

Socialization Techniques in Virtual 3D Space

George Marian Hotca, Teodor Stefanut, Dorian Gorgan

Computer Science Department, Technical University of Cluj-Napoca

Str. G. Baritiu 28, 400027, Cluj-Napoca, Romania

hotca_marian@yahoo.com, {teodor.stefanut, dorian.gorgan}@cs.utcluj.ro

ABSTRACT

This paper presents the LivingRoom platform for online socialization within the 3D virtual space. The user is represented by an avatar that is moving around the virtual space and can communicate with other avatars. The audio and video communication among avatars is triggered by their interactions similarly with the human action within the real world.

Author Keywords

Virtual Space; Visual Techniques; User Interaction Techniques; Audio Communication; Video Communication; Avatar.

ACM Classification Keywords

H.5.2 User Interfaces, I.3.7 Three-Dimensional Graphics and Realism, I.6.8 Types of Simulation.

INTRODUCTION

The saying of Eric Hoffer “Creativity is the ability to introduce order into the randomness of nature” states the basic idea of the LivingRoom platform.

A socialization platform is an ensemble of web services that allow users to build a public profile within a limited system, to create a list of other users sharing a connection, and to visualize and navigate through the list of connections. LivingRoom is a socialization platform concerning mainly on providing users with the possibility to make new connections in the social life, even with unknown persons.

The main characteristic of the social platforms is the user profile [6], which is public and includes a list of friends (i.e. connections) that are users of the same system as well. In the classic social networks the user adds and enhances new personal information to his profile, by the answers fulfilled during the creation of the account. The LivingRoom platform does not impose the enrichment of the user’s profile with irrelevant information or a list of human connections. This is just for allowing a high level of information confidentiality.

The interaction with others [10] is natural, and actually is the most preferred way of using an application, motivated by competition, collaboration desire, organizing virtual meetings with others and self-exploration. Because in the LivingRoom platform the competition is not a priority, the last three motivations are mainly considered.

The collaboration [10] is the opposite way to competition through which we can interact with another user. The collaborative environments are preferred because they involve more users.

The games and the interactive applications allow us to join the group of friends even though this activity is accomplished within the virtual space. The interactive applications are ways through which two persons could consolidate or start a friendship.

The paper is structured as follows. The next section is an exploration of the literature, highlighting the related achievements. Then three sections detail the LivingRoom solution as a collaborative platform within the 3D virtual space. The last section concludes on the development and evaluation of the LivingRoom platform, and makes recommendation for the future research directions.

RELATED WORK

Instant Messaging Systems

Instant messaging systems (Instant Messaging or "chat") are systems that allow instantaneous exchange of text messages or multi-media content with one or more people via personal computers. The advantage of such a system is the fluency in messaging, with no significant delays, creating the impression of a real conversation but without direct interaction between the interlocutors.

Instant messaging systems are based on client-server architecture. The client (or the user) connects to a central server sending messages, which are interpreted by the server and then redirected to the destination user. Most systems involve the use of contact lists. Each user is forced to add a contact to the list, his request needs approval from the other contact, after which it is allowed to submit messages.

Web Chat Applications

A web chat system [9] allows real-time inter-user communication through an easy-to-use web interface. Online messaging gives the user instant access, the only requirement being the use of a web browser, which is an advantage because it will always access the latest version of the application, without requiring installation of updates.

Omegle [4] is the most representative web chat application. The application allows users to communicate without the need of creating an account. It randomly connects two users

in a one-to-one chat through which they can communicate anonymously. If it was originally created to be able to communicate only through text messages, subsequently a video mode was implemented which allows users to communicate using the device's microphone and webcam.

At the end of the conversation, users have the option to save their conversation history and to share it using a generated link. The conversation is not fully private as the user's personal information could be distributed over the network without his approval.

Chat Systems based on Virtual World

The virtual world known under the term MMOW (Massively Multiplayer Online World) [3] is a computer simulated environment, populated by users who are represented by avatars and can explore the virtual world simultaneously and independently, participate in activities or communicate among them.

