Gait-Based Human Recognition by Classification of Cyclostationary Processes on Nonlinear Shape Manifolds

David KAZISKA and Anuj SRIVASTAVA

We study the problem of analyzing and classifying human gait by modeling it as a stochastic process on a shape space. We consider gait as a evolution of human silhouettes as seen in video sequences, and focus on their shapes. More specifically, we define a shape space of planar, closed curves and model a human gait as a stochastic process on this space. Due to the periodic nature of human walk, this process is naturally constrained to be cyclostationary, that is, its mean path is assumed to be cyclic. We compare two subjects using a metric that quantifies differences between average gait cycles of each subject. This computation uses several tools from differential geometry of the shape space, including computation of geodesics, estimation of means of observed shapes, interpolation between observed shapes, and temporal registration of two gait cycles. Finally, we apply a nearest-neighbor classifier, using the gait metric, to perform human recognition, and present results from an experiment involving 26 subjects.

KEY WORDS: Biometrics; Gait recognition; Shape analysis; Shape classification; Statistics on shape manifolds.

1. INTRODUCTION

In this work we study the problem of analyzing videos of humans with a goal of recognizing them by analyzing their gait. In the field of biometrics, there is a strong need to recognize subjects from a distance, especially in noncooperative environments. In this situation images of faces, fingerprints, or irises may not be available; a common solution is to use the style of walking, called *gait*, to recognize people. Human gait is valuable as a biometric because it can be observed from a distance and requires no physical contact. Gait recognition has been a problem of interest since the seminal work of Niyogi and Adelson (1994), where the authors took advantage of the periodic nature of gait and expressed an individual's gait as a combination of a canonical walk and individual variation. In the subsequent literature, the approaches have been divided into shape-based and motion-based approaches. In the shape-based approaches, researchers use landmark representations or binary images of the subjects at various points in the gait cycle (Liu, Malave, and Sarkar 2004a). Representations of gait using a mean shape on a (landmark-based) shape manifold have been used (Wang, Ning, Hu, and Tan 2002; Liu et al. 2004b), but without exploiting the gait's periodic nature. Motion-based methods often use parametric representations of certain bodily movements to represent gait. Some motion-based models combine parametric representations of arm and leg swing, stride length and cadence and combine this information with static data, such as height (Cuntoor, Kale, and Chellappa 2003; BenAbdelkader, Cutler, and Davis 2002, 2004; Foster, Nixon, and Prugel-Bennett 2003). Kale, Rajagopala, Cuntoor, Krueger, and Chellappa (2004) constructed a hidden Markov model that combines the shape- and motion-based methods to some extent.

Our study of gait recognition uses a shape-based approach based on discriminating shapes of silhouettes of humans during walking. It is an extension of a traditional statistical shape analysis, in which single, individual shapes are compared, to a comparison of sequences of shapes. Shape analysis traces its origins to the work of D. G. Kendall (1984) who used a representation of shape by an indexed set of landmarks (salient points on objects) in the plane. Book-length treatments of this approach include those of Small (1996) and Dryden and Mardia (1998). Landmark-based representations have been successful in many applications, especially in physician-assisted medical image analysis where landmarks are readily available, and have led to seminal advances in statistical analysis of shapes (Dryden and Mardia 1998; Kent and Mardia 2001; Holboth, Kent, and Dryden 2002). Asymptotic theory for this approach has been well developed by, among others, Bhattacharya and Patrangenaru (2002, 2003). Recently, it has been extended to make it view-independent by constructing a manifold of shapes that are within a projective transformation of one another (Mardia and Patrangenaru 2005), a thorough development of parametric distributions on high-dimensional landmark shape spaces was provided by Dryden (2005). The study of stochastic processes on shape manifolds traces back to work on diffusion on shape and Brownian motion on shape manifold by D. G. Kendall (1977, 1984, 1991), W. S. Kendall (1988, 1990), and Le (1991, 1994).

Our approach to shape analysis essentially follows the framework of Klassen, Srivastava, Mio, and Joshi (2004) and Mio and Srivastava (2004). The basic idea is to find a set of all relevant closed curves, quotient out all shape-preserving transformations from this set, and use the resulting quotient space for statistical shape analysis. Those authors constructed shape manifolds of parameterized curves, with individual parameterized curves represented by their angle functions and speed functions, to represent and analyze shapes. Similar constructions were also used by Younes (1998, 1999). Michor and Mumford (2004) studied the problem of comparing shapes of closed curves under various metrics. Using the representation and metrics described by Klassen et al. (2004), Srivastava, Joshi, Mio, and Liu (2005) presented techniques for clustering, learning, and testing planar shapes. The work of Klassen et al. (2004) was restricted to arc-length parameterization of curves; Mio and Srivastava

David Kaziska is Assistant Professor, Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH 45433 (E-mail: *david.kaziska@afit.edu*). Anuj Srivastava is Associate Professor, Department of Statistics, Florida State University, Tallahassee, FL 32306 (E-mail: *anuj@stat.fsu.edu*). This research was supported in part by grants ARO W911NF-04-01-0268 and ARO W911NF-04-1-0113. This article was presented in part at the 9th European Conference on Computer Vision, Graz, Austria, May 2006.

Journal of the American Statistical Association
???? 0, Vol. 0, No. 00, Applications and Case Studies
DOI 10.1198/

(2004) extended this by relaxing to permit variable-speed parameterization, thus introducing a representation that allows both bending and stretching of shapes to match them with one another. The analysis resulting from this approach seems more natural because interesting features, such as corners, are better preserved while constructing statistics in this approach. Statistical shape models on this manifold have been studied by Srivastava, Jain, Joshi, and Kaziska (2006).

Tools used in shape analysis often come from differential geometry and statistics. To help the reader with some relevant concepts from differential geometry, we present a short introduction in Appendix A. We define important quantities, such as geodesics, exponential maps, group actions, and Karcher means, all in a general setting. In Section 2 we construct the specific shape manifold of interest (i.e., the shape space of elastic curves), and in Section 3 we particularize tools from differential geometry to this shape manifold. In Section 4 we present our framework for gait recognition, in which we explore the concept of cyclostationary processes on the shape manifold. We define cyclostationary processes and develop a method for their classification. We also present important ingredients of our framework, including automatic detection of gait cycles, registration of gait sequences, computation of a mean gait cycle, and matching using geodesic distances. In Section 5 we illustrate our experimental setup using a infrared (nighttime) video camera and present some classification results. We successfully matched 17 of 26 individuals for a set of gait sequences that we extracted from the infrared video.

2. SHAPE SPACE OF ELASTIC CURVES

Our approach to gait recognition is based on modeling the evolution of human silhouettes as a stochastic process on a shape space. This space, comprising shapes of all simple closed curves in $\mathbb{R}^2$, is a nonlinear space, and tools from differential geometry are needed to perform calculus on this space. To help the reader understand geometric terminology (especially a reader not familiar with differential geometry), we provide a short introduction of relevant ideas from differential geometry in Appendix A.

The particular shape manifold that we use in our application is the space of elastic curves introduced by Mio and Srivastava (2004). This space is an extension of the space constructed by Klassen et al. (2004) who used the direction functions to represent individual shapes associated with planar, simple closed curves. To develop this representation, start by denoting each curve with a differentiable function $\alpha : \mathbb{R} \to \mathbb{R}^2$, which gives parameterized coordinates. That is to say at a time s, the vector $\alpha(s) \in \mathbb{R}^2$ gives the Cartesian coordinates of the curve. We further restrict our attention to curves that are periodic with period 2π (because periodicity makes curves closed). A closed curve α is called a *simple* closed curve if it does not intersect itself. Our representation of curves is in terms of a pair of functions (ϕ, θ), $\phi, \theta : [0, 2\pi] \mapsto \mathbb{R}$, such that at any point $s \in [0, 2\pi]$, we have $\alpha'(s) = \exp(\phi(s)) \exp(j\theta(s))$, where $j = \sqrt{-1}$. $e^{\phi(s)}$ is the instantaneous speed of α at s, and thus ϕ is called the *log-speed function*. The $\theta(s)$ is the angle formed by the vector $\alpha'(s)$ with the positive x-axis, and, thus, θ is called the direction function of the curve α. For an arc-length parameterized curve, ϕ is identically 0 but generally is a nonzero function.

