

CLASSIFICATION OF CONTINUOUS SHAPE DATA

by

Gregory J. Lobser

M.Sc., University of Colorado at Denver, 1998

M.Sc., Cleveland State University, 1978

B.A., Case Western Reserve University, 1971

A thesis submitted to the
University of Colorado at Denver
and Health Sciences Center
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Applied Mathematics

2006

© by Greg Lobser

All rights reserved

This thesis for the Doctor of Philosophy

degree by

Gregory J. Lobser

has been approved

by



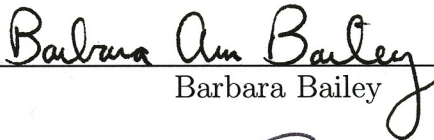
Karen Kafadar



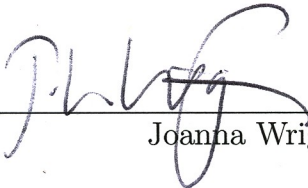
Mark Fitzgerald



Weldon Lodwick



Barbara Bailey



Joanna Wright

MAY 29, 2006

Date

Lobser, Gregory J. (Ph.D., Applied Mathematics)

Classification of continuous shape data

Thesis directed by Professor Karen Kafadar

ABSTRACT

Methods for discrimination when data are continuous curve shapes are investigated through application to three different problems where the data range from strictly functional data to curve data where the data are closed contours. Spectra of upper atmosphere winds are characterized by a shape that follows a power law, but this characteristic shape is contaminated by noise that is introduced by different methods of measurement. The wind spectra are strictly functional data. The spectra are modeled by principal component shapes which are then modified to adaptively remove the noise. The second application considers the vertical acceleration time series captured by an accelerometer worn on a person's chest. The data are not closed contours, but they share many of the same characteristics. Individual steps are segmented and extracted from the time series. Different methods for registration of the steps and their effect on discrimination are investigated. In the third application data are the closed contour shapes of dinosaur tracks. The gains that are realized from using continuous curves over landmarks are investigated.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed Karen Kafadar
Karen Kafadar

DEDICATION

This thesis is dedicated to my extraordinarily patient family who have supported me in this endeavor.

ACKNOWLEDGMENT

I would like to acknowledge the significant role of the following people in supporting this thesis.

- My adviser, Professor Karen Kafadar, for her suggestions and for her unending patience.
- Lockheed Martin Corporation for their support both financially and otherwise.
- Dr. Jeff Richardson, for providing the substantial stride data.
- Prof. Joanna Wright, for her suggestions and for the dinosaur track data.

CONTENTS

Figures	xi
Tables	xiii
<u>Chapter</u>	
1. Introduction	1
1.1 Motivation	1
1.1.1 Upper Atmosphere Winds	1
1.1.2 Stride	2
1.1.3 Tracks	2
2. Literature Review	3
2.1 Principal Components	3
2.2 Splines	3
2.3 Classification	4
3. Methodology	5
3.1 Special Treatments Required by the Different Applications	5
3.1.1 Wind Data: The SCARF Filter	5
3.1.2 Acceleration Time Series: The Stride Data	6
3.1.3 Continuous Closed Contours: Dinosaur Track Data	7
4. Upper Atmosphere Winds	8
4.1 Modeling the Wind Spectra with Component Shapes	10
4.1.1 Distribution of the Coefficients	11

4.2	Smoothing Spectral Estimates by Truncating Component Shapes	11
4.3	Selecting a Cutoff Frequency	14
4.4	Properties of the Scarf Filter	15
5.	Stride Time Series	17
5.1	Background of the Data	18
5.2	Data Characteristics	19
5.3	Data Quality Issues	21
5.3.1	Clipped Data	22
5.3.2	Mislabeled Data and Detection of Outliers	23
5.4	Discrimination Based on Size Parameters	25
5.5	Discrimination Based on Shape Parameters	25
5.5.1	Fisher’s Linear Discrimination Function	26
5.5.2	Discrimination between Walking and Running for All Subjects	28
5.5.3	Calibration to the Individual	30
5.5.4	Registration of Steps	32
5.6	Conclusions and Future Directions for Stride Data	33
6.	Closed Contours: Track Data	37
6.1	The Hands Experiment	37
6.1.1	Preparation of the Hands Data	37
6.1.2	Representation of a Track Contour	39
6.1.3	Calculation of Principal Shapes	39
6.1.4	Comparison of Registration Methods	41
6.1.5	Reconstruction of an Outline	42
6.1.6	Analysis of Variance of Shape Data	44

6.2	Dinosaur Tracks	47
6.2.1	Preparation of the Data	49
6.2.2	Principal Shapes	51
6.2.3	Registration of Shape Data	51
6.2.4	Interpretation of Principal Shapes	54
6.2.5	Approximation with PC Shapes	55
6.2.6	Plotting the PC Coefficients	56
6.2.7	Shape of the Trackway	58
6.2.8	Correlation of Manus and Pes Tracks	59
6.2.9	Internal Structure of Tracks	59
6.2.10	Efficient Methods to Collect Data	60
6.3	Mathematical Issues and Questions	60
6.3.1	The Registration Problem	60
6.3.2	Masking of Information	61
6.3.3	Missing Data	61
7.	Conclusions and Future Work	62
<u>Appendix</u>		
A.	Plots related to the SCARF filter	63
<u>References</u>		69

FIGURES

Figure

4.1	PSD Comparison: Typical Wind	12
4.2	Mean and Component Spectral Shapes of Wind Velocity Spectra	13
5.1	Typical segment of acceleration profile while running	17
5.2	Step Boundaries at the Zero Crossings of a Moving Average Filter	21
5.3	Example of clipped data	22
5.4	Multiple groups based on RMS	23
5.5	Example of a step outlier	25
5.6	Typical walking step compared with a running step	26
5.7	Mean shapes: walking step compared with a running step	27
5.8	Global discriminant function for walking -vs- running	29
5.9	Mean shapes of running steps at different speeds	29
5.10	Running at 7 mph - Mean shapes for three different registration methods	32
6.1	Preparation of Hand Data – Registration and Normalization	38
6.2	Representation of Hand Data – Tangent Angles	40
6.3	Representation of Hand Data - x,y Coordinates	41
6.4	Decomposition of Hand Shapes with Principal Components (Tangent Angles)	42
6.5	Decomposition of Hand Shapes with Principal Components (x,y Co- ordinates)	43

6.6	Reconstruction of hand shapes: shape -vs- Fourier coefficients . . .	44
6.7	Matrix of Hands	46
6.8	Dinosaur effect -vs- scientist effect	47
6.9	Developing a Dinosaur Track Shape from a Contour	49
6.10	Working with a Halftone Image	50
6.11	Dinosaur Trackway	51
6.12	Dinosaur Track	52
6.13	Registration of Dinosaur Tracks	53
6.14	Principal Shapes of the 17 Iguanodon Tracks	54
6.15	Approximating a Track Shape with a Subset of Principal Shapes . .	56
6.16	Plotting the PC Coefficients in a 2D Space	57
6.17	Locating the Central Path of a Trackway	58
6.18	Dinosaur Tracks relative to the Central Path	59
A.1	Component Spectral Shapes of Wind Spectra	63
A.2	Mean Spectral Shapes	63
A.3	Mean PSD's	64
A.4	Mean Transfer Function	64
A.5	PSD Comparison: Turbulent Wind	65
A.6	Smoothed Wind Profiles: Turbulent Wind	66
A.7	Smoothed Wind Profiles: Typical Wind	67
A.8	Impulse Response Function: Turbulent Wind	68
A.9	Impulse Response Function: Typical Wind	68

TABLES

Table

5.1	Discriminant success rates for modeling all subjects	30
5.2	Discriminant success rates for modeling individually for each subject	31
5.3	Discriminant success rates for different registration methods	33
6.1	Analysis of variance table for a single response point	45
6.2	Multivariate analysis of variance table with one principal shape . .	48
6.3	Multivariate analysis of variance table with two principal shapes . .	48

1. Introduction

A common theme of this thesis is the analysis of shapes which are fundamentally continuous curves, but whose representation of the curves is necessarily discrete. A progression of shape data will be considered through the study of three applications that differ in the types of curves that are acquired. The progression begins with shapes that are single-valued functional data where the x-axis is clearly defined (e.g. time), and runs to closed contours where the shapes must be aligned and registered.

Shape analysis has been associated with the study of biological forms. A set of points representing a biological form is typically understood as landmarks that have a corresponding biological meaning. This thesis compares the information content between the discrete landmark points and the continuous contour curves that connect them through a consideration of classification performance, but it will concentrate on the latter.

1.1 Motivation

1.1.1 Upper Atmosphere Winds

The first application is related to wind velocities as a function of altitude. The shapes are of the spectra of the wind velocities. These represent the first step, in that they are single-valued, highly correlated functional data. The data come in the form of x,y vectors where the x-axis is fixed.

1.1.2 Stride

The second application involves acceleration time series collected from subjects who are walking, running, or engaging in other activities. The shapes that will be considered are the acceleration profile corresponding to a single step. As these data are acquired in the form of time series evenly spaced in time, individual steps occur over different durations. It is then required of the step data that they be resampled so that they can be analyzed in a “step” domain. With time series data that arrives in a discrete form, to be resampled, the data must be interpolated, and they must be aligned. An additional decision, required for a proper analysis, is how to identify the dividing point where one step ends and the next step begins. Once this decision is made, the resulting resampled data can be treated similarly to the purely functional data considered above.

1.1.3 Tracks

The third application is actually the initial motivation for this research. The shape data here are the continuous closed curves that represent dinosaur tracks. A track actually corresponds to one member of a family of closed curves. Once a single curve is acquired for a track, contours for a family of tracks must be rotated and resized. To obtain a discrete representation of a contour, decisions must be made about the initial point (corresponding to the final point) in the vector and also the spacing of the discrete points around the contour.

2. Literature Review

2.1 Principal Components

Principal components analysis (PCA) has been around for at least a century. It is generally accepted that the earliest description of the technique was given by Karl Pearson in 1901, and the method is a fundamental tool of statistics today. An excellent current summary of the topic can be found in the monograph by Jolliffe ([18]).

Continuous curve data are infinitely dimensional. In practice, continuous curve data generally are measured and represented at a finite number of discrete points and hence are of finite dimension. Nevertheless, this dimension is generally very high. The problems that follow from this high dimensionality of the data are discussed in the paper by Leurgans et al. ([24]). For problems dealing with continuous curve data, the dimensionality usually must be reduced, and PCA is the primary tool that is used in this thesis to accomplish this.

A principal components analysis is a multivariate technique that works with an estimate of a mean vector and an estimate of a covariance matrix. Both of these estimates are sensitive to outliers, and much recent research with PCA is concerned with robust estimates of principal components. Discussion of robust techniques with PCA and with multivariate techniques in general can be found in Rousseeuw ([36], [37], and [39]).

2.2 Splines

The literature on smoothing and interpolating splines is very extensive. Splines find broad application in this thesis with both the stride time series data and

the closed contour track data. For example, with the stride data, steps need to be extracted from the time series data and aligned. Splines provide the means for resampling the time series to different time intervals.

Silverman ([40]) provides an excellent discussion of the frequency domain characteristics of the smoothing spline. Splines are shown to be equivalent to the Butterworth filter, and the connection is made between the smoothing parameter of the smoothing spline and the frequency cutoff of the Butterworth filter. In this thesis, we rely extensively on the methods for fitting splines, once the data have been sampled, aligned, and registered.

2.3 Classification

The literature on classification is vast. A good discussion of classification error can be found in McLachlan ([27]). An historical perspective of robust techniques of discrimination can be found in papers by Lachenbruch ([22], [23], and [21]).

3. Methodology

Continuous curves are infinite dimensional spaces. Representations of these continuous curves with discrete points are no longer infinite, but they generally are much more numerous than warranted by the information content in the data. The dimension of a vector space is related to the amount of information contained in the data. Although the number of points used to represent continuous curves may be large, resulting in a high dimensional vector space, the data are usually very highly correlated, and hence the dimension of the space does not reflect the amount of information.

On the other hand, maintaining a large number of points provides high fidelity to the shapes that are being modeled. Choices for representation of curves with a finite set of basic functions include wavelets and Fourier series. Principal component analysis (PCA) provides a means for maintaining the high fidelity to the shape data and for reducing the dimension of the associated vector spaces to be more consistent with the informational content of the data.

3.1 Special Treatments Required by the Different Applications

While the principal component method is applied throughout, each application will require some sort of special treatment in making the transition from the continuous to a discrete form where the shapes can be represented in a vector space.

3.1.1 Wind Data: The SCARF Filter

Spectra of upper atmosphere wind velocities have a known family of shapes, characterized by a straight line on a log-log plot, i.e. they follow a power law.

True wind velocities are not precisely known, however. They are measured by tracking the movement of weather balloons by radar, and these data are further processed before an altitude velocity profile is produced. When these spectral shapes are calculated from measured data, it is seen that the shapes are contaminated by noise associated with the different types of balloons that are used for measuring the wind velocities.

PCA supports a cluster analysis of the wind data to identify distinct groups within the data. The analysis revealed the important finding that some winds are mislabeled. Some groups are removed before proceeding with modeling. The SCARF filter is introduced here as a method for removing the noise associated with the balloon type.

3.1.2 Acceleration Time Series: The Stride Data

Although the underlying time series here is again continuous, people's strides, when measured are necessarily digitized, usually at evenly spaced times unless, for example, the recording instrument fails for a period of time. The data elements of interest in this study are the size and shape parameters of the acceleration profile associated with a single step. The shapes can not be represented in a vector space with data taken directly from the raw time series. Steps are not all of the same duration, and therefore, segments of the time series corresponding to individual steps are necessarily of different length.

The time series must be modified before the data can be represented in a vector space. Interpolating cubic splines are utilized for modeling the raw time series with a continuous function which is then redigitized to transform the time series representation of a step to a "step domain" representation. Properties of

interpolating splines are reviewed to show how the informational content of the time series may be distorted in this process.

Historically, step data have been modeled by the overall duration and average acceleration during a step. These correspond to size parameters, and we will be interested in the additional information contained in the shape of a step profile beyond the information conveyed by the size parameters.

Resampling the cubic spline representation of the time series and extracting a segment corresponding to a single step requires some method for choosing the boundary when one step ends and the next step begins. The choice of boundary point is closely associated with the concept of a landmark point. Landmark points are important in the study of shapes of biological forms, and they will be discussed more extensively in association with the track data. Study of the stride data, however, provides an introduction to the relationship between discrete landmark points and the continuous curves that connect between them.

3.1.3 Continuous Closed Contours: Dinosaur Track Data

With closed contours, even more decisions are required before the data can be represented in a vector space. Each decision is associated with some error in the chosen outcome. A closed continuous contour is only a sample from a family of contours that could be selected to represent the track.

Once a contour has been selected, representation in a vector space requires the selection of not only a starting (and ending) point, but also the distribution of points around the contour.

4. Upper Atmosphere Winds

The **S**pectral **C**omponent **A**daptive **R**eshaping **F**ilter (Scarf) models a database of spectra of upper atmosphere winds with principal component shapes. The shape of these spectra is expected to follow a α/f power law, but they generally deviate from this shape with noise introduced by the measurement system. Wind velocity is typically measured by tracking weather balloons with radar, and noise can result from sporadic movements of the balloons and also from errors in the radar tracking. The spectra are then reshaped by reconstructing them with a modified set of the principal component shapes.

A typical wind velocity profile consists of a u-component wind velocity (east-west) and a v-component (north-south) sampled every 100 feet. Depending on the maximum altitude eventually reached by the balloon, the velocity profiles stop at different altitudes. Spectra are estimated from a subset of the entire altitude range, high enough so that they are representative of the upper altitudes that are more turbulent, and low enough so that the number of profiles reaching the required altitude is adequate.

