

VoiceScan Technology In Clinical Practice

By Kristin Bravo, Gloria Goguen, RDMS, RVT,
Brooks Hawley, Cindy Owen, RDMS, RVT
Michael Washburn

In 2003, GE Healthcare was the world's first to introduce VoiceScan, an advanced technology that provides voice-activated control of GE's LOGIQ 9 ultrasound system. With a wired or wireless headset and the latest command-and-control speech recognition technologies, users can interact with the LOGIQ 9 during exams to perform more than 150 functions. VoiceScan uses intuitive voice commands which gives users the freedom to perform other important scanning activities. Studies have proven that VoiceScan eliminates more than 60% of manual button commands and streamlines workflow, especially for difficult to scan patients. By incorporating VoiceScan into their routine, users can perform their exams with greater ease, speed and comfort.

Speech Recognition Technology

Research in the area of speech recognition technology began in the early 1930s, but the technology did not become commercially feasible until recently. Today, the speech recognition technology is used in many industries and touches many aspects of our everyday lives. Automated telephone and banking systems, computer games, educational software and dictation systems alike provide hands-free control and convenience. By combining the latest in speech recognition technologies and GE's expertise in Ultrasound, VoiceScan was developed to make scanning a much better experience for clinicians, sonographers and patients.

At the heart of VoiceScan is a speech recognition engine that takes audio input via a wired or wireless microphone and translates it into a text string that the speech-aware application understands. This can be a complex task. Unlike the human brain, a computer does not have the advantage of understanding speech in context, or being able to use voice tone, body language or facial expression to help decipher the word meaning. However, it does have the ability to do rapid computations and comparisons.



For the computer to process the captured raw audio input, the signal must be converted from analog to a binary signal that the software can process. This is referred to as the digitization phase. This digital signal not only contains the spoken words, but also the noise within the environment. The computational engine works to filter out the noise so the remaining signal is a digital representation of the spoken words. This digital signal must be broken down further into the individual sounds or "phonemes" that make up speech. This is done by converting the digital signal into a frequency representation and then comparing the "frequency signatures" of various phonemes to the content of the signal. Once the individual sounds are identified they are assembled to create words. The group of sounds is compared to a dictionary or "grammar" of words in order to find a match. As you might expect, many words have the exact same sounds but different spelling and meaning. For example, "two", "to" and "too" can only be determined correctly when put into context of the words around them. All of these steps need to be done fast enough to keep up with the user's speech. Because of this, advancements in the computational speed of computers play a key role in the performance of speech recognition by enabling more sophisticated algorithms for identifying and combining phonemes.

There are two types of speech engines commonly used with speech recognition applications:

dictation engines and command-and-control. A dictation engine, like those used by physicians to record patient notes, records and writes every word to a word processing document. The grammar in a dictation engine is usually restricted to a specific domain such as the grammar contained in a medical dictionary. The speech engine used in the LOGIQ 9 is command-and-control. In this case, the user is allowed to speak using words or phrases from a list that the system is “trained” to hear. This allows the user the flexibility to speak naturally when giving the command. Unlike command-and-control, other speech engines require users to first say a key word such as “command” for example, “command print” or “command freeze”, in order to initialize speech recognition. This can disrupt the natural flow of speech for the user.

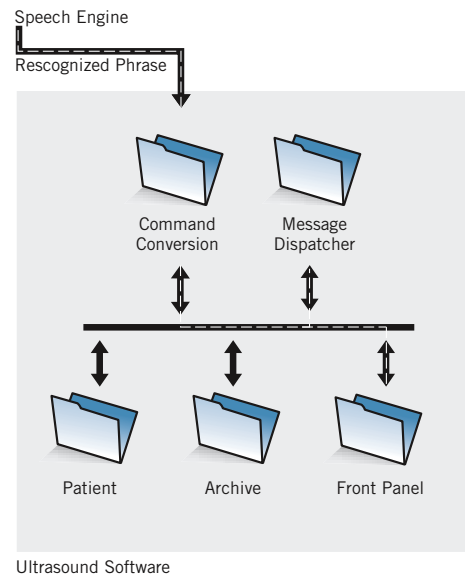
Breaking the speech into individual sounds and identifying those sounds is the core of any speech recognition engine. While there are many common traits across the frequency signatures of sounds, there are obviously differences from person to person such as accents, speed variation, pitch volume, and inflection.

