

Construction Quality Control through Voice Recognition Technology

Luh-Maan Chang* and Tao-ming Cheng**

- * Associate Professor, Division of Construction Engineering and Management, School of Civil Engineering, Purdue University, West Lafayette, IN 47907**
- ** Ph.D. Candidate, Division of Construction Engineering and Management, School of Civil Engineering, Purdue University, West Lafayette, IN 47907**

ABSTRACT

Voice recognition technology is one form of the Automatic Identification system. This automatic data collection through voice has been successfully applied in many other industries and has the potential to be used in construction quality control. Using this technology, the operator can enter information simply by speaking the desired data while eyes and hands are free for other quality control tasks. This paper is to present the potential use of this technology in the construction industry. The process which adopts voice recognition technology for bridge painting quality control systems will be illustrated. In addition, this paper will explain the basics of the technology and provide information on implementing a voice recognition system.

1. INTRODUCTION

To assure that the owner receives a quality facility with expected service life from its contractor, an efficient quality assurance (QA) system is needed. However, it is not uncommon that a sophisticated QA system involves a lot of paper work. The quality engineers often spend a considerable amount of time to collect, record, and process data. Many times the paper work distracts the engineers from their major assignment of obtaining quality products. Therefore, a research project was conducted at Purdue University to explore the viability of voice recognition technology to minimize the paper work for steel bridge painting inspection¹.

The application of voice recognition in a bridge painting quality control program depicted in this paper is based on the painting quality acceptance system developed by Hsie and Chang² for Indiana Department of Transportation. This applied research is to construct a process using voice recognition to accomplish data collection of the painting inspection. This quality acceptance system requires an inspector to take a lot of samples and record the tested data on different painting stages. Most of the data are simple and similar, thus they can be input by voice easily and accurately. Moreover, the inspector's

hands and eyes are free while the data is recording. Hence, in order to get real-time data entry and to avoid human error of double data entry, the voice recognition technology was studied.

The purpose of this paper is to present the research results. The procedures for painting inspection on steel bridges will be described. The background of voice recognition technology will be provided. The software and hardware required for the application will be illustrated. Then, an example will be depicted following the comparison of advantages and limitations. Finally, a conclusion will be drawn before applications of the technology are pointed out.

2. PAINTING INSPECTION

In field painting, the procedure to inspect the painting quality of a steel bridge is generally divided into three stages; surface blasting, priming, and finishing. Each inspection stage needs to record the test data which are the thickness of paint or profile. The unit for thickness is recorded in mils. The inspector needs to measure the thickness of each beam in different stages. Each data set contains ten data in this study. For each beam, the inspector measures its first data set at ten randomly selected spots. If the value of the tested data is less than the required thickness, it is counted as defect. If the sum of the defect number for whole data set is less than or equal to one, the lot will be regarded as passing the inspection. If the total defect number is equal to two, the inspector needs to take a second sample and to measure the thickness. If the total defect number is more than two, the lot will be rejected. For the situation which requires a second sample to be taken, the inspector still needs to take ten thickness measurements along the beam. The rule for this situation is, first to count the twenty data together, then sum the total defect number. If the total defect number is more than four, the painting of the beam will be rejected and the work needs to be repainted, otherwise the beam passes the inspection². The inspection rules are depicted graphically in Figure 1.

Prior to painting operation, the inspector needs to check those resources necessary for the operation. A check list can be provided to remind the inspector to bring the appropriate instruments, tools, equipment, and so on. In other words, many things need to be well prepared before inspection.

3. BACKGROUND

Voice recognition is a way that computers use voice as a means of sending and receiving information. There are two phases of operation in the voice recognition system. One is the training phase and the other is the recognition phase. In the training phase, each voice input will be analyzed and stored digitally in the memory of the computer in order to establish the reference templates of the spoken word. Herein, statistical modeling process is used to select the appropriate form of probability distribution for a given word or specific sound. After the training phase has been completed, the voice recognition system can begin to operate. In the recognition phase, any voice input is analyzed to match the closest reference template previously stored in

