Multimed Tools Appl
DOI 10.1007/s11042-011-0918-7

Automatic speech recognition- an approach for designing inclusive games

Moyen Mohammad Mustaquim

Abstract Computer games are now a part of our modern culture. However, certain categories of people are excluded from this form of entertainment and social interaction because they are unable to use the interface of the games. The reason for this can be deficits in motor control, vision or hearing. By using automatic speech recognition systems (ASR), voice driven commands can be used to control the game, which can thus open up the possibility for people with motor system difficulty to be included in game communities. This paper aims at find a standard way of using voice commands in games which uses a speech recognition system in the backend, and that can be universally applied for designing inclusive games. Present speech recognition systems however, do not support emotions, attitudes, tones etc. This is a drawback because such expressions can be vital for gaming. Taking multiple types of existing genres of games into account and analyzing their voice command requirements, a general ASRS module is proposed which can work as a common platform for designing inclusive games. A fuzzy logic controller proposed then is to enhance the system. The standard voice driven module can be based on algorithm or fuzzy controller which can be used to design software plug-ins or can be included in microchip. It then can be integrated with the game engines; creating the possibility of voice driven universal access for controlling games.

Keywords Design for all · Inclusive design · Universal design · Accessibility · Game

1 Introduction

1.1 Computer games in society today

Games are the vehicle with which society is and will change the computer. But how will the games themselves are changed by society? We can expect two processes to affect games:

M. M. Mustaquim (✉)
Department of Informatics and Media, Uppsala University, Ekonomikum (plan 3),
Kyrkogårdsg 10 Box 513, 751 20 Uppsala, Sweden
e-mail: moyen.mustaquim@im.uu.se

Published online: 03 November 2011

the mass market and the flowering of heterogeneity. In some ways, these processes work against each other [3].

As participation in gaming increases around the world and across a larger part of the population, it is obvious that games are not only an entertainment media for children or young people anymore but games can be used to serve several purposes for all kinds of people in the society. Today's technologically inclined people are going to be older tomorrow, and they would not like to be excluded from future services for instance games, just because of design issues associated with the game. Also, anyone can suffer temporary or permanent disabilities regardless of age and it would be unfortunate to exclude them from using games. Moreover, both the human death rates and birth rates are generally decreasing, which results in an age distribution with a large number of older people (disabled or not) in the society in the near future. So the game industries should be concerned with such issues, so that they can include these groups of people and offer them properly designed services.

Automatic speech recognition systems are not a new concept in modern computing. However the possibility of using automatic speech recognition systems to control games is one option to offer inclusiveness in game design. People with limited motor control ability can benefit from the opportunity of using their voice to drive the commands of the game. This paper proposes such speech recognition system for game which is one of the objective in achieving universal design in computer gaming.

1.2 Why inclusive game in needed

The philosophy underlying inclusive design specifically extends the definition of users to include people who are excluded by rapidly changing technology, especially the elderly and ageing, and prioritizes the role and value of extreme user groups in innovation and new product and service development. It also prioritizes the context of use, both physical and psychological, and the complexity of interactions between products, services and interfaces in specific contexts of use, such as independent living [14].

Two major demographic trends underscore the importance of considering technology adoption by the older adults: the aging of the population and rapid dissemination of technology within most societal contexts. In the past decade, developments in computer and information technologies have occurred at an unprecedented rate, and technology has become an integral component of work, education, healthcare, communication and entertainment. Game has twisted a different viewpoint in past decade in a way that, game is not only an amusement entity but it has other possibilities to serve too. While we are witnessing explosive developments in technology, the population is aging. In 2003 people aged 65+ years, in the United States, numbered about 35 million and represented approximately 13% of the population [4]. By 2030, this number is expected to increase to about 71 million, representing 20% of the population (Fig. 1). Moreover, there will be a dramatic increase in those aged 85+ years, from about 4 million in 2000 to nearly 21 million by 2050 [4].Similar trends are occurring worldwide. By 2030, the percentage of people aged 65+ in Europe will be about 24%, and about 12% in Asia and Latin America [4].

Similar trends exist in the European Union (EU). In the EU, 48% of the households had access to the Internet in 2005, and 23% had a broadband connection. However, a gap still remains between users and non-users according to age. Specifically, the proportion of computer or Internet users among those 16–24 years of age is three times higher than among persons 55–74 [6]. Recent data also indicates that older adults use other forms of technology such as ATMs and DVDs less than younger adults [4].

