Vtronix Incorporated

Simon Fraser University
Burnaby, BC
V5A 1S6
vtronix-inc@sfu.ca

March 9, 1999

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

Re: ENSC 370 project Voice Activated Control System design specifications

Dear Dr. Rawicz,

The attached document, *Voice Activated Control System Design 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 design requirements of the system as well as the components selected for use in the system. The modules of the system are the signal acquisition unit, signal conditioning unit, voice analysis unit, processor/controller, and device interface units for 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 Design Specifications*

Vtronix Incorporated

Voice Activated Control System Design Specifications

Submitted by Vtronix 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 March 9, 1999

Executive summary

The VOICE (Voice Operated Interface Control Electronics) System by Vtronix, formerly known as Aqua-Acoustic, is a generalized control system for the operation of any number of machines and appliances. As indicated by the name of the product this control system is designed to be operated by voice commands. In essence this system will introduce a new interface, a voice command interface, to any electronically controlled device.

The VOICE System was created primarily for the use of disabled people to aid in the operation of common devices which, unfortunately, were not designed for use by the disabled. It also has the potential to be used in other areas including remote operation and improving human-machine interfaces.

This system can be used with any electronically controlled device by adding software specific to the device. The VOICE System has the advantage of being small, portable and modular.

Table of Contents

1. INTRODUCTION 1

2. SYSTEM OVERVIEW 2

3. INTERFACE 3

3.1 USER INTERFACE 3
3.2 DEVELOPER INTERFACE 3

4. VOICE ACQUISITION UNIT 4

5. VOICE PROCESSING UNIT 5

5.1. FUNCTIONAL OPERATION 6
5.1.1. Training the Voice Processing Circuit 6
5.1.2. Recognition mode for the Voice Processing Circuit 6
5.2. HM2007 SPEECH RECOGNITION CHIP 6
5.2.1. Speaker Dependent Recognition 6
5.2.2. Speaker Independent Recognition 6

6. CONTROL UNIT 7

6.1 MICROCONTROLLER 7
6.2 PROGRAMMABLE LOGIC DEVICE 7
6.3 FIRMWARE 7

7. MOTOR CONTROLLER ELECTRONICS 8

7.1 GENERAL INFORMATION 8
7.2 FUNCTIONAL BLOCKS 8
7.2.1 Optoisolation 8
7.2.3 Current Limiting Circuitry 9
7.2.4 L298 Dual Full Bridge Driver 9
7.2.5 Power Dissipation and Heat Sinking 10
7.2.6 Clamping Diodes 11

8. POWER REGULATION 12

9. TELEPHONE CONTROL UNIT 13

10. SUBMERSIBLE 14

11. TESTING 15

11.1 VOICE PROCESSING UNIT 15
11.2 CONTROL UNIT 15
11.3 DEVICE INTERFACE TESTING 15
11.3.1 Telephone 15
11.23.2 Submersible 15

12. CONCLUSION 16

List of Figures and Tables

Figure 2.1: System Functional Block Diagram 2
Figure 4.1: Block Diagram of Voice Acquisition Unit 4
Figure 5.1: Voice Processing Unit Functional Block Diagram 5
Figure 7.1: The Internal Structure of the L298........ 9
Figure 7.2: Power dissipation in the L298 10
Figure 8.1: Power Regulator Block Diagram 12
Table 8.1: Power Requirements of Components........ 12
Figure 9.1: The Telephone Keypad Matrix 13
Figure 10.1: The Submersible viewed from above........ 14
Figure 10.2: Pulse Width Modulation........ 14

1. Introduction

Majority of devices in our everyday life relies heavily on tactile control or movement: such as computers, lights, T.V, phone, and any electronic device. Such a heavy demand on our physical system can cause many physical ailments, such as Carpal Tunnel Syndrome, and can also be a barrier to the physically disabled who have less tactile control or mobility. A solution is to have another means of controlling devices. There are device control systems developed or being developed using means of control that are unintuitive, due to current technology limitations, requiring electrodes or sensors to be attached to our bodies, such as in the case of using brainwaves or eye movements. Vtronix (formerly known as Aqua-Acoustic) believes that the voice is a natural and intuitive method of control. It requires a sensor that does not need to be worn and is not invasive (with our current technology) and can easily translate our thoughts to actions by words. We feel that the voice is the remaining undeveloped means of controlling devices. Other utilization of control systems using voice have mainly been:

- voice dependent, only one user's voice is recognized;
- large and bulky, requires either a computer or largecomponents;
- device specific, configured to control only that device;
- expensive, most require computers that not everyone has or can take everywhere;
- on and off control only; and
- pre-installed, such as most central control systems that accept computer and voice commands.

