eNTERFACE'13

João Manuel Ferreira Gameiro

Kinect-Sign

Project Proposal for eNTERFACE'13

FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

[February, 2013]

Abstract

This project main goal is to assist in the communication between deaf and non-deaf people. To do so, the project will consist in the development of a serious game that teaches sign language through *Kinect* NUI. The game will be composed of two modes: *Teaching* and *Competition*, where the user can learn sign language in the *Teaching Mode* and then apply what he has learned in this mode in the multiple games available in the *Competition Mode*.

The game, will be built with the *Unity* game engine to create a more realistic ambience for the game, also there will be an avatar inside that game that will respond to the user gestures and movements that are recognized with a *Kinect* sensor. This *Kinect* sensor will also be used on sign language recognition, recognizing letters, numbers, words and expressions, these being the most important part of the game, since these are the main goal to achieve during development process.

After the game is fully developed the objective is to make it available online for any deaf association or school to use it. This way assisting in the teaching of sign language.

Keywords: Sign Language, Serious Game, Kinect, Unity, Gesture Recognition

Index

1. PROJECT OBJECTIVES 1
2. BACKGROUND INFORMATION 2
2.1 Sign Language 2
2.2 Technical 2
2.2.1 Kinect 2
2.2.2 Unity 2
3. DETAILED TECHNICAL DESCRIPTION 3
3.1 Technical description 3
3.1.1 Game 3
3.1.2 Natural User Interface 4
3.1.3 Ambience 5
3.2 Resources needed 5
3.2.1 Equipment 5
3.2.2 Software 5
3.3 Project management 5
4. WORK PLAN AND IMPLEMENTATION SCHEDULE 6
5. BENEFITS OF THE RESEARCH 7
6. PROFILE OF TEAM 8
6.1 Leader 8
6.1.1 Education 8
6.1.2 Others 8
6.2 Other researchers needed 8
7. REFERENCES 9
7.1 General Knowledge 9
7.1.1 Sign Language 9
7.2 Software 9
7.2.1 Kinect 9
7.2.2 Unity 9
7.2.3 Integration Kinect with Unity 9

1. Project objectives

The project's main goal is to assist in the communication between deaf and non-deaf people. To do so, the project will consist in the development of a serious game that teaches sign language through *Kinect'* NUI. Other goal for this project will be the development of a base software for handling gestures and movements through the *Kinect* sensor that can be evolved into a software for handling virtual objects.

This way, the serious game will be composed of two modes: *Teaching* and *Competition*. The *Teaching mode* is for teaching sign language in the most traditional way. This means the user is embedded inside a classroom where he will learn sign language, from the basics, like the alphabet, to some complex expressions. The user will be able to try way he was learned in a fun way, through the *Competition mode*. This mode will be divided into two games: *Quiz* and *Lingo*, where the user will be able to play and try to outrank other users with better scores.

A side goal is to develop a software to handle gestures and movements through the *Kinect* sensor that is easy to understand and easy to implement in any other project, so that it can be developed for handling virtual objects in the future. So the base goal for this software is to record gestures and movements, recognize gestures and movements and inform the main project, in this case being developed in the *Unity* game engine, the movement that the user is doing.

2. Background information

Throughout the entire project there will be required knowledge regarding natural user interfaces, Unity development and machine learning, based on neural networks.

2.1 Sign Language

Today's world faces an increasing deafness problem. Two main reasons can be identified: Genetic and Behavioural:

- **Genetic** – Genetics represent 50% of the causes of deafness. Two babies per thousand births have a severe deafness. In Portugal, for example, there are 180.000 people with hearing impairment and this number tends to increase.
- **Behavioural** – More and more teenagers take a position of breaking rules, which leads them to nights at the disco, video games, listening to very loud music, among other things. These factors lead to the exhaustion of a very important sense, the hearing, since it's being taken to the extreme.

