FIVEPOINT
technologies

Functional Specifications for the
Voiceture system

Translation of American Sign
Language into Speech

October 18, 2005

Project Team:
David Brayden
Kjell Eggen
Phoenix Yuan
Ganesh Swami

Submitted to:
Andrew Rawicz
Steve Whitmore
Mike Sjoerdsma
Brad Oldham

Simon Fraser University
School of Engineering Science
ENSC 305/340

October 18, 2005

Dr Andrew Rawicz
School of Engineering Science
Simon Fraser University
Burnaby, British Columbia

Re: ENSC 340 Functional Specifications

Dear Dr. Rawicz,

The following document the functional specifications for the Voiceture System. This is part of our project for ENSC 340. Our goal is to create a device that will translate sign language into speech using a commercially available motional capture device to recognize hand gestures.

The system has already begun the development phase with potential sensor research and design, preliminary motional capture software programming, and algorithm development. The system will feature a hand mountable motional capture device with mounted sensors and a standard personal computer.

These specifications outline the operational, technical, system, regulatory, and environmental requirements that the fully developed Voiceture System must accomplish in order to be accepted by the end user.

Fivepoint Technologies is comprised of four motivated Engineering students from Simon Fraser University. These are Ganesh Swami (President and CEO), David Brayden (Vice President, Research and Development), Phoenix Yuan (Chief Operating Officer), and Kjell Eggen (Chief Technical Officer). Questions and concerns can be directed to 604- 992-1404 or e-mail to ensc340-Fivepoint@sfu.ca.

Sincerely,

Ganesh Swami

Ganesh Swami
President and CEO
Fivepoint Technologies
Vancouver, BC

Enclosure: Functional Specifications for the Voiceture System

Executive Summary

Being deaf or mute in today's society is not easy. Something as trivial as ordering food in a restaurant can be a slow and frustrating task. In North America, few people know how to communicate properly with a deaf person, and even fewer understand American Sign Language (ASL). This is surprising, considering that the number of people using ASL as their primary language in the United States is estimated to be up to 2 million.

The Voiceture system aims to fill this gap in basic human communication by being the first economical and reliable technological device to translate signs and gestures into speech. The first stage of Voiceture will be the development of the sensing and analysis system consisting of core technology features:

1. motion capture device to sense hand orientation and positions
2. standard computer interfacing for data acquisition
3. software with the capability to sample and monitor hand positions
4. algorithms to analyze single character gestures
5. audible and visual translation outputs

Following this, the second stage of development would focus on the following requirements:

1. full-fledged word and grammar translation
2. device portability (user mobility)
3. quality and reliability

The first phase of development is already underway, with a scheduled deadline to be completed and demonstrated by December 14, 2005. This is what will be demonstrated and delivered as our ENSC 340 project.

Table of Contents

List of Figure and Tables

---

1. Introduction

The Voiceture is a motion capture system to translate American Sign Language (ASL) signs and gestures into readable English on a computer screen. The device consists of two components, the motion capture device and the gesture capturing and translation software.

Communicating with English speaking people has never been easy for a deaf or mute person as to date no economical and robust system exists outside of the traditional translator. This device will bridge that gap, proving to be a milestone towards independent real-time communication for the deaf and hard of hearing.

The development proposed herein will need to be done in stages. The first product iteration will address the fundamentals needs for such a device and showcase the core technology.

1.1. Scope

The purpose of this specification is to describe the functional requirements Voiceture as required by the potential user. Only functional requirements for the core technology will be presented as required to build the first product iteration.

The finished production device and technology is assumed to meet all the following specification plus build upon the technology through continual engineering.

1.2. Acronyms

ASL   American Sign Language

DAQ   Data Acquisition

FAQ   Frequently Asked Questions

PC   Personal Computer

USB   Universal Serial Bus

1.3. Intended Audience

This document will serve as a guideline while developing the Voiceture system by the design engineers of Fivepoint technologies.

Project analyzers will use these specifications to determine project execution and design objectives.

This document will also serve as the basis for product verification at end of the first design cycle.

1.4. Objectives

These functional requirements are intended to serve as the basis for the design and implementation for the American Sign Language translation device, the Voiceture.

Specific functional requirements will be noted using a number system according to priority marked as *(#)*. Functional requirements that will be tentatively attempted (due to timing and non-critical for core technology showcase) in the first iteration are denoted with a *. These requirements are assumed to be mandatory that must be fully satisfied by the end product.

