Chapter 11
ZatLab:
Programming a Framework for Gesture Recognition and Performance Interaction

André Baltazar
Catholic University of Portugal, Portugal

Luís Gustavo Martins
Catholic University of Portugal, Portugal

ABSTRACT

Computer programming is not an easy task, and as with all difficult tasks, it can be faced as tedious, impossible to do, or as a challenge. Therefore, learning to program with a purpose enables that "challenge mindset" and encourages the student to apply himself in overcoming his handicaps and exploring different theories and methods to achieve his goal. This chapter describes the process of programming a framework with the purpose of achieving real time human gesture recognition. Just this is already a good challenge, but the ultimate goal is to enable new ways of Human-Computer Interaction through expressive gestures and to allow a performer the possibility of controlling (with his gestures), in real time, creative artistic events. The chapter starts with a review on human gesture recognition. Then it presents the framework architecture, its main modules, and algorithms. It closes with the description of two artistic applications using the ZatLab framework.

INTRODUCTION

There is so much information in a simple gesture. Why not use it to enhance a performance? We use our hands constantly to interact with things. Pick them up, move them, transform their shape, or activate them in some way. In the same unconscious way we gesticulate in communicating fundamental ideas: stop; come closer; go there; no; yes; and so on. Gestures are thus a natural and intuitive form of both interaction and communication (Watson, 1993). Children start to communicate by gestures (around 10 months age) even before they start speaking. There is also ample evidence that by the age of 12 months children are able to understand the gestures other people produce (Rowe & Goldin-meadow, 2009). For the most part gestures are considered an auxiliary way of

DOI: 10.4018/978-1-4666-7304-5.ch011

communication to speech, tough there are also studies that focus on the role of gestures in making interactions work (Roth, 2001).

It is also important to understand that whereas all gestures derive from a chain of movements, not all movements can be considered gestures (Kendon, 1994). Gestures are the principal non-verbal, cross-modal communication channel, and they rely on movements for different domains of communication (Volpe, 2005). Looking at the Merriam-Webster dictionary[1], one will find the word "gesture" means a movement usually of the body or limbs that expresses or emphasizes an idea, sentiment, or attitude, as well as the use of motions of the limbs or body as a means of expression.

Gestures and expressive communication are therefore intrinsically connected, and being intimately attached to our own daily existence, both have a central position in our (nowadays) technological society.

However, the use of technology to understand gestures is still somehow vaguely explored, it has moved beyond its first steps but the way towards systems fully capable of analyzing gestures is still long and difficult (Volpe, 2005). Probably because if in one hand, the recognition of gestures is somehow a trivial task for humans, in other, the endeavor of translating gestures to the virtual world, with a digital encoding is a difficult and ill-defined task. It is necessary to somehow bridge this gap, stimulating a constructive interaction between gestures and technology, culture and science, performance and communication. Opening thus, new and unexplored frontiers in the design of a novel generation of multimodal interactive systems.

This chapter describes the entire process of learning how to program and implement a framework that enables the recognition of gestures in real-time and their use for artistic purposes. Therefore, first one will review the literature on gesture research, followed by the framework proposal, implementation and application.

BACKGROUND

Introduction

As Godoy (Godøy & Leman, 2009) refers, there is no clear definition of what a gesture is: "Given the different contexts in which gestures appear, and their close relationship to movement and meaning, one may be tempted to say that the notion of gesture is too broad, ill-defined, and perhaps too vague." This framework is focused on gesture recognition, so there is intrinsically a demand for the explanation and definition of the terms that are not well clarified.

This section is dedicated to the understanding and definition of a gesture and how it can be captured and recognized. It will also discuss the previous works published on this research field and present a review and technologically comparison of the different Motion Capture (MoCap) systems available nowadays. This section will provide valuable input for the development of the proposed framework.

Gestures

The human movement (Zhao & Badler, 2001) can be involuntary, subconscious, that occurs for biological or physiological purposes (e.g. blinking, breathing, balancing), or voluntary, conscious like those task-driven actions such as speaking or running to get somewhere. There is also a wide class of movements that fall in between these two, having both the voluntary and involuntary qualities. Such movements are the ones that occur in an artistic performance or music concert and perhaps unconsciously with other activities. These can range from leg and foot coordination enabling walking, till the communicative gestures, such as facial expressions, expressive limb gestures and postural attitude. The communicative gestures are the focus of this work and thus, their definition is of central importance.