Consider the space $\mathcal{C}$ of all closed curves of length 2π and average direction π in $\mathbb{R}^2$ given by

$$\mathcal{C} = \left\{ (\phi, \theta) \Big| \int_0^{2\pi} e^{\phi(s)} e^{j\theta(s)}\, ds = 0, \right.$$
$$\left. \int_0^{2\pi} e^{\phi(s)}\, ds = 2\pi, \frac{1}{2\pi} \int_0^{2\pi} \theta(s) e^{\phi(s)}\, ds = \pi \right\}.$$

$\mathcal{C}$ is called the *preshape space*. Note that the variability generated by shape-preserving transformations (e.g., rotation, translation, and scale), are already removed in this representation. In other words, if a curve α is rotated, translated, or scaled, then its representation in $\mathcal{C}$ remains unchanged. But, the variability resulting from different placements of origin on the closed curve α and different reparameterizations of $[0, 2\pi]$ remain. In other words, if we place $s = 0$ at different points along the curve, or simply reparameterize the interval $[0, 2\pi]$, then we will get a new pair $(\tilde{\phi}, \tilde{\theta}) \in \mathcal{C}$, but the shape of the underlying α will remain unchanged. We identify all elements of $\mathcal{C}$ that have the same shape using algebraic equivalence as follows. The variability resulting from different placements of origin can be modeled as a group action of $\mathbb{S}^1$, the unit circle, on $\mathcal{C}$, and that reparameterization results from the group action of $\mathcal{D}$, the set of all automorphisms $\{\gamma : [0, 2\pi] \mapsto [0, 2\pi]\}$, on $\mathcal{C}$. For a definition of a group action on manifolds, refer to Appendix A. The action of $\mathbb{S}^1$ on $\mathcal{C}$ is given by the following: For a $s_0 \in \mathbb{S}^1$, the curve $(\phi(s), \theta(s))$ becomes

$$s_0 \cdot (\phi(s), \theta(s)) \equiv \big(\phi(s - s_0), \theta(s - s_0)\big).$$

An automorphism γ changes the representation of a curve according to

$$(\phi, \theta) \circ \gamma = (\phi \circ \gamma + \log \gamma', \theta \circ \gamma).$$

All of the curves generated by changing the origin or reparameterizing the same curve are considered to be of the same shape. Therefore, the shape space is defined to be a quotient space $\mathcal{S} = \mathcal{C}/(\mathbb{S}^1 \times \mathcal{D})$. In $\mathcal{S}$, each point represents a distinct shape, and $\mathcal{S}$ becomes the space in which statistical analysis of shapes is to be performed. In the next section we particularize tools from differential geometry to this shape space $\mathcal{S}$ for use in our application on gait recognition.

3. TOOLS FOR ANALYSIS ON SHAPE SPACE $\mathcal{S}$

Our representation of a human gait consists of a temporal sequence of simple, closed curves, which are the silhouettes of individuals in video sequences. Focusing on the shapes of these silhouettes, we consider them to be points on the shape space $\mathcal{S}$. Our method of gait recognition exploits the differential geometry of $\mathcal{S}$ various ways, including interpolating gait sequences, registering gait sequences, computing a mean gait cycle, and performing nearest-neighbor matching for gait recognition. In all of these tasks, our understanding of the differential geometry of $\mathcal{S}$ plays a central role. We note that the our preshape space $\mathcal{C}$ is a complete Riemannian manifold and, consequently, between any two points of this space, a geodesic can be defined and computed. Furthermore, because the shape space $\mathcal{S}$ is a quotient space of $\mathcal{C}$, geodesics in $\mathcal{S}$ are computed as shortest geodesics connecting equivalence classes in $\mathcal{C}$.

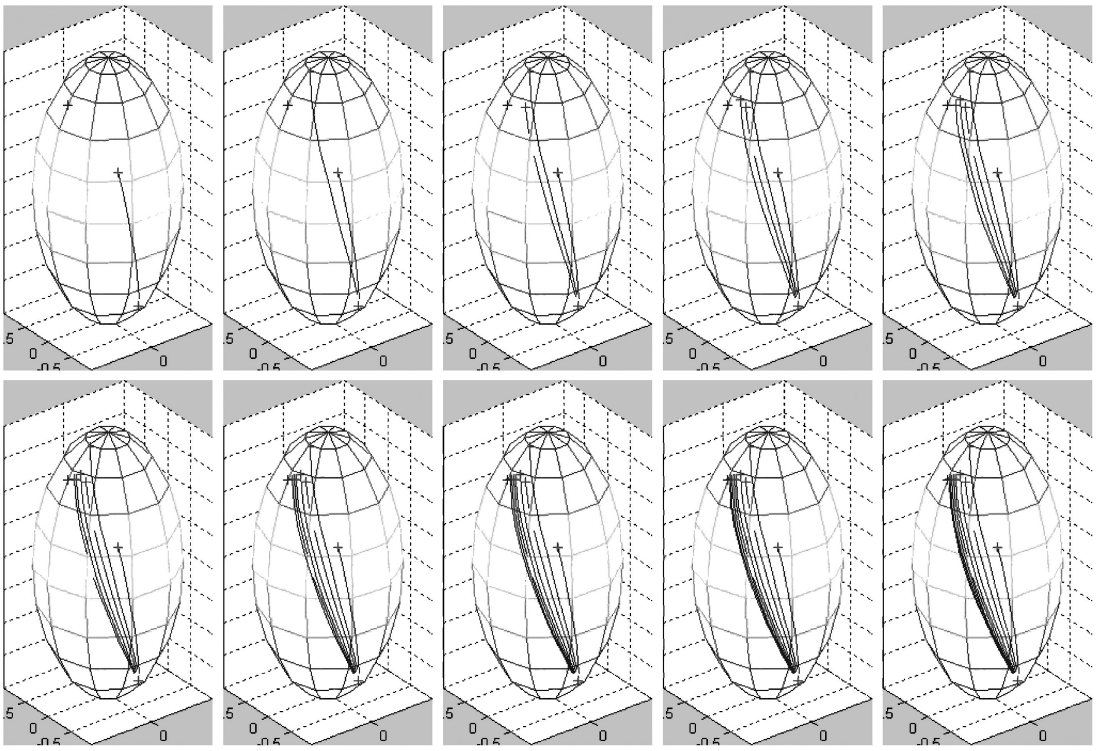

Figure 1. Ten iterations of the shooting method for computing geodesics. In each iteration, the shooting direction is corrected to get closer to the target point.

3.1 Geodesic Paths on Shape Space

An efficient technique for quantifying shape differences is to compute geodesic paths in $\mathcal{S}$ connecting shapes, and then use their lengths to quantify shape differences. Because $\mathcal{S}$ is a quotient space of $\mathcal{C}$, the geodesic length between any two shapes in $\mathcal{S}$ is given by

$$d_s((\phi_1,\theta_1),(\phi_2,\theta_2)) = \min_{s_0\in\mathbb{S}^1,\gamma\in\mathcal{D}} d_c\big((\phi_1,\theta_1), s_0\cdot(\phi_2,\theta_2)\circ\gamma\big).$$

Here d_c and d_s are the distances between points on $\mathcal{C}$ and $\mathcal{S}$. This equation states that we fix one shape (ϕ_1,θ_1) and seek the reparameterization of the other shape (ϕ_2,θ_2) that minimizes the geodesic distance from (ϕ_1,θ_1) in $\mathcal{C}$. The minimization over s_0 can be easily performed using an exhaustive search over $\mathbb{S}^1$, but the search for optimal γ deserves a closer look. We seek a reparameterization function $\gamma\in\mathcal{D}$ of (ϕ_2,θ_2) such that it minimizes the matching cost

$$H(\gamma) = \int_0^{2\pi} \big(\lambda\big\|(\phi_1(s),\theta_1(s)) - (\phi_2(s),\theta_2(s))\circ\gamma\big\|^2 + (1-\lambda)|\gamma'(s)|^2\big)\,ds,$$

$0<\gamma<1$, and for a fixed $\lambda>0$. The minimization is performed using the dynamic programming algorithm, which is briefly outlined in Appendix B. Figure 2 shows some examples of this matching. Note that in the figure, the matching process works well whether the legs are apart or together, hands are visible or not visible, and so on. This estimation of γ is also called the *registration* of one curve to another. This registration of curves is similar in concept to the registration of gait sequences that we introduce in the next section, although the two cases differ in the actual quantities being registered.

