SYSTEM DEVELOPMENT IN SEARCH DATA BOOK LIBRARY USING VOICE COMMANDS

Bambang Yuwono
Informatic Department UPN "Veteran" Yogyakarta
Jl. Babarsari no 2 Tambakbayan 55281 Yogyakarta Telp (0274)-485323
Email : masbambangy@yahoo.com

Abstract

Library is a place where people read and find the reference. There are many references of knowledge, good general knowledge, kodekteran, science and technology, novels, and so forth. Currently, each library has a system there is a search for books you want the visitors, but the equipment used is still standard, uses the mouse and keyboard. Visitors type in a word for word to search for books using the keyboard. In the current era such as this, it is necessary that all the applications automatically and efficiently, both to ease the visitors, but also to follow the flow of technology is very rapid. With this application, the visitor does not need to type for word, just the book title only, the system will automatically display the books you want.Book Search Data System in the Library Using the Voice Command was created using this programming language Visual Basic 6.0 and Microsoft Access 2003 database and program as a supporting vote SAPI 5.1. Method using the Waterfall development system. This system is tested by using black box testing method. Black box method is a method of testing to test a system without having to know the internal process that is on the system.Results from this research include data search system in the library using books as a medium for voice search. This application provides information about books that will be sought, with only mention the book title in English through a microphone.

Keywords: Searching, Library, Voice Command

1. Introduction

Voice recognition technology has grown rapidly, so that many applications that have been adapting technologies. The voice recognition technology makes it easier when the user interacts with the application that supports voice commands. Convenience is the form of voice commands directed to an application, so that the input speed compared with the pengetikan for word.

Data search system in the library books in general, use only the tools of mouse and keyboard as a medium for interaction, so not everyone is able to use these applications, especially those who lack the physical, for example, disabled hand. Some examples of applications that have been using the voice recognition technology as a medium for interaction between the user and the computer is a document editor, voice mail applications, games, edutainment, application for handycapped people, and others.

Library is a place where people read and find the reference. There many different book categories, such categories of agriculture, medicine, public, computer, magazines, newspapers and so forth. Currently, each library is equipped with computer equipment, that allows the user to process the search in books, but usually the computer is only equipped with standard equipment such as mouse and keyboard. Equipment, supplies will not be useful if the user is a physical disability, for example, disabled hand.

Research on the making of book data search system in the library using the voice command is very useful to eliminate the dependence on user mouse and keyboard tools. In addition, to help users of the defect, so that confusion does not need to seek help for the book you are looking for. This application has 2 search method, namely using the mouse and keyboard or use voice commands in the search process.

In using voice commands, this application requires SAPI (Speech Application Programming Interface) is used as a tool for processing analog voice and then converted into text. Minimum equipment that is used in running this application is a set of computers equipped with microphone.

2. Overview references

2.1. Voice

Sound is mechanical compression or longitudinal waves that spread through the medium (Anonymous, 2006). Most of the sound is a combination of various signals, but pure sound can be described theoretically with the speed or the oscillation frequency measured in Hertz (Hz) and amplitudo or the sound pressure measurement in decibel. Quoting the voice can be heard by the human ear is approximately from 20 Hz to 20 kHz in general amplitudo with different variations in kurva responsnya. Voters over 20 kHz is called ultrasonic and below 20 Hz is called infrasonik.

2.2. SAPI Technology (*Speech Application Programming Interface*)

SAPI 5.1 (Speech Application Program Intervace) 5.1 is one of the API (Application Programming Interface) that are provided as an introduction to programming using the method of speech recognition and Text To Speech (help, overview SAPI 5.1). With the SAPI allows the system to implement the user's voice by using a different engine without changing the application that was created [1].

There are two kinds of voice mode on or wicara, namely [1]:

1. *Dictation Mode*
   This method on computer users can say words or sentences, which will then be recognized by the computer and converted into text data. The possibility that the number of words recognized is limited by the number of words that have been found in the database. Introduction to Speaker Mode Diktasi is dependent. Accuracy introduction this mode depends on the pattern of the sound and accent of the speaker and the training was done.
2. *Command and Control Mode*
   On this method, computer users say the word or sentence that has been didifinisikan first on the database and will then be used to run a command on a particular computer application. The number of commands that can be recognized depending on the application that has difinisikan first database on the types of commands that can be dieksekusikan. This mode is a Speaker Independent because the number of known words is usually limited and there is a possibility once the speaker does not need to do training on the system before.

