Interactive Voice Response for Educational Institution

Roshini Shanbag
Dept. of Electronics and Communication
CIT, Kodagu, Karnataka, India

Akshay
Dept. of Electronics and Communication
CIT, Kodagu, Karnataka, India

Pooja K
Dept. of Electronics and Communication
CIT, Kodagu, Karnataka, India

Kanthi
Dept. of Electronics and Communication
CIT, Kodagu, Karnataka, India

Abstract— We Interactive voice response systems (IVRs) acts as bridge between people and computer databases connect telephone users with the information they need, from everywhere at any time. These systems have been around for more than a decade, and in modern times they are playing vital role in all real time applications. Most of today's IVR and transaction processing applications provide a user interface which employs a touchtone or dual-tone multi frequency (DTMF). However, for ease of use and for better user interface, applications that allow callers to use their own voice rather than DTMF inputs or a regular touch-tone to complete transactions are known as speech enabled interactive response systems (SEIVRS) and they are rapidly rising as the latest innovation in telephony-based remote self-service.

Keywords— GSM, LCD, Renesas microcontroller, sd card, FN-M16P.

I. INTRODUCTION

The IVR systems were the first IT application to start up government services to the public 24X7 hours. During busy hours we find it more difficult to reach operators and other key department staff. But IVR can save the public and private industry's time. The input modality used is DTMF in phone based systems. There are certain types of tasks which are classified as linear tasks i.e. hearing message in the order they are received and another task is known as nonlinear task where we can here the message from a particular sender in random order.

Research studies indicate that DTMF was more effective and efficient for linear tasks, whereas for nonlinear tasks, speech was given higher priority. Due to this reason speech was preferred over DTMF by a majority of users. As customer satisfaction is given utmost priority in customer service providers, there is constant need to upgrade the technologies to provide better grade of service. As compared to DTMF based IVR, speech enabled IVR not only handles more difficult operations but also enhance the customer satisfaction.

Interactive Voice Response systems can play a significant role in providing efficient customer service. Properly implemented, they can increase customer satisfaction, lower costs and offer new services. The return on investment (ROI) on these systems is also quite amazing, making them the most popular Computer Telephony systems in the world. Compare them to a call center.

The price for the extra "human touch" translates into a huge running cost in the form of Agents, Supervisors, infrastructure maintenance, training, call center performance & discipline reviews, etc. World over, the first systems that any company deploys with a view towards enhancing customer satisfaction are IVR's. Call centers come much later. IVR's can provide information to callers in one of two ways:

A. Pre-recorded Information.

Common examples are audio movie snippet previews (e.g. at PVR). Though it is possible to build these IVR's through live information from databases (using text-to-speech engines), one doesn't get the voice variations, which are so important for the moviegoer. Other examples are around procedural (or "how to") information dissemination like Income tax filing procedures, bank account opening or credit card application procedures, etc.

B. Live information from databases.

These IVR's get information from databases, convert to voice, and speaks it back to the caller. Practically all industry segments are potential users for this, and examples include Phone banking (where you call in, dial in your account number & TPIN and can hear your account balance on phone) Courier packet trace (where you call in, dial the AWB number, and the system tells you whether the packet has been delivered, if it is in transit, etc)

II. LITERATURE SURVEY

[1] Paper on "Assamese Spoken Query System to Access the Price of Agricultural Commodities" by the authors Shahnawazuddin, Deepak Thotappa, B D Sarma, S R M Prasanna and R Sinha EEE Dept IIIT Guwahati. Technical method was on 1.IVR 2.ASR (HMM Based) Both are open source had a research objective on Enquire systems for Agricultual Commodities. [2] Paper on "Deploying Usable Speech Enabled IVR Systems for Mass Use" by the authors Chitralekha Bhat, Mithun BS, Vikram Saxena, TCS Innovation Labs - Mumbai, Yantra Park". Technical method was on 1.Wizard Oz(WoZ) if ASR fails to Recognize 2.TTS had a research on Speech Enable Railway Enquire Systems (SEREC). [3] Paper on "Experimental Comparison of Speech and DTMF for Voice XML Based Expert Systems" by the authors Uwadia Charles Onuwa Akinwale Adio Taofeek3 . Technical method on 1.Voice Objects Desktop for Eclipse 9