Once the two curves are registered, they can be treated as elements of $\mathcal{C}$, and the computation of a geodesic is based on a shooting method described in Appendix A. To illustrate the resulting geodesics, Figure 3 shows three examples of geodesic paths between shapes of human silhouettes. In each row, the shapes denote equally spaced points along a geodesic connecting the starting and ending shapes. For any two shapes (ϕ_1,θ_1) and (ϕ_2,θ_2), the length of geodesic path between them forms a natural tool for comparing them and is denoted by $d_s((\phi_1,\theta_1),(\phi_2,\theta_2))$. Also, we use the function $\Psi(t)$ to denote

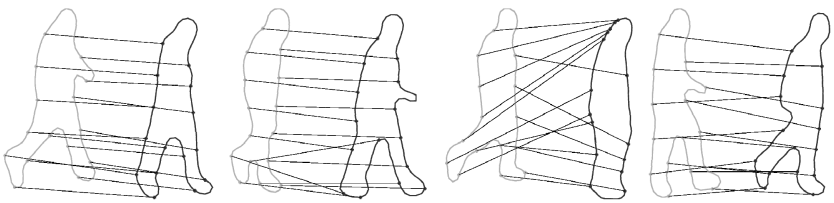

Figure 2. Shape matching. Each pair shows a γ that minimizes the matching cost between the two shapes; the lines show the corresponding points between the two curves.

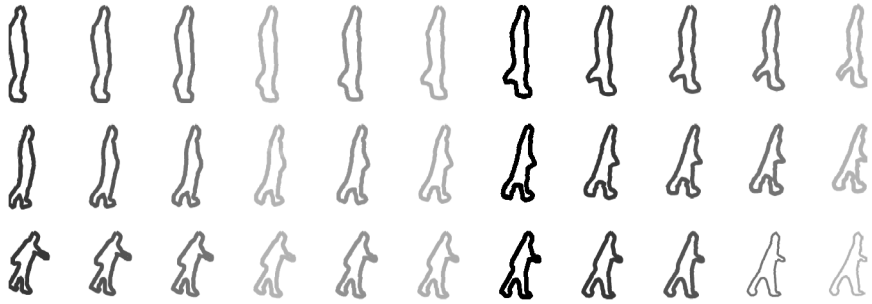

Figure 3. Examples of geodesic paths between human shapes in $\mathcal{S}$. In each row, the first and the last shapes are given, and the intermediate shapes depict points along geodesics connecting these given shapes.

the geodesic path, so that $\Psi(0) = (\phi_1, \theta_1)$ and $\Psi(1) = (\phi_2, \theta_2)$. Here $\dot{\Psi}(0)$, the initial velocity vector along this path, is useful in computing mean shapes.

3.2 Piecewise-Geodesic Interpolations Between Shapes

Our framework considers the underlying gait sequence as a stochastic process on $\mathcal{S}$. In practice, one has only a finite, discrete set of observations along a sample path when the human walk is captured using a video camera. For the purpose of aligning (or registering) different sequences, as described later, it will be useful to fill in between the observed shapes to result in dense time samples. We use a piecewise-geodesic interpolation to fill in between the observed shapes. In other words, we connect each successive pair of observed shapes with geodesic paths on $\mathcal{S}$, and in this way we can obtain an arbitrarily dense time sampling. Interpolation is an important tool in our framework because it allows us to analyze a gait process at arbitrary time resolutions. To compare two gait sequences, we use a metric, given in (3), that requires samples of sequences at simultaneous, discrete points in time. When the time points do not correspond to the times at which shapes are observed, we will need to interpolate on $\mathcal{S}$ to estimate the silhouette at those times. To guard against bias that can be introduced in this interpolation, we performed a limited amount of interpolation. We performed only the amount of interpolation needed to standardize the number of silhouettes in a gait cycle.

Suppose that we wish to interpolate between two consecutive observed silhouettes having representations (ϕ_i, θ_i) and $(\phi_{i+1}, \theta_{i+1})$. Using the shooting method for computing geodesics, we compute a geodesic between the two points. The resulting geodesic Ψ from (ϕ_i, θ_i) will reach the target shape $(\phi_{i+1}, \theta_{i+1})$ in unit time, $\Psi(0) = (\phi_i, \theta_i)$ and $\Psi(1) = (\phi_{i+1}, \theta_{i+1})$. We can then evaluate $\Psi(t)$ for any $t \in (0, 1)$ to continuously interpolate between the two observed shapes.

To demonstrate this idea experimentally, we present an example in Figure 4. We start with a gait sequence consisting of six silhouettes, which represents a half-cycle of gait for this individual. To illustrate geodesic interpolation, we first drop the third shape in this sequence and use the second and fourth shapes to estimate it using interpolation. The first contour in the second row is this interpolated shape. Note its closeness to the actual observed third shape in the original sequence. The same experiment is repeated for the fourth shape; it is estimated by interpolating between the third and fifth shapes. The estimated shape in shown as the second shape in the second row. Because alternate techniques in the literature involve binary images, we performed an interpolation on the binary images produced from the second and fourth silhouettes; the resulting in-

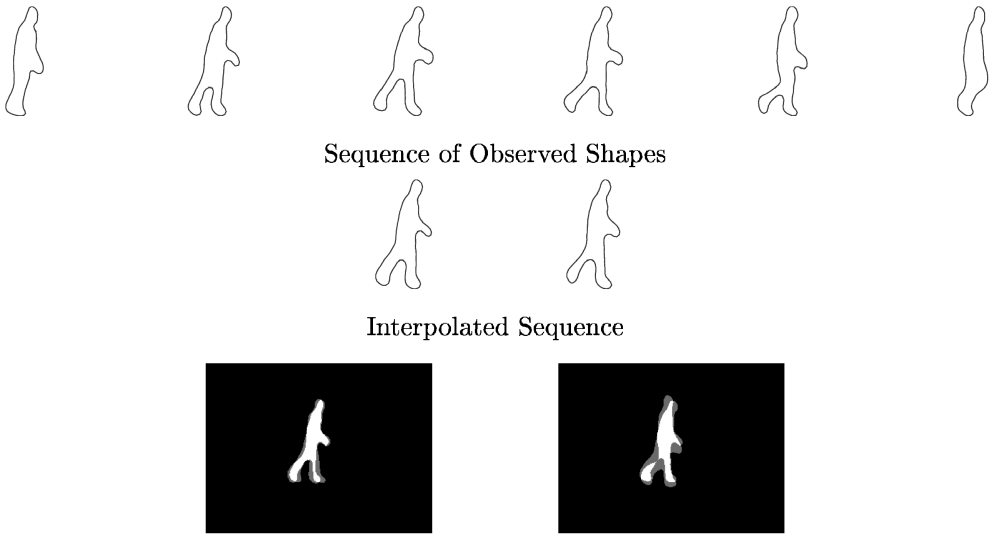

Figure 4. Illustration of gait cycles. The sequence from legs together to right leg forward to legs together is the first half-cycle. The sequence from legs together to left leg forward to legs together is the second half-cycle.

