

Doctor-Computer Interface using Gestures

Juan Wachs¹, Helman Stern¹, Yael Edan¹,
Michael Gillam², Craig Feied², Mark Smith² and Jon Handler²

¹Department of Industrial Engineering and Management, Ben-Gurion University of the
Negev, Be'er-Sheva, Israel, 84105,
{juan, helman, yael}@bgu.ac.il

²Institute for Medical Informatics, Washington Hospital Center, 110 Irving Street, NW,
Washington, DC, 20010,
{feied,smith,handler,gillam}@medstar.net

1 System Overview

The motivation of this system is the sterile control for the interaction between doctors/surgeons and medical imaging systems. A web-camera placed above a screen (Fig. 1) captures a sequence of images of the hand. The hand is tracked by a tracking module which segments the hand from the background using color and motion cues. This is followed by black/white thresholding and various morphological image processing operations. The location of the hand in each image is represented by the 2D coordinates of its centroid. The system is implemented through a two level architecture. The lower level, named Gestix, provides tracking and recognition functions, while the higher level, named Gibson, manages the user interface.

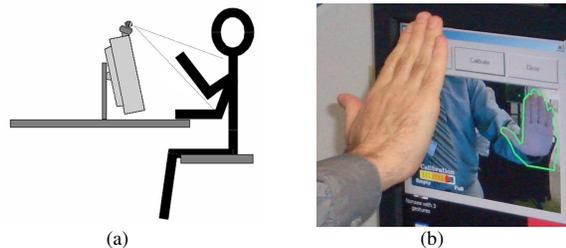


Fig. 1. Gesture capture system

The Gibson image browser is 3D visualization medical tool that enables examination of images, such as CT scans and X-rays. To interface the Gestix gesture recognition routines with the Gibson system, information such as, the centroid of the hand (used for image navigation controls), its size representing zoom, and orientation for rotation angle, are used to enable screen operations in the Gibson GUI. Both the gesture interface and the Gibson image browser are embedded in ActiveX controls which are communicated using messages and windows events.

2 Hand Tracking and Operation Modes

Gesture operations are initiated by a calibration procedure in which a skin color model of the users hand is constructed. Control between dynamic gestures used for browsing through images and pose gestures (used for rotation and zoom) are affected by mode switch gestures. Superimposed over the image is a rectangular frame. The area inside the frame is called the "neutral area". Movements of the hand across the boundary of the rectangle constitute directional browser commands. When a doctor decides to perform a specific operation on a medical image, he/she places the hand in the 'neutral area' momentarily, and an attention window event is called. The spatio-temporal information and other attributes of the posture are sent to a "mode detector" to determine whether a zoom or rotation pose gesture is presented, see Fig 2. Interaction is designed in this way because the doctor will often have his hands in the 'neutral area' without intending to control the Gibson data browser. Only when a flick gesture is moved towards one of the four quadrants (left, right, up, down), is the image browser moved in the direction of the flick, see Fig 3.

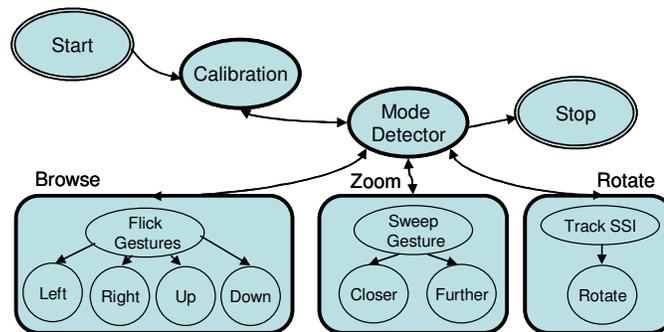


Fig. 2. State machine for the gesture-based medical browser

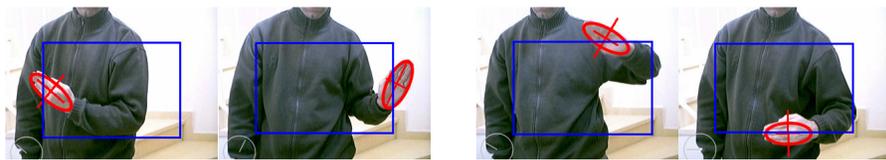


Fig 3. Four quadrants mapped to browsing operations

3 System Implementation

The software requirements of Gestix are Windows XP with Visual Studio 6.0 installed. Hardware requirements include a regular USB webcam connected to a Laptop with an Intel Pentium processor (1.8 GHz) and 500 Mb RAM. Artificial fluorescent light is recommended.