Aqua-Acoustic Incorporated
Simon Fraser University
Burnaby, BC
V5A 1S6
aqua-acoustic@sfu.ca

February 16, 1999

Dr. Andrew Rawicz
School of Engineering Science
Simon Fraser University
Burnaby, BC
V5A 1S6

Re: ENSC 370 project *Voice Activated Control System Functional Specifications*

Dear Dr. Rawicz,

The attached document, *Voice Activated Control System Functional Specifications*, outlines the requirements for our project. Our project is to design a voice control system which can be adapted to control different devices. We will demonstrate the system by using it to control a telephone and a submersible in a fresh water tank.

This document lists the functional requirements of the system as well as the specifications for the components of the system. The components of the system are the signal acquisition unit, signal filtering unit, voice analysis unit, processor/controller, and a device interface unit sending information to each device the system controls.

Aqua-Acoustic Incorporated consists of four enthusiastic third-year engineering students: Scott Emery, Amy Lu, Sean Nicolson and David Peterman. If you have any questions or concerns about our proposal, please contact David Peterman by phone at 526-4724 or by e-mail at dpeterma@sfu.ca.

Sincerely,

Scott Emery Amy Lu Sean Nicolson David Peterman

Aqua-Acoustic Incorporated

Enclosure: *Voice Activated Control System Functional Specifications*

Aqua-Acoustic Incorporated

Voice Activated Control System Specifications

Submitted by	**Aqua-Acoustic Incorporated** **Scott Emery, Amy Lu, Sean Nicolson, David Peterman**
Contact	**David Peterman** **School of Engineering Science** **Simon Fraser University** dpeterma@sfu.ca
Submitted to	**Andrew Rawicz** **School of Engineering Science** **Simon Fraser University** **Steve Whitmore** **School of Engineering Science** **Simon Fraser University**
Date	**February 16, 1999**

Executive Summary

One of the problems for humans using machines is the lack of intuitive interfaces between them. The most intuitive mode of communication for humans is voice, so it makes sense to use voice to interface with machines. We are developing a voice-activated system which can be used to control a variety of devices.

The VOICE (Voice Operated Interface Control Electronics) System will allow precise control beyond the level of basic commands such as On/Off. The system will use a speaker-independent voice recognition system, so that anyone may use it without having to retrain it.. To demonstrate the versatility of the voice-controlled system, it will be tested on a telephone and robot submersible, also constructed by Aqua-Acoustic.

The VOICE System will use a voice recognition chip to identify spoken commands. A code corresponding to the identified command will be sent to the controller unit, which will output a control signal through the device interface to the appropriate device. The interface will convert the signal into the proper form to activate the device.

This document describes the functions of the VOICE System and its components.

Table of Contents

List of Figures and Tables

Introduction

One of the problems for humans using machines is the lack of intuitive interfaces between them. Since our most intuitive mode of communication with other humans is through voice, so it makes sense to use voice to interface with machines. We are developing a voice-activated system which can be used to control a variety of devices.

The VOICE (Voice Operated Interface Control Electronics) System will allow precise control beyond the level of basic commands such as On/Off. The system will use a speaker-independent voice recognition system, so that anyone may use it. To demonstrate the versatility of the voice-controlled system, it will be tested on a telephone and robot submersible, also constructed by Aqua-Acoustic. The VOICE system will regulate the speed of the submersible, thus demonstrating that fine control is possible with a voice recognition system. The telephone will demonstrate that the system can interface with other devices.

This document will describe the functional requirements of the VOICE system and its components, as well as the physical, environmental, electrical, and safety requirements of the system.

System Overview

The VOICE system will take in a set of voice commands to control any device for which the proper firmware has been written. Figure 1 shows the overall design and structure of the system.