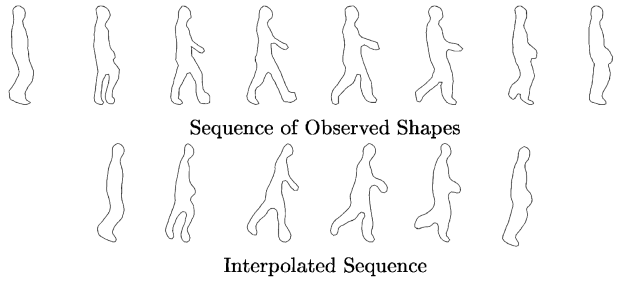

Figure 5. Detection of gait cycles. On the left is an ordered sequence of silhouettes. On the right is a plot of the geodesic distance from the first shape to each subsequent shape. The peaks at shapes 5, 12, 18, and 25 correspond to the beginnings and ends of cycles.

terpolated gray-scale image, is shown in the third row of Figure 4. Similarly, we interpolated between the third and fifth silhouettes and compared both methods with the observed fourth silhouette.

3.3 Computation of Mean Shapes

In the proposed statistical framework for gait representation and analysis, another important ingredient is computation of the statistical mean of observed shapes. In other words, given a finite collection of points on $\mathcal{S}$, we would like to define and compute a quantity that represents the central tendency of that dataset. For this purpose, we use the notion of a Karcher mean, which is essentially the centroid of the sampled data points on $\mathcal{S}$ using the geodesic distance d_s. This quantity has also been called the Frechet mean or the intrinsic mean in the literature. Given a set of sample shapes $(\phi_1, \theta_1), \ldots, (\phi_n, \theta_n) \in \mathcal{S}$, the sample Karcher variance is a function of $(\phi, \theta) \in \mathcal{S}$ and is given by $V(\phi, \theta) = \sum_{k=1}^{n} d_s((\phi, \theta), (\phi_k, \theta_k))^2$. The Karcher mean set of $(\phi_1, \theta_1), \ldots, (\phi_n, \theta_n) \in \mathcal{S}$ is the set of minimizers of $V(\theta)$. Several authors, including Klassen et al. (2004), have used an iterative algorithm for computation of the sample Karcher mean.

4. HUMAN IDENTIFICATION USING GAIT ANALYSIS

In this section we present a statistical framework for performing gait analysis and its use in human recognition. Human gait is modeled as a stochastic process on the shape space $\mathcal{S}$. With this formulation, we are representing silhouettes of people walking using simple closed curves. We recognize that there can be times when armswing or leg movements prevent a silhouette from being a simple closed curve. We did not observe such cases in practice, but in this framework we would represent this situation by the simple closed curve obtained from the outermost contour.

Because human walk is quite naturally a periodic process, we restrict ourselves to a cyclostationary process, defined next.

Definition 1 (Cylostationary). A stochastic process is called *cyclostationary* with a period τ if the joint probability distribution of the random variables $X(t_1), X(t_2), \ldots, X(t_n)$ is same as that of $X(t_1 + \tau), X(t_2 + \tau), \ldots, X(t_n + \tau)$, for all $t_1, t_2, \ldots, t_n$ and for all n. In particular, the random quantities $X(t)$ and $X(t + \tau)$ have the same probability distribution.

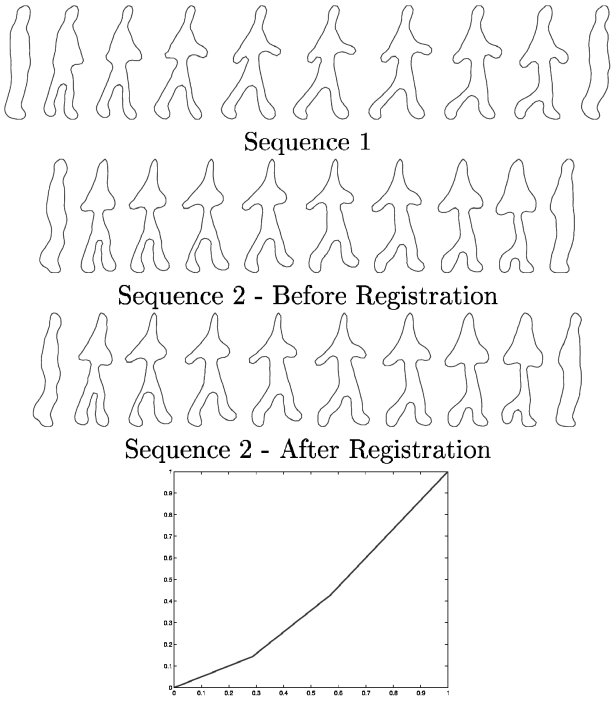

Figure 6. Observed gait cycles for two different people. The correspondence problem determines which shapes in the first and second sequences are to be compared.

Previous applications of cyclostationary processes have included the fields of signal processing and meteorological science. These applications have mostly involved low-dimensional signals in Euclidean spaces. To the best of our knowledge, ours is the first use of the cyclostationary structure on infinite-dimensional, nonlinear manifolds.

We model the shape process generated by temporal evolution of human silhouettes as a cyclostationary process on $\mathcal{S}$. For any cyclostationary process, the notion of a cycle is central to its analysis. For a gait process, we consider a *full cycle* as the period starting from when legs and hands are all together to the time of return to a similar state, as shown in Figure 7. The top row of Figure 7 shows the first half-cycle where the left foot goes forward and the right foot catches up, and the bottom row shows the second half-cycle where the right foot moves first. In our illustrations in this section and for our classification results in the next section, we use half-cycles. We assume that gait sequence associated with a person is a cyclostationary process on $\mathcal{S}$. The duration of a cycle corresponds to the period τ of the process. Like any cyclostationary process, it suffices to study a gait sequence within the period $[0, \tau]$. As a result of the cyclostationary nature of gait, in our notation we generally consider the time t to be modulo τ, so that $t \in [0, \tau)$. Given two stochastic processes, our main goal is to quantify differences between them. Let $X(t)$ and $Y(t)$ be two gait processes on the shape space $\mathcal{S}$, with periods τ_x and τ_y. We seek a metric $d_p(X, Y)$ with certain desired properties. There are two possibilities for such a metric: mean of squared distances and squared difference between average paths.

Mean of Squared Distances. The first idea is to compute the expected value of the squared distance between the sample

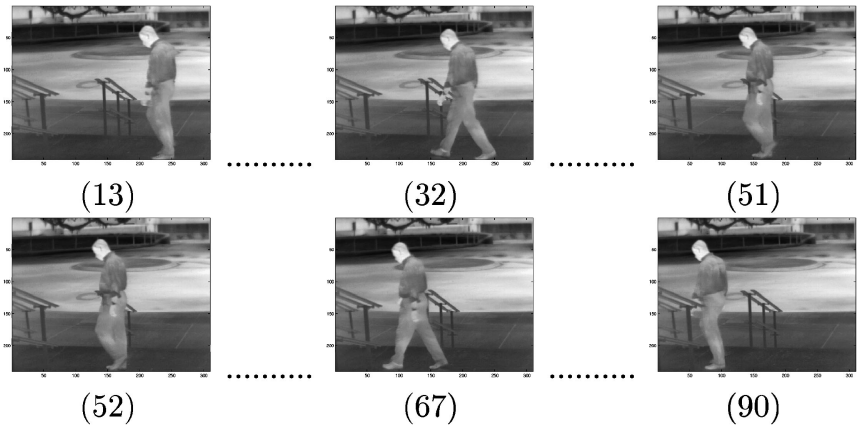

Figure 7. Demonstration of interpolation between points in $\mathcal{S}$. The first row shows the original gait sequence of six shapes. The second row shows results of interpolation on the shape space for the third and fourth shapes. The interpolated shapes are drawn directly below the original third and fourth shapes for comparison. In the third row we interpolate the corresponding binary images, producing the gray-scale images that correspond to the same shapes.

paths of X and Y,

$$d_p(X,Y) = \min_{\kappa\in[0,\tau_y],g} E\left[\left(\int_0^{\tau_x} d_s\big(X(t), Y(\kappa+g(t))\big)^2\,dt\right)\right]^{1/2}, \quad (1)$$

where E denotes the expected value, $d_s(\cdot,\cdot)$ denotes the geodesic length metric defined on $\mathcal{S}$, g denotes a smooth mapping between $[0,\tau_x]$ and $[0,\tau_y]$, and κ denotes a possible relative time shift between the two observed gait sequences. The mapping g is needed to register shapes along the two sequences before comparison.