Special Issue - 2018

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCESC - 2018 Conference Proceedings

2.Voxeo Prophecy 8 3. XLite soft phone had a research objective on Health Dialogue Expert System. [4] Paper on "Telephony Speech Recognition System: Challenges."by the authors 1.Joyanta Basu 2. Rajib Roy 3. Milton S. Bepari 4. Soma Khan. Technical method on 1.CMU-SPHINX for ASR 2.Asterisk PBX for acts as Telephone Exchange. Had a research objective on, to identify the Challenges in Telephony Speech. [5] Paper on "IVRS For College Automation" by the authors Santosh A. Kulkarni Dr. A.R. Karwankar Technical method on 1.Goertzel algorithm, 2.Dual-tone multi frequency signaling (DTMF), 3.speech synthesizer. Had a research objective on Enquire systems for College Automation. [6] Paper on "Speech Vs Touch: A Comparative Study of the Use of Speech and DTMF Keypad for Navigation" by the author Kwan Min Lee Jennifer Lai. Technical method on 1.Statistical Based Language model for ASR 2.TTS for speech synthesis. Had a research objective on Mobile Assistant systems developed by IBM.
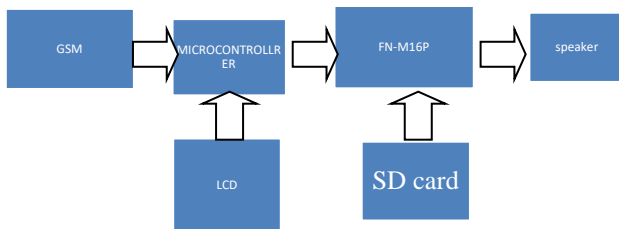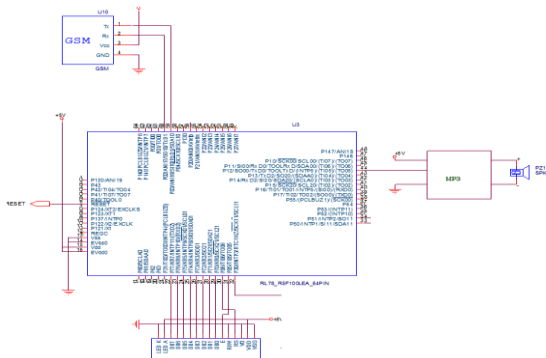
III. METHODOLOGY

Fig. 1. Basic Block Diagram

Fig. 2. Circut Diagram

This is a basic block diagram of the interactive voice response for educational institute. Initially all the device are connected to the microcontroller . The microcontroller used in the device is Renesas microcontroller. At one side of the microcontroller GSM device is connected on the other side is lcd display is connected to it. FN-M16P is connected to it. It even has the reset pin. Initially the sim is inserted in the gsm. Where the college number will be inserted in it. Futher it is connected to the microcontroller as shown in block diagram. In the microcontroller lcd display is connected to it. In the lcd display recorded audio to be displayed it displayed in the lcd.

The FN-M16P contains the SD card where all voice are recorded in it. This voice recorded is further connected to the speaker. This is the device that is built up. The reason for
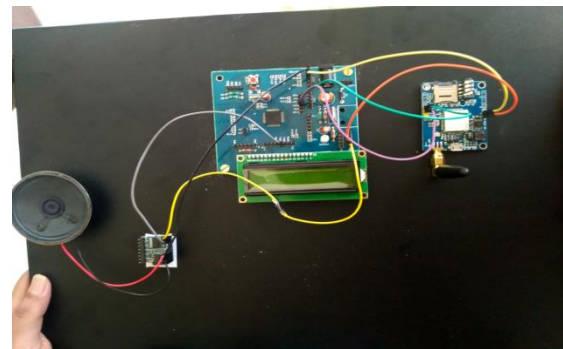
building up this this device in this format. Initially all the instruction will be given from college saying that "9876543210" is the college number to which you can your children progress report, attendance as well as fee structure.

The standardized number is inserted in the gsm. For example if one of the parents wants to know there children 1st internal marks parents need to send sms saying that last two digit of the USN and 1 for 1st interenal i.e, 121 , 12 is the usn number and 1 represents internal marks of all subject as well as fee structutre will be announced. Since it is the starting if the semester. Later 122, similarly 12 is the last two digit of the usn and 2 represents the 2nd internal marks. 123 similarly 3 represents the 3rd internal marks as well as attendance because 3rd internal will be conducted at the end of the semester therefore overall attendance of the student will be displayed. In this we can store as many students record as we want and reduces the human effort. This facility can be made available at any time.