The user accesses a simulated world that presents perceptual stimuli [13] that can manipulate elements of the world and thus offers a high degree of telepresence experience. Telepresence refers to a set of technologies that enable a person to feel that it would be physically present.

"3Dchat" application is the most representative in the niche of virtual worlds. It combines social networking concepts with online virtual world concepts, allowing users to interact with each other through avatars in a three-dimensional space. Although interaction is achieved largely through mouse and keyboard, the system allows voice communication within a certain area in the scene.

SOCIALIZATION PLATFORM IN VIRTUAL 3D SPACE

"LivingRoom", unlike other social platforms, aims to use the concept of communication cluster. The communication cluster is a group composed of all users who adhere to it. The application allows the user to change the cluster dynamically. If the cluster to which the user wants to be connected does not exist, it is created.

The purpose of the cluster is to group the users who wish to belong to particular geographical settlements (village, city, country, continent, etc.). Thus, a user can only interact with users in the same cluster, the interaction with users in other clusters is not allowed.

Each user is represented by an "avatar", a term with a wide range of meanings. In the Hindu religion, "the avatar is an incarnation of divine beings in the form of man or animal. Usually, a deity that exists at a higher level of reality chooses to manifest physically in the reality of people's lower level for specific purposes." [1].

In IT, the avatar is "a graphical representation of the user or the user's alter ego (character). They may have a three-dimensional form, as in games or virtual worlds, or a bi-

dimensional form, such as an icon for Internet forums or online communities "[1].

In the socialization platform, the user is represented by an avatar, which is a three-dimensional character. Each avatar is placed in the 3D space corresponding to the cluster he is associated with.

A 3D scene is a group of objects and entities in a well-defined virtual three-dimensional space [16], and can be characterized by three geometric parameters of the physical universe (excluding the fourth dimension - time). These parameters can be labeled by a combination of three elements chosen among the terms: length, width, height, depth and breadth.

The initial position of the character in the scene is obtained by placing it in a direction perpendicular to the vertical axis, and amplitude which will tend to zero (irrespective of the metric used) to the origin of the plane orthogonal to the vertical axis.

Camera in the field of gaming is the point of view of a character or the world view, in the context of the application scene. Most applications use one or more types of camera.

A view frustum [7] is defined as the volume of space that contains all the visible elements of interest in the three-dimensional scene. The X axis is oriented to the right and Y axis is pointing upwards. Z-axis direction depends on the graphics system used.

"LivingRoom" uses two types of perspective: "First Person" and "Tracking".

The "First Person" perspective [5] combines the camera with the character point of view, which means that the user sees through the eyes of the avatar (Figure 1). The user's viewing perspective corresponds to the viewing perspective of its avatar. This increases the realism of the application, even within the virtual environment.

The "Tracking" perspective [5] involves the tracking of the character along a predefined line in the 3D space. The camera can perform operations such as rotation, speed up and slow down the movement relative to the point that it aims, or change its position synchronously with the tracked character. The user can dynamically change the viewing direction by using specific control actions.

Animation [12] is the concept that visually brings to life the character in the application. Changing the position involves an animation of the avatar. Character animation can be seen as a state machine, and depending on the state of the character it will play an appropriate animation. Characters will have two types of animation: "idle" and "walking". The "idle" animation will show how the character is at rest, and the "walking" animation will occur when the character is in a motion state.