Vtronix's VOICE (Voice Operated Interface Control Electronics) System will be voice independent, small and portable, and universally adaptable to any device. The VOICE System's key feature is its open architecture, and precision control. The VOICE System is currently geared towards the disabled and elderly people, and the scientific/research community. In future, industrial and general consumer applications are foreseeable.

The following sections outline the design specifications of the VOICE System and two sample applications: an underwater submersible, and telephone - demonstrating the system's control of motors and electronics.

2. System Overview

The system is composed of six functional blocks:
1) Interface
2) Voice Acquisition Unit
3) Voice Processing Unit
4) Control Unit
5) Motor Controller Electronics
6) Power Regulation Unit

The user supplies a command through the user interface (microphone), this is how the voice signal will be transferred to the voice processing unit. The command is then translated into a 13-bit address and sent to the control unit where the command will be decoded and executed. The Control Unit is responsible for logical operations, including decision making, data transfer, and timing. It forwards control signals to the appropriate device interface, which performs the task associated with these commands/control signals. See Figure 2.1.

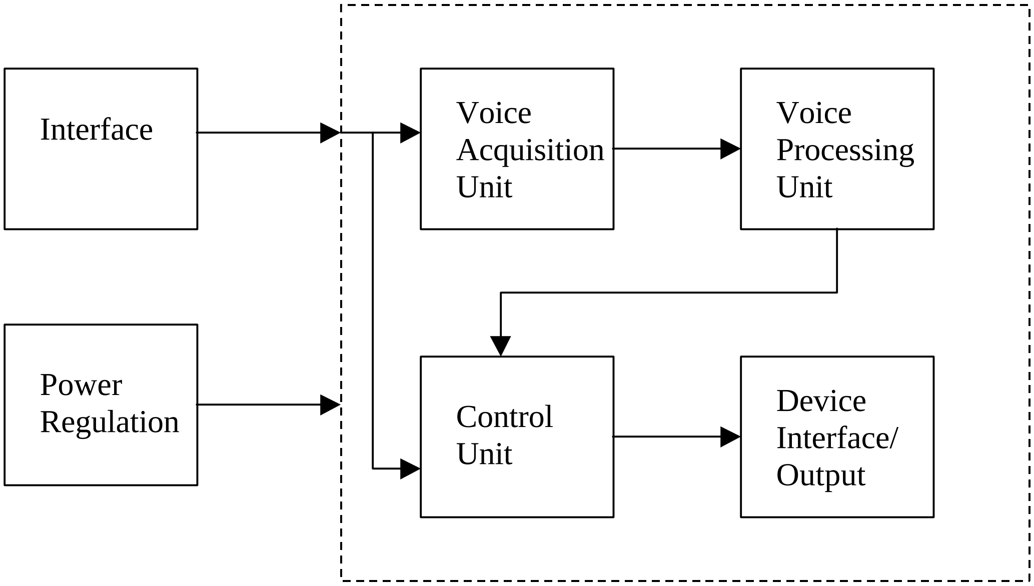

Figure 2.1: System Functional Block Diagram

To demonstrate the VOICE System we decided to control two devices: a telephone and a submersible. These two devices required either modification or construction and will be discussed in later sections.

3. Interface

The system interface consists of two parts: a User Interface, and a Developer Interface.

3.1 User Interface

The user interface is what can be seen outside the system housing. It consists of:
- a microphone,
- a Ready LED,
- a Power LED,
- an Error LED,
- a plug-in AC adapter,
- and an ON/OFF power switch.

The microphone we are currently using in our prototype is a small Electret omni-directional microphone, because it was inexpensive and easy to test. A larger handheld microphone can easily replace the current microphone, requiring only the addition of a connector.

3.2 Developer Interface

The Developer Interface consists of:
- two 7-segment displays,
- a numeric keypad,
- and debugging LEDs.
- serial bus

This interface was design to allow a developer to have a more detailed idea of how the VOICE system is operating allowing for easier programming. The 7-segment displays allow the developer to see exactly which command word corresponds to which eight-bit signal. The keypad is used in the programming of new command words. The debugging LED's are for general purpose debugging. The serial bus is used for downloading and uploading programs from the control unit.

4. Voice Acquisition Unit

Figure 4.1. shows the block diagram of our Voice Acquisition Unit.

