Hand Gesture Analysis & Recognition:

An Internet of Things Interface Design Solution

Jason Mc Clafferty

BSc (Hons) in Computing with Games Development 2018/2019



Computing Department, Letterkenny Institute of Technology, Port Road, Letterkenny, Co. Donegal, Ireland.

Hand Gesture Analysis & Recognition:

An Internet of Things Interface Design Solution

Author: Jason Mc Clafferty

Supervised by: Lecturer’s Name

A thesis submitted in partial fulfilment of the requirements for “Research in Computing with Emerging Technologies” and “Project Development”

BSc (Hons) in Computing with Games Development.

Submitted to Quality and Qualifications Ireland (QQI)  
Dearbhú Cáilíochta agus Cáilíochtaí Éireann May 2019

# Declaration

I hereby certify that the material, which l now submit for assessment on the programmes of study leading to the award of BSc (Hons) in Computing with Games Development, is entirely my own work and has not been taken from the work of others except to the extent that such work has been cited and acknowledged within the text of my own work. No portion of the work contained in this thesis has been submitted in support of an application for another degree or qualification to this or any other institution.

I hereby certify that the material on which I have relied on for the purpose of my assessment is not deemed as personal data under the GDPR Regulations. Personal data is any data from living people that can be identified. Any personal data used for the purpose of my assessment has been pseudonymised and the data set and identifiers are not held by the LYIT. Alternatively personal data has been anonymised in line with the Data Protection Commissioners Guidelines on Anonymisation.

I understand that the dissertation produced for this BSc may be stored electronically by LYIT’s library, may be stored electronically by relevant academic staff for review and/or assessment purposes and a printed copy may be retained within the Department of Computing after the student has completed their studies.

Signature of Candidate Date

# Acknowledgements

I would like to thank…

# Abstract

Written here is no more than 250 words of summary of the problem, hypothesis and main conclusions. The abstract will entice people to read the rest of the document.

# Acronyms

|  |  |  |
| --- | --- | --- |
| Acronym | Definition | Page |
| HCI | Human-Computer Interaction | 1 |
| IoT | Internet of Things | 1 |
| UX | User-Experience | 1 |
| UI | User Interface | 1 |
|  |  |  |

# Table of Contents

[Declaration 4](#_Toc429471704)

[Acknowledgements 5](#_Toc429471705)

[Abstract 6](#_Toc429471706)

[Acronyms 7](#_Toc429471707)

[Table of Contents 8](#_Toc429471708)

[Table of Figures 9](#_Toc429471709)

[Table of Tables 10](#_Toc429471710)

[Table of Code Listings 11](#_Toc429471711)

[1. Introduction 1](#_Toc429471712)

[1.1. Purpose 1](#_Toc429471713)

[1.2. Background 1](#_Toc429471714)

[1.2.1. One liner 1](#_Toc429471715)

[1.2.2. Sub-Item 1](#_Toc429471716)

[1.3. Research Question 1](#_Toc429471717)

[1.4. Report Outline 2](#_Toc429471718)

[2. Introduction 3](#_Toc429471719)

[2.1. Problem Context 3](#_Toc429471720)

[2.2. Report Outline 3](#_Toc429471721)

[2.3. Tables 4](#_Toc429471722)

[2.4. Code and Formulae 5](#_Toc429471723)

[Appendices 7](#_Toc429471724)

[Appendix A: References i](#_Toc429471725)

[Appendix B: Code Listing ii](#_Toc429471726)

# Table of Figures

[Figure 1 Letterkenny Institute of Technology Logo 13](#_Toc429427538)

[Figure 2 Letterkenny Institute of Technology Logo 13](#_Toc429427539)

# Table of Tables

[Table 1. Table Formatting Guidelines 15](#_Toc429429027)

[Table 2 Second Sample Table 16](#_Toc429429028)

# Table of Code Listings

[Code Listing 1 MDBean Message Handling 16](#_Toc429428842)

# Introduction

Research in human-computer interaction (HCI) has experienced a renaissance recently as computing technology develops beyond the desktop & mouse model, smartphones and tablets, and towards the Internet of Things, Smart Home technology & the Digital Mesh. Although these technologies have advanced at an accelerating rate the communicative interfaces have only in the last 8-10 years begun to be explored significantly.