2. Linguistic Requirements

Research indicates that the median reading comprehension of deaf and hard-of-hearing people is at the elementary/middle school level (Gallaudet Research Institute, 1996). It is also estimated that there are 160 deaf or hard-of-hearing people for every sign language interpreter in the UK (RNID.org). There is no reason to believe the situation is any better in North American (considering very few people understand ASL even though in the United States it is estimated that the number of people using ASL as their primary language is up to 2 million). These facts form the basis for the user requirements:

[1]   The system should have the ability to adapt to specific user gesture orientation

[2]   The software shall translate at a rate no slower than two letters every second

[3]   The system must understand at least the standard ASL 26-character alphabet gestures

[4]   Translation algorithms must ignore intermediate gestures between intended characters (ie, detecting a 'C' while moving towards an 'O')

[5]   Device must conform to the K.I.S.S. principle (Keep It Simple Stupid)

3. Motion Capture Device Requirements

The Voiceture system is intended to translate hand ASL gestures. Therefore, the motional capture device needs to be hand wearable and the motion capture device must then satisfy both technical and human physical requirements

3.1. Physical Requirements

[6]   The device shall be wearable on a human hand with a secure & comfortable fit

[7]   The device cannot impede the user from accomplishing normal hand gesture movement

[8]   The device must be adjustable so as to fit nominal adult human hand sizes

[9]   The device cannot exceed 250g in weight

[10]   All device wires must run away from the user in order to not impede movement*

3.2. Technical Requirements

[11]   The device must be able to track hand position and location as well as the 5 finger positions

[12]   The device must be capable of sampling with at least 20Hz of resolution

[13]   The sensors for finger positions must be accurate within 5 degrees of finger movement

[14]   Hand position tracking must be accurate to within 5 degrees of movement

[15]   Hand location sensors must be accurate to within 3cm

[16]   For connectivity, the device must be Microsoft Windows XP compatible and preferably use a serial or USB connection

[17]   Plug and Play connectivity is required for quick and easy access to translation software*

[18]   The device driver must be open sourced

[19]   The device needs to bundle all wires into a single cable*

4. Computer System Requirements

4.1. System overview

The Voiceture system consists of the motion capture device connected to a Microsoft Windows based PC with audio visual multimedia capabilities. The motion capture device incorporates the sensors that outputs hand and finger positional data at a certain sampling rate that is inputted into the computer. The computer is then responsible for all data processing and analysis and to give output to the user.
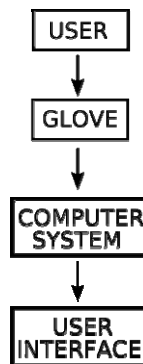

Figure 1: System Block Diagram

4.2. Physical requirements

[20] The system will consist of a standard size PC (desktop or laptop) plus the motional capturing device

[21] The system must use no more than 3 cables more than the standard PC to ensure easy setup

[22] Cable length must be sufficient enough to allow the user to be able to move hands and arms freely in a standing position while using the device

4.3. Computer Performance requirements

[23] x86 Class architecture (Standard Windows/Intel based PC)

[24] The PC must respond to user input no slower than 100ms

[25] The PC must be capable of sampling the motion capture device at least 20Hz

FIVEPOINT
technologies

[26]   For recognition purposes, the PC must detect a sign or gesture no slower than 250ms after the gesture has been made by the user

[27]   8-bit sound suppose required for voice synthesis and translation capabilities

4.4.   Software requirements

[28]   Must be capable of running Java SDK 1.4.2

[29]   Software must be executable using motion capture device without the need of additional input device beyond the motion capture device (ie, software function is independent of motional capture device)

[30]   Software must be stand-alone executable.

4.5.   Environmental requirements

[31]   The device electronics must be protected from the environment so that it can ensure operation in light rain conditions

[32]   Temperature operating range must be within -10C to 40C*

[33]   The motional capture device must withstand normal human daily use including shaking and up to 6 ft drops.

4.6.   Service & Reliability

The Voiceture system reliability mostly relies on that of the PC it is using. The PC reliability is outside the scope of Fivepoint technologies. Thus, only the software and motion capture device requirements are given here.

[34]   The unit shall be able to run continuously as long as power is supplied via the USB cable

[35]   In the event of motion capture or software failure, the Voiceture system must be capable of restarting through the software and by power cycling the motion capture device

[36]   The system must be able to handle erratic user movement such as a child having a tantrum without crashing

[37]   The system shall be upgradeable via software and motion capture device firmware and driver.*

[38]   Upgrading the software and firmware shall be able to do done by the end user*

[39]   The system should have a lifetime of at least 10 000 hours of nominal usage.*

[40]   The device must last longer than a cockroach in the event of a 5-megaton hydrogen bomb detonation *

5. User Interface Requirements

The user interface will consist of a software GUI that is controlled by the motion capture device. Program commands and options will be access via menus and prompts. The requirements laid out here are described in terms of overall function.