Squared Distance Between Average Paths. The second idea is to use squared distance between the mean cycles,

$$d_p(X,Y) = \min_{\kappa\in[0,\tau_y],g}\left(\int_0^{\tau_x} d_s\big(E[X(t)], E\big[Y(\kappa+g(t))\big]\big)^2\,dt\right)^{1/2}, \quad (2)$$

where expectations $E[X(t)]$ and $E[Y(t)]$ are computed on the shape space $\mathcal{S}$.

We note that these functionals may not be symmetric; however, in either case a true metric can be constructed using $d(X,Y)+d(Y,X)$. Although both distances are relevant, the latter is more efficient from a computational standpoint. The first definition requires computing distances between several random realizations of gait cycles of X and Y. In other words, we need to compute distances between all training and test cycles associated with X and Y. In contrast, the second definition requires computing an average gait cycle each for X and Y, then computing a distance between the average cycles. In the context of human recognition using a distance-based classifier, X denotes the gait process of a known person (training data) and Y denotes that of an unknown person (test data). Therefore, once the average gait cycle for the known person is computed during training, it can be simply stored for future classification purposes. It can be viewed as a *template gait cycle* for this person; this is similar to the idea of *deformable template* theory for classification.

Consequently, we choose the second metric and use its discrete form,

$$d_p(X,Y) = \min_{\kappa\in[0,\tau_y],g}\left(\sum_{t=1}^{\tau_x} d\big(\overline{X}(t), \overline{Y}(\kappa+g(t))\big)^2\right)^{1/2}, \quad (3)$$

where $\overline{X}$ and $\overline{Y}$ are sample Karcher means of the processes X and Y. With a slight abuse of notation, we continue to use τ_x and τ_y as lengths of gait cycles in the discrete case as well. Now $g:\{1,2,\ldots,\tau_x\}\mapsto\{1,2,\ldots,\tau_y\}$ is a mapping that registers points across the two sequences. Estimation of κ and g is discussed in the next two sections.

4.1 Automatic Detection of Cycles

The first issue in computing the distance given in (3) is the extraction of gait cycles for each person. We typically have large sequences of human silhouettes generated from videos, and we need to extract a few gait cycles for analysis. Because the process is repetitive, the geometry of the shape space provides a method for automatically detecting cycles, because similar shapes should occur at the same point in the cycle. In our application to human gait, we noticed that the silhouettes with the arms and legs together are far away, in terms of geodesic distance, from the points at which the limbs are extended. To identify the beginnings and ends of cycles in a sequence of observed shapes, we begin with a silhouette with the limbs extended, then compute the geodesic distances from that first shape to all of the following shapes. Because the shapes with arms and legs together are far from shapes with limbs extended, then the distance that we compute shows peaks at the shapes with the limbs together. We detect the beginnings and ends of cycles by looking for these peaks. An example of this is shown in Figure 8.

This automated detection of cycles obviates the need for minimizing over κ in (3). The resulting metric is now

$$d_p^*(X,Y) = \min_{g}\left(\sum_{t=1}^{\tau_x} d\big(\overline{X}(t), \overline{Y}(g(t))\big)^2\right)^{1/2}. \quad (4)$$

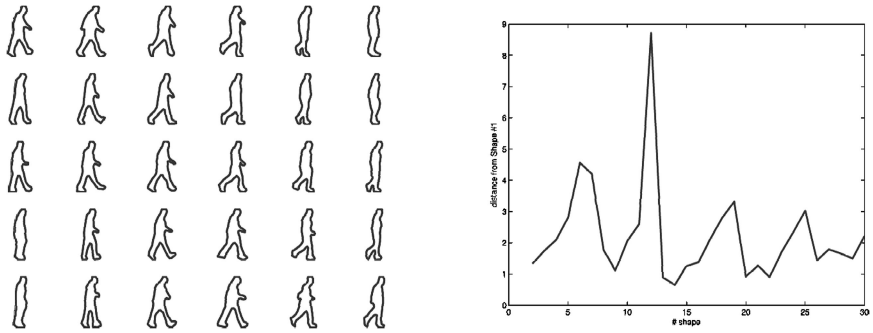

Figure 8. Example of interpolation by linear time scaling. The first row shows a sequence of eight silhouettes taken from a video clip. The second row shows the interpolation and the uniform resampling of this sequence to result in a sequence of length six.

4.2 Registration of Gait Cycles

Once we have a method for automatically extracting cycles, the next important problem in comparison and recognition of processes is the registration of points along two processes. To formalize, let τ_x and τ_y be the periods of two processes, X and Y, and let $g : [0, \tau_x] \mapsto [0, \tau_y]$ be a map that is invertible, with both g and g^{-1} having piecewise continuous derivatives. Our goal is to find g's that minimize energy functions such as those given in (1)–(4). In a discrete implementation, the registration of any two discrete cycles amounts to solving the equation

$$\hat{g} = \underset{g}{\operatorname{argmin}} \sum_{t=1}^{\tau_x} d\big(\overline{X}(t), \overline{Y}(g(t))\big)^2.$$

To explain this problem further in the context of human gait cycles, given samples of shapes along two observed walks, we need to determine which shapes to compare. Even though the shapes form an ordered sequence, there may be a time scaling, time warping, and/or time shifting between the two sequences. Figure 9 shows examples of observed gait cycles for two people. One cycle contains 10 shapes, and the other contains 8 shapes. To compare these two sequences, we need to determine which shape in the second sequence corresponds to a given shape in the first sequence. The shape that we need may be an observed shape from the second sequence or may be a shape occurring between two shapes from the first sequence. In the latter case, we may need to interpolate to estimate the shape that we need. Thus two issues necessitate the registration of processes:

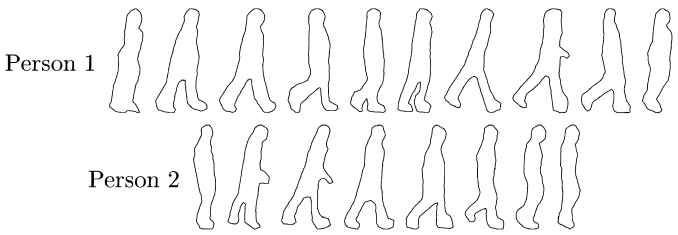

Figure 9. Example of registration by DTW. The first row shows a sequence of 10 silhouettes from a training sequence. The second row, shows a test sequence that we wish to register. The third **???** row shows the sequence after registration and the function ϕ, found by dynamic programming, used in the registration.

- Registration of two mean cycles to compare them. In computing the metric in (4), we compare two mean cycles at points in time. It may occur that process speeds within the cycles vary between the mean sequences that we compare. If so, then registering the mean cycles may provide a more accurate match.
- Registration of processes to compute a mean cycle. Registration of observed processes is necessary to compute a mean cycle for a particular observed process. The cycles that we observe in an observed process may have different lengths if the process changes speed. If the process changes speed within a cycle, then we may wish to correct for the speed change. For both of these purposes, we must register the observed sequences within the sequence before estimating the mean cycle.

We achieve registration using one of two techniques: linear time scaling or time warping using dynamic programming.

