Gait sequence modelling and estimation

using Hidden Markov Models

Presented by:
Kouame Hermann Kouassi

Prepared for:
Fred Nicolls
Dept. of Electrical and Electronics Engineering
University of Cape Town

Submitted to the Department of Electrical Engineering at the University of Cape Town in partial fulfilment of the academic requirements for a Bachelor of Science degree in Electrical and Computer Engineering

November 1, 2017

Declaration

---

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the IEEE convention for citation and referencing. Each contribution to, and quotation in, this report from the work(s) of other people has been attributed, and has been cited and referenced.
3. This report is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as their own work or part thereof.

Signature:............................
Kouame H. Kouassi

Date:...............................

Acknowledgments

Abstract

- Open the **Project Report Template.tex** file and carefully follow the comments (starting with %).
- Process the file with **pdflatex**, using other processors may need you to change some features such as graphics types.
- Note the files included in the **Project Report Template.tex** (with the .tex extension excluded). You can open these files separately and modify their contents or create new ones.
- Contact the latex namual for more features in your document such as equations, subfigures, footnotes, subscripts & superscripts, special characters etc.
- I recommend using the **kile** latex IDE, as it is simple to use.

Contents

Chapter 1

Introduction

1.1 Background to the study

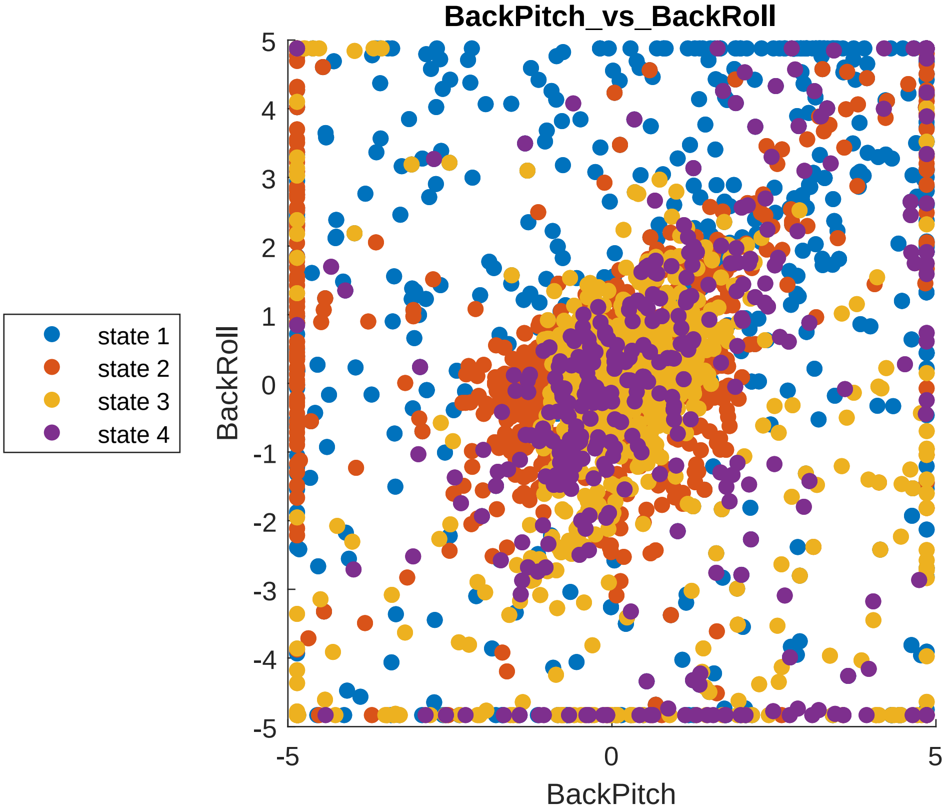
A very brief background to your area of research. Start off with a general introduction to the area and then narrow it down to your focus area. Used to set the scene [1].

Bio-inspired robotics uses nature to inform real-world engineering systems. Research has been conducted at UCT to investigate the manner in which a cheetah uses its tail for stability during high acceleration, quick turns and sudden braking, with an aim to incorporating identified mechanisms into sophisticated robot designs. One way to acquire useful data is to strap an inertial measurement unit (IMU) to an animal, and log the sensor data while certain actions are being performed. We currently have such a dataset of a dog moving, along with corresponding video data.

