

## ABSTRACT

Title of Thesis: HUMAN ACTIVITY CLASSIFICATION  
BASED ON GAIT AND  
SUPPORT VECTOR MACHINES

Amon Brigoli Ducao II, Master of Science, 2008

Dissertation directed by: Professor Rama Chellappa  
Department of Electrical and Computer Engineering

Presented is a method to characterize human gait and to classify human activities using gait. Slices along the x-t dimension of a spatio-temporal sequence are extracted to construct a gait double helical signature (gait DHS). A DHS pattern is a compact description that encodes the parameters of human gait and shows inherent symmetry in natural walking (without encumbered limb movement). The symmetry takes the form of Frieze groups, and differences in DHS symmetry can classify different activities. This thesis presents a method for extracting gait DHS, and how the DHS can be separable by activity. Then, a Support Vector Machine (SVM) n-class classifier is constructed using the Radial Basis Function (RBF) kernel, and the performance is measured on a set of data. The SVM is a classification tool based on learning from a training set, and fitting decision boundaries based on an output function. This thesis examines the effect of slicing at different heights of the body and shows the robustness of DHS to view angle, size, and direction of motion. Experiments using real video sequences are presented.

HUMAN ACTIVITY CLASSIFICATION BASED ON GAIT AND  
SUPPORT VECTOR MACHINES

by

Amon Brigoli Ducao II

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park in partial fulfillment  
of the requirements for the degree of  
Master of Science  
2008

Advisory Committee:  
Professor Rama Chellappa, Chair/Advisor  
Professor Robert Newcomb  
Professor Adrian Papamarcou

© Copyright by  
Amon Brigoli Ducao II  
2008

## Dedication

*Dedicated to everyone who said this was a good idea...*

*and especially those who said it was a bad idea.*

## Acknowledgments

First and foremost, I am extremely thankful to my advisor, Dr. Rama Chelappa, for his constant support and guidance. I am especially thankful for the seemingly endless opportunities he has provided for me since I was an undergraduate at the University of Maryland.

I also must thank Dr. Yang Ran, who developed gait DHS and the groundwork program that my thesis is based on, and Mr. Gavin Rosenbush, who studied and implemented gait DHS into a live program with me, allowing the programs I use to actually work. Both have become good friends and ping pong rivals.

I would like to thank Dr. Robert Newcomb and Dr. Adrian Papamarcou for agreeing to be on my thesis defense committee as well as teaching me during my undergraduate years.

I wish like to thank all my family and friends for their love, support, and encouragement, especially my parents, who *implied* that graduate school was not a choice, it was *the* choice.

Finally, I wish to thank the unknown poster on a Japanese OpenCV wiki for posting the code to patch a thesis-crippling flaw in OpenCV. I would like to thank them most of all.

# Table of Contents

List of Tables	vi
List of Figures	vii
1 Introduction	1
1.1 Motivation . . . . .	1
1.2 Previous Work . . . . .	3
1.3 Contribution . . . . .	4
1.4 Organization of Thesis . . . . .	4
2 Activity Sequence and Human Body Model	6
2.1 The Data Sequences . . . . .	6
2.1.1 Activity Sequence . . . . .	6
2.1.2 Data Sets . . . . .	6
2.2 Human Body Characteristics and Motion Model . . . . .	8
2.3 Human Body Model . . . . .	10
2.4 Activity Descriptions . . . . .	13
2.5 Description of Algorithm . . . . .	14
3 GaitDHS Construction	15
3.1 Background Subtraction . . . . .	15
3.2 Size Filtering and Target Tracking . . . . .	17
3.3 Image slicing and DHS construction . . . . .	17
3.4 DHS patterns of Activities and Symmetries . . . . .	20
3.5 Symmetry: Frieze Groups . . . . .	20
3.6 DHS patterns . . . . .	21
3.6.1 Symmetry and Periodicity of Arm DHS patterns . . . . .	26
4 Gait SVM Construciton	27
4.1 SVM derivation . . . . .	27
4.1.1 Linearly Separable Dataset . . . . .	27
4.1.2 Non-Linearly Separable Data Set . . . . .	31
4.1.3 Non-Linear Kernel SVM Method . . . . .	32
4.1.4 n-class classifiers . . . . .	35
4.2 Training on the DHS patterns . . . . .	35
5 Results	37
5.1 DHS Under Changes in Walking Angle . . . . .	37
5.2 SVM results . . . . .	39
5.2.1 Determining Parameters through Test Set Validation . . . . .	39
5.2.2 Arm Performances . . . . .	40
5.2.3 Leg Performances . . . . .	42

6	Conclusion and Future Work	45
6.1	Modification . . . . .	45
6.2	Extension . . . . .	46
	Bibliography	48

## List of Tables

5.1	Selected Parameters . . . . .	41
5.2	Confusion Matrix for Slices Taken at <b>Elbow Height</b> . . . . .	42
5.3	Confusion Matrix for Slices Taken at <b>Wrist Height</b> . . . . .	43
5.4	Confusion Matrix for Slices Taken at <b>Knee Height</b> . . . . .	43
5.5	Confusion Matrix for Slices Taken at <b>Ankle Height</b> . . . . .	44



## List of Figures

2.1	<i>Example of an Activity Sequence as a Stack of Frames</i> . . . . .	7
2.2	<i>Subject Walking Fronto-Parallel to Camera:</i> A subject is walking fronto-parallel to the camera such that his sagittal plane (blue) is parallel to the camera's image plane. This is also known as a profile view. A subject is said to be "walking 45° to the camera" if their sagittal plane makes a 45° angle with the image plane. . . . .	7
2.3	The Average relative heights of specific points on the body[15] . . . . .	9
2.4	<i>Image of the Average Gait Cycle:</i> Includes duration of phases relative to the total gait cycle time [12] . . . . .	10
2.5	<i>Flow Chart of the Proposed Algorithm:</i> This chart shows how the method to be described in this thesis is laid out. . . . .	14
3.1	<i>Background Subtraction on the Same Frame:</i> The left image shows the original frame (with bounding box and lines drawn across slice heights (described later). The top right image shows the monochromatic, Gaussian background subtraction scheme, which was sensitive to shadow or near-foreground-chroma areas, while the bottom right image shows background subtraction using chroma gain and off-unity thresholding. The grey line is the center line of the body, and not used in background subtraction. . . . .	16
3.2	<i>Gait DHS construction process:</i> The activity sequence (a) and the silhouettes produced (b) are used to create slices at specific heights like the ankle (c). These slices are stacked over time to create the DHS band (d). . . . .	19
3.3	<i>Gait Parameters Captured by the DHS pattern.</i> . . . . .	20
3.4	<i>Frieze Groups:</i> Examples of the seven Frieze symmetry groups: (1) Hop, (2) Step, (3) Jump, (4) Sidle, (5) Spinning Hop, (6) Spinning Sidle, (7) Spinning Jump . . . . .	22
3.5	<i>DHS patterns for Arm Activities:</i> Common DHS patterns for the three types of arm loading conditions are shown for slices taken at approximate elbow and wrist height. The red box is the bounding box drawn by the DHS construction program. . . . .	24

3.6	<i>DHS patterns for Leg Activities:</i> Common DHS patterns for the two types of leg loading conditions. In the second image, the subject's right leg is being loaded by three five pound leg weights. . . . .	25
3.7	<i>Periodicity and Symmetry Across Arm Activities:</i> Under the same view (fronto-parallel), all arm activities display translational symmetry (type 1 Frieze group), except for the "Two Arms Occupied" case. The difference comes in periodicity. Slices are taken at wrist height. .	26
4.1	<i>Linear Separable Case:</i> black dots are one class, white stars are another. The input space is separable by a simple hyperplane or in the 2D realm, a boundary line. . . . .	29
4.2	<i>Non-Linear Separable Case:</i> Outliers may exist, but not addressing them could alter the hyperplane. Adding an error margin relaxes the constraints on the boundary and maintains a good separating hyperplane. . . . .	31
4.3	<i>Separation With a Complex Boundary:</i> The input space can be mapped to a higher dimension feature space where a simple separating hyperplane exists. Mapping back to the original input space creates a more complex and accurate boundary. . . . .	33
5.1	<i>DHS patterns for Arm Activities at an Angle:</i> The first row shows the arm DHS patterns as in figure 3.5, while the second row shows the DHS patterns of the same activity, but for the subject walking at 45°, 45°, and 40° from left to right. Similar patterns exist for loaded arm conditions, but a different pattern exists for the "No Hands Occupied" case. . . . .	38
5.2	<i>Average Detection Rate Across a Range of Gammas for Arm Height Slices</i> . . . . .	40
5.3	<i>Average Detection Rate Across a Range of Gammas for Leg Height Slices</i> . . . . .	41

# Chapter 1

## Introduction

### 1.1 Motivation

Human activity recognition is used in fields like medical research, sports, and increasingly in surveillance. Specifically, the ability to analyze gait opens up the potential of studying physiological ailments, optimize motion, and classify and model certain human actions. There are many modes in which to capture a gait sequence, among them both marker and marker-less video capture, and force data [5].

The simple act of human bipedal motion is of interest in areas such as the effects of alzheimers on gait, to modeling forces and stride for implementation in robots. Human gait can be seen as a simple periodic signal: limbs flex and swing in a pendulum-like fashion, the silhouette of the body deforms and returns to a starting shape in a regular tempo. Research in this area has extended to identification of a person based on his/her unique walking pattern.

In this work, we attempt to automatically declare what load conditions a person is under while walking in front of a single camera. We can accomplish this by analyzing the normal gait cycle of a person under no load on his back or limbs, carrying no object, and without any physical afflictions. We then analyze how gait under different conditions, like any of the previously named restrictions, will change the gait signature.

The classification problem can be divided into two important steps:

1. Characterizing the **features** of the subject/activity to be classified
2. Developing a method both to **distinguish** among those characterizations as well as predict what an unknown characterization would be **classified** as

Toward the first problem, we present a characterization of human gait that encodes body articulation and appearance of human motion as well as reveals natural symmetry of the body in motion. We accomplish this by decomposing sequences of frames into a stack of x-t slices. The periodic swing of limbs creates an image similar to the double helical twist of DNA, and so we call this stack of slices the gait **Double Helical Signature**, or gait DHS. The majority of this work was developed from concepts developed by Ran and Chellappa [11, 10].

