Comparing video and avatar technology for a health education application for deaf people

Ionuț Adrian Chiriac \textsuperscript{a}, Lăcrămiocara Stoicu-Tivadar \textsuperscript{b}, Elena Podoleanu \textsuperscript{a}

\textsuperscript{a}University of Medicine and Pharmacy ”Carol Davila”, Bucharest, Romania
\textsuperscript{b}University Politehnica Timișoara, Romania

Abstract. The article describes the steps and results of a parallel research investigating e-health systems design and implementation for deaf people both in avatar and video technology. The application translates medical knowledge and concepts in deaf sign language for impaired users through an avatar. Two types of avatar technologies are taken into consideration: Video Avatar with recorded humans interface and Animated Avatar with animated figure interface. The comparative study investigates the data collection, design, implementation and the impact study. The comparative analysis of video and animated technology for data collection shows that the video format editing requires fewer skills and results are obtained easier, quicker and less expensive. The video technology supports an easier to design and implement architecture. The impact study for 2 deaf students communities is under development and for the time being the video avatar is better perceived.

Keywords. sign language, health education, health care, oral health, dental care

Introduction

The paper presents results of research progress in developing a health education application for deaf patients in Romanian Sign Language for oral health education and dental care. This objective may be achieved using video or animated avatar technology. The work to develop the application started with a comparative study between avatar and video technology in order to evaluate the load and complexity of the development process of such applications and also to assess the acceptance degree of the application in the specific community.

Previous work\textsuperscript{1} contains a state of the art prospective research worldwide for the domain of animated avatar. No systematic comparative research between animated and video technology avatars was found. The international projects were focused mainly on automatic translation and better integration of the deaf people in social activities. As applications concerning Romanian sign language the investigations returned modest results only related to dictionaries for sign language using video technology and no projects with animated avatars or e-health for deaf people oral healthcare education.\textsuperscript{2}
1. Methods

The general research objective is the comparative study between systems of medical education for deaf people based on video avatar technology and animated avatar technology.

Two types of avatars are taken into consideration: the Animated Avatar will display an animated character, and the Video Avatar will display recorded humans. The parallel comparative analysis between video technology and animated technology used in e-health applications for deaf people starts with the creation of the data collection and continues with the design, implementation and results analysis phases. (Figure 1)

This paper presents the results of the first two upper levels in Figure 1.

2. Results

The next paragraphs synthetically present the results of the research for the first two phases of the project (Figure 1).

2.1. Medical Data Collection - Video vs. Animated

A collection of keywords and phrases from oral health domain was selected with oral health specialists from University of Medicine and Pharmacy "Carol Davila". Working with teachers experienced in Romanian Sign Language the collection was next recorded in video format as signs and signs phrases performed by two different interpreters from different regions of Romania, discussing different regional signs.
The human signs recorded earlier were edited with eSIGN Editor\(^3\) so the final result was a collection of 262 key signs in HamNoSys\(^4\) and 96 phrases as SiGML\(^5\) files.\(^1\)

![Figure 2. Signs represented parallel Video vs. Animated](image)

Figure 2 presents snapshots for parallel signs first represented by the two different interpreters in video format and the same signs represented by the animated avatar. The results of both recordings were afterward filtered and selected in collaboration with Romanian Sign Language specialists, activity resulting in a collection of 101 words and syntagmas, 37 for Ana and 64 for Andreea.\(^1\)

<table>
<thead>
<tr>
<th>Table 1. Medical Data Collection – Video vs. Animated</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Avatar Technology</strong></td>
</tr>
<tr>
<td>Data Format – Animated files</td>
</tr>
<tr>
<td>SiGML</td>
</tr>
<tr>
<td>Translation/Conversion Module</td>
</tr>
<tr>
<td>– XML, Esign, HamNoSys</td>
</tr>
<tr>
<td>Interface – Java, Esign</td>
</tr>
<tr>
<td>Animated figure</td>
</tr>
</tbody>
</table>

Table 1 shows in a very condensed form the results of the comparative analysis of the first phase of the research – the parallel composition of data collections.