1.2 Objectives of this study

The objective of this project is to design, implement, and test Hidden Markov Models (HMM) for estimating gait sequence from Inertia Measurement Unit (IMU) data.

so that specific models can be formulated and their parameters estimated and interrogated. The project can be extended to include any other useful analysis of gait patterns from similar sensor measurements



1 - formulate model 2 - estimate its parameters 3 - Interrogate its parameters 4 - Useful analysis of gait patterns from IMU measurements

1.2.1 Problems to be investigated

Description of the main questions to be investigated in this study.

The main questions to be answered are the following:

1. How well can HMM model gait sequence dynamics using IMU data, in the abscence of enough training samples?
2. Can dimensionality reduction cause an increase in performance of HMM models when there is not enough training data?

1.2.2 Purpose of the study

Give the significance of investigating these problems. It must be obvious why you are doing this study and why it is relevant.

1.3 Scope and Limitations

Scope indicates to the reader what has and has not been included in the study. Limitations tell the reader what factors influenced the study such as sample size, time etc. It is not a section for excuses as to why your project may or may not have worked.

1 - Does not include data collection 2 - Focus on design of HMM only 3 - Focus on analysis of the model 4 - Focus on impact of dimensionality reduction

1.4 Plan of development

Here you tell the reader how your report has been organised and what is included in each chapter.

I recommend that you write this section last. You can then tailor it to your report.

Chapter 2

Literature Review

Once upon a time engineers and researchers believed... In this area of research, they used the following methods... [2]

Write this section first as it will take you the longest. I suggest you start writing this as soon as you have done your initial research at the beginning of your project. You can then return to it once you have completed your work to edit and adjust it.

A literature review forms the theoretical basis of your project. You need to read a large number of journal papers, sections in books, technical reports etc. relevant to your work at the start of project. This will give you a good idea of the field of research.

When writing your review start of with the general concepts and move to the more specific aspects explaining the necessary theory as you go. This section is NOT a copy and paste from others work or a rewrite-but-change-one-word section. I suggest you read all your material, and then put it down and write this section, referring back to the work only when you need to check something.

See your PCS textbook for more details on how to write a literature review.

If you include a figure or a table in your text please see the example in Fig. 3.1 as to how to caption it. Please make sure that all text in your figures is readable and that you reference your figures if they are from another source.

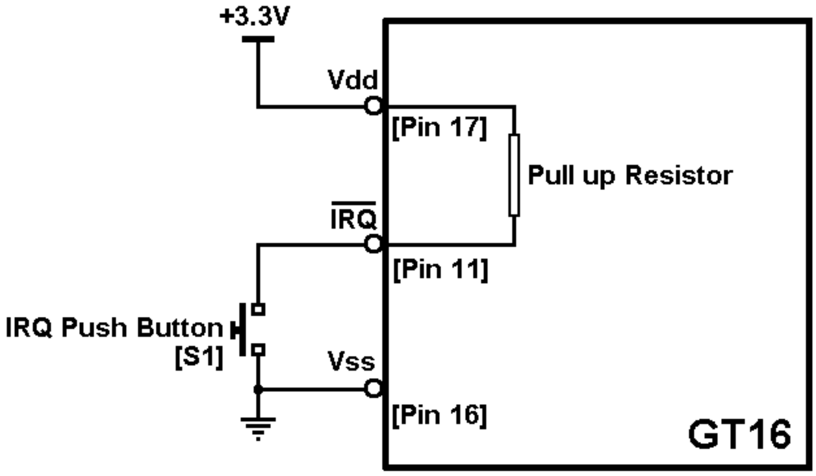


Figure 2.1: A block diagram illustrating the connections to the IRQ pin on the MCS08GT16A microcontroller (Please note that your headings should be short descriptions of what is in the diagram not simply the figure title)

2.1 Gait sequence modelling and estimation

2.1.1 Quadrupede gait modelling

Periodicity

2.1.2 Quadrupede gait estimation

2.2 Case study: Inertia Measurement Unit

2.3 Hidden Markov Models

Assumption of statistical model: Signal can be parametrised as a parametric random process and that the parametrers of the stochastic process can be determined/estimated in a precise, well-defined manner.
Observation is a probabilitistic function of the state - doubly embedded stochastic process.
Each state characterised by the probability distribution of observations, and transitions between states are characterised by a state transition matrix.