“It has been widely believed that the computing, communication and display technologies progress further, but the **existing techniques may become a bottleneck in the effective utilization of the available information flow**. To efficiently use them, most computer applications require more and more interaction. For that reason, human-computer interaction (HCI) has been a lively field of research in the last few years.” (Rautaray and Agrawal, 2015).

Today it is possible to turn on a washing machine from a voice-controlled unit in an entirely separate location, but the main means of communicating with technology remains the WIMP paradigm. WIMP stands for Windows, Icons, Menus, and Pointers, was a state-of-the-art graphical interface in 1995, 23 years ago. Moore’s Law, redefined in a modern context by Moore, (2006) in Layman’s terms states that technological advancement will accelerate year on year, but this has not held true for human-first user interfaces (Turk and Ed, 2000). The rate at which technology is advancing is accelerating, but since 1980 there has been almost no major deviation from the WIMP protocol for user interface or human computer interaction.

The author believes that the fundamental reason for this phenomenon lies in one of the following possibilities:

(i) The HCI mechanics developed for/using 80’s/90’s technology are still the most appropriate way to use technology today.

OR

(ii) The industry has favoured ‘tried and tested’ design instead of committing to the innovation of the user’s experience by investing in technology to develop that experience fully.

The author understands why ‘tried and tested’ design is favourable – it’s tested, designers know that it works, it’s as pervasive in technology as computers themselves. Computer software is designed with mouse and keyboard in mind. However, as the possibility to communicate through almost any modality (A modality is an input/output channel for information flow. Humans have 6 modalities – Vocal, Aural, Olfactory, Visual, Gesture-based & Haptic) \*\* Need to Reference Simon’s Lecture Reference\*\* has become a reality in the past few years, few of the possibilities to implement a more natural human-computer interaction have been realised.

\*\*List the modalities, the main attempts at integrating them into HCI, and how successful they were/are\*\* Matthew Turk Paper – Multimodal Interaction: a review

Virtual personal assistants such as Siri, Alexa and Google Assistant have become more involved in our day to day lives.

“A widely accepted prediction is that computing will move to the background, weaving itself into the fabric of our everyday living spaces and projecting the human user into the foreground. To realise this prediction, next-generation computing should develop anticipatory user interfaces that are human-centred, built for humans and based on naturally occurring multimodal human communication. These interfaces should transcend the traditional keyboard and mouse and have the capacity to understand and emulate human communicative intentions as expressed through behavioural cues, such as affective and social signals.” (Pantic *et al.*, 2008)

The general topics relevant to the author’s discussion are User Experience and Human-Computer Interaction (HCI), more specifically gesture analysis and its feasibility as an interface. The IoT is a catch all term for technologies that implement sensors, data processing, communication and user interface in order to communicate with one another and form a ‘Digital Mesh.’ (McClelland, 10/10/2018) The IoT, as a technology, was featured on Gartner’s Top 10 2018, and Smart Spaces - the next iteration of IoT & the Digital Mesh expanded to city scale, was featured on Gartner’s Top 10 2019.

## Purpose

This paper is focused on the areas of gesture-based communication and interface design, the factors that encourage and discourage the use of gesture, and the use of computer vision technology to record and analyse hand-gesture. The information is synthesised and presented to theoretically support the assumption that a gesture-based interface is a more natural mode of communication between human and computer.

## Background

There is a reasonable body of research in this field, but it is growing rapidly and the author expects the research body to expand rapidly as the benefits of gesture as a means of HCI are illuminated.

## Aims

### To recognise hand gestures through a visual-input.

### To analyse hand gestures and discern a lexicon, syntax and semantic structure capable of being understood as input by a software program. This is achieved by collecting qualitative data from a practical number of subjects and assessing their responses.

### To design a user-first interface for controlling smart lights.

### To establish the limits of the research in this area and provide a platform for further study.

## Objectives

### To devise a research strategy capable of delivering the required information enabling the author to provide credible references and conclusions to the research question/hypothesis posed.