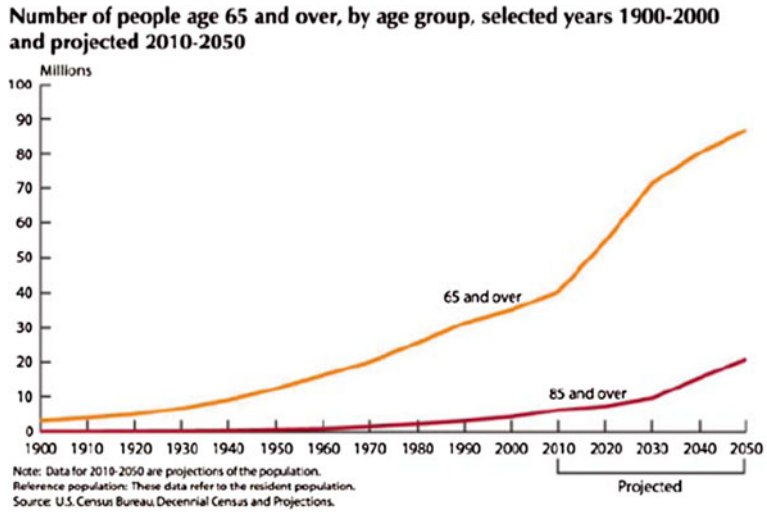

Fig. 1 Projected growth of people 65+ in The United States. Data source: federal interagency forum on aging-related statistics, 2005

To make technology useful to and usable by older adults, a challenge for the research and design community is to "know thy user" and better understand the needs, preferences and abilities of older people [4]. It is fairly well established that many technology products and systems are not easily accessible to older people. There are of course a myriad of reasons for this, such as cost, lack of access to training programs, etc. However, to a large extent lack of accessibility is due to the fact that designers are unaware of the needs of users with varying abilities, or do not know how to accommodate their needs in the design process [2]. For example, findings from an observational study of designers involved in a design competition for older people [7] found that the designers tended to restrict their use of user information and user involvement in the design process [2].

As described in the previous section it is apparent that, as game is becoming more popular in a different standpoint and the user of technology becoming elder, it is now time to think about what will happen in the future while the users will be technologically inclined but hemmed in to be restricted from playing game just because the game is not properly designed for him or her? Or regardless of this issue, any age group people can suffer from certain physical disability which makes them excluded from playing games-again because of the design issue. Hence the importance of designing inclusive games is going to be more and more highly valued in the future. If people with physical limitations or people with disabilities due to age can access game; because it is designed according to their need, universal access of game design can be achieved. This paper takes an approach of designing inclusive games for achieving universal design of gaming and for giving accessibility and interaction, use of speech recognition system is the topic of research here.

1.3 Future game design

As society ages and the care gap increases, technology is envisaged as a means of allowing older people to live independently for longer. Age on its own is not a disability, but older people are more likely to experience disabilities of various kinds. These include both single severe impairments and combinations of multiple minor

sensory, cognitive and physical impairments, which can combine to make products and services much harder to use [8, 9].

In the upcoming future where computer games are a major recreational activity as well as a tool for attaining tasks, a mass market of computer games not too different from what we now have should focus on 'properly designed' games, where proper design unlocks the possibility, to make a game complete and self sufficient.

1.4 Significance and nature of the interaction in games

Interaction in gaming is important for several reasons. First, it injects a social or interpersonal element into the event. It transforms the challenge of the game from a technical one to an interpersonal one. A puzzle will always present the player with exactly the same challenge. But a game opponent reacts to the player's actions, and presents different challenges in each game. This difference has major emotional significance [3].

What is important about the modes of interaction in gaming is not their mechanical quality but their emotional significance. Thus, the degree of interaction provides a useful index of gaminess" [3, 11]. It is then important to use the right interaction technique in right level during the design of voice interaction technique for controlling game.

The individual user may have excellent ability in some areas and yet be poor in others. For the population as a whole, there can be a wide variability in any one attribute. The complexity of the problem increases dramatically as more attributes are considered. In general, attributes deteriorate with ageing, whereas the variability increases [15].

1.5 Using human modality in inclusive design

The advent of the Information Society, along with the emergence of novel technological paradigms such as mobile or ubiquitous computing, ambient intelligence, and the "disappearing computer", raises new and stimulating scientific challenges in research on multimodal human computer interaction (HCI) [1]. New input/output modalities and forms of multimodality are hence needed to provide easy and effective access for control to all users, including the very young and the elderly, as well as people with various types of disabilities; in all contexts of use, especially in mobile interaction with embedded systems, wearable computers, and augmented everyday objects integrating aspects of intelligent interactive behavior [1]. It is thus paramount to use human modalities for controlling game.

2 Background

2.1 Game controller

A game controller is a device used for controlling the playable character or objects or otherwise providing input in a computer game. The controller is usually connected to a game console or a computer by means of a wire, chord and nowadays also by means of wireless communication. Typical game controllers are keyboards, mice, joysticks, game pads etc. There are also special purpose devices such as steering wheels or foot pedals. The principal function of a game controller is to control the movement or actions of a playable body-object or otherwise influence the events in a video- or computer game. There is health concerns related to game controllers such as risk for injuries like repetitive strain injuries or carpal tunnel syndrome hence they need to be designed ergonomically to give best possible

relaxation of hands and mind. However, certain groups of people have limited accessibility or no accessibility at all to such game controllers because of their physical limitations. Alternative ideas for game control are therefore important to consider for such groups of people.