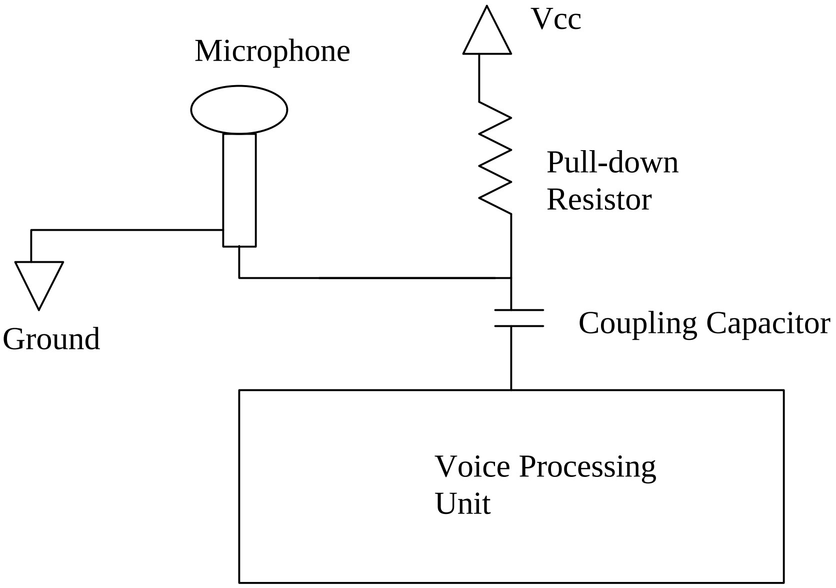

Figure 4.1: Block Diagram of Voice Acquisition Unit

The microphone we chose had two leads, one as ground as one as output. The Vcc specified was 5 V, as in the rest of the control system. And the pull-down resistor value was 1 kΩ. There is also a decoupling capacitor of value 0.1 μF to remove noise.

5. Voice Processing Unit

The functional block diagram of the Voice Processing unit of the VOICE system is shown in Figure 5.1. This unit takes in the analog voice signals from the Voice Acquisition Unit and outputs a digital code to the micro-controller unit of the VOICE system for further processing.

The unit uses:

- the HM2007 speech recognition chip,
- a DS1225Y nonvolatile SRAM,
- a 74LS373 D-Latch,
- two 7447 BCD to 7 segment display converters, and
- two 7 segment displays - common anode.
- 1 green READY LED

