

LITERATURE SURVEY

Abstract

The use of doctor-computer interaction devices in the operation room (OR) requires new modalities that support medical imaging manipulation while allowing doctors' hands to remain sterile, supporting their focus of attention, and providing fast response times. This paper presents “*Gestix*,” a vision-based hand gesture capture and recognition system that interprets in real-time the user's gestures for navigation and manipulation of images in an electronic medical record (EMR) database. Navigation and other gestures are translated to commands based on their temporal trajectories, through video capture. “*Gestix*” was tested during a brain biopsy procedure. In the in vivo experiment, this interface prevented the surgeon's focus shift and change of location while achieving a rapid intuitive reaction and easy interaction.

Introduction

In this paper, we suggest the use of hand gestures as an alternative to existing interface techniques, offering the major advantage of sterility. Even though voice control also provides sterility, the noise level in the operating room (OR) deems it problematic. In this work we refer to gestures as a basic form of non-verbal communication made with the hands. Psychological studies showed that young children use gestures to communicate before they learn to talk. Manipulation, as a form of gesticulation, is often used when people speak to each other about some object. Naturalness of expression, non-encumbered interaction, intuitiveness and high sterility are all good reasons to replace the current interface technology (e.g., keyboard, mouse, and joystick) with more natural interfaces.

Overview

In two brain surgeries at the Neurosurgery OR at the Washington Hospital Center, procedures were observed by the authors to gain insights about the use of current technologies and how they affect the quality of the surgeon's performance. We found that: (a) surgeons kept their focus of attention between the patient and the surgical point of interest on the touch-screen navigation system; (b) a short distance between the surgeon and the patient was maintained during most of the surgery; (c) the surgeon had to move close to the main control wall to discuss and browse through the patient's MRI images.

The hand gesture control system “*Gestix*” developed by the authors helped the doctor to remain in place during the entire operation, without any need to move to the main control wall since all the commands were performed using hand gestures.

Tracking Algorithm

After a short calibration process, where a probability colour model of the doctor's hand is built, images of the surgeon's hand gesturing are acquired by video-camera and each image is back-projected using a colour model. The hand is then tracked by an algorithm which segments it from the background using the colour model back-projection and motion cues.⁷ This is followed by black/white thresholding, and a sequence of opening and closing morphological operations resulting in a set of components (blobs) in the image. The location of the hand is represented by the 2D coordinates of the centroid of the biggest blob in the current image.

“Gibson” Image Browser

The “*Gibson*” image browser is a 3D visualization medical tool that enables examination of images, such as: MRIs, CT scans and X-rays. The images are arranged over a multiple layer 3D cylinder. The image of interest is found through rotating the cylinder in the four cardinal directions. To interface the gesture recognition routines with the “*Gibson*” system, information such as the centroid of the hand, its size, and orientation are used to enable screen operations in the “*Gibson*” graphical user interface.

Hand Tracking and Operation Modes

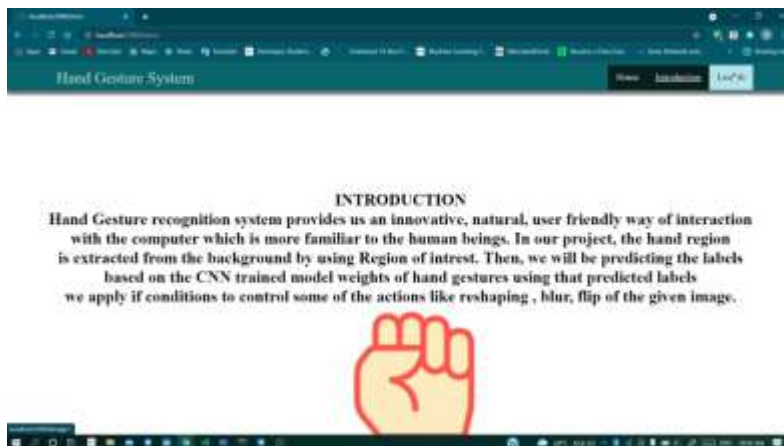
Gesture operations are initiated by a calibration mode in which a skin colour model of the user's hand or glove, under local lighting, is constructed. In a browse mode, superimposed over the image of the camera's scene is a rectangular frame called the “neutral area.” Movements of the hand across its boundary constitute directional browser commands. When a doctor/surgeon wishes to browse the image database, the hand is moved rapidly out of the “neutral area” toward any of four directions, and then back again. When such a movement is detected, the displayed image is moved off the screen and replaced by a neighbor image.

