Gesture-Based Control of the Scanning Electron Microscope Using a Low-Cost Webcam Sensor

B Wong[1], BC Breton[1], DM Holburn[1] and NHM Caldwell[2].

[1.] Department of Engineering, University of Cambridge, Cambridge, United Kingdom
[2.] School of Science, Technology and Engineering, University of Suffolk, Ipswich, United Kingdom

This work forms part of our ongoing research into enhancing and improving microscopy and image processing technologies, and was one of a number projects undertaken to explore the potential of gesture-based interface control. This project focused on gesture-based control of a scanning electron microscope (SEM), Carl Zeiss model 1430VP, using a low-cost webcam as the sensor to detect and identify gestures.

With the inexorable advances in computing power and adoption of connected devices, there has been a surge of interest in developing gesture control for many devices and apparatus in a wide range of applications. A number of these have already achieved commercial success. For example, the Microsoft Xbox and the Nintendo Wii both use gesture controls for video gaming to enhance the user experience.

The objective of this project was to investigate the feasibility and practicality of developing a hand gesture control system for a scanning electron microscope and develop a proof of concept system. Apart from the on-site usage of the SEM by using a keyboard, mouse and joystick, there is nowadays the possibility of remote access to an SEM, where the user remotely connects to the instrument and controls it using standard mouse and keyboard commands, with images shown on a remote screen. Interacting remotely with a SEM through gestures may be more intuitive and can offer a better user experience.

Up to now, most implementations of gesture control have made use of hardware custom-designed for the gaming industry, such as the Microsoft Kinect and the Leap Motion Sensor. These two devices target different styles of interaction. Kinect focuses on capturing body pose, but there is limited software support for developers. Leap is a short-range hand-gesture capture device using a stereo infra-red camera, with a range of programming interfaces available to developers. Cater et al have described how the Leap Motion Sensor can be used to recognise a selection of gestures, and hence to allow control of a Scanning Electron Microscope [1]. However, they identified a trend in the industry towards the deprecation of "gestures" in the most recent Leap Motion software releases; this may mean that gesture control of software and instrumentation could be limited to emulation of touchscreen gestures scaled up from smartphone or tablet computers for the foreseeable future.

This project has taken a different approach and explores the possibility of developing a real time hand gesture-control system for SEM or similar applications using low-cost hardware, based on an ordinary webcam of the type commonly used for web communications, video conferencing or security, coupled with computer vision technology implemented using an open-source vision software library.

The programming approach taken was based on the use of the OpenCV computer vision library [2], to discriminate the subject's hand from the background, and to recognise gestures by extracting features of the hand and fingers using digital image processing techniques. In the first stage, to locate the subject's hand, skin segmentation is carried out. This is a challenging exercise, since skin appearance is affected

by ambient illumination and camera characteristics, as well as individuals' skin tone. To reduce the effects of illumination, the raw frames in RGB (red, green, blue) colour space are transformed to a cylindrical colour space that is less sensitive to fluctuations in illumination. A calibration technique was developed based on an array of sampling boxes in the field of view, over which the subject places the hand; the median values recorded in the boxes are used to determine appropriate segmentation thresholds, by means of which a binary frame showing only the hand may be produced. This image is subjected to morphological operations to remove unwanted noise pixels and improve the discrimination. In the following stage, a contouring operation is carried out to identify the perimeter of the hand and fingers. Finally, using a technique described by Sklansky [3], the binary representation is decomposed into a convex object associated with a set of so-called convexity defects, from which information about the disposition of the hand and fingers can be obtained.

A number of considerations needed to be taken into account when choosing gestures, some physiological, others related to the detection process. For a user facing the webcam sensor, hand gestures with palm out, hand down were awkward and inconvenient. By and large, users preferred to present the back of the hand, and to use the left hand to point to the right, and the right hand to point to the left. Secondly, hands and fingers should not overlap, or erroneous gestures may be recorded, leading to ambiguity. Finally, gestures involving one, two or four fingers were convenient and natural, whereas gestures involving three fingers were less so. The gesture control system developed was designed to control three key SEM parameters of the SEM: the focus, through adjustments to the objective lens current (up or down), magnification (up or down), and stage movement in the x-y plane of the SEM (up, down, left and right). Accordingly, the gestures chosen for control were distinguished by the numbers of fingers presented to the webcam, plus the orientation of the user's hand. These choices were arrived at by experiment and through consideration of the robustness of the gesture recognition process.

Task	Associated Gesture
Focus up / down	Present one finger up / down
Magnification up / down	Present two fingers pointing up / down
Move specimen up / down (Y axis)	Present four fingers pointing up / down
Move specimen left / right (X axis)	Present four fingers pointing left / right

Table 1: Mapping of SEM Tasks to Gestures

The developed system was able to recognise all eight designed gestures performed by a user in real time, with modest computational cost, easily within the reach of a desktop or laptop computer. Integration of the gesture control system with the SEM Application Programmer's Interface (API) developed by Carl Zeiss was also undertaken, allowing gesture-driven commands to control the microscope directly. The low-cost hardware gesture control system described has demonstrated proof of concept and shows great potential for development into a robust and advanced gesture control system for the SEM [4].

References:

[1] S Cater *et al*, Microsc. Microanal. 23 (Suppl 1) July 2017, p. 228.
[2] https://en.wikipedia.org/wiki/OpenCV (accessed on 15 February 2018).
[3] J Sklansky, IEEE Transactions on Computers, Issue 12, **21** (1972) p. 1355.
[4] Leap Motion is a trademark of Leap Motion, Inc. Kinect is a trademark of Microsoft Corporation.
This research was supported by funding from Carl Zeiss Microscopy.