center of classes of feature vectors in low-dimensional subspace.

$$D\left(\hat{R}_p, R_i\right) = \frac{1}{n_R}\sum_{j=1}^{n_R}\left\|\hat{\mathbf{r}}_j - m_{ri}\right\|, \qquad i = 1,\ldots,c \tag{10}$$

where $m_{ri} = \frac{1}{n_i}\sum_{r\in R_i}\mathbf{r}$ is the mean of feature vectors of ith class and $\|\cdot\|$ is the norm of vector. We assign given probe templateto k^{th} class if

$$D\left(R_p, R_k\right) = \min_i D\left(\hat{R}_p, R_i\right), \qquad i = 1,\ldots,c\,. \tag{11}$$

To recognize the gait, the Euclidean distance between all probe sets and gallery isfirst computed based on Equation (10). Then gaits in probe sets are matched to gallery according to minimum distance (Equation (11)). More details on measuring the template similarities could be found in Han's work [3].

V. Experimental Results

In this section, we will review our experiments in detail. This includesevaluating the effectiveness of our GSI model in human gait representation. Meanwhile, the performance of gait recognition based on our GSI template will be compared with recent published methods including; Baseline algorithm [6], HMM [36], 2DLDA [25], GTDA [25], Gabor+GTDA [25], DATER [24], MTP [27], DNGRA [8], GEI+Real [3], and CGI+Real [12].


Table 1
The Specifications of Gallery and Probe Sets

Data Label	Size	Variance	Differences	period / # period
Gallery	122	G, A, R, NB	--	35 / 5
Probe A	122	G, A, **L**, NB	Shoes/ View	35 / 5
Probe B	54	G, **B**, R, NB		35 / 5
Probe C	54	G, **B**, **L**, NB		35 / 5
Probe D	121	**C**, A, R, NB	Surface + Shoes/ View	33 / 5
Probe E	60	**C**, **B**, R, NB		33 / 5
Probe F	121	**C**, A, **L**, NB		33 / 5
Probe G	60	**C**, **B**, **L**, NB		33 / 5
Probe H	120	G, A, R, **BF**	Briefcase + Shoes/ View	35 / 6
Probe I	60	G, **B**, R, **BF**		35 / 6
Probe J	120	G, A, **L**, **BF**		35 / 6
Probe K	33	G, A/**B**, R, NB, **T**	Time + Shoes/ View Surface	34 / 6
Probe L	33	**C**, A/**B**, R, NB, **T**		34 / 6

Abbreviation Note: G-Grass, C-Concrete, A-Shoe A, B-Shoe B, R-Right View, L-Left View, NB-No Briefcase, BF-Briefcase, T-Time, period-the mean of periods of individuals in given set, # period-the number of periods repeated in sequencesof given set.

The benchmarks of evaluation are "Rank1" and "Rank5" performance, which Rank1 performance means the percentage of the correct subjects ranked first, while Rank5 performance means the percentage of correct object placed in the top five rank of the ranked list.

A. Gait Databases

We choose USF HumanID database (silhouette version 2.1) [6] and CASIA Gait Database (Dataset B) [7] as two benchmarks to evaluate proposed GSI model. Some examples of these databases are shown in Figure 6.

The USF database consists of 122 persons (individuals) walking in elliptical path in front of a camera. The walking conditions are on concrete and grass surfaces, with/ without a briefcase, with different shoe types, and with respect to elapsed time. Considering these 5 challenging factors, Sarakar et al. [6] select the sequence with "Grass, Shoe Type A, Right Camera view, No Briefcase, captured in Time t_1 (May)" for the gallery set and develop 12 different experiments for probe sets. The conditions in all probe sets with number of individuals in each one are listed in Table 1. The conditions in the gallery and each of the probe sets are unique and there are no common sequences between the probe sets. The whole experiments in database could be divided into four distinctive groups. The difference between each group to the gallery set, difference within the groups,and thelist of probe sets belongs to each group are shown in Table 1.

In this database, the sequences of aligned silhouettes along with the period of gait (as mentioned in subsection III-A) have been also provided. The rightmost column in Table 1 shows the mean of period of individuals for each set. Furthermore, the number of periods in given set is also noted as "# period" in Table 1. This parameter shows maximum number of gait templates we could compute for each set.