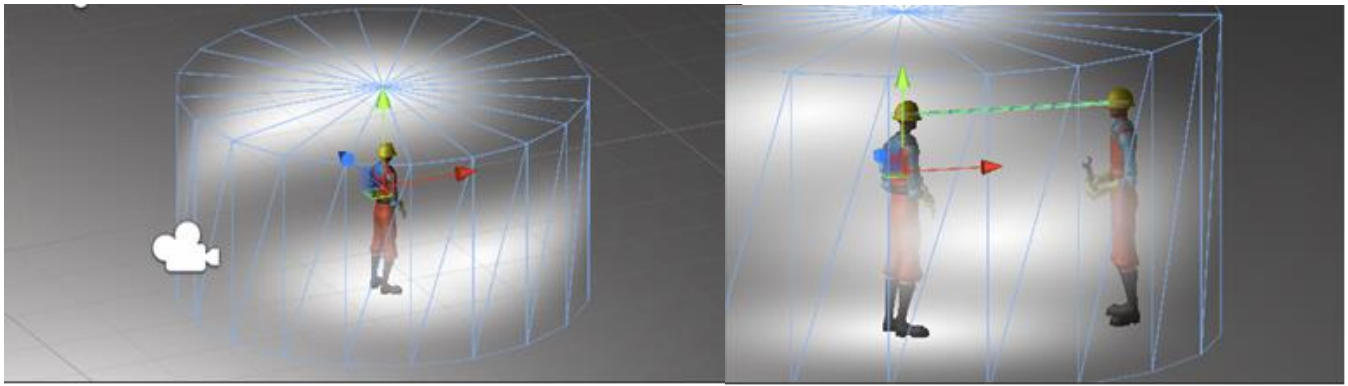


Figure 1. Audio and video communication techniques: (left) audio collision, and (right) video collision.

Most interactions between persons may take the form of verbal communication, such as talking, or non-verbal such as gestures or body language. The avatars are representations that take human characteristics in appearance, interaction manner and decision making, being used in a large number of applications that require interaction as natural interfaces.

One of the difficult problems is the development of a character that looks and behaves like a human, without another behavioral failure known as “Uncanny Valley” [14]. “Uncanny Valley” is a principle defined by Masahiro Mori in 1970 to describe robots and characters that resemble human creatures but do not behave like them. The feeling of frustration and mistrust is increasing

proportionally with the non-realism of the human aspect and behavior.

The interactions between characters are classified into: audio interaction and visual interaction (Figure 1). Audio interaction between avatars is achieved by fulfilling the condition that an avatar is situated at a predetermined distance to another avatar, similar to the real world. Conceptually, this can be seen as a sphere centered in the avatars position.

Visual interaction between avatars is achieved by fulfilling the condition that an avatar “looks” into the direction of the destination avatar’s eye. It means that the camera orientation is toward the eye of the other avatar, and the

