Voice Recognition Software with Emotion Discerning Capabilities

Mohammed MoawadAlenazi

Tabuk University
Computer Engineering

Abstract- The purpose of this project is to develop a voice recognition system that distinguishes different emotions. Innovative headways have made it conceivable to make various software applications for speech recognition. However, currently available software – commercial as well as research-based voice recognition - is not advanced enough to correctly discern and categorize the emotional status of a speaker (Wheeler et al., 2017). The aim of this project is to create software using MATLAB that will not only identify and categorize the emotions of a speaker but also provide appropriate feedback by the determined moods/emotions.

Index Terms: Matlab program, voice recognition system, software, samples of voices.

I. Introduction

Software solution is ideal to this project because it is possible to program almost anything if the process is well understood and is also accurately defined. In this project, the programmers will define various emotions and moods using well-recognized conventions. For instance, happy people whisper and laugh frequently. On the contrary, disturbed individuals talk angrily; they raise their voices and may swear several times. The voice recognition software will be programmed to monitor voice intonation, identify voice stress patterns, and determine the nature of the words uttered to capture the moods or emotional status of the speaker. The aim is generating content that is mirrors their emotional state.

For a start, the application will be programmed to distinguish between two common emotions: happiness or joy, and sadness or agitation. When a person is happy, the program will output relevant content that will encourage the individual to share his or her happiness. It will also encourage them to continue talking. When the software recognizes that a person is sad, it will provide comforting words to help the person deal with his or her ordeal. By focusing on two emotions at first, it will be possible to gain significant knowledge and skills regarding how to improve the application in the future. It is also easier to design a simpler version of a pilot program.

II. Description of How the Engineering Design Process is Incorporated into the Project

The success of the program will depend on the utilization of efficient procedures. In this regard, the engineering design process will be strictly followed to ensure thoroughness and clean implementation. The methods of engineering design are the establishment of objectives and criteria, synthesis, analysis, construction, testing, and evaluation (Fuggetta& Di Nitto, 2014).

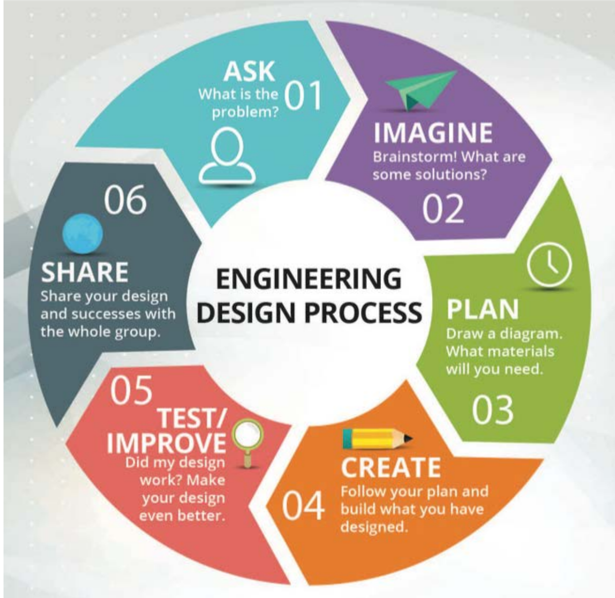

Fig 1: Engineering Design Process As shown in Fig 1, the fundamental elements of engineering design process are

establishment of objectives and criteria (asking what the problem is), synthesis (imagining or brainstorming some solutions), analysis (planning), construction (construction by following the plan), testing (to see if the design works), and evaluation (which includes sharing the success of the project with others).

The primary objective of this project is to create voice recognition software that can discern the emotions of a speaker. After the system successfully classifies the type of emotions the user exhibited, it can authentically engage the user by accessing the internet, and – through speech technology – utter relevant stories whose similarities would either enforce positive emotions or mitigate negative ones. The voice recognition system will possess some characteristics of an intelligent system. This socially aware system is capable of forging some form of semi-natural emotional bonds with the people. The system will most likely find numerous applications in robotics and other areas.