Toward the second realm, we use a machine-learning tool used for classification and regression problems: the support vector machine (SVM). Given a set of input feature vectors  $\{x_1, x_2, \dots, x_i\}$  and class labels for each vector,  $\{y_1, y_2, \dots, y_i\}$ , an SVM can learn the structure of the hyperplane(s) that defines a boundary between each of the input vectors in the feature space, and so can assign a class label to a new feature vector. SVMs have found applications in areas like hand-written letter recognition and gender based on face.

In this thesis, only a small number of human actions is measured, specifically in relation to arm movement and as it is affected by weights on the leg.

## 1.2 Previous Work

A great deal of work has already been done on gait and classification, either in the area of markerless gait feature extraction, gait symmetries, pattern analysis of a sequence of temporal slices, and use of SVMs for classification/identification using gait.

Yoo, Nixon, and Harris also used a markerless set up, but used angle data as their feature set [18]. They would first calculate the body contour of the walker, then perform line fitting by a least-squares method in order to the femur and tibia/fibula of the legs. The joint angle of the knees could be calculated from these line segments using the slope of the lines segments [18].

Mowbray and Nixon used a similar system of markerless video capture and analysis [8]. But rather than examine the limbs individually, treating the leg as separate oscillating entity, they used Fourier descriptors of the entire foreground mask. Fourier descriptors have "long been established as a method for representing a two-dimensional shape's boundary" [8]. Human gait now becomes a cycle of deformed shapes, upon which a two-dimensional Fourier series representation can be obtained. The use of Fourier analysis is beneficial because the natural periodicity and frequencies of motion become encoded in the feature space. However, it is harder to analyze separate limb pairs, specifically arms versus legs, separately.

Nandy performed similar analysis on angle data using a marker-based data set [9]. He further extended this data to simulate human motion for different activities (normal walking, limping, walking with a backpack). His analysis using auto-

regressive and moving average models and dynamic time warping shows that these activities are both very similar among subjects and very different between activity classes [9].

In [17], SVM's are used on gait data to determine the two-class problem of gender classification. In their work, the SVM's were trained and classified subjects based on angle joint data at three places: the knees, the ankles, and at the hips. They achieved very good recognition rates, but did no work on activity classification.

Support vector machines have also been used on gait data in [16], this time for the purpose of subject identification. The feature they use, however is a discrete wavelet transform (DWT) decomposition of the binary silhouette of the subject, which loses body articulation.

### 1.3 Contribution

In this thesis, we capture many characteristics of gait in a compact manner using gait DHS. We also show how human gait under different loads change the basic DHS pattern. Finally, we try to discriminate between these conditions by classifying DHS signatures using support vector machines.

### 1.4 Organization of Thesis

In this chapter, we laid the ground work for what we wish to develop, a system that can characterize and classify human activity, as well as introduced the problem behind it and previous work concerning the solution. In chapter 2, we introduce the

activity sequence, which is decomposed to construct the DHS, and develop a human body model to analyze. Chapter 3 shows how the DHS pattern is constructed from an activity sequence, and shows the common patterns for the activities we analyze. In chapter 4, we provide a derivation of the SVM classifier, then show how parameters are selected. Chapter 5 provides a summary of the results. Finally, in chapter 6, we draw conclusions and suggest future work on this topic.

## Chapter 2

### Activity Sequence and Human Body Model

#### 2.1 The Data Sequences

##### 2.1.1 Activity Sequence

Each point on the human body can be plotted in the three-dimensional space, and as a movement is performed, a trajectory can be plotted in four-dimensional  $X, Y, Z, t$  space. We define this as the **Activity Volume**.

A camera system projects the four-dimensional activity volume into a sequence of two-dimensional images. In [11], an **Activity Sequence** is defined for a human activity, where the activity is used synonymously with gait. An activity sequence is created by stacking the two dimensional video frames, creating a three dimensional volume  $(x, y, t)$ . Motion of a point of interest in a series of video frames can be plotted in this space. In subsequent discussions, we use the term activity sequence and sequence interchangeably and both to mean a captured video sequence, and Figure 2.1 shows an example of this.

##### 2.1.2 Data Sets

We used two sources of data to extract DHS patterns from subjects. The first is the University of South Florida (USF) gait data set. The data set consists of a





Figure 2.1: *Example of an Activity Sequence as a Stack of Frames*

subject walking a wide circle in front of a camera. The camera is positioned at body height. The scenes are then varied based on ground type (grass or concrete), one of two shotypes, whether the subject is carrying a briefcase in their hand or not, either a left or right positioned camera is used, and date of data collection (May or November 2001). This results in 32 different possible combinations of conditions.

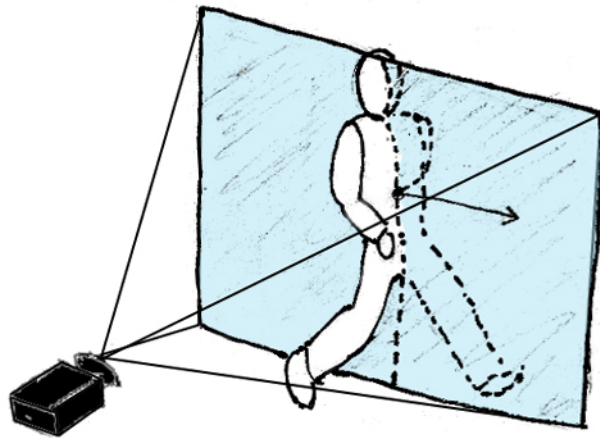


Figure 2.2: *Subject Walking Fronto-Parallel to Camera*: A subject is walking fronto-parallel to the camera such that his sagittal plane (blue) is parallel to the camera's image plane. This is also known as a profile view. A subject is said to be "walking 45° to the camera" if their sagittal plane makes a 45° angle with the image plane.

The second source consists of various activity sequences we captured at the University of Maryland, taken in various places both indoors and outdoors, taken during various times between 2004 and 2007. Subjects were asked to walk into frame, either walking normally or with a load on either an arm(s) or a leg, exit out the other side, then return through the frame, capturing both directions of travel as well as both sides of the subject. Subjects were directed to walk fronto-parallel to the camera, as shown in Figure 2.2, as well as to walk at an angle to the image plane, varying from thirty to forty-five degrees.

## 2.2 Human Body Characteristics and Motion Model

In order to analyze human body motion, we must first examine the general structure of the human body. When constructing a gait DHS pattern, we slice the silhouette image at specific heights on the body. Therefore, we are very interested in the dimensions of the body. Specifically, we wish to know the lengths of segments like the forearms and their relation to the overall height of the body.

Studies involving the measurements of such segments have been presented in medical literature [15] and presented in other gait signature analysis papers [18]. Figure 2.3 shows heights of the body joints as percentages of the overall height. We use this information to determine where to position the slices on the body.

We adopt a model for the gait cycle that can be divided into different states. Though the body goes through many discrete sub-articulations, including pronation and supination of the foot, we focus on three general states [7]:

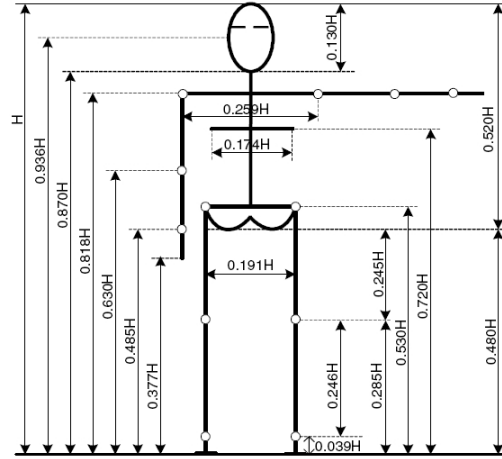


Figure 2.3: The Average relative heights of specific points on the body[15] .

1. **Double Support** - where both feet have contact with the ground while forward motion carries weight distribution from the toes of one foot to the heel of the other
2. **Left Swing** - weight is supported on the right leg as the left leg swings forward
3. **Double Support** (again)
4. **Right Swing** - like the left swing phase.

Figure 2.4 shows the transitions between these states. Though studies in [12] shows there is an inherent asymmetry in muscle propulsion, we assume articulation in the two swing and two double support phases mirror each other during a normal walk.

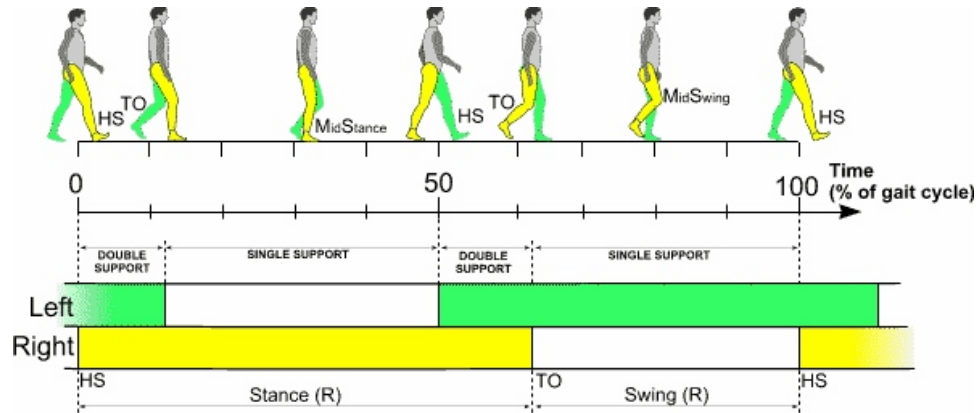


Figure 2.4: *Image of the Average Gait Cycle*: Includes duration of phases relative to the total gait cycle time [12] .

