Gait Sequence Analysis using Frieze Patterns *

Yanxi Liu, Robert Collins and Yanghai Tsin
The Robotics Institute, Carnegie Mellon University
{yanxi,rcollins,ytsin}@cs.cmu.edu

Abstract. We analyze walking people using a gait sequence representation that bypasses the need for frame-to-frame tracking of body parts. The gait representation maps a video sequence of silhouettes into a pair of two-dimensional spatio-temporal patterns that are near-periodic along the time axis. Mathematically, such patterns are called "frieze" patterns and associated symmetry groups "frieze groups". With the help of a walking humanoid avatar, we explore variation in gait frieze patterns with respect to viewing angle, and find that the frieze groups of the gait patterns and their canonical tiles enable us to estimate viewing direction of human walking videos. In addition, analysis of periodic patterns allows us to determine the dynamic time warping and affine scaling that aligns two gait sequences from similar viewpoints. We also show how gait alignment can be used to perform human identification and model-based body part segmentation.

1 Motivation

Automated visual measurement of human body size and pose is difficult due to nonrigid articulation and occlusion of body parts from many viewpoints. The problem is simplified during gait analysis, since we observe people performing the same activity with certain time period. Although individual gaits vary due to factors such as physical build, body weight, shoe heel height, clothing and the emotional state of the walker, at a coarse level the basic pattern of bipedal motion is the same across healthy adults, and each person's body passes through the same sequence of canonical poses while walking [6]. We have experimented with a simple, viewpoint-specific spatio-temporal representation of gait. The representation collapses a temporal sequence of body silhouette images into a periodic two-dimensional pattern. This paper explores the use of these frieze patterns for viewing angle determination, human identification, and non-rigid gait sequence alignment.

2 Related Work

Many approaches to analyzing gait sequences are based on tracking the body as a kinematic linkage. Model-based kinematic tracking of a walking person was pioneered by Hogg [11], and other influential approaches in this area are [2, 4].

* This work is supported in part by ONR N00014-00-1-0915 and NSF # IIS-0099597.

These approaches are often brittle, since the human body has many degrees of freedom that cannot be observed well in a 2D image sequence. Our work is more closely related to approaches based on pattern analysis of spatio-temporal representations. Niyogi and Adelson delineate a person's limbs by fitting deformable contours to patterns that emerge from taking spatio-temporal slices of the XYT volume formed from an image sequence [17]. Little and Boyd analyze temporal signals computed from optic flow to determine human identity from gait [14]. Analyzing features over a whole temporal sequence is a powerful method for overcoming noise in individual frames.

Liu and Picard [15] propose to detect periodic motions by treating temporal changes of individual pixels as 1D signals whose frequencies can be extracted. Seitz and Dyer [18] replace the concept of period by the instantaneous period, the duration from the current time instant at which the same pattern reappears. Their representation is effective in studying varying speed cyclic motions and detecting irregularities. Cutler and Davis [5] also measure self-similarity over time to form an evolving 2D pattern. Time-frequency analysis of this pattern summarizes interesting properties of the motion, such as object class and number of objects.

3 A Spatio-Temporal Gait Representation

Consider a sequence of binary silhouette images $b(t) \equiv b(x, y, t)$, indexed spatially by pixel location (x, y) and temporally by time t. Form a new 2D image $F_C(x, t) = \sum_y b(x, y, t)$, where each column (indexed by time t) is the vertical projection (column sum) of silhouette image $b(t)$, as shown in Figure 1. Each

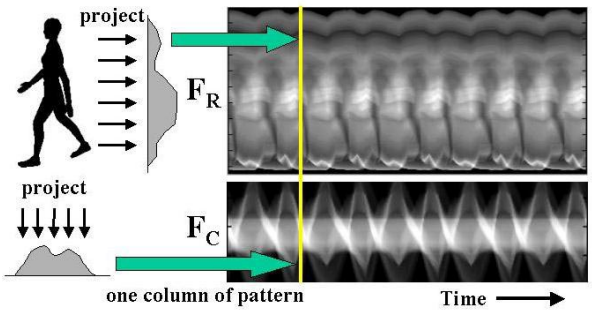


Fig. 1. Spatio-temporal gait representations are generated by projecting the body silhouette along its columns (F_C) and rows (F_R), then stacking these 1D projections over time to form 2D patterns that are periodic along the time dimension. A 2D pattern that repeats along one dimension is called a "frieze" pattern.

value $F_C(x, t)$ is then a count of the number of silhouette pixels that are "on" in column x of silhouette image $b(t)$. The result is a 2D pattern, formed by stacking

column projections together to form a spatio-temporal pattern. A second pattern $F_R(y,t) = \sum_x b(x,y,t)$ can be constructed by stacking row projections. Since a human gait is periodic with respect to time, F_C and F_R are also periodic along the time dimension. A two-dimensional pattern that repeats along one dimension is called a *frieze* pattern in the mathematics and geometry literature, and group theory provides a powerful tool for analyzing such patterns (Section 4.1).