There are four main processes in the voice recognition system, both at the Dictation mode or Command and Control, namely [1]:

1. *Word separation*
   The separation process is said to memisahakan sound spoken by the user into several sections. Each section can be a sentence or just a word. There are three kinds of methods that can be used in the process of separation of the word, namely [1]:
   a. *Discrete Speech*
   On this method, users are required to sentence the terpenggal with the momentary pause between words. Pause is used by the system to detect the beginning and end of a word. Discrete Speech has a few advantages, namely resource (within the computer memory) used by the system to detect the sound, but have weaknesses that inconvenience to the user in a sentence.
   b. *Word Spolling*
   In a spoken sentence that the user, the system detects only the words that are held in the treasury, and ignore the other words it has not. So even though the user to a different sentence but the sentence is that there is a word vocabulary and the same system, the result of the introduction will be the same. The weakness of this method is likely the system will make a mistake in introducing the meaning of the sentence. But this method has the advantages that the user can say the sentence to normal without having to stop between the words.
   c. *Continuous Speech*
   The system will recognize and process each word spoken. This method will result in accuracy to identify a user greeting. But besides that, this method requires a resource (computer memory within the process) is large in the process. On this method the system should be able to detect the beginning and end of each word in the sentence without any pause between words, and after successfully separating the word, the next step is matching with the vocabulary that the.

2. *Dependence on the user*
   The voice recognition system has some character, seen from the dependence of users, namely [1]:
   a. *Speaker Dependent*
   The system requires training for each user who will use the system. The system will not be able to identify users who have not been doing the training.
   b. *Speaker Independent*
   Users do not need to do training before you can use the system, because the system is able to identify the user's voice does not sound and color depending on the dialect used.

3. *Speaker Adaptive*
   Speaker is an adaptive combination of speaker dependent and speaker Independent, where the user does not need to do the training and introduction of the accuracy of the system will increase if the same user to work continuously for several time

4. *Matching words*
   Matching process is the word for word speech matching successfully identified with the database that the system.
   There are two methods that can be used in the process of matching words, namely [1]:
   a. *Whole-Word Matching*
   The system will search for data dibasis said the exact same results with the word user utterance.
   b. *Phoneme Matching*
   The system has a dictionary phoneme. Phoneme is the smallest and unique voice that's a form of words.

5. *Vocabulary*
   Vocabulary is the last in a voice system. There are two things to note on the vocabulary, namely the size and accuracy. If the vocabulary of many a system will be easy to match in words, but with the increasing number of vocabulary, the number of words that have almost the same speech also mengingkat, which reduce the accuracy of this introduction. And so vice versa, if a system has a small vocabulary, then the accuracy will be higher because at least words that are similar, but will be more and more words are not recognized [1].

For the system with the introduction pembicaaarn mode Command and Control, will be better if the number of vocabulary slightly (less than 100 words), but for Dictation mode will require more amount of vocabulary a lot.

On this research, the applications are text-based application to receive diktansi language indonesia. Due to the application, then the mode is command and control process with continuous speech and speaker on adaptive users.

Next SAPI *(Speech Application Program Interface)* 5.1 the outline can be divided into two main structures that function to process Speech to Text or Text To Speech 1 image sections, namely:

1. *Aplication Programming Interface* (API)

   On the voice recognition application in Microsoft Word, the API function of Microsoft Word application with the SAPI Runtime.

2. *Divace Drive Interface* (DDI)

   DDI provides a function to receive voice data from the SAPI to return to the Runtime phrase in the introduction of SAPI most basic level. There are two engines that are used by the DDI recognition engine that functions as a modifier to be text or voice writing and TTU engine (Text to Speech Engine). In designing the application using the recognition engine is only just, because TTU only work for the implementation of applications based Text To Speech.