### 2.3 Human Body Model

A variety of models exist to represent the human shape, from ellipses to stick figures to non-rigid models. But all create a basis upon which to analyze movement. For this project, a model was constructed of rigid limbs connected by rotating joints, creating a **kinematic chain**. A single chain can consist of an upper arm, an elbow, and a forearm, or a thigh, a knee, and a lower leg. The human model is constructed by connecting these chains (arms, legs, and head) onto a central trunk through additional joints (shoulders, pelvis, neck). Ran *et al* [11] compose such a human model,  $H$ , consisting of rigid limbs  $L_1, L_2, \dots, L_k$  and joints  $J_0, J_2, \dots, J_k$ , connected to a static (in space) base body in the order  $J_0, L_1, J_1, L_2, \dots, J_k, L_k$ , and a system of coordinates  $\bar{X} = (X, Y, Z)$  with origin on the base body. Figure 2.3 illustrate this, where the open circles represent the joints. In this system, a point in the kinematic chain model  $(X, Y, Z, t)$  is related to it's previous position in time by a product of

the previous point and a series of transform matrices:

$$X_P(t) = T(t) \cdot X_P(t) = \prod_{i=0}^{k_0} T_i(t) \cdot X_P(t) \quad (2.1)$$

Where  $T_i(t)$  is the 4x4 transformation matrix of the  $J_i + 1$  articulation. This matrix can be further decomposed into a chain of multiplications of a position vector,  $D_i$  and 3x3 rotations  $R_{i-1,1}$ , for the  $i^{th}$  relative to the  $(i - 1)^{th}$  part [11]:

$$X_p(t + 1) = R_{0,1}(t) \cdot D_1(t) \cdot R_{1,2}(t) \cdot D_2(t) \cdot \dots \cdot R_{k-1,k}(t) \cdot X_p(t) \quad (2.2)$$

In the most simplest of models, and assuming an activity like "normal" walking (i.e. no limp, no load on any limb), the articulation parameters for a pair of limbs (one left, one right)  $\theta = \theta_l, \theta_r$ , are the same save for differences in phase[10]. We can assume a periodicity of  $T$ , giving the new constraints on motion:

$$\theta_l(t) = \theta_l(t + T) \quad (2.3)$$

$$\theta_r(t) = \theta_r(t + T) \quad (2.4)$$

$$\theta_l(t) = \theta_r(t - T/2) \quad (2.5)$$

This periodicity of limb movement and phase shift between pairs of limbs gives rise to a specific pattern, the **Double Helical Signature** (or DHS). Assume a given activity sequence,  $A$ , of a person walking normally (no load or gait impediments). If we slice  $A$  horizontally, at specific heights corresponding to movement, a "view-dependant twisted pattern" can be seen [10].

Ran *et al* [11] define an **Activity Signature** as "the set of shapes,  $S = S_1, S_2, \dots$  formed by slicing [the activity sequence  $A$ ] at all heights covering the whole human body during a complete stride," where each shape,  $S_i$  is a DHS.

By slicing the entire frame, it is also possible to track one or more persons, possibly through occlusions, as well as aid in background subtraction [11]. A vertically aligned activity signature of a single person is obtained by taking slices centered at points on the body. Like the kinematic chain model, the body can be taken as a fixed reference point in space. From the profile view, a vertical line can be extended through the center of the body, and limb activity would be similar to pendulums, swinging with steady tempo and deflection off the center line. We note that if the pair of legs (or arms) is modeled as a simple pair of pendulums with rigid swinging segments and joined at the pelvis, that DHS pattern at different heights vary only by amplitude.

In *Ran et al* [11], a theorem was presented to show the power of this observation, specifically the compactness of a DHS analysis, a summary of the proof of which is presented:

**Theorem 2.3.1** [11] *There exists a finite set of DHS as a compact representation for the hip-to-toe activity volume  $A$ .*

**Proof** Suppose an activity sequence  $G$  with articulation parameters  $\theta$ , and a limb/rigid structure,  $L$  vertically aligned with the y-axis. We can mark the structure at different heights,  $y_0, y_1, \dots, y_n$ , and each consecutive pair mark a sub-interval of the limb,  $L_i = [y_{i-1}, y_i)$ . For any two points in this interval, the DHS pattern is identical save some linear scaling along the x-axis, thus for any one point in the interval, defined at height  $\bar{Y}_0 \in [y_{i-1}, y_i)$ , all other DHS patterns can be reconstructed by a simple scaling. This holds true for all other intervals, and so a complete DHS can

be represented by a set  $\bar{Y}_0, \bar{Y}_1, \dots, \bar{Y}_{n-1}$ .

The theorem applies not just to hip-to-toe activity but shoulder-to-finger activity, or any other periodic activity that can be modeled thusly. As such, it is possible to accurately model the motion of the entire system by taking DHS patterns at the shared joint and at the ends of the pendulums. This encodes more information with less processing.

## 2.4 Activity Descriptions

There are many possible actions the human body can perform (sitting, bending over, etc.), but this project focuses on actions performed during the walking cycle. We first recorded a *regular gait*, one without any load on the subject or with any possible gait impediments (like a limp). This provides a gallery of "normal" images that we can train upon later as well as use a reference against pattern changes that occur when different activities are performed.

With regard to the arms, two activities were analyzed. We captured a subject with a *load in one hand*, be it a briefcase with some weight, or holding the strap of a courier bag. The weight was added to accurately represent real world conditions. Then the subject walked *with both arms occupied*: either carrying something close to their chest, or holding onto the straps of a back pack, or possibly tucked in their pockets. Finally, we had limited trials with a person *walking with weight on one leg*.

## 2.5 Description of Algorithm

With a definition of the activity sequence and a well defined model for human motion developed, we present an algorithm to decompose and classify a gait sequence:

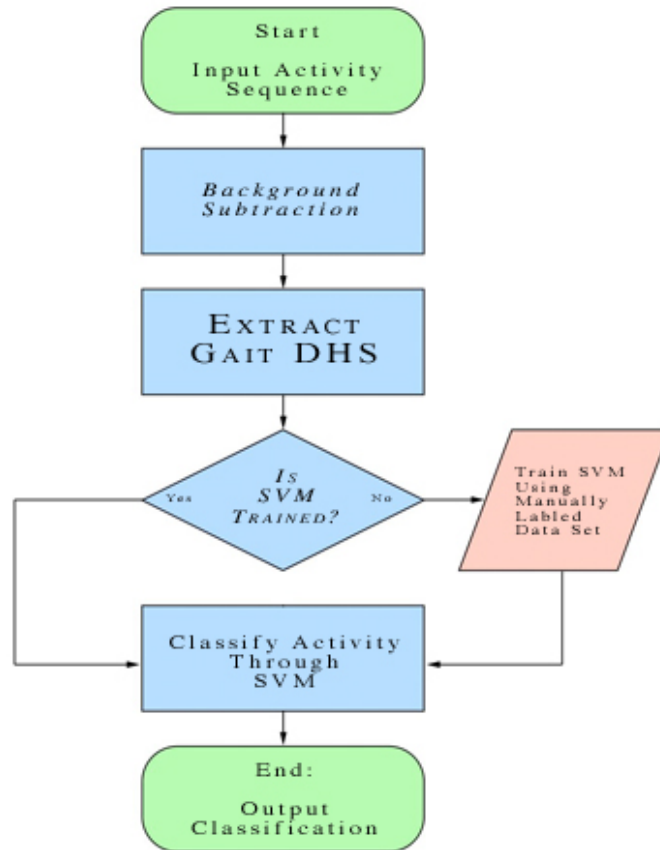


Figure 2.5: *Flow Chart of the Proposed Algorithm*: This chart shows how the method to be described in this thesis is laid out.



## Chapter 3

### GaitDHS Construction

#### 3.1 Background Subtraction

The DHS is constructed using the silhouette of the target human, taken from the foreground mask. To achieve background subtraction, each pixel (converted to greyscale) was modeled simply as a Gaussian distribution. A running accumulator calculates the mean,  $\mu$ , and variance,  $\sigma^2$ , of each greyscale pixel over a set of frames. This model is appropriate because the camera was always stationary, leaving the background static over the sequence of frames. Every pixel for each new frame was compared against the background model, and pixels with greyscale intensity values outside of the mean-plus-variance range are considered to be foreground pixels, marked as a white pixel in the foreground mask.

In some sequences, shadows changed pixel intensity enough to be considered foreground pixels, and if the subject were walking close to a wall, the shadow next to them could alter the foreground mask considerably. In this case, a mean background image was constructed, but in three-channel RGB instead of greyscale. The gain between the current frame and the background was calculated with the knowledge that unaffected background pixels have a gain that is (very close to) unity, while shadowed pixels have a gain close to unity, but at a distance that could be thresholded between background and shadowed background pixels shown in figure 3.1.

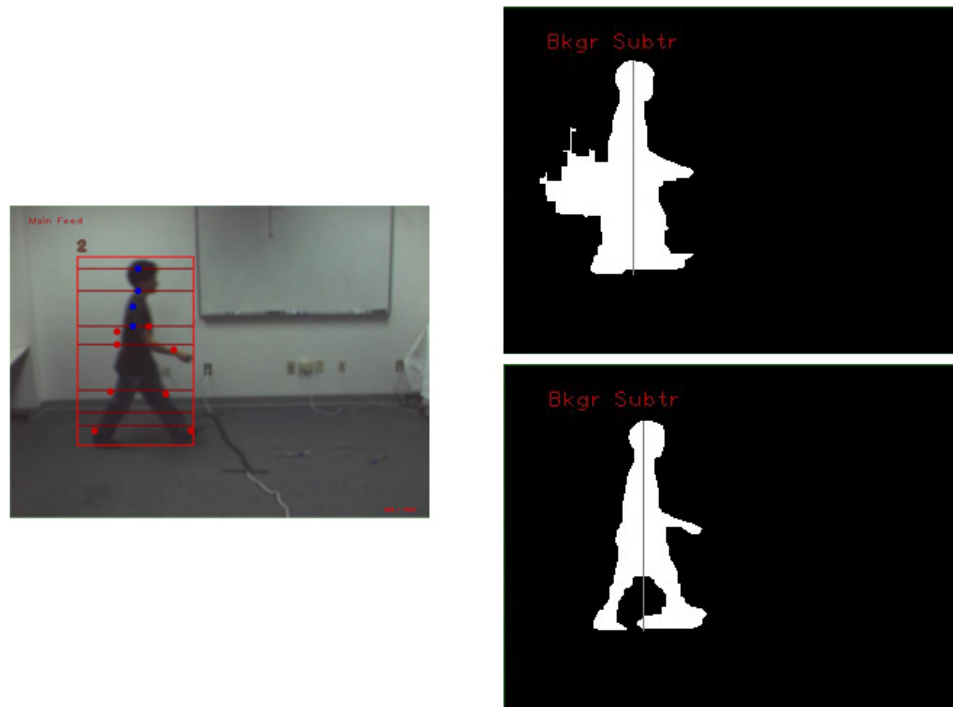


Figure 3.1: *Background Subtraction on the Same Frame*: The left image shows the original frame (with bounding box and lines drawn across slice heights (described later)). The top right image shows the monochromatic, Gaussian background subtraction scheme, which was sensitive to shadow or near-foreground-chroma areas, while the bottom right image shows background subtraction using chroma gain and off-unity thresholding. The grey line is the center line of the body, and not used in background subtraction.