Fig. 3. Implemented Board

IV. EXPERIMENTAL DETAILS

A. Alpha Numeric display

A liquid crystal display (LCD) is a flat panel display, electronic visual display, based on on Liquid Crystal Technology. A liquid crystal display consists of an array of tiny segments (called pixels) that can be manipulated to present information. Liquid crystals do not emit light directly instead they use light modulating techniques.

LCDs are used in a wide range of applications, including computer monitors, television, instrument panels, aircraft cockpit displays, signage, etc. They are common in consumer devices such as video players, gaming devices, clocks, watches, calculators, and telephones.

LCDs are preferred to cathode ray tube (CRT) displays in most applications because of

- The size of LCDs comes in wider varieties.
- They do not use Phosphor; hence images are not burnt-in.
- Safer disposal
- Energy Efficient
- Low Power Consumption

It is an electronically modulated optical device made up of any number of segments filled with liquid crystals and arrayed in front of a light source (backlight) or reflector to produce images in color or monochrome.

Special Issue - 2018

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCESC - 2018 Conference Proceedings

B. GSM

GSM stands for Global System for Mobile Communications formerly called as Group e Special Mobile. This is a standard set developed by the European Telecommunications Standards Institute (ETSI) to describe technologies for second generation (or "2G") digital cellular networks. The GSM standard initially was used originally to describe switched circuit network for full duplex voice telephony to replace first generation analog cellular networks.

The standard was expanded over time to include first circuit switched data transport, then packet data transport via GPRS (General packet radio service). Packet data transmission speeds were later increased via EDGE. The GSM standard is succeeded by the third generation (or "3G") UMTS standard developed by the 3GPP. GSM networks will evolve further as they begin to incorporate fourth generation (or "4G") LTE Advanced standards. "GSM" is a trademark owned by the GSM Association.

GSM networks operate in a number of different carrier frequency ranges (separated into GSM frequency ranges for 2G and UMTS frequency bands for 3G), with most 2G GSM networks operating in the 900 MHz or 1800 MHz bands. Where these bands were already allocated, the 850 MHz and 1900 MHz bands were used instead (for example in Canada and the United States). In rare cases the 400 and 450 MHz frequency bands are assigned in some countries because they were previously used for first-generation systems.

Regardless of the frequency selected by an operator, it is divided into timeslots for individual phones to use. This allows eight full-rate or sixteen half-rate speech channels per radio frequency. These eight radio timeslots (or eight burst periods) are grouped into a TDMA frame. Half rate channels use alternate frames in the same timeslot. The channel data rate for all 8 channels is 270.833 kbit/s, and the frame duration is 4.615 ms. The transmission power in the handset is limited to a maximum of 2 watts in GSM850/900 and 1 watt in GSM1800/1900.

Fig. 4. GSM SIM

One of the key features of GSM is the Subscriber Identity Module, commonly known as a SIM card. The SIM is a detachable smart card containing the user's subscription information and phone book. This allows the user to retain his or her information after switching handsets. Alternatively, the user can also change operators while retaining the handset simply by changing the SIM. Some operators will block this by allowing the phone to use only a single SIM, or only a SIM issued by them; this practice is known as SIM locking. We are be using SIM300 GSM Module in our Project.
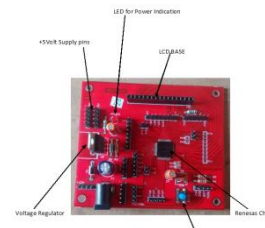
C. Microcontroller

Fig. 5. 64 Pin Microcontroller Based

Fig. 6.

