Advanced HCI and 3D Web over Low Performance Devices

David Oyarzun, Arantza del Pozo, John Edgar Congote, Igor G. Olaizola
Vicomtech Research Centre
{doyarzun, adelpozo, jcongote, iolaizola}@vicomtech.org

Iñaki Sainz
Aholab Group, Basque Country University
Ingeniaritza Goi Eskola Teknikoa
inaki@bips.bi.ehu.es

Igor Leturia
Elhuyar Foundation
i.leturia@elhuyar.com

Xabier Arregi
IXA Group, University of the Basque Country
Informatika Fakultatea
xabier.arregi@ehu.es

Oscar Ruiz
Universidad EAFIT
oruiz@eafit.edu.co

ABSTRACT
This position paper presents the authors' goals on advanced human computer interaction and 3D Web. Previous work on speech, natural language processing and visual technologies has achieved the development of the BerbaTek language learning demonstrator, a 3D virtual tutor that supports Basque language students through spoken interaction. Next steps consist on migrating all the system to multidevice web technologies. This paper shows the architecture defined and the steps to be performed in the next months.

Categories and Subject Descriptors
D.2.2 [**Software Engineering**]: Design tools and techniques – *user interfaces.*

General Terms
Algorithms, Performance, Design, Standardization, Languages

Keywords
Advanced HCI, Virtual Characters, WebGL, Speech, Natural Language

1. INTRODUCTION
During the last 3 years, the authors of this paper have been working in BerbaTek, a strategic research project on speech, language and visual technologies for Basque, promoted by the Basque Government.

Copyright © 2012 for the individual papers by the papers' authors. Copying permitted only for private and academic purposes. This volume is published and copyrighted by its editors.
Dec3D2012 workshop at WWW2012, Lyon, France

For the field of education, we have created a demo of a personal tutor in language learning that provides a natural way to human computer interaction, by means of 3D virtual characters, speech and natural language (see Figure 1).

The tutor is a 3D avatar that shows emotions, developed by Vicomtech-IK4, and which speaks Basque and understands what is said in Basque using Aholab's technology. The tutor is capable of assisting the user in the following tasks:

- carrying out grammar (e.g. verb conjugation, word inflection) and reading comprehension exercises (e.g. fill in gaps in a text, choosing from several options), that are created automatically from texts using technology from IXA;
- evaluating the quality of its pronunciation through Aholab technology;
- and giving aids to write texts (e.g. inflection of words, writing of numbers or querying dictionaries) by means of technology from IXA and Elhuyar.

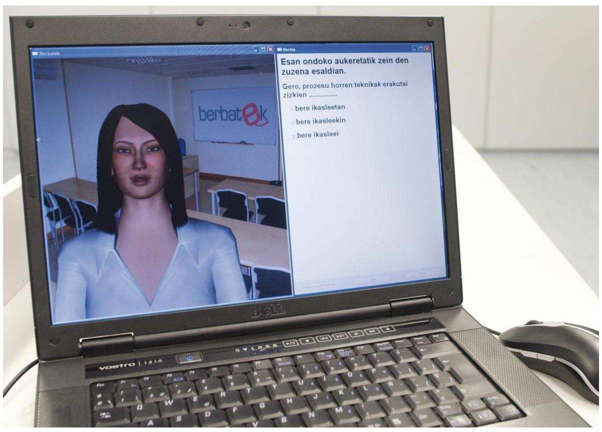

Figure 1: Appearance of the BerbaTek language learning tutor

This position paper presents the current efforts to migrate this language learning demonstrator resulting from the BerbaTek project into 3D Web technologies using the new WebGL API [1, 2].

The resulting prototype will fulfill these features:

- Avoid the use of plug-ins. Web technologies capable of mixing several media contents in a native way will be exploited to avoid the use of plugins.
- Keep all the BerbaTek demo functionalities. Both 3D virtual character and advanced HCI technologies will be included.
- Be multidevice and multiplatform. The result will be multiplatform and multidevice and designed as to be accesible from low performance devices such as standard mobile phones.

2. ARCHITECTURE OVERVIEW

After the review of the state of the art, a system architecture that takes into account the targeted multidevice capabilities has been defined. Figure 3 shows the architecture schema.

Basically, in the server side, the BerbaTek modules and databases (both internal and external) are managed. When a new device connects to the system, it automatically discovers its capabilities and decides which is the most suitable way to perform the rendering of 3D contents.

3. SERVER MANAGEMENT

The goal of this part of the architecture is to decentralize the storage of the modules.

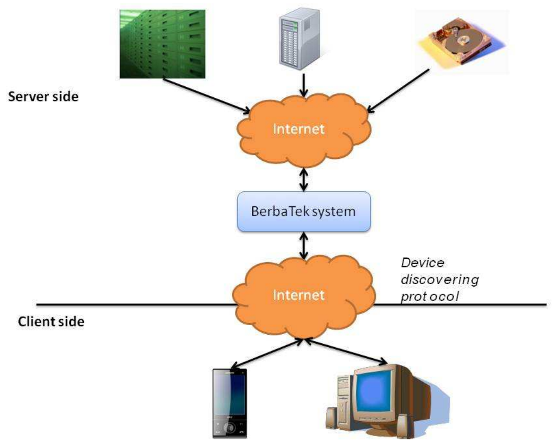

Figure 3: Architecture schema

Each arrow in the architecture represents an HTTP call. Therefore, the physical location of each part of the system is not relevant.

A typical process will be as follow:

- The user connects to the system
- The device features are detected
- The virtual character is sent via streaming or rendered in the device depending on the detected system capabilities.
- The user interacts with the system, ideally via voice. If the device does not support it, other input/output paradigms can be used.

4. MULTIDEVICE FEATURES

One of the main goals of the migration is to obtain a real multidevice system. In order to achieve that, protocols for device discovering and automatic content adaption will be implemented.

These protocols will allow the server to know which kind of device is connected and which its performance features are. With this information, the server will be able to decide the best way to send the multimedia information to the device (in-device rendering or interactive streaming). In the case of interactive streaming, standard protocols like RTP/RTCP will be used to obtain a better performance.

As a consquence, the resulting prototype will allow a complete description of immersive environments that will be automatically adapted to the specific visualization properties of the end device (e.g. from a totally immersive 360º with haptic interaction in a cave, to a small representation in auto-stereoscopic handheld devices with simple input controls).

Moreover, the interaction paradigm will be adapted to the device features too. Depending on the end-device, not only speech-based paradigms, but also new paradigms that better fit into the specific device interaction features will be automatically applied.

5. CONCLUSIONS AND NEXT STEPS

In this position paper, the goals in the evolution of the BerbaTek language learning demonstrator have been presented. The conclusions of a technological review and an architecture design have been explained. The goal of this initial work is to obtain an advanced HCI prototype based on Web 3D and HTML5 technologies.

The next steps will consist on the implementation of the system described in the previous sections.

- In a first stage, the architecture will be implemented using local infrastructure. Both PC and mobile devices will be used to test the discovery protocol.
- In a second stage, advanced interaction and visualization devices will be included, in order to implement extra features like stereoscopy.

The results of these stages will be disseminated among the scientific community by means of publications, technical workshops and ad-hoc developed web sites.

6. REFERENCES

[1] Khronos Group, “WebGL Specification,” 2011, Retrieved from https://www.khronos.org/registry/webgl/specs/latest/

[2] Behr, J., Eschler, P., Jung, Y. and Zöllner M. X3DOM: a DOM-based HTML5/X3D integration model. Proceedings of the 14th International Conference on 3D Web Technology, ACM, 2009, pp. 175-184.