The CASIA Gait Database (Data set B) consists of 124 individuals. Each person walks under ten distinct gait situations as six sequences under normal conditions, two sequences with carrying a bag and two sequences with different coats (named as NM-01 to NM-06, BG-01, BG-02, CL-01, CL-02, respectively). In addition each experiment is captured in 11 different view directions, from 0 to 180 degrees with steps of 18 degrees between two nearest view directions. More details about this database could be found in [7].

Since we experiment GSI model on data captured from sidewall view direction, we only use the data provided from 90 degrees.

Although in this database the silhouettes have been provided after background subtraction, the silhouettes are not aligned already. To align silhouettes, we align horizontal centroid and cut the silhouette images into 160×100 (similar to conditions provided in CGI [12]). Most of our experiments have been evaluated for USF database due to ….

….

….

….

B. GSI Templates

To demonstrate how GSI modelis representative, we apply a single 1-Nearest Neighbor (1-NN) classifier on the original GSI. For this purpose, real GSI templates are computed from the gallery/probe sets and matched directly without using Principle Component Analysis/ Linear Discriminant Analysis (PCA+LDA).The reasons for direct comparisons are: 1) The input templates of gait construct a high-dimensional input space; so if the template models are discriminative enough in input space, the high recognition rate in low-dimensional feature space will be guaranteed,2) The performance of any gait recognition system suffers from noises and exterior factors.If the gait model is invariant enough inhigh-dimensional input space to such noises, the robustness of the model in low-dimensional space will roughly assured. Based on the mentioned reasons, the performance of direct template

matching provides a useful toolfor evaluating the effectiveness of the model. To evaluate the GSI template, the results of 1-NN classifier are comparedwith the baseline algorithm [6], GEI [3] and CGI [12] in the USF HumanID database. The results are listed in Table 2. It can be noted from Table 2 that; 1) The proposed GSI model achieves the best average Rank 1 performance among all other algorithms. 2) Compared withboth baseline and GEI, GSI is very robust to Shoe/View/Briefcase and Time conditions in experiments A, B, C, H, I, J, K and L. Specially, the Rank 1 accuracy is improved in most challenging conditions (probes H through L) by about 20-30 percent.3) Compared with CGI, the Rank 1 performance in GSI is improved in most of the experiments from 1 to 16 percents(in probe K). 4) Compared with bothbaseline and GEI, GSI has better Rank5 performance while the resultsare near to CGI. GSI is vulnerable to surface conditions through experiments D, E, F, G and L where Rank1 accuracy falls down to maximum of 8percents in probe F. It should be noted that baseline algorithm has better performance in surface conditions while the results of both GEI and CGI template modelsare weak accordingly.It means that, compared with GEI and CGI templates, our GSI template proposes more representative model of gait in video. This improvement could be easily followed in Table 2 since the accuracy of Rank1 accuracy in 7 out of 12 conditions are better than others.

Table 2
Comparison ofthe Recognition Performance for Direct GSI Template Matching with other comparable methods on the USF HumanID Dataset Using1-NN

Exp.	Rank1 Performance (%)				Rank5 Performance (%)			
	Baseline [6]	GEI [3]	CGI [11]	**GSI**	Baseline	GEI	CGI	**GSI**
A	73	84	87	**91**	88	93	96	**97**
B	78	87	94	**93**	93	94	94	**95**
C	48	72	72	**75**	78	93	93	**91**
D	32	19	17	**18**	66	45	41	**33**
E	22	18	25	**24**	55	53	45	**42**
F	17	10	12	**9**	42	29	32	**20**
G	17	13	13	**12**	38	37	35	**27**
H	61	56	78	**86**	85	77	91	**97**
I	57	55	80	**90**	78	77	97	**94**
J	36	40	54	**65**	62	69	82	**93**
K	3	9	6	**22**	12	15	30	**34**
L	3	3	9	**7**	15	15	27	**19**
Avg.	41.0	41.1	48.6	**52.3**	64.5	61.4	66.8	**65.4**

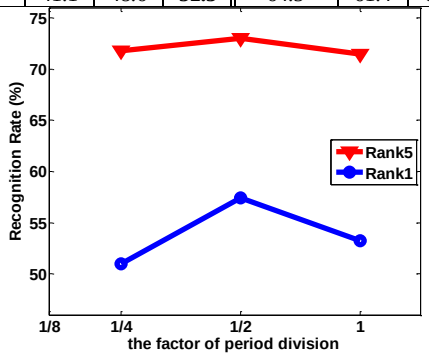


Figure 7.Performance of gait recognition using GSI model with different values of 1/*p*. Here, horizontal axis shows the factor for period division and vertical axis shows the recognition rate.

