



January 16, 2004

Lakshman One
School of Engineering Science
Simon Fraser University
Burnaby, British Columbia
V5A 1S6

Re: ENSC 440 Project Proposal –Voice Recognition System in an MP3 Player

Dear Mr. One:

The attached document is a *Proposal for a voice recognition system in MP3 player*. We are currently working with Start Labs Inc., whose product, an MP3 player, is to be controlled by the voice of the user. Our design is the voice recognition module of the product. We will ensure that the design meets Start Labs Inc.'s expectation and needs in most effective ways.

This proposal includes a system overview, a tentative budget and funding. We have found a few viable solutions or designs and they are discussed and compared in the system overview. A tentative schedule of the project progress is also added in this document

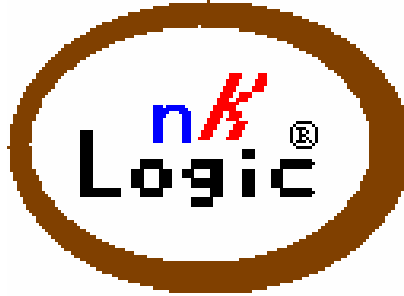
nK Logic consists of two experienced senior engineering students: Won Kang and Gareth Kim. We are looking forward to your feedback and suggestions. Please feel free to contact me by phone at (604) 785-5933 or by e-mail at gkim@sfu.ca. Thank you for your attention.

Sincerely,

A handwritten signature in black ink, appearing to read 'Gareth Kim', with a long horizontal flourish extending to the right.

Garet Kim
nK Logic

Enclosure: *Proposal for a voice recognition system in a MP3 player*



Proposal for a Voice Recognition System in MP3 Players

Project Team: Won Kang
Garet Kim

Contact Person: Garet Kim
gkim@sfu.ca

Submitted to: Lakshman One – ENSC 440
Nakul Verma – ENSC 440
Mike Sjoerdsma – ENSC 305
School of Engineering Science
Simon Fraser University

Issued date: January 16, 2004

Revision: 1.1



Executive Summary

nK Logic is a student team in ENSC 440 working with Start Labs Inc. to develop the voice recognition and processing module, later to be integrated with other modules to create a wireless MP3 player. To enter and thrive in today's MP3 player market as a startup company, not only design with advanced technology matters, but also competitive price of the product is vital to survive in the market.

Customers in the consumer electronics market are exceptionally cost-sensitive. For MP3 players, whose price typically ranges from \$100 to \$400, even \$10 difference in the unit price can significantly impact a customer's perception of the product. However, when a company is exceedingly concerned with lowering the unit price and not caring about the product's design and functionality enough, the reputation of the company suffers. Then, it is often costly to recover the lost reputation.

At nK Logic, two senior engineering students bring their excellent working-ethics combined with experiences in engineering companies to guarantee that the voice recognition and processing module for Start Labs Inc. gets realized in a cost-efficient manner and functions in fully expected ways.



Table of Contents

Executive Summary.....	ii
Introduction.....	1
System Overview.....	1
Comparison Chart.....	3
Budget and Funding.....	4
Schedule.....	4
Conclusion.....	5
Reference.....	6



1. Introduction

Start Lab Inc. is a company specializing in wireless MP3 player with a voice activated remote controller. Their MP3 player utilizes Bluetooth for connection between the main unit and the remote controller. The remote controller features a voice activated commands so as to give the users some degree of autonomy.

The key characteristics which distinguish an MP3 player among many competitors include price, quality, size, battery lifetime, and unique features like voice activated controller. nK Logic participates in this project as a group which is responsible for the research and development of this voice activated controller unit.

Our specific duties are research and comparison of the possible solutions, detailed implementation plans, and development of the working prototypes. Software components will be evaluated in terms of the efficiency, quality, reliability, and portability of the implementation. Hardware components will be evaluated by memory space for data and RAM, processing power required, and the interface with the main units.

The ultimate solution of this project will yield voice activated controller units which will have a debugging tool in connection with a PC; an easy to implement interface with the main units utilizing one of I2C, UART, and SPI; and an expandable demonstrating circuit for the future enhancement.

