

Technical Brief ■

A Gesture-based Tool for Sterile Browsing of Radiology Images

JUAN P. WACHS, PhD, HELMAN I. STERN, PhD, YAEL EDAN, PhD, MICHAEL GILLAM, MD,
JON HANDLER, MD, CRAIG FEIED, MD, PhD, MARK SMITH, MD

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. Data from two usability tests provide insights and implications regarding human-computer interaction based on nonverbal conversational modalities.

■ *J Am Med Inform Assoc.* 2008;15:321–323. DOI 10.1197/jamia.M241.

Introduction

Computer information technology is increasingly penetrating into the hospital domain. A major challenge involved in this process is to provide doctors with efficient, intuitive, accurate and safe means of interaction without affecting the quality of their work. Keyboards and pointing devices, such as a mouse, are today's principal method of human—computer interaction. However, the use of computer keyboards and mice by doctors and nurses in intensive care units (ICUs) is a common method for spreading infections.¹ 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.²

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 technol-

ogy (e.g., keyboard, mouse, and joystick) with more natural interfaces.

This paper presents a video-based hand gesture capture and recognition system used to manipulate magnetic resonance images (MRI) within a graphical user interface. A hand gesture vocabulary of commands was selected as being natural in the sense that each gesture is cognitively associated with the notion or command that is meant to represent it. For example, moving the hand left represents a "turn left" command.

The operation of the gesture interface was tested at the Washington Hospital Center in Washington, DC. Two operations were observed in the hospital's neurosurgery department and insights regarding the suitability of a hand gesture system was obtained. To our knowledge, this is the first time that a hand gesture recognition system was successfully implemented in an "in vivo" neurosurgical biopsy. A sterile human—machine interface is of supreme importance because it is the means by which the surgeon controls medical information avoiding contamination of the patient, the OR and the surgeon.

Medical Gesture Interfaces

By the early 1990's scientists, surgeons and other experts were beginning to draw together state-of-the-art technologies to develop comprehensive image-guidance systems for surgery, such as the *StealthStation*.³ This is a free-hand stereo-tactic pointing device, in which a position is converted into its corresponding location in the image space of a high-performance computer monitor. In a setting like the OR, touch screen displays are often used, and must be sealed to prevent the buildup of contaminants. They should also have smooth surfaces for easy cleaning with common cleaning solutions. These requirements are often overlooked in the busy OR environment.

Many of these deficiencies may be overcome by introducing a more natural human-computer interaction mode into the hospital environment. The bases of human-human commu-

Affiliations of the authors: Department of Industrial Engineering and Management, Ben-Gurion University of the Negev (JPW, HIS, YE), Be'er-Sheva, Israel; Institute for Medical Informatics, Washington Hospital Center (MG, JH, CF, MS), Washington, DC.

The authors would like to thank Robert Irving, John Gillotte, and Alan Fischer for testing and software support. This work was partially supported by the Paul Ivanier Center for Robotics Research and Production Management, and by the Rabbi W. Gunther Plaut Chair in Manufacturing Engineering, Ben-Gurion University of the Negev.

Correspondence: Juan P. Wachs, Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Be'er-Sheva, Israel, 84105; e-mail: <juan@bgu.ac.il>.

Received for review: 02/07/07; accepted for publication: 01/22/08.



Figure 1. Gestix setup.

nication are speech, hand and body gestures, facial expression, and eye gaze. Some of these concepts have been exploited in systems for improving medical procedures. In *FAce MOUSE*,² a surgeon can control the motion of the laparoscope by simply making the appropriate face gesture, without hand or foot switches or voice input. Current research to incorporate hand gestures into doctor-computer interfaces appeared in Graetzel et al.⁴ They developed a computer vision system that enables surgeons to perform standard mouse functions (pointer movement and button presses) with hand gestures. Another aspect of gestures is their capability to aid handicapped people by offering a natural alternative form of interface and serving as a diagnostic tool.⁶ Wheelchairs, as mobility aids, have been enhanced as robotic vehicles able to recognize the user's commands through hand gestures.⁵

Method

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.

Architecture

The sterile gesture interface consists of a Canon VC-C4 camera, whose pan/tilt/zoom can be initially set using an infrared (IR) remote. This camera is placed just over a large flat screen monitor (Figure 1). Additionally, an Intel Pentium

IV, (600MHz, OS: Windows XP) with a Matrox Standard II video-capturing device is used.

A two layer architecture is used: In the lower level "*Gestix*" provides tracking and recognition functions, while at the higher level a graphical user interface called "*Gibson*" manages imaging visualization.

The Tracking Algorithm

After a short calibration process, where a probability color 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 color model. The hand is then tracked by an algorithm which segments it from the background using the color 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 color 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/counterclockwise (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. Dropping the hand (stop-tracking command) is associated to the idea of 'stop-playing', while the waving gesture ("wake-up" command) is associated with 'greeting a new person'.