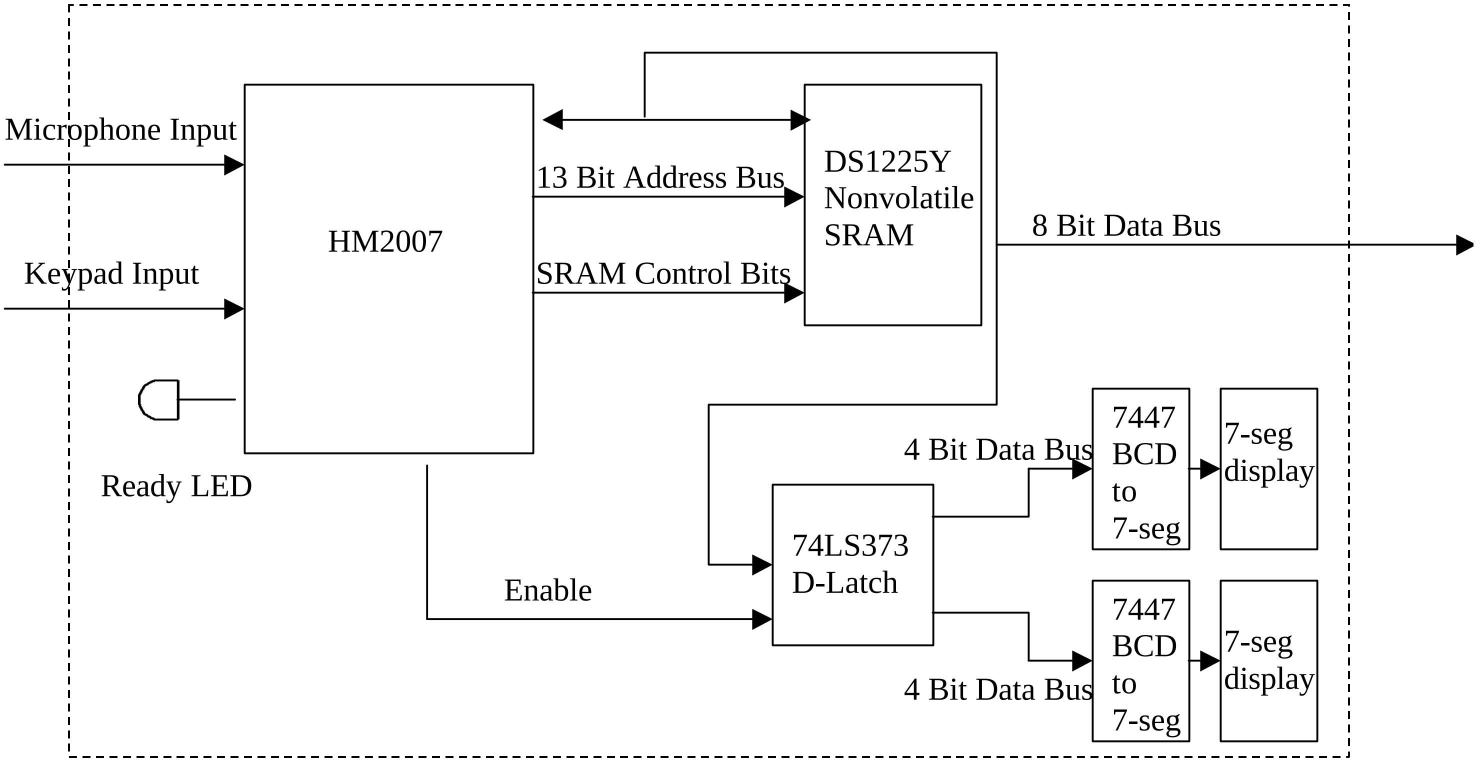

Figure 5.1: Voice Processing Unit Functional Block Diagram

We chose the HM2007 speech recognition chip because of its:

- affordable price (~$25 US),
- high recognition success rate (advertised as 95%),
- and large vocabulary (10 words independent, 40 words dependent)

The HM2007 chip required an SRAM to store the voice signals. Since it was important to us that the Voice Processing Unit retain its information we looked for a nonvolatile

SRAM. The DS1225Y nonvolatile SRAM was inexpensive (~$10 Cdn) and easily accessible (purchased at a local electronics store). Other components of the Voice Processing Unit are standard parts used for development purposes.

5.1. Functional Operation

The Voice Processing Unit can operate in one of two different modes: Training mode and Recognition Mode.

5.1.1. Training the Voice Processing Circuit

The voice processor starts in recognition mode and the Ready LED will be on. In order to switch to training mode one must use the keypad to type in the address where the new word is to be stored and then press the "#" key. For example, if you wanted to program word 01 you would type "0" then "1" then "#", also note that the number 01 will be displayed on the 7-segment displays. When you pressed the "0" the Ready LED would turn off and when the "#" is pressed the Ready LED will turn back on. The Voice Processing Unit is now waiting for the new word. Speak the word into the microphone. The Ready LED will flash off and on to indicate that the word has been learned.

5.1.2. Recognition mode for the Voice Processing Circuit

As mentioned above the Voice Processing Unit starts in recognition mode. In this mode the Unit will compare any word spoken to it with all the words in it memory. If there is a match then the address of that word will be displayed on the 7-segment LED displays. For example, assume word 32 is "assimilate". If "assimilate" is spoken into the microphone the Voice Processing Unit will recognize it as word 32 and display the number 32 on the 7-segment displays.

5.2. HM2007 Speech Recognition Chip

The voice processor we selected for use in the VOICE system is the HM2007 speech recognition chip. This chip is capable of both speaker dependent and speaker independent recognition.

5.2.1. Speaker Dependent Recognition

Speaker dependent recognition has each word trained once - each word is stored in one address. It is important that the manner of speaking remain the same for recognition to occur. The recognition is really a match between the word spoken and the word stored.

5.2.2. Speaker Independent Recognition

Speaker independent recognition is achieved by training each word four times, storing four different persons articulation/vocalizations of the same word to a different address with the same significant digit each time. For example "assimilate" spoken by person 1 can be stored in address "01", person 2 would store his/her utterance in address "11", person 3 in address "21" and person four in address "31" notice that the first digit "1" is common to all spoken versions of the same word.

6. Control Unit

6.1 Microcontroller

In the VOICE system we are using the MC68HC11E2 microcontroller from Motorola. We selected it because it was provided to us free of charge, and because the 68HC11 series is easy to use, common and well supported. The MC68HC11E2 has internal EEPROM for storing code and 5 ports for interfacing to other devices. These include a serial port for downloading code, a multi-use port, a port that can be used as an A/D converter, and two address ports.