Figure 2 shows the column projection frieze pattern F_C extracted from a roughly 30 second long sequence of a person walking along a test course. Note the changes in appearance of the frieze pattern as the walking direction changes. In our experiments, body silhouette extraction is achieved by simple background subtraction and thresholding, followed by a 3x3 median filter operator to suppress spurious pixel values. Silhouettes across a gait sequence are automatically aligned by scaling and cropping based on bounding box measurements so that each silhouette is 80 pixels tall, centered within a window 80 pixels wide by 128 pixels high. Background subtraction is a commonly used method for extract-

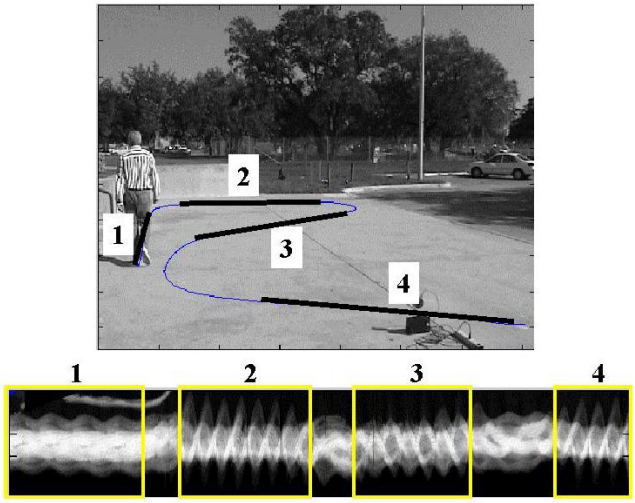


Fig. 2. Frieze pattern extracted from a 30 second long walking sequence. Note the changes in appearance of the frieze pattern as the walking direction changes.

ing body silhouettes from a stationary background scene [7, 20]. It is difficult for kinematic trackers to automatically identify and fit individual limb positions from such data. This is because background subtraction often yields noisy silhouettes with holes, fragmented boundaries, and extra parts due to background clutter and shadows. However, by distilling a sequence of silhouettes into a periodic pattern that can be smoothed and analyzed using robust signal analysis techniques, our "holistic" approach to gait analysis is better able to deal with noisy silhouette data.

4 Model-Based Gait Analysis

With the aid of a 3D walking humanoid model, we have studied how the spatio-temporal frieze patterns described above vary with respect to camera viewpoint. Our model of human body shape and walking motion is encapsulated in a VRML/H-Anim 1.1 compliant avatar called "Nancy". [1] Nancy's 3D polyhedral body parts were generated by a graphics designer, and the gait motion, specified by temporal sequences of interpolated rotations at each joint, is based on the motion sequence of a real person in "The Human Figure in Motion" by Eadweard Muybridge. We have ported Nancy into an open-GL program that generates 2D perspective views of the avatar given a camera position and time step within the gait cycle. Figure 3 illustrates variation of the column projection frieze patterns F_C defined in Section 3 when Nancy's gait is seen from different viewing directions. The diversity inspires us to seek an encoding for these different types of frieze patterns in order to determine viewpoint from frieze group type. One candidate for categorizing frieze patterns is by their symmetry groups.

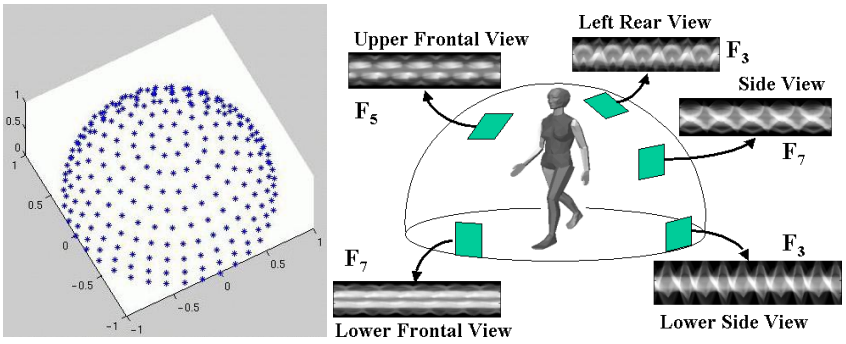


Fig. 3. (a) A database of gait sequences is generated from 241 sample viewpoints. The subject is a walking humanoid avatar (motion sequence is from a real person). (b) Some gait patterns of the avatar "Nancy" viewed from different directions.

4.1 Frieze Symmetry Groups Classification

Any frieze pattern P_i in Euclidean space R^2 is associated with a unique symmetry group F_i, where $i = 1..7, \forall g \in F_i, g(P_i) = P_i$. These seven symmetry groups are called *frieze groups*, and their properties are summarized in Figure 4 and Table 1. Five different types of symmetries can exist for frieze patterns: (1) translation, (2) 2-fold rotation, (3) horizontal reflection (4) vertical reflection, and (5) glide-reflection. A frieze pattern can be classified into one of the 7 frieze groups based

[1] ©1997 Cindy Ballreich, 3Name3D / Yglesias, Wallock, Divekar, Inc.

on what combination of these 5 primitive symmetries are present in the pattern [16] (Table 1).

