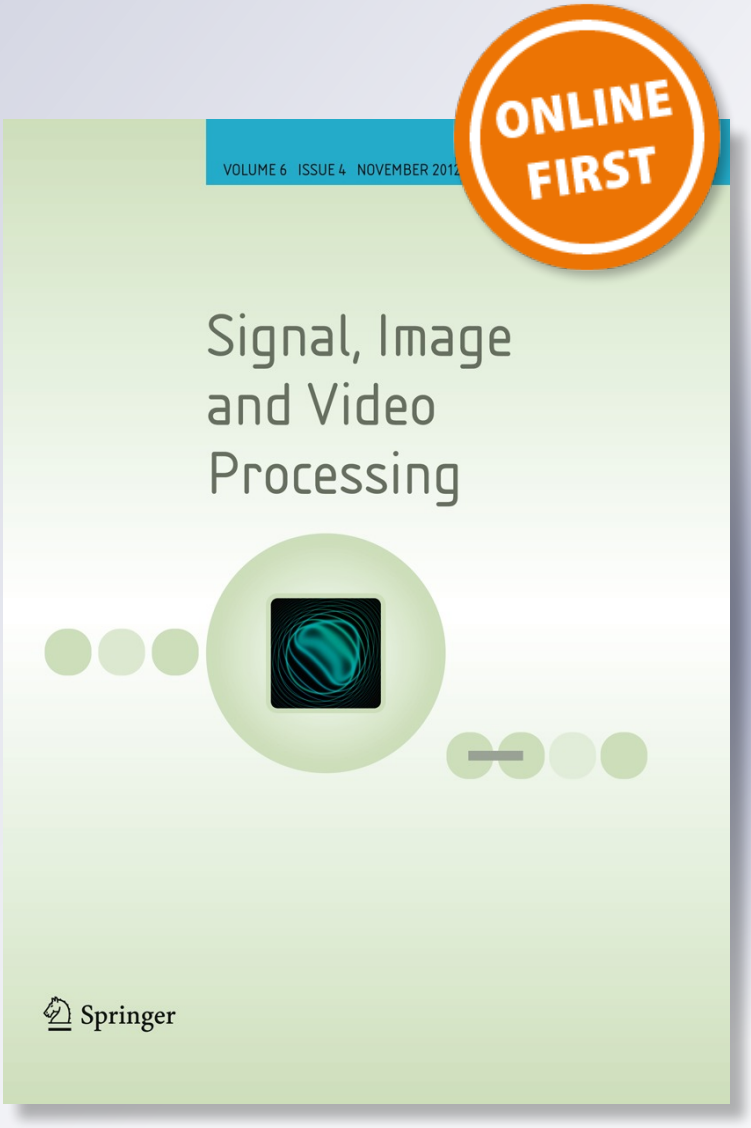
3D Head pose estimation and camera mouse implementation using a monocular video camera

Masoomeh Nabati & Alireza Behrad

Signal, Image and Video Processing

ISSN 1863-1703

SIViP
DOI 10.1007/s11760-012-0421-2



Springer

SIViP
DOI 10.1007/s11760-012-0421-2

ORIGINAL PAPER

3D Head pose estimation and camera mouse implementation using a monocular video camera

Masoomeh Nabati · Alireza Behrad

Received: 27 November 2011 / Revised: 8 December 2012 / Accepted: 10 December 2012

Abstract In this paper, we present a novel approach to estimate 3D head pose using a monocular video camera for the control of mouse pointer and generating clicking events. Our approach proceeds in four stages. First, the face area is detected using Haar-like features and AdaBoost algorithm. Second, the point features are extracted and tracked over video frames by KLT algorithm. Third, by employing the tracked point features and 2D motion model of the face area, we estimate the 3D rotation matrix and translation vector between web camera and the head position. Finally, the 3D rotation matrix and translation vector are employed to calculate the mouse pointer location on the PC screen and generating clicking events. Furthermore, we propose eye wink detection as an alternative for clicking event implementation.

Keywords Camera mouse · 3D head pose estimation · Mouse pointer control

1 Introduction

Nowadays, different support devices and care equipments have been developed to help the handicapped people. One of the main support devices for handicapped people with severe disabilities is an instrument for communication with computers [1]. A camera mouse system is a non-intrusive method that helps the handicapped people to interact with computers. A camera mouse system is generally composed of one or multiple video cameras for capturing video frames and a processing unit like a PC which uses an image processing algorithm to convert the motion events in video frames to mouse operations.

Different algorithms have been proposed for the implementation of the camera mouse which most of them used head pose or head motion as well as facial features like eyes, nostrils and mouth. Camera mouse system based on eye motion utilizes eye gaze or motion for camera mouse implementation. Different algorithms for eye gaze detection have been proposed using active or passive sensors [2,3]. In [2], an algorithm based on eye motion and eye blinking was proposed for the camera mouse implementation. The system is limited in resolution for camera mouse implementation because of the restricted number of gaze directions.

Shin et al. [4] proposed a camera mouse based on multiple facial features. The method implemented mouse movements using the user's eyes motion, while clicking events were implemented based on the user's mouth shapes, such as opening/closing. In this approach, user face was assumed to be in the frontal state which limits the generality of the algorithm. In [5] joint use of head pose and eye location information is proposed for human computer interfacing.

Some other algorithms tracked nose and nostril to provide camera mouse functionality [6,7]. The main problem with using facial features for camera mouse is their small area in the input image; therefore, high-resolution images or proper zoom ratio are needed. Tu et al. [8] introduced a camera mouse using the 3D face model based on visual face tracking. The method suffers from the accumulative tracking error and manual initialization. The camera mouse system of Manresa-Yee et al. [9] was also based on face tracking. However, 2D motion of the face or facial features was utilized for the camera mouse implementation. Since the head rotation generally produces 3D rotation, the algorithm detects face or facial features again in the case of lost features.

M. Nabati · A. Behrad (✉)
Faculty of Engineering, Shahed University,
Persian Gulf Express Way, Tehran, Iran
e-mail: behrad@shahed.ac.ir

M. Nabati
e-mail: masoomeh_nabaaty@yahoo.com

Published online: 27 December 2012