## 3.2 Size Filtering and Target Tracking

Using the foreground mask created from one of the two methods mentioned above, morphological operations (dilation, then erosion) were performed on the pixel blobs to close any gaps between very near groups. This was done in the event that a neck was accidentally labeled as a background, resulting in the head and the body as two separate blobs. This is important because the resulting blobs were filtered by size, in order to eliminate small foreground pixel groups that may have been due to moving leaves on trees or other artifacts from the video encoder.

A minimum size bounding box was drawn around the foreground blob, and given a target identification number depending on if it was recognized as a new target or not. A target was considered previously tracked if its bounding box overlapped with that of a target from the previous frame. If it did, the new bounding box received the target identification number of the target in the previous frame, otherwise the target was given a new number.

## 3.3 Image slicing and DHS construction

The video sequences showed that a person walking at a natural pace takes approximately two seconds to complete two complete gait cycles. One complete cycle can be thought of as the time from when the right heel makes ground after swinging the leg forward to the time the left heel touches the ground. For each target, a buffer of the last foreground images was kept, as well as its age as described by the number of frames it had been in consecutively. The video sequences were

encoded at 30 frames per second, so when a target reached the age of 64 frames, which approximates two gait cycles, the gait DHS was constructed.

A single row of gait DHS consists of a binary, single-pixel-height row of data of the targets foreground mask, taken from the region of interest of the frame dictated by the bounding box. A slice is taken at height,  $y$ , such that for over all target height,  $H$ , and a percentage of that height,  $p \in [0, 1]$ :

$$y = p * H$$

such that  $p = 0 \Rightarrow y = 0 =$  top of the head

and  $p = 1 \Rightarrow y = H =$  bottom of the feet.

For each of the 64 buffer images, a slice was taken at the same  $p$  value, and these data rows were stacked vertically, creating a  $x-t$  spatio-temporal representation of the activity at that height.

Figure 2.3 shows the average heights of specific joints of the body, though with the  $p = 0$  reference point taken at the feet instead of the top of the head. Slices were taken at vales  $p=\{0.064, 0.182, 0.37, 0.47, 0.71, 0.83, 0.9\}$ , which correspond to the head, shoulders, elbows, wrists, knees, shins, and ankles.

The widths of bounding boxes of the 64 frames are adjusted so that the bounding box is centered horizontally with respect to the shoulder. The shoulder was chosen because it was always found along the vertical centerline of the body, unlike the head which moved in some sequences, and remained horizontally static, unlike the limbs. Of these 64 bounding boxes, the largest width was found and all other

bounding boxes are adjusted to this width. The subsequent DHS pattern is now vertically aligned and normalized to the maximum bounding box width. In the video sequences, the maximum width was often dictated by the largest stride taken. DHS images were then resized to a standard image size of 102 x 64. Shown are two such DHS patterns taken at the knee height. Figure 3.2 shows the process of background subtraction, slicing, and stacking to create a DHS pattern. From the DHS pattern, characteristics of gait, such as stride (step length) and cadence, are extracted as shown in figure 3.3.

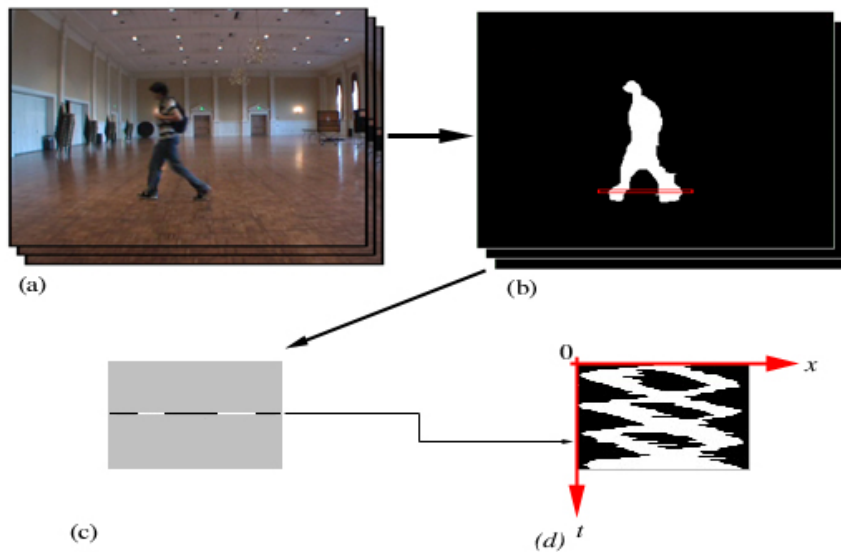


Figure 3.2: *Gait DHS construction process*: The activity sequence (a) and the silhouettes produced (b) are used to create slices at specific heights like the ankle (c). These slices are stacked over time to create the DHS band (d).

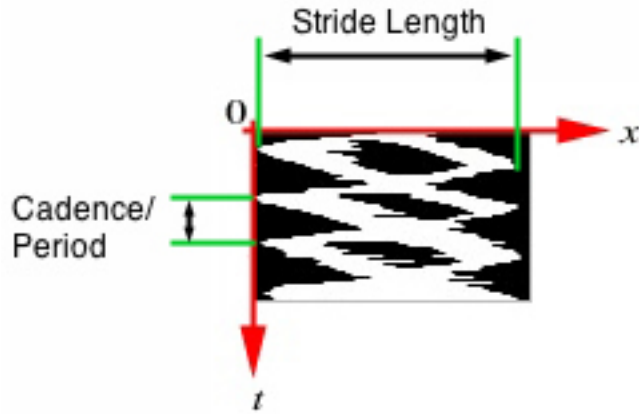


Figure 3.3: *Gait Parameters Captured by the DHS pattern.*

### 3.4 DHS patterns of Activities and Symmetries

Research in human activity classification shows that for a given method of measurement, there exists a measurable similarity in the way multiple persons perform the same action, and that there is a measurable difference between the two different actions. We can show that as different activities are performed by a person, there will be a change in the periodic/aperiodic behavior and symmetry of the DHS.

### 3.5 Symmetry: Frieze Groups

Occurring frequently in areas like art decoration and architecture, symmetry underlies visual elements that show recurrence. Studied in Geometric Group Theory[14], symmetric patterns in the two-dimensions belong to a class of isometries on the 2D plane. It can be shown that every isometry (rotation, translation, glide, or reflection) is the product of at most three reflections [3]. Groups of these isometries are called Frieze groups. These groups can be enumerated by seven forms

(the names in parentheses are the Orbifold notation):

1. Translation only (hop)
2. Glide and Translation (step)
3. Reflection across the horizontal axis and Translation and Glide (jump)
4. Translation and Reflection across the vertical axis (sidle)
5. Translation and  $180^\circ$  rotation (spinning hop)
6. Reflections across the vertical axis, Translation, Glide, and Rotation (spinning sidle)
7. Translation, Glide, Reflection (in both axes) and Rotations (spinning jump).

Examples of these groups are show in figure 3.4. *Liu et al* using frieze patterns to Frieze group classification to determine viewing camera angle and align gait signatures using dynamic time warping [6]. We show that DHS patterns display these types of symmetries.

### 3.6 DHS patterns

We now present common DHS patterns for the tested activities. It should be noted that the images of the subjects have been resized (either stretched or shrunk) to be uniform and visible. However, the DHS patterns are as they are originally outputted, so the sizes of the subjects will not reflect the silhouette DHS bands.

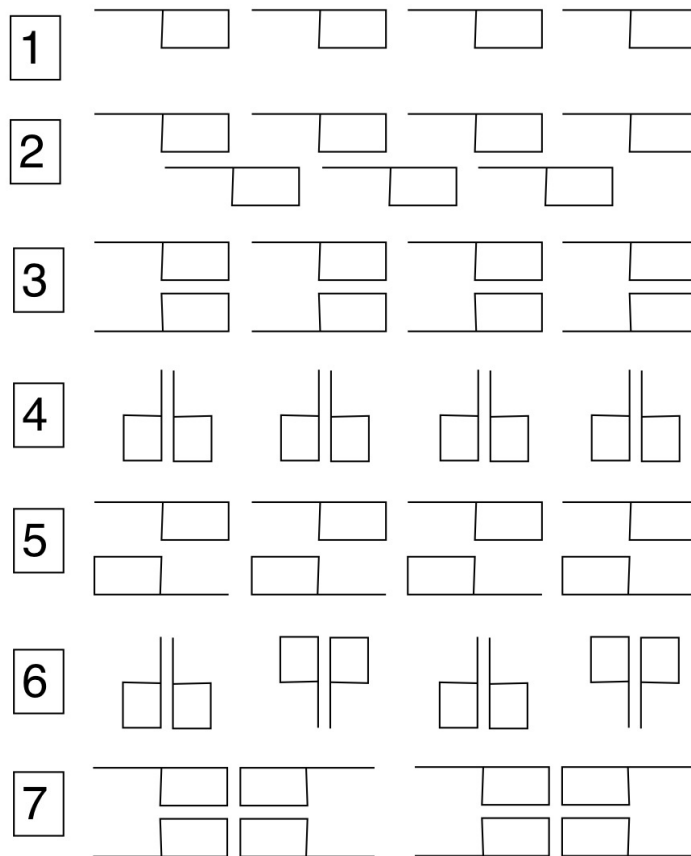


Figure 3.4: *Frieze Groups*: Examples of the seven Frieze symmetry groups: (1) Hop, (2) Step, (3) Jump, (4) Sidle, (5) Spinning Hop, (6) Spinning Sidle, (7) Spinning Jump



Figure 3.5 shows the DHS patterns of slices taken at the elbow and wrist heights among the three activity classes, taken for a subject walking fronto-parallel to the camera. The subject in the first picture is walking naturally, without anything in his/her hands. The periodic bumps on the left side of the slice are caused by the alternating swing of the subject's arms (left arm, right arm, repeat) extending past her body. Not present in many DHS patterns taken for a subject walking fronto-parallel are bumps on the other side, like pictured. These appear if the subject has noticeable back swing to their arms. From this angle, the subjects body occluded the backswing of her right arm.