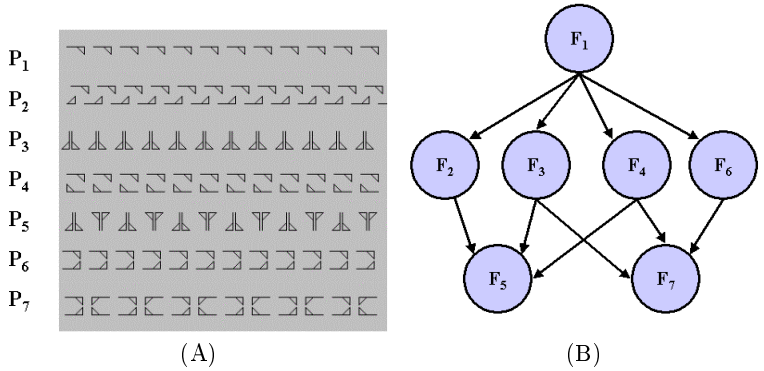


Fig. 4. (A) The seven frieze patterns ($P_1 ... P_7$) in Euclidean space R^2. (B) The subgroup relationship among the seven frieze symmetry groups ($F_1 ... F_7$ in Table 1). $F_i \rightarrow F_j$ means F_i is a subgroup of F_j.

We are interested in classifying imperfect and noise-contaminated frieze patterns generated from real human gaits. There are two important and intertwined computational issues for frieze symmetry group classification: 1) given an imperfect frieze pattern, how to decide whether or not it has certain types of symmetries; and 2) given the symmetry measures for a pattern, how to give each of the seven frieze groups an equal chance to be chosen as the symmetry group of the pattern, since these groups are *not disjoint*. The first issue is addressed by establishing a distance measure between an imperfect periodic pattern and frieze patterns. The second issue is addressed by using geometric AIC [12, 13] for symmetry group model selection.

Distance to the Nearest Frieze Patterns We define the symmetry distance (SD) of an approximately periodic pattern P to the set of all frieze patterns $\{P_n\}$ with frieze group F_n as

$$SD_n(P) = \min_{Q \in \{P_n\}} \{ \sum_{i=1}^{tN} \left(\frac{p_i - q_i}{s_i} \right)^2 \} \tag{1}$$

where N is the number of pixels in a tile (smallest 2D repeating region), t is the number of tiles being studied, p_i and q_i are intensity values of corresponding pixels of pattern P and $Q \in \{P_n\}$ respectively, and s_i is the standard deviation of the frieze pattern at pixel i. For independent Gaussian noise, the distance SD_n has a χ^2 distribution with tN degrees of freedom.

Table 1. Symmetries of frieze pattern tiles (N is number of pixels in one tile)

Symmetry Group	translation	2-fold rotation	Horizontal reflection	Vertical reflection	Glide reflection	Degrees of Freedom
F1	yes	no	no	no	no	N
F2	yes	no	no	no	yes	N/2
F3	yes	no	no	yes	no	N/2
F4	yes	yes	no	no	no	N/2
F5	yes	yes	no	yes	yes	N/4
F6	yes	no	yes	no	no	N/2
F7	yes	yes	yes	yes	no	N/4

The symmetry distance measure is defined with respect to a frieze pattern $Q \in \{P_n\}$ that has the minimal distance to P. We can show that this pattern Q can be constructed as follows: (1) For $t > 1$ and $n = 1$, Q is the pixel-wise average of all the tiles in P. (2) For $t = 1$ and $n > 1$, Q $= \frac{(\mathcal{O}_n(P)+P)}{2}$, where $\mathcal{O}_n(P)$ is the pattern obtained by applying the set of symmetry operations in F_n to P. (3) For $t > 1$ and $n > 1$, Q is the pixel-wise average of each Q obtained above. Our definition of frieze pattern symmetry distance in pixel intensity space is analogous to that of Zabrodsky et.al. [21, 13] for polygon distance in vertex location space.

Geometric AIC for Frieze Group classification The frieze symmetry groups form a hierarchical structure (Figure 4B) where frieze group F_1 is a subgroup of all the other groups and so on. For example, a frieze pattern P_3 (with vertical reflection symmetry) is a more general pattern type than P_5 or P_7, since any P_5 or P_7 frieze with more complicated symmetries also has vertical reflection symmetry. But this implies that the distance of a pattern P to P_3 is always no greater than the distance to P_5, since the set of P_5 patterns is a subset of the P_3 patterns. If no care is taken, a symmetry group classification algorithm based on raw symmetry distance scores will always favor P_3 over P_5. To address this problem, we adopt the concept of Geometric-AIC (G-AIC) proposed by Kanatani [12, 13]. Given two possible frieze patterns whose symmetry groups have a subgroup relationship, G-AIC states that we should prefer F_m over F_n if

$$\frac{SD_m}{SD_n} < 1 + \frac{2(d_n - d_m)}{r(tN) - d_n} \tag{2}$$

where d_m and d_n are the degrees of freedom for frieze patterns of F_m and F_n respectively, and r is the codimension. Since the data space (the intensity space) is dimension one, and our model space (point in multidimensional intensity space) dimension is 0, the codimension $r = 1 - 0 = 1$.