Linear Time Scaling. The simplest idea is to consider g simply as a linear time scaling, $g(t) = \beta t$, where $\beta > 0$ is a scalar. When the endpoints of the two cycles are known, β is simply the ratio τ_y/τ_x. The underlying assumption here is that speeds of two processes X and Y are constant during their observed cycles, and we can match them by simply rescaling timewise. We illustrate this idea with an example in Figure 5. Consider that a half-cycle, Y, shown in the top row, has come from an observed sequence, and we wish to register this sequence in time to another half-cycle of length 6 X (not shown). We obtained the silhouettes in the top row at eight uniformly spaced points in time, which we denote by $t = 0$–7. Using $\beta = 5/7$, we can register the first half-cycle to the second half-cycle, but we need silhouettes at the six uniformly spaced points, $t = 0, \frac{7}{5}, \frac{14}{5}, \frac{21}{5}, \frac{28}{5}, 7$, for process Y. To estimate the shapes occurring at noninteger units of time, we use piecewise geodesic interpolation on Y; for example, to get a point at time $t = \frac{7}{5}$, we travel for time .4 along the geodesic path from the shape at $t = 1$ to the shape at $t = 2$. The remaining shapes in the second row of Figure 5 are computed similarly.

Time Warping Using Dynamic Programming. In cases where speeds of processes vary within their cycles, the dynamic time warping (DTW) approach (Bellman 2003) provides an alternative to linear interpolation that may provide a more accurate match. Given processes X and Y on intervals $[0, \tau_X]$ and $[0, \tau_Y]$, seeks a diffeomorphism, $g : [0, \tau_X] \to [0, \tau_Y]$, such that

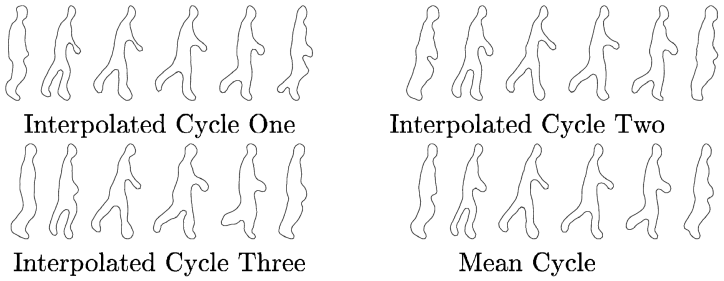

Figure 10. Computation of a mean cycle. The first three consist of gait cycles registered using linear interpolation. The mean cycle is in the fourth set. Each shape in the fourth row is the Karcher mean of the three corresponding shapes.

the metric in (5) is minimized. Rather than a linear form of the function g, we generalize to a function that can have the effect of speeding up or slowing down an observed cycle, as necessary. We may need to do this before computing a mean cycle if the subject changes speeds during a cycle. This technique also may improve our ability two match two observed mean cycles. Using techniques described earlier, it is often possible to directly detect the start and end of a shape gait cycle in a given sequence. Therefore, once the cycles have been detected, the need to estimate κ is obviated, and we can directly estimate g using dynamic programming. Several authors have used this DTW idea in different contexts. This solves for

$$\hat{g} = \underset{g}{\operatorname{argmin}} \sum_{t=1}^{\tau_x} d\big(X(t), Y(g(t))\big)^2 \, dt. \tag{5}$$

Figure 6 illustrates DTW showing one half-cycle from two different gait sequences X and Y for which we wish to register the second sequence to the first. We show the registered sequence and note that it provides a visually better match in time for the first sequence, and we plot the function g that we used to perform the registration.

4.3 Computation of Mean Gait Cycles

To use (4) for comparing and classification of processes, we need to estimate the Karcher mean shape, $E[X(t)]$, for relevant times in a cycle. Assuming that we have observed multiple observations cycle of each process, the first task is to register the shapes across cycles, as described earlier, and then compute the means of the corresponding shapes in $\mathcal{S}$. The mean shape $E[X(t)]$ is defined as the sample Karcher mean shape at time t and computed as explained in Section 3.3. Figure 10 shows an example of a calculated mean gait cycle.

5. EXPERIMENTAL RESULTS

5.1 Experimental Setup

As an application of gait analysis, we used a night-vision or infrared, (IR) video camera to observe human gait. In addition to collecting data in noncooperating environments, and from a distance, an IR camera provides the added benefit of working in dark. This is especially well suited to surveillance and urban battlefield scenarios, where a major portion of imaging is done using IR sensors. Figure 11 shows an example of video sequence obtained using an IR camera. The task of extracting human silhouettes from video sequences was semiautomated; that is, it required some human intervention in addition to automated programs. We do not describe this process of silhouette extraction and assume that the data are available as sequences of silhouettes for each human subject. Our experimental results are based on a collection of IR video clips of 26 student and faculty volunteers from Florida State University. We collected at least two clips of each person and formed disjoint training and tests. We performed a gait-matching experiment following these steps:

- For each of training and test sequence, we extracted three half-cycles, performed registration using linear time scaling, and then computed an average gait cycle.
- For each test sequence, we computed the metric in (3) for each training sequence and sought the nearest match.

We observed anecdotally that the silhouettes in individuals' gaits were different depending on which leg was leading, usually due to different armswings. As a result, the half-cycles that we used were from a single side for each individual.

5.2 Classification Results

The classification results are summarized in Table 1. Under the nearest-neighbor (NN) criterion, we obtain a successful match for 17 of the 26 test sequences. For three-NN classifiers, where the correct class in the training set must be among the three neighbors of the test cycle, the success rate was 21 out of 26. An example of a correct match is shown in the top row of Figure 12, whereas an incorrect match is shown in the bottom row.

We implemented a few current approaches to compare them with our results. First, we implemented a simple method, called

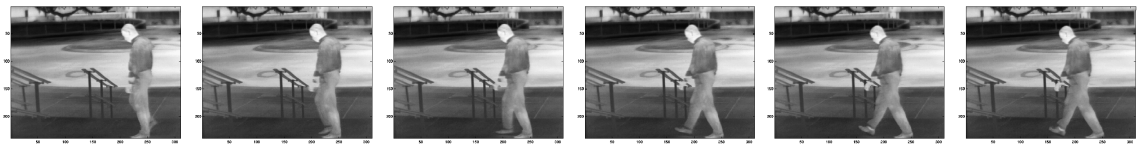

Figure 11. A small portion of a sample IR video sequence taken at 30 frames/second.

Table 1. Recognition performance

i	1	2	3	4	5	6	7	8	9	10
Our approach	17	20	21	21	21	22	22	23	24	24
Mean-shape approach	13	14	16	19	19	20	20	20	20	22
Landmark-based approach	10	11	14	17	19	19	20	20	21	22

NOTE: For each test cycle, we see whether the correct class is in the first i classes, ranked according to the metric in (3). This table shows the number of test cycles out of 26 for which that is the case, plotted against i.

the mean-shape *approach*. Some works in the literature suggest that gait recognition can be achieved using merely a mean shape of the cycle, rather than the full cycle (as in Liu et al. 2004b and Wang et al. 2002). For each person in the training set, these methods compute a single mean shape. Then, for a test sequence, a single mean shape is computed, and the best match in the training set is sought. Surprisingly, decent performance has been reported with this simplified method. With a slight variation, we computed two means for each gait sequence rather than a single mean. We computed a mean for each person when the arms and legs were together and a second mean of the silhouettes when the arms and legs were extended. Figure 13 shows a mean sequence of length 6 as used in the previous method, along with the two means for the same person that we computed for the present method. Table 1 summarizes results from this approach under the mean shape method. Finally, we also computed recognition performance using the landmark-based shape analysis of boundary curves. Although the general approach here is same as our method, the choices of shape space $\mathcal{S}$, geodesic lengths $d(\cdot,\cdot)$, Karcher means, and so on are different (Dryden and Mardia 1998). In this computation, we represent each silhouette by 100 uniformly spaced pseudolandmarks. Recognition results based on this method are also reported in Table 1.