2.2 Speech technology

Speech technology seems to provide new opportunities to improve the accessibility of electronic services and software applications including game, by offering compensation for the limitations of specific user group. These limitations can be quite diverse and originate from specific sensory, physical or cognitive disabilities—such as difficulties to see icons, to control a mouse or keyboard [15]. Such limitations have both functional and emotional aspects that should be addressed in the design of user interfaces [16]. Speech technology can be an 'enabler' for understanding both the content and 'tone' in user expressions, and for producing the right information with the right tone [15].

2.3 Automatic speech recognition system

Automatic speech recognition (ASR) is the automatic conversion of human speech to a sequence of words. The aim of ASR is to recognize automatically what has been said. Nowadays, ASR systems are generally based on the hidden Markov models (HMM) for modeling the acoustics of speech and use either statistic language models (n-grams) or rule-based grammars to model the language component. Furthermore, there are many techniques that normalize the acoustic features of the signal and adjust the acoustic models to a specific speaker or different recording conditions [15].

If a speech recognizer is part of an application, usually it is used as an input technology or technique [10]. Various other performance measures can augment the word error rate, depending on the application. For instance, van Leeuwen et al. [17] mentions the following issues that are important to consider when estimating the performance of an ASR application: the type and quality of feedback, error correction strategy, dealing with out-of-domain words, speed and response time, the user's situational awareness in the dialog structure, dialog/task success rate, and subjective impression of the overall performance of the system [18].

2.4 Fuzzy logic

Fuzzy logic is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PCs or workstation-based data acquisition and control systems. They can be implemented in hardware, software, or a combination of both [19]. Fuzzy logic provides a simple way to arrive at a definite conclusion based upon vague, ambiguous, imprecise, noisy, or missing input information and fuzzy logic's approach to control problems mimics how a person would make decisions, only much faster. Fuzzy logic was first conceived as a better method for sorting and handling data but has later proven to be an excellent choice for many control system applications since it mimics the logics of human control. It can be built into anything from small, hand-held products to large computerized process control systems. It uses an imprecise but very descriptive language to deal with input data more like a human operator. It is very robust and forgiving of non-perfect operator performance and data input, and often works when first implemented with only little or no tuning.

2.5 Fuzzy logic controllers

Fuzzy logic control may be viewed as a branch of intelligent control. It can be seen, as an emulator of the human decision making process in that it is approximate rather than exact. The conventional control systems work either in a linear or a non-linear manner. Linear systems are simple and well explored but not too flexible. The non-linear systems on the other hand can be powerful but are often sensitive to modeling errors. A fuzzy controller is effective in situations where the control process is too complex to analyze by conventional quantitative techniques and also when the available sources of information are interpreted qualitatively, inexactly or uncertainly. A fuzzy logic controller has several advantages over a traditional controller. It is more flexible, and easy to understand. A fuzzy controller uses human linguistic terms for control. To design a fuzzy controller for parallel or distributed control, multiple fuzzy rules, or complex non-linear systems can be used. However, there are more parameters to tune in the fuzzy controller and the stability of the fuzzy controller is sometimes difficult to analyze mathematically. Regardless of its complexity a fuzzy logic controller can be used where it is difficult or impossible to model mathematically and where traditional strategies are too complex or non-linear to be controlled. Figure 2 shows the basic structure of a fuzzy logic controller.

A typical design of a fuzzy controller involves some sequential process starting from determination of state and control variables. Then an inference method and a fuzzification method are selected. The next step is to normalize the state variable space while the shapes of the fuzzy sets are determined too. Once this is done the fuzzy rule base is constructed and a defuzzification strategy is selected to convert the fuzzy sets of data to the crisp data set. Of course there are further steps involved like testing, fine tuning of the controller and construction of a lookup table so that the lookup table can be used in order to save the inference and defuzzification time. The proposed fuzzy controller described in the proposed system section is based on the above description.

2.6 Requirements of voice controlled system for games

While the traditional controlling approach of game is using input devices, people with certain limitation of doing so would like to use other alternatives for control. The approach presented in this paper is to use speech for controlling the game. People with motor system problem, limitation of moving the upper limbs, problems with hand muscles, hand joints, muscular stress, and many other neurological disorders, can be highly benefited by using voice driven command system in gaming. Age related issues can make the hand of a person weak which is quite common and hence use of speech is reasonable for such users for game control. Also the use of speech does not require any special hardware. Ordinary computers of today are able to carry speech transforming operation from user to the process handler of game, hence there is no need

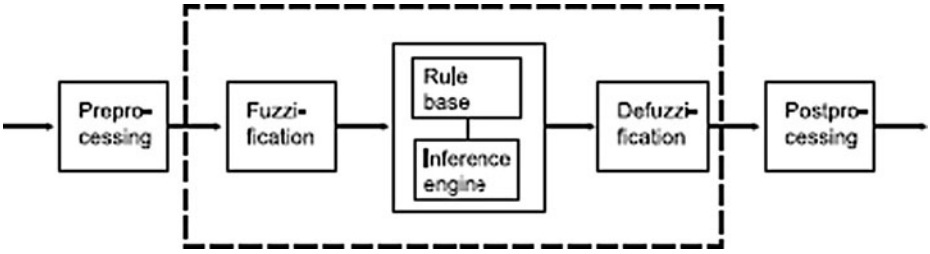

Fig. 2 Structure of a fuzzy controller

of any extra hardware and disabled people do not need to wear or use any special device with them to use speech system for controlling games. For this reason, the use of the human speech modality is a constructive move toward achieving universal design of games.

