

Development of the Voice Recognition Function SMART Voice

Shimadzu Corporation, Medical Systems Division
Yumi Tsutsumida

1. Introduction

The Trinias™ series angiography system with SCORE™ Opera provides the latest treatment environment in the ever-evolving field of medical technology and practice by offering the application platform SCORE Link. Recently, we have developed a voice control feature, SMART Voice, as one of the applications of SCORE Link, and we would like to introduce its details.

1.1 SCORE Link

SCORE Link is a subscription service that allows users to access the latest applications under a fixed-term contract (renewed annually). Software updates are conducted upon contract continuation and renewal, enabling users to benefit from the latest applications and software enhancements. By selecting and renewing contracts from multiple packages each year according to usage and frequency, users can utilize the latest applications while minimizing costs.

1.2 SMART Voice

The angiography system typically requires operators to control touch panels or consoles to select images, save fluoroscopy, and change acquisition modes. Consequently, operators have had to divert their gaze from the monitor to perform these operations or request assistance from other medical personnel. SMART Voice enables the operation of the angiography system through voice recognition by speaking into a microphone (**Fig. 1**), thereby supporting labor-saving and efficiency improvements in the medical field. The developed SMART Voice supports both Japanese and English, recognizing 13 types of commands related to image saving, control reference images, and activate the applications (**Table 1**).



Fig. 1 Voice Control Feature: SMART Voice

Table 1 13 Commands Currently Supported by SMART Voice

Save Image	Save Fluoroscopy
	Direct Fluoroscopy Save
	Save last Image
	Register Reference image
Activate Application	StentView
	StentShot
	SCORE Chase
Reference image control	Previous/Next Image
	Previous/Next Frame
	Play
	Stop

2. Features of SMART Voice

2.1 Basic Operations and Display

In developing SMART Voice, we added a microphone icon to the information display area of Live monitor and a recognition status icon displayed on both Live monitor and reference monitor (**Fig. 2**). Live monitor shows real-time images displayed during fluoroscopy or acquisition, while the reference monitor displays and allows control of previously fluoroscoped images. During procedures, the system captures audio within the examination room. The audio capture status is displayed as a volume indicator bar in the information display section of Live monitor. If the voice volume is too

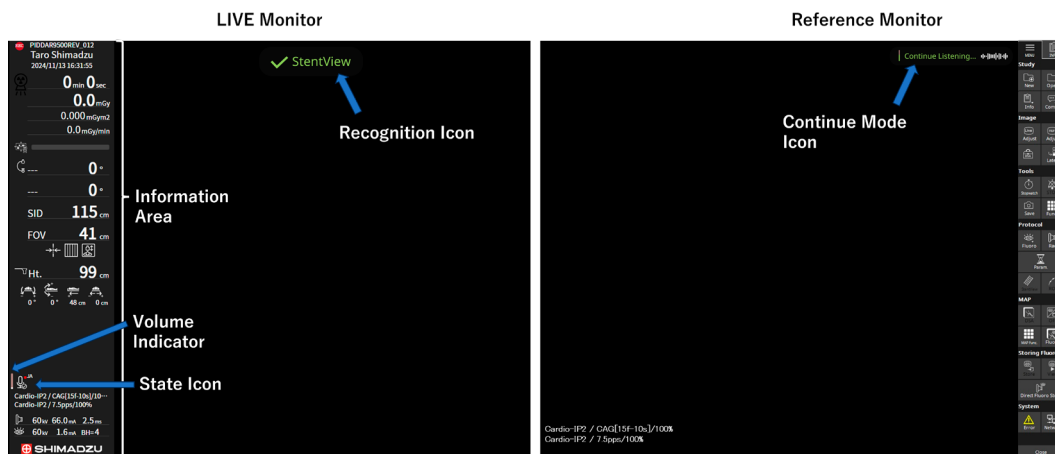


Fig. 2 Icons Displayed During SMART Voice Use

low, it appears white; if at an appropriate level, it appears green; and if too loud, it appears red (Fig. 3).



Fig. 3 Volume Indicator Bar During Audio Capture

SMART Voice uses "Trinias" as its wake word, a special word that activates the voice recognition feature. When the operator calls out "Trinias" to the microphone, a "Listening" icon appears on Live monitor. While this icon is displayed, the system is in command reception mode. During this time, if a command is spoken, the recognized command will be displayed, and the corresponding operation will be executed. The command reception mode automatically ends after a certain period, preventing unintended command recognition at inappropriate times (Fig. 4).