6. SUMMARY

In this work we have presented a novel framework for gait recognition, considering gait as a cyclostationary process on a shape space of simple closed curves. Geometric tools allowed us to perform interpolation between shapes, registration of gait cycles, averaging of gait cycles, and comparisons of gait cycles for human recognition. By comparing mean cycles rather than the cycles themselves, we suppressed intraclass variability and improved the classification performance. We have demonstrated the classification technique on a set of 26 individuals.

APPENDIX A: DIFFERENTIAL GEOMETRY: BACKGROUND

In this section we provide a short introduction to ideas from differential geometry that are relevant to our approach. (For more detailed discussion, see, e.g., Spivak 1979; do Carmo 1976; Lang 1999.)

A manifold M is a set that, among other properties, has the important property of being locally Euclidean; that is, for any point $p \in M$, there exists a one-to-one mapping between a neighborhood of p and an open subset of a Euclidean space. For any point $p \in M$, consider the collection of all parameterized curves, $c : (-\varepsilon, \varepsilon) \to M$, with $c(0) = p$. Then T_pM is the collection of all vectors $c'(0)$, and it is a vector space even when M is not a vector space. If we can define an inner product on T_pM that varies smoothly with p, then M is called a Riemannian manifold, with that inner product as its Riemannian metric. A group G is said to act on M if there exists a mapping $G \times M \mapsto M$, denoted by $g \circ M$, that satisfies the following properties:

- We have $e \circ p = p$ for all $p \in M$, where e is identity element of G.
- If $g_1, g_2 \in G$, then $g_1 \circ (g_2 \circ p) = (g_1 \cdot g_2) \circ p$ for all $p \in M$. Here $g_1 \cdot g_2$ denotes the group operation in G.

A distance measure on a Riemannian manifold M is indispensable for comparing elements of M. Because distances on M are realized as lengths of geodesics, we start by introducing geodesics. A geodesic on a Riemannian manifold is a path that is locally length-minimizing. For two points $p, q \in M$, we let $c : [0, 1] \to M$ be a differentiable curve on M connecting p and q, that is, $c(0) = p$ and $c(1) = q$. If we let $c'(t)$ be the derivative of c, then the length of the curve c from p to q is

$$L_0^1(c) = \int_0^1 \|c'(t)\|_g \, dt,$$

where $\|\cdot\|_g$ is the norm induced by the Riemannian metric on M. The distance between any two points p and q is then

$$(p, q) = \min_c \{\mathcal{L}_0^1(c) : c(0) = p, c(1) = q\}.$$

The curves c for which this minimum is achieved are geodesics of the manifold M. Generally, geodesics on a M are constant-speed curves on M, but this class of curves may contain curves that are not minimizing. For example, on the manifold $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$, arcs of great circles are geodesics, but a great-circle arc is minimizing only if its length is π or less. We often denote $\Psi(p, v, t)$ to be the point obtained by starting at p and traveling for a time t along a geodesic path in the direction of $v \in T_pM$.

On simpler manifolds, such as a sphere, analytical expressions for geodesics are readily available. However, on other manifolds, including our shape space $\mathcal{S}$, analytical derivation of a geodesic is intractable, and we must use a computational technique. One possibility is a shooting method, a numerical technique for computing geodesics on complicated nonlinear manifolds. Given two points $p_1, p_2 \in M$, the goal is to compute a geodesic path between p_1 and p_2. This objective is

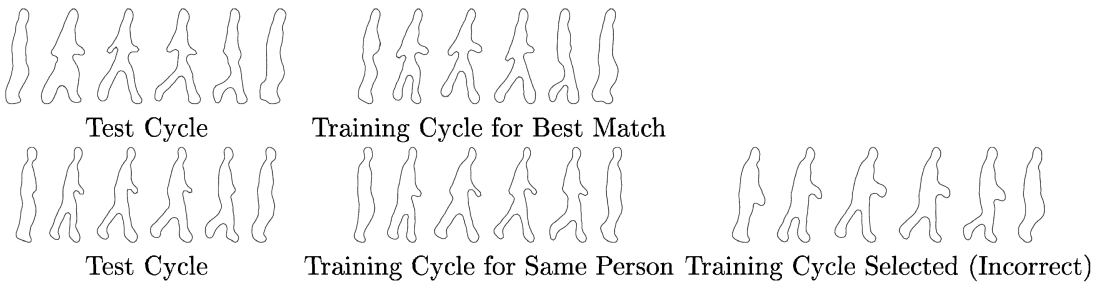

Figure 12. Examples of a correct match and an incorrect match.

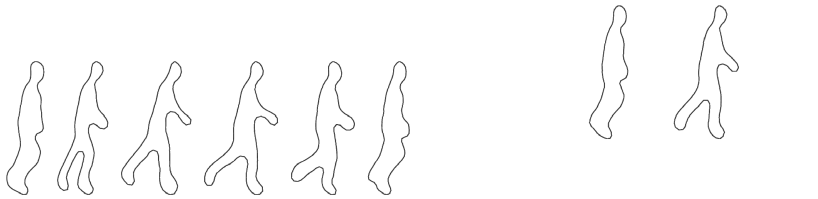

Figure 13. An example of the training sequence used in mean-shape–based classification. On the left is a mean cycle of six shapes used in method 1. On the right are the two mean shapes used in the mean-shape approach.

achieved by calculating a vector $v \in T_{p_1}M$ such that $\Psi(p_1, v, 1) = p_2$. In the shooting method, we begin with an initial approximation of v, then use a gradient method to minimize the miss distance between successive estimates of $\Psi(p_1, v, 1)$ and the target point p_2. The shooting method for this minimization provides a sequence of approximations to the path that we seek. Figure 1 gives an example of implementation of the shooting method on an ellipsoid, showing successive approximations to the path between two points.

A method by which we can go back and forth between M and T_pM is the exponential map. This is formally defined as follows.

Definition 2 (Exponential map). Let M be a manifold and let $p \in M$ and $v \in T_pM$. Let $c:[0,1] \to M$ be the unique parameterized geodesic with $c(0) = p$ and $c'(0) = v$. The exponential map of v at p is $\exp_p(v) = c(1)$.

Definition 3 (Inverse exponential map). Let M be a manifold and let $p, q \in M$. We let $c:[0,1] \to M$ be a minimizing geodesic with $c(0) = p$ and $c(1) = q$. The inverse exponential map of q at T_pM is any vector v such that $\exp_p(v) = q$.

We also need the ability to compute a sample mean shape on our shape space. On a manifold, the sample mean shape is taken to be the sample Karcher mean defined as follows.

Definition 4 (Karcher mean). Suppose that $Y_1, \ldots, Y_m$ is a set of points on M. Then the sample Karcher mean set is the set of minimizers of the function

$$F(q) = \frac{1}{m}\sum_{i=1}^{m} d(q, Y_k)^2, \qquad q \in M.$$

We say that the sample Karcher mean exists if the minimizer is unique.

The Karcher mean is sometimes called the intrinsic mean, because it uses an intrinsic metric on a manifold as opposed to an extrinsic mean computed by embedding a manifold in an ambient Euclidean space. There exist iterative techniques for finding the Karcher mean that use the gradient of F to update the estimate until convergence. A commonly used method updates the mean μ as

$$\mu \to \exp_\mu(v), \quad \text{where } v = \frac{1}{n}\sum_{i=1}^{m} \exp_\mu^{-1}(Y_i).$$

APPENDIX B: DYNAMIC PROGRAMMING

Dynamic programming is an important tool in the registration of curves and sequences, and it can help solve the following general problem. For two given functions $f:[0,a] \mapsto \mathbb{R}$ and $h:[0,b] \mapsto \mathbb{R}$, find a function $g:[0,a] \mapsto [0,b]$ that minimizes the functional,

$$\int_0^a \big(f(x) - h(g(x))\big)^2\, dx.$$

In addition, g is assumed to be differentiable and to have a positive derivative at each point. An overview of the DP algorithm is as follows. We divide the region $[0,a] \times [0,b]$ into a grid with $N \times N$ nodes. Explicitly, the node (i, j) in the grid represents the point $(\frac{i}{N-1}a, \frac{j}{N-1}a)$. We obtain a piecewise-linear approximation to g on $[0,a]$ and construct a g that passes through nodes on the grid.