3 Proposed system

3.1 Description of the proposed system

Figure 3 shows the block diagram of the proposed system. The proposed system consists of a standard automatic speech recognition system where the vocabulary array is built by fetching the commands from the most commonly used game platform. The user input is processed through the soundcard which converts the analog data to digital format and then checks with the user word model from the vocabulary register. Once the word uttered by the user matches with the word from the vocabulary list, it then sends the binary signal corresponding to that word to the process handler of the game. Users see that the action has been taken according to their given voice command, through the GUI of the game. However, it is important to utilize the user's emotion to control the character of game in different manner. So a command voiced in low pitch should result doing something different than when it is said in a higher pitch.

3.2 The traditional controller

The idea of using a controller between the word recognition system and the communication handler of the game engine is, that the user's command can be interpreted in a different

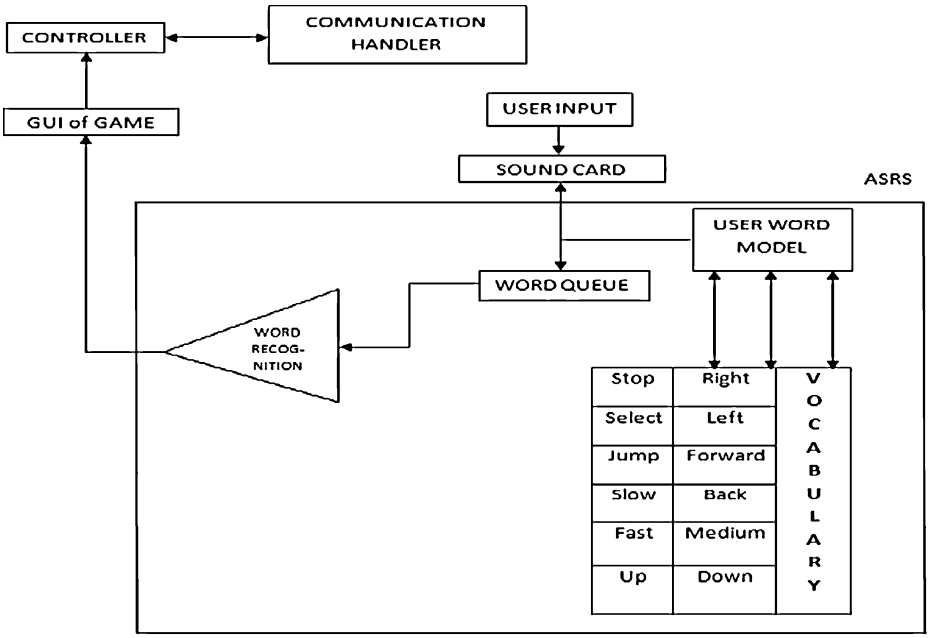

Fig. 3 Proposed automatic speech recognition system model for controlling game

manner; that is, the way users send the command for executing certain operation. The controller proposed here can be everything from simple logic based controller to a sophisticated fuzzy controller. A probability based model like HMM [5] can also be used to construct algorithm which is showed in the result section. A logic or probability based controller following certain algorithms derived by the method of this research can be programmed on a single micro chip (a PIC microcontroller for example) which can be used for universal voice access prospect for controlling game.

4 Method

4.1 Selection of a set of commands

To initiate this research computer game has been classified while thinking in mind that both old and disable people of any age-present and future can be the user of the game. The classification of computer game that was picked in this research was: action, first person shooter, strategy, board, action-adventure and sports. At least three games from each of the group of classified games were played and observed. The most frequently used commands from each game were filtered and listed. Then from each category of game, after analyzing the three different game's filtered commands, the best possible commands were picked up for each category of game. The same technique is used to filter and fetch best commands for all classified groups of game.

Two groups of users were involved in playing and observing the games and were told to write down the command they are mostly using by the input devices. They were asked which command they would like to use if the game is to be controlled by voice command. Their written commands and personal opinion of using voice commands were closely matched before creating the set of commands from different categories.