Training a speech recognition system helps the software to better know the frequency signature of a particular individual sounds so that they can be more accurately identified. First time users of VoiceScan create their own voice signatures by reading a short story which the system records. With this recording, the system computes a statistical average of multiple samples of the same word and stores the averaged sample as that users voice profile. The voice profile is stored by the system as a user preset. Every time users log on, the system accesses this personal voice profile to improve recognition accuracy.

In addition to the voice profile, each speech-aware application has a pre-defined set of grammar that the engine references to help in the recognition process. The LOGIQ 9 grammar language has a pre-defined list of commands from which users can choose including “freeze”, “print” and “split screen”. In order to control the system, the user must provide voice commands that are specific to the ultrasound domain. For example, “gain up four” is a typical command that will increase the gain of the image in the

current mode. When spoken, the command is instantly captured and filtered by the speech engine, which searches through its grammar list and finds the command that is the best match. The command is then converted into text and passed on to the LOGIQ 9’s command conversion software instantaneously.

Figure 1. VoiceScan software implementation



The command conversion software process (shown in Figure 1) takes the recognized phrase, determines the desired action, parses the phrase, and identifies that the command is “gain” with the parameters “up” and “four”. The command conversion software creates a command based on multi-parameter information, and depending on the type of command, it dispatches it to the appropriate software module. In this case, the “gain” command is handled by Front Panel module, which receives the information and increases the image gain by four.

In a command-and-control system like GE Healthcare’s VoiceScan, the grammar is important because it allows the use of natural speech commands while differentiating one command from another. However, in Ultrasound it is common for the person scanning to say the word “breathe” to the patient during an exam. And the word “freeze” is a natural command for telling the ultrasound system to stop scanning. Without special consideration, saying “breathe” to the patient could result in the ultrasound

machine freezing the image. By putting “breathe” in the grammar, the command-and-control engine can decide if the command sounds more like “breathe” or “freeze”. The speech engine passes the result to the LOGIQ 9 and it stops scanning in the case of a “freeze” command and simply ignores a “breathe” command.

Clinical Applications

VoiceScan allows the user to control virtually every function available on the Ultrasound console. Currently the system can perform over 150 voice commands with a great degree of accuracy. Commands range from the obvious such as “freeze”, “print”, “cine” and “start scanning” to more complex commands such as an annotation “type right prox CCA” or “vessel mid CCA” for assigning a measurement in the worksheet. This gives the user incredible flexibility, efficiency and in some cases even allows for better body mechanics while scanning.

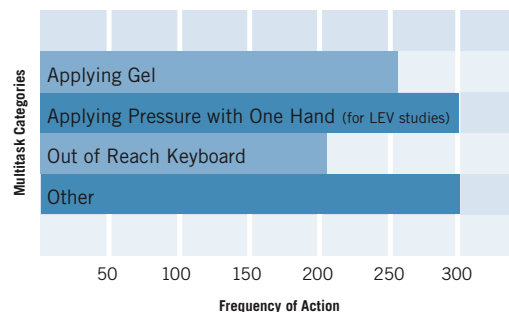
In addition to improved scanning ergonomics, VoiceScan allows the user to handle multiple tasks simultaneously. A prospective study conducted with Baptist Memorial Hospital - Desoto, a 200-bed hospital located in Southaven, Mississippi, identified a wide range of multitasking benefits with VoiceScan technology. The study involved 87 examinations, which indicated a total of 18 multitasking activities – 15 of them routinely performed. For all examinations it was found that on average, VoiceScan eliminated 60% of manual button pushes.

Technology in Practice

Input from clinicians, sonographers and patients played an important role in the development of GE’s VoiceScan technology. We continuously collect feedback from users and have cited comments from three customers that are representative of our customers’ satisfaction with the feature.

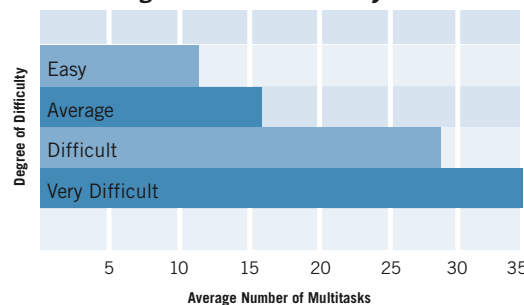
Michelle Brantman, RDMS of St. Francis Hospital, Wisconsin has been using VoiceScan for portable and venous studies. Brantman states, “I like the idea of being hands-free. I am not limited to standing near the unit. I can even be on the other side of the patient’s bed for things