For an altitude band of length T , the power spectrum is estimated by (see [1]).

$$G_x(f) = 2 \lim_{T \rightarrow \infty} \frac{1}{T} \mathbf{E}[|\mathbf{X}(f, T)|^2] \quad (4.1)$$

where $\mathbf{X}(f, T)$ is the finite Fourier transform of the velocity profile $x(t)$.

$$\mathbf{X}(f, T) = \int_0^T x(a) e^{-j2\pi f t} da \quad (4.2)$$

where $x(a)$ is equivalent to a time series, but is treated as a function of the altitude parameter a , measured in feet. For the work done here, each spectrum is calculated for the u- and v-velocity components over the altitude range of 25000 to 50500 feet. The Fourier transform ($X(f, T)$) can then be calculated with the fast Fourier transform (FFT). The power spectrum (Eq 4.1) is estimated as the average of Fourier transforms that are calculated for subsets of the altitude band. The subsets, consisting of 64 points, overlap by 50%, and to reduce variance in the estimate of the power spectrum, each subset is multiplied by a Hanning window.

To model the spectra optimally, principal components of the spectra vectors are derived from the database of spectra. Each spectrum is necessarily estimated from a limited amount of data, and therefore the spectral estimate is rough. A convenient method for smoothing the spectra is accomplished by truncating the spectral components. For a given wind profile, a vector coefficients for the truncated component shapes is calculated. The resulting smooth approximation of the spectral shape doesn't necessarily follow a power law. The Scarf filter achieves a spectral shape that does follow a power law by keeping the same coefficients, but applying them to a modified set of component shapes. The modified component shapes are derived from the original components. Because the noise is more dominant in the high frequency portion of the spectrum each component is modified by replacing the high frequency end of the component with an extrapolation from the low frequency end that does follow a power law.

One advantage of this approach is that the original components and the modified components can be stored. When the filter is applied to a new wind, a

new vector of coefficients is calculated from the stored set of original components, and it is then applied to the stored set of modified components.

4.1 Modeling the Wind Spectra with Component Shapes

Principal component shapes are calculated as the eigenvectors of the covariance matrix of the wind spectra. $G_x(f)$ denotes the power spectral density (PSD), or autospectrum, of a velocity profile x as a function of frequency (f). In the remainder of this section we will drop the f , and will work with the log of $G_x(f)$.

Let

$$g_x = \log(G_x) \quad (4.3)$$

where the subscript x denotes a wind velocity profile from the winds database. A design matrix is then constructed that contains the centered database of spectra.

$$X = \begin{bmatrix} (g_{x_1} - \bar{g}_x)' \\ (g_{x_2} - \bar{g}_x)' \\ \dots \\ (g_{x_n} - \bar{g}_x)' \end{bmatrix} \quad (4.4)$$

The principal component shapes of the spectra are taken as the columns of V from the singular value decomposition (SVD) of X :

$$X = UWV' \quad (4.5)$$

The matrix of component shapes is orthogonal, so that $V'V = I$. The product $\Gamma = UW$ is the matrix of principal component shape coefficients for the spectral database.

The singular value decomposition can be partitioned as

$$X = X_s + X_r = \begin{bmatrix} U_s | U_r \end{bmatrix} \begin{bmatrix} W_s & \mathbf{0} \\ \mathbf{0} & W_r \end{bmatrix} \begin{bmatrix} V_s | V_r \end{bmatrix}' \quad (4.6)$$

where the subscripts s and r denote *smooth* and *rough*.

4.1.1 Distribution of the Coefficients

If we drop the rough part of the model,

$$X_s = U_s W_s V_s' + \varepsilon_r \quad (4.7)$$

where ε_r are the residual differences between the raw spectrum and the smooth approximation. Let's assume that the residuals are iid $\varepsilon \sim N(0, \Sigma)$. The coefficients $\Gamma_s = U_s W_s$ are then

$$U_s W_s = (X_s - \varepsilon_r) V = X_s V - \varepsilon_r V \quad (4.8)$$

The expectation is $E(U_s W_s) = E(X_s V)$. The residuals are a linear combination of the iid normal residuals ε_r . The coefficients are therefore multivariate normal.

4.2 Smoothing Spectral Estimates by Truncating Component Shapes

A smooth approximation of X (the database of wind spectra) is calculated by using only a subset of the principal components of the wind spectra.

$$X \approx X_s = U_s W_s V_s' = \Gamma_s V_s' \quad (4.9)$$

Or, the smooth estimate of a single spectrum g_x is constructed by

$$\hat{g}_x = \gamma_s V_s' \quad (4.10)$$

An example of this reconstruction is seen in Figure 4.1(b), comparing the reconstructed “model” with “raw”.

The goal is to replace this smooth spectral estimate with a spectral shape that matches the estimate at the low frequency end of the spectrum but replaces

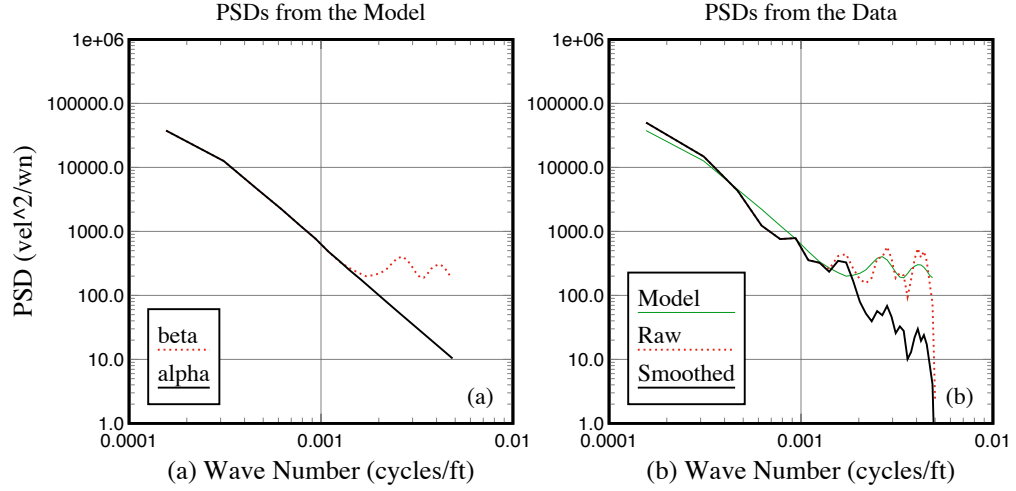


Figure 4.1: PSD Comparison: Typical Wind

the high frequency end of the spectrum with a spectral shape that follows the power law.

This is summarized in Figure 4.2. The coefficients that are calculated for the raw spectrum can then be applied to modified shapes. Let $B = V_s$ denote the component shapes *before* modification, and let $A = V_s^*$ denote the component shapes *after* modification. The coefficients that model the spectrum g_{x_i} are estimated by

$$\hat{\gamma}_{x_i} = B'(g_{x_i} - \bar{g}_x) \quad (4.11)$$

The smooth estimate of the spectrum, denoted here by \hat{g}_{x_i} is constructed by applying the coefficients to the principal component shapes.

$$\hat{g}_{x_i} = B\hat{\gamma}_{x_i} + \bar{g}_x \quad (4.12)$$

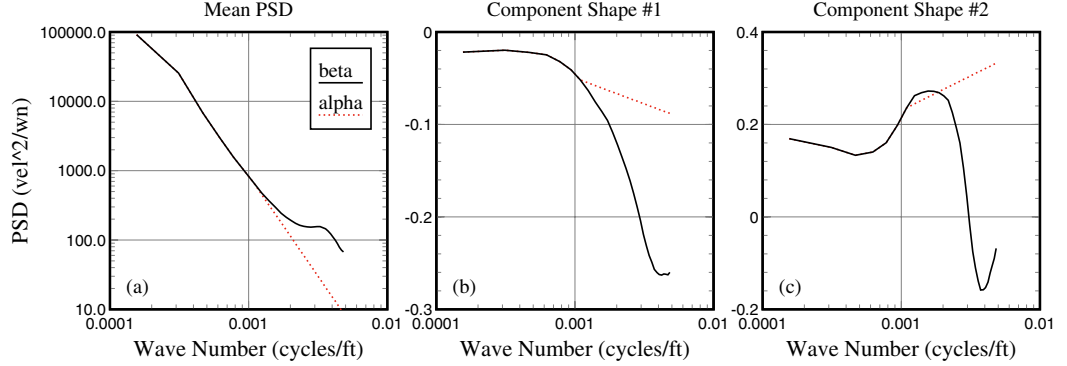


Figure 4.2: Mean and Component Spectral Shapes of Wind Velocity Spectra

The modified spectral estimate that follows the power law is denoted here by \hat{g}^* . It is calculated by applying the same coefficients to the modified principal component shapes.

$$\hat{g}_{x_i}^* = A\hat{\gamma}_{x_i} + \bar{g}_x^* \quad (4.13)$$

The transfer function ($H(f)$) is used to smooth the wind profile is the ratio of the two PSD's:

$$\hat{H}_{x_i} = \sqrt{\frac{\hat{g}_{x_i}^*}{\hat{g}_{x_i}}} \quad (4.14)$$

$$= \exp [((A - B)\hat{\gamma}_{x_i} + (\bar{g}_x^* - \bar{g}_x)) / 2] \quad (4.15)$$

Any values in the vector H_{x_i} greater than one are set to one so that no values are amplified by application of the filter. To summarize, the algorithm to construct the wind smoothing transfer function is implemented by equations (4.11) and (4.14). The component shape matrices for the original spectra (B), the component shape matrices for the modified spectra (A), and the mean spectrum vectors \bar{g}_x and \bar{g}_x^* are calculated up front from the winds database. They are then stored for subsequent modeling and filtering of new wind profiles.

4.3 Selecting a Cutoff Frequency

An important parameter in the construction of the Scarf model is the cutoff frequency, denoted by, say, f_c . The principal component shapes are estimated using the entire database, but frequencies less than f_c remain unadjusted. The portion of the spectral components at frequencies greater than f_c is reconstructed by extrapolating, following a power law, from the lower frequencies.

The frequency cutoff is a constant that is used for the construction of the modified principal component shapes. Because it is the same set of stored principal component shapes that are applied in the filtering of any new wind velocity profile, the frequency cutoff is a constant parameter of the filter, and its value is selected prior to construction of the modified component shapes.

Selection of a filter cutoff is a bias-variance tradeoff. Noise affects all frequencies of the wind velocity spectra. But because the true spectra are expected to follow a power law and are necessarily of lower level at higher frequencies, and also because of the source of the noise (weather balloon oscillations) in the wind velocity data, the noise is much more predominant at higher frequencies. If too low of a frequency cutoff is selected, then construction of the modified component shapes is based on less data (higher variance). If too high of a frequency is selected, then too much noise remains in the wind velocity profiles (higher bias).

An appropriate value to use for the frequency cutoff was estimated by an evaluation of the database of wind spectra. Initially, a value of $f_c = 0.001$ was selected where the distribution of spectral shapes appeared to begin to

diverge from the expected power law shape. Subsequent studies comparing different noise characteristics resulting from different types of balloons appeared to support this originally selected value.

4.4 Properties of the Scarf Filter

A complex transfer function is constructed by assuming zero phase, and the impulse response function is calculated as the inverse Fourier transform. Although the filter is modeled from the altitudes at 25000 to 50500 ft, the filter is then applied to the entire wind profile.

Consider the spectrum of the residual differences between the raw and smoothed profiles. The spectrum of the residuals can be calculated directly, knowing the transfer function H_s and the spectrum g_β of the raw velocity profile:

$$g_r = (1 - H_s^2)g_x \quad (4.16)$$

Because H_s is constrained to 1.0 for frequencies $f < f_c$ at lower frequencies below a cutoff frequency, the transfer function has no effect at the zero frequency, and therefore the residual differences will have zero mean.

H_s is also constrained not to exceed 1.0. This leaves open the possibility that H_s is a constant 1.0 at all frequencies and therefore has no effect on the velocity profile, and the residuals are all zero. Because of this constraint, the frequency cutoff (f_c) should be considered as a minimum frequency cutoff. The effective frequency cutoff for a specific wind is determined as that frequency where the modified spectrum ($\hat{g}_{x_i}^*$) begins to fall below the unmodified reconstructed spectrum (\hat{g}_{x_i}).

Consider that the raw velocity profile is the sum of the true wind velocity plus residuals. Because the spectrum of each of these components is smooth, the velocity profile can therefore be characterized as wide band stationary (WSS). The spectrum of the sum is also WSS. The transfer function is constructed as the difference between a smooth approximation to the measured velocity profile and the power law modification of the component shapes. The spectrum of the residual differences can then be estimated directly from the transfer function and is also WSS, although only at the higher frequencies. The residuals are therefore iid

$$\varepsilon_r \sim N(0, \Sigma) \tag{4.17}$$

The Scarf transfer function is calculated from spectra that are estimated for a limited altitude band of data. Outside of this altitude band, the spectrum of the velocity profile after passing through the Scarf filter will be biased to the extent that the spectrum of the altitude band differs from the spectrum for the altitude band on which the transfer function was based.

5. Stride Time Series

The second application of this thesis moves beyond the strictly functional data of the first application. The stride data are acceleration time series measured for a variety of individuals during a variety of activities. An example of a short segment is illustrated in Figure 5.1. To consider shapes of individual step profiles, these data must be segmented into time periods that correspond to a single step. Once this is accomplished, step shapes share characteristics with closed contour shapes to be considered in the third application. After segmentation, the steps can be discriminated first by the size parameters of step duration, and second by overall acceleration (variance). Traditionally, studies within the fitness community of acceleration profiles are limited to these parameters. The novel work in this thesis considers the additional information provided by the shape of individual step profiles.

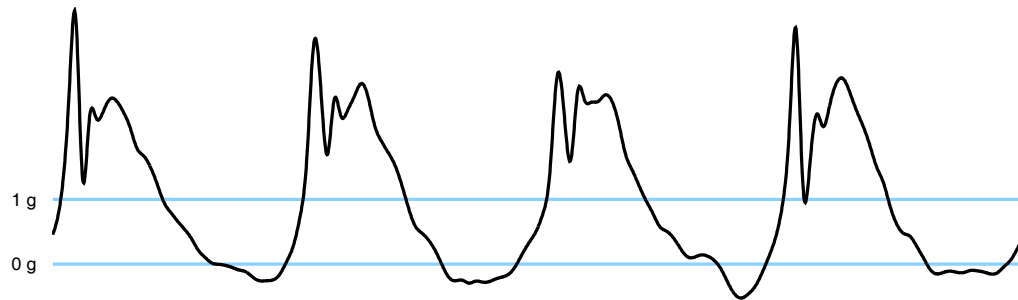


Figure 5.1: Typical segment of acceleration profile while running

The shapes of individual step profiles can be characterized by their continuous outline and, while perhaps not as clearly as with closed contour shapes, they can also be characterized by landmarks. Landmarks are, according to [3] “loci that have names . . . as well as Cartesian coordinates.” They can be identified in terms of their biological interpretation, and not just by a mathematical properties. Study of step shapes provides the opportunity for investigating the merits of characterization of shapes with landmarks compared to characterization with the continuous profile outline.

5.1 Background of the Data

The stride data were collected as part of a project fund by an SBIR (small business innovation research) grant from the Department of Health and Human Services. The eventual objective of this research is the development of a device for monitoring energy expenditure with the use of the dual measurements of heart rate and the vertical acceleration measured by an accelerometer worn on a subject’s chest.

From [35],

Several reviews have suggested that the recent trends in obesity are more strongly related to decreases in energy expenditure (EE) than to increases in energy intake. . . . The link between physical activity and obesity in individuals is the energy balance equation, the relationship between energy intake and energy expenditure. In order to promote physical activity and understand its role in the energy balance equation, the need for improved physical activity instrumentation is broadly recognized as an important topic of research. . . . A next-generation, low-cost monitoring capability

designed specifically for use in large-scale epidemiological studies is needed to help us better understand the EE side of the energy balance equation. ... In sum, to be suitable for broad-based use, an EE monitoring device must be accurate, reliable, sensor-based, practical and affordable. In spite of considerable research efforts and major advances in technology, no current physical activity methodology adequately meets this combination of criteria.

While the research direction of the grant study requires the development of a model that incorporates the simultaneous use of a heart rate monitor and the measurement of vertical acceleration, the objective in this thesis focuses on the investigation of the dependence of energy expenditure on the shape of the acceleration profiles associated with individual steps.