We are using the microcontroller in expanded memory mode, which gives us 256 bytes of RAM and 2K bytes of EEPROM to store our code. We suspect we will need more than 2K bytes of EEPROM, so we are using an external EEPROM device, the X68C64, to obtain an additional 8K bytes of EEPROM. The X68C64 is designed to interface to the 68HC11 microcontroller to expand its EEPROM. The X68C64 also has security features that we plan to use to implement our firmware as an operating system with separate drivers for each device that the system controls.

6.2 Programmable Logic Device

We are using a programmable logic device (the MAX7000S7128) to interface the outputs from the microcontroller to the devices that the system controls. The MAX7000S7128 has 60 I/O pins and can be used to implement any logic functions between the inputs and outputs. The MAX7000S7128 is connected to the MC68HC11E2 through an 8-bit data bus. The 8-bit output of the HM2007 voice recognition chip is connected to the MAX7000S7128, which will transfer it through the 8-bit data bus to the MC68HC11E2. The outputs of the MAX7000S7128 are connected to the device being controlled, in this case the telephone or submersible.

6.3 Firmware

The firmware to control the VOICE system consists of an operating system with driver programs for each device that the system controls. The operating system will consist of boot code that will start up the system in the proper state. Once the system is initialized, it will run in a continuous loop in which it checks for a change in the output command from the voice processing unit. When a change is detected, the firmware will process the new command appropriately.

Each device driver will consist of a case statement that specifies the output to the device for each input command. There will be a keyword that is spoken to switch to each of the devices that the system controls. When a device selection keyword is recognized, the firmware will switch to the proper driver. When a control command for the device is recognized, the firmware will find the command in the case statement and send the corresponding control signal to the device.

7. Motor Controller Electronics

7.1 General Information

The MCE, an improved version of the Mirosot motor control circuitry, is designed to drive three DC motors using a single power supply. Each motor may draw up to 2.5A, and may be pulse width modulated by external circuitry to obtain variable speed and direction settings. The outputs for each of the three motors may by connected in parallel, to provide a 7.5A max output current. The MCE is designed to dissipate 50W of power, and has internal current limiting circuitry that disables the outputs in the event that more than 2.5A are drawn from any one output. The MCE contains a +5V regulator to supply on board logic with a noise free power supply. Noise is prevented from travelling to external circuitry through the use of optoisolation on all control and status signals.

Features:

- variable power supply (10V to 20V)
- 50W power dissipation
- 2.5A maximum current per output
- internal current limiting circuitry
- 2A fast recovery schottky transient suppressors
- internal voltage regulator
- optoisolation

7.2 Functional Blocks

7.2.1 Optoisolation

The TTL compatible optoisolators U4 and U5 shield external circuitry from noise generated by the motors, or motor power supply. The ILQ74 TTL compatible optoisolators are inexpensive and readily available parts proven to work on the Mirosot motor controller. The pull-up resistors R1-R7 ensure that all signals received by the bridge drivers are logic low at power up or reset. The series resistors R20-R26, and R29-R31 limit the current through U4, U5, and U11 to 6mA.

7.2.2 Control Logic

The schmitt trigger inverters U2 and U3 add hysterisis to all control and status signals. Without hysterisis, noise could generate unwanted toggling of the control signals. The 74HCT08 and 74HCT14 were chosen for the control logic because they are TTL compatible with the ILQ74 optoisolators. Also, these parts are stocked by the SFU engineering lab, which makes them readily available.

7.2.3 Current Limiting Circuitry

The current limiting circuitry uses a sense resistor to detect motor over-current. Current from the motor flows through the sense resistor (R17-R19), generating a voltage across it. The LM339 comparator (U9) compares the voltage across the 0.33Ω sense resistor to a reference voltage of 0.825V generated using the voltage divider R27-R28. Note that R8-R16 add 0.5% hysterisis to the comparison circuitry, which prevents noise from toggling the comparator output.

Ordinarily, all the comparator outputs are low, which makes the output of U1A, B, and C the same as the output of U3A. However, should any one motor draw more than 2.5A, one comparator output will rise, and the AND gate U1C will drive the ENABLE inputs on the motor driver low, cutting off current to that particular motor. The other two motors will remain full operational.

We decided to use the LM339 because it provides four comparators in one package, which limits the physical size of the current limiting circuitry. Also, previous experience with the LM339 allowed us to quickly integrate the part into our design. The part is also readily available from most electronics distributors.