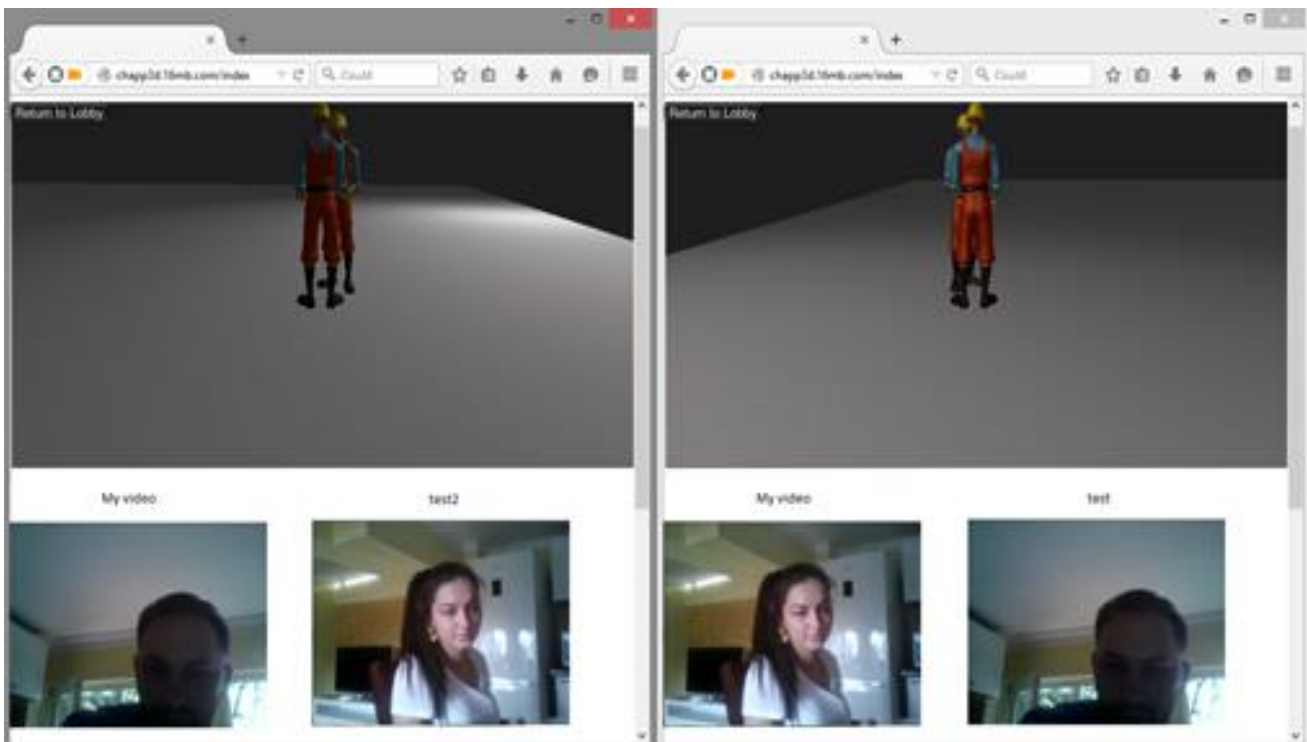


Figure 2. Video communication between two users.

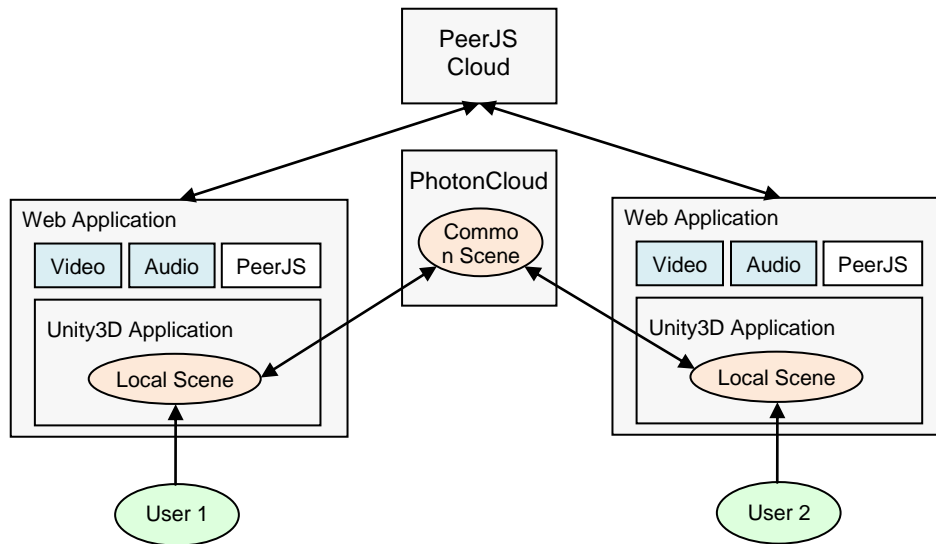


Figure 3. LivingRoom system architecture.

distance between them is equivalent or less than the threshold value. The Ray-Casting algorithm is used to identify the visual interaction.

The visual collision assumes that there is already an audio collision. The distance constraint between the two avatars has to be satisfied as well, which makes communication more naturally and similar to human communication mode.

AUDIO AND VIDEO COMMUNICATION TECHNIQUES

To perform actions within the system, the user must be authenticated and have specific permissions of his user type.

The main functionality of the system consists of the audio and video calls.

The scenario to start a video communication (Figure 2) is the following:

Start video-communication

1. The user controls his character in order to reach the audio collision with the other actor;
2. The user directs his visual perspective toward the eye of the other character;
3. The system detects intention to initiate a video call;
4. The system checks if the two users have the required resources;
5. The system asks for permission to use the hardware video resources;
6. The emitter sends a request for video transmission to the receiver;

7. The receiver accepts the video call;
8. The emitter responds with the video transmission;

- 8.1. One of the two users closes the communication

End video-communication

SYSTEM ARCHITECTURE

The application requires the use of two independent servers: Photon Cloud [8] and Cloud PeerJs [2]. PhotonCloud is used to synchronize characters (position, rotation) and all components attached to them in common scene (Figure 3). The scene is a common 3D place where all users of a cluster are represented by their characters. Each cluster has an individual scene.