4.2 Obtaining user voice command samples

Next step was to record and monitor the frequency of each commands uttered by the users. Six users were selected in these test- three males and three females. Their voice uttering the commands were recorded and tested in two different environments. One was fairly silent environment which was assumed to be the normal settings when we usually play a game in our computer. Another was a little noisy atmosphere. The idea was to see, how the level of frequency of command varies with respect to different environment. Also each user was told to give their voice command three different times in two different environments. They were requested to utter commands this way: one as the way they want, one with some emotion on it and one with no emotion at all. Therefore, one particular command has been recorded from one user for analysis, total 6 times in two different environments and one particular command had 36 different data; 18 from three male users and 18 from three female users for further analysis. For this test, the audacity audio editor and recorder (available in: http://audacity.sourceforge.net/) for windows was used which is a free multilingual audio editor.

4.3 Analysis of voice data

The frequency domain filtering uses a Fourier transform filtering technique implemented in MATLAB. In this filtering technique noise is basically considered as unwanted high frequency content in the sample data and applying a low pass filter to the noisy data allows

low-frequency components to remain unchanged while high-frequencies are smoothed or attenuated. After feature extraction using frequency domain filtering method the best data for one particular command recorded by three male and three female users in two different settings was selected which was a satisfactory average value of frequency and intensity for each command and repeating the same procedure frequency and intensity values for all 15 commands were picked up.

5 Results

5.1 Set of commands

Once the individual category's selected commands were found, a set of commands from 'all type of games' were created. If a command from one category was member of the main set, then it was picked up. This is how the best possible 15 commands were picked up from all these categories of games, considering that they are mostly used and quite obligatory at controlling any of the categorized games.

Table 1 shows a sample data collection from three different users while involving in selecting best commands from different types of games.

5.2 User voice samples

The next step was to analyze the recorded data. For doing this, each individual command from six different users were plotted in graphs to analyze the spectrum. The frequency (Hz) and intensity (dB) values were picked up from the graph and frequency domain filtering method was used to eliminate the noisy dataset. Figure 4 shows the spectrum analysis of the command 'Select' by one male and one female user.

5.3 Analysis of voice data

Figure 5 shows how the frequency domain filtering was used to remove noise from the data obtained from the spectrum. The left most curve shows the filter function that was constructed. Then, to filter the data in the frequency domain, the Fourier transform of the data is multiplied by the frequency response of a filter and then an inverse Fourier transform was applied to return the data to the spatial domain. The third curve shows that the same filter function can also be used as a high pass filter that is, allowing only the high frequency or noise components through. MATLAB was used to generate these curves.

The curve in the middle shows the effect of using a low pass filter on the noisy dataset while the right curve shows the use of a high pass filter for the same function which allows

Table 1 Sample data from three different users

Game type	User 1	User2	User3
Action	Fire	Right	Left
First person shooter	Select	Fire	Back
Strategy	Stop	Select	Forward
Board	Select	Back	Forward
Action-adventure	Jump	Fire	Select
Sport	Stop	Slow	Forward

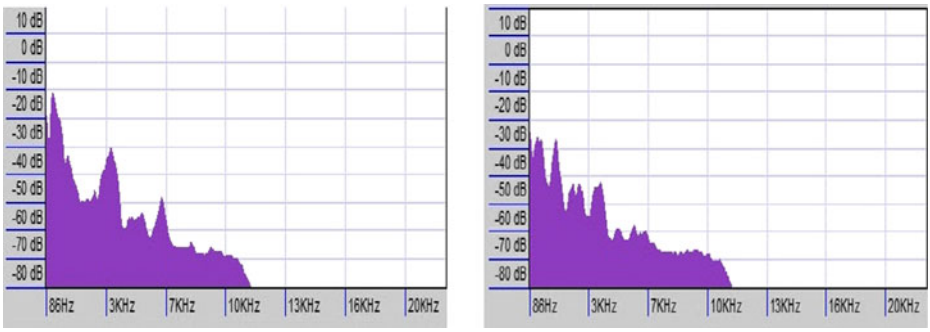

Fig. 4 Left spectrum is obtained by a male player and the right spectrum is obtained by a female player for saying same command 'Select'

only the high frequency or noise components through. The high pass and low pass filtering was simultaneously used for male and female voice data according to the number of high or low noisy level in data.

5.4 Algorithm

The frequency values well filtered and obtained from different commands were further analyzed to take the average, which is the mean value. The Hz and dB level was ranged based on the filtered data and was then used to construct a controller which is connected between the automatic speech recognition system and the game communication handler. The ASRS system detects and matches the stored vocabularies which in this case are the commands, while the controller checks the emotion of the uttered command (frequency and intensity) to instruct the game communication handler what to do according to the received pitch from a user. It is possible to construct a simple algorithm using the straightforward logic based loop where the purpose will be served but the run time might be quite long. Hidden Markov model (HMM) [5] is the broad-spectrum basis on modern form of speech recognition system and is a statistical model where the states are not directly visible to the user. This model can use the result of this research. The idea of using the frequency values obtained from the different users commands in HMM which combines many other algorithms for temporal pattern recognition in the game (speech in this case), can be viable. An example algorithm from this research for deciding a command between 'Fire' and 'Jump', based on statistical HMM may look like as follows:

```
States=('Jump', 'Fire')
Observations=('Hz, 'dB')
```