As far as the criteria are concerned, voice-based characteristic such as pitch and frequency will be used to define the mechanism for judging the emotions based on how words are uttered. Visual cognition is another tool that might be employed in MTALAB. With the computer vision tool box that is available in MTALAB, capturing facial features such as the mouth or eyes orientation is a very effective way of identifying emotions. But essentially, judgments of the emotions of an individual will be based on intonation, and the specific words said. It is expected and assumed that people interacting with the software will follow globally recognized – cross cultures - conventions most of the time. For instance, it is anticipated that a person filled with happiness will talk pleasantly and use kind words. They will also laugh often and be more attentive. On the other hand, a frustrated individual will be expected to swear frequently, or at least use negative words. The software will pick these variations in voice and intonation to determine the moods of a person accurately. The program will also contain learning capabilities so that it can increase its effectiveness over time.

Synthesis is the next step in the engineering design process. It is about looking at the idea and putting it all together to create a product that meets the desired goals. The synthesis may involve an analysis of the feasibility of the product and its appropriateness (Tayal, 2013).

In this project, integration will include considering the viability of adding new feature to current project, and how it can be used shortly to solve global problems. It is always important to ensure that projects are not only doable but also economically viable. Previously, it has been noted that a piece of software that can accurately discern the moods of a person can be integrated into socially aware robots to improve how they relate to humans.

The analysis is the detailed examination of the elements or structure of the new product with the aim of understanding its inner workings and making any necessary improvements. For this project, the analysis will focus on the accuracy of the definition of emotions. Errors in definition will lead to inaccurate interpretations of emotions and moods. Therefore, it is important to be as accurate as possible and to ensure that the working definitions for the program are useful and reliable. In this project. Additionally, psychologists and language experts will be consulted to offer their inputs. Using the services of cultural and language professionals will increase the effectiveness of the definitions provided, and reduce the likelihood of error.

Construction is the actual method of building the system. In this project, construction will be done using Matrix Laboratory (MATLAB) programming language. Carefulness should be highly maintained during the construction phase to avoid critical errors that can render the system functionless and in inappropriate (Moore, 2014). The use of an appropriate and resilient system can reduce the level of mistakes during construction. MATLAB is the preferred construction method because it is the language of technical computing. Voice recognition is a specialized function that requires a flexible programming language that is optimized for solving technical problems. MATLAB is considered the most dominant and preferred language for expressing computational mathematics in a natural way (Moore, 2014). Other reasons for choosing MATLAB include the fact that it allows engineers to run their analyses on big sets of computational data and easy integration with other programming languages. Finally, MTALAB has powerful built-in tools that save hours of development time, especially during the prototyping phase.

Once construction is complete, testing must be carried out to see whether the new system achieves the previously identified goals. The testing primary function is to determine the resilience of the scheme, and what developers may need to modify to improve the performance. Testing for this voice recognition software will try to determine whether the program can identify emotions and respond accordingly or not. The testing phase is also critical because it is used to determine whether any errors were introduced during coding. The testing team can then immediately fix such errors before the system is deployed.

A system that has been constructed, tested, and found to meet all requirements as outlined in the project proposal must go through an evaluation to allow the builder to make decisions about the nature and appropriateness of the final output. After evaluation, the builders can either embrace or enhance the project through another cycle. For the speech recognition software, the evaluation will determine its effectiveness and appropriateness.

IV. V-chart for System Design and Verification

The V-chart - or model as used in software engineering - is comparable in some respect to the waterfall model (Chandra, 2016). The V-chart shows the software development process and how different phases are related to each other (Mc Hugh et al., 2013). Fig. 2 represents the V-chart model for the current project.