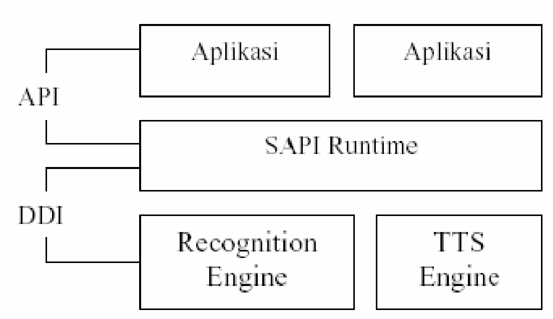
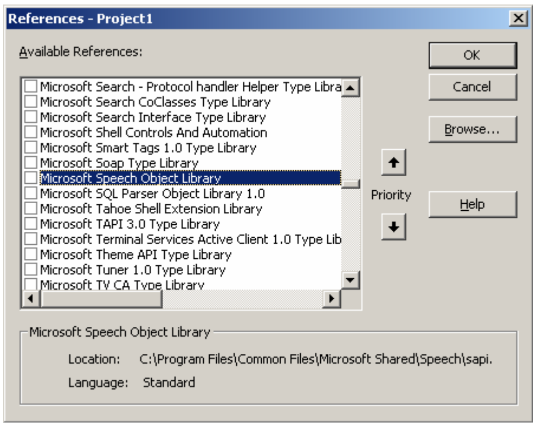


Figure1 Structure SAPI SDK 5.1

2.3 SAPI.dll

SAPI.dll file is the file in the main cattle 5.1, this file should be placed in the system directory. Berekstensi the dll file (Dynamic Link Libraries) are used so that interfaces and objects that have elements of property, method and event of the cow (Speech Application Programming Interface) can be accessed, which interfaces with objects and elements in the cow is used to build an application that can recognize the human voice input [2]. To build an application required SAPI.dll files that have been combined with a particular programming language. In making this application using the programming language Visual Basic 6.0. To add a file SAPI.dll in Visual Basic 6.0 can be done by choosing Microsoft Speech Object Library in the menu Project - References. Appearance can be seen in Figure 1


References - Project1

Available References:

□ Microsoft Search - Protocol handler Helper Type Libra
□ Microsoft Search CoClasses Type Library
□ Microsoft Search Interface Type Library
□ Microsoft Shell Controls And Automation
□ Microsoft Smart Tags 1.0 Type Library
□ Microsoft Soap Type Library
□ Microsoft Speech Object Library
□ Microsoft SQL Parser Object Library 1.0
□ Microsoft Tahoe Shell Extension Library
□ Microsoft TAPI 3.0 Type Library
□ Microsoft Terminal Services Active Client 1.0 Type Lib
□ Microsoft Theme API Type Library
□ Microsoft Tuner 1.0 Type Library
□ Microsoft TV CA Type Library

Microsoft Speech Object Library
Location: C:\Program Files\Common Files\Microsoft Shared\Speech\sapi.
Language: Standard

Figure 2 Dialog Box *References*

Interfaces and objects of SAPI 5.1 used in the making of this system is as follows:

1. *ISpeechRecoGrammar interface*

   ISpeechRecoGrammar interface is the interface that allows an application to organize and arrange the words to be recognized by Speech Recognition engine. In the interface there is some method, and method used in the making of this system is as follows.

   a. *CmdLoadFromFile method*

      CmdLoadFromFile method used to open a command and control grammar (command and control) of the file that has been set.

   b. *CmdSetRuleIdState method*

      CmdSetRuleIdState method used to activate or deactivate grammar. Introduction of speech so that the system can work, then the grammar must be enabled.

2. *ISpeechRecoResult interface*

   Interface this function to return the information about the introduction of the business. Results returned by the introduction of recognition in the context of a successful or an unsuccessful recognition recognition (or the introduction of the wrong).

3. *ISpeechPhraseInfo interface*

   This interface contains properties and methods that can be used to obtain information that contains the order-word order. Interface in this method there is a gettext function to return the text of the introduction that has been done, and the text is a single string.

4. *SpSharedRecoContext object*

   SpSharedRecoContext object is an object class that has a relationship (relationship) with IspRecoContext interface. The object recognition context is used to identify when an application will be active, context recognition is also possible to start or end the introduction, set the grammar, and receive important notifications. Recognition recocontext used so that applications can interact with SAPI (Speech API). IspRecoContext interface can be used at the time SpSharedRecoContext object was created. IspRecoContext interface is the main interface used in creating a voice application. IspRecoContext is the application to receive notification of an event the introduction of speech.