The degrees of freedom (DOF) of a frieze pattern depends on how the intensity of each pixel on the pattern is constrained. For frieze patterns with translation symmetry only, the only constraint for each of the tN pixels is to have the

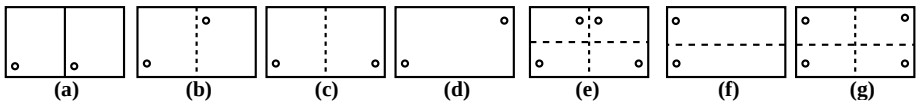


Fig. 5. Determining the degrees of freedom of frieze patterns by how many constraints a pixel intensity has to satisfy. The figure shows the corresponding pixels that must have the same intensity values in (a) two tiles of a P_1 pattern; (b)-(g) a tile from frieze pattern $P_2 ... P_7$ respectively.

same intensity value as the pixel t units to the left. Thus its DOF is N. On the other hand, pixels on a P_3 pattern have to satisfy a vertical reflection symmetry constraint, and thus half of the pixel intensities need to be the same as the other half. So the DOF of a P_3 pattern is $N/2$. The last column of Table 1 and Figure 5 explain the DOFs of the seven frieze groups. In summary, we would prefer to classify a pattern P as having frieze group F_m rather than F_n if

$$\frac{SD_m(P)}{SD_n(P)} < \frac{t}{t-1}, \text{ for } m = 2, 3, 4, 6 \text{ and } n = 1 \tag{3}$$

$$\frac{SD_m(P)}{SD_n(P)} < \frac{2t}{2t-1}, \text{ for } m = 5, 7 \text{ and } n = 2, 3, 4, 6 \tag{4}$$

$$\frac{SD_m(P)}{SD_n(P)} < \frac{2t+1}{2t-2}, \text{ for } m = 5, 7 \text{and } n = 1 \tag{5}$$

4.2 View Direction Estimation

To study the effects of viewpoint on human gait appearance we have generated a database of 241 walk sequences, indexed by viewing direction azimuth and elevation, by sampling the view sphere at roughly every 10 degrees (Figure 3). This type of detailed study is perhaps only possible given a generative model, since the cost of gathering such data experimentally would be prohibitive. Figure 6 shows the frieze groups associated with the 241 column projection frieze patterns F_C generated from Nancy's gait when viewed from these directions. Row projection frieze patterns F_R exhibit less variation with respect to group structure, and are not used in this section.

The underlying assumption in the current approach is that the distribution of symmetry groups of the gait patterns from different views of the computer model can provide guidance for determining the viewing angle of an observed human subject's gait. Figure 7 shows a comparison among corresponding avatar and two individual human subject (one male, one female) frieze patterns, viewed from six different viewing angles. One can observe that frieze patterns from the same view point share the same frieze symmetry group, and their tiles have a similar appearance. We also observe that the avatar frieze patterns are visually similar to the patterns extracted from gait video of real human subjects. In another word, the gait frieze patterns so computed are more similar across different subjects than across viewing directions.

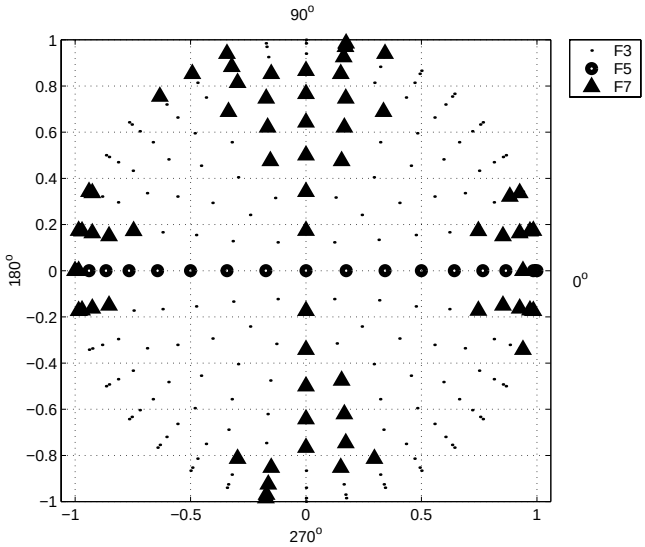


Fig. 6. Symmetry group labelings of the frieze patterns of a humanoid avatar viewed from different directions on the upper hemisphere (Figure 3). Each sample point on the hemisphere is projected to the plane of elevation 0 degrees. Since the shape and motion of the avatar is based on a real person, the symmetry group distribution map above is NOT perfectly symmetrical with respect to the 0^o (Frontal view) and 180^o (Back view) line.