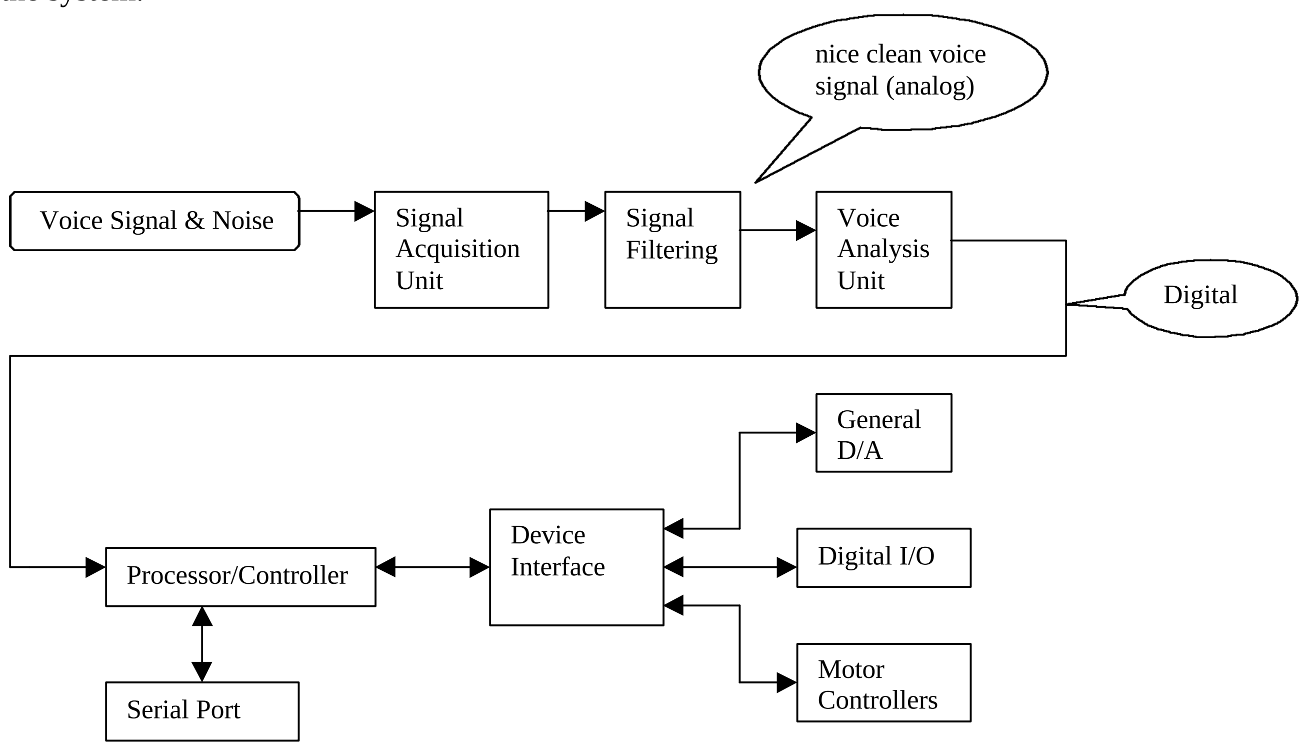

Figure 1: System Schematic

Each of the sub-systems shown in the diagram will be explained in greater detail in the following sections.

The VOICE System will:

- use voice as the standard input.
- have a maximum response time of 2 seconds.
- use a standard wall socket as power supply.
- have an effective and intuitive user interface.

System Specifics

Signal Acquisition

The first stage of the VOICE system acquires the voice signal from the user. Figure 2 shows the block diagram of this part of the system.

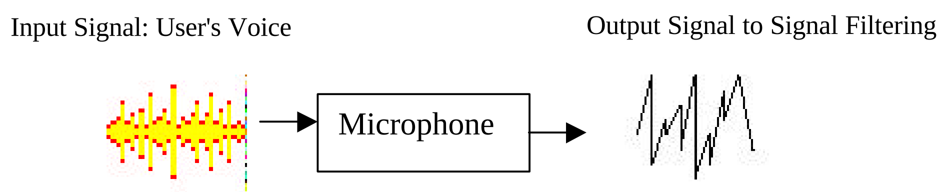

Figure 2: Signal Acquisition Block Diagram

A microphone will be used to collect the voice signal, which is the input for the system.

The specific microphone should

- be able to collect frequencies of 50-15,000 Hz (human voice range)
- have a high signal-to-noise ratio.
- attach easily and non-permanently or be mountable
- be comfortable to use and of minimal disturbance to the user.
- pass the acquired signal to the signal filtering stage of the system.

Signal Filtering

The second stage should amplify and filter the signal from the signal acquisition stage. Figure 3 shows the block diagram for this second stage.

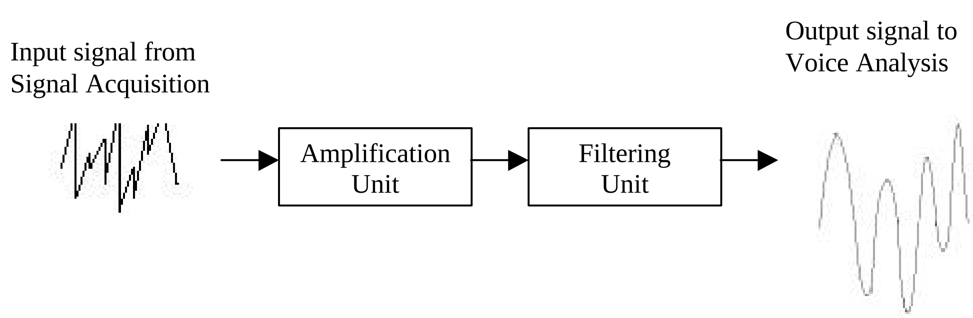

Figure 3: Signal Filtering Block Diagram

The Signal Filtering stage will:

- receive input from the signal acquisition stage.
- provide output to the voice analysis unit.
- amplify the desired signal to the level required by the electronics in the voice analysis unit.
- filter out all noise signals.
- provide background noise reduction/suppression.