This document is a proposal of such a device by nK Logic including a system overview, the budget plan, and the project schedule. Throughout the project, we hope to improve various engineering and entrepreneurial skills as well as team dynamics. The members of nK Logic are proud to be involved in this project and we expect to contribute to the successful completion of the project.

2. System Overview

Start Labs Inc. is currently building an MP3 player integrated in a headphone. This device can be controlled through either the user's command via the microphone attached to it or a watch that the user wears on his/her wrist. Our goal is to design the voice activated controller, which is an integral part of the MP3 player. The basic functions of the controller are as follows:

- Receive voice command from the user through the microphone
- Recognize the command
- Send the corresponding signal to the MP3 main unit (or PC debugger)

Figure 1 illustrates the high-level design of the system.

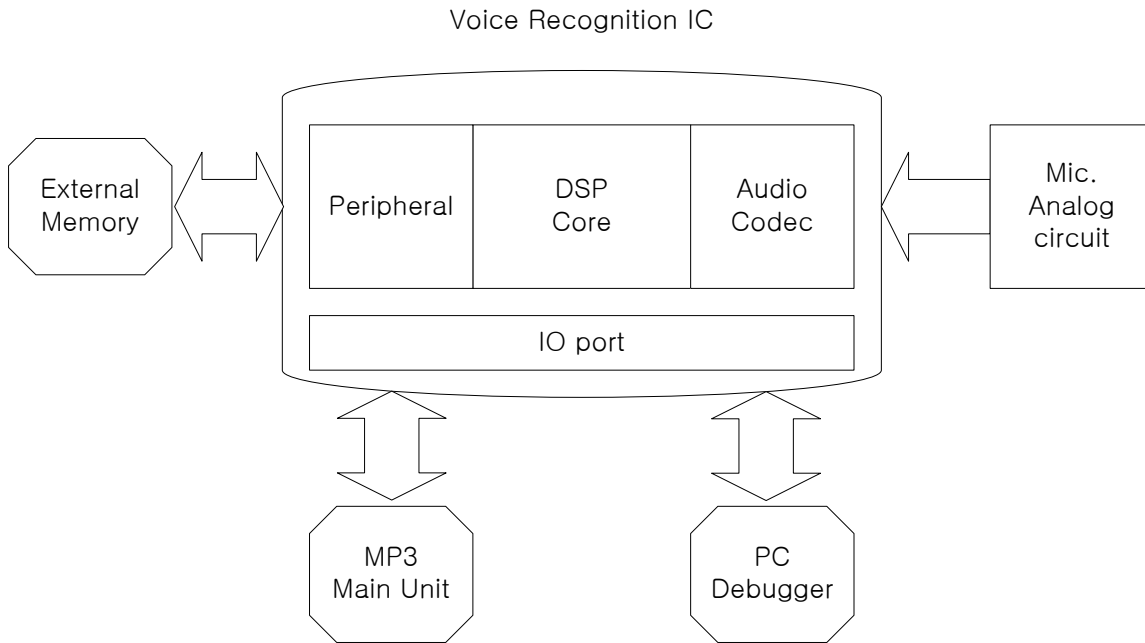


Figure 1: System Overview

Start Labs Inc. has also suggested a certain set of constraints that the design should demonstrate:

- Acknowledges about 200 command lines
- Operate in Speaker Dependent mode. In other words, the user must be able to “train” the system to achieve high voice recognition ratio.
- Low Power
- Small Package
- Inexpensive cost
- I/O interface

We should make sure that we put these constraints under consideration when choosing a particular voice recognition chip.



3. Comparison Chart

So far, we have found 3 chips specialized in voice recognition as shown in Table 1.