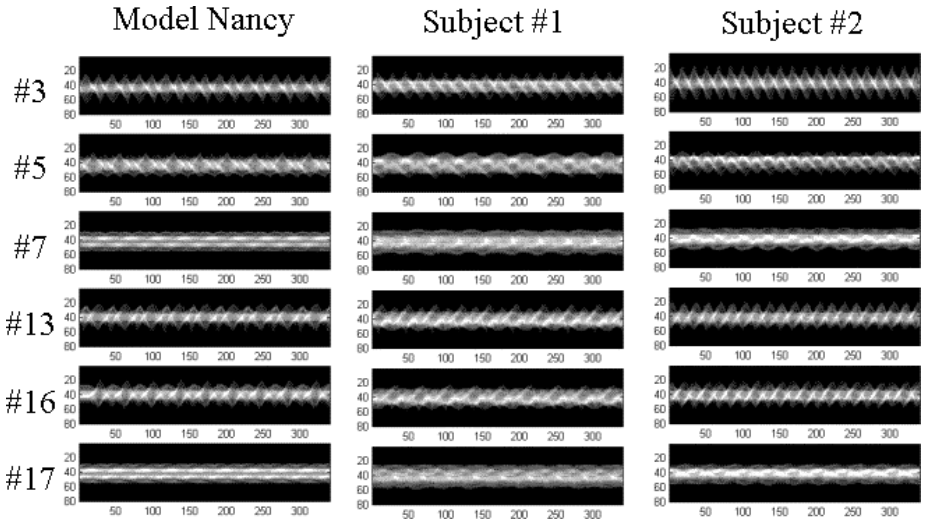


Fig. 7. View of F_C frieze patterns from six different angles. Left: Avatar Nancy (motion sequence is extracted from a real female person). Middle: human subject # 1 (male). Right: human subject # 2 (female).

Given an observed human subject gait pattern $P = F_C$ (Section 3), we use a moment-based method (Section 5.1) to align the model friezes P_i from each of the 241 candidate viewing directions to the subject frieze. Applying PCA to a typical tile from P and taking the non-dominant PCA components that are most sensitive to discriminate pattern variations, the closest K nearest neighbors are found in this subspace. We used a dual elimination method to decide which angle values from these K neighbors we can count on. The first condition is that P and P_i have the same symmetry group. The second condition is that corresponding pixels of tiles from P and P_i must have similar intensities. Results for classifying viewing direction for two human subjects is listed in Table 2.

Table 2. View direction estimation using frieze groups, for two real human subjects viewed from six different cameras

Camera	Sym.	Ground truth			subj 1 estimate		subj 2 estimate	
ID	Group	View Dir.	elevation	azimuth	elev	azim	elev	azim
3	F7	L. side	15.4	83.2	50	75	10	80
5	F3	L. front	12.0	37.4	30	160	80	45
7	F5	Frontal	25.0	359.8	20	0	20	0
13	F3	R. back	11.4	234.9	20	200	20	240
16	F3	R. front	11.9	314.5	40	334	10	20
17	F5	Back	26.5	181.4	20	180	20	180

In this framework we have assumed affine camera models, thus only one viewing direction (specified by azimuth and elevation) is estimated for a ray directed towards the center of the person. However, the data used in the experiment comes from a perspective camera, and due to the proximity of the subject, the difference in viewing ray elevation between their head and feet is roughly 28 degrees. This partly explains why estimation of azimuth angles tends to be more accurate than elevation angles. Furthermore, much more accurate estimations of viewing angles are achieved for frieze patterns with non-F_3 groups (Table 2). This can be explained by the multiple possible angle ranges for P_3 patterns (Figure 6). We have dealt with this using majority votes and robust median estimators. Although having the same symmetry group is only a *necessary* condition for two gait patterns to share the same viewing direction, our initial experiments show that this condition yields better estimation accuracy than not using symmetry groups at all. Accuracy would be further improved by using these cues in combination with other constraints, for example geometric constraints on azimuth and elevation provided by approximate knowledge of ground plane and direction of travel [19].

5 Spatio-Temporal Gait Alignment

Consider two gait sequences, represented by two pairs of frieze patterns $F_C(x,t)$, $F_R(y,t)$ and $F'_C(x',t')$, $F'_R(y',t')$. We seek to align the patterns temporally and spatially, as a precursor for further correspondence-based analysis. Temporal alignment of gait sequences amounts to aligning frieze patterns horizontally, thereby determining a mapping between time variables t and t'. Spatial alignment means finding a mapping between pixel locations (x,y) in sequence 1 and (x',y') in sequence 2. We restrict this to a four parameter affine mapping, and show that it can be found by aligning the corresponding row and column friezes along their vertical dimensions.

Spatio-temporal alignment of two video sequences is typically treated within a framework of 3D space-time volumetric warping [3]. However, representing human activity using line projection frieze patterns collapses the problem down to 2D spatial pattern alignment. Temporal alignment of these frieze patterns is further simplified by the periodic nature of the patterns themselves, allowing us to use simple periodic signal analysis in place of expensive dynamic time warping procedures [8].

5.1 Moment-Based Gait Alignment

It is well known that the first and second moments of two binary silhouettes can be used to determine an affine transformation that coarsely aligns them, and that some of the moments of a silhouette image can be computed from its row and column projections [1]. This forms the basis of our gait alignment method.