5.2 Data Characteristics

Acceleration time series data were collected from three different studies, or “protocols,” at constant time increments of 0.01 seconds. Acceleration was measured with an accelerometer that was worn at a subject’s chest. Energy expenditure is also measured during both protocols by measuring oxygen consumption with a metabolic cart. The metabolic cart measures the carbon dioxide exhaled by the subject, calculates the amount of oxygen consumed, and then calculates the corresponding energy expenditure.

The study involved twenty subjects, but not all subjects were available throughout all protocols. The initial sets of data are measured for a subject walking and running on a treadmill at known treadmill speeds. Each subject

walks on the treadmill at 2, 3, and 4 miles per hour, and runs at 5, 6, 7, and possibly 8 miles per hour.

Additional data to be considered here were collected from the third protocol, which involved some of the same subjects, with additional treadmill speeds at 2.5, 3.5, 5.5, and 6.5 miles per hour. Additional activities also were added in the third protocol. Stride profiles are measured with the subjects walking at a 10% grade, climbing and descending stairs, and riding a bicycle.

Different models will be constructed for estimating energy expenditure depending on whether the subject is running, walking, or neither (e.g. bicycling). Discriminating between running and walking is easily accomplished by considering the overall acceleration level (overall g-rms). But determining different speeds, and subsequently energy, within the categories of walking or running is much more difficult. The shape of the acceleration profiles will then be investigated to improve the discriminatory power of the model.

A typical acceleration profile for a subject when running is shown in Figure 5.1. The overall acceleration is calculated directly from the raw time series by calculating the standard deviation for n acceleration samples with a known mean level of $\mu = 1.0$, where n is chosen so that the running rms values are reasonably smooth. The approach taken here, however, will be different. Because this study investigates the use of the step shapes to estimate energy expenditure, first the time series will first be segmented into individual steps, and then the overall g-rms level will be calculated for each step.

A visual inspection of each step to determine its start and end times is impractical and feasible, so an automatic procedure (algorithm) had to be de-

veloped. An obvious approach may be the times when the acceleration profile crosses some threshold, say 1.0 g. Although not evident in Figure 5.1, it is possible that the profile may, or may not, cross this threshold multiple times for a single step because of high frequency features of the profile. Filtering the time series to remove the high frequency content provides a simple solution with a high enough smoothing parameter (bandwidth). This is illustrated in Figure 5.2.

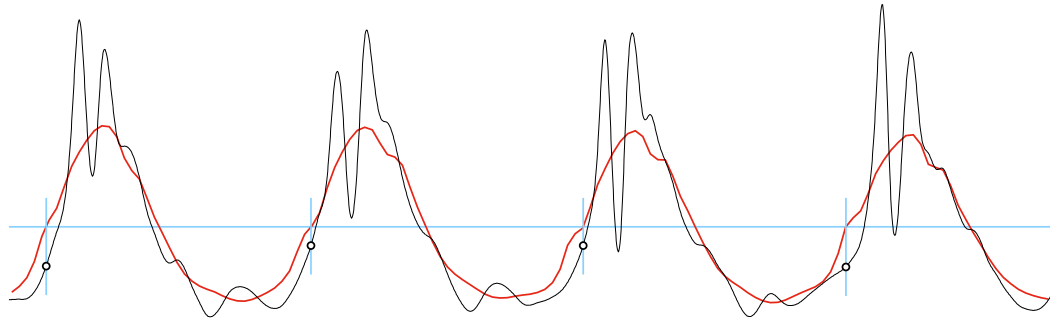


Figure 5.2: Step Boundaries at the Zero Crossings of a Moving Average Filter

5.3 Data Quality Issues

For each individual the data consist of periods of time when the subject is engaged in a single identified activity. We want to use the data to develop a model so that, given the time series, we can predict the corresponding activity. For this prediction to be accurate, the measurement of the time series must be representative. Several issues with the data arose which had to be addressed before this could be done.

In some of the protocol 3 data, the time series appear to have been clipped. Depending on the type of accelerometer used, and how the acquisition system

processes the data, it reached the limit of its capability; i.e. the data are “Win-sorized” to an upper limit, sometimes denoted “right-censored” - all that is known is that the actual value exceeded an upper limit, which could differ from person to person and from step to step. The indicated acceleration is then truncated from the actual acceleration level to the limit of the measurement. Also, in many time segments for a given activity some outlier steps were identified. One cause for these outliers could be the placement of the accelerometer worn on a subject’s chest. Although the accelerometer is attached very snugly, there is inevitably some variability in how tightly it is attached. Some features of the time series can be associated with accelerometers that appear to be moving with a higher level of acceleration than would be possible for the subject.

5.3.1 Clipped Data

The accelerometers can measure only over some preset range. If the actual acceleration goes outside of this range, the values are “clipped,” or set to the extremes of the range. Figure 5.3 provides an example of clipped data.

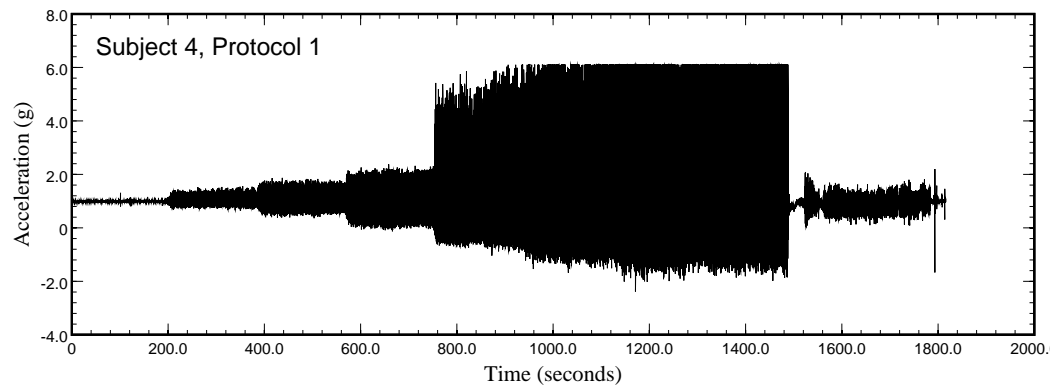


Figure 5.3: Example of clipped data

5.3.2 Misabeled Data and Detection of Outliers

The identity of the activities are necessarily coarse, and typically consist of the start and end time when a subject is, say, walking on a treadmill at a known speed. Although a treadmill is at a constant speed, not all steps are equal. Some steps may actually be faster or slower than what the treadmill might indicate. In developing a model, outlier steps are best left out of the family of steps for purposes of modeling a given activity.

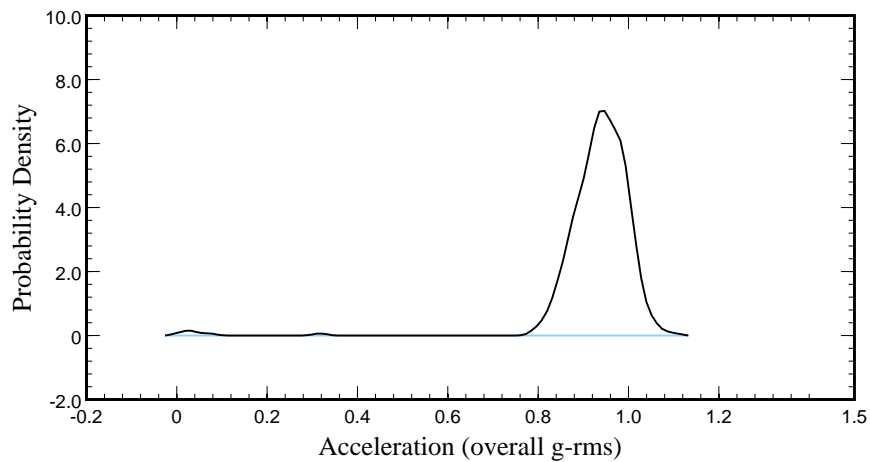


Figure 5.4: Multiple groups based on RMS

One method for detecting an outlier step estimates the probability that the step could have occurred by chance in a random sample of n steps. This probability is estimated here based on a robust estimate of the ratio of the peak magnitude to the standard deviation, using a robust estimate of the standard deviation calculated as the pseudo-sigma. (see [16])

As implemented here, the pseudo-sigma is calculated by using corresponding quantiles of the Normal distribution. The quantile corresponding to $\sigma = -1$ is $q_{-1} = F_N(-1) = .1587$, and the quantile corresponding to $\sigma = 1$ is $q_1 = F_N(1) = .8413$. Let F_e be the empirical distribution based on the sorted data. Then the pseudo-sigma can be estimated as

$$s^* = [(F_e)^{-1}(q_1) - (F_e)^{-1}(q_{-1})] / 2 \quad (5.1)$$

The peak magnitude standardized to the pseudo-sigma is

$$x_{(n)}^* = \frac{x_{(n)} - \bar{x}}{s^*}. \quad (5.2)$$

A critical value can be estimated for this statistic by assuming that the data follow a Normal distribution. The cumulative distribution of the maximum order statistic is then F_N^n . A critical value is then calculated from

$$F_N^n(c_r) = 1 - \alpha \quad (5.3)$$

where α is a small probability of exceedance, say $\alpha = .01$. The critical value is then calculated as

$$c_r = F_N^{-1} [(1 - \alpha)^{1/n}] \quad (5.4)$$

A sample of $n = 500$ and $\alpha = .01$, for example, results in a critical value of $c_r = 4.1$. Values of $x_{(n)}^*$ that exceed this would then be rejected. Similarly, values of the standardized minimal order statistic $x_{(1)}^* < -4.1$ would also be rejected. Figure 5.5 shows an example of a step, occurring in the middle of a series of ordinary steps, that has been identified as an outlier.

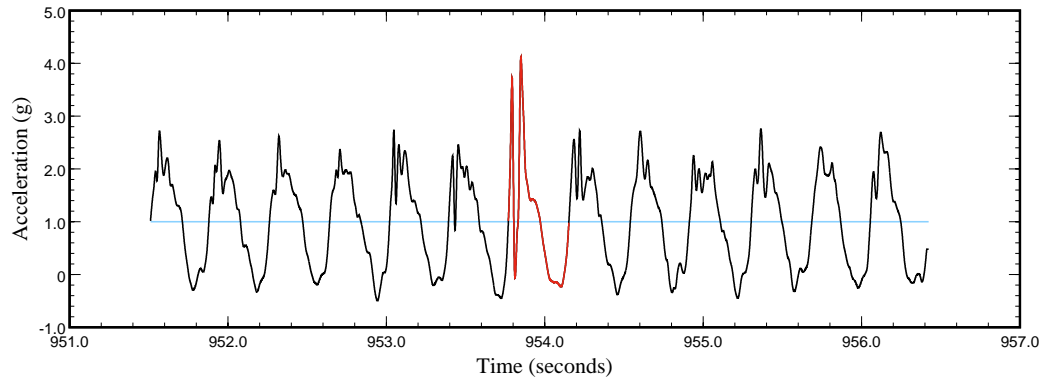


Figure 5.5: Example of a step outlier

5.4 Discrimination Based on Size Parameters

Figure 5.6 shows a typical running step compared with a walking step. Not surprisingly, the running step occurs over a shorter period of time, and the overall level of the acceleration is larger. Both of these parameters characterize the scale or size of the steps, and discrimination based on scale will be examined in this section.

5.5 Discrimination Based on Shape Parameters

Shape is what remains after location and scale parameters have been taken into account. The stride data have been segmented into time periods corresponding to a single step. If we rescale the resulting time interval to $[0, 1]$, then we have accounted for the scale of the step duration. If we normalize the continuous acceleration profile to have a standard deviation of 1.0, then the scale parameter

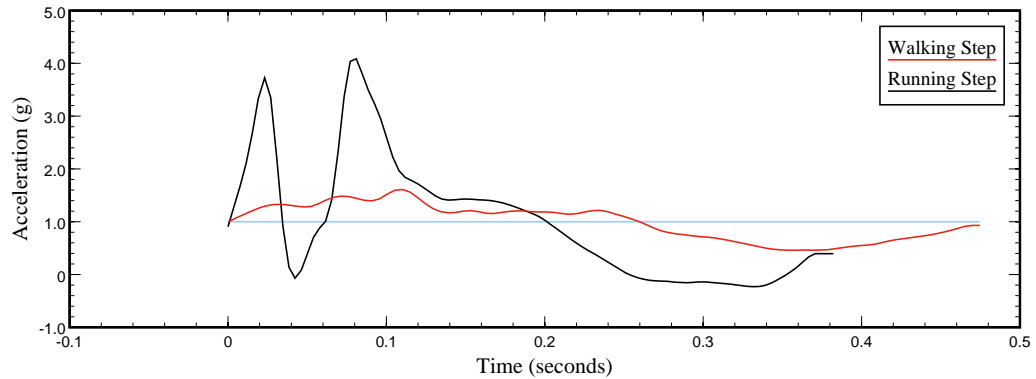


Figure 5.6: Typical walking step compared with a running step

of overall acceleration (g-rms) has been discounted.

Figure 5.7 summarizes the mean step shapes for running and walking after the contributions of the size parameters have been normalized out. There is some difference between the two curves, but they appear to be very similar overall.

5.5.1 Fisher's Linear Discrimination Function

This section summarizes Fisher's linear discrimination function (LDF) as it can be applied to the stride shapes. The LDF works quite successfully with the stride data. The development of the LDF as it is implemented here will work with components of shape. If the LDF is applied directly to the time series stride data, the resulting discriminant function will be overdetermined. It is necessary to first smooth the stride data, and this is accomplished by using a reduced set of principal shapes, or principal components.

Each time series step can be modeled by a set of coefficients applied to a small set of principal shapes. Let y be the vector of coefficients of the principal

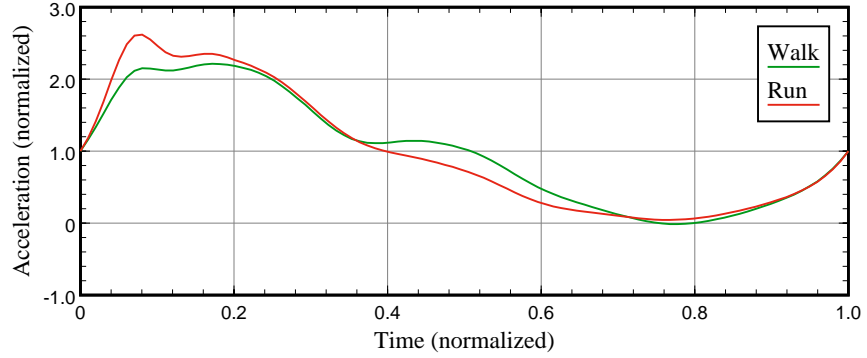


Figure 5.7: Mean shapes: walking step compared with a running step

shapes. Fisher's discriminant function is then calculated by

$$D_{pc} = (\bar{y}_1 - \bar{y}_2)' S_{\text{pooled}}^{-1} \quad (5.5)$$

where S_{pooled} is the combined, or pooled, estimate of the covariance matrix Σ :

$$S_{\text{pooled}} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{(n_1 + n_2 - 2)} \quad (5.6)$$

This optimally separates the two populations. The point of division between the two populations is calculated as

$$\hat{m} = \frac{1}{2} D_{pc} (\bar{y}_1 + \bar{y}_2) \quad (5.7)$$

The decision of whether to assign a vector y to group 1 or group 2 is defined by

$$\text{If } D_{pc}y - \hat{m} > 0 \text{ then } y \in \text{Group 1} \quad (5.8)$$

The discriminant function that is applied directly to the raw data is obtained by applying the principal shapes (V):

$$D_{pc}V'x - [D_{pc}V'\bar{x} + \hat{m}] > 0 \quad (5.9)$$

and the decision for the raw step time series is then stated as

$$\text{If } D'x - C > 0 \text{ then } x \in \text{Group 1} \quad (5.10)$$

where

$$D = (\bar{y}_1 - \bar{y}_2)' S_{\text{pooled}}^{-1} V' \quad (5.11)$$

and

$$C = D_{pc} V' \bar{x} + \frac{1}{2} D_{pc} (\bar{y}_1 + \bar{y}_2) \quad (5.12)$$

5.5.2 Discrimination between Walking and Running for All Subjects

Fisher's linear discriminant function can be applied to the principal components, but it is extended here to apply directly to the step shapes. Robustness is a consideration, but in this thesis outlier steps are first identified and then removed from the database before a discrimination model is developed. Robustness will be an important consideration later in the implementation of the model, but that is beyond the scope of this thesis and will be part of future work.