Voice Analysis

The voice analysis unit will:

- identify approximately 40 words.
- be speaker independent.
- have processed the clean voice signal to a digital signal.
- have a generalized vocabulary which allows for increased range in applications.

Processor/Controller Unit

The Processor/Controller Unit (PCU) Interface

The PCU coordinates the operation of the entire voice recognition system. Two types of signals are transmitted/received by the PCU: device signals and system signals. These signals are processed by system firmware, which is also divided into two levels: device code, and operating system code. Device code is firmware dedicated to controlling a particular device, and thus processes device signals. Operating system code handles all aspects of system control not related to a particular device, and thus processes system signals.

Table 1: PCU signals

System Signals		**Direction**
	system error signals	Input
	component failure signals	Input
	indicator lights, sounds, displays	Output
	serial port I/O	Bi-directional
	some voice commands	Input
Device Signals		
	All DIU I/O	Bi-directional
	most voice commands	Input

General PCU Functions

The PCU will:

- interface with the standard serial port.
- receive digital data identifying the command words recognized by the Voice Processing Unit.
- at power ON, initiate the startup sequence that places the system into a usable state.
- coordinate command word training.
- provide a data storage location which does not erase at power OFF.
- detect and handle system errors and component failures.
- provide the user with visible and audible feedback.
- for debugging purposes, provide visible feedback indicating system status.
- send digital commands to the Device Interface Unit.
- receive digital/analog signals from the Device Interface Unit.
- run a firmware program for each device under control.
- provide a basic operating system.

Operating system code will:

- take precedence over device code whenever necessary.
- handle all system errors and failures.
- coordinate device programs, especially if multiple devices are being controlled simultaneously.
- handle the system startup sequence.
- not be modified or overridden by device programs.
- manage system memory.
- reside in system memory, even after power OFF.

A device program will:

- handle all device errors and failures
- be programmable by Aqua Acoustics customers, should they wish to do so.
- be shut down by the operating system if it generates an exception.
- reside in system memory, even after power OFF.

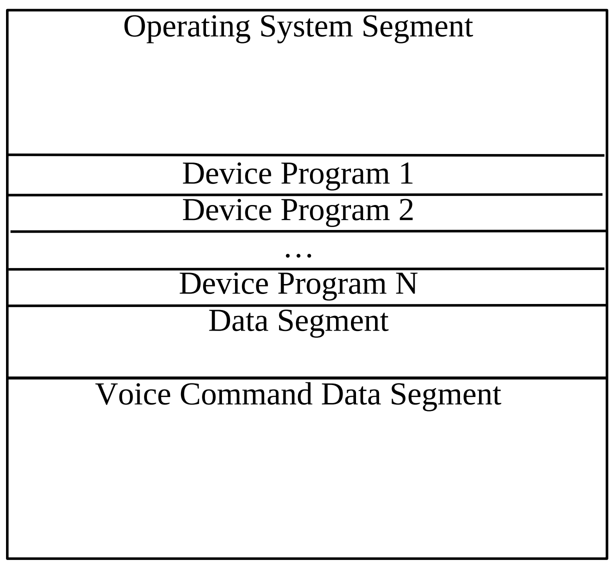

Figure 4: PCU Memory Organization

Device Interface

The VOICE System will be able to operate up to two devices. The PCU will output an index to indicate which device is being used. The device interface will accept digital commands from the PCU and send the appropriate signal to the device.

Figure 5 shows the functional block diagram for the device interface unit.

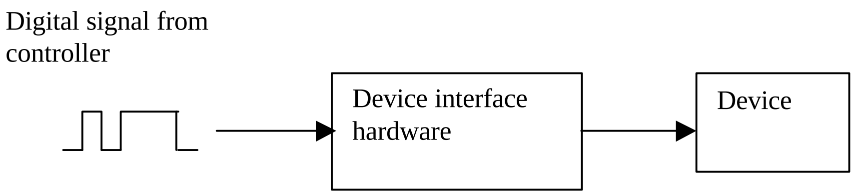

Figure 5: Device Interface Block Diagram

The device interface unit will:

- accept a digital signal from the controller through a serial interface.
- interpret the signal.
- send the appropriate actuation signal to the selected device.

Tables 2 and 3 show the input-output relationships for the interface units for a telephone and submersible.