We have not yet chosen a manufacturer for the sense resistors. However, they are to be 1Ω, non-inductive, 7.5W resistors.

7.2.4 L298 Dual Full Bridge Driver

The L298 is the component that is used to control the speed of the motors, its internal structure is illustrated in Figure 7.1

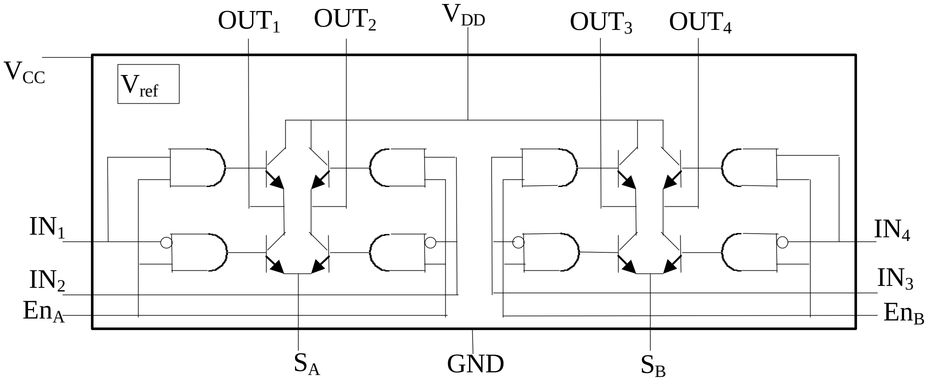

Figure 7.1: The Internal Structure of the L298

The L298 contains two separate bridge drivers, each capable of delivering 2A. The submersible motors require 2.5A each, so the two channels of the L298 are connected in parallel. The control signals are also connected in parallel, as shown in Figure 3.3. The En_A and En_B pins can be used to enable the entire device, while the IN_1, IN_2, IN_3, and IN_4 pins can be used to supply the pulse width modulation signals that control the speed of the motors. The motor current flows out of the sense terminals S_A and S_B, across the sense resistors, R17, R18, and R19, and into ground.

7.2.5 Power Dissipation and Heat Sinking

Due to the large motor currents flowing through R17, R18, R19, U6, U7, and U8, these components must be capable of dissipating large amounts of power. The sense resistors are 0.33Ω, and handle a maximum of 2.5A of current. Thus the maximum power dissipation is given in [7.1].

$$P_D = 2.5^2 \times 0.33 = 2.1W \qquad [7.1]$$

The sense resistors we chose are capable of dissipating 5W.

The power dissipated by the L298 is the product of the motor current flowing through the bridge, and the collector-emitter saturation voltage of the bridge transistors.

$$P_D = I \times (V_{CE_1} + V_{CE_2}) = 2.5 \times (2.7 + 2.3) = 12.5W \qquad [7.2]$$

Figure 7.2 shows a model of the L298 dissipating 12.5W of power. Note that the L298 is capable of dissipating up to 25W of power. However, even 12.5W could cause the L298 to overheat if no heatsink is used. The maximum junction temperature is given in the L298 data sheet, along with the junction-case thermal resistance. These values are, respectively, 130°C and 3°C/W. Figure 7.2 shows a junction temperature 10°C lower that the absolute maximum rating.

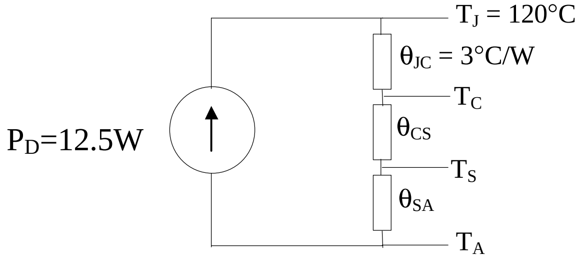

Figure 7.2: Power dissipation in the L298

Using [7.3], the maximum allowable case temperature is calculated to be 82.5°C.

$$T_C = T_J - P_D q_{CS} \qquad [7.3]$$

Again, using [7.3] the thermal resistance from the case to ambient is calculated to be 3.4°C/W. A typical value for θ_{CS} is 1.0, therefore, we need a heatsink that has θ_{SA} = 2.4°C/W.

7.2.6 Clamping Diodes

To suppress transient voltages generated by the motors, two 1N5822 200ns recovery Schottky clamping diodes are connected to each motor output. These diodes are capable of handling up to 3A of current, and have a reverse breakdown voltage of 40V.