- General-purpose register: 8 bits × 32 registers (8 bits × 8 registers × 4 banks)
- ROM: 512 KB, RAM: 32 KB, Data flash memory: 8 KB
- On-chip high-speed on-chip oscillator
- On-chip single-power-supply flash memory (with prohibition of block erase/writing function)
- On-chip debug function
- On-chip power-on-reset (POR) circuit and voltage detector (LVD)
- On-chip watchdog timer (operable with the dedicated low-speed on-chip oscillator)
- I/O ports: 16 to 120 (N-ch open drain: 0 to 4)
- Timer → 16-bit timer: 8 to 16 channels, Watchdog timer: 1 channel
- Different potential interface: Can connect to a 1.8/2.5/3 V device
- 8/10-bit resolution A/D converter (VDD = EVDD =1.6 to 5.5 V): 6 to 26 channels
- Power supply voltage: VDD = 1.6 to 5.5v
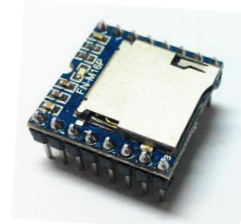- D. FN-M16P

Fig. 7. UART device

- Supports MP3 and WAV decoding.
- Supports FAT16 and FAT32 file system.
- 24-bit DAC output and supports dynamic range 90dB and SNR 85dB.
- Supports AD key control mode and UART RS232 serial control mode.
- Supports maximum 32GB micro SD card and 32GB USB flash drive.
- Audio files are sorted by folders; supports up to 99 folders, and each folder can be assigned to 255 sound files.
- Supports inter-cut advertisements.
- Supports playback of specifying folders.
- Support random playback.
- Built-in 3W amplifier that can direct drive a 3W/8Ohm speaker.

Special Issue - 2018

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCESC - 2018 Conference Proceedings

- 30 levels adjustable volume, and 6 levels adjustable EQ.

D. Speaker

Fig. 8. Speaker

A loudspeaker (or "speaker") is an electro acoustic transducer that produces sound in response to an electrical audio signal input. Non-electrical loudspeakers were developed as accessories to telephone systems, but electronic amplification-by-vacuum-tube made loudspeakers more generally useful. The most common form of loudspeaker uses a paper-cone supporting a voice coil electromagnet acting on a permanent magnet, but many other types exist. range. Miniature loudspeakers are found in devices such as radio and TV receivers, and many forms of music players. Where high fidelity reproduction of sound is required, multiple loudspeakers may be used, each reproducing a part of the audible frequency loudspeaker systems are used for music, sound reinforcement in theatres and concerts, and in public address systems. Technical Pro is a brand of speakers, amplifiers and other components for beginner and semi-professional disc jockey, live music, and home audio use. Technical Pro speakers are either active (built-in amplifier) or passive, with a variety of configurations and cabinet types.

The most common type of driver, commonly called a dynamic loudspeaker, uses a lightweight diaphragm, or cone, connected to a rigid basket, or frame, via a flexible suspension, commonly called a spider, that constrains a voice coil to move axially through a cylindrical magnetic gap.

When an electrical signal is applied to the voice coil, a magnetic field is created by the electric current in the voice coil, making it a variable electromagnet. The coil and the driver's magnetic system interact, generating a mechanical force that causes the coil (and thus, the attached cone) to move back and forth, thereby reproducing sound under the control of the applied electrical signal coming from the amplifier. The following is a description of the individual components of this type of loudspeaker.

V. SOFTWARE REQUIRMENT

- Integrated Development Environment (IDE) CubeSuite offers the ultimate in simplicity, usability, and security for the repetitive editing, building and debugging of codes.
- Easy to Install and operate.
- CubeSuite offers a highly user-friendly development
- Environment featuring significantly shorter build times. The robust line up of expanded functions and user support functions ensures a dependable environment for all users.

A. Application
- Auto Receptionist
- IVRS Telephonic Alerts
- Customer Care Automation
- IVRS Surveys
- IVRS Inventory Control
- IVRS Status Information

B. Advantages & Disadvantages
1) Advantages
- Ease and Accessibility
- Better Customer Service
- Unlimited Customer Access
- Wider Personalization
- Create a Better Company Image
2) Disadvantages
- Menus are too long
- There is too much information. When writing a script for IVR systems, start with the least amount of extraneous information possible.

VI. CONCLUSION

In India 17 language are officially recognized and we need to handle hundreds of dialects and thousands of accents. Clearly a single universal system would not able to service the entire country because of such a system would need to cater to all Indian languages, rendering the system extremely complex. Hence we would like to design a specific application oriented Speech Enabled Interactive System(SEIVRS) which can give more customer (students and parents) satisfaction as if they are interacting with human rather than machine in their own regional language(Telugu).