Table 1: Comparison of Voice Recognition Chips

Manufacturer	Sensory Inc.	Sensory Inc.	Voiceware
Product	Voice Direct II	Voice Extreme	ZVSR 600/620
Core	8-bit CPU	8-bit CPU	16-bit Fixed Point 100MIPS DSP
Additional Memory Req'd	2MB Flash	2MB Flash	N/A
External Memory Bus	Flash	Flash	N/A
Maskable ROM	0	64KB	N/A
Internal ROM	N/A	N/A	Internal 64KB ROM
Speech Duration (Max.)	40 sec.	100 sec. (ext. flash)	N/A
RAM	2.5 KB	2.5 KB	8KB
I/O	0	14	16
Key Technologies	SD, CL	SI, SD, SV, CL	SI, SD
SI words on chip	0	350 (ext. flash)	unlimited
SD/SV words on chip	60 (ext. flash)	1900 (ext. flash)	N/A
Packages	TQFP-64	TQFP-64	TQFP-80
100k die price	<\$1.50	<\$1.50	N/A
Power dissipation	3.0V, 10mA	3.0V, 10mA	3.0V, 10mA

Our recommendation is to use Voice Extreme. Voice Direct may not be adequate in a system that needs to deal with 200 commands. On the other hand, ZVST 600 would be too powerful for our need. The optimal choice is Voice Extreme for our applications to handle 200 words in a power-efficient, small-packaged chip.



4. Budget and Funding

Table 2 indicates our tentative budget for the voice recognition and processing module. Acoustic accessories refer to equipments such as microphones and speaker. Fifteen percent of contingency fund has been put into account.

Table 2: Tentative Budget

Item	Cost (in CAN \$)
Voice Recognition Toolkit	170.00
Development Software	50.00
Acoustic accessories	20.00
Cables	20.00
Case	20.00
Contingencies (15%)	40.00
Total	320.00

Start Labs Inc. may purchase these items for us. Alternately, we may apply to the Engineering Student Society Endowment Fund (ESSEF) and Whigton Development Fund.

5. Schedule

The following figures illustrate the project schedule. With this project, we are strongly focusing on the testing and troubleshooting stages as it is very important to have a final product ready to be used by Start Lab without any modification or alteration.