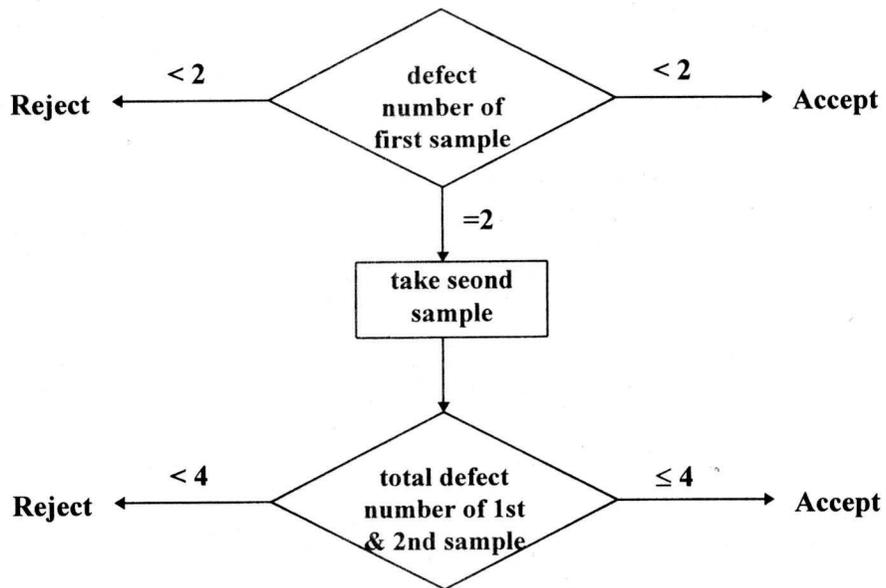


Figure 1 The decision tree for accepting or rejecting a lot

the training phase. After finding the closest template, the system can execute the requested task or output the corresponding text of voice on the PC monitor. The voice recognizing process is shown in Figure 2. Feature extraction captures the acoustic elements, then the endpoint detection detects the beginning and the end of the unknown word or phrase. Finally, the system searches the closest reference template and outputs the best match³.

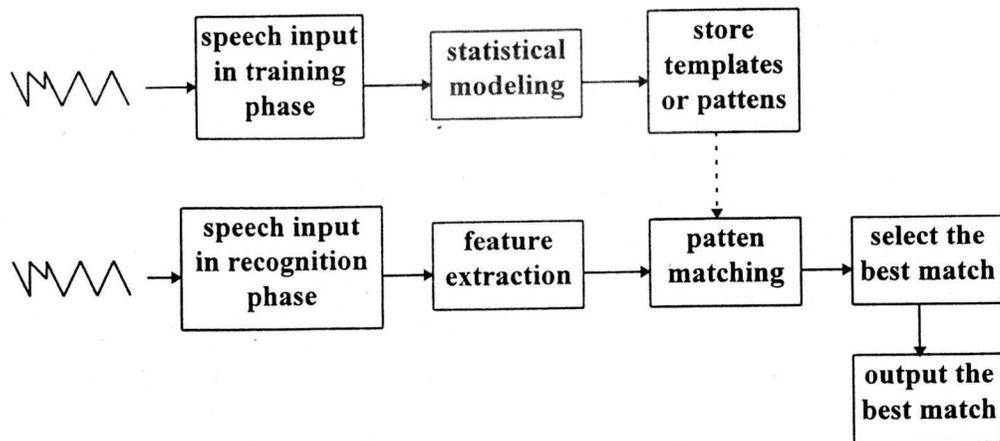


Figure 2 Voice recognition process

Voice recognition systems can be broken into two types: Speaker-Dependent Recognition (SDR) and Speaker-Independent Recognition (SIR). For SDR, the system must be trained to be familiar with the user's voice pattern. It recognizes the voice of one who trained the computer.

Unlike SDR, the users don't need to train the system to recognize their voice. "SIR allows a system to recognize a fixed set of words from a wide range of speakers⁴." Theoretically, SIR can understand any human speaker. In SIR systems, the computer requires collection of many voice samples or patterns and learns to ignore personal voice characteristics. Then, the computer can recognize and respond to words spoken in isolation or in continuous sentences. Hence, there must be a large data base of voice samples stored in templates. The computer compares voice input with voice samples and matches the words, phrases, or sentences. However, it is difficult for the computer to recognize different people's voice patterns because everyone has his/her different way of saying or pronouncing the same word. For this reason, the SIR system usually has less accuracy in recognizing the voice pattern than the SDR system does. Therefore, for SIR to develop the capability of recognizing different operators' utterance is still the greatest challenge to the manufacturers of the voice recognition systems^{4,5}.