Moreover, due to gray level representation of the GSI model (rather than color level in CGI) it is more convenient to apply it in gait recognition system.

C. Tuning Parameters

Generally, to compute the GSI model, one may need to tune filtering and template parameters. The parameters of filtering affect the quality of responses (or salient maps) while the others impress the quality of final template. In this subsection we study the effect of these factors on performance of gait recognition.

In this paper, three parameters are introduced: spatial scale (σ in Equation 2), temporal scale ($k=1/\tau$ in Equation 3), and height of kernels ($t$in Equation 3).

In out experiments, it is found that the value of k=0.17 ($\tau \approx$ 6) generates acceptable performance. It means that higher orders of temporal resolution could better model the human's walking style.

To adjust other parameters (σ andt_{max}), some considerations should be taken into account. With increasing the σ (or spatial scale), the contour of templates are smoothed and vice versa. Moreover, with decreasing t_{max} the motional information are localized to short duration of time. By considering these rules, the performance of gait recognition based on GSI will be studied with the values of σ= {3, 6, 8} and t_{max}={T/2, T}.The resultsofgait recognition using GSI with different parametershas been shown in Table 3.However, to choose the best parameters for GSI, the Rank1/ Rank5 accuracy has been compared with CGI+Real temporal template.

It is clear from Table 3 that; 1)with increasing the σ (or reducing the spatial resolution), it is required to rise up the time resolution (increasing t_{max}) to achieve the same performance. 2) With considering any configurations, the Rank1 performance of proposed GSI is better than CGI+Real in most conditions while the Rank5 performance is almost near to CGI+Real.3) The average Rank1/Rank5 accuracy in all configurations is almost near to each other. It could be inferred from the results that our GSI model is robust against the parameters. However, we choose (3, 0.17) for pairs of (σ, k) with t_{max} = T/2in this paper.

The otherparameter for computing the GSI template is the period division factor,p.Although the values of $p$have similar effects with the parameter t_{max}, but also we inspect the effect of this parameter on performance.For this purpose, we evaluate the gait recognition model with different values of p.The performance of recognition against different values of 1/p is shown in Figure 6. It is clear that the best performance would be achieved with p=2, which means that the sequence of gait in one period should be divided into two equal cycles.The half-division of period makes easy the computational processsince the p in CGI has been set to 4.

Table 3
Comparison of Recognition Performance of CGI and GSI Using Different Parameters

Exp.	Rank1 Performance (%)							Rank5 Performance (%)						
	CGI+Real [12]	σ=3; k=0.17		σ=6; k=0.17		σ=8; k=0.17		CGI+Real [12]	σ=3; k=0.17		σ=6; k=0.17		σ=8; k=0.17	
		t_{max}†		t_{max}		t_{max}			t_{max}		t_{max}		t_{max}	
		$T/2$	T††	$T/2$	T	$T/2$	T		$T/2$	T	$T/2$	T	$T/2$	T
A	90	**94**	89	96	92	82	92	95	**99**	98	100	98	89	97
B	91	**95**	95	95	93	78	93	94	**95**	95	95	95	88	95
C	80	**80**	80	84	82	67	78	93	**95**	95	94	93	76	91
D	33	**32**	36	33	42	27	42	66	**62**	59	55	64	51	62
E	33	**30**	35	32	39	25	37	65	**57**	56	59	54	52	59
F	18	**19**	18	31	22	19	29	48	**40**	37	42	47	38	51
G	22	**17**	20	24	27	22	25	52	**45**	42	42	50	32	54
H	84	**85**	79	60	70	59	66	96	**92**	92	88	92	81	87
I	80	**84**	74	64	65	64	69	97	**92**	90	87	87	80	85
J	61	**74**	65	74	65	60	68	87	**90**	90	92	92	74	91
K	3	**19**	10	16	16	16	16	24	**43**	40	34	34	22	28
L	6	**16**	25	13	10	7	10	27	**31**	40	25	22	16	22
Avg.	54.49	**57.60**	55.34	55.70	55.67	47.08	56.21	75.05	**73.69**	72.64	72.00	74.04	62.81	73.52

† t_{max} is the length of sequence which is height of directional kernels.
†† T is the gait period.