In the second picture, the subject is holding a black briefcase in his left hand. The weight of the brief case is such that stability of motion is maximized and energy input is minimized so long as the weight is kept at the lowest part of the swing i.e. the arm is kept down beside the body. The resulting bumps come from the swing of the right hand.

The third image shows the subject holding onto the straps of a (negligibly weighted) backpack, preventing either arms from swinging. The subjects displays no bumps like the first two activities. Also, the DHS pattern at the elbow slice is thicker than at the wrist, unlike the previous two. This is because taking DHS slices at elbow height includes the backpack, but the wrist height is close to the waist and lower body, below the backpack.

Figure 3.6 shows the DHS patterns for slices taken at knee and ankle height among the two loading conditions for the leg, taken for subjects walking fronto-

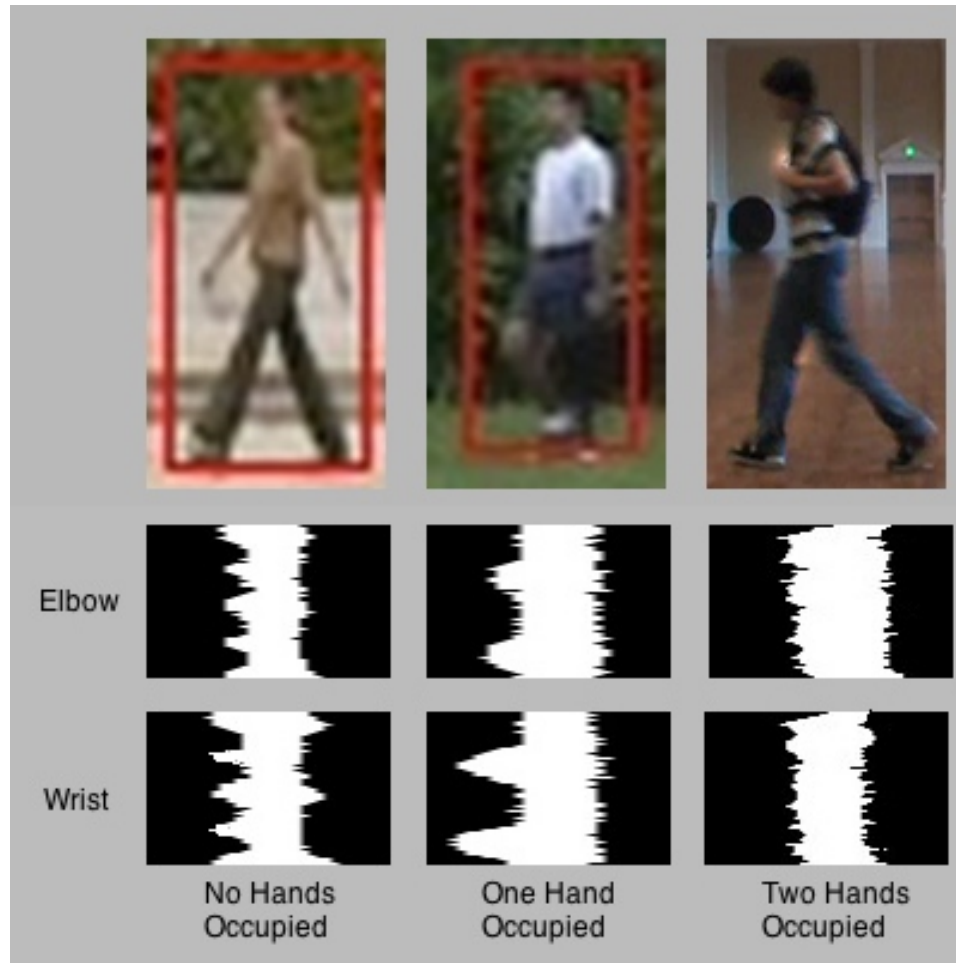


Figure 3.5: *DHS patterns for Arm Activities*: Common DHS patterns for the three types of arm loading conditions are shown for slices taken at approximate elbow and wrist height. The red box is the bounding box drawn by the DHS construction program.

parallel to the camera. For the first loading condition, slices showed a consistent cadence and stride length. For the second loading condition, stride length changed. Swing from the leg without load created the longer strides. This is most noticeable in the ankle-height slices.

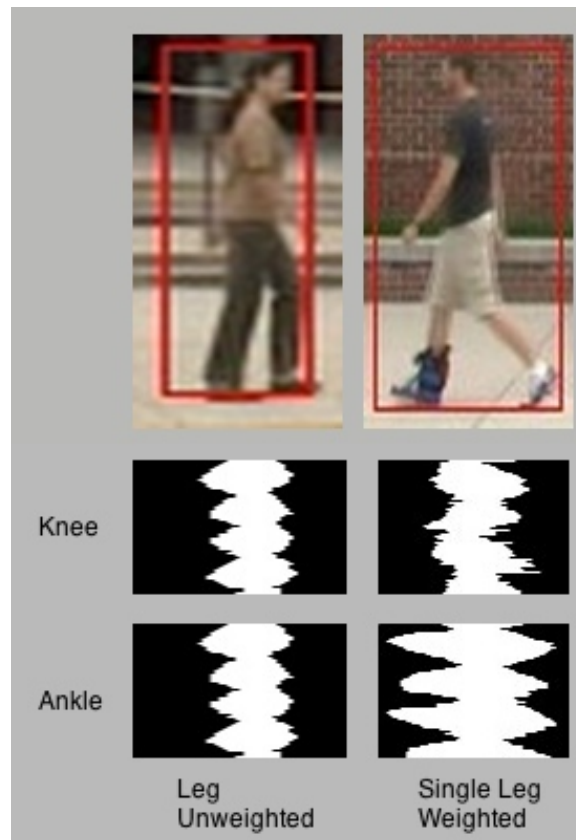


Figure 3.6: *DHS patterns for Leg Activities*: Common DHS patterns for the two types of leg loading conditions. In the second image, the subject's right leg is being loaded by three five pound leg weights.

### 3.6.1 Symmetry and Periodicity of Arm DHS patterns

The leg slices show mainly a translational symmetry, but keep a constant period. The different arm loading conditions, however, can change in periodicity and/or symmetry, depending on the activity and view angle. Figure 3.7 shows how differences in load change the periodicity of the subject. We recall that the arms swing out of phase with the legs. If we define the period of leg swing as  $T$ , then arm swing periodicity becomes a multiple of  $T$ .

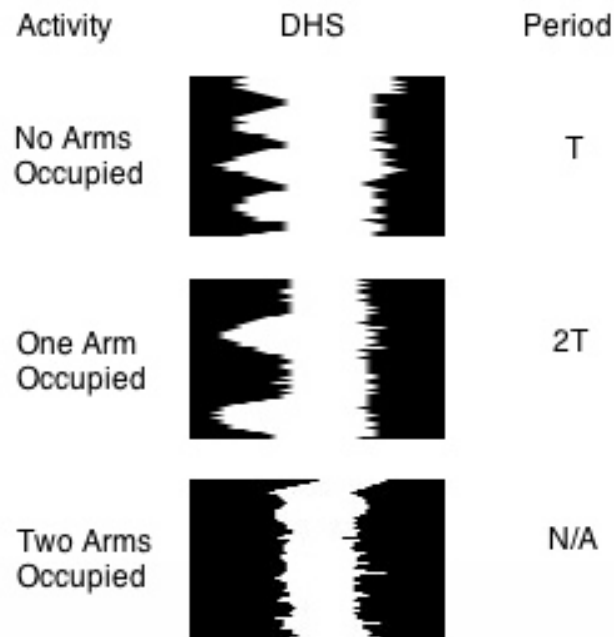


Figure 3.7: *Periodicity and Symmetry Across Arm Activities*: Under the same view (fronto-parallel), all arm activities display translational symmetry (type 1 Frieze group), except for the "Two Arms Occupied" case. The difference comes in periodicity. Slices are taken at wrist height.

## Chapter 4

### Gait SVM Construciton

#### 4.1 SVM derivation

Support vector machines are increasingly used for pattern classification, primarily in the presence of labeled exemplars. Its creation is attributed to Vladimir Vapnik who, with Alexy Chervonenkis, developed V-C theory for computational learning[1]. We have shown that DHS patterns characterizing different activities are distinguishable. We use this fact to train an SVM for classifying DHS patterns. In the following, we introduce the basic linear kernel SVM for a two class, linear separable set of data, then show how to handle the more realistic non-separable data set. Finally we generalize to the non-linear kernel SVM, which maps the input vectors to a high-dimension feature space before estimating the minimal enclosing hypersphere. This derivation can also be found in [4] We also address the case of the n-class classifier ( $n \geq 2$ ), which is what all our SVMs will be.

##### 4.1.1 Linearly Separable Dataset

For the a set of N points,  $x_i$ , belonging to one of two classes, we can assign a label,  $y_i$  to each point, indicated by -1 or 1. The dataset can be written as:

$$\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_N, y_N)\} \quad (4.1)$$

We construct a method to predict the sign of  $y_i$ . In the linearly separable case, there exists a combination  $\{\mathbf{w}, b\}$ , which dictate a hyperplane, such that:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 0 \quad (4.2)$$

for every input point. In the geometric representation, all  $\mathbf{x}_i$ 's of one sign are strictly on one side of the hyperplane. The sets of points can be surrounded by convex hulls. There can exist a family of separating hyperplanes between the hulls, but we choose the hyperplane that is maximally furthest from both.

This is done by determining the point on each hull such that the length of the line connecting the two points is the minimum. The hyperplane is laid perpendicular to the connecting line at the midpoint, making it far away as possible from both. Figure 4.1 shows an example of this.

It is possible to scale  $\mathbf{w}$  and  $b$  by a positive number and still satisfy (4.2). We can choose these values such that for any  $\mathbf{x}_i$ :

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq D \quad (4.3)$$

where  $D$  is a constant that can be chosen as  $D = 1$ . Now, we can select points  $x_j$  and  $x_k$ , one each from the different classes that achieve equality when their labels are -1 and 1, respectively, placing them on opposite sides of the hyperplane. These points are, among their clusters of like-class data points, the closest to the hyperplane (note: there might be several such points from either class). Plugging in these points into (4.3) and adding gives  $\mathbf{w} \cdot (\mathbf{x}_j - \mathbf{x}_k) = 2$ .