4. IMPLEMENTATION

Since the purpose of this research is to apply voice recognition in bridge painting quality control, all the inspection process is computerized and combined with the voice recognition system. A demonstrated program is developed based on two software packages, Quattro-Pro (spreadsheet) and Verbex voice recognition system. The painting inspection forms are first computerized into a spreadsheet format. By means of voice recognition, the user is allowed to enter commands or data into the spreadsheet by normal voice .

The program is designed to allow the user to input data by voice following a series of instructions through voice synthesis. The user can hear the prompt from the computer, then follow the instruction from the system and speak the commands or the data to record the tested results. The system is also capable of repeating any command or data which it "heard" from the operator, thus the operator can get the response from the system to check the computer whether it receives the correct commands or not. After the user finishes his/her data entry, the tested results will be spoken out automatically by the computer. Therefore, the user can keep his/her eyes and hands free all the time.

5. SYSTEM DESCRIPTION

Hardware: A 80486 personal computer operating at 66MHz was used in this applied research project. The heart of the voice recognition system is the voice input module #0625 which comes from Verbex Voice System, Inc. The module has the ability to transfer the voice input into a digital message into the computer. The module #0625 contains two processors. One is the control processor which is an Intel 80286 CPU having 1MB memory, operating at 12.5MHz speed. The other is the recognition processor that

is a TM320C31 CPU which has 512 Kilobyte memory and Operating at 27MHz speed. The word which can be active during the voice recognition process is called vocabulary. Up to 2050 vocabularies can be stored in the control processor and up to 300 active vocabularies can be loaded into recognition processor. However, in this study there are only 112 vocabularies needed to be defined.

The headset contains earphones and microphone. Voice responses can be heard from the computer through earphones and voice commands can be sent to the computer through the microphone. A printer is used to print out the information which is saved in the computerized data collection forms. A portable narrow band Radio Frequency (RF) device is connected with the computer which helps the inspector record the data without wires.

Software: There were three types of software used by this application. The spreadsheet software Quattro-Pro is used for recording the data. The interfacing software, Softkey, provides the function which one can mimic the keyboard by voice. In other words, through Softkey all the action that one can do from keyboard can be done by voice. The development software include "V-Update" and "Convert". Convert allows one to compile one's own voice recognition grammar program which is generated from any editor and saved in ASCII code. Just as the English language is governed by a grammar which specifies the proper orders and patterns for forming sentences, the voice recognition grammar defines what is and what is not a proper statement⁶. According to the grammar rules of the voice recognition system, any statement which is not defined or improper, will be ignored by the computer. V-Update offers some commands that can work under the DOS (Disk Operating System) environment to load the grammar file and voice pattern file into the recognizer.

Figure 3 describes the procedure of voice recognition used by this system. In order to let the system be ready for voice recognition process, the voice pattern file and grammar file must be loaded into the recognition processor through the V-Update software. The voice can be digitized after the recognition processor receives the voice command. However, the command must have been defined in the grammar and match the voice pattern saved in the voice pattern file. Then, the interfacing software Softkey gets the recognition results stored in ASCII code, and sends the messages to the keyboard buffer. In other words, through the recognition processor and Softkey, the voice generated commands are like the commands produced from the keyboard.

6. EXAMPLE

An example will be introduced in this section to demonstrate how to use the program. When the operator starts the computer system, the operator is asked to choose one of the three stages where he/she wants to record data. It is assumed that the operator chooses stage-II because primer inspection is scheduled to be done. Then, operator speaks "stage-II." Immediately, the computer screen will switch to the desired stage. There is a menu bar which lists "input", "save", "print", "menu", and "2nd-data" commands shown on the upper-left corner. Through the earhpones, the operator will hear the voice prompt which is "please enter the following commands: input, save, print, menu, or second-data." If the user would like to record data, the user can speak the