D. Other Gait Environments

xxx

E. Computational Issues

Now, we want to prove that the time and space complexities in our GSI template are not quite more than the introduced gait approaches. Here, a simple gait model, e.g. CGI temporal template [12], and a sophisticated method, e. g. DNGRA [7],have been chosen as two benchmarksfor comparisons.

To generate the GSI templates for each training and test data, the most timing part is due to spatiotemporal filtering of each silhouette (Figure 4). Here, the timing costs of proposed opponent-based filtering consist of three major parts: complexities of convolution (in the order of $\Theta(4I_{filt})$), normalization process (with the order of $\Theta(6I_{norm})$), and other mathematical operations (in the order of $\Theta(4I_{math})$). By taking these costs into account, the whole filtering process for one single silhouette takes computational ordering of $\Theta(4I_{filt} + 6I_{norm} + 4I_{math})$. It is clear that $\Theta(I_{filt})$ is much more than $\Theta(I_{norm})$ and $\Theta(I_{math})$, so it would be dominated. With this assumption, the time of computing all GSI templates in one gait period, T, approximately takes $\Theta(4TI_{filt})$. Generally, the time complexity of measuring all GSI templates for the training and test data would be $\Theta(4N_{tr}TI_{filt} + 4N_{te}TI_{filt})$. Here, N_{tr} and N_{te} are the number of gait periods in the training set and test set, respectively. If the input silhouette has size of $W\times H$ (in USF database it is 88×128) and spatiotemporal kernels has size of $w\times h$ ($7\sigma\times T/2$ in our work), the filtering of one gait silhouette takes $\Theta(I_{filt}) \approx \Theta(WHwh)$ [38]. With this assumption,the total complexity of GSI templates will be $\Theta(4(N_{tr} + N_{te})TWHwh)$. Considering our complexity, the CGI model [11] with k channels (i.e. 3) takes $\Theta(k(N_{tr} + N_{te})TWH)$. It could be easily proved that compared with the CGI, the computational overhead of our model would be increased with the order of $\Theta(4/k wh) \approx \Theta(wh)$. This issue could be relieved efficiently if we employ fast filtering techniques.

Moreover, the complexity of GSI is considerably small in comparison with sophisticated algorithm, e.g. pHMM has been employed in DNGRA. Specifically, generating dynamics-normalized stance-frame for each training and test data with pHMM takes $\Theta(N_{tr}TWHI_{Kmeans}+N_{tr}TWHN_s^2I_{pHMM}+N_{te}TWHN_s^2)$ [8, 12]. Here, $N_s$is the number of states in the pHMM model in DNGRA, while I_{Kmeans} and I_{pHMM} are the numbers of iteration for K-means clustering and pHMM training, accordingly.

To simplify the complexities [12], let $S_{tr}=N_{tr}TWH$ and $S_{te}=N_{te}TWH$ be the size of the training and test data respectively. We can rewrite the time complexity of GSI, CGI and DNGRA into

$\Theta(4(S_{tr}+S_{te})wh)$, $\Theta(k(S_{tr}+S_{te}))$, $\Theta((I_{Kmeans}+N_s^2I_{pHMM}) +N_s^2S_{te})$.

Note that $k$is set to 3 in CGI [12], while N_s=20 andI_{Kmeans}>10 in DNGRA (I_{pHMM} is not mentioned in [8]). It is obvious that GSI and CGI are much faster than DNGRA and GSI is quite slower than CGI, while the recognition performance provided by GSI is competitive. In other words, GSI can process more than 20 frames per second[1], which with fast filtering techniques it could be utilized for teal scenarios.

[1] The time is measured with MATLAB code running on a machine with Intel Core2 Duo CPU P8400 2.20 GHz and 2 GB of DDR2 memory.