The sign language was created in order to combat the major problem that comes with deafness, the communication between peers. This language was created from fundamentals of spoken language and is based on the spontaneity of communication between deaf and non-deaf individuals. Referring to this, it should be noted that not all the population is familiar with this type of language and that the knowledge on sign language is not mandatory.

2.2 Technical

In order to develop this serious game some technical background in the fields of natural user interfaces and/or 3D game development is required. To support this fields there will be used the *Kinect* sensor and the *Unity* game engine.

2.2.1 Kinect

Kinect, a motion sensing input device, enables users to control and interact with the Xbox 360 and the computer without the need to touch a game controller. This control is made through a natural user interface (NUI) with the support of video, gestures and spoken commands.

This device features a colour sensor (RGB Camera), a depth sensor (infrared laser projector + monochrome CMOS sensor) and a multi-array microphone which provides the NUI support described.

The gesture feature provides full-body 3D motion capture and facial recognition, assisting in the control and interaction with devices stated.

2.2.2 Unity

Unity is a cross-platform game engine and IDE, from where we will be developing the project into a desktop/console video game.

This platform supports integration with *3ds Max*, *Maya*, *Softimage*, *Blender*, among others. The graphics engine uses *Direct3D*, *OpenGL*, *OpenGL ES* and proprietary APIs, supporting multiple graphics effects, such as dynamic shadows using shadow maps, reflection mapping and render-to-texture.

Game scripting comes via *Mono*, the open-source implementation of the *.NET* framework. To do the scripting *Unity* ships and uses a customized version of *MonoDevelop* assisting in debugging scripts.

3. Detailed technical description

In the following chapter, a detailed description of the desired game will be presented. It will be described the game features, objectives, what sort of support the game offers and what are the requirements both to play and to develop the game.

3.1 Technical description

3.1.1 Game

The game will be divided in the following categories:

– *Teaching Mode*

In this mode, the user is presented with a classroom, where the main goal for the user is to learn sign language. To do so, the learning will be divided into the following categories: **letters**, **numbers**, **words** and **expressions**. For each category there will be groups of five lessons with a final evaluation at the end of each group. The number of groups will vary according to the categories, for example the **numbers** category will have at least two groups since there are 10 digits and the **letters** category will have at least six groups since there are 26 characters.

– *Competition Mode*

For *Competition Mode*, the user will be inside a TV Show scenario, where he can play the following two games:

- *Quiz*:

As the name implies *Quiz* is where the game makes a questionnaire of five questions to the user, one at a time, giving four options for the user to answer correctly. The questions will be sign language related and the answers must be given in sign language. This questionnaire will have a 10 minute time limit to be answered and afterwards the score will be determined and posted in a scoreboard, in case of outranking other scores previously made.

Example:

The game makes the question: “What is the correct symbol for the letter ‘A’?” and with it gives the images that represents the symbols of the letters ‘B’, ‘F’, ‘A’ and ‘P’, in this order. Then the user must repeat the letter, in sign language, to give the answer.

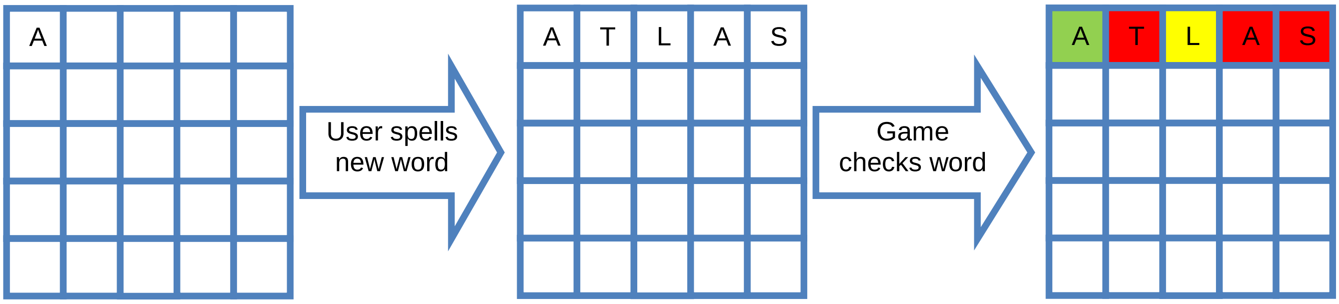
- *Lingo*:

Lingo is like the game *MasterMind*, but is based on words with 5 characters. This means that the user must find the correct word within a 5 limit number of trials, where the game will be saying if the characters are correct and if they are in the right place, according to the colours red, for not existing word, yellow, for existing word in the wrong place, and green for existing word in the right place. So for the user to insert a new word, this word must be spelled through sign language.

To complete a full game of *Lingo* the user must determine 5 words within a 15 minute period of time. After this a score is calculated, according to the number of correct words found and the time the user needed to complete the game, putting the user in a scoreboard, given the fact that he outranked other users.

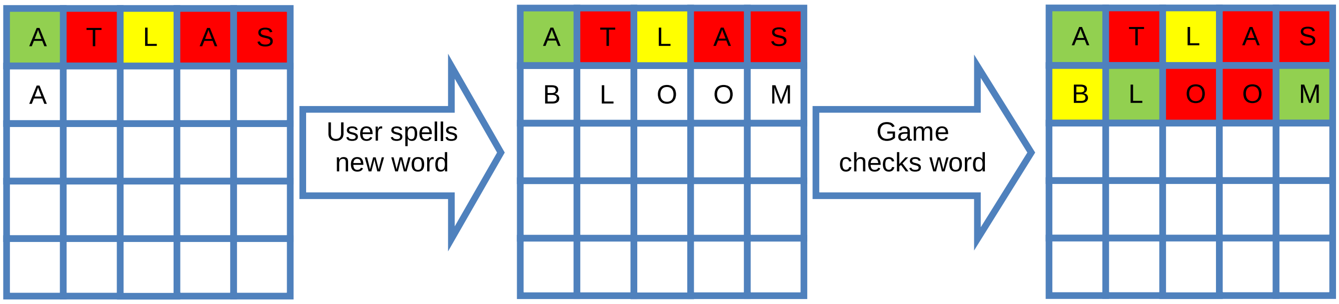
Example:

The game has the word “Album” to be correctly found, giving, as an initial foothold, the ‘A’ letter, since is the first letter of this word. Now it’s the time for the user to spell a word that he thinks might be the correct one, for example “Atlas”. After that the game determines the validity of the characters giving the result in the following illustration.



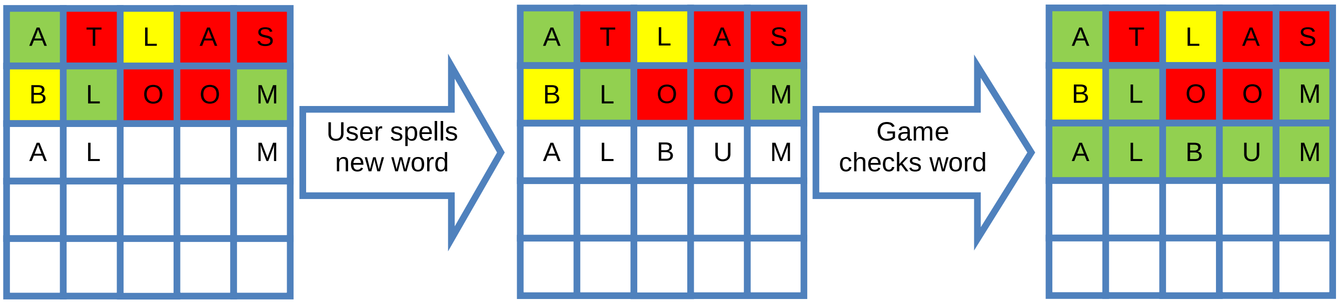
As it can be seen in the previous illustration the first letter ‘A’ is correct and in the right place, the letter ‘L’ exists in the word but is in the wrong place and the other characters do not exist, in the wanted word “Album”, appearing in red. In case of the wanted word containing another ‘A’ the letter in the board should have been green if it was in the right place or yellow if it was in the wrong place.