First, we generalize the concept of moments of a binary image to cover a time series of moments computed from a sequence of binary images. Define a *moment sequence* as $m_{ij}(t) = \sum_x \sum_y x^i y^j b(x,y,t)$, which is a sequence of single-frame binary silhouette moments, indexed by time. Note that $m_{00}(t)$ is just the area of the binary silhouette over time, while $\bar{x}(t) \equiv m_{10}(t)/m_{00}(t)$ and $\bar{y}(t) \equiv m_{01}(t)/m_{00}(t)$ are the coordinates of the silhouette centroid over time. Similarly, define a *central moment sequence* as $\mu_{ij}(t) = \sum_x \sum_y (x-\bar{x}(t))^i (y-\bar{y}(t))^j b(x,y,t)$, which is a sequence of moments measured after translating each silhouette so that its centroid is at the origin. The second central moments measure the spread of silhouette pixels about the centroid, and can be used to derive the principal axis of the silhouette shape.

Since we are summarizing each sequence of silhouettes with frieze patterns, we are concerned only with moments that can be computed from row and column projections. For example, consider silhouette area

$$m_{00}(t) = \sum_x \sum_y b(x,y,t) = \sum_x \left(\sum_y b(x,y,t) \right) = \sum_x F_C(x,t)$$

which can thus be computed from the frieze pattern as well as the original silhouette sequence. Any moment sequence $m_{ij}(t)$ or central moment sequence $\mu_{ij}(t)$ with either i or j (or both) equal to zero can be computed from frieze patterns

$F_C(t)$ and $F_R(t)$. In the present case, we will use $m_{00}(t)$, $m_{10}(t)$, $m_{10}(t)$, μ_{20}(t), and $\mu_{02}(t)$. Note that the second central moment $\mu_{11}(t)$ can not be determined from the two frieze patterns, and we will therefore not be able to adjust skew or principle axis rotation when aligning silhouette shapes using friezes alone.

We now present an algorithm for moment-based gait alignment. To a first approximation, the temporal alignment between the two periodic gait sequences can be represented as $t' = \rho\, t + \phi$, where ρ corrects for the relative stride frequency and ϕ corrects for the relative phase difference (position within a stride). The average stride frequency of each gait sequence is found by taking signal $m_{00}(t)$, "whitening" it by subtracting its mean and dividing by its standard deviation, then autocorrelating to find peaks occurring at a fundamental frequency. From some viewpoints this is the stride frequency, and from others it is half the stride frequency (e.g. a bipedal gait viewed from the side looks self-similar halfway through a full stride). Whether the autocorrelation of $m_{00}(t)$ yields peaks at half the stride frequency is viewpoint dependent, and can be calibrated using the walking avatar model. Let f and f' denote the average frequencies of the two gait sequences, computed from m_{00} of sequence 1 and m'_{00} of sequence 2. Then $\rho = f'/f$. To determine the relative phase, we crop a subsequence of temporal length f from $m_{00}(t)$, expand or contract it by ρ, then correlate with m'_{00}. The average lag of prominent peaks of the correlation result determines the relative phase. There may be a two-fold ambiguity in the phase from those viewpoints for which the autocorrelation of m_{00} yields peaks at half the stride frequency. For people close to the camera, the perspective effects are usually enough to uniquely determine the phase. For people far away, however, it can be difficult to distinguish between left foot forward or right foot forward on the basis of silhouette moment information alone.

After determining the temporal mapping between t and t', we now align the frieze patterns spatially. Given the moments that we can compute from frieze patterns, we determine the two translations and two scale factors that relate (x, y) and (x', y') for corresponding time steps in the two sequences. Dropping the time variables from the notation, this affine transformation is found to be

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \sqrt{\frac{\mu'_{20}\, m_{00}}{\mu_{20}\, m'_{00}}} & 0 \\ 0 & \sqrt{\frac{\mu'_{02}\, m_{00}}{\mu_{02}\, m'_{00}}} \end{bmatrix} \begin{bmatrix} x \;-\; m_{10}/m_{00} \\ y \;-\; m_{01}/m_{00} \end{bmatrix} \;+\; \begin{bmatrix} m'_{10}/m'_{00} \\ m'_{01}/m'_{00} \end{bmatrix}$$

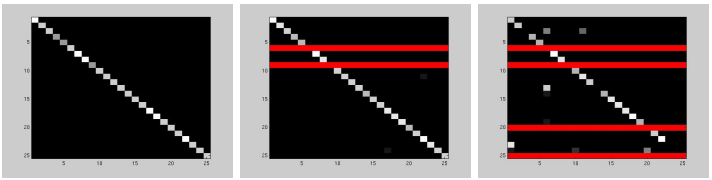
Whether to allow both scale factors to vary independently for each time step, to enforce their ratio to be constant, to compute a temporal average for each, or other variations depends on the application and on the amount of noise one can expect in the underlying silhouette data.

5.2 Applications of Gait Alignment

We illustrate the utility of moment-based frieze alignment with two applications. The first involves comparing frieze tiles to classify a walking person's identity given a prior training set of gait data. The second application concerns matching

a walking humanoid model to gait silhouette data from a real person, in order to locate specific body parts in each frame.