For any node in the grid, define the cost $H(i, j)$ at node (i, j) to be the minimum cost of reaching (i, j) on a strictly increasing path from $(0, 0)$, and set $H(0, 0) = 0$. Now our problem is to find a path that achieves $H(N-1, N-1)$. To minimize the cost, $H(i, j)$, of reaching (i, j), we first denote $N_{(i,j)}$ to be the set of nodes that are permissible predecessors to the node (i, j). Elements of $N_{(i,j)}$ are selected to ensure that the slope of a candidate path g is positive at each point and are close to (i, j).

For $(k, l) \in N_{(i,j)}$, the cost of traveling to (i, j) through a path that contains (k, l) is the minimum cost of traveling to (k, l), plus the cost of traveling from (k, l) to (i, j) along a linear path. Letting $E((i, j), (k, l))$ denote the cost of traveling to (i, j) through a path that contains (k, l), this says that $E((i, j), (k, l)) = H(k, l) + L((k, l), (i, j))$, where $L((k, l), (i, j))$ is the cost of traveling from (k, l) to (i, j) in a straight line connecting those two points. Now the cost of reaching node (i, j) is $H(\hat{k}, \hat{l})$, where

$$(\hat{k}, \hat{l}) = \operatorname*{argmin}_{(k,l) \in N_{(i,j)}} E((k, l), (i, j)).$$

We find $\hat{g}$ by computing $H(n-1, n-1)$ in an iterative manner. In doing so, we obtain nodes $(u_1, v_1), \ldots, (u_k, v_k) = (N-1, N-1)$, representing the critical path.

[Received May 2006. Revised January 2007.]

REFERENCES

Bellman, R. (2003), *Dynamic Programming* (Dover Edition), **?????**: Dover.

BenAbdelkader, C., Cutler, R., and Davis, L. (2002), "Person Identification Using Automatic Height and Stride Estimation," presented at the **?????** International Conference on Pattern Recognition, **?????**.

——— (2004), "Gait Recognition Using Image Self-Similarity," *EURASIP Journal on Applied Signal Processing*, 4, 572–585.

Bhattacharya, R., and Patrangenaru, V. (2002), "Nonparametric Estimation of Location and Dispersion on Riemannian Manifolds," *Journal of Statistical Planning and Inference*, 108, 23–35.

——— (2003), "Large-Sample Theory of Intrinsic and Extrinsic Sample Means on Manifolds I," *The Annals of Statistics*, 31, 1–19.

Cuntoor, N., Kale, A., and Chellappa, R. (2003), "Combining Multiple Evidences for Gait Recognition," presented at the **?????** International Conference on Acoustics, Speech, and Signal Processing, Hong Kong, **?????**.

do Carmo, M. (1976), *Differential Geometry of Curves and Surfaces*, Englewood Cliffs, NJ: Prentice-Hall.

Dryden, I. L. (2005), "Statistical Analysis on High-Dimensional Spheres and Shape Spaces," *The Annals of Statistics*, 33, 1643–1665.

Dryden, I. L., and Mardia, K. V. (1998), *Statistical Shape Analysis*, Chichester, U.K.: Wiley.

Foster, J., Nixon, M., and Prugel-Bennett, A. (2003), "Automatic Gait Recognition Using Area-Based Metrics," *Pattern Recognition Letters*, 24, 2489–2497.

Holboth, A., Kent, J. T., and Dryden, I. L. (2002), "On the Relation Between Edge and Vertex Modeling in Shape Analysis," *Scandinavian Journal of Statistics*, 29, 355–374.

Kale, A., Rajagopala, A. N., Cuntoor, N., Krueger, V., and Chellappa, R. (2004), "Identification of Humans Using Gait," *IEEE Transactions on Image Processing*, 13, 1163–1173.

Kendall, D. G. (1977), "The Diffusion of Shape," *Advances in Applied Probability*, 9, 428–430.

——— (1984), "Shape Manifolds, Procrustes Metrics, and Complex Projective Spaces," *Bulletin of the London Mathematical Society*, 16, 81–121.

——— (1991), "The Mardia–Dryden Distribution for Triangle: A Stochastic Calculus Approach," *Journal of Applied Probability*, 28, 225–230.

Kendall, W. S. (1988), "Symbolic Computation and the Diffusion of Triads," *Advances in Applied Probability*, 20, 775–797.

——— (1990), "The Diffusion of Euclidean Shapes," in *Disorder in Physical Systems*, eds. G. R. Grimmett and D. J. A. Welch, Oxford: Oxford University Press, pp. 203–217.

Kent, J. T., and Mardia, K. V. (2001), "Shape, Procrustes Tangent Projections and Bilateral Symmetry," *Biometrika*, 88, 469–481.

Klassen, E., Srivastava, A., Mio, W., and Joshi, S. (2004), "Analysis of Planar Shapes Using Geodesic Paths on Shape Spaces," *Pattern Analysis and Machine Intelligencer*, 26, 372–383.

Lang, S. (1999), *Fundamentals of Differential Geometry*, New York: Springer.

Le, H.-L. (1991), "A Stochastic Calculus Approach to the Shape Distribution Induced by a Complex Normal Model," *Mathematical Proceedings of the Cambridge Philosophical Society*, 109, 221–228.

——— (1994), "Brownian Motions on Shape and Shape-and-Size Spaces," *Journal of Applied Probability*, 31, 101–113.

Liu, Z., Malave, L., and Sarkar, S. (2004a), "Studies on Silhouette Quality and Gait Recognition," in *Conference on Computer Vision and Pattern Recognition*, June 2004, **?????**: **?????**, pp. 704–711.

——— (2004b), "Simplest Representation Yet for Gait Recognition: Averaged Silhouette," presented at the International Conference on Pattern Recognition, **???**.

Mardia, K., and Patrangenaru, V. (2005), "Directions and Projective Shapes," *The Annals of Statistics*, 33, 1666–1699.

Michor, P. W., and Mumford, D. (2004), "Riemannian Geometries on Spaces of Plane Curves," *Journal of the European Mathematical Society*, 8, 1–48.

Mio, W., and Srivastava, A. (2004), "Elastic-String Models for Representation and Analysis of Planar Shapes," presented at the **?????** IEEE Conference on Computer Vision and Pattern Recognition, **?????**.

Niyogi, S., and Adelson, E. (1994), "Analyzing Gait With Spatiotemporal Surfaces," presented at the **?????** Conference on Computer Vision and Pattern Recognition, **?????**.

Small, C. G. (1996), *The Statistical Theory of Shapes*, New York: Springer.

Spivak (1979) **?????**.

Srivastava, A., Jain, A., Joshi, S., and Kaziska, D. (2006), "Statistical Shape Models Using Elastic-String Representations," in *Proceedings of the Asian Conference on Computer Vision*, **?????**: **?????**, pp. 612–621.

Srivastava, A., Joshi, S., Mio, W., and Liu, X. (2005), "Statistical Shape Analysis: Clustering, Learning, and Testing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 590–602.

Wang, L., Ning, H., Hu, W. and Tan, T. (2002), "Gait Recognition Based on Procrustes Shape Analysis," presented at the International Conference on Image Processing, **?????**.

Younes, L. (1998), "Computable Elastic Distance Between Shapes," *SIAM Journal of Applied Mathematics*, 58, 565–586.

——— (1999), "Optimal Matching Between Shapes via Elastic Deformations," *Journal of Image and Vision Computing*, 17, 381–389.