Figure 5.8 shows the discriminant function D that discriminates between walking at 4 mph and running at 5 mph for the all subjects in the study. As with discrimination based on the size parameters, the shapes still carry a significant amount of information to discriminate between walking and running steps.

Performance of the discriminant function is evaluated by doing a cross-validation study. Although the discriminants in Figure 5.8 are based on data from all subjects, the cross-validation study is performed by constructing the discriminant using data for all subjects except for one. The model is then used to discriminate on the steps for the subject that was left out, and the number

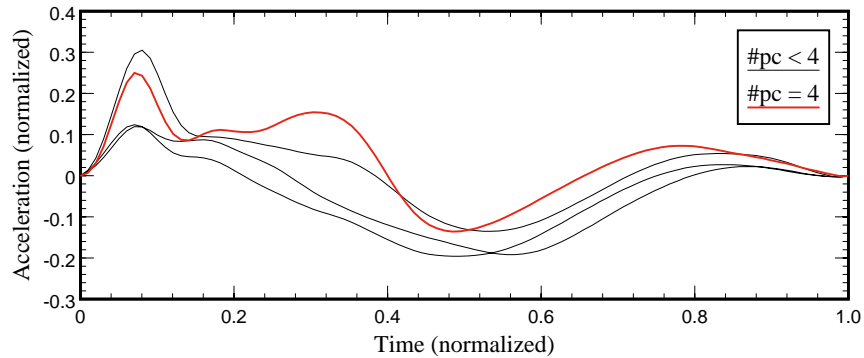


Figure 5.8: Global discriminant function for walking -vs- running

of correct and incorrect evaluations is recorded. This process is repeated for all subjects. The results of the cross-validation study are shown in Table 5.1.

The reader is reminded here that a success rate of 50% for discriminating between two groups is no better than flipping a coin. Shape data provide no additional information for discriminating between different speeds while running. Discriminating between different speeds while walking is not much better.

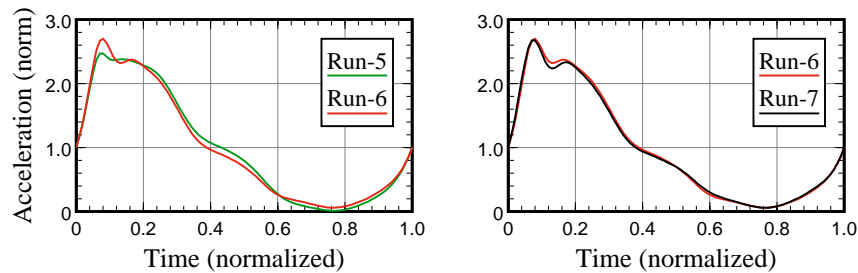


Figure 5.9: Mean shapes of running steps at different speeds

Figure 5.9 illustrates the difficulty of discriminating between different speeds while running. When averaged over all subjects in the study, the mean shapes,

	Discrimination speeds (mph)				
# pc's	2-3	3-4	4-5	5-6	6-7
1	58.6%	65.7%	85.0%	58.0%	53.6%
2	66.6%	63.5%	82.4%	59.5%	51.4%
3	66.2%	62.7%	85.3%	58.4%	50.7%
4	63.3%	61.5%	90.0%	58.3%	50.3%
5	60.9%	65.0%	90.8%	58.1%	50.5%
6	63.3%	65.6%	91.1%	56.1%	50.4%

Table 5.1: Discriminant success rates for modeling all subjects

especially between 6 and 7 mph appear to be nearly identical.

5.5.3 Calibration to the Individual

One of the ground rules for the development of an energy monitoring device is that it be calibrated for each individual. How such a calibration would be best implemented will not be investigated here. But the feasibility of calibrating to the individual can be evaluated by constructing a discrimination model for an individual using only that individual's data.

Testing the models is accomplished a little differently. Instead of performing a cross-validation study, the models are constructed using only half of the data. The other half of the data are then used for testing the performance of the model. Table 5.2 shows, not surprisingly, a much higher rate of success. But this higher success rate comes with the added cost of calibration for each individual.

	Discrimination speeds (mph)				
Subject	2-3	3-4	4-5	5-6	6-7
2	95.3%	99.1%	100.0%	95.8%	79.8%
5	97.2%	97.9%	100.0%	97.6%	90.7%
6	85.4%	92.4%	100.0%	92.7%	86.1%
7	66.9%	90.9%	100.0%	97.9%	70.0%
8	99.2%	95.8%	98.6%	93.7%	89.2%
9	87.7%	84.0%	99.8%	94.5%	86.7%
10	94.6%	96.9%	100.0%	99.3%	94.5%
11	93.0%	97.2%	100.0%	87.2%	70.5%
13	87.9%	92.2%	99.8%	93.0%	89.3%
14	80.5%	98.9%	100.0%	93.8%	95.6%
15	81.0%	98.8%	100.0%	87.4%	78.0%
Overall	86.7%	94.9%	99.8%	94.0%	84.6%

Table 5.2: Discriminant success rates for modeling individually for each subject

5.5.4 Registration of Steps

The step shapes considered to this point are constructed from segments of the time series that are bounded by those times when the measured acceleration levels ascend across the 1-g acceleration level.

It was considered that different methods of registration may be more or less successful in capturing features of the step shapes that carry information about the activity. Two additional methods were also considered, and the three methods are illustrated in Figure 5.10. This shows the mean step shape for running at 7 mph averaged over all subjects.

Registration “a” is actually the simplest, and it is based on the times when the filtered profile ascends across the 1-g acceleration level. Registration “c” is based on the times when the first major acceleration peak occurs after the times identified with the registration “a” method.

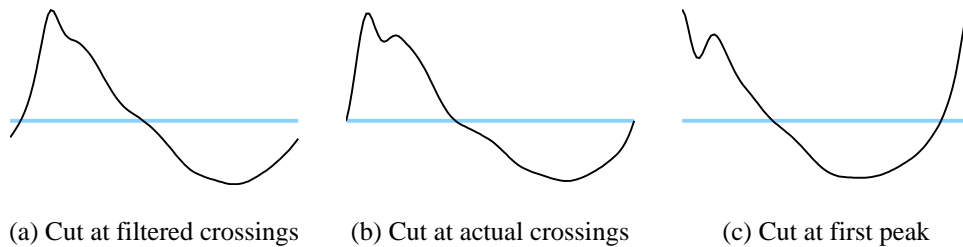


Figure 5.10: Running at 7 mph - Mean shapes for three different registration methods

The performance of the discriminants developed with the three different registration methods are summarized in Table 5.3. Registration methods (a)

	Discrimination speeds (mph)				
Registration	2-3	3-4	4-5	5-6	6-7
a	87.3%	96.4%	99.8%	93.2%	82.1%
b	86.7%	94.9%	99.8%	94.0%	84.6%
c	80.7%	89.9%	98.7%	89.7%	79.2%

Table 5.3: Discriminant success rates for different registration methods

and (b) are similar in performance. Method (c) has significantly poorer performance. The time of the first major peak in the acceleration profile is essentially a landmark that carries information about the step. The reduced performance of this method suggests that this information is being discarded in the registration process, lining up each step at the first peak.

5.6 Conclusions and Future Directions for Stride Data

An important issue to consider when analyzing stride data is the advantage of the continuous time series data over an approach based only on selected “landmarks” of the trace. Two types of landmarks can be considered. The first type of landmark corresponds to a point on the continuous curve with some clear biological meaning. For example, in Figure 1, the first of the positive peaks occurs when the foot strikes the ground with maximum impact. The second type of landmark is the type that is defined mathematically. For example, in the first analysis of the stride data, summarized in Figure 2, landmarks will be taken where an 11-point moving average crosses over the 1-g line from below.

Landmark points may not be needed at all. In attempting to discriminate between running steps and walking steps we could, for example, simply select random segments of the time series of equal duration. These segments could, however, have different numbers of steps, and the steps will not necessarily correspond in time of occurrence from one segment to another. But we could estimate coefficients of a Fourier transform, and we could base a classification strategy on these coefficients.

By introducing a single landmark point at corresponding positions for each step, we can align the steps. The dataset can then be considered as a collection of steps, with a continuous curve sample for each step. A mean curve corresponding to a step can be calculated, and principal shapes can be calculated. If we choose a landmark time for each step, and align the steps accordingly, then the elementary element in the dataset consists of a curve corresponding to a single step, and also the time of occurrence. We probably are not really interested in the actual time of occurrence of a step, since the time when we start is arbitrary. But we are interested in the relative time between steps. We need to ask ourselves, what data must we store with each step so that the original raw time series can be reconstructed? In our first application, a landmark time is associated with each step so that the steps can be aligned. But the first discriminant shape function was constructed without any use of the step duration. Similar to biological shapes, the step duration corresponds to a size variable, and the difference between using and not using the step duration in the classification procedure should be investigated.

So at this point, we start out with continuous data, and develop a classification procedure. Then we introduce a landmark at each step. We are then presented with three possibilities: 1) classify with the continuous data only, 2) classify with the landmark only, and 3) classify with both the landmark and continuous data.

Before adding any additional landmarks, there is yet another variable to consider. In the first application, a landmark was introduced at a 1-g crossing of the smoothed step time series. The resulting discriminant function was especially informative. The parts of the function that are furthest from the zero line indicates the portions of the step that are important in discriminating between running and walking. In this first application, the discriminant function suggested that the important areas, unsurprisingly, were at 1) the initial heel strike, 2) the subsequent toe action, and 3) the loft. In the initial study a few years ago, a landmark point was introduced not at the 1-g crossing, but at the initial heel strike. This landmark by itself does not carry any more information. We still have just the same step duration. But now the continuous step curves are aligned at the heel strike. If the relative time of occurrence of the heel strike was important before, this information may now be lost since the heel strike is now forced to occur at the same time. On the other hand, this information may be shifted to the other parts of the curve. The question is, how does a change of landmark position effect the classification performance? The answer is not obvious.

Does it make any sense to add additional landmarks? The next landmark could be an additional point. But it could also be the amplitude of the existing

landmark point. In this case the step curve would be normalized to start at the same amplitude. Each step curve would then be paired with a step duration and an initial step amplitude. The original raw time series could be reconstructed from this data, so no information is lost.

The problem of classifying running versus walking steps should be relatively easy. In fact, the first attempt is close to 100%. But the stride data do offer a number of other possibilities: 1) Can we discriminate between subjects? Some studies in the literature ([10], [28], and [29]) suggest that individuals can be identified by their strides. 2) Can we discriminate between different speeds? 3) Can we discriminate between running/walking and not running/walking? These issues will be considered in future work that goes beyond the scope of this thesis (analysis of functional data).

6. Closed Contours: Track Data

To gain a better understanding of the sources and magnitudes of variation in the paleontologists' traces of dinosaur tracks, an experiment was devised that involved traces in which the various sources of error were included and characterized explicitly. This experiment is described in the next section as a preamble to the dinosaur track data described in a subsequent section.

6.1 The Hands Experiment

A high school class was enlisted for an experiment with contours. Seven students provided their left and right hands as prototypes for "dinosaur" tracks. Three students served as "scientists." Each scientist made three tracings of each of the fourteen hands. This resulted in a dataset of 126 tracings.

6.1.1 Preparation of the Hands Data

Tracing of hands is not equivalent to the tracing of a dinosaur track, but a few words may provide some perspective to the variability of the process. Each tracing was made with one student placing a hand on a sheet of paper. The other student (scientist) would then trace an outline of the hand with a pencil. The lines tended to be uneven and light, and were difficult to digitize. Many tracings also had multiple lines and were often not closed curves. In other words, this raw data set was a little "messy."

A follow-up tracing was then made for each outline. The raw tracings were scanned and printed. An additional three people were then drafted, each matching up with one of the scientists. They retraced the original tracings with a

heavy black pen and with the objective of ending up with a well-behaved contour: smooth, no multiple lines, and closed. Each “scientist” in this study is therefore actually a pair of people - one who made the the original tracing, and one who followed up with the revised tracing.

This second set of tracings was then scanned and digitized. Figure 6.1(a) is an example of a contour digitized from one of these tracings. Two contours actually resulted from each tracing - one for the inside edge, and one for the outside edge. The inside edge contour was used in the subsequent work.

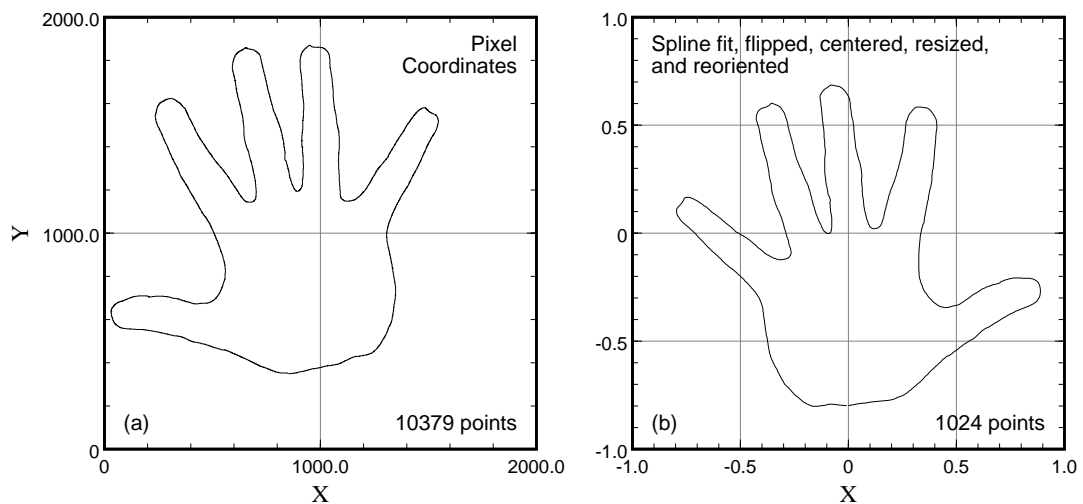


Figure 6.1: Preparation of Hand Data – Registration and Normalization

After being initially digitized from an image, the closed contour is fit with a spline. A set of contours are then resampled to have the same number (1023) of points. They are reoriented and registered so that the contours are aligned in roughly the same direction, and the contour vectors all begin at a corresponding point. They are then normalized (centered about zero with a unit area), and

flipped (right hand becomes left).

6.1.2 Representation of a Track Contour

Building a model with which we can make inferences about track shapes requires that we represent a contour in a way that can be manipulated mathematically. In particular, we aim for some vector representation of a contour. Each contour must then have a consistent length, or number of points. Two methods of representation were considered: 1) tangent angles, and 2) x, y coordinate pairs.

The tangent angles are the angle of a tangent line to a contour (0 to 2π) with an axis. The “flattened” tangents are the difference between the contour tangent and the corresponding tangent of a circle at the same arc length. Both of these tangent angle representations of a track are illustrated in Figure 6.2 for a typical hand. Both of these tangent angle representations of a contour are equivalent to the x, y coordinate representation which is illustrated in Figure 6.3 in the sense that any of these representations can be constructed from any other. For subsequent work, the x, y representation was used.

6.1.3 Calculation of Principal Shapes

It is necessary to place the same number of points around the contour for each track, and we would like the points to correspond to similar features, or landmarks, on different tracks. Mean functions μ_x and μ_y are then calculated separately for the x and y coordinates.

