A Framework for Human-like Behavior in an Immersive Virtual World

Fons Kuijk
Distributed and Interactive Systems
CWI
Amsterdam, Netherlands
e-mail: Fons.Kuijk@cwi.nl

Sigurd Van Broeck
Bell Labs
Alcatel-Lucent Bell NV
Antwerp, Belgium
e-mail: sigurd.van_broeck@alcatel-lucent.com

Claude Dareau
Blitz Games Studios
Leamington Spa, United Kingdom
e-mail:
cdareau@blitzgamesstudios.com

Brian Ravenet, Magalie Ochs
CNRS-LTCI
Telecom-Paristech, Paris, France
e-mail:
brian.ravenet@telecom-paristech.fr
magalie.ochs@telecom-paristech.fr

Konstantinos Apostolakis,
Petros Daras
Information Technologies Institute
Centre for Research and Technology - Hellas
Thessaloniki, Greece
e-mail: kapostol@iti.gr, daras@iti.gr

David Monaghan, Noel E O'Connor
CLARITY: Centre for Sensor Web Technologies
Dublin City University
Dublin, Ireland
e-mail: david.monaghan@dcu.ie,
Noel.OConnor@dcu.ie

Julie Wall, Ebroul Izquierdo
Multimedia and Vision Research Group
Queen Mary, University of London
London, United Kingdom
e-mail: julie.wall@eecs.qmul.ac.uk,
ebroul.izquierdo@eecs.qmul.ac.uk

Abstract—**Just as readers feel immersed when the story-line adheres to their experiences, users will more easily feel immersed in a virtual environment if the behavior of the characters in that environment adheres to their expectations, based on their life-long observations in the real world. This paper introduces a framework that allows authors to establish natural, human-like behavior, physical interaction and emotional engagement of characters living in a virtual environment. Represented by realistic virtual characters, this framework allows people to feel immersed in an Internet based virtual world in which they can meet and share experiences in a natural way as they can meet and share experiences in real life. Rather than just being visualized in a 3D space, the virtual characters (autonomous agents as well as avatars representing users) in the immersive environment facilitate social interaction and multi-party collaboration, mixing virtual with real.**

Keywords—autonomous agent; avatar; collaboration; emotional engagement; humanoid; immersive; interaction; semi-autonomous avatar; tele-immersion.

I. INTRODUCTION

Teleconferencing systems enable participants in different locations to share a common experience, where specially designed tables and real-time high-definition video increase the feeling of proximity. Current-generation teleconferencing systems allow participants to talk to each other as if they were in the same location. Their shortcoming is that participants cannot collaboratively perform a task together: they remain captives in a 2D screen projection. The next challenge in tele-presence is tele-immersion, which will enable individuals who may be geographically apart to interact naturally with each other in a shared 3D synthesized environment. Where teleconferencing allows participants to share a common space, tele-immersion allows them to share an activity.

REVERIE[1] is a project that aims at such an immersive environment: an ambient, content-centric Internet-based environment where people can work, meet, participate in live events, socialize and share experiences as they do in real life, but without time, space and affordability limitations. Within the context of this project, the framework presented in this paper serves to model the mind of the virtual human as well as analyzing the captured data from the real human and modeling the interaction between virtual and real.

The paper is organized as follows. Section II describes a use case for the framework, Section III reviews related work on immersive systems and virtual human-like characters. Section IV presents our framework for human-like characters in an immersive virtual environment. In Section **Error! Reference source not found.** we draw conclusions and describe future work.

II. USE CASE

A teacher may take his class on a trip for an object lesson on the architecture of a historical building. Rather than having this trip in the real world, the teacher can give this lesson in the virtual world of REVERIE; he and the students can stay in the classroom, sitting behind their desks. They don't have to make a lengthy, costly and unsafe trip to the building at hand, which at that time may even be crowded with tourists. Each of them has a PC with a capturing device at their disposal, which can be a simple webcam or any RGB camera. Using their capturing device they can create a look-alike avatar of themselves with just a few mouse clicks. Being represented in the virtual world by their personalized avatar, the students and the teacher see on their display what their representation in the virtual world is supposed to see, the gaze of the avatar corresponds to whom or what the student is looking at on the screen. The teacher introduces his class to the virtual environment. In that

---

[1] http://www.reveriefp7.eu

environment a virtual guide, a fully autonomous agent, is at their disposal to take them on a tour through the building. This autonomous agent is part of the REVERIE system and becomes active when the users enter the virtual environment. In a "follow-me" mode, this agent takes care of navigation of the whole group through the environment, explaining what is to be seen and at times consulting the teacher on the progress, reacting upon the teacher's spoken instructions. By raising a hand (noticeable in the virtual world because it is mimicked by their avatar), students get the opportunity to pose a question. Besides this type of gestures, also the facial expression, the gaze and emotional behavior of the participants are analyzed and reflect in the appearance of their avatar. This may give rise for the agent to adapt his presentation, or for the teacher to instruct a student to pay attention. All interactions take place in the virtual world. After the tour the students can have their avatar wander around in the environment, assisted by embedded path planning and collision detection software. In this way they can inspect objects or elements of the building to be able to answer the list of questions the teacher has prepared. In this virtual world they can have private discussions with the teacher or fellow students they meet, their avatar showing their frustration, enthusiasm or indifference, just as in the real world.

