UbiHand: A Wearable Input Device for 3D Interaction

Farooq Ahmad [*]
University of Alberta
Edmonton, Canada

Petr Musilek [†]
University of Alberta
Edmonton, Canada

Abstract

Increasing interest in virtual and augmented reality is encouraging the development of unobtrusive, intuitive input interfaces for gesture recognition and 3D interaction. Current approaches require the user to wear instrumented gloves, or rely on fixed cameras that lack precise finger tracking. The UbiHand is an input device that uses a miniature wrist-worn camera to track finger position, providing a natural and compact wearable input interface. A hand model is used to generate a 3D representation of the hand, and a gesture recognition system can interpret finger movements as commands.

1 Introduction

Hand gesture interface development has followed two paths: data-glove based approaches which require the user to wear instrumented gloves; and machine vision based approaches that use a fixed camera placed in front of the user to segment the hands and determine hand position using a kinematic model. Glove based approaches incorporate several flex sensors to provide accurate measurements of the bending of each finger joint, but are cumbersome and expensive. On the other hand, vision based approaches do not require the user to wear any special equipment. However, video based motion capture using a fixed camera generally provides less precise finger tracking, and restricts the users movement to the cameras' field of view.

The approach described here is a fusion of the two approaches; a wrist-worn camera is used to determine the position of each finger using skin color segmentation, and the finger tracking data is used to generate a 3D model of the hand. The system can record 10 degrees of freedom of motion, corresponding to the flexion and abduction of four fingers (8DOF), and wrist pitch and yaw (2DOF). The device can be trained to recognize hand gestures to switch between modes; for example, object selection and manipulation, and drawing modes.

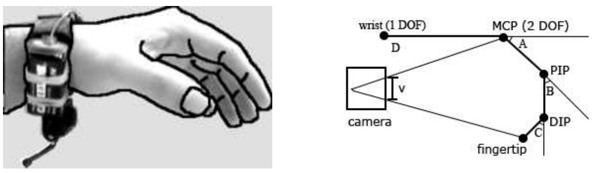

Figure 1: The prototype system and kinematic hand model. Bend $\angle A = \angle B = 1.5\angle C \propto v$ (metacarpal to fingertip image distance)

[*]e-mail: farooq@ualberta.ca
[†]e-mail: musilek@ece.ualberta.ca

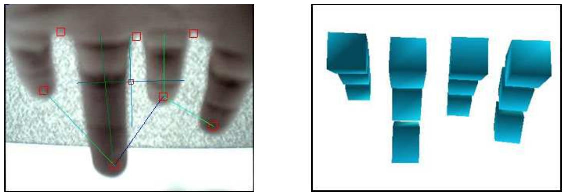

Figure 2: Hand Motion Capture and Generated Model

2 System Overview

In [Ahmad and Musilek 2006], a text input device based on a wrist worn camera is described. The work described here extends the approach to include a hand model and 3D finger pose estimation, for use within virtual and augmented reality environments. The prototype system is shown in figure 1. A wireless color camera is worn under the users' wrist and continuously tracks finger motion. The hand contour is extracted from the images by skin color discrimination. After applying morphological operators to remove noise, the peaks and valleys of the hand contour are recorded, corresponding to the four fingertips and metacarpal joints visible within the cameras' field of view. Finger motion constraints allow the determination of the abduction angle (side-to-side flexion) of each finger, as well as the most likely bend angles of the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints. More specifically, the hand model illustrated on the right side of figure 1 illustrates the relative bending constraints of each joint, and considers bend angles to be directly proportional to the image distance v between the metacarpal joint and fingertip. Using this model, the finger pose is determined and used to control a virtual hand model as shown in figure 2.

In addition to finger tracking, a 2D marker integrated into the wristband allows hand tracking in three dimensions using an additional camera. The second camera can be static, or integrated into a head mounted display, allowing use of the system as a 3D input device for augmented reality applications.

Gesture recognition allows the user to input discrete data such as text and commands. The fingers are tracked continuously, and rapid finger movements that exceed a threshold distance from a rest position are detected as keystrokes. When a keystroke is detected, the position and speed of each finger is recorded. During a brief training phase, keystrokes are mapped to commands. For example, keystroke gestures can be used to select among several interface tools, such as a paintbrush, eraser, or selection tools which can be used to perform actions in the virtual environment.

References

AHMAD, F., AND MUSILEK, P. 2006. A keystroke and pointer control input interface for wearable computers. In *Proceedings of Fourth IEEE International Conference on Pervasive Computing and Communications (PerCom'06)*, IEEE, 2–11.