The mean is then subtracted from each track. A vector can then be constructed by concatenating the y 's with the x 's, and a design matrix is then

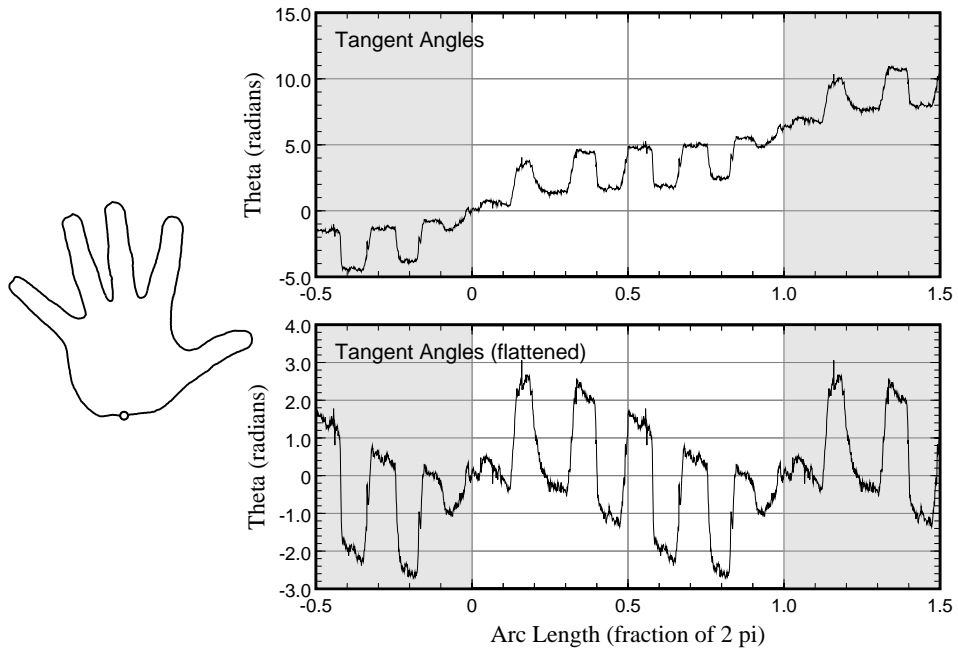


Figure 6.2: Representation of Hand Data – Tangent Angles

constructed from these concatenated vectors.

$$X = \begin{bmatrix} s_1 \\ s_2 \\ \dots \\ s_n \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_1 \\ \dots & \dots \\ x_n & y_n \end{bmatrix} \quad (6.1)$$

The singular value decomposition exists for any matrix.

$$X = U W V' \quad (6.2)$$

where V' is the transpose of V . U and V are both orthogonal, i.e. $U U' = I$ and $V V' = I$ where I is the identity matrix. W is a diagonal matrix with the eigenvalues along the diagonal.

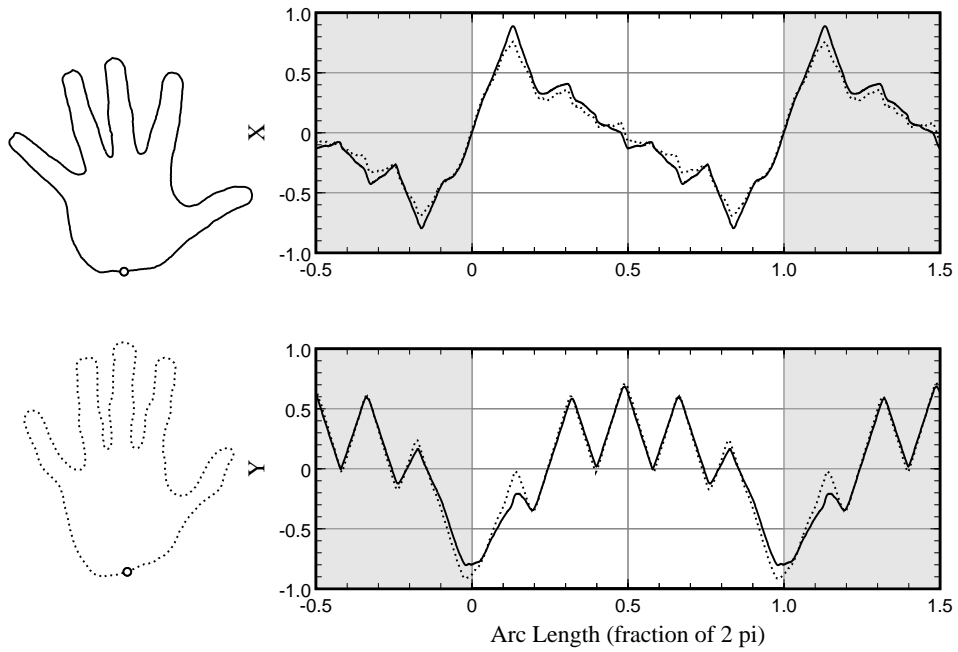


Figure 6.3: Representation of Hand Data - x,y Coordinates

Let's consider, for example, how this is constructed for the contour curves in this study. Each contour was represented with 1024 points. Each concatenated shape vector s is then 2048×1 . With 126 contours, the design matrix X is then a 126×2048 matrix. U and W are then both 126×126 , and V is 2048×126 .

6.1.4 Comparison of Registration Methods

The tangent angle representation of a track has the advantage of shorter vectors with only a single tangent angle (one number) at each point compared to x and y coordinates (two numbers) at each point. The tangent angles are also unaffected by the size of the track, and therefore no normalization to a unit area is required. The principal shapes that are summarized in Figures 6.4 and 6.5 for

the two methods highlights a problem with the tangent angle method for the hands study.

The first two principal shapes have approximately the same explanatory power between the two methods, about 53%. The tangent angle method, however, has principal shapes that are characterized by isolated spikes. These correspond to sharp changes in angle at certain locations around the contour, and it is therefore expected that this method will result in increased uncertainty in the modeling of the tracks using principal shapes. For this reason, the x, y coordinate pair method was used for representation of tracks.

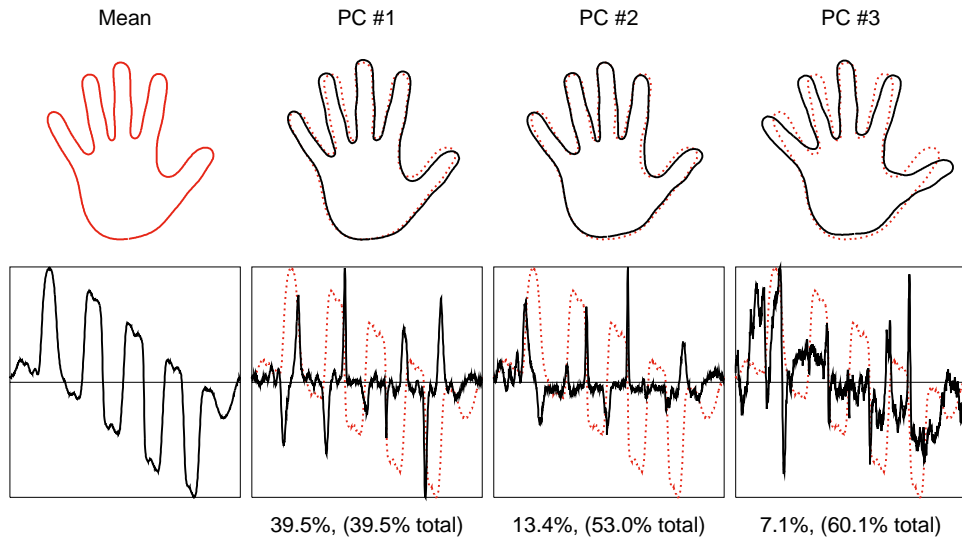


Figure 6.4: Decomposition of Hand Shapes with Principal Components (Tangent Angles)

6.1.5 Reconstruction of an Outline

Once principal shapes are determined, an approximation to the original shape can be calculated. If the shape is represented with x, y coordinate pairs, we first

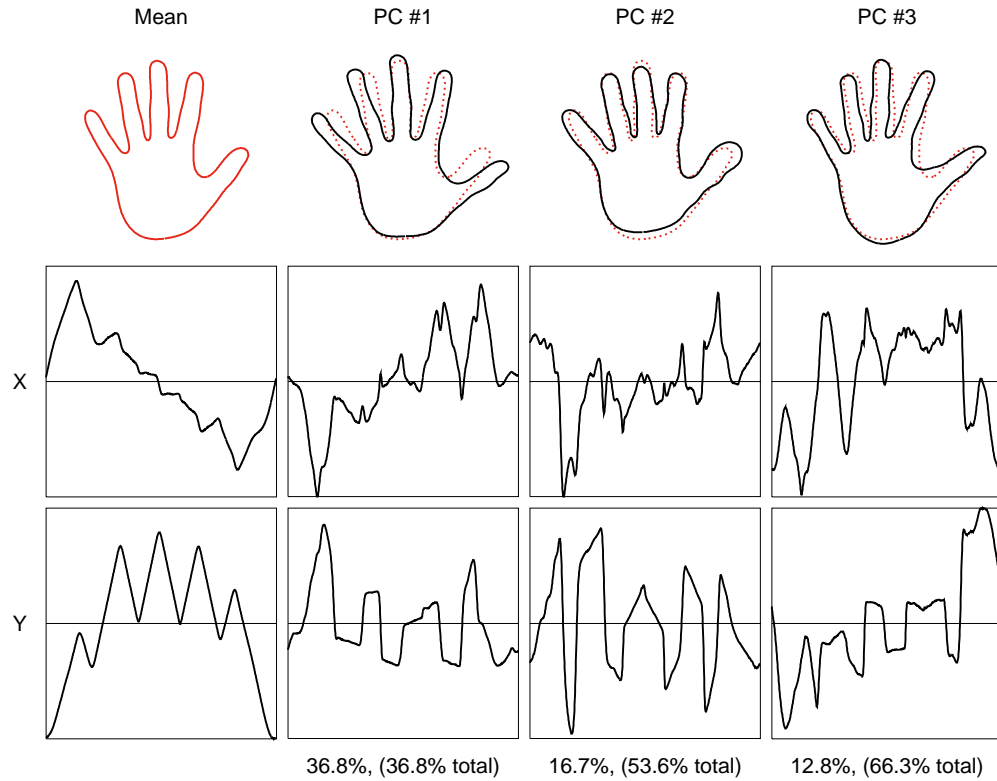


Figure 6.5: Decomposition of Hand Shapes with Principal Components (x,y Coordinates)

construct a $2n \times 1$ vector by concatenating the x and y vectors. If \mathbf{x}_c is this concatenated vector, then

$$\hat{\mathbf{x}}_c = \bar{\mathbf{x}}_c + \mathbf{V}\mathbf{V}'(\mathbf{x}_c - \bar{\mathbf{x}}_c)$$

For a given number of vectors in the basis the principal shapes are optimized to minimize the error between the reconstructed vector and the raw data. An illustration of this can be seen in Figure 6.6. An outline is reconstructed using seven principal components. An alternative basis is the Fourier basis. The example also uses seven coefficients in both x and y components. The Fourier

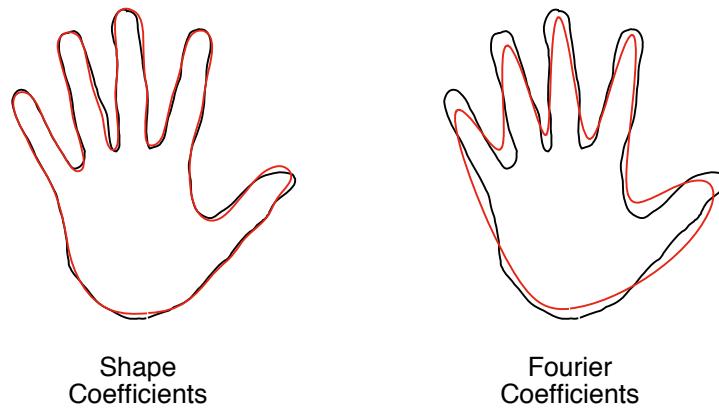


Figure 6.6: Reconstruction of hand shapes: shape -vs- Fourier coefficients

example actually uses double the number of coefficients, because coefficients are applied in both the x and y directions. This example illustrates how the principal shape basis models data more closely than other models.

6.1.6 Analysis of Variance of Shape Data

A closed contour shape, such as the hand outlines being considered here, are infinite dimensional. The hands outlines naturally lend themselves to an ANOVA study, but the dimensionality presents a problem.

$$Y_{ijk} = \mu_{...} + \alpha_i + \beta_j + \varepsilon_{ijk} \quad (6.3)$$

where α is the “scientist” effect, and β is the “dinosaur” effect. Two approaches can be taken for this problem. Although the response variable, the contour outline, is infinite dimensional, we can work with one point of the contour at a time and perform a univariate two-way ANOVA.

For example, consider the hypothesis that the scientist means are all equal: $\mathbf{H}_0 : \alpha_1 = \alpha_2 = \alpha_3$. A univariate ANOVA table is summarized in Table 6.1 for a

Source	SS	df	MS	F	Prob
Mean	0.12648	1	0.12648	221.69	0.0000
Scientists	0.00131	2	0.00066	1.15	0.3209
Dinosaurs	0.03418	6	0.00570	9.99	0.0000
Interaction	0.00661	12	0.00055	0.97	0.4864
Error	0.05991	105	0.00057		

Table 6.1: Analysis of variance table for a single response point

single point on the contour. The large p-value (0.3209) for the scientists indicates that the null hypothesis (equal means) can not be rejected for this contour point. Similarly, the table indicates that there is no significant interaction between scientists and dinosaurs. The null hypothesis can be rejected for the dinosaurs, i.e. their mean contours, considered point by point, are different. The pattern along the rest of the points of the contours remains the same for the dinosaurs and for interaction. The pattern changes, however, for the scientists. At some points their means are equal, and at other points, their means are different.

The mean shapes for both the dinosaurs and scientists are summarized in Figure 6.8. The scientist contours are highlighted at those points where the hypothesis of equal means is rejected. The univariate functional analysis of variance is then seen as a useful technique for identifying those features where contour shapes differ.

Although it may be useful to know where the contour shapes differ, it still does not tell us whether the overall effects of the scientists on shape are signif-

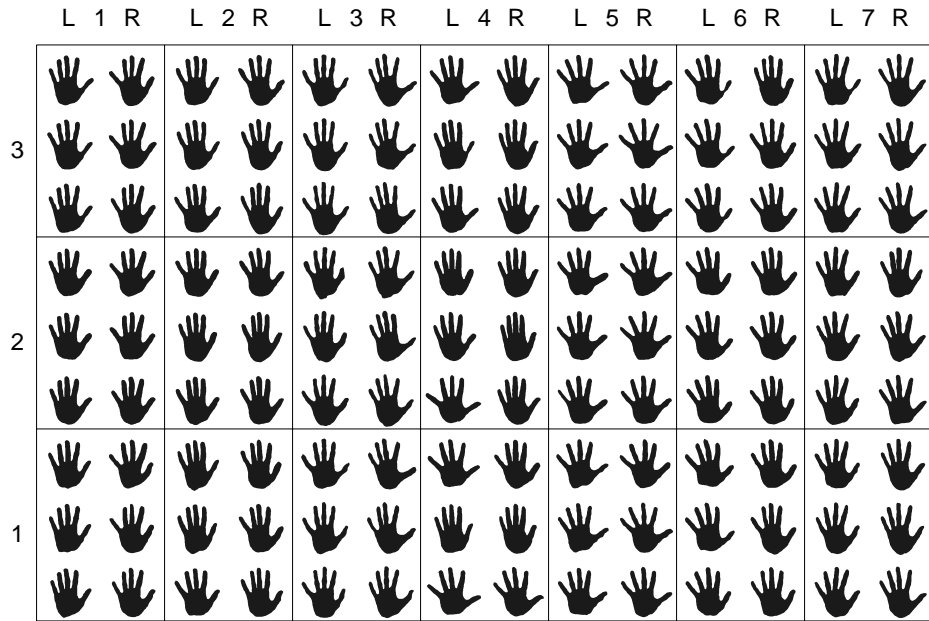


Figure 6.7: Matrix of Hands

icantly different. A multivariate analysis of variance tests the null hypothesis that the entire mean contour shapes are equal. The high dimensionality can be controlled by working with a subset of the principal shapes, and using the coefficient vector as the response variable.