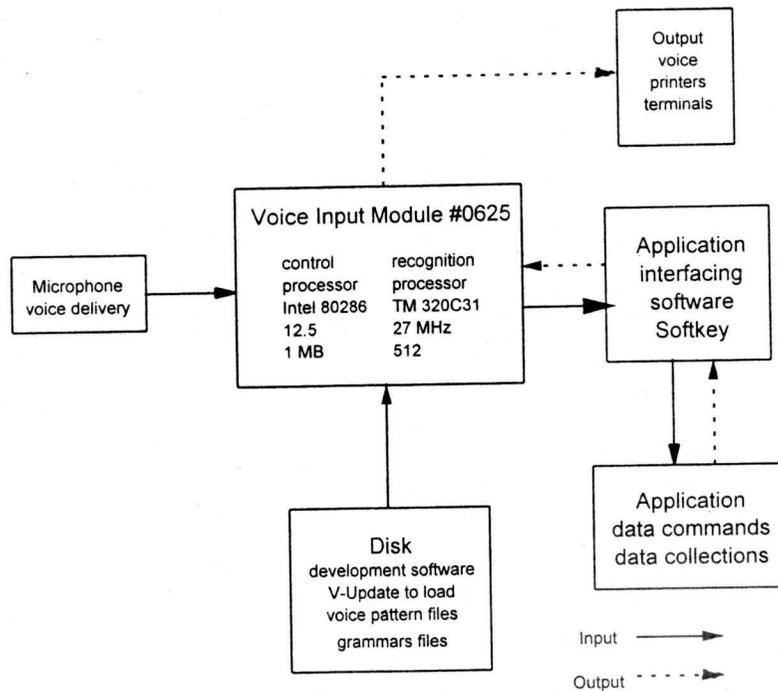


Figure 3 The procedure of the voice recognition of the painting quality control system

command "input" and then the spreadsheet prompt will jump into the right position ready for inputting data. Again, the computer will give the operator prompt "please enter the data one by one." Then, the user can speak the data one by one. After each data was spoken, the user could get the voice reconfirmation response from the system to prevent inputting the wrong data. For instance, the operator said "4.2" to record this data and the operator would hear "you just say 4.2" from the computer.

In this case, suppose that the entered 10 data are 2.6, 3.6, 5.7, 2.2, 3.5, 2.1, 3.0, 4.4, 3.6, and 4.0 respectively. According the rule of thumb, when there are two data values less than the require value, the operator needs to take another sample lot which also includes 10 data. If the required thickness for the primer painting is 2.5 mils, obviously, the first data set did not pass, so a second data set needs to be taken. Thus, the "judgement" was shown "take the second sample" immediately on the screen after the user finishing the input of data. At the same time, the computer would also speak "according to the rule, you need to take another 10 data." During this time, the screen has already been back to the ready state for waiting another command input. Thus, the operator can speak "second data" and then the spreadsheet would get into the position for recording the information. It was assumed that the data set of the second sample lot are 2.5, 3.6, 4.3, 5.0, 2.4, 6.4, 3.1, 2.9, 3.5, and 4.2. Now the total defect value of these two sample lots would be 3 (2.2 and 2.1 in the first data set and 2.4 in the second data set). Again, according to the rule, the total number of defected data were less than 4, then the judgement "Accept" would be spoken and shown on the screen.

With the data recording done, the menu bar would show up again and the system would wait the next command to be input from voice or keyboard. The dialogue in the example aforementioned is listed in Figure 4.

Q:	Please enter the follow commands: input, save, print, menu or second-data.
A:	Input
Q:	Please enter the required thickness.
A:	2.5
Q:	Please enter the beam number.
A:	I345
Q:	Please enter the data one by one.
A:	2.6, 3.6, 5.7, 2.2, 3.5, 2.1, 3.0, 4.4, 3.6, 4.4
Q:	You have already enter 10 data. According to the rule, you need to take another 10 data.
	Please enter the following commands: input, save, print, menu or second-data.
A:	Second-Data
Q:	Please enter the Data one by one.
A:	2.5, 3.6, 4.3, 5.0, 2.4, 6.4, 3.1, 2.9, 3.5, 4.2
Q:	You have already enter 10 data. According to the rule, the tested result is accept.
	Please enter the following commands: input, save, print, menu or second-data.

Figure 4 A sample dialogue of the example

7. ADVANTAGES AND LIMITATIONS

The advantages of the voice activated bridge painting quality control system developed in this research are described as follows:

- (1) The inspector does not need to have much computer experience, thus, it reduces the training on the system.
- (2) The inspector can enter the data by normal voice and get responses or prompts from the computer, therefore, his/ her hands and eyes are free and can concentrate on other tasks.
- (3) The voice data entry is in real-time; it eliminates recording after the inspector gets the data.
- (4) The system offers a wireless working environment.