Human Identification Given a dataset of gait sequences collected from one camera viewpoint, we want to analyze a new sequence to determine which person it is. Our approach is to create row and column silhouette projection friezes for each sequence, warp them all temporally to a canonical frequency and phase using the first half of the above alignment procedure, then cut out several tiles corresponding to individual strides from each sequence. These aligned frieze tiles are compared using normalized correlation, and subject classification is performed by nearest neighbor matching on correlation scores. This approach implicitly captures biometric shape cues such as body height/width ratio, body-part proportions, stride length and amount of arm swing.

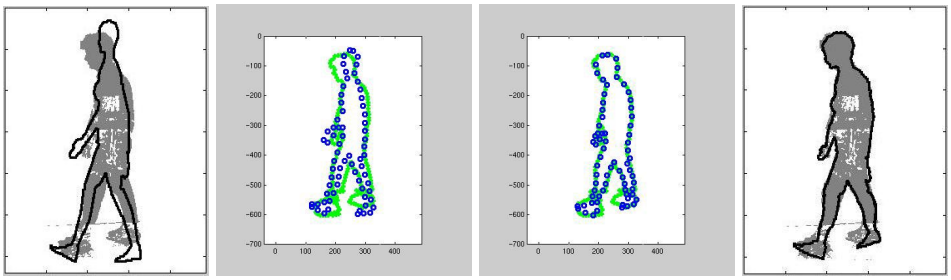


(a) (b) (c)

Fig. 8. Confusion matrices for nearest neighbor classification of 25 human subjects using gait frieze pattern tiles. (a) Result from training and testing on non-overlapping slow walking gait sequences. Classification rate is 100%. (b) Training on slow walk, testing on fast walk. Classification rate is 100%. (c) Training on slow walk, testing on walking carrying a ball (to inhibit arm swing). Classification rate is 81%. Blank rows in (b) and (c) denote subjects for which there is no corresponding test data available.

To test this approach, we use the CMU MoBo database [9], which contains motion sequences of 25 subjects walking on a treadmill. Each subject is recorded performing four different types of walking: slow walk, fast walk, inclined walk, and slow walk holding a ball (to inhibit arm swing). Figure 8 shows results achieved for side views, for gait combinations slow-slow, slow-fast and slow-ball. For the slow-slow experiment, the gallery consisted of tiles from the first five seconds of each subject's slow walk gait sequence, and the probe set consists of tiles from the last five seconds of the same sequences. For both slow-fast and slow-ball, the classification algorithm is trained on all tiles from the slow walk sequences, and tested on all tiles from the other two gait sequences. We see that the results are quite good, even across different gait types. Although the match similarity metric is simple normalized correlation, each tile succinctly represents both spatial and temporal information from an entire stride subsequence.

Model-Based Body Part Analysis Assume that we know the camera viewpoint, and have rendered a walking humanoid model from that viewpoint. We now have a sequence of model body silhouettes that can be matched against a real gait sequence. After spatio-temporal gait alignment, the temporal pairing of each frame of the data sequence with a corresponding model frame is known, along with a four parameter affine transformation that aligns those two binary silhouettes. Thus, for each frame, we can project the model silhouette contour onto the data silhouette image. A sample frame showing an overlayed model contour found through automatic gait sequence alignment is shown in Figure 9A. The aligned model contour does not exactly coincide with the person's body



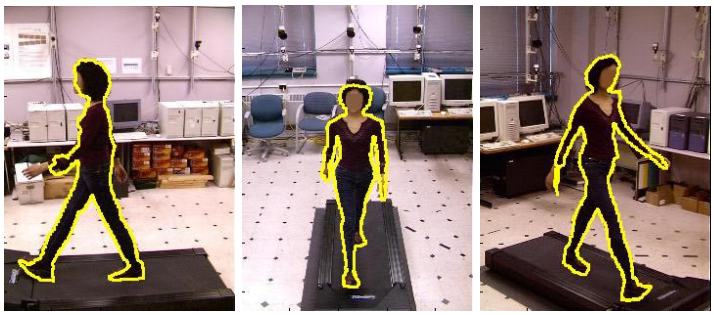
(a) (b) (c) (d)

Fig. 9. (a) Moment-based alignment of model and data silhouettes. (b) Sampled points from model and data silhouette contours. (c) Results of non-rigid thin-plate spline alignment of the two sets of sample points. (d) Model silhouette warped by thin-plate spline transform, overlayed on data silhouette.

outline due to a variety of factors, including differences in body shape and joint angle kinematics between the avatar and the human being (e.g. body proportions and amount of arm swing), as well as small differences in camera perspective between the model and data viewpoints. However, note that the overall temporal and spatial alignment is quite good, in the sense that the aligned model tells us what body parts should be visible, and roughly where they should appear in the image. More importantly, we know which body parts are occluded and should not be considered for further analysis in this frame.