This server works most of the time in emitter mode, sending to all Unity3D clients [11] within the same cluster, the positions of all characters. Otherwise, it works as a listener (receiver), respectively, expecting information about the position and rotation of a user's character. Then, it broadcasts them to all clients in that cluster.

PeerJs Cloud server is used for peer-level networking allowing the users to achieve the audio and video calls. Each user gets a peer connection after a successful authentication into the application. The user accesses the application through a Web browser. The application consists of two main components: the front-end application (client-side web scripting) and Unity3D application (plugin).

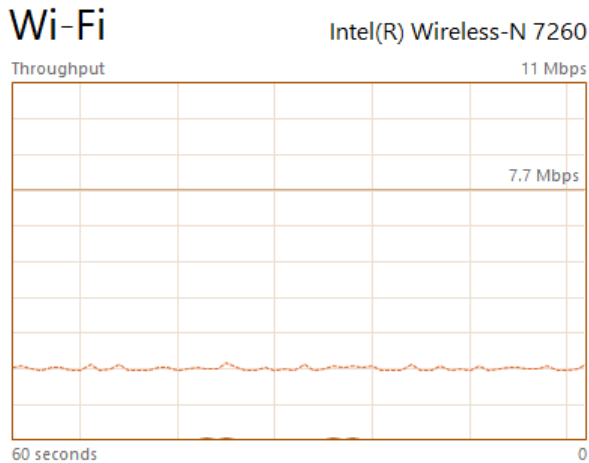


Figure 4. ChatRoulette bandwidth consumption

The Unity3D application renders the graphics - local scene and the character within the scene. Through this application, the user is able to control his character by a set of predefined controls. Conceptually, the local scene displayed for each user is a copy of the common virtual 3D scene. The user controls his character within the local scene, and the PhotonCloud server synchronizes the characters' position in all the other local scenes.

EVALUATION

Hardware resources

Running an instance of the application and having an audio transmission in progress requires about 100-120 Mb of RAM, with 25-30% more than the "SmallWorlds" application which is one of the most representative virtual chat software, but this surplus is justified by the advanced graphics that the "LivingRoom" has.

LivingRoom uses about 0.6-0.7% CPU because of the Unity3D application which uses hardware acceleration. This number is about double than the CPU power required by the "SmallWorlds" application which has a 2.5D graphic system, avatars represented as sprites and basic interaction, by far less interactive.

Bandwidth

"ChatRoulette" application (Figure 4) is similar to "Omegle" (presented in this paper) and for a video resolution of 640x480 (VGA) requires 500-550 Kbps bandwidth.

"LivingRoom" application (Figure 5) requires about 2Mbps of bandwidth for video transmission through WebRTC, the video resolution being 1280x1024 (~1.3 megapixels). The

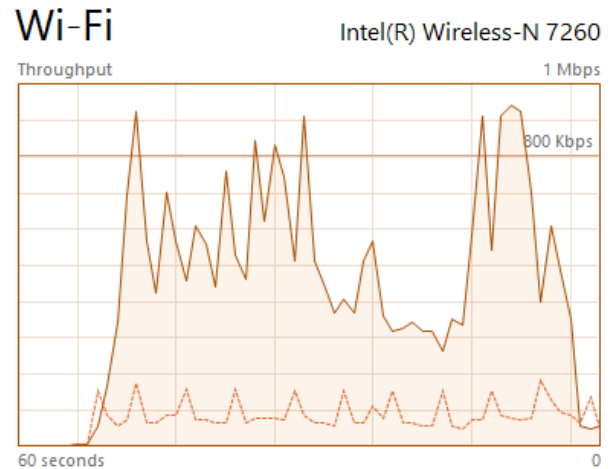


Figure 5. LivingRoom bandwidth consumption

result is very good if we consider that the bandwidth required includes PeerJs and PhotonCloud communication as well.

The transmission efficiency is about equal if we consider that VGA resolution is about 4 times smaller than a resolution of 1.3 mpx specific to the HD web camera used.