5.1. Program main screen

This is the screen the user will see when the main program executable is running

[41]   Upon initial launch, options to use existing profile, new profile, and delete profiles shall be shown

[42]   a "quit program" option shall be accessible in the bottom right hand corner from the main screen

[43]   Past the main screen, the "quit program" option shall be replaced with a "log out" option which will take the user back to the main screen

[44]   When a new profile is created, a step-by-step wizard will begin to walk the user through the calibration.

5.2. Gesture Recognition & Translation screen

The user will spend most of their time in this screen and thus usability and clarity is key.

[45]   A large portion of the middle of the screen will be dedicated to displaying and sounding characters as they are identified

[46]   Motion Capture device status shall be displayed

[47]   A pause translation command shall be available at all times.

[48]   A part of the screen will be dedicated to show at least the least 10 gestures translated

[49]   Recognized and translation characters will stay displayed until the next translation occurs

5.3. User Controls

The user will be required to navigate the software's commands and options via the following constraints

[50]   All commands and options must be accessible using the motional capture device

[51]   No input device other than the motion capture device shall be mandatory for access to all software features

[52]   The user shall be able to enable and disable the motion capture device (ie, motion capture device is effectively off)

[53]   A button mounted on the motion capture device will control whether translation occurs to minimize unwanted translations.

5.4. Profile & User Calibration

Profile creation and calibration serves the purpose of creating a baseline for the software to use for comparative purposes in order to correctly detect and translate gestures.

[54]   The user shall have the option to input their desired profile name (if keyboard present, otherwise a default name will be given)

[55]   Calibration shall consist of walking the user step-by-step through the 26-character ASL alphabet

[56]   Calibration data shall be saved to an external data file named after the user profile name

[57]   Once calibration is completed, the main program screen will reappear.

6. Regulatory Requirements

The Voiceture system is for home indoor/outdoor use and must comply with the appropriate government and regulatory association standards.

[58] Comply with the limits for a Class B digital device pursuant to Part 15 of FCC rules*

[59] Sensor LEDs must conform to Class 1 per IEC 60825-1-am2 (2001-01)*

[60] It will follow the CSA C22.2 No 1010.1-92 safety specification, IEC 1010-1:1990+A1:1992+A2:1995, and UL 3111-1 (3)*

[61] In order to be marketable in Europe, the product manufacturing will strive to be RoHS compatible*

The equipment may generate, use, and radiate radio frequency energy, and if not installed and used in accordance with the user manual, may cause interference with radio communications.

7. Documentation Requirements

The users of the Voiceture system will be adolescent-adult deaf or hard-of-hearing people. To make this product accessible do all people within this demographic, the documentation must be writing towards middle school level readers.

[62] Step-by-step instructions will be given for: driver & software installation, putting on and adjusting the motion capture device, and how to calibrate the glove

[63] The product will also come with a 10-30 page manual on how to use the software interface and to properly use the motional capture device to get best results.

[64] All documents will be written so that a user with no electronic knowledge can successfully use the product

[65] The manual will also include an ASL chart, FAQs, troubleshooting section, and complete product device specifications and limitation section.

[66] First stage product training and setup will be done with Voiceture system developers on hand

[67] Additional information, including software and drivers updates, that is released after the product will be available via the product website.

8. Conclusion

The Voiceture system requirements described in this document are accurate and rigorous to the best of Fivepoint technologies engineering team knowledge. Notice that these specifications are subject to change during the product development process due to either new user requirements or technical requirements that are likely to happen due to the nature of this product being the first of its type. However, any specifications will only serve to enhance the device as Fivepoint technologies believes in delivering a customer/user focused device with high regard for quality and usability.

The proof of concept device being delivered for December 2006 will focus on the core technologies as outlined in the requirements. All requirements will be fully met with the exception of the tentative specifications marked by a *.

9. Reference Documents

[1]   Project proposal for Voiceture System. Fivepoint technologies,
[2]   American Sign Language Browser. Michigan State University.
       Communication Technology Laboratory.
       http://commtechlab.msu.edu/sites/aslweb/browser.htm
[3]   Java SDK 5.0 documentation. J2SE platform.
       http://java.sun.com/j2se/1.5.0/docs/
[4]   Interfacing the Serial RS232 Port.
       http://www.beyondlogic.org/serial/serial.htm