Figure 2. Frequently used multitasks



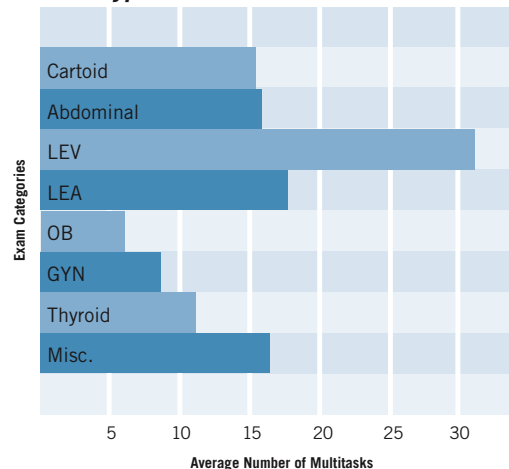
Of the 18 possible multitasks used by clinicians at Baptist, Figure 2 shows the most frequently performed multitasks.

Figure 3. Average multitasks used based on degree of exam difficulty



As one would expect, the more difficult the exam the more multitasking activities were performed. Figure 3 shows that as the degree of procedure difficulty increase, so does the benefit of multitasking capabilities with VoiceScan.

Figure 4. Average multitasks based on exam type



Of the variety of exams completed, it was shown that lower extremity venous scanning had the highest number of average multitasks performed. This is due, in part, to the fact that the sonographer typically uses both hands - one to hold the transducer, and the other to augment the patient’s leg. From the 87 procedures studies represented in Figure 4, multitasking with VoiceScan was done most often for lower extremity venous exams (LEV). Carotid, abdominal, and lower extremity arterial (LEA) exams also showed significant multitasking activity.

like portable exams. With VoiceScan I have the ability to do needed functionality without having to get help. It is most important when I don't have an extra set of hands." Brantman has also experienced the productivity and ergonomic benefits of VoiceScan, "VoiceScan saves me time, especially when doing venous Doppler exams. It also allows for better body mechanics, which can save the Sonographer the strain of reaching."

Donovan Schmitt of Inland Imaging in Spokane, Washington uses VoiceScan for all exams. Schmitt remarks, "I like it when doing DVT, shoulder exams (a dynamic study that requires use of both hands) and other musculoskeletal work. It is good for portables when you are not scanning under ideal situations and helps control system functionality when the Sonographer cannot reach the controls. It is good for invasive procedures where the Sonographer needs to scan on the other side of the patient, but can still control the system with VoiceScan." When asked what he likes most about the VoiceScan feature, Schmitt stated "It increases efficiency. I miss it when I work on other systems that do not have the VoiceScan capability."

Victoria Walker of Baptist Memorial Hospital Desoto in Southaven, Mississippi notes the many ergonomic benefits of VoiceScan. "It enables me to get into a position to optimize an image and not strain my body in the process. Overall, body ergonomics is better. Particularly, it helps me to enjoy my profession a bit more because I am not straining so much." Walker also remarked on the benefits physicians can enjoy from VoiceScan:

"When we go into a special procedure where physicians are in the room with us, they really like it because they can use the machine without having to know where the buttons are."

Patients were also excited about the use of this technology during their examinations. In a poll of 50 patients done at Baptist Hospital, none had objections to VoiceScan being used during their studies. In fact, 95% of the patients polled gave positive feedback about being scanned with a voice-activated system. Overall, patients were excited about the technology commenting that "this would help get patients in and out faster", "the technology is impressive" and "I think it was fun. This technology is wonderful."

Conclusions

For many years ultrasound users have inquired about the possibility of incorporating voice activation technology into ultrasound equipment. GE Healthcare responded to this request, becoming the first ultrasound company to bring this exciting technology to fruition.

By adding GE Healthcare's VoiceScan technology to the ultrasound exam routine, users can control up to 150 system functions, thereby reducing manual commands by approximately 60%. VoiceScan allows users to incorporate better body mechanics during scanning, gives them an extra hand for difficult exams, and allows for a high degree of multitasking. This new technology has truly revolutionized the way clinicians and sonographers approach ultrasound.

GE Healthcare

Waukesha, Wisconsin U.S.A.
www.gehealthcare.com

General Electric Company reserves the right to make changes in specifications and features shown herein, or discontinue the product described at any time without notice or obligation. Contact your GE Representative for the most current information.

© Copyright 2004 General Electric Company
GE Healthcare, a division of General Electric Company.

04-xxxx 10/04 Printed in USA

imagination at work