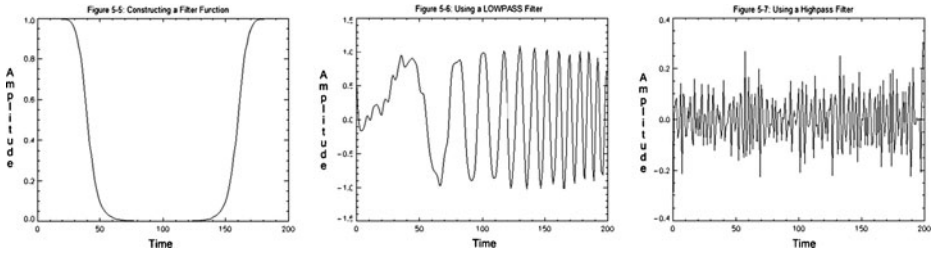

Fig. 5 Frequency domain filtering method for eliminating noise data obtained from voice spectrum

```
start_probability={'Fire': 0.6, 'Jump': 0.4}
transition_probability={
'Fire' : {'Fire': 0.7, 'Jump': 0.3},
'Jump' : {'Fire': 0.4, 'Jump': 0.6},}
emission_probability={
'Fire' : {'Hz': 0.7, 'dB': 0.3},
'Jump' : {'Hz': 0.6, 'dB': 0.4},}
```

From the above piece of code it can be understood that, if a user tends to use higher tone to utter a command it is possibly 'Fire' with higher probability assigned in the function start_ probability. The probability of the command 'Fire' to be 'Fire', when 'Fire' really is uttered is high when it is uttered in a high pitch otherwise it is 'Jump' command and this logic is assigned in the function transition_probability. The last function called emission_probability uses the frequency (Hz) and intensity (dB) value's average (in this example) or limit and decides the uttered command received from user and sends it to the game communication handler for the execution of the command on game environment and show it on the game's GUI.

5.5 The fuzzy controller

The proposed fuzzy controller's block diagram is shown in Fig. 6. The fuzzification process involves handling voice command variables (d1,d2….dn), output variables (ω and v) and transforming numerical variables into fuzzy sets. Two output variables ω and v are used respectively for determining the angular velocity and the linear velocity of an object. Hence state variables (d1,d2….dn), are decided with the error and change of error by the user's voice input. Here, error is determined by the error from the processed output minus the error from the last processed output. The controlled variables ω and v are decided, as said before to determine the angular and linear velocity of an object which is important to consider for the different movement based on the user-given voice command. The fuzzification process converts the crisp input values to fuzzy set values.

An inference method was chosen to design this fuzzy controller which is the Mamdani inference method [19]. Also the fuzzy singleton [13] fuzzification method is selected as it measures the state variables without uncertainty. The inference process defines the connective implication and rule combination operations. The controller uses minimum connectives and a singleton sum product inference mechanism. This choice relies on two

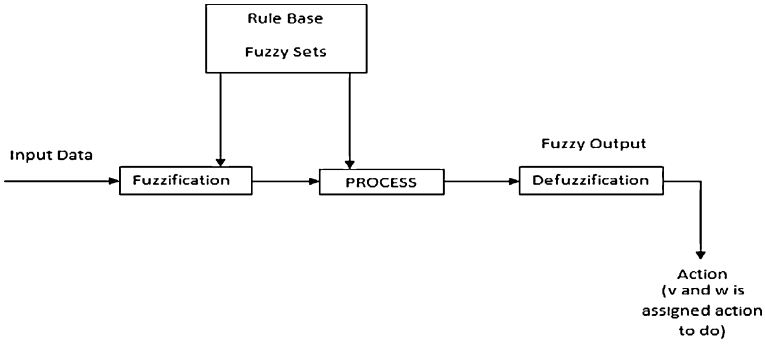

Fig. 6 Proposed fuzzy logic controller system

reasons. First, the product preserves the shape of the output fuzzy set and second, with the sum, the result is influenced by difficult rules, reaching the same conclusion.

Output variables (ω and v) are also modeled as discrete fuzzy set as it increases speed in inference and defuzzification process. Figure 7 shows the shape of fuzzy input sets where d1, d2, d3……dn are considered to be the voice sensor data received from the user. The angle θ is used as a guidance operator.

That is, the way d is sensed from the user input, triggers the value of and puts a decision value on the output variable describing at which angle the object should move. The value of is discritized and normalized within the range of [−180,180]

Figures 8 and 9 shows the shape of fuzzy output sets where limit of v and ω are discretized and normalized within the range of [0, 8] for v and [−8, 8] for ω.

The numerical values obtained from the two output sets are sent for a combination of two sets, and then sent to the communication handler of the game to execute the desired action. The content of the fuzzy set are linguistic terms which is the rule based fuzzy set used to measure the pitch of the voiced command to act differently from the character of interest.

The center of gravity (COG) [12] method is used for defuzzification strategy, as in this problem it results less processing time with no degradation of defuzzified value. Also COG method simplifies the computation of inference mechanism.