For the next round, the correct characters are shown in the right place, unrelated to the fact that those characters where used in the last word. To show this, we will choose the word “Bloom”, for the next word that the user says.



As it can be seen in the previous illustration the word “Bloom” contains the ‘L’, previously used, and the ‘M’ letters in the right place in relation to the wanted word “Album”. Another correct letter that appears is ‘B’, although this is in the wrong place in relation to the wanted word, this way bearing the colour yellow.

So for the next round, the game will give the letters ‘A’, ‘L’ and ‘M’ in the right place, since the user already found the right place for these letters. Knowing the position of this three letters and knowing that there is a letter ‘B’ in the remaining two slots available the user can now say with some degree for certainty that “Album” is the correct wanted word, spelling this word in sign language.



After the game has checked that the user spelled the wanted word correctly, it gives the entire word at green, as it can be seen in the previous illustration, and moves on to the next word.

3.1.2 Natural User Interface

The game is to be implemented in the most natural way possible with the assistance of the *Kinect*. So the objective is for the user to interact directly with the game environment through gestures and movements, for example to go from the classroom to the TV Show scenario by the user making movements that will reflect in its in-game avatar.

The most important usage of the natural interface will be the sign language recognition. To do this recognition a mix of image processing and depth data processing will be used. More accurately the depth data will be used to record the hands position in space and to record the movement patterns. The image processing will be to recognize the current gesture that the hand is making and mixing with the movement patterns will be used to recognize the multiple gestures and features.

3.1.3 Ambience

The final game will support an ambience, developed with assistance of the **Unity Game Engine**, which will embed a classroom, for the *Teaching Mode*, and a TV Show scenario, for the *Competition Mode*. During the game the user will be represented as an avatar that reproduces the user movements. According to those movements, the user will be able to move around the scenario and pick what he wishes to play.

The idea behind the classroom will be to show a chalkboard where the signs appear for the user to repeat them. Other decorative elements for the classroom will be designed, such as the schoolteacher table, chairs and others.

For the TV Show scenario there will be needed to build a dark room with a tribune where the user will play the various games presented in this mode. To increase realism there will be added an audience filled with stickmen cheering and supporting the user. Some decorative elements will be required, such as light bulbs, cameras and any other that seem to be necessary.

3.2 Resources needed

3.2.1 Equipment

There will be needed at least one Kinect Sensor.

3.2.2 Software

For developing the serious game the following software is required:

- Kinect for Windows SDK
- Kinect for Windows Developer Toolkit
- Unity 4

3.3 Project management

The Project will be divided into three main groups: **Natural User Interface** (*Kinect*), **Ambience** (*Unity*) and **Game Progress**. As it's explained in the sub-chapter 3.1 the **Natural User Interface** group is to develop the sign language recognition and the interaction with the game through an avatar. This avatar will be developed in cooperation with the **Ambience** group, also responsible to develop the game ambience such as the classroom, the TV Show scenario and the overall linking background. Later on the two groups will merge to form the **Game Progress** that will be responsible for developing the main game events, like the lessons that occur in the classroom or the *Lingo* game occurring in the TV Show scenario.

4. Work plan and implementation schedule

The following timetable details' the work that must be done by the end of each week, for each group.