The test statistic that is used here is Wilk's lambda, following the formulation outlined in [17]. A summary of the analysis using the coefficients for only a single principal shape is provided in Table 6.2. Because only a single coefficient is used in the response vector, the analysis is essentially a univariate ANOVA. But the analysis is easily extended in Table 6.3 to incorporate the coefficients for the first two principal shapes. Reviewing the first table shows that there is no interaction, and the differences between scientist mean shapes are

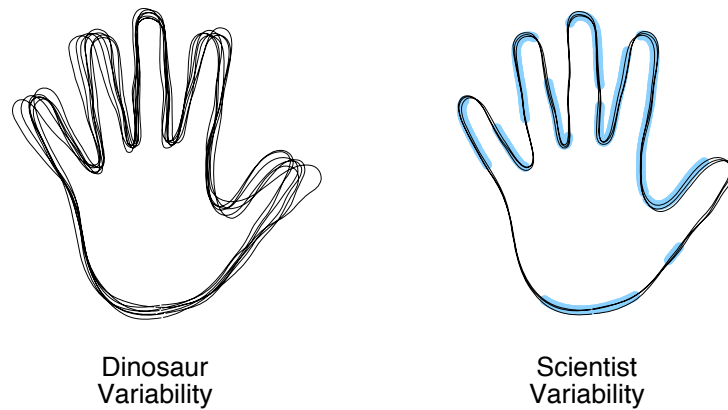


Figure 6.8: Dinosaur effect -vs- scientist effect

marginally significant. This conclusion changes, however, when an additional principal shape is incorporated into the model. Both the scientist and dinosaur effects, as well as their interaction, are now significant with the larger model. Clearly the additional principal shape has provided additional information.

6.2 Dinosaur Tracks

The ideas of functional data analysis are demonstrated as they might be applied to the study of dinosaur tracks. Although the method used here to collect raw data may not be optimal, it did work surprisingly well for the purposes of this proof-of-concept study. A halftone photographic image of a dinosaur trackway was scanned into a computer file with the idea that calculating contours from a photographic image is not unlike calculating topographic contours from three dimensional data. Shapes evaluated in this study consists of the external track outlines for seventeen tracks from a single trackway.

	SS	SS	Wilk's		Num	Den	
Source	res	res+fac	Lambda	F	df	df	Pr> F
Scientists	46.42	50.23	0.9242	4.3031	2	105	0.0160
Dinosaurs	46.42	106.08	0.4377	22.4863	6	105	0.0000
Interaction	46.42	50.96	0.9111	0.8538	12	105	0.5955

Table 6.2: Multivariate analysis of variance table with one principal shape

	SS	SS	Wilk's		Num	Den	
Source	res	res+fac	Lambda	F	df	df	Pr> F
Scientists	309.56	709.27	0.4364	134.2901	1	104	0.0000
Dinosaurs	309.56	4651.57	0.0665	291.7520	5	104	0.0000
Interaction	309.56	400.47	0.7730	2.7767	11	104	0.0034

Table 6.3: Multivariate analysis of variance table with two principal shapes

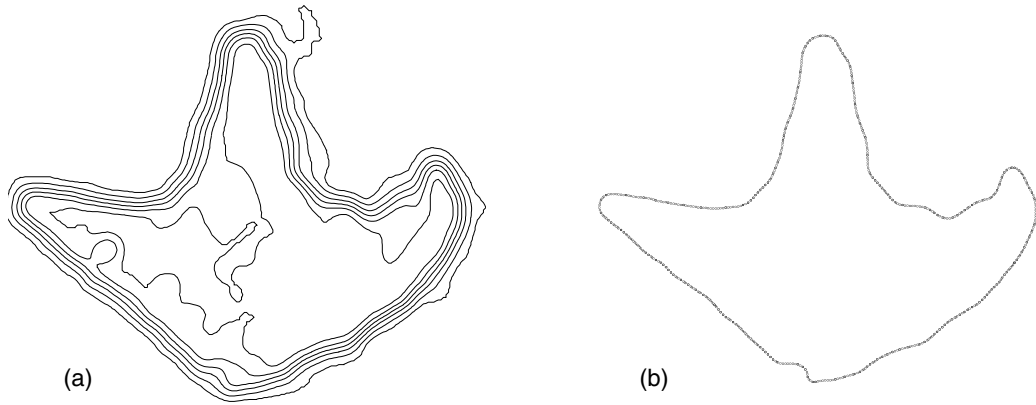


Figure 6.9: Developing a Dinosaur Track Shape from a Contour

6.2.1 Preparation of the Data

External outline shapes of dinosaur tracks can be calculated as contours from three dimensional data. The choice of level from which to develop the contour is somewhat subjective with different contour shapes resulting from different choices of level. The contour that is illustrated in Figure 6.9 was actually developed from a grayscale photographic image published in [43], with the gray intensity (an integer from 0 to 255) substituting for depth.

While not clear to the eye when viewing the journal article, the halftone dots (Figure 6.10a) are an obstacle when attempting to follow a contour. The halftone image was scanned at a high resolution, and there were about eight pixels from one halftone dot to another. With an image processing program, "Gaussian blurring" was then applied with a radius of 8. This effectively got rid

of the dots. Although the resulting image in Figure 6.10b appears out of focus, a smooth contour can now be traced around the image. This is illustrated in Figure 6.10c over a closeup view of the image. The final contour is shown in Figure 6.9b.

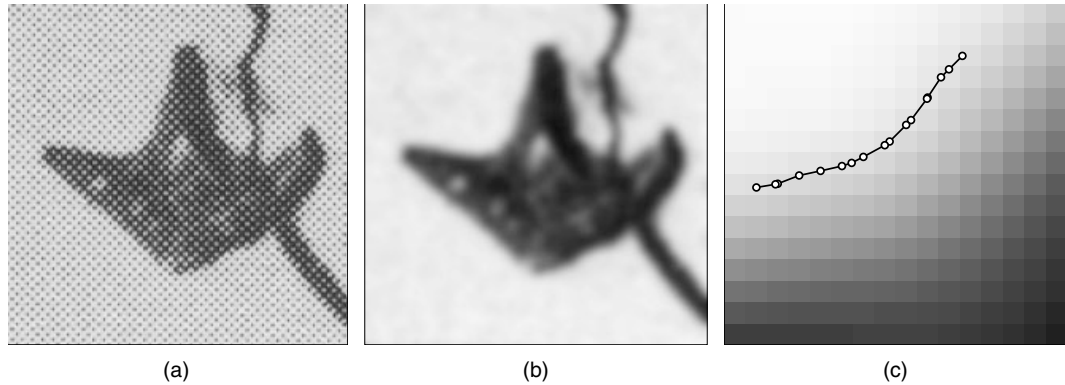


Figure 6.10: Working with a Halftone Image

What is not immediately apparent here is the treatment of cracks contaminating the prints. To expedite the analysis, an image processing program was also used to edit the halftone image by simply erasing the cracks and defining (subjectively again) a clear path around the track. This was done to the halftone before applying the blurring.

The entire image was processed into a single file. As a result, the coordinates of the contours then contained the relative positions of the different tracks. A one meter bar included in the photo allowed for distances in meters to be calculated from the pixel coordinates. The contour outlines that were obtained for all seventeen tracks are shown in a plot of the complete trackway in Figure 6.11.

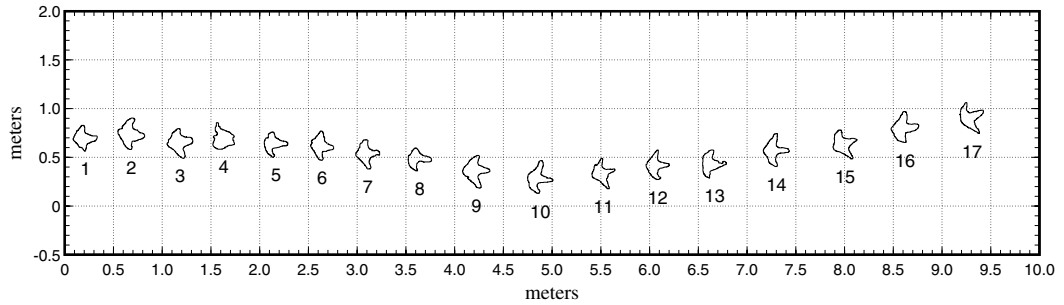


Figure 6.11: Dinosaur Trackway

6.2.2 Principal Shapes

The outline of a track is a continuous curve, and although the tracks are represented here with discrete points, these points are understood to be an approximation to the continuous curve. A textbook summarizing the recent field of functional data analysis is [32]. Starting with a set of track outlines, a decomposition of the shapes can be constructed that models the variability in the set.

6.2.3 Registration of Shape Data

The first task when constructing the principal components for a function is to calculate a mean function. At a given t_0

$$\mu(t_0) = \frac{1}{n} \sum_{i=1}^n f_i(t_0) \quad (6.4)$$

For a function this calculation is straightforward. When given a t_0 for one function, it is generally clear what is the corresponding point on a second function.

But with a closed contour, the correspondence is not so clear. The problem is illustrated in Figure 6.12. When a contour is first constructed with a contouring

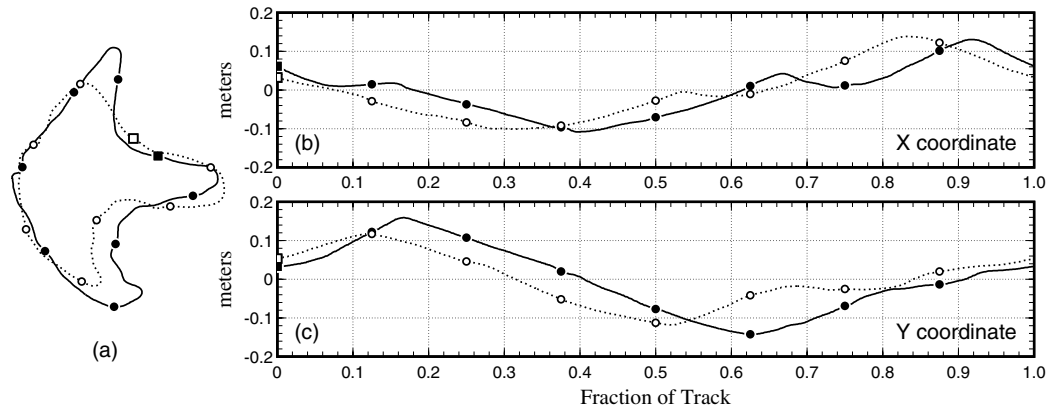


Figure 6.12: Dinosaur Track

algorithm, the discrete points are somewhat irregularly spaced, being located on either the horizontal or vertical centerlines of the pixels in the image. Two different contours will generally be composed of a different number of points, and will have different lengths. The starting point is arbitrary with no correspondence between contours for two different tracks.

In Figure 6.12(b&c) points are scaled to range evenly from 0 to 1. A symbol is then drawn at eight evenly spaced intervals around the track for reference. Just by chance, the starting points are relatively close. But features, e.g. ends of toes, do not correspond between the two shapes.

In [42] shapes are independent of location, scale, and orientation information. The registration of track outlines is closely related to normalizing out scale and orientation information. This was not done here, but the tracks were nearly

the same scale (they were produced by the same dinosaur) and very nearly the same orientation.

In Figure 6.12(b&c) the x and y coordinates are plotted separately. Four landmarks were chosen here as those points that correspond to the minimum and maximum x and the minimum and maximum y . This worked reasonably well, picking points that are at least close to the ends of the three toes and to the heel. After fitting a spline function to both the x and y curves, 250 points were then placed evenly between the landmarks. The result is shown in Figure 6.13. It should now be clear that the eight points located on the contours correspond

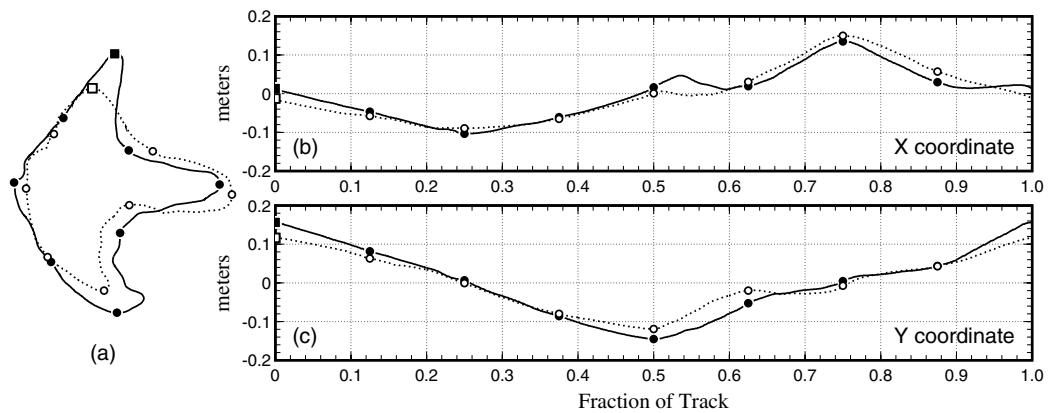


Figure 6.13: Registration of Dinosaur Tracks

much better. Also the x and y curves are much closer. After the shapes are properly registered, or aligned, with each other, it is seen that the distribution of the discrete points along the curve must also be decided.

6.2.4 Interpretation of Principal Shapes

Each column of V is a principal shape. It is a 2002×1 vector, and as before, is constructed with the x portion of the shape followed by the y portion. If we let $C = UW$, so that

$$X = CV' \tag{6.5}$$

then C contains the coefficients for each shape. The first row c_1 of C , for example, contains the 16 coefficients for the first track. The track can then be constructed exactly by applying each of its coefficients to each of the principal shapes.

$$s_1 = c_1V' \tag{6.6}$$

To plot principal shapes that are closed curves, it is best to first add them back to the mean curve. The first five shapes are shown in Figure 6.14. The mean curve is shown for reference as a dotted line with each shape.

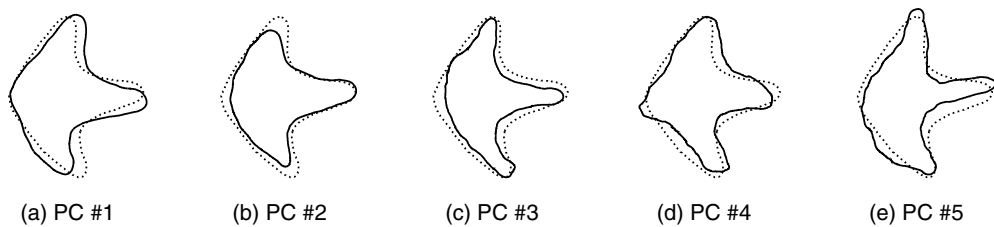


Figure 6.14: Principal Shapes of the 17 Iguanodon Tracks

These shapes do appear to allow some interpretation. The first shape corresponds to a rotation of the track. In this plot, relative to the mean curve,

the principal shape appears to be rotated to the right. A positive coefficient for the first shape would then correspond to a rotation to the right, and a negative coefficient would correspond to a rotation to the left.

The second shape appears to correspond to the size of the main body of the foot relative to the middle toe. This may also correspond to the amount of erosion or deterioration of the shape.

The third shape seems to correspond to the width of the toes. This is probably an artifact of the contouring algorithm since the contour shapes were not calculated with a consistent level for each track. A level closer to white resulted in fatter toes, and a level closer to black resulted in skinnier toes. This effect could then probably be removed by simply setting the coefficient for this shape to zero for all tracks.

The fourth and fifth shapes appear together to model the asymmetry of a track, whether the middle toe is closer to the right or left toe.

6.2.5 Approximation with PC Shapes

The shapes eventually progressively turn to noise. From a dinosaur initially making an impression on the ground, to its finally being retrieved from a small halftone photographic image, some noise is not unexpected. The decomposition of the shape into a sequence of principal shape vectors provides a useful way of separating the noise from the true information.

$$s_1 = \begin{bmatrix} c_{11} & c_{12} \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix} \quad (6.7)$$

Then an approximation to X_1 can be made by keeping the partition

$$s_1^* = c_{11}V_1' \quad (6.8)$$

For example, if we keep two of the principal shapes in our model. And, say, α_1 and β_1 are the coefficients for the first track, then an approximation of the first track can be calculated from the first two principal shapes by

$$s_1^* = \alpha_1 v_1 + \beta_1 v_2 \tag{6.9}$$

An approximation using principal shapes 1,2,4 & 5 is shown in Figure 6.15 for three different tracks. Principal shape 3 was left out of the model because it was judged to model an artifact of the contouring algorithm.