First order Markov model - current and predecessor only considered

2.3.1 Transition Probability Matrix

2.3.2 Emission Probability Matrix

2.3.3 Initial distribution

2.3.4 Elements of an HMM

1. Number of states, N
2. Number of distinct observation symbols per state, M
3. The initial state distribution, pi

2.3.5 Three fundamental problems for HMM design

1. Number of states, N
2. Number of distinct observation symbols per state, M

2.3.6 Types of HMM

Ergodic model

Left-Right model or Bakis model

Evaluation of the probability of a sequence of observations

The determination of a best sequence states

The adjustment of model parameters to account for observed signal

2.4 k-Nearest Neighbour

2.5 Dimension reduction

2.5.1 Feature selection

2.6 Sufficiency of Training Data

2.7 Techniques to increase Training Data

2.7.1 Mirroring

Chapter 3

Hidden Markov Model design

This section focuses on the design of the HMM used to test the hyphotheses postulated above.

3.1 Description of available dataset

The available dataset was acquired from a moving dog using Inertia Measurement Units. Two inertial measurements units (IMU) were straped to the front and back of a dog. Each unit has an accelarometer, a gyroscope and a magnetometer. The dataset contains calibrated measurements of a dog running, walking, or trotting then walking; together with the footfalls. The footfall is a binary value that indicates the state of the dog's leg: if it is on or above the ground, at a particular instant in its gait sequence. The four variables representing the footfalls effectively constitutes the ground truth informing us about the state in which the dog is at any time of the gait sequence.

The dataset can be retrieved from nine different matlab files. Each file contains twenty four matlab variables. The variables of interest are listed below.

Observations	Footfalls		
	Accelerometer	Gyroscope	Magnetometer
Front	accFrontX	FrontPitch	magFront_cal
	accFrontY	FrontRoll	magFront_cal2
	accFrontZ	FrontYaw	magFront_cal3
Back	accBackX	BackPitch	magBack_cal
	accBackY	BackRoll	magBack_cal2
	accBackZ	BackYaw	magBack_cal3

The observations are continous and the statistical property are assumed to be stationary, i.e, the do not vary over time. In this sta

stationary: statistical property do not vary over time or non-stationary: properties vary over time

pure or corrupted?

3.1.1 Gait sequence modelling

One of the objectives of this project is to effectively model the gait sequence dynamic of the dog using IMU measurements. The gait sequence was modelled as a succession of hidden or latent states observed by the IMU measurements.

States and observation

The state space is

$$S = S_i = \{(LF, RF, LB, RB)\} = \{0000, 0001, 0010, ..., 1111\}. \tag{3.1}$$
$$i = 1, 2, ..., 16 \tag{3.2}$$

is made up of 16 different states that steam from the combination of the four binary footfalls as shown in equation. 3.1

The states sequences are observed through the IMU measurements that form the feature space. An observation instance is a row vector of K dimensions. The initial K value is

18 from the 18 IMU measurements. Thus the observation sequence O is an TxK matrix of continuous values as presented in 3.3. T is the number of observations in a sequence.

$$O = \{Ok_t\} = O1_t, O2_t, ..., O18_t. \tag{3.3}$$

$$k = 1, 2, ..., 18. \tag{3.4}$$

$$t = 1, 2, ..., T. \tag{3.5}$$

Splitting the design into front and back module

To reduce the HMM model's complexity, it was broken down into front and back sub-HMM. These two models may be combined to get back the holistic 16-states model. This simplifies the model by reducing the number of states from 16 down to 4 states. This design design was motivated by the fact that it is a simpler task to classify 4 classses than discriminating between 16 classes. The remainig of this section will focus on the design of the 4-states HMM model.

Transition between states

This design assumes that a dog can transition from one state to any other possible state. So, for any transtion from S_i to S_j both S_i and S_j may be any of the element of the state space

$$S = \{S1, S2, S3, S4\}$$

.

