Brazilian Sign Language Multimedia Hangman Game: A Prototype of an Educational and Inclusive Application

Renata C. B. Madeo
University of São Paulo
Av. Arlindo Béttio, 1000
São Paulo, SP, Brazil
renata.si@usp.br

ABSTRACT

This paper presents a prototype of an educative and inclusive application: the Brazilian Sign Language Multimedia Hangman Game. This application aims to estimate people, specially children, deaf or not, to learn a sign language and to help deaf people to improve their vocabulary in an oral language. The differential of this game is that its input consists of videos of the user performing signs from Brazilian Sign Language corresponding to Latin alphabet letters, recorded through the game graphical interface. These videos are processed by a computer vision module in order to recognize the letter to which the sign corresponds, using a recognition strategy based on primitives - hand configuration, movement and orientation, reaching 84.3% accuracy.

Categories and Subject Descriptors

I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*Games*; I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems—*Natural language interfaces*; K.4.2 [**Computers and Society**]: Social Issues—*Assistive technologies for persons with disabilities.*

General Terms

Experimentation

Keywords

Fingerspelling Applications, Educational Application, Inclusive Application, Gesture Recognition, Sign Language.

1. INTRODUCTION

In recent years, there has been a growing concern about minorities social inclusion, including people with disabilities. Computer science can contribute greatly for this group of people through the development of assistive technologies, aiming to promote autonomy, independence, life quality and social inclusion for people with disabilities.

This paper presents the Brazilian Sign Language (BSL) Multimedia Hangman Game[1]: an assitive technology application aiming the social inclusion of deaf people, specially children. This application has two main goals: estimulating children, deaf or not, to learn BSL; and helping deaf children to improve their vocabulary in Portuguese, since, despite BSL is an official language in Brazil, knowing Portuguese is really important to overcome communication barriers.

The BSL Multimedia Hangman Game uses a simple webcam to record videos of the user performing signs belonging to the manual alphabet of BSL. The recorded video is analyzed through video processing and pattern recognition techniques aiming to identify which letter corresponds to the sign in the video. Such letter is the input for the game itself.

[1]In the Hangman Game, the user shall discover a secret word by guessing the letters which compose it.

ASSETS'11, October 24–26, 2011, Dundee, Scotland, UK.
ACM 978-1-4503-0919-6/11/10.

2. GAME ARCHITECTURE

The game architecture (Fig. 1) is composed by a Graphical User Interface (GUI) and a Computer Vision (CV) module, including video processing and recognition modules.

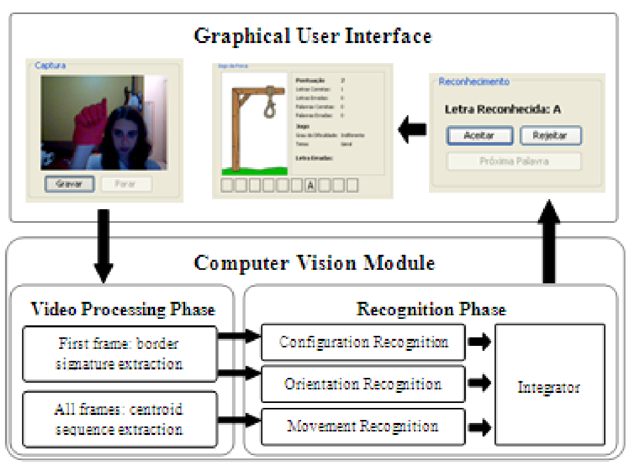

Figure 1: Game architecture: a video is recorded through a webcam. The decision about which sign was performed is used as input to the game logic.

The GUI (Fig. 2) was implemented in Java™. It allows the user to record videos, submit these videos for processing, viewing the result and accept or reject it, since the CV module can recognize a different sign than what the user meant to perform. For improving its educational role, the user can see the sign in BSL or an image illustrating the word while guessing it, as a hint, and after guessing it, in order to associate the word in Portuguese with the sign in BSL. It also includes the game logic, i.e., all functionalities needed to support a Hangman Game. Thus, once an answer

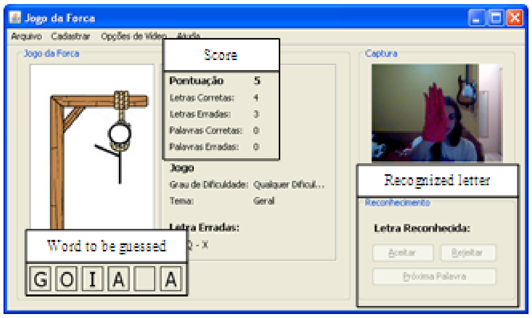

Figure 2: Graphical User Interface.

from CV module is accepted, the GUI shows if that letter is present in the word to be guessed. If it is, the letter is shown in the right position in the word; else, the letter is shown as a wrong guess and another part of the stickman appears in the gibbet. There are other general functionalities: counters to count how many letters and words were guessed correctly or wrongly in the game; a score based on these counters, which can motivate children to play; and functionalities for including new words to the game.

