**Accesi Vend: A Closer Look into the Design of Vending Machines for the Handicapped**

Patrick Gandolfo, Joshua Wetzel, Ryan Craft, Joseph Schenck

Advisor: Kuo-pao Yang

Computer Science and Industrial Technology Department

Southeastern Louisiana University

Hammond, LA 70402 USA

**Abstract—**

The accessibility of everyday machines is explored as the typical vending machine is reinvented to make it easier to use for both able-bodied and handicapped individuals. NodeJS, Microsoft Bing Speech, and Azure Cognitive Services are utilized on a Raspberry Pi are utilized to bring this concept to life.

**Accessible Hardware; Speech-to-Text; Facial Recognition**

**I. INTRODUCTION**

When looking at examples of hardware, there are few examples as prolific yet unseen as the everyday vending machine. They are utilized across the world every day and bring in tons of money. While the design of the vending machine is perfect for dispensing snacks and drinks, there are a few improvements that can be made to the accessibility of the machines themselves [10]. The solution utilized here was engineered to make these useful devices even more accessible to a wider audience without sacrificing any of the simple elements that make the machines useful in the first place.

On the hardware side, a Raspberry Pi was utilized. Not a difficult decision to make, as a Pi is modular and has enough power to execute code and power external devices without much hassle. The Pi’s power was not enough for this project so an external power supply had to be added to the project in order for the peripherals to receive the electricity they needed to function fully. Attached to these power-supplied are three servo motor that rotate upon receiving a command to drop the desired product to the user. Attached to the Raspberry Pi itself are a USB microphone and a small camera which are needed to utilize speech-to-text and facial recognition respectively.

On the software side, we decided to go with Node.js not only because it runs on JavaScript but also because the node packages available for it make implementing new functionalities easier on a conceptual level as well as in execution. For the speech-to-text Microsoft’s Bing Speech was utilized, allowing for consistent and functional detection of speech and accurate end results from the audio provided. For the facial recognition from the camera Azure Cognitive Services API provided accurate readings and even allowed for certain faces to be stored and recognized for repeat customers.

While this project does explore the interactions of hardware and software, the idea of accessible technology and making everyday technology more accessible to handicapped individuals is explored. The problems surrounded such implementations are explored on the software level as well as the hardware layer. The implementations of these concepts as well as the reasons for certain implementations are discussed and explore. Lastly the conclusions derived from the project are presented as well as issues that future implementations and research can solve.

**II. RELATED WORK**

With vending machines being such a ubiquitous sight in modern society, it is no surprise that its design has been questioned and reconfigured by many computer scientists. Solutions found by them however, have a much different end goal in mind compared to the scope of this project which was made with accessibility to the handicapped in mind.

The first solution to be looked at is the RFID technology implementation in vending machines. [2] Through implementation of RFID this project aims to eliminate the need for coin machine, as the user will just use a chip with their account on it as payment, making the machine more accessible to those who would have trouble with putting cash in the machine or having the ability to bring cash around with themselves. However, this solution is less good than the one than the solution presented in this paper because with the one presented here the only thing a user has to do is stand in front of the machine and speak to it, making the addition of scanning a card and pressing a button less effective than simply being activated by proximity to a user. The addition of a chip still requires the pressing of buttons which can be problematic for those with lesser motor skills.

One more solution was the idea VendScreen which replaced traditional buttons with a touch screen display, this made accessing what product you wanted as simple as pressing a picture on a screen [9]. This eliminates the need to press a button combination and makes selecting a product much easier and more accessible, however this solution is much more expensive as a touch screen the size of a vending machine would be too costly and would be easy to break and difficult to maintain in the long run.

Another method much closer to the one that is present here is called NuiVend, and much like this one it recognizes user presence and even hears voice commands [1]. Using an Xbox Kinect it detects a user nearby and recognizes their speech and gesturing, and by seeing these it can assess what product the user wants from it. These are certainly goals accomplished by this project as well, however this project is designed with more capable individuals in mind. By requiring gesture interaction as well as voice commands, the vending machine alienates individuals with restricted range of motion just as a much as a typical vending machine does. By reading off of simple face recognition and speech to text the machine can help anyone who can be seen and heard greatly increasing the accessibility of the machine and its practicality for everyday use.

**III. IMPLEMENTATION**

1. **Face Recognition API**

The concept of facial recognition has two respective phases: detection and identification.