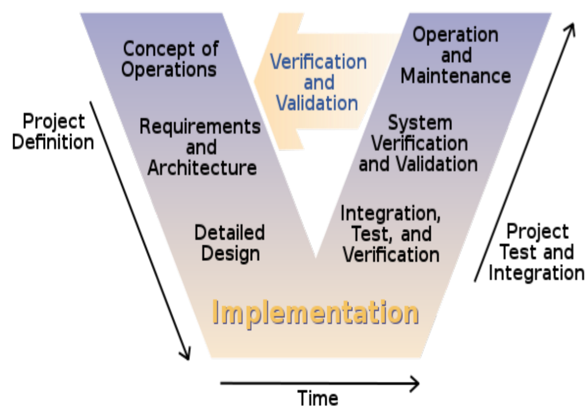

Fig 2: V-Chart Model

V. Summary of How Realistic Design Constraints Are Being Incorporated into the Project

The current project is enormous, and there is considerable hope that it will contribute positively towards the advancement of knowledge especially in speech recognition technologies. However, the economic, environmental, ethical, health, safety, social, political, sustainability, and manufacturability problems also plague the project. The economic challenges are being addressed by seeking sponsorship and reducing wasteful and expensive processes. Environmental issues are being addressed through multiple approaches, including efficient utilization of resources. Proper steps are also being taken to ensure there are no ethical concerns like an infringement of intellectual property or lack of compensation for those who are part of this great process. Health and safety issues are managed by using appropriate inputs and ensuring that outputs are not hazardous.

VI. Description of the Deliverables of the Project and Their Final Status

The final deliverable for the project is a voice recognition software application that can accurately categorize the mood or emotional state of an individual.

1) Product Development

Product development will commence immediately after all the necessary details have been confirmed. It is important to have all details in order before beginning the actual development to ensure a smooth process. A team of software engineers, who are conversant with MATLAB to ensure the achievement of the best outcomes, will develop the product. Effective development of the product will depend on exploratory conceptualization, plans and requirements, requirement baseline, product design, and detailed design. The exploratory conceptualization is to develop a prototype of the voice recognition software. The requirements will include the ability to program and define the key variables of the software. The software will recognize voices and classify them according to moods.

2) Testing

Testing of the product will be performed after every significant milestone of the project. Regular testing ensures that errors are detected and removed early. Testing will look at the effectiveness of the product and determine what needs to be done to align the final product with the aims and objectives of the project. Tests for the product will include unit tests, integration and system test, acceptance test, validation and verification, and installation. Unit tests will check every piece of software produced to ensure it meets the project requirement. Integration and system tests will ensure that the software is carefully and appropriately put in the electronic devices. Thereafter, the user acceptance test will aim to ensure aspect is executed according to

plan. Lastly, the user acceptance tests validate the new system and verify it before installation. During the installation process, the engineers will maintain all records of both proper software performance and of any failure encountered. The revision of the system to compensate for errors detected during these tests will follow the similar procedures and controls as for any other software change.

VII. Other Considerations

The speech recognition software will be devised in the English language. In other words, the software will only recognize words that are spoken in English. The program will also capture the feelings or moods of the speaker – through voice and visual cognition - and produce consistent output to either encourage the speaker to keep sharing information or comfort the speaker. This new development will become an important component especially in applications such as the socially conscious robots.

1) Economic

Economic considerations are important in every project. In the current one, measures will be taken to ensure that there are some economic advantages to it. For instance, it should be easy to manage and use. The product should also be affordable, and the benefits of the project should surpass the cost of development. It should also be easy to distribute and utilize the new software after it is completed. The current plan is practical and makes economic sense because it can be used in robotics and the manufacture of automobiles. Additionally, machine learning capabilities should be incorporated into the project. This will lay the foundation to further develop the product to be utilized many other diverse walks of life such as mental institutions.