8. Power Regulation

When considering the power supply of the VOICE System we had to tabulate the current requirements of all components, see Table 8.1, and find a common supply voltage (Vcc). Figure 8.1 shows a block diagram of the power regulator.

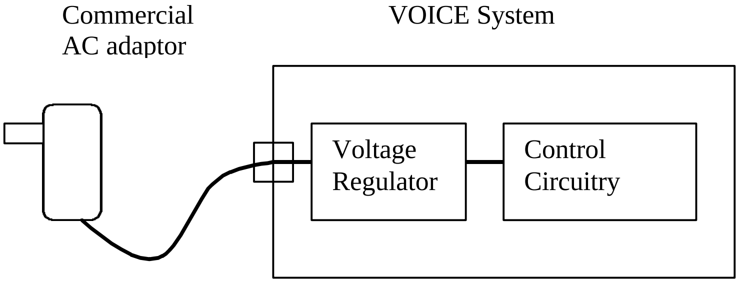

Figure 8.1: Power Regulator Block Diagram

Table 8.1: Power Requirements of Components

Component	Voltage Required	Total Current
Schmitt Trigger	5 V	0.1 mA
RS 232	5 V	15 mA
LS 373 (2 @ 40mA each)	5 V	80 mA
SN 7447 (2 @ 1mA each)	5 V	2 mA
X 68C64	5 V	60 mA
DS 1225Y	5 V	75 mA
HM 2007	5 V	15 mA
EPLD	5 V	120 mA
HC11	5 V	27 mA
Decoupling Capacitors		50 mA
LED's	5 V	50 mA

The total current required is 494.1 mA. Since the above is an approximation of the current required we felt a 500 mA supply may be insufficient. We will be using a commercial plug-in AC adaptor (CSA approved), 12V @ 800mA, with 2.1mm coaxial power plug (center negative). In order to provide an exact 5 V to our system we will be using a commercial voltage regulator.

We chose a commercial power supply for safety reasons because it would minimize our user's and our exposure to shock and electrocution.

Any appliance or machine connected to the VOICE System is responsible for its own power supply and should not rely on the VOICE System's power regulation system for its power supply. The motor controller electronics is responsible for its own power supply, a commercial regulated DC power supply (14 V @ 7-10 A).

9. Telephone Control Unit.

The telephone control unit is designed to operate a commercially available telephone with only minor modifications to the telephone keypad and dial tone button. A keypad works by means of making or breaking a connection between two wires. A connection is made when a finger presses a conductive material across a gap between two wires. The wires are arranged in the form of a matrix, so that row and column indicate which button is being pressed. Figure 9.1 shows how the matrix in a typical telephone keypad is arranged.

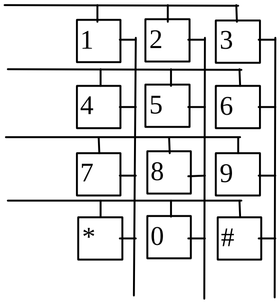

Figure 9.1: The Telephone Keypad Matrix

The telephone control unit uses digitally controlled analog switches to make and break the connections on the keypad. Experiments show that the maximum allowable resistance to successfully make a connection is 22Ω. Thus, we need an analog switch with an on resistance of 22Ω or less. Each analog switch will be toggled on and off using one I/O pin of the EPLD. The dial tone button on a telephone works in exactly the same fashion as the keypad. Thus we can control it using the same analog switch.

There are 12 buttons on the telephone keypad, and 1 dial tone button. Therefore, 13 analog switches are needed to control one telephone.

Currently, the VOICE system uses two microphones: the telephone microphone, and the voice command microphone. Using the same microphone for both purposes would make the system easier to use. However, if the user spoke a command word during a telephone conversation, the VOICE system might respond to the command word, even if the user did not intend it as a command per se. We have not yet found a solution that will enable the VOICE system to detect the context of the command word. However, once we have a working prototype, we will be in a better position to experiment with potential solutions.

10. Submersible

One of the devices we will use to demonstrate the VOICE system's operation is a small submersible which operates in a freshwater tank. The tank is approximately 4 meters across in both dimensions and about 2 meters deep. The submersible is connected to the world outside the tank by a tether which contains wires that supply power to the submersible's motors. The submersible is powered by three DC motors driving propellers, which operate at 12 Volts and can draw up to 2.5 Amps each. The frame of the sub is made of 1/2 inch diameter plastic piping connected in a square 30 centimeters on each side. Four 10-cm diameter spherical plastic floats filled with air are attached to the frame to create buoyancy.

