**INTRODUCTION**

Speech recognition technology is one from the fast-growing engineering technologies. It has a number of applications in different areas and provides potential benefits. Nearly 20% people of the world are suffering from various disabilities; many of them are blind or unable to use their hands effectively.

The speech recognition systems in those particular cases provide a significant help to them, so that they can share information with people by operating computer through voice input. This project is designed and developed keeping that factor into mind, and a little effort is made to achieve this aim.

Our project is capable to recognize the speech and convert the input audio into text; it also enables a user to perform operations such as “save, open, exit” a file by providing voice input. It also helps the user to open different system software such as opening Ms-paint, notepad and calculator. At the initial level effort is made to provide help for basic operations as discussed above, but the software can further be updated and enhanced in order to cover more operations.

**BACKGROUND OF THE STUDY** **INTRODUCTION**

The development for speech recognition system has been for a while. The recognition platform can be divided into three types. Dynamic Time Warping (DTW) (SAKOE, 1978), the earliest platform, uses the variation in frame's time for adjustment and further recognition. Later, Artificial Neural Network (ANN) replaced DTW. Finally, Hidden Markov Model was developed to adopt statistics for improved recognition performance.

Besides the recognition platform, the process of speech recognition also includes: recording of voice signal, point detect, pre-emphasis, speech feature capture, etc. The final step is to transfer the input sampling feature to recognition platform for matching.

In recent years, study on Genetic Algorithm can be found in many research papers (Chu, 2003a; Chen, 2003; Chu, 2003b). They demonstrated different characteristics in Genetic Algorithm than others.

For example, parallel search based on random multi-points, instead of a single point, was adopted to avoid being limited to local optimum. In the operation of Genetic Algorithm, it only needs to establish the objective function without auxiliary operations, such as differential operation. Therefore, it can be used for the objective functions for all types of problems.

Because artificial neural network has better speech recognition speed and less calculation load than others, it is suitable for chips with lower computing capability. Therefore, artificial neural network was adopted in this study as speech recognition platform. Most artificial neural networks for speech recognition are back-propagation neural networks. The local optimum problem (Yeh, 1993) with Steepest Descent Method makes it fail to reach the highest recognition rate. In this study, Genetic Algorithm was used to improve the drawback.

Consequently, the mission of this chapter is the experiment of speech recognition under the recognition structure of Artificial Neural Network (ANN) which is trained by the Genetic Algorithm (GA). This chapter adopted Artificial Neural Network (ANN) to recognize Mandarin digit speech. Genetic algorithm (GA) was used to complement Steepest Descent Method (SDM) and make a global search of optimal weight in neural network. Thus, the performance of speech recognition was improved. The nonspecific speaker speech recognition was the target of this chapter. The experiment in this chapter would show that the GA can achieve near the global optimum search and a higher recognition rate would be obtained. Moreover, two method of the computation of the characteristic value were compared for the speech recognition.

However, the drawback of GA used to train the ANN is that it will waste many training time. This is because that the numbers of input layer and output layer is very large when the ANN is used in recognizing speech. Hence, the parameters in the ANN is enormously increasing. Consequently, the training rate of the ANN becomes very slow. It is then necessary that other improved methods must be investigated in the future research.

The rest of this chapter is organized as follows. In section 2, the speech pre-processing is introduced. Then, in section 3 we investigate the speech recognition by ANN which is trained by genetic algorithm to attain global optimal weights.

The use of information and communications technology is becoming increasingly embedded in and interconnected with our everyday living. From basic operations such as creating and editing documents, working with data in spreadsheets and databases, or creating PowerPoint presentations to advanced skills required for specialized work, individuals and organizations continue to rely on more computer software than they have ever done before.

This reliance has not only continually transformed lives but highlights the importance of digital literacy as a skill. Today, digital skills are becoming a near-universal requirement. In the United Kingdom, the government is exponentially increasing its investments in digital skills to avoid being left behind and position its citizens for global recognition and competition. Several organizations are supporting this endeavor by listing basic digital skills as an employment requirement. An analysis of 9.4 million UK job postings reveals that 77% of positions required basic digital skills.

While research indicates that the number of people with fundamental digital skills has increased in recent years, concerns remain about one marginalized group –– people with disabilities. An ONS (Office of National Statistics) survey in 2020 found that 81.4% of disabled respondents were recent users of the internet compared to 95.5% of non-disabled respondents.

The key factor playing a constraining role is the unavailability of content in accessible format. Consequently, all hands are on deck towards building inclusive digital communities. This global commitment is evident in the initiatives and solutions developed in recent years to address some of these specific challenges faced by people with disabilities.

One of such initiatives is to incorporate a speech recognition component into a communication app for disabled users.

Speech recognition, also known as speech-to-text, is a notable interdisciplinary subfield of computer science and computational linguistics.

Through methodologies and technologies that allow computers to recognize and translate spoken language into text, users primarily benefit in the scope of productivity, mobility, and essentially inclusion. It provides users with varying application options such as voice dialing (e.g., “Call Mary”), call routing (e.g., “I would like to make a call”), domotic appliance control, search key words, (e.g., find a podcast where certain words were spoken), simple data entry (e.g., entering a number), preparation of structured documents, speech-to-text processing (e.g., emails and word processors), and aircraft ( often termed as direct voice input).

Speech recognition has a long history,

**1.1   Project Objective**

•      To understand the speech recognition and its fundamentals.

•      Its working and applications in different areas

•      Its implementation as a desktop Application

•      Development for software that can mainly be used for:

•      Speech Recognition

•      Speech Generation

•      Text Editing

•      Tool for operating Machine through voice.