To evoke a zoom mode, the open palm of the hand is rotated within the “neutral area” clockwise/counter clockwise (zoom-in/zoom-out). To avoid the tracking of unintentional gestures, the user may enter a “sleep mode” by dropping the hand. To re-arouse the system the user waves the hand in front of the camera. The selection of these gestures was designed to be intuitive, expressing the “natural” feeling of the user. For example, the left/right/up/down gestures evoke the actions used to turn pages in a book left/right, or flip notepad pages up/down. The rotation gesture (zoom-in/zoom-out commands) reminds one of a radio knob to increase or decrease volume.

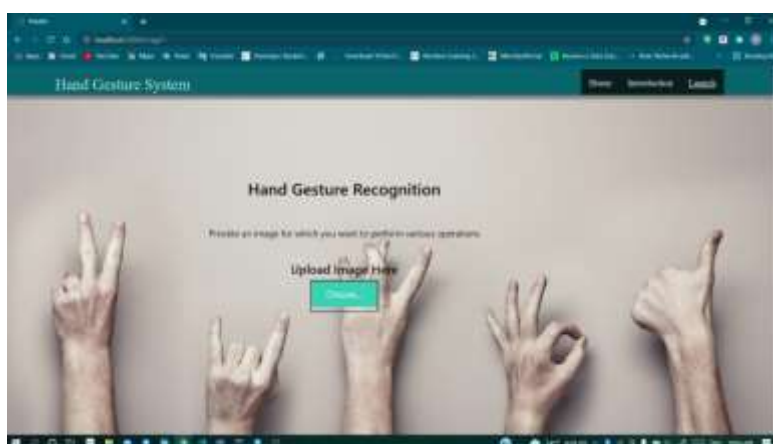
Home page



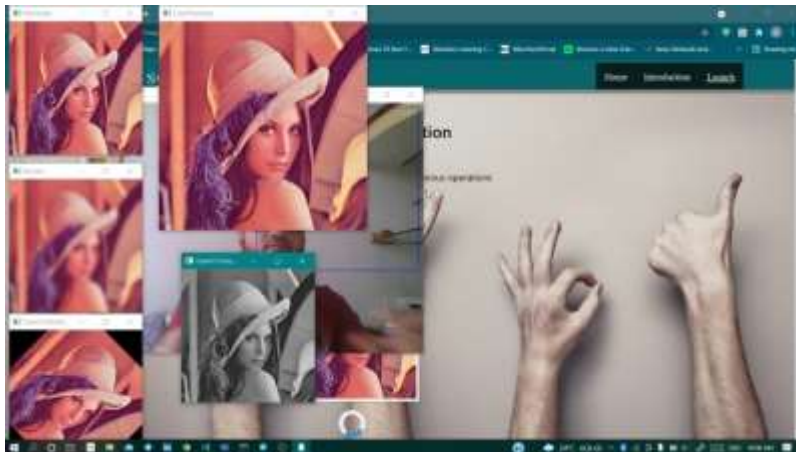
Introduction page



Model Launch Page



Predicting results using random image



Real time example



In this project we have used Convolutional Neural Network to first train the model on the images of different hand gestures, like showing numbers with fingers as 0,1,2,3,4,5. Then we made a web portal using Flask where user can input any image on which he wants to perform the operations. After uploading the image, our portal uses the integrated webcam to capture the video frame using OpenCV. The gesture captured in the video frame is compared with the Pre-trained model and the gesture is identified. If the prediction is 0 - then images is converted into rectangle, 1

- image is Resized into (200,200), 2 - image is rotated by -45° , 3 - image is blurred , 4 - image is Resized into (400,400) , 5 - image is converted into grayscale.

References

- 1) Wachs JP, Stern HI, Edan Y, et al. "Real-Time Hand Gesture Interface for Browsing Medical Images" *Int. J Intel. Comp. Med. Sci. Image Proc*
- 2) Graetzel C, Fong TW, Grange S, Baur C. "A non-contact mouse for surgeon-computer interaction," *Technol Health Care*
- 3) Smith KR, Frank KJ, Bucholz RD. "The NeuroStation- a highly accurate, minimally invasive solution to frameless stereotatic neurosurgery," *Comput Med Imaging Graph*