Experiments

"Gestix" Setup

Prior to the initiation of the surgery, the "*Gestix*" system was placed in front of the main surgeon, midway between the

patient's bed and the main control wall. The VC-C4 camera was attached over a large flat screen. Before the use of "Gestix," a calibration process is conducted to capture a sample of the gamut of colors of the hand or surgical glove. The setup time for the whole "Gestix" system was approximately 20 minutes. Instead of a nurse assistant, the "Gestix" system was used for the biopsy planning procedure during the time that the biopsy results were being obtained in the lab.

Usability Tests

Three different types of usability tests were conducted with the "Gestix" system in the OR: (i) a contextual interview; (ii) an individual interview; and (iii) a subjective satisfaction questionnaire. The main result of the contextual interview which was based on watching and listening to the users while they work indicated that the main surgeon had to remain sterile, close to the patient, and avoid distractions (change in focus of attention), which could be life threatening. The main issues found in a 20 minute interview of the main surgeon were; (a) the need of replacing the plastic adhesive cover from the touch-screen monitor for every new surgery in order to keep it sterile; (b) the delay caused by frequent visits of the surgeon to the main control wall and to return to the patient's side; and (c) the surgeon preferred hand gesture control since it is based on interaction with hands, in which he/she is most proficient. At the end of the entire operation procedure, the main surgeon filled in a questionnaire to measure overall satisfaction about usability. This questionnaire was designed to measure satisfaction on a five point scale, similar to the ASQ created by Lewis.⁸ This questionnaire included questions on task experience, ease of task, task completion time, and overall task satisfaction. The surgeons found the "Gestix" system easy to use, with fast response and quick training times. Subsequent tests, using students, resulted in gesture recognition accuracy of 96% (for the eight gestures used in the system). Learning tests when tested by students for the eight gestures used and the learning tests required only 10 trials to converge with an average task performance time of 22 seconds.⁷

Discussion

A hand gesture system for MRI manipulation in an EMR image database called "Gestix" was tested during a brain biopsy surgery. This system is a real-time hand-tracking recognition technique based on color and motion fusion. In an in vivo experiment, this type of interface prevented surgeon's focus shift and change of location while achieving, rapid intuitive interaction with an EMR image database. In addition to allowing sterile interaction with EMRs, the "Gestix" hand gesture interface provides: (i) ease of use—the system allows the surgeon to use his/her hands, their natural work tool; (ii) rapid reaction—nonverbal instructions by hand gesture commands are intuitive and fast (In

practice, the "Gestix" system can process images and track hands at a frame-rate of 150 Hz, thus, responding to the surgeon's gesture commands in real-time), (iii) an unencumbered interface—the proposed system does not require the surgeon to attach a microphone, use head-mounted (body-contact) sensing devices or to use foot pedals, and (iv) distance control—the hand gestures can be performed up to 5 meters from the camera and still be recognized accurately. The results of two usability tests (contextual and individual interviews) and a satisfaction questionnaire indicated that the "Gestix" system provided a versatile method that can be used in the OR to manipulate medical images in real-time and in a sterile manner.

We are now considering the addition of a body posture recognition system to increase the functionality of the system, as well as visual tracking of both hands to provide a richer set of gesture commands. For example, pinching the corners of a virtual image with both hands and stretching the arms would represent an image zoom-in action. In addition, we wish to assess whether a stereo camera will increase the gesture recognition accuracy of the system. A more exhaustive comparative experiment between our system and other human-machine interfaces, such as voice, is also left for future work.

References ■

1. M. Schultz, J. Gill, S. Zubairi, R. Huber, F. Gordin, "Bacterial contamination of computer keyboards in a teaching hospital," *Infect Control Hosp. Epidemiol.* 2003;4(24):302-3.
2. A. Nishikawa, T. Hosoi, K. Koara, D. Negoro, A. Hikita, S. Asano, H. Kakutani, F. Miyazaki, M. Sekimoto, M. Yasui, Y. Miyake, S. Takiguchi, and M. Monden. "Face MOUSE: A Novel Human-Machine Interface for Controlling the Position of a Laparoscope," *IEEE Trans. on Robotics and Automation.* 2003;19(5):825-841.
3. K. R. Smith, K. J. Frank, R.D. Bucholz, "The NeuroStation- a highly accurate, minimally invasive solution to frameless stereotactic neurosurgery," *Comput Med Imaging Graph.* 1994;18: 247-56.
4. C. Graetzel, T.W. Fong, S. Grange, and C. Baur, "A non-contact mouse for surgeon-computer interaction," *Technol Health Care* 2004;12(3):245-57.
5. Y. Kuno, T. Murashima, N. Shimada, and Y. Shirai, "Intelligent Wheelchair Remotely Controlled by Interactive Gestures." *Proceedings of 15th International Conference on Pattern Recognition.* 2000;4:672-5.
6. T. Starner, J. Auxier, D. Ashbrook, and M. Gandy. "The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring". *Fourth Intl. Symp Wearable Comp.* 2000;87-94.
7. 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.* 2007;1(3):175-85.
8. Lewis J. R. Psychometric evaluation of an after scenario questionnaire for computer usability studies: The ASQ. *SIGCHI Bulletin.* 1991;23:78-81.