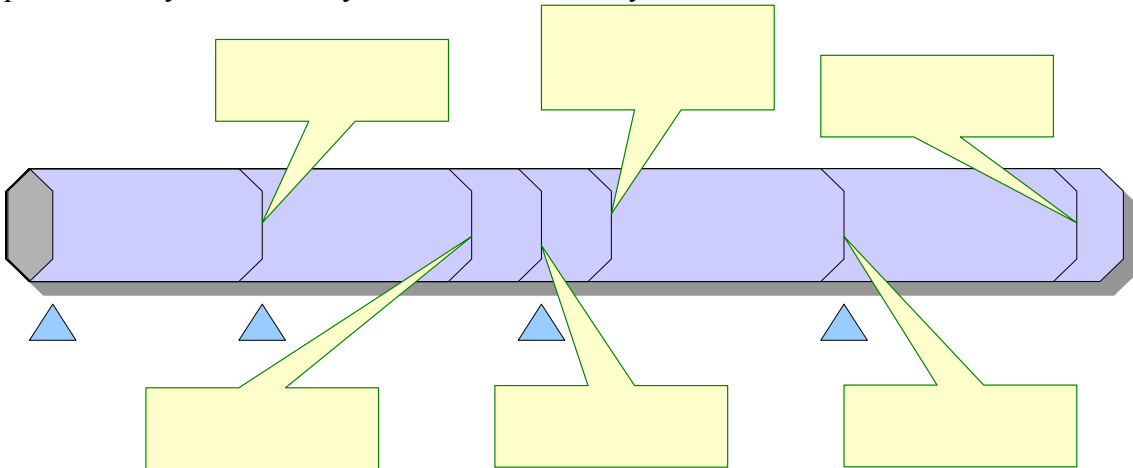


Figure 2: Timeline Milestone

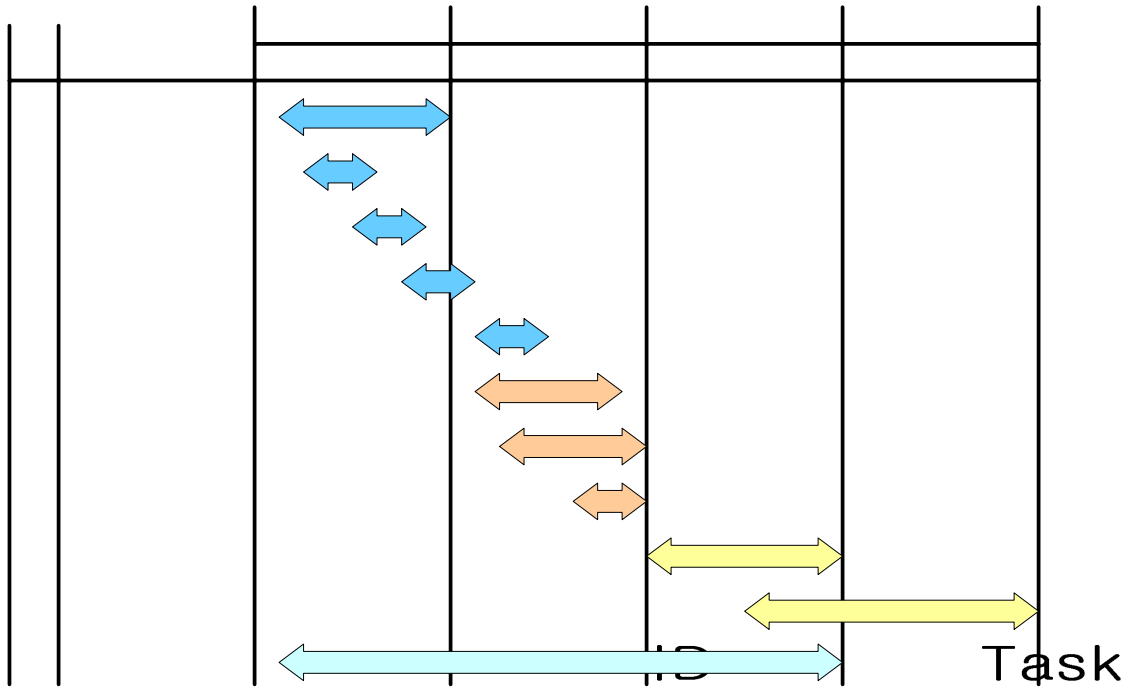


Figure 3: Gantt Charts

6. Team Profile

nKLogic consists of 2 senior engineering students at SFU, Won Kang and Garet Kim. They have known each other for more than 4 years. They previously worked in a group for ENSC 151 project. They now form a team nKLogic to assist StartLabs in their product development.

Won Kang

Won is an n-th year student graduating SFU engineering this summer. Between 2000 and 2003, he worked at iCable System, Seoul, and S. Korea as a researcher. Through the employment at iCable, Won specialized in many aspects of developing Cable/ADSL modem, such as DOCSIS, RTOS, VoIP, and programming DSP. Won brings great knowledge that is directly related to this particular project which will lead nKLogic to success.

Garet Kim

Garet is also an n-th year engineering student at SFU.

1 Project Research

2 Project Proposal

3 Functional Spec.

4 Design Spec.

5 Progress Report

6 Hardware development

Software



He did his coop at Nuvation Labs, San Jose, CA as a hardware engineer. Nuvation Labs is a small consulting firm in embedded systems. Working at a consulting company taught him general problem-solving skills and gave him insight in system development. Garet is naturally a hard worker, and he will help nKLogic to accomplish its goals on time.

7. Conclusion

Based on our expertise, knowledge, skills, and commitment that we have, nK Logic will make sure to complete the project on time without a failure. We are pleased to have this opportunity with Start Lab Inc. and hope that this project will benefit both nK Logic and Start Lab in great deal upon the completion of the project.

8. Reference

Bell Lab Speech Recognition Project

(<http://cm.bell-labs.com/cm/ms/departments/sia/project/speech/index.html>)

Voiceware: (<http://www.voiceware.co.kr>)

Sensory Inc. (<http://www.sensoryinc.com>)