IspRecoContext in the interface there are several methods and events. Method and the event's IspRecoContext interface used in the making of this application are as follows (Anonymous, 2000).

a) CreateGrammar method

This method is used to create ISpRecoGrammar.

b) Recognition event

This event is used at the Speech Recognition engine produces an introduction.

c) Sound Start event

This event is used when the initial vote in the audio input stream by the Speech Recognition engine.

d) SoundEnd event

This event occurs when the final votes that have been spoken in the audio input stream by the Speech Recognition engine.

5. *SpVoice Object*

SpVoice Object is an object that brings ability of Text To Speech engine to the application using the Speech API. This object allows an application to conduct the conversion process from the input text into a form of voice output.

Object interfaces and its methods and its event is needed in making the voice recognition system. At the time the object SpSharedRecoContext created, then the application will be able to relate to the audio input, Speech Recognition engine and grammars through Speech API, this is because the SAPI is a bridge / mediator between the application engine with Speech Recognition and Text to Speech engine

At the time the word or sentence spoken, the word or phrase will be sent to the Speech Recognition engine to recognize through the SAPI. In recognizing the expressions, Speech Recognition engine to use grammar to enhance its ability to recognize in the word or sentence. To call a grammar that has been made, the first step that must be done is to create objects ISpRecoGrammar first, this can be done by using ISpRecoContext:: CreateGrammar. Furthermore, the grammar that has been made is placed in memory (loaded), which can be done with the call ISpRecoGrammar:: CmdLoadFromFile. Then the grammar must be enabled in order for the introduction of speech system can work, this can be done with the call ISpRecoGrammar:: CmdSetRuleIdState.

Method Development System Waterfall

System development methodology is used, or model sekuensial linier also often called the classic life cycle or waterfall model. Waterfall model proposes an approach to software development and the systematic sekuensial a start on the level and progress of the whole system analysis, design, code, test, and maintenance. Waterfall model encompass activities as follows [4]:

1. Engineering and modeling

   In this phase services, constraints and objectives resulted from consultation with users system. Because the software is always part of a larger system, work begins to build with the requirements of all elements of the system and allocating some subset of the needs to the software.

2. Requirement Analysis

   The process of intensif needs and focus, especially on the software. To understand the nature of the program that was built, the analysis must understand the domain of information, behavior, performance, and inter-face is necessary. The need for better systems and software documented.

3. Design

   Design software is a multi-step process that focuses on the four attributes of a different program, data structures, software architecture, interface representation and procedural algorithms. The process of translating the design requirements / needs to representation in the software that can be estimated for the quality of the code before the eruption began. As the requirements, design documented and become part of the configuration software.

4. Coding

   Code generation should be translated into the form of a machine that can read. Then the design was changed in the form of program code. If the design is done in a manner that is complete, making the code can be mechanically.

5. Testing

   Once the code is created, a test of the program can be done. The process of testing focuses on the internal logic of software, ensuring that all the statements has been tested, and on the external drive that is functional testing to find errors and to ensure that input is limited akan provide the actual results in accordance with the required results.

6. Maintenance

   Software changes after delivered to the user. The changes will occur because of errors is determined, because the software must be adjusted to accommodate changes in the environment eksternalnya (for example, changes that are required as a result of device peripherals or a new operating system or because the user requires the development of functional or performance. Maintenance devices again apply the software program before each phase and does not create a new one again.

3. Research Method

The methodology used in this research, among others; Literature Studies, SDLC (System Development Life Cycle), which includes the Analysis, Design, Implementation, Testing and Maintenance.