The motors are mounted on the sub's frame as shown in Figure 10.1. The two rear propellers are mounted so that they thrust forward and the up/down propeller is mounted so that it thrusts downward. By reversing the input voltage to the motor, the direction of rotation of the propellers can be reversed so that they can provide thrust in the backward or upward directions. Steering of the sub to the left or right is done by running one of the rear propellers in the forward direction and the other rear propeller in the reverse direction or by running one propeller when the other is off. Pulse width modulation of the input voltage to the motors will be used to achieve variable speed. Pulse width modulation involves changing the duty cycle of a square wave voltage input to the motor. The amount of power the motor produces is determined by the percentage of the duty cycle that is at the high level. Pulse width modulated signals are shown in Figure 10.2.

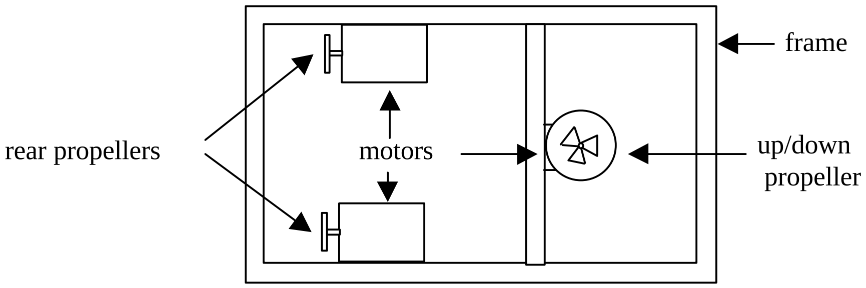

Figure 10.1: The Submersible viewed from above

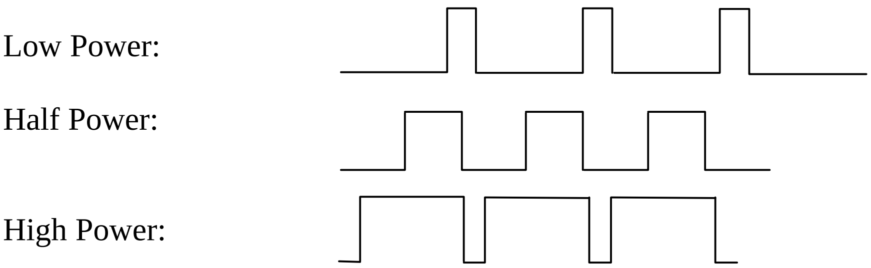

Figure 10.2: Pulse Width Modulation

11. Testing

Our testing procedure is split into three sections based on the three major units: Voice Processing Unit, Control Unit, and Device Interface. This hierarchical testing procedure allows us to effectively debug knowing that the lower level sections are reliable.

11.1 Voice Processing Unit

The Voice Processing Unit will be tested independent of other units. Testing the Voice Processing Unit requires that we program our command words into the voice control chip, see 5.1.1 Training the Voice Processor Circuit. We will probe the address bus looking for corresponding values of the entered address and verifying the displayed address on the BCD. To test the recognition aspect, we speak words that were stored and an unknown word. The BCD display should accurately display the address of the stored word or display "77" for the unknown word.

11.2 Control Unit

The Control Unit will be tested independent of the Voice Processing Unit or the Device Interface. We will monitor the internal states due to changes in the input, which will be from the Voice Processing Unit address bus. The appropriate output from the Control Unit will be determined and verified.

11.3 Device Interface Testing

The Device Interface testing will be accomplished by testing it independently, tested with the device and also testing it with all the other components as a final system test. In the final system test, switching control between devices and individual device functionality will be tested.

11.3.1 Telephone

The telephone functions of answering a call, making a call, and ending a call will be tested. This will be repeated as specified above: testing just the interface and probing its output, testing the interface with the telephone, and testing the integrated system.

11.23.2 Submersible

Testing the control of the submersible includes testing all directions of mobility at various speeds, and ensuring the manual control override works properly.

12. Conclusion

The VOICE System has been designed with readily available components. The design incorporates, and improves upon, existing electrical systems, such as the Mirosot motor control board. Thus, we can use the experience of other engineers to speed up prototype development. Where new design work is necessary, the VOICE System is simple, which will also make prototyping easier. Since no prototype has yet been built, some parts of the design may not function as intended. Furthermore, a few remaining design problems need a working prototype before we can develop solutions. Thus, although incomplete, the design presented here will be used to develop the first working prototype of the VOICE System.