The CV module includes a video processing phase, which extracts the frames from each video and performs a image segmentation phase that will allow the image processing and recognition phases. To allow the recognition, the user shall use a red glove during recording, aiming to facilitate segmentation. After extracting the frames from each video using Java Advanced Imaging API, the images are processed by the segmentation routine implemented with Matlab®. The image processing phase extracts features from the videos first frame to create a border signature, which is used to recognize the configuration and orientation primitives; and from all frames to create a centroid sequence, which is used to recognize the movement primitives. This feature extraction process was already explained in previous works [1, 2].

The recognition phase was implemented in Java™ and has already been described in [2, 3]. It is composed by four subphases: three specific recognition phases, one for each primitive and an integration phase. The recognition phase uses a neuro-fuzzy and a heuristic approach, providing fuzzy output for the recognition process, i.e., membership degrees are associated between the analysed sign and each primitive. The integration phase, also described in [3], is responsible for joining these fuzzy outputs through a fuzzy grammar and provide a final decision, i.e., which letter corresponds to the sign performed in the video. The complete CV module implementation has achieved 84.3% accuracy.

It is important to note that it is possible to adapt the CV module to recognize another sign language, provided that the new signs use the same primitives - same types of hand configuration, movement and orientation - as the previous signs, and are performed with only one hand. It would be necessary just to change the rules in the fuzzy grammar.

3. APPLICATION TESTS

To get real results on the perception of users about the game developed, the application was tested by four users. These tests were performed in two stages: first, the users tested the application by themselves and wrote freely their perceptions about the game; then, users were accompanied by the developer while playing so that the source of recognition errors could be identified.

In the first stage, users reported that the recognition process was slow and presented errors frequently. Despite the concern with processing time during development, which lead the developer to reduce video quality, the recognition process time of 10 seconds[2] is still too high for users. It was also suggested that there could be some feedback during image processing and recognition, so the user could know what the system is doing and how long it is taking. Despite the problems reported, users pointed the way of interaction through shortcuts (to operate the basic funcionalities of the game, such as starting and stopping video recording) as a good feature of the system: since the dominant hand is busy performing the signs, shortcuts make it easier to play.

In the second stage, users played the game monitored by the developer. At each recognition error, segmented images were verified to assess if the recognition phase was responsible for the error or if segmentation phase had achieved lower quality than expected. From this evaluation, it was possible to conclude that about 70% of all CV module errors were caused by problems with the segmentation.

4. FINAL CONSIDERATIONS

The presented application is still a prototype and needs some improvements. Firstly, it is necessary to improve image processing, specially image segmentation, making it faster and more accurate. Secondly, the results achieved by the recognition phase has great improvement potential: any refinement in a primitive recognition could lead to a better accuracy in the complete CV module. Lastly, when the prototype tests present better results with available testers, it will be possible to organize a test with a group of children (deaf or not) to assess their opinions on the application and, finally, make the application available to these children.

5. REFERENCES

[1] D. B. Dias, R. C. B. Madeo, T. Rocha, H. H. Bíscaro, and S. M. Peres. Hand movement recognition for brazilian sign language: A study using distance-based neural networks. In *International Joint Conference on Neural Networks*, pages 697–704. IEEE, 2009.

[2] R. C. B. Madeo, S. M. Peres, H. H. Bíscaro, D. B. Dias, and C. Boscarioli. A committee machine implementing the pattern recognition module for fingerspelling applications. In *Proceedings of the 25th Annual ACM Symposium on Applied Computing 2010*, pages 954–958. ACM Press, 2010.

[3] R. C. B. Madeo, S. M. Peres, D. B. Dias, and C. Boscarioli. Gesture recognition for fingerspelling applications: An approach based on sign language cheremes. In *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 261–262. ACM, 2010.

[2]Run on a Intel®Core™2 Duo 2.2GHz with 2GB RAM.