**Detection**

During the detection phase, our system will determine if a user is present for use of the system. It does so by using a front-mounted camera that takes an image on an interval, every 3000 milliseconds. Each image is saved to the hard drive and POSTed to the Microsoft Azure Cognitive Services Face API, specifically targeting the Facial Detection endpoint. After computation, we expect to receive a response consisting of a Boolean value (userWasDetected) and some facial data if true. If a user was detected in the image, the system begins phase two, identification.

**Identification**

To identify the user in the photo, we POST the same image to the Face Identification API of Cognitive Services, where it will compare the photo-user against a preexisting database of users and their facial data. If a match is found, the API will return the corresponding user’s ID. This is similar to processes that can recognize a face on smartphones and similar devices [5].

Upon request resolution, the system will compare the returned userID against user’s within our system database to locate a potential customer. The process that recognizes the structure of a face draws distances and can recognize light gradients and other effects on faces [6]. Upon successful recognition, the system creates and stores a new transaction (receipt) for the user and their drink selection. After storing the completed transaction, the system will decrement the remaining quantity for the drink selection for inventory purposes.

1. **Speech API**

The process of speech consists of playing an audio prompt that is prerecorded to the user so they are aware of what can be ordered and that the machine is ready for input. Then the device begins recording audio input from the user, sending recorded audio to the Azure Speech API speech to text (figure B) which requires an authorization token (figure A). The interpreted text is then evaluated by the Azure Speech API Conversation (figure C), which returns a single word intent paired with a confidence of the intent, and this intent is then used in a switch statement to determine what segment of code to activate [4]. The segment of code logs information to our API to record who ordered which drink and what drink they ordered, requests the code that performs the vend action to the proper drink row, then returns back to the main loop of the program and awaits its next initiation. The process can be configured to work with different languages and accents if necessary as it is modular and can be integrated into several different programming languages[7].

**Figure A:**

POST https://api.cognitive.microsoft.com/sts/v1.0/issueToken

content-type: application/x-www-form-urlencoded

Ocp-Apim-Subscription-Key: [Key]

**Figure B:**

POST <https://speech.platform.bing.com/speech/recognition/interactive/cognitiveservices/v1?language=en-US>

Accept: application/json;text/xml

Content-Type: audio/wav; codec="audio/pcm"; samplerate=16000

Authorization: [AccessToken]

requestid: [unique ID]

locale: en-US

**Figure C:**

GET 'https://westus.api.cognitive.microsoft.com/luis/v2.0/apps/[SubscriptionID]?subscription-key=[SubscriptionKey]&verbose=true&timezoneOffset=0&q=[string]';

**C. Motors and Pins**

It was determined early on that the vending machine would be moved by motors that would be activated by the Raspberry Pi. However early on in our test runs it was discovered that the Pi did not have the sufficient ability to run the motors at the rate we desired which resulted in a Pi getting burnt out, rendering it unusable. Soon after that this an external power source that connected to the i2cbus slots on the Pi was procured, giving enough power that we were free to continue the project as planned.

In order for the motors to be moved, first a pulse width modulation had to be set. A pulse width is the frequency in which the device sends power to the servo motors attached to it [3]. With this in mind the pulse width was set to go to a maximum and then soon after return to its minimum. This allowed the drink time to fall into the PVC tubing and be dispensed without much risk of getting trapped.

Once the speech to text receives a valid response the pins receive a request to fire at whichever pin the soda requested corresponds to on the external power source. A predetermined range of motion is given to the pins beforehand, so once this command reaches them they know exactly where to move, where to move to, and when to move back to original location.

**D. Hardware**

The hardware chassis was decided by our team to be prototyped in wood, PVC, black metal grating, and plastic. The system is configured as three separate vertical chutes aligned over a plastic ramp. Each can holster was cut to 4 ½ inches from 2x6 wood to fit the cans with a ¼ inch of space respectively. When the user turns the vending machine on via the facial recognition center, the PVC canister cylinder rotates, releasing a single can into the container. This in turn is rotated 180 degrees causing the can to be dropped onto the plastic ramp, completing the process.

The creation of the cylinders out of PVC was created using saws and angle grinders to cut the proper slit for the can to sink into the container. The ramp was sliced from black roofing ventilation ducts. The metal grating that displays the front of the drinks to the user was cut to match the can holster with a wire cutter. Wiring was aligned and spaced behind the machine while the facial recognition sensor and speaker are glued to the side of the chassis as a temporary solution.

The process behind building the structure vertically and using the chute system was to simplify the mechanics needed to select soda, making it vastly simpler than modern vending machines. This is because the project is essentially focus testing the usage of voice and facial recognition sensors on vending machines, not reinventing the wheel on specific mechanics for dispensing the soda.