To illustrate what can potentially be done given this initial alignment between model and data silhouettes, we uniformly sample points from along each silhouette contour and use a program for non-rigid point matching to compute a thin-plate spline transformation between them [10]. Figure 9 shows, from left to right, the initial model contour alignment, the two sampled point sets, the resulting point sets after warping non-rigidly by a thin-plate spline, and the new warped model contour overlayed over the data silhouette. The agreement between contours is now much improved. The success of the non-rigid point matcher in this case is due in large part to the accuracy of the model silhouette

topology, as determined by moment-based alignment of gait frieze patterns. More examples are shown in Figure 10. Model-based gait analysis using frieze patterns offers an efficient alternative to kinematic body part tracking for determining the location of individual body parts in each frame of a gait sequence.



(a) (b) (c)

Fig. 10. Spatially and temporally aligned model silhouette overlayed on original image for three views taken simultaneously by synchronized cameras. We plan to use results like these for further model-based body part analysis, ultimately leading to 3D body reconstruction and motion capture of a walking human.

6 Summary

We have presented a periodic pattern representation for analyzing gait sequences. Silhouette row and column projections are stacked over time to form frieze patterns that can be analyzed using the mathematical theory of symmetry groups. With the help of a walking humanoid avatar, we have studied the correlation between the seven frieze symmetry groups and gait viewing direction, and have developed practical techniques for classifying imperfect frieze patterns. Our future work will explore methods for more efficient and accurate viewpoint estimation from frieze patterns, and extend our mathematical methods for imperfect pattern analysis to patterns that are periodic along two dimensions. We have also presented a moment-based method for aligning frieze gait patterns both temporally and spatially. The method has applications in determining human identity from gait biometrics, and it provides an efficient alternative to frame-by-frame tracking approaches for locating and delineating body parts.

References

1. B.K.P.Horn. *Robot Vision*. MIT Press, 1986.
2. C. Bregler and J. Malik. Tracking people with twists and exponential maps. In *Proc. IEEE Computer Vision and Pattern Recognition*, pages 8–15, 1998.
3. Y. Caspi and M. Irani. A step towards sequence-to-sequence alignment. In *IEEE Computer Vision and Pattern Recognition*, pages II:682–689, 2000.
4. T.J. Cham and J.M. Rehg. A multiple hypothesis approach to figure tracking. In *Proc. IEEE Computer Vision and Pattern Recognition*, pages II:239–245, 1999.
5. R. Cutler and L. Davis. Robust real-time periodic motion detection, analysis and applications. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 22(8):781–796, August 2000.
6. E.Ayyappa. Normal human locomotion, part 1: Basic concepts and terminology. In *Journal of Prosthetics and Orthotics*, volume 9(1), pages 10–17. The American Academy of Orthotists and Prosthetists, 1997.
7. A. Elgammal, D. Harwood, and L. Davis. Non-parametric model for background subtraction. In *European Conference on Computer Vision*, pages 751–767, 2000.
8. M.A. Giese and T. Poggio. Morphable models for the analysis and synthesis of complex motion patterns. *International Journal of Computer Vision*, 38(1):59–73, June 2000.
9. R. Gross and J. Shi. The CMU motion of body (MoBo) database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon University, 2001.
10. H.Chui and A.Rangarajan. A new algorithm for non-rigid point matching. *IEEE Computer Vision and Pattern Recognition*, pages 44–51, 2000.
11. D. Hogg. Model-based vision: A program to see a walking person. *Image and Vision Computing*, 1(1):5–20, 1983.
12. K. Kanatani. *Statistical Optimization for Geometric Computation : Theory and Practice*. North-Holland, 1996.
13. K. Kanatani. Comments on "Symmetry as a Continuous Feature. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(3):246–247, 1997.
14. J.J. Little and J.E. Boyd. Recognizing people by their gait: The shape of motion. *Videre*, 1(2), 1998.
15. F. Liu and R. W Picard. Finding periodicity in space and time. In *IEEE International Conference on Computer Vision (ICCV)*, 1998.
16. Y. Liu and R. T. Collins. A Computational Model for Repeated Pattern Perception using Frieze and Wallpaper Groups. In *Computer Vision and Pattern Recognition Conference (CVPR'00)*, pages 537–544, Los Alamitos, CA, June 2000. IEEE Computer Society Press. (http://www.ri.cmu.edu/pubs/pub_3302.html).
17. S.A. Niyogi and E.H. Adelson. Analyzing and recognizing walking figures in xyt. In *Proc. IEEE Computer Vision and Pattern Recognition*, pages 469–474, 1994.
18. S.M. Seitz and C.R. Dyer. View-invariant analysis of cyclic motion. *International Journal of Computer Vision*, 25:1–23, 1997.
19. T.N. Tan, G.D. Sullivan, and K.D. Baker. Recognizing objects on the ground-plane. *Image and Vision Computing*, 12(3):164–172, April 1994.
20. K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and practice of background maintenance. In *International Conference on Computer Vision*, pages 255–261, 1999.
21. H. Zabrodsky, S. Peleg, and D. Avnir. Symmetry as a continuous feature. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(12):1154–1165, December 1995.