5.6 Fuzzy linguistic terms

Sample linguistic terms for the rule base fuzzy set for few game commands are shown below. The knowledge of these commands is picked up from most widely used commands of controlling game using voice.

R	Right
SR	Right turn slow
MR	Right turn medium
FR	Right turn fast
VFSR	Right turn very fast
L	Left
SCT	Select
NS	Not select
ST	Stop
U	Up
MV	Move
S	Slow
M	Medium
FS	Fast
VFS	Very fast

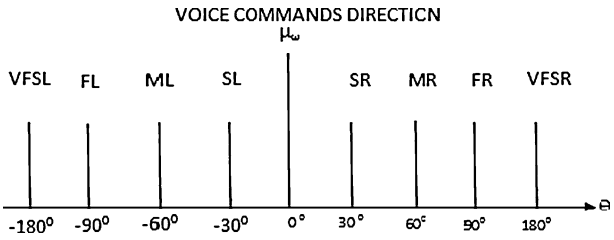

Fig. 7 Shape of fuzzy input sets deciding direction of object

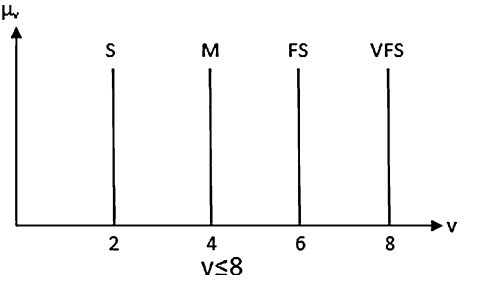

Fig. 8 Shape of fuzzy output sets for linear velocity

5.7 Sum-product inference

Degree of each rule is determined by the following rule:

βl=min(μAi(di)…… μAi (dk)). Ai are the input fuzzy sets that is the defined linguistic terms and dj (j=1….k) are the input variables.

Output fuzzy set Bl′ is constructed using following rule:

For each rule l, we use product operation μ βl′(v)=βl μBl(v). Here Bl are the output fuzzy sets and v is the output variable.

Combination of output fuzzy sets to single fuzzy set B′:

$$\text{Sum operation} : \mu B'l(v) = \mu B'l(v) + \ldots + \mu B'_{N}(v)$$

5.8 Construction of fuzzy rule

The following example fuzzy rules are the representation of the gathered knowledge of user's voice command for game control.

Turn Left:

If θ is SCT and d0 is L and d1 is SCT Then v is ST and ω is L

Stop:

If θ is MV and d0 is SCT and d1 is ST Then v is ST and ω is ST

Right Turn Fast:

If θ is SL and d0 is R and d1 is FR and d2 is MV Then v is ST and ω is FR

Select and Move Forward

If θ is NS and d1 is SCT and d2 is FW and d3 is MV Then v is MV and ω is ST

5.9 Implementation- hardware and industry perspective

Use of PIC (Programmable Interface Controller) microcontroller, to program the logic instructions derived in this research can be a crucial initiative. For example, if a soundcard of the computer comes up with a built in chip with the certain instructions of voice

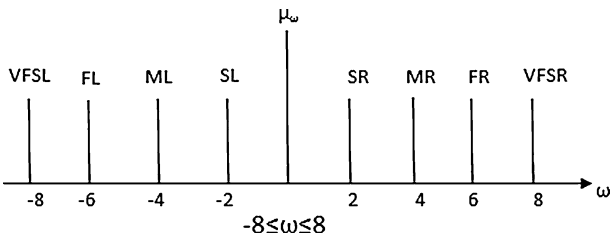

Fig. 9 Shape of fuzzy output sets for angular velocity

command, then the soundcard can be called inclusive design supported soundcard, designed for controlling games with specific voice commands. Game designers then only need to integrate their control structure of the game with that hardware. There are two advantages of implementing the results in hardware form. First, it can be highly effective at replacing relatively complex discreet logic. Secondly, if probability based algorithm can be implemented and programmed in chip, it may replace the idea of using a speech recognition system at all; as long as we focus on certain voice input parameters. The memory type in the PIC is EPROM hence PIC can alter the data in the memory and can retain its value, even when the power is removed. The company (Microchip technology) that designed PIC provides a freeware IDE package called MPLAB that can be used to program PIC. Using 'Programmer' which is hardware to configure programmable non-volatile circuit such as EPROM, a PIC can easily be programmed with certain instructions to accomplish desired task.

It is now understood that, hardware implementation of the proposed system have several constructive issues that are going to come to pass. First, the overall system will be more compact and integrated. The command execution time will be reduced. The speech recognition system might not be required for pattern matching of word from a vocabulary array, as it can be replaced by the stylish use of algorithms to find out selected commands where the algorithm comes to a decision of the instruction based on the frequency level and intensity values from the user. This can lead to an industrial and engineering challenge. While the hardware satisfies certain voice commands in built-in format, the game industries are then going to be under press to integrate with the service from such hardware with their designed games. A universal design standard thus can be produced and games can be benchmarked in terms of their inclusivity and also can be standardized in some structure.