CONCLUSIONS

One of the most difficult tasks was synchronizing the characters in the local scenes, since each user specifies the commands in an asynchronous mode. The experimental results have proved the effectiveness of collision detection in Unity, considering that these computations are executed every frame, and the application runs up to the value of 50-60 fps. This means that all the computations of the collision run 50-60 times per second, so every 1.5 milliseconds.

The audio and video transmissions are effective, clear and comparable with the ones offered by applications such as Skype (Figure 2). The experimental results are satisfactory since the free versions of the PeerJs Cloud and PhotonCloud servers offer some limitations of the number of connections (e.g. 50 peer connections, and respective 20 concurrent users), and lower bandwidth.

In conclusion, the LivingRoom platform meets the objectives of a social platform by inter-connecting users in groups called clusters, and by implementing a multimedia communication system. The users initiate audio and video communications by interaction techniques in natural manner.

FUTURE DEVELOPMENT AND RESEARCH DIRECTIONS

Google Chrome announced the withdrawal of support for NPAPIs (Netscape Plugin Application Programming

Interface), which means that Unity Web Player applications will no longer be supported.

The main objective is to create a cross-browser and cross-platform application, in order to solve this technology issue. One possible solution is to use WebGL, a graphic programming interface for the Web (OpenGL for the Web), as a rendering platform due to its compatibility in all HTML5 browsers.

Furthermore, the social interaction between the participants can be further developed through additional gestures and interaction techniques that can be implemented.

ACKNOWLEDGMENTS

The research has been carried out in the Computer Graphics and Interactive Systems Laboratory of the Computer Science Department, in the Technical University of Cluj-Napoca.

REFERENCES

1. Avatar in computer applications, [http://en.wikipedia.org/wiki/Avatar_\(computing\)](http://en.wikipedia.org/wiki/Avatar_(computing))
2. Badass Javascript, "PeerJs : A peer to peer networking library in Javascript using WebRTC", 2013, <http://badassjs.com/post/43090030238/peerjs-a-peer-to-peer-networking-library-in>
3. Bartle Richard, "Designing Virtual Worlds", New Riders, 2003
4. BeWebSmart.com, "What is Omegle? Is Omegle okay for kids?", 2015, <http://www.bewebmart.com/internet-safety/what-is-omegle-is-it-okay-for-kids/>
5. Camera Comes First [Game Design], <http://www.whatgamesare.com/2011/10/camera-comes-first-game-design.html>
6. Danah M. Boyd, Nicole B. Ellison, "Social Network Sites: Definition, History and Scholarship", *Journal of Computer-Mediated Communication*, pp.210-215
7. Eric Lengyel, "Mathematics for 3D Game Programming and Computer Graphics" Third Edition, Course Technology, a part of Cengage Learning, 2012
8. Exit Games, "Photon Unity Networking", <https://www.exitgames.com/en/PUN>.
9. Gao Kevin, "5 Best Practices for Increasing Sales with Live Chat Software", Salesforce, 2014.
10. Jesse Schell, "The Art of Game Design – A Book of Lenses" Second Edition, CRC Press
11. Jon Brodtkin, "How Unity3D Became a Game-Development Beast", 2013, <http://insights.dice.com/2013/06/03/how-unity3d-become-a-game-development-beast/>
12. Kyle D'Aoust, "Unity Game Development Scripting", Packt Publishing, 2014
13. Marvin Minsky, "Telepresence", Omni, 1980
14. Patrick G. Kenny, Thomas D. Parsons, Albert A. Rizzo, "Human Computer Interaction in Virtual Standardized Patient Systems", Institute for Creative Technologies, University of Southern California
15. Stay In Touch With Yahoo Pager, 1998, <https://web.archive.org/web/20101220214553/http://yho.client.shareholder.com/press/releasedetail.cfm?ReleaseID=173501>
16. Three-dimensional space coordinate system, http://en.wikipedia.org/wiki/Three-dimensional_space
17. Yahoo! Messenger, http://ro.wikipedia.org/wiki/Yahoo!_Messenger