REFERANCES

[1] Kwan Min Lee, Jennifer Lai, "Speech vs. Touch: A Comparative Study of the Use of Speech and DTMF Keypad for Navigation", International Journal of Human Computer Interaction IJHCI, Vol. 19, No. 3, 2005Ding, W. and Marchionini, G. 1997.

[2] Delogu,C.Di Carlo, A.Rotundi,P.& Sartori,D.(1998), "Acomparison between DTMF and ASR IVR services through objective and subjective evaluation" (FUB Rep. No. 5D01398).Proceedings of IVTTA'98, November 30– December 4,1998,

[3] M. J. "Constraint Satisfaction and Debugging for Interactive User Interfaces", Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington, 1994

[4] Foster,J.C.McInnes,F.R.,Jack, M.A.,Love, S.,Dutton, R. T., Nairn, I. A., et al. (1998), "An ex perimentalevaluation of preference for data entry method in automated telephone services", behaviour & Information Technology,17,1998, pp. 82–92

[5] Goldstein,M.Bretan,I.Sallnas,E.L&Bjork,H, "Navigational abilities in voice-controlled dialogue structures", Behaviour & Information Technology,18,1999, pp. 83–95.

[6] Thotappa, Deepak ; Sarma, B D ; Deka, "Assamese Spoken Query System to Access the Price of Agricultural Commodities", A Published in: IEEECommunications (NCC),National Conference . International Journal of Computer Engineering and Applications, Volume IX, Issue X, Oct. 15 www.ijcea.com ISSN 2321-3469, 2013.

[7] Mohan Dholvan and Dr. Anitha Sheela Kancharla, "Deploying usable speech enabled IVR systems for mass use", Human Computer Interactions (ICHCI), 2013

[8] N. Patel, S. Agarwal, N. Rajput, A. Nanavati, P. Dave, and T. S. Parikh. "A comparative study of speech and dialed input voice interfaces in rural India"In CHI '09:Proceedings of the 27th

Special Issue - 2018

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCESC - 2018 Conference Proceedings

international conference on Human factors in computing systems, New York, NY,USA, 2009, pp. 51–54.

[9] M. Plauch´e, U. Nallasamy, J. Pal, C. Wooters, and D. Ramachandran. "Speech recognition for illiterate access to information and technology", In Proc. IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, May 2006.

[10] A. Sharma, M. Plauch´e, E. Barnard, and C. Kuun. HIV health information access using spoken dialogue systems:Touchtone vs. speech. In Proc. IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, April 2009

[11] J. Sherwani, N. Ali, S. Mirza, A. Fatma, Y. Memon,M. Karim, R. Tongia, and R. Rosenfeld, "Healthline:Speechbased access to health information by low-literate users", In Proc. IEEE/ACM Int'l Conference on Information and Communication Technologies and Development,December 2007

[12] J. Sherwani, S. Palijo, S. Mirza, T. Ahmed, N. Ali, and R. Rosenfeld. "Speech vs. touch-tone: Telephony interfacesfor information access by low literate users". In Proc.IEEE/ACM Int'l Conference on Information and Communication Technologies and Development, April 2009.

[13] Avaaj Otalo - "A Field Study Of An Interactive Voice Forum For Small Farmers In Rural India", IEEE Trans. Acoust. Speech Signal Process, Vol. 37, no. 7, pp. 984_95, Jul. 1989

[14] Steve Young, Gunnar Evermann, Mark Gales, Thomas Hain, Dan Kershaw, Xunying (Andrew) Liu, Gareth Moore, Julian Odell, Dave Ollason, Dan Povey, Valtcho Valtchev Phil Woodland, "The HTK Book (for HTK Version 3.4)", December 2006

[15] Arthur Chan, Evandro Gouvˆea, Rita Singh, Mosur Ravishankar, Ronald Rosenfeld, Yitao Sun, David HugginsDaines, and Mike Seltzer, "(Third Draft) The Hieroglyphs: Building Speech Applications Using CMU Sphinx and Related Resources", March 2007

[16] Kalika Bali, Partha Pratim Talukdar, "Tools for the development of a Hindi Speech Synthesis System", 5 th ISCA Speech Synthesis Workshop, Pittsburgh, 2004, pp.109- 114,