We wish to maximize the distance of the points (and so the hulls) to the hyperplane such that a new feature vector maybe properly classified, even if falls

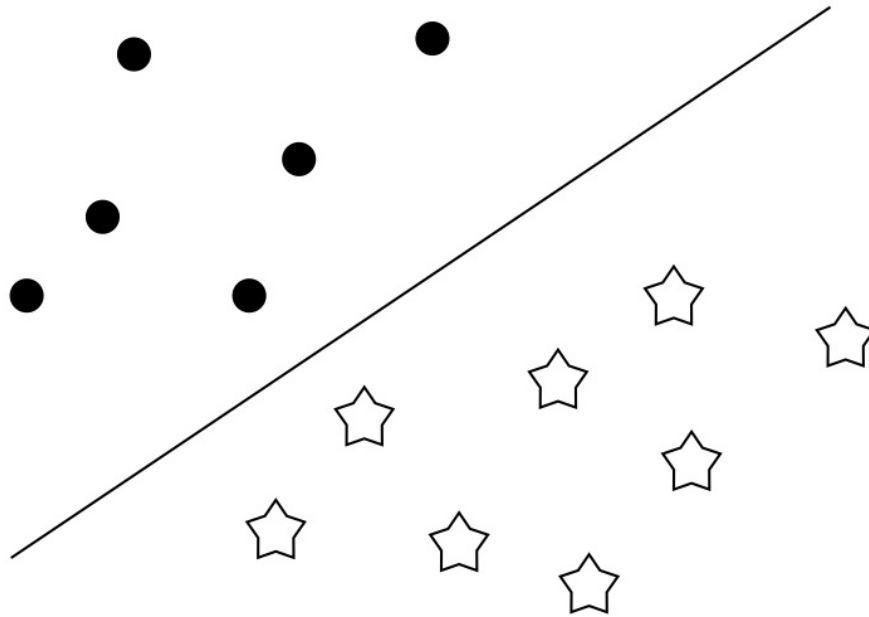


Figure 4.1: *Linear Separable Case*: black dots are one class, white stars are another. The input space is separable by a simple hyperplane or in the 2D realm, a boundary line.

near but outside the convex hull. Using the sum we just obtained, we get:

$$\begin{aligned}
Dist(x_j, hyperplane) + Dist(x_k, hyperplane) &= \left( \frac{\mathbf{w}}{|\mathbf{w}|} \cdot \mathbf{x}_j + \frac{b}{|\mathbf{w}|} \right) - \left( \frac{\mathbf{w}}{|\mathbf{w}|} \cdot \mathbf{x}_k + \frac{b}{|\mathbf{w}|} \right) \\
&= \frac{\mathbf{w}}{|\mathbf{w}|} \cdot (\mathbf{x}_j - \mathbf{x}_k) \\
&= \frac{2}{|\mathbf{w}|}
\end{aligned} \tag{4.4}$$

In other words, to maximize the distances from the hyperplane is to minimize  $\frac{|\mathbf{w}|}{2}$ , subject to (4.3). These two constraints establish a problem that can be optimized using Lagrange multipliers:

$$L(\mathbf{w}, b) = \frac{|\mathbf{w}|}{2} - \sum_{i=1}^N \alpha_i (y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1) \tag{4.5}$$

The optimal solution must satisfy the Karush-Kuhn-Tucker (KKT) conditions, where taking the partial derivatives of L with respect to  $\mathbf{w}$  and  $b$  gives:

$$\mathbf{w} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i \tag{4.6}$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \tag{4.7}$$

Substituting back into (4.5) and manipulating, we obtain the *dual* problem of maximization:

$$\begin{aligned}
&\text{maximize: } \sum_{i=0}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i (y_i y_j \mathbf{x}_i \cdot \mathbf{x}_j) \alpha_j \\
&\text{subject to: } \alpha_i \geq 0 \\
&\text{and } \sum_{i=1}^N \alpha_i y_i = 0
\end{aligned} \tag{4.8}$$



And classification becomes:

$$\begin{aligned} \text{class}(x) &= \text{sign}(\mathbf{w} \cdot \mathbf{x} + b) \\ &= \text{sign}\left(\sum_{i,j=1}^N \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x}_j + b\right) \end{aligned} \tag{4.9}$$

#### 4.1.2 Non-Linearly Separable Data Set

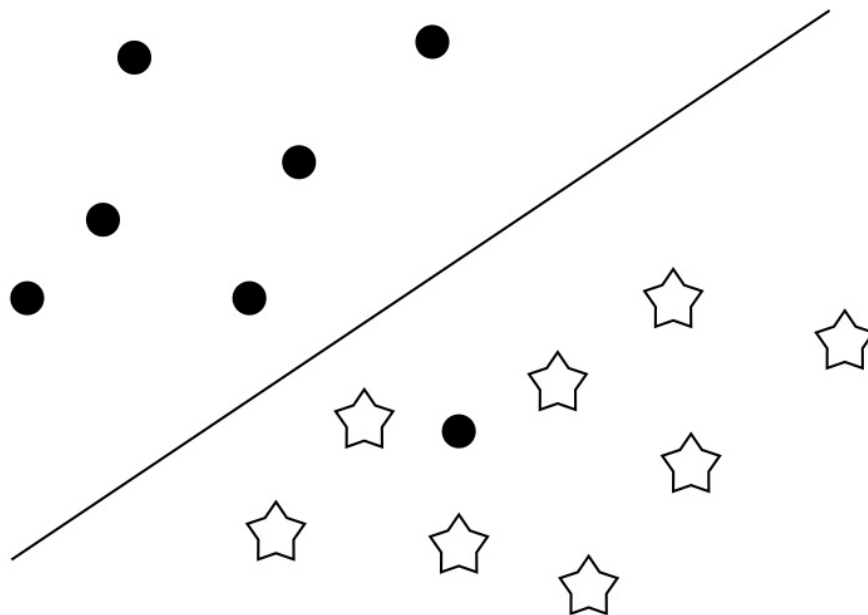


Figure 4.2: *Non-Linear Separable Case*: Outliers may exist, but not addressing them could alter the hyperplane. Adding an error margin relaxes the constraints on the boundary and maintains a good separating hyperplane.

In reality, the data is rarely cleanly separable in two events: (1) the case of outliers for each class and (2) the separation boundary is modeled better by a hypersphere or some shape that is not a straight hyperplane. Figure 4.2 shows an example of an outlier in the input space. In the first case, one can introduce a slack

variable,  $\nu$ , which relaxes the constraint and gives:

$$y_i(\mathbf{w} \cdot \mathbf{x}_1 + b) \geq 1 + \nu \quad (4.10)$$

The minimization problem is altered now as:

$$\begin{aligned} \text{minimize: } & \frac{|\mathbf{w}|}{2} - C \sum_{i=1}^N \nu_i \\ \text{subject to: } & y_i(\mathbf{w} \cdot \mathbf{x}_1 + b) \geq 1 + \nu_i \\ & \text{and } \nu_i \geq 0 \end{aligned} \quad (4.11)$$

Where  $C$  is the cost weight of constraint violations. The dual problem of maximization holds the same form as before, but with the altered constraint that  $C \geq \alpha_i \geq 0$ .

### 4.1.3 Non-Linear Kernel SVM Method

Toward the second problem of a more complex boundary shape, we use a non-linear kernel to map the input data to a higher dimensional feature space. In this space, a separating hyperplane is easier to construct, which translates back into a more complex geometric form in the original feature space, like the separating boundary in figure 4.3.

We define a mapping for  $\mathbf{x}'_i = \phi(\mathbf{x}_i)$ . The optimization problem becomes:

$$\begin{aligned} \text{maximize: } & \sum_{i=0}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i (y_i y_j \mathbf{x}'_i \cdot \mathbf{x}'_j) \alpha_j \\ & = \sum_{i=0}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i (y_i y_j \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)) \alpha_j \end{aligned} \quad (4.12)$$

subject to the same constraints on  $\alpha_i$  and  $y_i$  as before. Likewise, the classifier

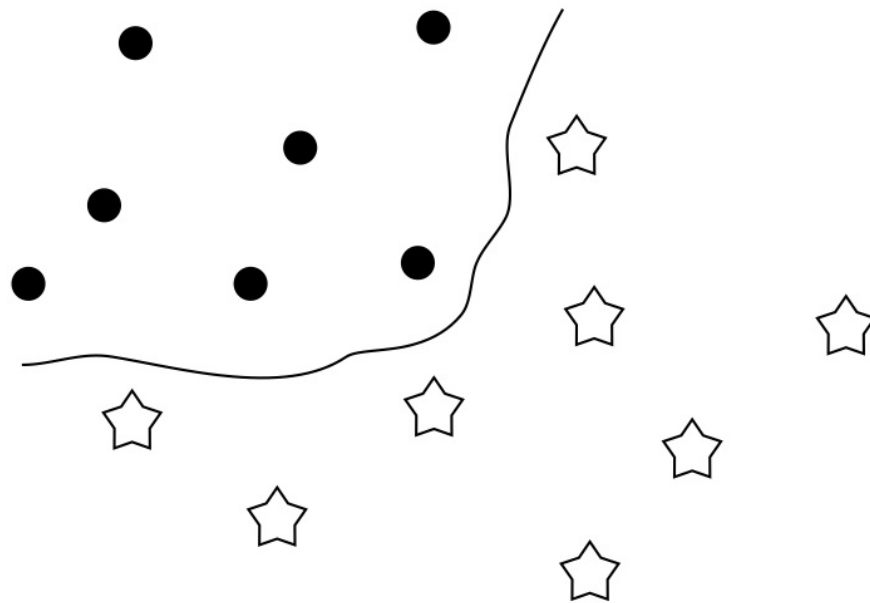


Figure 4.3: *Separation With a Complex Boundary*: The input space can be mapped to a higher dimension feature space where a simple separating hyperplane exists. Mapping back to the original input space creates a more complex and accurate boundary.

becomes:

$$\begin{aligned}
 class(x) &= sign\left(\sum_{i,j=1}^N \alpha_i y_i \mathbf{x}'_i \cdot \mathbf{x}'_j + b\right) \\
 &= sign\left(\sum_{i,j=1}^N \alpha_i y_i \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) + b\right)
 \end{aligned} \tag{4.13}$$

Here, we apply the "kernel trick", which basically states that we can use a kernel  $K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{y})$  if  $\phi(\mathbf{x}) \cdot \phi(\mathbf{y})$  is positive. Kernel functions are "a computationally efficient technique to map the data into the induced feature space..." [2]. The kernel function must, however, satisfy Mercer's Theorem [13], which states that "in order for a kernel function"  $K(\mathbf{x}, \mathbf{y})$  to take on an eigenfunction expansion:

$$K(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle = \sum_{k=1}^{\infty} b_k \phi_k(\mathbf{x}) \phi_k(\mathbf{y})$$

must be positive semi-definite:

$$\int K(\mathbf{x}, \mathbf{y}) g(\mathbf{x}) g(\mathbf{y}) d\mathbf{x} d\mathbf{y} \geq 0, \forall g(\cdot) \in L_2$$

We can now use a common positive semi-definite kernel to provide a more robust SVM functions. The maximization problem (4.12) and classifier function (4.13) become:

$$\begin{aligned}
 \text{maximize: } & \sum_{i=0}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i (y_i y_j \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)) \\
 class(x) &= sign\left(\sum_{i,j=1}^N \alpha_i y_i \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j) + b\right)
 \end{aligned}$$

In this project, we used the Radial Basis Function (RBF) as the kernel:

$$K(\mathbf{x}, \mathbf{y}) = exp(-\gamma \|\mathbf{x} - \mathbf{y}\|)$$

The RBF is used in applications like artificial neural networks. In [2], the RBF kernel is used to detect anomalies in hyper-spectral imagery, while [16] uses

the RBF kernel along with dynamic time-warping to identify people from gait. The parameter  $\gamma$  controls the strictness of the classifier on input data: the larger the  $\gamma$  value, the more likely the SVM rejects a true value as not belonging to the class.