2.2. Application Design&Architecture - Video vs. Animated

The second phase of the research was dedicated to the study of the architecture for both parallel versions of the application. Originally a general architecture was composed for both types of applications and then building the two separate variants, studying the differences.

![Figure 3. Application Design&Architecture - Video vs. Animated](image)

Figure 3 represents parallel architecture snapshots for both cases, with video and animated technology\(^6\)
The comparison is looking for quality of the sign associated movements, expressivity, detail, flexibility, and storage space needed for data collection, and is intended to guide future developers of applications in the domain. (Table 2)

Table 2. Application Design&Architecture – Video vs. Animated

<table>
<thead>
<tr>
<th>Avatar Technology</th>
<th>Video Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower quality (mechanical movements)</td>
<td>higher quality (smoothness of movement)</td>
</tr>
<tr>
<td>less expressivity of the face</td>
<td>more expressivity of the face</td>
</tr>
<tr>
<td>less precision for details</td>
<td>more precision for details</td>
</tr>
<tr>
<td>phrases require time and specialists to work with the editing tools</td>
<td>HD quality is less expensive and very accessible</td>
</tr>
<tr>
<td>flexibility in editing changes</td>
<td>less flexibility in editing changes</td>
</tr>
<tr>
<td>less space used on storage devices</td>
<td>more space used on storage devices</td>
</tr>
</tbody>
</table>

Table 2 shows in a very condensed form the results of the comparative analysis of the second phase of the research – the parallel comparison of the Application Design&Architecture.6

3. Discussion

3.1. Medical Data Collection - Video vs. Animated

The two collections differ as data format and this determines the main characteristics of the two applications.

For the video avatar format, the recordings were made first in MOV format. Because MP4 became in the last years almost a standard and has more widespread support than MOV, the recorded files were further converted in MP4 format. Operations such as crop and proportions changes were sometimes required in the editing process. Because of the widely available Adobe Flash Player the FLV format is now the most widespread de facto standard for embedded video on the web and is viewable in most operation system, the final conversion was made in FLV format.1

The collection of files used for the case of the animated avatar contains words and phrases in SiGML animated format. Every sign and phrase edit process is complex and involves time and dedicated work. For some of the signs that are more complex the editing process may require several days of work as in the case of “carbonated drink”.

The SiGML format editing is more complex and requires more skills, time, effort and costs to be prepared. The video format requires skills easier to be learned and the results can be obtained easier, quicker and more inexpensive.

3.2. Application Design&Architecture - Video vs. Animated

As for the development architecture, the two alternatives differ in the Conversion/Translation module in the Application Level, respectively the Avatar Interface in the Avatar Interface Level concerning the display module of the avatar.
The Video Avatar will display recorded humans, the Animated Avatar will display an animated figure.\(^6\)

In terms of design, the two versions are different mainly regarding the data format, and this influences differences between the two architectures through the used software in order to run the applications. The collection of medical data is stored in a different format. In the animated avatar variant the collection of files contains phrases in SiGML animated format. In the video avatar variant format the data is stored as files in video format (FLV, MP4 or MOV).\(^6\)

Analyzing Figure 3 and Table 1 it follows that the architecture in the case of the animated avatar is more complex and complicated and may involve more expenses and specialists to work with. From the analysis resumed in Tables 1 and 2 results that the architecture in the case of the video technology is more common and easier to be designed and implemented.

3.3. Future Work

The research will continue with the implementation process for both cases and the study of the impact for the two types of applications tested on the impaired users on different groups on different age categories and from different regions in Romania. For now the impact study shows that the video avatar is better perceived.

References


[3] Rubén San-Segundo Hernández Translated by Robert Smith – “ Extract from Improvement and expansion of a system for translating text to Sign Language”- Chapter 5. Representation of the signs downloaded from

http://vhg.cmp.uea.ac.uk/tech/hamnosys/An%20intro%20to%20eSignEditor%20and%20HNS.pdf