	Natural User Interface	Ambience	Game Progress
Week 1	- Sign Language: - Hand position storage, record the sequences of position of the hand - Hand image storage - Hand image preparation for comparison	- Classroom: - Background - Classroom elements, such as chalkboard and desk/chairs - TV Show scenario: - Background - Tribune for the avatar - Avatar: - Generic avatar for Kinect experimentations	
Week 2	- Sign Language: - Letters and numbers recognition	- Classroom: - Professor avatar for assistance during lesson - TV Show scenario: - Audience development with Stickman's	
Week 3	- Sign Language: - Integration of letter and number recognition in the game - Words and expressions recognition - Avatar: - Basic movements of the avatar within the game	- Classroom: - Drawings in the chalkboard according to the lessons - TV Show scenario: - Board for *Quiz* - Board for *Lingo* - TV show presenter avatar for assistance during the games	- *Teaching Mode* - Idealize the lessons during the game - Implement few lessons in the game - *Competition Mode* - Idealize levels and respective questions for *Quiz* - Idealize levels and respective words for *Lingo*
Week 4	- Sign Language: - Integration of word and expressions recognition in the game - Error correction - Avatar: - Complete repetition of the user movements	- Classroom: - Error correction - TV Show scenario: - Error correction	- *Teaching Mode* - Implement evaluations in the game - *Competition Mode* - Implement *Quiz* game - Implement *Lingo* game - Error correction

5. Benefits of the research

This project will bring benefits in natural user interfaces and teaching sign language through a serious game.

The world is following to a more natural usage of computers, so with the *Kinect* sensor and the gesture recognition system, this project can help in future applications or games that require handling virtual objects, providing the base software for those applications or games.

The serious game, which is the final product from this project, has the objectives to assist schools, in teaching sign language, and public places, where a translator can be a derivative of this game.

6. Profile of team

6.1 Leader

Name	João Manuel Ferreira Gameiro
Email	j.gameiro@campus.fct.unl.pt
Nationality	Portuguese

6.1.1 Education

Dates	1/9/2006 – 12/2013
Qualification	Masters in Electrical and Computer Engineering
Organization	Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Dates	1/10/2012 – 27/11/2012
Qualification	Neural Networks for Machine Learning
Organization	Coursera

Dates	1/10/2010 – 1/10/2011
Qualification	First Certificate in English (FCE)
Organization	University Cambridge ESOL Examinations

6.1.2 Others

Competition	Imagine Cup 2012 – Kinect Fun Labs
Project	K4D – Kinect 4 Deafness
Result	Top 100

6.2 Other researchers needed

For developing this project there will be needed at least:

- 2 person with experience in *Unity* and/or *Mono*
- 1 person with experience in *Kinect*

Any other person with experience in *Blender*, machine learning and neural networks is also welcomed.

7. References

7.1 General Knowledge

7.1.1 Sign Language

- Wikipedia: http://en.wikipedia.org/wiki/Sign_language
- American Sign Language Dictionary: http://www.aslpro.com/cgi-bin/aslpro/aslpro.cgi

7.2 Software

7.2.1 Kinect

- Wikipedia: http://en.wikipedia.org/wiki/Kinect
- Microsoft: http://www.microsoft.com/en-us/kinectforwindows/
  - Download: http://www.microsoft.com/en-us/kinectforwindows/develop/developer-downloads.aspx
  - Developer How-To Videos: http://www.microsoft.com/en-us/kinectforwindows/develop/tutorials.aspx
- Channel 9 – Kinect For Windows Quickstart Series: http://channel9.msdn.com/Series/KinectQuickstart

7.2.2 Unity

- Wikipedia: http://en.wikipedia.org/wiki/Unity_(game_engine)
- Unity: http://unity3d.com/
  - Download: http://unity3d.com/unity/download/
  - Tutorials: http://video.unity3d.com/channel/1649856/tutorials
  - Manual: http://docs.unity3d.com/Documentation/Manual/index.html

7.2.3 Integration Kinect with Unity

- Unity Wikipedia: http://wiki.etc.cmu.edu/unity3d/index.php/Microsoft_Kinect_-_Microsoft_SDK
- Channel 9: http://channel9.msdn.com/coding4fun/kinect/Unity-and-the-Kinect-SDK
- Nightmare Kitty: http://www.nightmarekitty.com/2011/10/28/unity-and-kinect-tutorial/