III. RELATED WORK

3D tele-immersion (3DTI) is a new video medium that creates 3D photorealistic, immersive, and interactive collaboration among geographically dispersed users. Over the last few years, 3DTI has shown great potential in a wide range of applications such as distance learning of physical activities, collaborative art performance, fire fighter training, and medical consultation. So far, only few commercial systems exist with not so advanced capabilities. Among them Immersivemedia[2] offers a camera system and packaged solutions, along with a variety of production and post-production services, such as planning, capturing, post-production and editing. Moreover, the company Immersion[3] supports interaction using only haptic devices dedicated to medical applications. Their virtual reality medical simulators combine sight, sound, and touch for highly effective training in minimally invasive procedures. The company Total Immersion[4] provides augmented reality solutions, based on their platform "D'Fusion". Their clients experienced augmented reality in Digital Marketing, Retail, Consumer Products, E-Commerce and Entertainment. Finally, Immersion France[5] produces new hardware solutions for interaction and visualization in virtual environments. The aforementioned hi-tech companies do support some of the provisioned REVERIE services, yet in an immature way since they don't support multi-party collaboration (the huge demand of networking and computing resources has restricted these systems to work with only two sites), without interactivity (except only using haptic devices) and they don't support the mixture of virtual with real. Nevertheless, the dynamism of the offered services is obvious and clearly demonstrates the potential of these services for a new market.

Expressive human-like virtual characters facilitate social interaction within immersive virtual environments, by triggering users' innate empathetic responses to the human form and behavior [1]. Research has shown that users can respond to such virtual 3D humans as if they were real, despite being aware that the characters are synthetic [2], [3]. However, a lack of expression in virtual human-like characters can also negatively impact communication within collaborative virtual environments [4], [5]. Moreover, many challenges exist in realizing believable virtual human-like characters that avoid falling into the 'uncanny valley' trap of raising expectations due to human-like appearance and associated behavior, only to fall short as users notice imperfections in the simulation [6], [7], [8]. Since emotion pervades human behavior, human perception is very sensitive to the behaviors and expressions of emotion in others. As a result, when virtual characters resemble humans, expectations are that their behavior will correspondingly be human-like, and any discrepancies can be disturbing [9], [1]. Traditional character animators rely on a variety of techniques, animation principles, to ensure readability of emotions. These include the careful crafting of poses so that they are readable in silhouette as viewed by a predefined camera view; exaggeration of factors such as timing, movement and expression to ensure the audience notices the targeted action/emotion; careful staging of the character and all other elements within the scene, in addition to deliberate toning down of any other possible movements within the scene which could draw the audiences' eyes away from the intended action [10], [11]. However, interactive immersive collaborative environments add extra challenges, as they do not provide the same luxury of a predefined linear story, pre-set camera view, and predetermined interactions with other characters that traditional character animators can exploit. The components of our framework dealing with emotion and behavior have to seek, identify, and process influencing factors in real-time to ultimately realize plausible behavior.