Table 4Comparison of the Recognition Performance(%) of Popular Methods and GSI on the USF HumanID Database

	A	B	C	D	E	F	G	H	I	J	K	L	Avg.
Rank1 Performance													
Baseline [6]	73	78	3	32	22	17	17	61	57	36	3	3	40.96
HMM [36]	89	88	68	35	28	15	21	85	80	58	17	15	53.54
2DLDA [25]	89	93	80	28	33	17	19	74	71	49	16	16	50.98
GTDA [25]	86	88	73	24	25	17	16	53	49	45	10	7	43.70
Gabor+GTDA [25]	84	86	73	31	30	16	18	85	85	57	13	10	52.51
DATER [24]	87	93	78	42	42	23	28	80	79	59	18	21	56.99
MTP [27]	90	91	83	37	43	23	25	56	59	59	9	6	51.57
DNGRA [8]	85	89	72	57	66	46	41	83	79	52	15	24	62.94
GEI+Real [3]	89	87	78	36	38	20	28	62	59	59	3	6	51.04
CGI+Real [12]	90	91	80	33	33	18	22	84	80	61	3	6	54.49
GSI+Real (our method)	**94**	**95**	**80**	**32**	**30**	**19**	**17**	**85**	**84**	**74**	**19**	**16**	**57.60**
Rank5 Performance													
Baseline [6]	88	93	78	66	55	42	38	85	78	62	12	15	64.54
HMM [36]	-	-	-	-	-	-	-	-	-	-	-	-	-
2DLDA [25]	97	93	93	57	59	39	47	91	94	75	37	34	70.95
GTDA [25]	100	97	95	57	54	34	45	75	80	70	25	25	66.15
Gabor+GTDA [25]	96	95	89	59	63	33	49	94	92	76	19	40	70.32
DATER [24]	96	96	93	69	69	51	52	92	90	83	40	36	75.68
MTP [27]	94	93	91	64	68	51	52	88	83	82	18	15	71.38
DNGRA [8]	96	94	89	85	81	68	69	96	95	79	46	39	82.05
GEI+Real [3]	93	93	89	65	60	42	45	88	79	80	6	9	68.68
CGI+Real [12]	95	94	93	66	65	48	52	96	97	87	24	27	75.05
GSI+Real (our method)	**99**	**95**	**95**	**62**	**57**	**40**	**45**	**92**	**92**	**90**	**43**	**31**	**73.69**

F. Performance Analysis

Now we compare our proposed model with several recently published algorithms on the USF HumanID database. We also report the average performance of recognition, by computing the ratio of correctly recognized individuals to the total number of subjects. The results are listed in Table 4. For fair comparisons, we evaluate real templates of GSI in this paper with real templates of GEI and CGI. We can see that the average performance of our GSI outperforms most of the algorithms. Furthermore, the GSI is robust in most of the challenging conditions due to Rank1/Rank5 improvement. Specifically, in 5 out of 12 conditions (probes A, B, H, J, K), the Rank1 recognition is the best among all other methods while in other experiments the accuracy is close to the best one. Generally, our method is robust in all conditions except in surface conditions in probes D, E, F and G. This result has been mentioned before when applying direct template matching using 1-NN. We can also see that in normal walking conditions, (probes A, B, C) the Rank1 provides better results compared with recent approaches (e.g. GEI [3] and CGI [12]). The Rank1 accuracy also has been improved in hard experiments (probes K, L) which the recognition rates in most of the algorithms are pretty low.

Moreover, the Rank5 results are competitive too where this accuracy is on the top or near to the top. The average Rank5 accuracy outperforms most of the algorithms and it is near to CGI Rank5 performance. However, our approach provides Rank1/Rank5 results close to DNGRA [8] with very lower computational complexity.

To conclude the results, it could be seen that the proposed gait recognition based on GSI template provides representative model of gait in video. Because in our model, an efficient motion model for walking conditions has been applied preserving spatiotemporal information of walking. To build GSI template, we need to tune a few parameters (just three scaling factors), where the performance of recognition is somehow invariant against different settings. These privileges make the model convenient to apply in real world scenarios.