Nevertheless, there are some limitations that may decrease the effect of adopting voice recognition technology for setting up the painting quality control system. First, the background noises from the field may affect the accuracy of voice recognition. These effects may highly restrict the use of voice recognition in some construction sites. Second, the user needs to train the computer to recognize his/her voice if the SDR system is used. Third, it takes a longer time for training the computer if the vocabularies required are larger. Fourth, a speaker's voice pattern will not only be affected by ambient conditions, but by the variations of the speaker's articulation and pronunciation, and the used signal transducers and microphones. Thus, while using the voice recognition system, if the

speaker's environment is not identical to the previous one used when he/she trained the system, it could influence the voice recognition accuracy and reliability.

8. APPLICATIONS

According to Rabiner, Juang, and Lee, Voice Recognition Technology has been used to⁷:

- (1) provide information or access to data or services over telephone lines from remote sites such as automatic call handling and limited banking services,
- (2) help user control office environments, interact with PC's/workstation, use softwares, enter preformatted forms, and take dictation,
- (3) aid in monitoring quality control on manufacturing assembly line and handling mail packages,
- (4) create various medical/legal reports and forms which contain repeatedly used technical terminology,
- (5) supply the voice control of wheel chair functionality for handicapped and game playing machines.

In the construction industry, there are a lot of hands-busy, eyes-busy, and repetitive operations which fit the application of the voice recognition technology; such as quality inspection, interface with PC's/ workstations, handling materials, control of equipment, and tools, form filling for contract administration, etc. In other words, there is indeed a great potential to transfer this technology into the construction industry.

Although the applications are very limited and in most cases are still experimental, the rapid development in signal processing, algorithmic methods for pattern recognition, computer architecture and hardware, will further advance voice recognition technology. It is foreseeable in the near future that technology will bring a much more simple, direct, natural voice communication between the user and the recognition system, enhance the system ability from speaker-dependent to speaker-independent, and supplement a self-adjusted system to adapt to changing speaker environments.

9. CONCLUSION

Voice recognition technology like bar coding or handwriting recognition technology is one form of automatic identifications. It fits to the situations when:

- (1) hand-free and eye-free operations are needed,
- (2) keypunch is a source of time delays and errors,
- (3) portability and real time communication are a concern⁸.

There is no doubt that voice recognition technology will find its place in the construction area. Voice recognition may be constrained by background noises which impair the recognition rate and decrease the effectiveness of adopting the voice recognition system in the field. However, compared with other methods for data entry, voice recognition system has the potential value to be developed.

Voice recognition technology can allow users to enter data by voice while they are using their hands and eyes, thus, increasing overall human potential. Compared with

other automatic identification technologies, voice recognition can be more powerful and useful for the construction industry in this regard.

The demonstrated program mentioned in this paper is just a prototype to show the possibility of applying voice technology in bridge painting quality control. Construction engineers can look forward to using this new technology to help them collect data and interact with machines; such as PC's, construction robots, manipulators, and facilitate real-time decision.

10. REFERENCE LIST

1. Chang, Luh-Maan and McCullough, Bobby G. (1992). "Implementation of the Developed Quality Acceptance System for Steel Bridge Painting Construction," Proposal For New Research Study, School of Civil Engineering, Purdue University, West Lafayette, Indiana.
2. Chang, Luh-Maan and Hsie, Machine (1995). "Developing Acceptance-Sampling Methods for Quality Construction," J. of Constr. Engrg. and Mgmt., ASCE, 121(2), pp. 246-253.
3. Lennig, M. (1990). "Putting Voice to Work in the Telephone Network," Computer, Vol. 23, No. 8, pp. 35-41.
4. Thompson, B. and Gottesman, K. (1988). "Voice Comes of Age," Instruments & Control System, Vol. 61, No. 12, pp. 29-31.
5. Award, Selim S., Wanger, David A., and Flaherty, Michael M., (1989). "A Voice Controlled Telephone Dialer," IEEE Transactions on Instrumentation and Measurement, Vol. 38, No. 1, pp. 119-125.
6. Verbex Voice System, Inc., (1992). "Grammar Development Manual," Edison, New Jersey.
7. Rabiner, L.R., Juang, B.H., and Lee, C.H., (1996). "An Overview of Automatic Speech Recognition," Automatic Speech and Speaker Recognition: Advanced Topics, edited by C. H. Lee, F. K. Soong, and K. K. Paliwal, Kluwer Academic Publisher, New York, pp. 1-30.
8. Rowings, James E., (1991), "Project-Controls Systems Opportunities," Journal of Construction Engineering and Management, ASCE, Vol. 117, No. 4, pp. 691-697.