SIViP
DOI 10.1007/s11760-013-0558-7

ORIGINAL PAPER

Spatio-temporal 3D surface matching for hand gesture recognition using ICP algorithm

Saeed Nasri · Alireza Behrad · Farbod Razzazi

Received: 19 February 2013 / Revised: 25 June 2013 / Accepted: 5 September 2013

Abstract With emerging new applications like virtual reality, different algorithms for human action and gesture recognition have been proposed. In this paper, a new method for the recognition of moving hand gestures is presented. The proposed algorithm is based on the representation of hand motion as spatio-temporal 3D surfaces. Then, 3D surface matching is used to recognize the hand gesture. To form the spatio-temporal 3D surface of hand motion, we first apply the necessary preprocessing to video frames and extract hand contours. Then, by normalizing and overlapping hand contours in different frames, we construct spatio-temporal 3D surface of the hand gesture. To recognize hand gesture, we match the input 3D surface with surfaces in the database. For this purpose, we utilize ICP algorithm to find and compensate for 3D transformation between surfaces as well as the similarity measure between them. In real-world applications, hand motion is continuous and results in a sequence of disjointed hand gestures, which is called continuous hand gesture. To recognize continuous hand gestures, we propose an algorithm which first estimates probable disjointed gestures in the continuous gesture and then divides iteratively continuous gestures to true disjointed gestures. Finally, by applying a robust algorithm, the continuous gesture is recognized. We tested the proposed algorithm with hand gestures of American sign language and results showed the recognition rate of 94 % for disjointed gestures and 93.9 % for continuous gestures. The experimental results showed the efficiency of the proposed algorithm for hand gestures with noise as well.

Keywords Gesture recognition · Hand gestures · Continuous gesture recognition · 3D surface matching · ICP (Iterative Closest Points) algorithm

S. Nasri (✉) · F. Razzazi
Electrical Engineering Department, Faculty of Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
e-mail: s_nasri@iaun.ac.ir

F. Razzazi
e-mail: razzazi@srbiau.ac.ir

A. Behrad
Electrical Engineering Department, Faculty of Engineering, Shahed University, Tehran, Iran
e-mail: behrad@shahed.ac.ir

1 Introduction

With emerging new applications like virtual reality, it is necessary to have more nearly perfect interfaces for human–computer interaction. In these applications, new tools are needed for users to interact with computer without mouse or keyboard 2D limitations [1]. Sensors and video cameras are two sample tools that have been used for this purpose. Sensor outputs need some filtering stages to eliminate temperature, pressure, and other noise effects. In addition to less accuracy, higher cost and limited flexibility of these devices because of wiring connections put them out of the reach for general use. Webcams are cheap and widespread. By applying proper video and image processing algorithm, they can be used for gesture recognition and interaction of humans with computer [2].

Vision-based hand gesture recognition is one of the most commonly used methods for human–computer interaction [3]. In this method, the motion information of hand is obtained using a video camera. Then, by processing video frames, motion type or hand gesture is recognized.

Figure 1 shows the block diagram of vision-based gesture recognition system. The system consists of two parts, including online and offline parts. In offline part, video images of different gestures are recorded in an offline manner, and

Published online: 13 October 2013