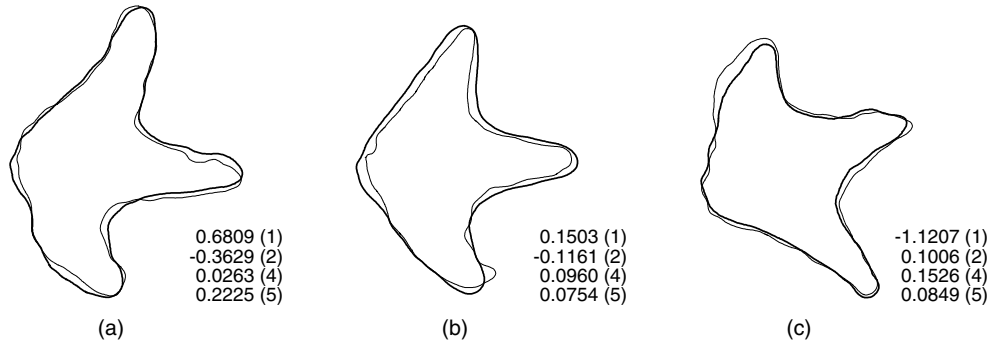


Figure 6.15: Approximating a Track Shape with a Subset of Principal Shapes

6.2.6 Plotting the PC Coefficients

The approximate shapes in Figure 6.15 are completely determined by the values of the four coefficients that were used in the model. Each track can then be thought of as a point in a four dimensional space. Using only two coefficients each track can be considered as a point in a two dimensional space. The shapes of the total sample of tracks can then be summarized as points on a plot, as in Figure 6.16.

In this plot each track is placed by the values of its first two coefficients. As discussed earlier, the first coefficient corresponds to a rotation of the track, with

positive values corresponding to the track being rotated to the right. Considering only the coefficient for PC #1, it is seen that the even numbered tracks (the left foot) tend to be oriented to the right, and the odd numbered tracks (the right foot) tend to be oriented to the left. Only two of the tracks (5 & 13) fail to follow this pattern. So it appears that our dinosaur walks slightly pigeon-toed.

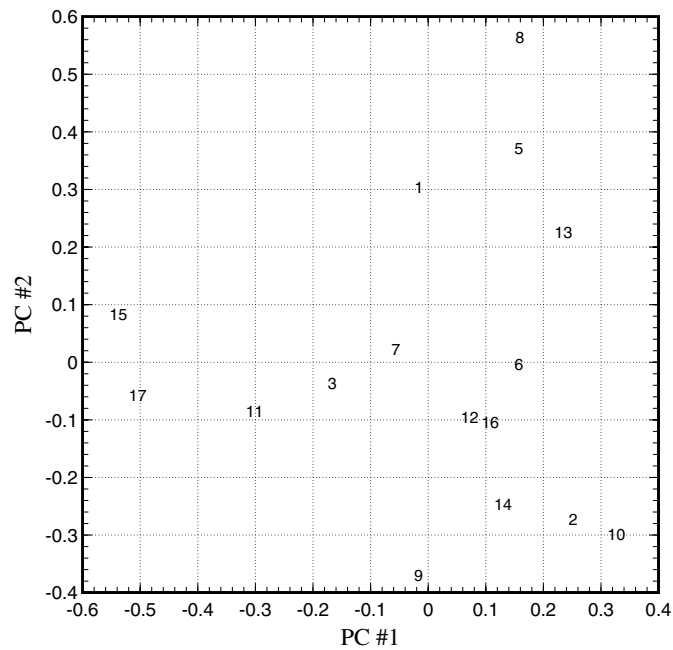


Figure 6.16: Plotting the PC Coefficients in a 2D Space

This plot is only suggestive of further possibilities. A major benefit of decomposing shape data into its principal shapes comes from separating the shapes that represent meaningful information from those shapes that model only noise. As a result, complex continuous shapes which are of infinite dimension, have been reduced to a relatively small set of discrete dimensions, and can be represented by a small set of coefficients.

Once this smaller dimensional space has been determined, other tools can be brought to the problem. A cluster analysis, for example, could be used to identify distinct groups of shapes within a dataset. Discrimination analysis could be applied to identify what species produced the tracks.

The principal shapes provide the basis for a vector space in which the population of tracks can be represented. Once this space has been determined, it is possible to find other bases, or other shapes, that also model the same space.

6.2.7 Shape of the Trackway

A simple approach was taken with the location of each track relative to each other. Referring to Figure 6.17, connecting a line between each two adjacent tracks, a point is located at the center. An interpolating spline was then fit through these center points. This procedure appears to identify the tracks of the left foot versus the right foot. Track #10 falls on the opposite of the spline curve, but this is apparently just an artifact of the procedure, because this particular track coincides with the extreme right of the trackway.

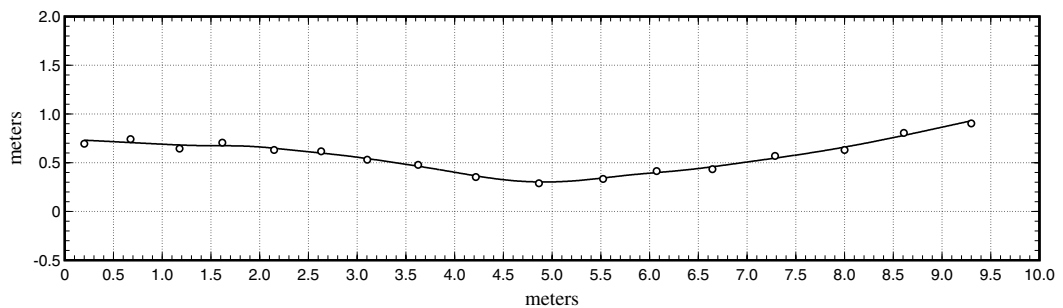


Figure 6.17: Locating the Central Path of a Trackway

Figure 6.18 attempts to illustrate the track shapes relative to their distance from the central trackway.

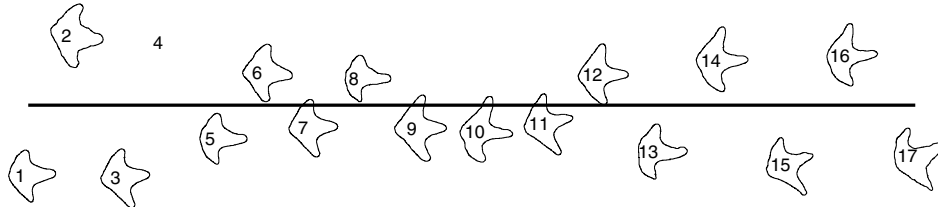


Figure 6.18: Dinosaur Tracks relative to the Central Path

6.2.8 Correlation of Manus and Pes Tracks

The topic of the [43] paper concerned the manus tracks and their relationship with the pes tracks. It would be interesting to look at this problem with the techniques used here, but it was difficult to identify the manus tracks from the small halftone image available. Canonical correlation might be used to study the relationship between the manus shapes and the pes shapes, as well as their location and orientation.

6.2.9 Internal Structure of Tracks

Apparently, no internal structure was evident in the tracks that were studied here. But this is not always the case. Structural data of the foot, e.g. pad shapes, location of joints, and so on, should also yield to the kind of analysis suggested in this study.

6.2.10 Efficient Methods to Collect Data

While the study of dinosaur tracks is necessarily observational, there is still a great deal of flexibility in deciding how tracks can best be measured to facilitate the subsequent analysis. The method that was used in this paper, scanning a small halftone image from a journal article, is crude. Working from a photograph is one approach. The application of polynomial texture mapping is another approach that would also provide a measure of depth into a track. There might be methodology that could measure the density of the underlying strata.

6.3 Mathematical Issues and Questions

Although functional data analysis is a relatively well developed field, the analysis of track data does present a number of issues specific to the study of track data which are interesting mathematically.

6.3.1 The Registration Problem

The registration of closed outline curves as well as the placement of discrete points to represent the curves presents problems not typically encountered with functional data which have a clearly defined independent variable. This is also closely related to the taking of landmark data. This approach is typical of the paleontology community, and it seems to make good sense. But it should be interesting to learn how to frame this as an optimization problem. What is the optimal way to orient two semi-irregular shapes relative to each other? What is the optimal way to distribute discrete points to characterize the shapes?

6.3.2 Masking of Information

Consider two scenarios. First, principal shapes are constructed for tracks which have not been normalized for size and orientation. In the interest of time, this is the approach that was taken in this study. In the second scenario tracks are first normalized for size and orientation, and then principal shapes are constructed. The principal shapes form a basis for a vector space, and may be thought of as a model for the data. In the first case, it would appear that this model must work harder to account for size and orientation. This is not necessary in the second case. What are the differences between the models? Will the second scenario somehow model the information that remains better than the first?

6.3.3 Missing Data

Generally with functional data, the entire function is available. While not addressed in this preliminary study, many of these tracks were interrupted with cracks or large missing sections of the underlying rock. In one case, in a different trackway that wasn't studied here, only one of the toes was evident. The quick solution taken here was to use image processing software to simply draw through the missing sections to define where outline of the track appeared to be. This appears to be an unusual missing value problem. When sections of the curves are missing, the subsequent distribution of points on the visible sections of a track outline is problematic.

7. Conclusions and Future Work

In this thesis I have developed a common methodology to the analysis of curve data ranging from curves which are strictly functional data to curves which are closed contour shapes.

Future research will be most immediately concerned with the stride data, and the immediate concern here is with implementation of the ideas presented in this thesis. In development of a statistical model, outliers can be liberally identified at leisure and removed from any model development. Implementation of the methodology for a monitoring device, however, requires that outliers be treated much more quickly, and robust techniques for working with the stride acceleration time series data will be essential.