### To gain insight into hand-gesture recognition/analysis, open cv and the Internet of Things technology.

### To provide the reader with a clear understanding of hand-gesture recognition/analysis, open cv and the Internet of Things technology.

### To deliver a working artefact that to display the proof of concept.

### To widen the author’s technical skillset by working with different languages and across different hardware and software protocols.

## Research Question

Is it feasible to use hand gestures and Open CV to control a Phillips Hue Smart-Light?

Can an AR/VR solution be used to address the problems of illumination and depth estimation in gesture analysis?

## Report Outline

The author expects the study to remain within the domains of communication and user experience, specifically HCI, interface design and gesture recognition/analysis.



Figure 1.1 Letterkenny Institute of Technology Logo

In order to provide a caption for an image, table or equation the item should be selected. The Microsoft References ribbon should be selected. From there the Insert Caption button should be selected. The label should be set to the most appropriate one. In this example the Figure label was selected. New labels can be created as necessary. Remember not to simply copy and paste from above. Instead insert the image into the file and select it. Add the caption as described.

Where an image does not have a clearly defined border, one should be added. Care should be taken to ensure that all details of images are clearly visible both when in print and when in electronic format. Careful selection of colours should be considered for this purpose. Images as shown in Figure 1.2 should always be referred to from the main text.



Figure 1.2 Letterkenny Institute of Technology Logo

# Literature Review

## Hand Gesture Analysis/Recognition

* According to Hall (1973), human communication consists of verbal communication (35%) and non-verbal, gesture-based communication (65%).
* Hand Gestures are a means of … non-verbal communication.
* A gesture is any movement of the face, arm or body that is meaningful or informative.
* Gestures may be articulated with any part or combination of parts of the body. Gestures as a means of communication may serve as an important means of HCI. Gestures are subjective across cultures – pointing finger rude in Asia for E.g. (Chakraborty *et al.*, 2018)
* Hands involved in 62% of body parts used for gesturing.(Chakraborty *et al.*, 2018)
* There are two types of theoretically defined hand gestures – static and dynamic. Static = position and orientation of hand in space with no movement, dynamic = moving hand.
* Metaphoric gestures and iconic gestures are sub-gestures and are important.(Rautaray and Agrawal, 2015)
* Gesture recognition, semantically meaningful commands, design and development of such systems that can identify explicit human gestures as input and process these gesture representations for device control through mapping of commands as output.
* The visual interface establishes channels for inferring intentions for inferring intentions from human behaviour, including facial expressions and hand gestures. (Chakraborty *et al.*, 2018)
* Models for vision-based hand gesture recognition focus on the hand orientation in space, and automate the temporal aspect of dynamic gesture recognition, dividing it into three phases (McNeil, 1992) – the pre-stroke phase (preparation phase), stroke phase (nucleus phase), post-stroke phase (retraction phase). Every phase corresponds to one or many transitions of spatial states of the 3d human model.
* Difficulties involved with video-based gesture recognition (VGR): Depth estimation, Illumination issues., cluttered backgrounds, synchronising hardware, large streams of data (real-time systems), user feedback for user learning (Chakraborty *et al.*, 2018).

The chapters should all begin on a new page. Page numbers appear at the bottom right of each page. Page numbers appear from Chapter 1 and onwards. Appendices should be paginated using roman numerals (I, II, etc.).

* Bullets should be aligned with the text.
* Bulleted items should have one blank line above and below.
* If there are only 1 or 2 items a bulleted list is not required.

Any paragraph after the list, image or table should resume its normal position for the given header. When using any acronym such as Some Silly Acronym (SSA) it must be expanded on its first occurrence within the text. All acronyms should appear in an acronyms list preceding the main chapters.