Table 2: I/O Relationships for a Telephone Interface

Input	Output
0000	Press key 0
0001	Press key 1
0010	Press key 2
0011	Press key 3
0100	Press key 4
0101	Press key 5
0110	Press key 6
0111	Press key 7
1000	Press key 8
1001	Press key 9
1010	Pick-up
1011	Hang-up
1100	Dial
1101	Re-dial

Table 3: I/O Relationships for a Submersible Interface

Input	**Output**
Directional control	
0000	Up
0001	Down
0010	Left
0011	Right
0100	Forward
0101	Backward
Speed Control	
0000	Off
0001	Speed 1
0010	Speed 2
0011	Speed 3
0100	Speed 4
0101	Speed 5
0110	Speed 6
0111	Speed 7
1000	Speed 8
1001	Speed 9
1010	Full Speed

User Considerations

Human Interface

Although the most important aspect of the human interface in the VOICE System is the human voice there are other important elements to an effective interface. The main purpose of the interface is for effective communication with the user.

The human to system interface will consist of:

- the user's voice.
- a microphone to capture the voice.
- a simple LED display to report system status.
- audio indicators for system status.
- LCD display for specific error messages and acknowledgement of settings.
- an LED display, within the enclosure used for debugging purposes.
- a large power button to turn system on and off.

Training

The training required for the VOICE system will be minimal. The only training that will be required is training in the use of the command words to get the desired response from the machine being controlled. The command words should be well enunciated.

Physical, Environment and Electrical Specifications

The VOICE system will have the following physical, environmental and electrical specifications as given in Table 4.

Table 4: Physical, Environmental and Electrical Specifications

Height	10 cm maximum
Length	30 cm maximum
Width	20 cm maximum
Weight	2 kg maximum
Power Supply	120V, 20A, AC source, a standard wall socket
Operating Temperature Range	-20 – 40 degrees Celsius
Enclosure Temperature	Will not exceed the ambient temperature by 10 degrees
Humidity	Will operate in all humidity levels

The box containing the VOICE system will be made of a non-conducting, waterproof plastic.

Safety Requirements

The VOICE system will meet the following safety requirements.

Enclosure

The enclosure will not have sharp corners or edges, or be electrically conductive thus not posing a danger to the user. It will also be water resistant as it will be used in the proximity of water.

Electrical Isolation

All inputs and outputs will be shielded and protected from external static voltage sources.

Emission

The system will emit minimal electromagnetic radiation.

Testing

Each functional block of the VOICE System will be tested separately to ensure that any faults in its operation are detected and corrected before the system is assembled. This way, problems can be isolated more quickly, and the adverse effects of those problems can be limited to smaller portions of the system. Whenever possible, the schematics for a functional block will be reviewed by engineers not involved in the design process. This way, design problems will be detected before functional blocks are assembled. Before releasing a design for prototyping, the design engineers will review the netlist, checking all text nets against the graphical schematics. This process, used extensively at Creo Products Inc., forces design engineers to view their schematics at the “connection” level, rather than at the conceptual level. Finally, several indicator LEDs shall be provided for debugging purposes.

Testing Criteria:

- Speaker independence will be tested using a variety of male and female speakers.
- Percentage success will be tested by repeatedly speaking command words into the microphone and recording the number of successful detections versus the total number of commands issued.
- Memory permanence will be tested by turning the system off.
- The startup code will be tested by repeatedly booting the system.
- All rail voltages are within ±5% of their nominal value.
- The supply can sustain an output of 300W through load resistors.
- Ripple on supply rails is ±0.1% or less.
- Noise rejection will be tested for two criteria: noise outside the human voice frequency band has been completely eliminated, and noise inside the human voice frequency band has been attenuated to the extent that the voice recognition processor can recognize command words with the same rate of success as when the noise is not present.
- The system will be tested for a maximum propagation delay of 2 seconds.
- System errors and component failures will be tested through simulation.
- Usability will be tested by people not involved in the design process. Special attention will be given to usability as it applies to aiding the disabled. Whenever possible, disabled people will be asked for their input concerning usability issues.

Conclusion

Aqua Acoustics' VOICE system is a versatile product that will be capable of controlling a wide range of devices. Digital and analog I/O will allow easy control of consumer electronics, while motor control capabilities make the voice recognition system ideal for controlling assistive devices such as wheelchairs. The system also has research applications, of which the voice controlled submersible is a prime example. Furthermore the voice recognition system is fully programmable, which means that the command set and firmware can be adapted to suit the user's needs and preferences.

Aqua Acoustics has also given considerable thought to usability issues. The system will operate from the standard wall socket, which makes it ideal for household applications. Audio and visual indicators provide both the blind and the deaf with adequate feedback.

Aqua Acoustics believes that the versatility of its VOICE System will make it the product of choice for research applications and control of assistive devices for the disabled.