2) Environmental

The project will bring about real relationship and interactions between human beings and machines. Since creating this voice recognition software involves coding, the project does not pose any direct threat to the environment. In fact, the software will possibly contribute towards preserving and improving the environment; an enhanced version of the software – through machine learning – could detect through their voice the likely hood of involvement in an altercation. By so doing, the software will calm the person and reduce the likelihood of the person destroying the environment due to anger. Or, it could alternatively call a relative to be present in the scene.

3) Ethical

The project is also acceptable when ethical considerations are made. Engineers will create the application with the aim of giving people some socially significant feedback. The voice recognition software does not “indulge” in any unethical processes. Instead, it encourages individuals to be more effective and more in-tune with the environment. The engineers will also uphold honesty in their work. Consequently, the final product will be caring, fair, and respectful just as programmed. Engineers will also respect the rights of others in every phase from programming to implementing the software.

4) Health and Safety

The product will be created such that it is appropriate and does not have any health or safety issues. The use and disposal of the product will not cause any health problem because it is essentially code. The hardware components required to implement the program will also meet the requirements created by the relevant health and safety protection agencies.

5) Manufacturability

To mass produce the software efficiently and economically, a legacy organization will be given the source code and told to implement it in a variety of hardware components. In the initial

stage, a company will be required to create software that is integrated into a simple hand-held gadget that people can talk to. The gadget will have speech recognition and speech output capabilities.

6) Political

The project will be created to follow existing regulations to avoid issues that may arise from failure to abide by existing laws. The engineers will look at all relevant legislation and regulations influencing software development and comply with every single requirement.

7) Social

The current software will be used for social purposes. In fact, the speech recognition software will be helpful in guiding individuals on what they need to do depending on their moods and emotions. Social concerns are always related to a broad grouping of people with common traditions, institutions, or corporate activities and interests, and it is always important to respect their values. On a different note, the product can be further developed to promote cross- cultural awareness. Since it will incorporate machine learning algorithms that will increase the product cultural awareness as it is being used.

8) Sustainability

The project is also sustainable as it does not utilize any natural resource once completed. It is a simple voice recognition software that will identify a person's emotions immediately when the user talks. This project is sustainable in a way that it does not deplete or permanently damage the use of a resource. Additionally, the product can be utilized in customer service settings. It can give feedback to employees regarding how they fared during interactions with customers. And fee collected from clients can be used to further sustain and develop the product.

VIII. Conclusion

Field recognition systems exist in the market today but none of them can be said to be effective in the discernment of the moods of the person talking with the aim of providing them with the needed advice. The current project will create voice recognition software with the capability of understanding the moods of the person with the aim of giving appropriate advice. Such voice recognition systems can be used to improve the human-machine interactions in the future and form a strong foundation for other improvements especially with regards to socially aware systems.

References

Chandra, V. (2015).Comparison between various software development

methodologies. *International Journal of Computer*

Applications, *131*(9), 7-10. Fuggetta, A., & Di Nitto, E. (2014, May).Software process.In*Proceedings of the on Future of*

Software Engineering (pp. 1-12).ACM. Mc Hugh, M., Cawley, O., McCaffcry, F., Richardson, I., & Wang, X. (2013, May).An agile v-

model for medical device software development to overcome the challenges with plan- driven software development lifecycles. In *Software Engineering in Health Care (SEHC), 2013 5th International Workshop on* (pp. 12-19). IEEE.

Moore, H. (2014). *MATLAB for engineers*. New York, NY: Prentice Hall Press. Tayal, S. P. (2013).Engineering design process. *International Journal of Computer Science and*

Communication Engineering, 1-5. Wheeler, J., Richardson, B., Amman, S., Ji, A., Huber, J., &Rangarajan, R. (2017). Systems

engineering approach for voice recognition in the car. *SAE International Journal of*

Passenger Cars-Electronic and Electrical Systems, *10*(2017-01-1599), 181-185. https://www.freelancer.in/contest/Engineering-Design-Process-Poster-350020-byentry-8093022.html