Appendix A. Plots related to the SCARF filter

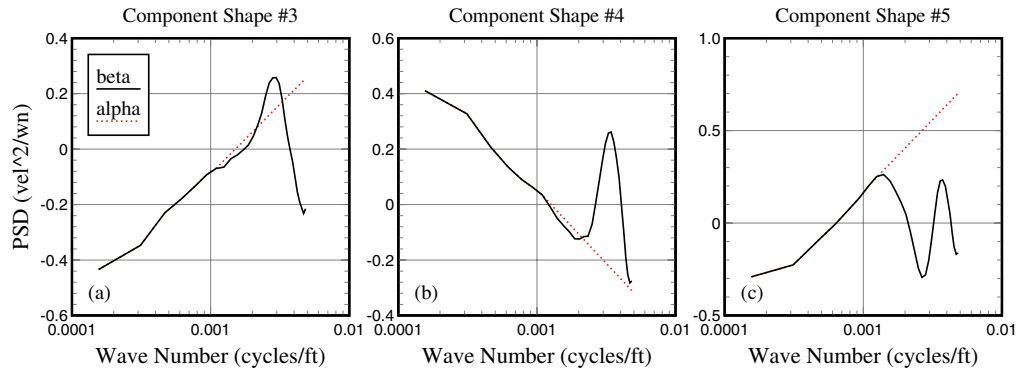


Figure A.1: Component Spectral Shapes of Wind Spectra

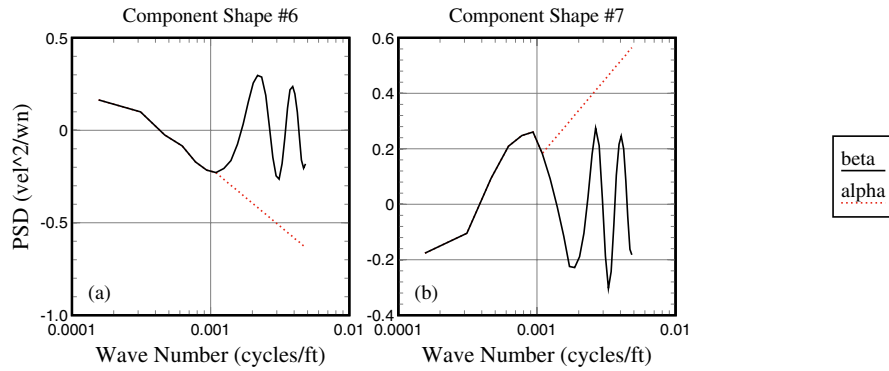


Figure A.2: Mean Spectral Shapes

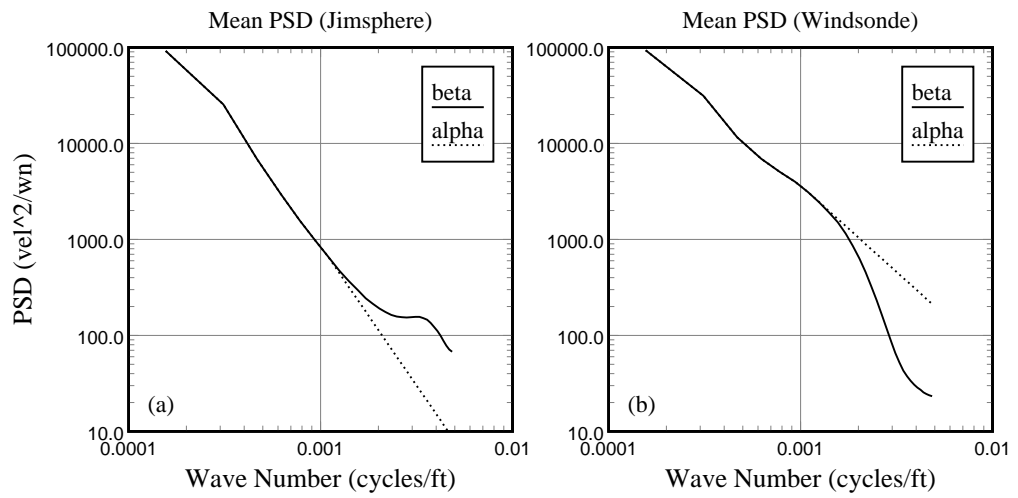


Figure A.3: Mean PSD's

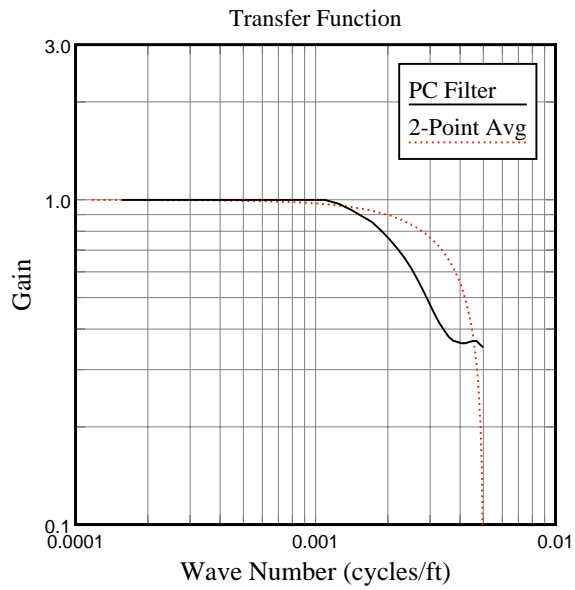


Figure A.4: Mean Transfer Function

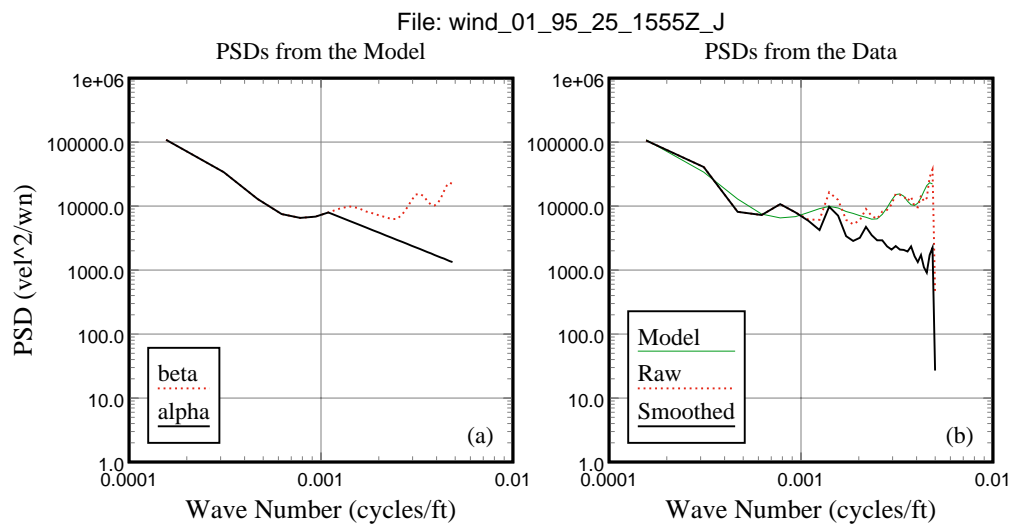


Figure A.5: PSD Comparison: Turbulent Wind

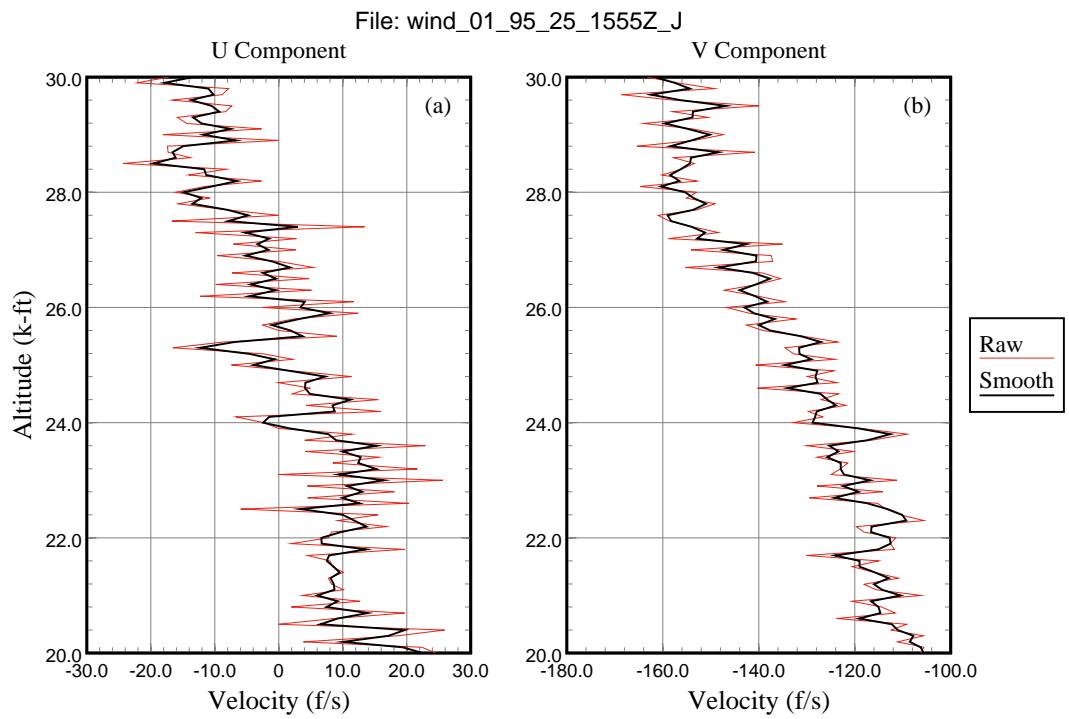


Figure A.6: Smoothed Wind Profiles: Turbulent Wind

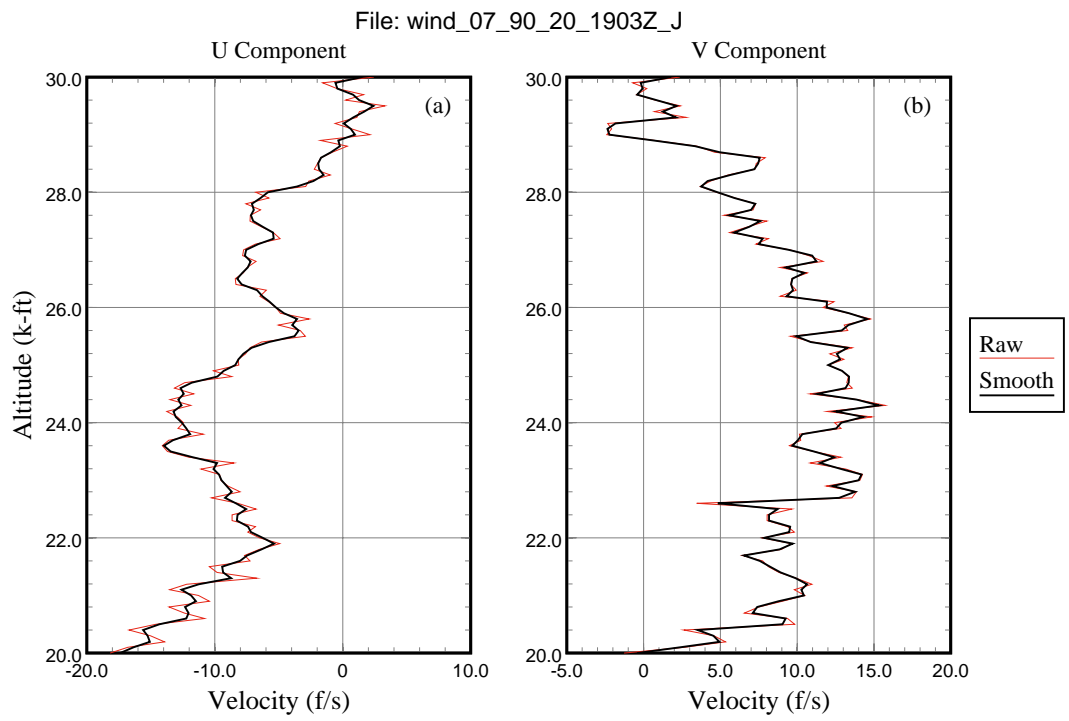


Figure A.7: Smoothed Wind Profiles: Typical Wind

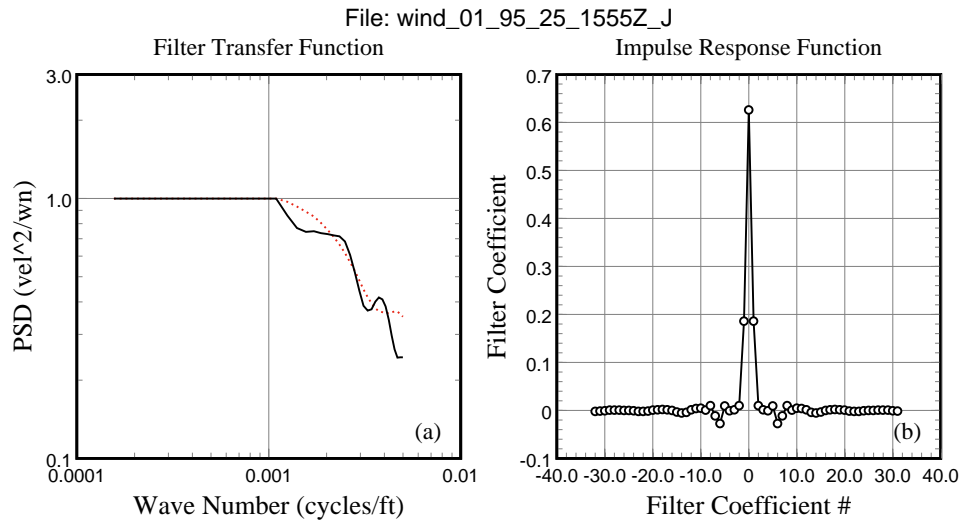


Figure A.8: Impulse Response Function: Turbulent Wind

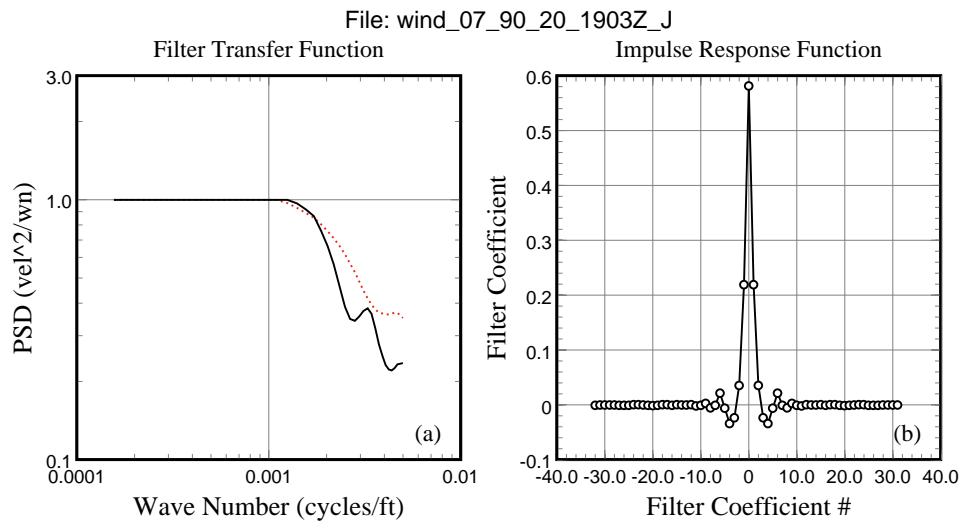


Figure A.9: Impulse Response Function: Typical Wind

REFERENCES

- [1] J. S. Bendat and A. G. Piersol. *Random Data: Analysis and Measurement Procedures*. John Wiley & Sons, New York, 3rd edition, 2000.
- [2] Fred L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1(2):181–222, May 1986.
- [3] Fred L. Bookstein. *Morphometric tools for landmark data*. Cambridge University Press, 1991.
- [4] Jorge F. C. L. Cadima and Ian T. Jolliffe. Size- and shape-related principal component analysis. *Biometrics*, 52:710–716, 1996.
- [5] I. Cohen and I. Herlin. Curves matching using geodesic paths. *IEEE Proceedings of Computer Vision and Pattern Recognition, Santa Barbara*, pages 741–746, June 1998.
- [6] Frank Critchley. Influence in principal components analysis. *Biometrika*, 72(3):627–636, 1985.
- [7] Frank Critchley and Ian Ford. On the covariance of two noncentral F random variables and the variance of the estimated linear discriminant function. *Biometrika*, 71(3):637–638, 1984.
- [8] Frank Critchley and Ian Ford. Interval estimation in discrimination: The multivariate normal equal covariance case. *Biometrika*, 72(1):109–116, 1985.
- [9] Frank Critchley and Cecilia Vitiello. The influence of observations on misclassification probability estimates in linear discriminant analysis. *Biometrika*, 78(3):677–690, 1991.
- [10] S. R. Das, R. C. Wilson, M. T. Lazarewicz, and L. H. Finkel. Gait recognition by two-stage principal component analysis. *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition (FGR'06)(IEEE)*, February 2006.

- [11] Kim-Anh Do and Katherine Kirk. Discriminant analysis of event-related potential curves using smoothed principal components. *Biometrics*, 55:174–181, 1999.
- [12] I. L. Dryden and K. V. Mardia. Size and shape analysis of landmark data. *Biometrika*, 79(1):57–68, 1992.
- [13] Randall L. Eubank. *Nonparametric Regression and Spline Smoothing*. Marcel Dekker, New York, 2nd edition, 1999.
- [14] Colin R. Goodall and Kanti V. Mardia. Multivariate aspects of shape theory. *The Annals of Statistics*, 21(2):848–866, 1993.
- [15] D. J. Hirst, I. Ford, and F. Critchley. An empirical investigation of methods for interval estimation of the log odds ratio in discriminant analysis. *Biometrika*, 77(3):609–615, 1990.
- [16] D. C. Hoaglin, F. Mosteller, and J. W. Tukey. *Understanding robust and exploratory data analysis*. John Wiley and Sons, Inc., New York, New York, 1983.
- [17] R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis*. Prentice Hall, New Jersey, 3rd edition, 1992.
- [18] I. T. Jolliffe. *Principal Component Analysis*. Springer-Verlag, New York, 2nd edition, 2002.
- [19] Abram Jujunashvili. *Angles between Infinite-dimensional Subspaces*. PhD thesis, University of Colorado at Denver, Denver, Colorado, 2005.
- [20] David G. Kendall. A survey of the statistical theory of shape. *Statistical Science*, 4(2):87–99, May 1989.
- [21] Peter Lachenbruch. Discriminant analysis when the initial samples are misclassified. *Technometrics*, 8(4):657–662, November 1966.
- [22] Peter Lachenbruch. Zero-mean difference discrimination and the absolute linear discriminant function. *Biometrika*, 62(2):397–401, 1975.
- [23] Peter Lachenbruch. Discriminant analysis. *Biometrics*, 35:69–85, 1979.
- [24] S. E. Leurgans, R. A. Moyeed, and B. W. Silverman. Canonical correlation analysis when the data are curves. *J. R. Statist. Soc. B*, 55(3):725–740, 1993.

- [25] K. V. Mardia and I. L. Dryden. The statistical analysis of shape data. *Biometrika*, 76(2):271–281, 1989.
- [26] K. V. Mardia and A. N. Walder. Shape analysis of paired landmark data. *Biometrika*, 81(1):185–196, 1994.
- [27] Geoffrey J. McLachlan. *Discriminant analysis and statistical pattern recognition*. John Wiley & Sons, 1992.
- [28] L. Middleton, A. A. Buss, A. Bazein, and M. S. Nixon. A floor sensor system for gait recognition. *Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies (Autold'05)*, March 1995.
- [29] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer. The gait identification challenge problem: Data sets and baseline algorithm. *Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02)(IEEE)*, January 2002.
- [30] Michael J. Prentice and Kanti V. Mardia. Shape changes in the plane for landmark data. *The Annals of Statistics*, 23(6):1960–1974, 1995.
- [31] J. O. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer-Verlag, New York, 1997.
- [32] J. O. Ramsay and B. W. Silverman. *Applied functional data analysis, methods and case studies*. Springer-Verlag, New York, 2002.
- [33] C. R. Rao and S. Suryawanshi. Statistical analysis of shape of objects based on landmark data. *Proc. Natl. Acad. Sci.*, 93:12132–12136, October 1996.
- [34] John A. Rice and B. W. Silverman. Estimating the mean and covariance structure nonparametrically when the data are curves. *J. R. Statist. Soc. B*, 53(1):233–243, 1991.
- [35] John J. Richardson. A novel device for monitoring energy expenditure. *Grant Application to the Department of Health and Human Services, Public Health Services*, 2004.
- [36] P. J. Rousseeuw, S. Van Aelst, and M. Hubert. Regression depth: Rejoinder. *Journal of the American Statistical Association*, 94(446):419–433, June 1999.
- [37] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection*. John Wiley and Sons, Inc., 2003.

- [38] Thomas B. Sebastian, Philip N. Klein, and Benjamin B. Kimia. On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(1):116–125, 2003.
- [39] Robin Sibson. Studies in the robustness of multidimensional scaling: Procrustes statistics. *J. R. Statist. Soc. B*, 40(2):234–238, 1978.
- [40] B. W. Silverman. Spline smoothing: The equivalent variable kernel method. *The Annals of Statistics*, 12(3):898–916, 1984.
- [41] B. W. Silverman. Incorporating parametric effects into functional principal components analysis. *J. R. Statist. Soc. B*, 57(4):673–689, 1995.
- [42] Christopher G. Small. *The Statistical Theory of Shape*. Springer-Verlag, New York, 1996.
- [43] Joanna Wright. Ichnological evidence for the use of the forelimb in iguanodontid locomotion. *Special Papers in Paleontology*, 60:209–219, 1999.
- [44] R. Zhang, P. Tsai, J. E. Cryer, and M. Shah. Shape from shading: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8), August 1995.