Artif Intell Rev (2013) 40:495–516
DOI 10.1007/s10462-011-9292-0

Hidden Markov model for human to computer interaction: a study on human hand gesture recognition

Sara Bilal · Rini Akmeliawati · Amir A. Shafie · Momoh Jimoh E. Salami

Published online: 28 December 2011

Abstract Human hand recognition plays an important role in a wide range of applications ranging from sign language translators, gesture recognition, augmented reality, surveillance and medical image processing to various Human Computer Interaction (HCI) domains. Human hand is a complex articulated object consisting of many connected parts and joints. Therefore, for applications that involve HCI one can find many challenges to establish a system with high detection and recognition accuracy for hand posture and/or gesture. Hand posture is defined as a static hand configuration without any movement involved. Meanwhile, hand gesture is a sequence of hand postures connected by continuous motions. During the past decades, many approaches have been presented for hand posture and/or gesture recognition. In this paper, we provide a survey on approaches which are based on Hidden Markov Models (HMM) for hand posture and gesture recognition for HCI applications.

Keywords HCI applications · HMM · Artificial intelligence · Hand posture recognition · Hand gesture recognition

1 Introduction

In recent years, many research have been conducted to determine natural approaches in HCI. The typical approaches currently are based on keyboard and mouse input. Hand gesture is one alternative approach, which has gained much attention for real-time HCI applications (Derpanis 2004; Mitra and Acharya 2007; Murthy and Jadon 2009). Vision based HCI systems have the ability of carrying a wealth of information in a natural way and at a low cost. Hand posture and/or gesture recognition systems can identify specific human hand posture and/or gesture and use them to interact with particular machines/computers.

S. Bilal (✉) · R. Akmeliawati · A. A. Shafie · M. J. E. Salami
Department of Mechatronics Engineering, International Islamic University Malaysia (IIUM),
Jl Gombak, 53100 Kuala Lumpur, Malaysia
e-mail: smosb@hotmail.com

R. Akmeliawati
e-mail: rakmelia@iium.edu.my

Hand posture recognition with no constraints on certain posture shape is an open issue because human hand is a complex articulated object consisting of many connected parts and joints. Considering the global hand pose and each finger joint, human hand has roughly 27 degrees of freedom (DOF) (Wu and Huang 2001). There are many HCI applications that require the recognition of hand posture and gesture. Both hand posture and gesture are combined together to perform a structure based applications such as Sign Languages (SL) translators. Non-structured gestures could be recognized using the same approaches as SL to control many devices. Computer games are other applications that are seeking natural control alternative such as hand.

Many traditional methods exist in the field of pattern recognition to achieve hand posture and gesture recognition (Derpanis 2004; Mitra and Acharya 2007) such as artificial intelligence techniques, statistical algorithms and other non traditional developed algorithms. For more details on the approaches used for visual human action recognition, readers can refer to the study by Del Rose and Wagner (2011).

HMM is a statistical method that has been used widely in pattern recognition applications. This paper surveys HMM topologies on hand posture and gesture recognition and their HCI application field. It highlights the advantages of HMM over the used artificial intelligence techniques in the field of hand gesture recognition. The organization of this paper is as follows; Sect. 2 gives a brief introduction to the computational tools used to manipulate HMM. Section 3 presents the HCI applications that have been developed using HMM for hand posture and/or gesture recognition. A comparison of HMM with other existing methods for hand posture and/or gesture recognition techniques is stated in Sect. 4. Section 5 concludes this paper.

2 HMM

HMM is a statistical model with a rich mathematical structure which can fit many applications requirement. Therefore, vast applications make use of the HMM-based recognition ranging from the automatic speech recognition (Rabiner 1989) to document processing in databases to image processing area (Aas et al. 1999). Their work provides some simple applications of HMM for image analysis in machine vision and video surveillance. These applications include text analysis, bottle recognition, vehicle detection and tumour classification using HMM.

In the research area of modeling and classifying dynamic gestures, HMM-based recognition has been a very popular technique, and mostly used in classification process because they offer dynamic time warping, a training algorithm, and a clear Bayesian semantics. The first approach for the recognition of human movements based on HMMs was introduced in Yamato et al. (1992a,b). It distinguishes between six different tennis strokes. The work by Rigoll et al. (1997, 1998) and Rigoll and Kosmala (1997) has a consequence of research development of gesture recognition using HMM. They have improved their system from Discrete HMM to Continuous HMM by extracting new statistical features, which has led to an improvement in the HMM outcomes. The extension of the system built by Rigoll et al. (1998) was able to recognize dynamic gestures in person and background independent mode. Depending on the type of the feature extracted from gestures, many HMM topologies have been extended from the conventional HMM approach such as Discrete HMM (DHMM), Continuous HMM (CHMM), and Partial HMM (PHMM) to handle more problems.

All HMM algorithms for HCI applications have been developed using computer programming languages such as C, VC++, MATLAB^{TM} etc as shown below.