VI. CONCLUSION

In this paper, a novel representation of gait has been proposed. The main contribution of GSI template is using a bio-inspired time-casual filtering to encode human's movement into a single template. In this model, input silhouettes are convolved with four spatiotemporal kernels each of themcan measure motion in one direction. The responses of filtering in each direction are then subtracted to find opponent motions of silhouettes in walking sequence. By averaging these opponent responses within the period of gait, the final GSI model will be generated. To verify effectiveness of our approach, several experiments have been performed on USF HumanID database. The results show that our approach provides competitive Rank1/Rank5 performance both in high-dimensional image space (using 1-NN) and low-dimensional feature space (usingPCA+LDA). Furthermore, our GSI model is somehow parameter-free since the performance of gait recognition is invariant against the parameters. Finally, we prove that the computational complexity of the GSI model is

comparable to the recent temporal template model (i.e. CGI [12]) and are lower than sophisticated method (i.e.DNGRA [8]).As a result, we can conclude that our GSI template provides better representation of gait without increasing the complexities in time and space. Consequently, this model is more robust than other proposed models and can be applied for real life applications.

REFERENCES

[1] M. S. Nixon, T. Tan, R. Chellappa, "Human Identification Based on Gait", *International Series on Biometrics*,US, Vol. 4, Springer-Verlag Publisher, 2006.

[2] G. Johansson, "Visual Motion Perception", *Journal of Scientific American*, 1976.

[3] J. Han, B. Bhanu, "Individual Recognition Using Gait Energy Image", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, Issue 2, Feb. 2006.

[4] S. A. More, P. J. Deore, "A survey on gait biometrics" *World Journal of Science and Technology*, Vol. 2, Issue 4, 2012.

[5] J. Wang, K. She, S. Nahavandi, A. Kouzani, "*A Review of Vision-Based Gait Recognition Methods for Human Identification*", IEEE International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp. 320-327, Dec. 2010.

[6] 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, Issue 2, Feb. 2005.

[7] S. Yu, D. Tan, and T. Tan, "*A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition*",Proceeding of 18th International Conference on Pattern Recognition (ICPR), Vol. 4, pp. 441-444, 2006.

[8] Z. Liu, S. Sarkar, "Improved Gait Recognition by Gait Dynamics Normalization", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 6, June 2006.

[9] A. Veeraraghavan, A.R. Chowdhury, R. Chellappa, "Matching Shape Sequences in Video with Applications in Human Movement Analysis",*IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 12, pp. 1896-1909, Dec. 2005.

[10] T. H. W. Lam, K. H. Cheung, J. N. K. Liu, "Gait Flow Image: A Silhouette-Based Gait Representation for Human Identification", *Journal of Pattern Recognition*, Vol. 44, Issue 4, 2011.

[11] K. Bashir, T. Xiang, S. Gong, "*Gait Recognition Using Gait Entropy Image*", Proceedings of 3rd International Conference on Imaging for Crime Detection and Prevention, 2009.

[12] C. Wang, J. Zhang, L. Wang, J. Pu, X. Yuan, "Human Identification Using Temporal Information Preserving Gait Template", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, Issue 11, Sep. 2012.

[13] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, Y. Yagi, "*Gait Recognition Using a View Transformation Model in the Frequency Domain*", Proceedings of 9th European Conference on Computer Vision, May 2006.

[14] E. H. Adelson, J.R. Bergen, "Spatio-temporal Energy Models for the Perception of Motion", *Journal of Optical Society of America*, Vol. 2, Issue 2, 1985.

[15] A. H. Shabani, J. S. Zelek, D. A. Clausi, "*Human Action Recognition using Salient Opponent-Based Motion Features*", Proceedings of IEEE Canadian Conference on Computer and Robot Vision, June 2010.

[16] W. Zeng, C. Wang, Y. Li, "Model-Based Human Gait Recognition via Deterministic Learning", *Journal of Cognitive Computation*, Springer, 2013.

[17] Z. Zhou, A. P. Bennett, R. I. Damper, "A Bayesian Framework for Extracting Human Gait Using Strong Prior Knowledge", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, Issue 11, Nov. 2006.

[18] L. Wang, T. Tan, H. Ning, and W. Hu, "Fusion of Static and Dynamic Body Biometrics for Gait Recognition," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 2, pp. 149-158, Feb. 2004.