1. A Literature Review

   This stage is the stage of collecting information and literature that are required for the system. The information and literature that is used among the Voice Recognizer, SAPI 5.1 and SAPI 4

2. Analysis and design

   At this stage, do the analysis and design required in the system, among DFD design, database design, and user interface design

3. Implementation

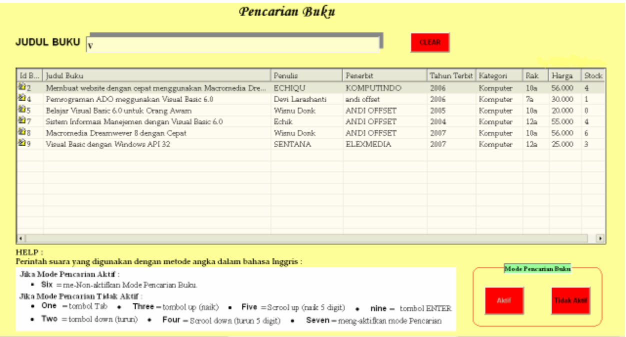
At this stage, the design of the system have been made will be implemented using Visual Basic and Microsoft Access and program support SAPI 5.1 and SAPI 5.

4. Testing and evaluation

At this stage, will be testing and evaluation of the system will be done and improvements are needed.

4. Discussion

Figure 3 shows the user form that is used to search for the desired data book. By using voice commands, the user does not need to type the book title for word. This application is using cow 2 (Speech Application Programming Interface), which is SAPI 4 and SAPI 5.1. At the time of active search mode, the user can specify the books you want, this system automatically displays all the books that have titles that match the words spoken by users. When the search mode is active, the application method Speech Dictation, where every utterance will be converted into text by a SAPI, the SAPI is a role here is 5.1. If the search method is not active, the application method to enable Voice Command, where every utterance will be processed by a SAPI with the assistance in the form of grammar. Txt files into a simple command to Windows, for example, the tab key, up, down, scroll up and scroll down. SAPI 4 that contribute to this method. Display search form the data shown in the figure 3.



Pencarian Buku

JUDUL BUKU v

Id B...	Judul Buku	Penulis	Penerbit	Tahun Terbit	Kategori	Rak	Harga	Stock
2	Membuat website dengan cepat menggunakan Macromedia Dre...	ECHIQU	KOMPUTINDO	2006	Komputer	10a	56.000	4
4	Pemrograman ADO meggunakan Visual Basic 6.0	Devi Larashanti	andi offset	2006	Komputer	7a	30.000	1
5	Belajar Visual Basic 6.0 untuk Orang Awam	Wisnu Donk	ANDI OFFSET	2005	Komputer	10a	20.000	0
7	Sistem Informasi Manejemen dengan Visual Basic 6.0	Echik	ANDI OFFSET	2004	Komputer	12a	55.000	4
8	Macromedia Dreamwever 8 dengan Cepat	Wisnu Donk	ANDI OFFSET	2007	Komputer	10a	56.000	6
9	Visual Basic dengan Windows API 32	SENTANA	ELEXMEDIA	2007	Komputer	12a	25.000	3

HELP :
Perintah suara yang digunakan dengan metode angka dalam bahasa Inggris :

Jika Mode Pencarian Aktif :
- Six = me-Non-aktifkan Mode Pencarian Buku.

Jika Mode Pencarian Tidak Aktif :
- One = tombol Tab
- Three = tombol up (naik)
- Five = Scrool up (naik 5 digit)
- nine = tombol ENTER
- Two = tombol down (turun)
- Four = Scrool down (turun 5 digit)
- Seven = meng-aktifkan mode Pencarian

Figure 3. Display Search Data Book

5. Conclusion

Based on the results of the analysis, design, and implementation, have successfully built the application data in a library book using voice commands that can facilitate a user in finding the desired book. Overall, the program has been running well. Only on the search process there is little data book shortages. Process the query using the English language, and the accuracy of arrest sensifitas vote by user programs are still lacking, so that the user needs to repeat several times in pengucapannya, so the need to practice speaking (English).

References

[1] Adipranata, Rudy, *Implementasi Sistem Pengenalan Suara Menggunakan SAPI 5.1 dan Delphi 5*, Jurnal Informatika, 2003, <http://www.petra.ac.id >

[2] Microsoft, *Microsoft Speech SDK 5.1*, <http://www.microsoft.com>

[3] Piniones, Raymond, *Inventory System using listview with ado connection*, http://www.a1vbcode.com, 2003

[4] Pressman, Roger, *Rekayasa Perangkat Lunak Pendekatan Praktisi*, Andi Offset, Yogyakarta, 2002.

[5] Priyan, *Voice Commander*, <http:// www.planetsourcecode.com, 2007