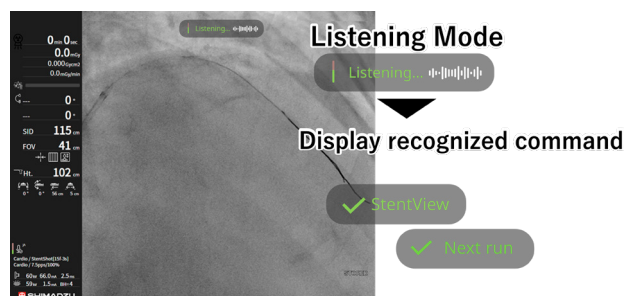


Fig. 4 Changing Recognition Status to Prevent Misrecognition of Commands Based on Mode

2.2 High Responsiveness

Since the angiography system is used during interventional procedures, it is crucial that operators can perform actions at the speed they require. This aspect was particularly emphasized in the development of SMART Voice. This feature is realized through the voice recognition and command operation components. By employing GPU-based voice recognition, the voice recognition component minimizes interference from ambient noise and achieves rapid voice recognition. Additionally, it can recognize audio accumulated over a certain period, ensuring that commands spoken before entering command reception mode are smoothly recognized.

2.3 Image Saving

Previously, the functions for saving fluoroscopic images and registering reference images were executed via mouse operations on the reference screen and through touch panel operations on the touch panel console (SMART Touch). Now, these functions can also be performed using SMART Voice. When an operator wishes to record a fluoroscopic image, they can say "Save fluoro / Store fluoro" to capture the most recent fluoroscopic image. To reserve the next fluoroscopic image for saving, the operator can say "Save next fluoro / Store next fluoro." By saying "Save last image / Store last image," the last frame of the most recent fluoroscopic image can be recorded. If the operator wants to set the captured image as a reference image, they can say "Save reference image / Store reference image / Add reference image" to register the currently displayed frame (Fig. 5).



Fig. 5 Fluoroscopy Saving and Reference Registration Operable via Voice Recognition

2.4 Application Activate

It is now possible to activate three applications that were previously executed via operations from SMART Touch or the reference monitor (**Fig. 6**). If an operator wants to emphasize the stent, they can say "StentView" to activate the SCORE StentView application, which supports real-time positioning of devices such as stents during PCI. To further emphasize the placed stent, the operator can say "StentShot" to activate the SCORE StentShot application, which can be used to evaluate the expansion and shape of the stent post-placement. If the operator wishes to check the entire lower limb, they can say "SCORE Chase" to activate the SCORE Chase application, which stitches together images captured in a single contrast injection. After each application is activated by voice, operators can get the image using the application immediately.

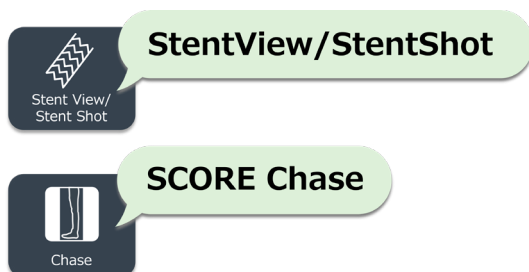


Fig. 6 Application Activate Also Possible via Voice Control

2.5 Image Control

Image selection, which previously required mouse, keyboard, or SMART Dial use, can now also be performed via voice commands (**Fig. 7**). By saying "Next image" or "Previous image," the operator can display the previous or next images on the configured monitor. Saying "Previous frame" or "Next frame" switches the displayed image

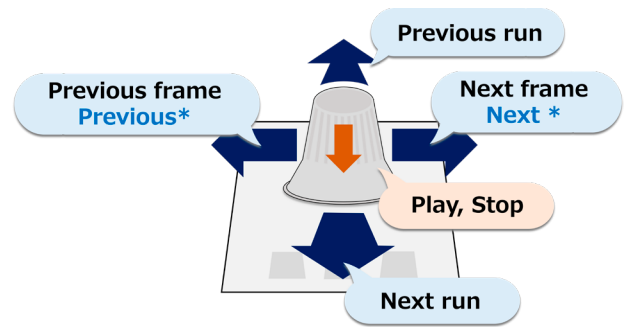


Fig. 7 Functions Frequently Executed, Such as Image control, Can Be Executed Without Saying "Trinias"

frame. The commands "Play" and "Stop" allow the operator to play or pause the dynamic image of the displayed frame. Since these image operations are often executed repeatedly, once the first execution is completed, the "Continuous Listening" icon (**Fig. 8**) will be displayed for a certain period, allowing for consecutive commands without having to say the wake word "Trinias." Additionally, during this continuous mode, the commands "next frame" and "previous frame" can be shortened to "next" and "back," allowing for smoother operation.



Fig. 8 Icon Displayed During Continuous Mode

3. Conclusion

We have developed the SMART Voice control function to create a stress-free procedure environment. This functionality allows operators to perform system operations such as image control, application activate, and fluoroscopic image saving without diverting their gaze from the patient or monitor. Moving forward, we will continue to develop additional voice commands based on customer feedback to enhance system usability and provide the latest procedure environments within the SCORE Link framework. We hope that these latest technologies will be of benefit to everyone.