For instance, if a dog has its left leg above ground and its right leg on ground, at time instance t, it may move to any of the 3 other possible positions or remain in the same state, in the next time instance, t + 1. This consideration yielded in an ergodic HMM where, all the transitions are possible. The graphical model of the simplified HMM is illustrated by figure 3.1

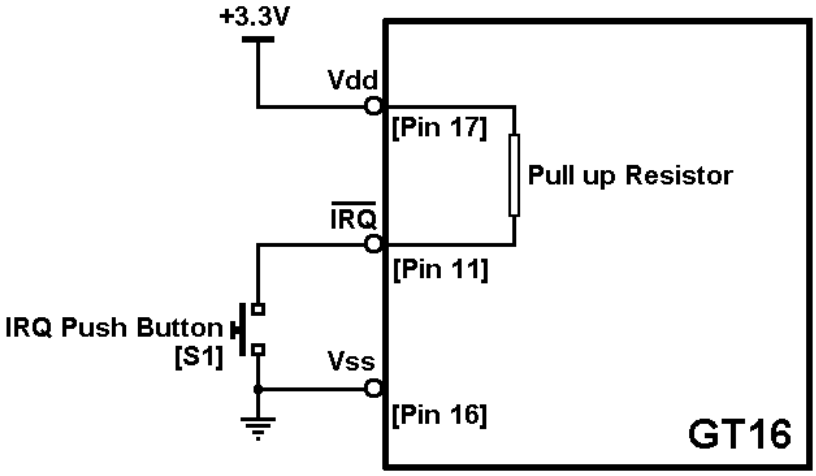


Figure 3.1: Ergodic HMM graphical model showing the hidden states, observation, and transitions between states

3.1.2 Data pre-processing

3.1.3 Parameters of the model

As a reminder, a continuous HMM model is completely specified by its initial state distribution: π transition matrix: A; the mean covariance matrices: μ, Σ which can be combined into Φ. If the observations are modelled with gaussian mixture distributions, an addition parameter is added for the initial mixture distribution: β. The next section how each parameters were estimated in this project.

Transition matrix: A

For each of the front and back 4-state HMM, the state transition matrix is a 4-by-4 matrix. Two different approaches were considered in the estimation of A. The two methods make use of expectation maximisation algorithm but differ in the modelling of the underlying mechanism.

1. Approach 1: Exploiting available ground truth
   This approach takes advantage of the ground truth for a labelled dataset to reduce the HMM model to a Markov Chain. Thus, the hypothetical observation sequence is made identical to the state sequence. Similar to discrete HMM, the transition matrix can be estimated using maximum likelihood algorithms. This approach

however makes two assumptions. It not only assumes assumes that the number of states is known but also requires labelled data. Fortunately, the second approach eliminates these constraints.

2. Approach 2: This method is the standard approach found in literature using expectation algorithm such as Welch-Baum algorithm.

Mean and covariance matrices: μ and Σ

Inital state distribution

1. Gaussian distribution
2. Gaussian mixture distribition

3.1.4 Dimension reduction

From 18 to 4 features using classification with knn

3.1.5 Continuous observations/features to Discrete observations/features

3.2 HMM implementation

Chapter 4

Results

These are the results I found from my investigation.

Present your results in a suitable format using tables and graphs where necessary. Remember to refer to them in text and caption them properly.

4.1 Aim

The objetive of this project is to design, implement and evaluate an algorithm, a model or a machine to predict the gait sequence of an animal (quadrupede or bipede) using Markov models. Thus, for a given state S, the model should be able to predict the next state S + 1, with a certain degree of confidence.

4.2 Apparatus

4.3 Methods

4.4 Results

4.5 Analysis

4.5.1 Discrete Probability density function duration d in state i

4.5.2 Expected number of observations (duration) in a state

4.6 Experimental Results

Chapter 5

Discussion

Here is what the results mean and how they tie to existing literature...

Discuss the relevance of your results and how they fit into the theoretical work you described in your literature review.

Chapter 6

Conclusions

These are the conclusions from the investivation and how the investigation changes things in this field or contributes to current knowledge...

Draw suitable and intelligent conclusions from your results and subsequent discussion.

Chapter 7

Recommendations

Make sensible recommendations for further work.

Use the IEEE numbered reference style for referencing your work as shown in your thesis guidelines. Please remember that the majority of your referenced work should be from journal articles, technical reports and books not online sources such as Wikipedia.

Bibliography

[1] M. S. Tsoeu and M. Braae, "Control Systems," *IEEE*, **vol. 34(3)**, pp. 123-129, 2011.

[2] J. C. Tapson, *Instrumentation*, UCT Press, Cape Town, 2010.

Appendix A

Additional Files and Schematics

Add any information here that you would like to have in your project but is not necessary in the main text. Remember to refer to it in the main text. Separate your appendices based on what they are for example. Equation derivations in Appendix A and code in Appendix B etc.

Appendix B

Addenda

B.1 Ethics Forms