6 Discussion

The performance of the proposed system depends on how the algorithm is implemented. If it is a clean code based controller, the run time will be interesting to observe to evaluate the system's performance. The proposed model using the frequency level to detect and generate commands can be made more sophisticated while issues like word error rate, single word error rate and command success rate will be verified and fine tuned. Use of Viterbi algorithm is reasonable as long as it does not increase the cost of the overall algorithm. As the system will mostly be dealing with old or disabled people, decoding the speech might be difficult sometimes, when the system is presented with a new utterance and it must then compute the most likely source word. Using Viterbi algorithm is thus prudent to find the best path which can be an added advantage in this system. Using a fuzzy logic controller between ASRS system and game controller is definitely going to make the command execution process faster. However in the proposed fuzzy controller, some issues have been considered in terms of performance. The measurement of uncertainty is the first issue where the classical rules to find the max or min value has been use. Second, the importance measurement was considered, that is the uncertainty measurement rule might be altered by the value of importance of the members in fuzzy set, which may not just give the priority to the maximum or minimum value in this case. The quality measurement rule was also considered. The quality of a rule is measured by multiplying the value obtained by a rule with the quality value assigned to a rule. The quality value assigned to a rule can be 0 or 1 where 0 describes a rule to be lower quality and 1 describes the rule to be of a good quality.

7 Conclusion

The use of speech to control games is only one step towards inclusiveness and thereby universal design and accessibility in game. Other human modalities while combined together with speech will give more optimal performance, overcoming any drawbacks of using only speech. The result from this paper has multiple possibilities in game design and computer industry. While any existing algorithm can be altered and new algorithm can be created for detecting user emotions for executing one command in different way in game playing; which is important for elderly adult or disable people, an electronic chip can be programmed and integrated with sound card (for instance a PIC), substantiating the lasting universal design of certain standard of voice command for playing game. This opens an enormous opportunity both for the hardware manufacturers and game industries. Voice support for the game, based on this paper's idea can thus shows the way towards finding an industry standard of universal design for the computer games.

References

1. Carbonell N (2003) Multimodality: a step towards universal access. Universal Access in the Information Society 2(2):89–90
2. Clarkson J, Coleman R, Keates S, Lebbon, C (eds.) (2003) Inclusive design: design for the whole population, pp 88–102
3. Crawford,C (1982) The art of computer game design
4. Czaja S, Lee C (2007) The impact of aging on access to technology. Universal Access in the Information Society 5(4):341–349
5. Eddy SR (1996) Hidden Markov models. Current Opinion in Structural Biology 6(3):361–365
6. Eurostat, Eurostat, Internet usage in the EU25 in 2005. 2006.
7. Goodman J, Langdon PM, Clarkson PJ (2006) Equipping designers for inclusive design. Gerontechnology 4:229–233
8. Goodman-Deane J, Keith S, Whitney G (2009) HCI and the older population. Universal Access in the Information Society 8(1):1–3
9. Hawthorn D (2000) Possible implications of aging for interface designers. Interact Comput 12(5):507–528
10. Hinckley K (2003) *Input technologies and techniques.* Lawrence Erlbaum
11. Holzinger A (2008) Universal access to technology-enhanced learning. Universal Access in the Information Society 7(4):195–197
12. http://farfromneutral.com/kaizen/center-of-gravity-method/comment-page-1/
13. http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol2/jp6/article2.html
14. Langdon P, Clarkson J, Robinson P (2007) Designing accessible technology. Universal Access in the Information Society 6(2):117–118
15. Neerincx M et al (2009) Attuning speech-enabled interfaces to user and context for inclusive design: technology, methodology and practice. Universal Access in the Information Society 8(2):109–122
16. Picard RW (1997) Affective computing. MIT, Cambridge
17. van Leeuwen DA, Martin AF, Przybocki MA, Bouten JS (2006) *NIST and TNO-NFI evaluations of automatic speaker recognition,* Computer. Speech Language 20:128–158
18. Van Wijngaarden SJ, Smeele PMT, Steeneken HJM (2001) *A new method for testing communication efficiency and user acceptability of speech communication channels.* Proceedings of Eurospeech 2001, Aalborg, September pp 1675–1678
19. Ying H (1993) The simplest fuzzy controllers using different inference methods are different nonlinear proportional-integral controllers with variable gains. Automatica 29(6):1579–1589

Moyen Mohammad Mustaquim is a doctoral researcher from Uppsala University. Moyen Mustaquim obtained his batchelor of computer science and engineering degree from Bangladesh and Masters degree from Uppsala University and presently working as a doctoral researcher and teacher in the Human Computer Interaction division on Department of Informatics and Media in Uppsala University. He has published so far in several conference proceeding. This is his first journal article. He lives in Uppsala, Sweden. He likes music and cooking besides research and teaching.