**IV. EVALUATION**

To test the solution different people in the group would interact with the machine and ask for drinks. Troubleshooting and bug detection was as simple as asking for an incorrect drink or simple saying nothing and seeing if the hardware would react in a desired way. With the only two user inputs being the detection of their face and what words they say to the machine, there were few points of failure that could destabilize the machine with errors.

Compared to the other similar solutions to problems like this one, the project here performed its intended purpose much to a greater extent. While the RFID chip idea functions similarly, its inability to eliminate the step of pressing buttons to order renders the machine less effective in the long run with the goals presented here in mind. And while NuiVend is even closer to the solution presented here because it requires gesturing and specific movements to dispense drinks it is simply less effective for people who are limited in the movements that they can produce. VendScreen has a good concept but something could easily break the screen and render it unable to function for anyone, not to mention the cost of repairing the vending machine sized touchscreen.

**V. CONCLUSION AND FUTURE WORK**

In this project, the problems of making everyday objects accessible to those who are unable to use them was explored as a way to implement technologies such as speech to text and face recognition. For the programming side NodeJS was used as the base software alongside Microsoft Bing Speech and Azure Cognitive Services. For hardware implementation, a Raspberry Pi was utilized alongside an external power source and servo motors. The implementation of these allowed for a complex and functional vending machine.

By looking at the way hardware is designed it is easy to forget that not every machine accessible by an average person can be used by someone who is disabled. Making software and hardware usable by everyone is a core necessity to the philosophies behind computer science design, and it is important to keep this in mind when creating usable machines. The major goal of this project was to prove that even with complex machinery there are design standards that must be met. However, while the solution here does exist there are still many different machines that could be improved in design to make accessibility greater and life easier for a great amount of the population. This voice recognition technology as well as speech could even be used to make everyday activities such as online learning, speech therapy, and learning new languages more accessible and affordable for a wider audience than ever before [8].

Further research into this area would require different every day devices to be analyzed with input from the handicapped to determine what areas of their life could be improved by user friendly and handicapped accessible technologies. Refrigeration is definitely a factor that could be added in in the future to improve the quality of the products being dispensed as well as a better casing for the vending machine to make it seem more inviting.

**REFERENCES:**

[1] R. Gruen and E. Liang, “NuiVend - Next Generation Vending Machine”, Computational Science and Computational Intelligence (CSCI), 2016 International Conference, 2016

[2] A. Ramzan, S. Rehman, and A. Perwaiz, “RFID technology: Beyond cash-based methods in vending machine”, Control and Robotics Engineering (ICCRE), 2017 2nd International Conference, 2017.

[3] H. Priya Vemaganti, D. Sreenivasarao, G. Siva Kumar, “Improved pulse-width modulation scheme for T-type multilevel inverter”, IET Power Electronics ( Volume: 10, Issue: 8, 6 30 2017 ), 2017.

[4] C. Desai and V. Colaco, “Development of a personalized integrated voice recognition and synthesis system”, ICWET '10 Proceedings of the International Conference and Workshop on Emerging Trends in Technology, 2010.

[5] C. Chu and S. Peng, “Implementation of Face Recognition for Screen Unlocking on Mobile Device”, MM '15 Proceedings of the 23rd ACM international conference on Multimedia, 2015.

[6] J. Kulandai Josephine Julina and T. Sree Sharmila, “Facial recognition using histogram of gradients and support vector machines”, Computer, Communication and Signal Processing (ICCCSP), 2017 International Conference, 2017.

[7] P. Iswarya and V. Radha, “Speech and text query based Tamil - English Cross Language Information Retrieval system” Computer Communication and Informatics (ICCCI), 2014 International Conference, 2014.

[8] R. Shadiev, Y. Huang, W. Hwang, and N. Shadiev, “Investigating the Effectiveness of Speech-to-Text Recognition Application on Learning Performance in Traditional Learning Environment”, Advanced Learning Technologies (ICALT), 2015 IEEE 15th International Conference, 2015

[9] McGlaun Shane, "VendScreen Gives Boring Plain Vending Machines Touchscreen Coolness", SlashGear. SlashGear 22 Sept. 2011. Web. 10 Oct. 2016.

[10] G C. Wenshan, H. Yanqun, L. Minyang, "Influential Factors of Vending Machine Interface to Enhance the Interaction Performance", 2015 8th International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 486-489, 2015.