doi:10.1017/S1431927621009685
Microsc. Microanal. 27 (Suppl 1), 2021

Voice Control of the Scanning Electron Microscope Using a Low-Cost Virtual Assistant

David Holburn[1], Bernie Breton[2], Tim Rowsell[3] and Cindy Lam[4]

[1]University of Cambridge, Cambridge, England, United Kingdom, [2]University of Cambridge, United Kingdom, [3]Rowsell Research, United States, [4]University of Cambridge, Hong Kong

This work forms part of our ongoing research into enhancing and improving microscopy and image processing technologies, and was one of a number of projects undertaken to explore the potential of voice-based interface control. This project focused on voice control of a scanning electron microscope (SEM), Carl Zeiss model 1430VP, using skills developed on a low-cost Virtual Assistant, specifically the Amazon Alexa family of devices [1], to receive commands and requests for information relating to the SEM.

With the inexorable advances in computing power and adoption of connected devices, there has been a surge of interest in developing voice control for instruments and apparatus in a wide range of applications [2]. While Alexa primarily targets domestic users, a number of business and other applications have now achieved commercial success. These rely on the voice-recognition and natural language understanding abilities common to all these devices. Devices with an integrated screen for display of images or movies give additional capabilities well-suited to the needs of imaging instruments. These are available for under £100.

The objectives of this project were to investigate the feasibility and practicality of developing a voice-controlled system for managing a scanning electron microscope and to explore its areas of application. There is considerable interest in the possibility of remote access to an SEM, where the remote user connects to the instrument and sees images shown on local screen. However, many of these systems are purely extensions of the standard instrument user interface and require considerable training. Interacting remotely with a SEM through voice control opens up the possibility for inexperienced users to get results using ‘smart’ automated procedures and can offer a better user experience.

The objectives were pursued through development of a voice user interface (VUI) skill for the Carl Zeiss SEM using the Amazon Alexa software development kit (SDK) and a variety of Alexa devices. The system comprises three main parts: (1) the Zeiss SEM with its server computer; (2) the Alexa device for use as the control instrument; and (3), the Amazon Cloud where Alexa skills and other relevant documents are stored. In order to achieve the objective, it was necessary to construct a complete network for data transmission between the different parts of the system.

Where a Personal Assistant such as Alexa is used to control a separate piece of equipment, a standard protocol is generally used. When a voice command is given, the Alexa device sends the request to the Amazon cloud. The command is identified, and a request sent by a secure protocol to the designated domain name corresponding to the device to be controlled or interrogated. The equipment responds appropriately, and may send back a report containing requested values or notifications via the Cloud, once the command has been executed. These may be sent to the requesting Alexa device to allow it to build and issue a speech response. When the instrument being controlled is as complex as an SEM, a number of other factors come into play. Firstly, interactions with the SEM may consume significant time. Parameters need to be checked to confirm consistency with safe operation. Although many SEMs are connected to a computer network, most are placed behind a secure network firewall. Accepting commands direct from the Amazon Cloud to control the instrument would be unacceptable in most cases, on grounds of security. In the present project, several challenges such as these had to be overcome, and a number of adaptations made to the protocol to achieve a working system. A proprietary network relay infrastructure had to be installed to ensure that only commands from specified sources were routed to the SEM. A web service based on the Flask development framework [3] was created to intercept and validate incoming commands. A password authentication system was devised to ensure that only authorized users could issue commands.

The Flask web service on the SEM computer accepts commands from the Amazon Cloud via the relay service, and arranges for the necessary calls to be made via the SEM Programmer's Interface (API). This returns requested values or notifications to the Amazon Cloud once the command has executed. After every command execution, a copy of the current SEM image is automatically captured and uploaded to the Amazon Simple Storage Service (Amazon S3) [4]. The Alexa device can then access the image by means of a standard object URL. These data are then used with the Alexa SDK to build up the speech response and Alexa device display. A special technique was devised to ensure the Alexa device would not break the connection during long operations such as moving the stage or carrying out alignment. The coding for the project was done in Python 3. In addition, suitable utterances intended to invoke a skill or functions within a skill were devised and coded using the Amazon online developer console. Figure 1 shows examples.

In operation, the voice-controlled system readily allowed microscope parameters to be interrogated and changed, and a range of commands to be given and the result seen or heard, totally hands-free. Covid-19 lockdown provided an unexpected opportunity to evaluate the system, when one co-author was confined at home in Hong Kong. In these circumstances standard domestic network connections were found to be perfectly adequate to operate the microscope satisfactorily.

This project has successfully demonstrated the feasibility of creating a low-cost voice control system for the SEM that allows a wide range of SEM commands to be executed, and with a suitable Assistant device, the resultant image can be shown on the screen. The system shows considerable potential for remote access to an SEM for many purposes. The underlying methodology could readily be applied to other classes of instrument.

Task	Associated Utterance
Get/set magnification	What is the magnification? Set the magnification to 5000.
Adjust brightness, contrast etc	Increase brightness. Decrease brightness.
Stage control	Move image to left. Move image up.
Auto functions	Auto focus. Auto align. Save image to cloud. Turn on the beam.

Figure 1. Figure 1: A small sample of SEM Tasks and the associated verbal commands

References

[1] Amazon.com Inc. 2020, Amazon Alexa official site, https://developer.amazon.com/en-GB/alexa - viewed February 25, 2021.
[2] Poongothai, M., K. Sundar, and B. Vinayak Prabhu. 2018, "Implementation of IoT based Intelligent Voice Controlled Laboratory using Google Assistant." International Journal of Computer Applications (0975 – 8887), Volume 182 - No. 16.
[3] Flask web development, one drop at a time, https://flask.palletsprojects.com/en/1.1.x/ - viewed on February 25, 2021.
[4] Amazon S3, https://aws.amazon.com/s3/ - viewed on February 25, 2021.

https://doi.org/10.1017/S1431927621009685 Published online by Cambridge University Press