## Problem Context

In some

## Report Outline

The text of any given chapter may refer to an interesting idea presented in another book, paper, journal or whitepaper. On-line sources should not normally constitute more than 50% of your references. ALL of your references must be peer-reviewed or whitepapers. Further details on this will be given in the Research Workshop and may be obtained from your supervisor. This is supported by research (Bloggs, 2012) carried out how best to reference. The idea is succinctly expressed by Murphy:

“A reference in a thesis should be of the previously demonstrated Harvard Style.” (Murphy, 2011a)

Notice that the quote has indentations on both sides and is surrounded by quotes. If the quote abstracts only part of a sentence double dots should be placed before or after to show where there is missing text. Further, where additions for clarification are used in the text square brackets should be used. According to Murphy:

“A reference in a thesis should be of… Harvard Style.” (Murphy, 2011b)

While the example above is provided for demonstration purposes it is obviously not a good idea to provide the same quote twice so for the purposes of this example we will assume this quote was taken from a different book by the same author. Further the general use of a single or a small number of sources multiple times is referred to as ‘over-reliance on a source’ and is deemed plagiarism.

## Tables

The text of any chapter may include tabular data. In order to aid legibility some simple guidelines should be adhered to. Refer to Table 2.1.

Table 2.1 Table Formatting Guidelines

|  |  |
| --- | --- |
| **Format** | **Description** |
| Size | The table should be able to fit into one page and should not overrun. |
| Margins | The table should not extend past the normal margins of the page |
| Colour | Colour may be used but consideration should be given to both on screen display and printed display. |
| Design | Simple designs are best. At all times consider that the information in the table is more important than the ‘flashy’ design. |

The title for an image or code must appear directly underneath and on the same page. If this is not possible then move the item within the text to ensure that the caption remains with the item. The title for a table must appear directly above the table.

Refer to Table 2.2 for the second short table sample.

Table 2.2 Second Sample Table

|  |  |
| --- | --- |
| **Format** | **Description** |
| Size | The table should be able to fit into one page and should not overrun. |

## Code and Formulae

Where code requires listing within the text it should be treated as an image in that it is sectioned off with a border and has a caption directly underneath. Refer to Code Listing 2.1 below.

…

TextMessage msg = null;

try {

if (message instanceof TextMessage) {

msg = (TextMessage) message;

System.out.println("A Message received in TMDB: "

+ msg.getText());

}

else {

System.out.println("Message of wrong type: "

+ message.getClass().getName());

}

} catch (JMSException e) {

e.printStackTrace();

mdc.setRollbackOnly();

} catch (Throwable te) {

te.printStackTrace();

}

…

Code Listing 2.1 MDBean Message Handling

Notice that only minimal commenting is provided within the text. The code is shown in Cambria, 10 point. This reduces the overall text size and clearly distinguishes it from the main text.

If a single line of formula is required it can be referred to within the text as formula (Equation 2.1) for example with the formula example shown slightly indented and with the formula number to the far right.

Equation 2.1

A single line space above and below the formula (Tsiolkovsky, 2000) also aids legibility. Note also that the font size is increased by one point. A further item to note is that equations may also be referenced.

# Appendices

# Appendix A: References

Chakraborty, B. K. *et al.* (2018) ‘Review of constraints on vision-based gesture recognition for human–computer interaction’, *IET Computer Vision*, 12(1), pp. 3–15. doi: 10.1049/iet-cvi.2017.0052.

McClelland, C. (no date) *IoT Explained — How Does an IoT System Actually Work?* Available at: 1 (Accessed: 10 October 2018).

Moore, G. E. (2006) ‘Understanding Moore’s Law: Four Decades of Innovation’, *Understanding Moore’s Law: Four Decades of Innovation*, (Figure 2), pp. 67–84. doi: 10.1086/529346.

Pantic, M. *et al.* (2008) ‘No Title’, *International Journal of Autonomous & Adaptive Communications Systems*, Vol. 01(No. 02), pp. 168–187.

Rautaray, S. S. and Agrawal, A. (2015) ‘Vision based hand gesture recognition for human computer interaction: a survey’, *Artificial Intelligence Review*. Springer Netherlands, 43(1), pp. 1–54. doi: 10.1007/s10462-012-9356-9.

Turk, R. 1 and Ed, M. (2000) *Proceedings of the 1998 Workshop on Perceptual User Interfaces*, *COMMUNICATIONS OF THE ACM*. Available at: http://www.cs.ucsb.edu/~mturk/Papers/CACM2000.pdf (Accessed: 22 October 2018).

# Appendix B: Code Listing