[19] A. Sundaresan, A. R. Chowdhury, R. Chellappa, "*A Hidden Markov Model Based Framework for Recognition of Humans from Gait Sequences*", IEEE International Conference on Image Processing (ICIP), Sept. 2003.

[20] S. D. Mowbray, M. S. Nixon, "*Automatic Gait Recognition via Fourier Descriptors of Deformable Objects*",Audio Visual Biometric Person Authentication, pp. 566-573, 2003.

[21] Zhang R, Vogler CH, Metaxas D. Human gait recognition at sagittal plane. *Image Vis Comput*.,25(3):321–30, 2007.

[22] I. Bouchrika and M.S. Nixon, "*Model-Based Feature Extraction for Gait Analysis and Recognition*," Proceedings of Third International Conferenceon ComputerVision/Computer Graphics Collaboration Techniques and Applications,pp. 150-160, 2007.

[23] L. Wang, T. Tan, H. Ning, W. Hu, "Silhouette Analysis-Based Gait Recognition for Human Identification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, Issue 12, Dec. 2003.

[24] D. Xu, S. Yan, D. Tao, L. Zhang, X. Li, H.-J. Zhang, "Human Gait Recognition with Matrix Representation", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, Issue 7, July 2006.

[25] D. Tao, X. Li, X. Wu, S. J. Maybank, "General Tensor Discriminant Analysis and Gabor Features for Gait Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, Issue 10, Oct. 2007.

[26] D. Xu, Y. Huang, Z. Zeng, X. Xu, "Human Gait Recognition Using Patch Distribution Feature and Locality-Constrained Group Sparse Representation", *IEEE Transactions on Image Processing*, Vol. 21, Issue 1, Jan. 2012.

[27] C. Chen, J. Zhang, and R. Fleischer, "*Multilinear Tensor-Based Non-Parametric Dimension Reduction for Gait Recognition*", Proceedings of 3rd International Conference on Bioemtrics, 2009.

[28] J. Zhang, J. Pu, C. Chen, R. Fleischer, "Low-Resolution Gait Recognition", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 40, Issue 4, Aug. 2010.

[29] B. Guo, M. S. Nixon, "Gait Feature Subset Selection by Mutual Information", *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, Vol. 39, Issue 1, Jan. 2009.

[30] A. B. Watson and A. J. Ahumada, "Model of human visual-motion sensing",*Journal of Optical Society of America*, 2(2):322–342, 1985.

[31] J. Han and B. Bhanu, "*Statistical Feature Fusion for Gait-Based Human Recognition*", Proceeding of IEEE International Conference on Computer Vision andPattern Recognition, vol. 2, pp. 842-847, 2004.

[32] C. Liu and H. Wechsler, "*Enhanced Fisher Linear Discriminant Models for Face Recognition*", Proceeding of IEEE International Conference on PatternRecognition, vol. 2, pp. 1368-1372, 1998.

[33] R.O. Duda, P.E. Hart, and D. Stork,"Pattern Classification",2nd Edition, *Wiley Publisher*, 2000.

[34] J. Ye, R. Janardan, and Q. Li, "*Two-Dimensional Linear Discriminant Analysis*", Neural Information Processing Systems (NIPS),pp. 1569-1576, 2005.

[35] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces versusFisherfaces: Recognition Using Class Specific Linear Projection",*IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, no. 7, pp. 711-720,July 1997.

[36] A. Kale, A. Sundaresan, A.N. Rajagopalan, N.P. Cuntoor, A.K.RoyChowdhury, V. Kruger, and R. Chellappa, "Identification ofHumans Using Gait," *IEEE Transactions on Image Processing*, vol. 13, no. 9,pp. 1163-1173, Sept. 2004.

[37] Y. Chai, Q. Wang, J.P. Jia, and R. Zhao, "A Novel Human GaitRecognition Method by Segmenting and Extracting the RegionVariance Feature," Proc. Int'l Conf. Pattern Recognition, vol. 4,pp. 425-428, 2006.

[38] S. Jayaraman, S. Esakkirajan, T. Veerakumar, "Digital Image Processing", *Tata McGraw-Hill Education*, 2011.