To increase the positive detection rate, we iterate over the  $\gamma$  and cost,  $C$ , parameters and use cross validation to find an approximate good value for each.

#### 4.1.4 n-class classifiers

The discussion thus far has been primarily about the binary classifier, but we need a classifier that can handle  $n \geq 2$  classes. The two common strategies for dealing with the multi-class problem starts with creating a series of binary classifiers (i) one-versus-all, where the output function gives out a strong value for vectors that belong to the class and small values for vectors that do not; the classifier with the highest output is automatically declared the winner, and (ii) one-versus-one, where inputs are compared against pairs of classes, and a voting tally is kept amongst the competitions; the class with the most votes wins.

## 4.2 Training on the DHS patterns

In this thesis, heavy use was made of the OpenCV library, an open source C library for computer vision. The library includes a machine learning class that handles n-class SVMs using the one-versus-one strategy. The steps to train are:

1. Extract DHS pattern
2. Label the data (e.g. "1" for two hands free, "2" for one hand occupied, "3"

for two hands occupied)

3. Crop DHS band so that the pattern fills a standard size (in our case, 102 x 64 pixels).
4. Pass DHS's and Labels to SVM training function

The reflection across the vertical axis of each slice is also trained to address the pattern differences between a subject walking left-to-right and right-to-left.

## Chapter 5

### Results

#### 5.1 DHS Under Changes in Walking Angle

While symmetry stays constant and periodicity change across different activities for the same view, both can change for different views. Figure 5.1 shows the three arm activities for the front view (at wrist height), as well as an example of the DHS of the same activity for a subject walking at an angle to the camera. The DHS pattern of the "Two Hands Occupied" case stays consistent. The "One Hand Occupied" case maintains the same  $2T$  periodicity and type 1 Frieze group.

The only thing that changes are arm stride and which side of the DHS the bumps occur. Arm stride, or the amplitude of the bumps, depends on the angle to the camera the subject is walking, where walking fronto-parallel shows the biggest arm stride, and walking straight on to the camera shows nearly no movement, like the two hands occupied case. If the person is walking away from the camera, the back swing of the arm may be captured. Shown in 5.1, the body occludes the forward swing of the left arm and the back swing of the right arm.

The loaded limb conditions create similar patterns at an angle to their fronto-parallel counterpart, but the DHS pattern for "No Hands Occupied" shows a greater change at an angle. Periodicity goes from  $T$  to  $2T$ , like the one arm occupied case, but symmetry falls into the type 3 Frieze group ("Jump"), where the bumps are

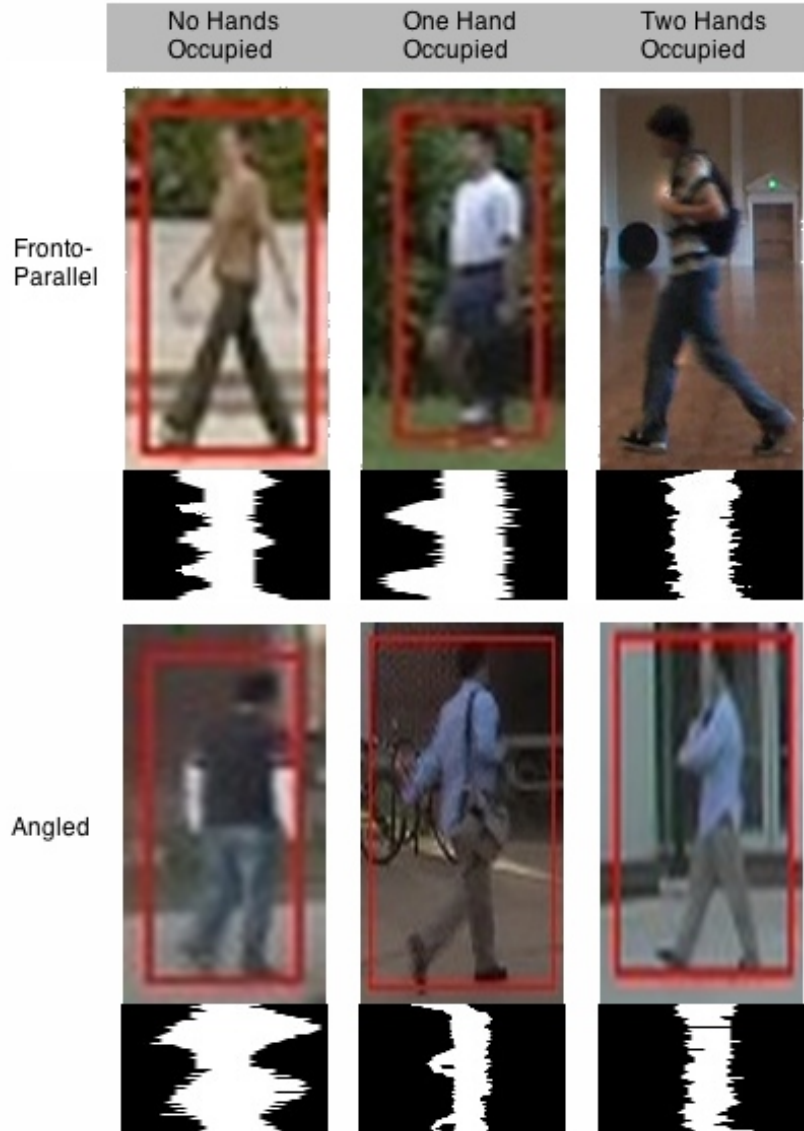


Figure 5.1: *DHS patterns for Arm Activities at an Angle*: The first row shows the arm DHS patterns as in figure 3.5, while the second row shows the DHS patterns of the same activity, but for the subject walking at  $45^\circ$ ,  $45^\circ$ , and  $40^\circ$  from left to right. Similar patterns exist for loaded arm conditions, but a different pattern exists for the "No Hands Occupied" case.



reflected across the vertical line of the DHS pattern. In this case the body occludes the half of the swing cycle, while pronouncing backswing during the other half of the cycle. In figure 5.1, the body occludes the forward swing of the left arm and the backswing of the right arm. Changes in DHS patterns under different viewing angles suggests the ability to predict camera viewpoint from the change in Frieze group symmetry, as was shown in [6].

## 5.2 SVM results

We now present the outcome of using SVMs to classify the different activities based on their DHS patterns. An SVM was created for each of the elbow, wrist, knee, and ankle slices. Tables 5.2 through 5.5 show the (non-symmetric) confusion matrices for each SVM. The output performance of the SVMs are listed horizontally across the top, while the true classification is listed along the leftmost vertical column.

### 5.2.1 Determining Parameters through Test Set Validation

Finding good values for  $\gamma$  and  $C$  are important to constructing a well functioning SVM. The process involves training the SVM on the training set with a selected pair of values, then evaluating performance over a smaller set of labeled test data. A grid search was performed on the  $\gamma$ 's, iterating first over magnitudes of ten, then over finer steps. For each  $\gamma$ , the  $C$  parameter was iterated over to find an optimal value.

The OpenCV library, in addition to constructing an SVM and using it to classify, also has a cross validation function. However, using it often returned parameter values that gave poor performance (one class may end up with zero positive detections). It was useful however in giving a base starting gamma around which to search. Figures 5.2 and 5.3 show the results of such searches. Table 5.1 shows the final selected parameters.

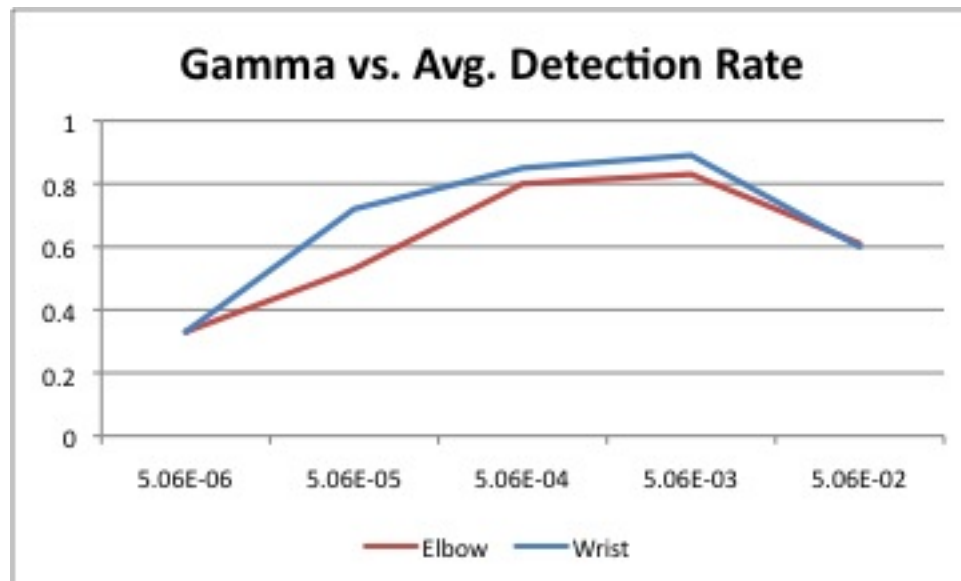


Figure 5.2: *Average Detection Rate Across a Range of Gammas for Arm Height Slices*

## 5.2.2 Arm Performances

Tables 5.2 and 5.3 show the SVM confusion matrices for DHS patterns at elbow and wrist height respectively. We see that at elbow height, the trained SVM

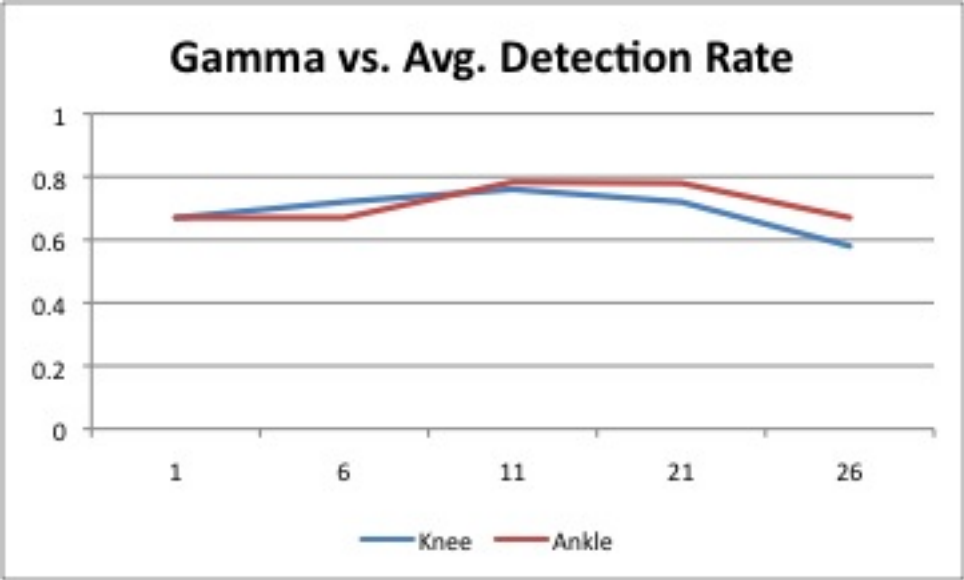


Figure 5.3: Average Detection Rate Across a Range of Gammas for Leg Height Slices

Table 5.1: Selected Parameters

Slice	$\gamma$	C
Elbow	$5.0625 \times 10^{-4}$	2.6
Wrist	$5.0625 \times 10^{-4}$	6
Knee	11	8
Ankle	11	8

had an average true positive detection rate of 75.87%, while an average rate of 80.53% was achieved at wrist height slices. While the average detection rate only differs by 4.66%, it hints that using the lower part of the arm acting as a better indicator of arm-loading conditions.

This is most likely due to the fact that the ends of the kinetic chains have the greatest swing length. This creates a more pronounced DHS pattern. We also note that the performance rates shown include subjects walking both fronto-parallel and at an angle to the camera, and that at wrist height, the SVM achieved 90.2% recognition for the "Two Arms Free" case, despite the different DHS patterns at an angle. This is due to proper training of the SVM, where so long as the training set covers a good spread of conditions, performance is still good.

Table 5.2: Confusion Matrix for Slices Taken at **Elbow Height**

Classified as	2 Arms Free	1 Arm Occupied	Two Arms Occu.	Number Tested
Two Arms Free	<b>75.5%</b>	24.5%	0%	129
1 Arm Occupied	14.1%	<b>83.5%</b>	2.4%	131
2 Arms Occu.	2.9%	28.4%	<b>68.6%</b>	69

### 5.2.3 Leg Performances

Tables 5.4 and 5.5 show the SVM confusion matrices for DHS patterns at knee and ankle heights respectively. We see that at knee height, the trained SVM had an

Table 5.3: Confusion Matrix for Slices Taken at **Wrist Height**

Classified as	2 Arms Free	1 Arm Occupied	Two Arms Occu.	Number Tested
Two Arms Free	<b>90.2%</b>	8.2%	1.6%	129
1 Arm Occupied	15.3%	<b>80.0%</b>	4.7%	131
2 Arms Occu.	2.0%	8.6%	<b>71.4%</b>	69

average true positive detection rate of 76%, while ankle height saw an average rate of 78%. The difference in average performance is even smaller than for the arms, so more testing is needed to suggest better performance of one slice height over the other.

We note however that detection rates for the leg weighted case are considerably lower than any other activity at any other height. The patterns between un-weighted and weighted leg are very similar. Changes in stride were not visually discernible in the DHS pattern when weights less than 15 pounds were applied. We suggest exploring other conditions for leg activity alteration: walking without a shoe on (to force a limp), or perhaps even walking with a cane or crutches.

Table 5.4: Confusion Matrix for Slices Taken at **Knee Height**

Classified as	Unweighted	Weighted	Number Tested
Unweighted	<b>93%</b>	7%	75
Weighted	41%	<b>59%</b>	75

Table 5.5: Confusion Matrix for Slices Taken at **Ankle Height**

Classified as	Unweighted	Weighted	Number Tested
Unweighted	<b>94.6%</b>	5.4%	75
Weighted	38.7%	<b>61.3%</b>	75

## Chapter 6

### Conclusion and Future Work

In this thesis, we presented a method to autonomously differentiate between human gait patterns. The classification problem can be divided into two sections: feature extraction and classification based on features. We explored gait DHS as a way to encode multiple features of gait, including symmetry and periodic characteristics. Then we used SVMs, for classifying DHS patterns.

Through test set validation, we found good base  $\gamma$  and  $C$  parameters for a RBF kernel-based SVM. We have shown through our results that the SVM is able to distinguish fairly well between different DHS patterns of arms under different loads.

We have developed a method to classify gait under different loaded limb conditions that shows good potential, but there are still avenues to explore and modifications to be tested.

#### 6.1 Modification

Because the DHS pattern is built from the silhouette, good background subtraction is very important. Modeling background pixels with a Normal distribution has worked well with this method, but there are events where if the skin color or clothing color of the arm, for example, is too similar to the background, it may be

cut off in the silhouette, and completely change the DHS pattern. Likewise, shadows can dramatically alter the silhouette, and while we have used a simple solution, there are better methods to detect and remove shadows from the foreground mask. Creating a silhouette that can capture the periodic deformation of the body is very important.

## 6.2 Extension

In addition to the gait activities described here, additional actions should be explored. This can include different loading conditions on the limbs or on the trunk of the body. We have presented classification results for a limited set of leg-loaded sequences, and have seen that regarding weights, most adult males do not display a significant change in DHS pattern unless loaded by approximately 15 pounds on one leg. Additional leg activities, like walking with a limp would be good to explore

The DHS pattern captures the x-displacement over time, and extending the DHS band beyond the 64 frames/two seconds creates a time line of motion. This provides the opportunity not only to focus on discrete activity sequences, but possibly to characterize the transitions between different activities (changing directions, picking up objects, etc.).

Also, periodicity in the leg slices could be used to classify a target as human, inanimate (like a rolling ball or a car), or perhaps quadrupedal (where two DHS strands would be created per band).

In this thesis, we have presented a nascent method of gait activity classifica-



tion, and believe that it performs well and can be honed and extended into a well functioning system.

## Bibliography

- [1] November 2006.
- [2] Amit Banerjee, Philippe Burlina, and Chris Diehl. A support vector method for anomaly detection in hyperspectral imagery. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(8):2282 – 2291, August 2006.
- [3] Judith N. Cederberg. *A Course in Modern Geometries*. Springer, 2001.
- [4] David Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, Upper Saddle River, NJ 07458, January 2003.
- [5] P. Kuchi, R. R. V. Hiremagalur, H. Huang, M. Carhart, J. He, and S. Panchanathan. DRAG: a database for recognition and analysis of gait. In J. R. Smith, S. Panchanathan, and T. Zhang, editors, *Internet Multimedia Management Systems IV. Edited by Smith, John R.; Panchanathan, Sethuraman; Zhang, Tong. Proceedings of the SPIE, Volume 5242, pp. 115-124 (2003).*, volume 5242 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 115–124, November 2003.
- [6] Yanxi Liu, Robert Collins, and Yanghai Tsin. Gait sequence analysis using frieze patterns. In *ECCV*, 2002.
- [7] Medi Tech Healthcare Ltd. Biomechanics, November 2008.

- [8] Stuart D. Mowbray and Mark S. Nixon. Automatic gait recognition via fourier descriptors of deformable objects. In Josef Kittler and Mark S. Nixon, editors, *Audio Visual Biometric Person Authentication*, pages 566–573. Springer, 2003.
- [9] Kaustav Nandy. Identification of human walking patterns using 3d-dynamic modeling. Master’s thesis, University of Maryland at College Park, 2006.
- [10] Yang Ran. *Human motion analysis in space and time: theory and experiment*. PhD thesis, University of Maryland at College Park, College Park, MD, USA, 2006. Adviser-Rama Chellappa and Adviser-Qinfen Zheng.
- [11] Yang Ran, Rama Chellappa, and Qinfen Zheng. Finding gait in space and time. In *ICPR ’06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 586–589, Washington, DC, USA, 2006. IEEE Computer Society.
- [12] Heydar Sdeghi, Paul Allard, and Morris Duhaime. Functional gait assymetry in able-bodied subjects. *Human Movement Science*, 16(2-3):243–258, April 1997.
- [13] V. N. Wapinik. *Statistical Learning Theory*. Wiley, 1998.
- [14] H. Weyl. *Symmetry*. Princeton University Press, 1952.
- [15] D. A. Winter. *The Biomechanics and Motor Control of Human Movement*. John Wiley and Sons, 2 edition, 1990.

- [16] Bo Ye and Yu-Mei Wen. Gait recognition based on dwt and svm. *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, 3:1382–1387, November 2007.
- [17] Jang-Hee Yoo, Doosung Hwang, and Mark S. Nixon. Gender classificaiton in human gait using support vector machine. In *Advanced concepts for intelligent vision systems*, volume 3708, pages 138–145, Antwerp, Belgium, September 2005.
- [18] Jang-Hee Yoo, Mark S. Nixon, and Chris J. Harris. Extracting human gait signatures by body segment properties. In *Proc. IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 35–39. IEEE Computer Society, April 2002.