IV. A FRAMEWORK FOR HUMAN-LIKE BEHAVIOR IN AN IMMERSIVE VIRTUAL WORLD

The framework we present in this paper is developed in the context of REVERIE. The objectives of that project are to develop an advanced system for immersive media capturing, representation, encoding and semi-automated collaborative content production, as well as transmission and adaptation to heterogeneous displays. It may serve as a key instrument to push social networking towards immersive collaborative environments that support realistic inter-personal communication. It will exploit technologies and tools to enable end-to-end processing and efficient distribution of 3D, immersive and interactive media over the Internet. Being part of that work, our framework deals with autonomous behavior, physical interaction and emotional engagement regarding the virtual humanoid characters. The framework relies on other parts of the REVERIE system to take care of 3D data acquisition and processing (capturing of the human user's gestures and emotional expressions), sound processing, networking, and real-time rendering.
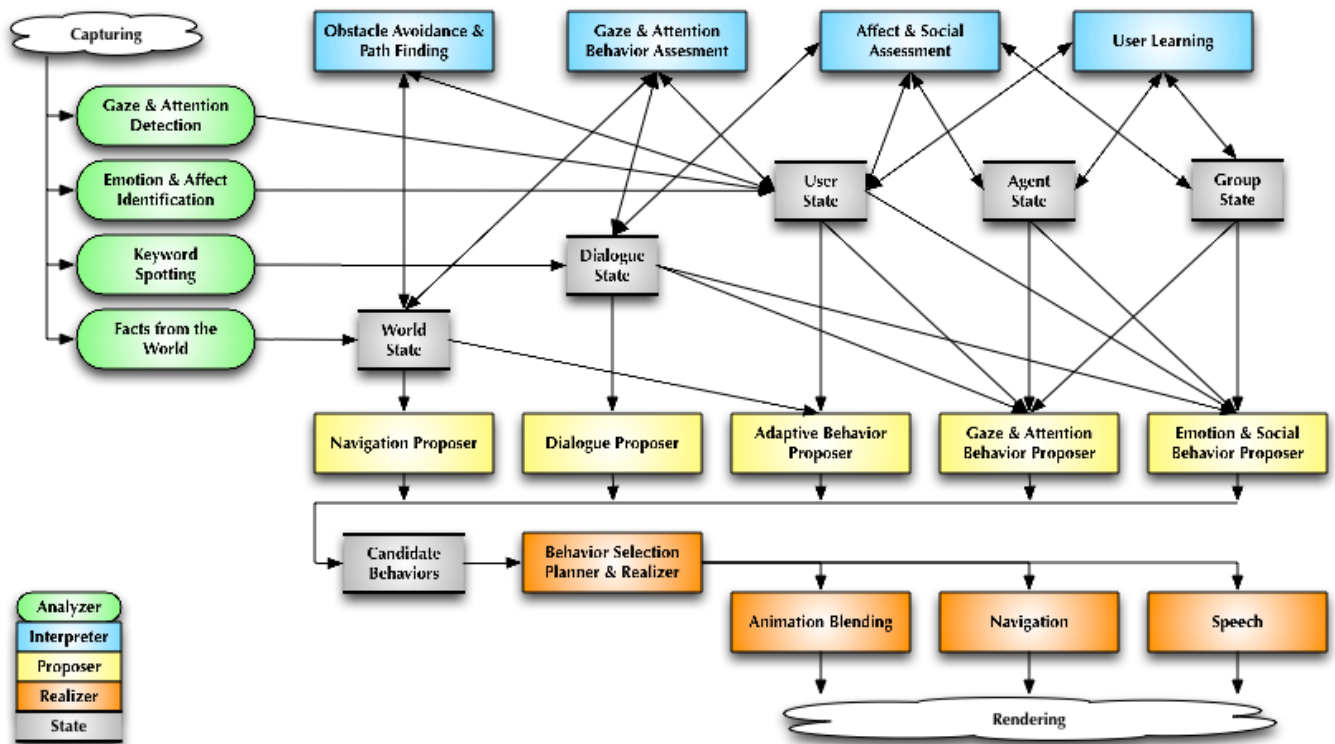
Figure 1. Functional model of the Framework, consisting of Analyzers, Interpreters, Proposers, Realizers and State modules.

Virtual humanoid characters living in REVERIE's virtual world come with different levels of autonomy. For the humanoids representing the teacher and the students we use the term *avatar*. The teacher and the students are sitting behind a desk, so only the upper part of their body can be captured. The REVERIE system adds control of un-captured features (such as their legs) to still obtain plausible behavior. In this way they are able to freely wander around in the 3D environment. The user specifies a destination; the system automatically blends in instructions to make the avatar walk towards that destination, avoiding collisions with static and dynamic objects. In that situation we use the specific term *semi-autonomous⁶ avatar*. Humanoids not representing a user, as the guide mentioned in the use case, are fully autonomous; we call these *autonomous agent,* or simply *agent*.

The objectives of the framework are:

- To achieve an immersive interactive experience through the coherent and natural behavior of (semi-) autonomous virtual humans.

- To allow user driven and autonomous virtual characters to plausibly interact in virtual spaces which represent real life counterparts.

- To enable autonomous virtual beings to engage in understandable social interactions with other users in an emotional and expressive manner.

- To develop techniques for interaction between a virtual human and user with responses based on and affected by the virtual human's (emotional) state.

Agents that interact with users in the real world and show emotional behavior are not uncommon [12], but in current implementations usually provide limited modalities and operate in controlled conditions. The challenge here is to combine information on all participants, the environment and the social context, and still come up with natural behavior of the agents and avatars. This adds up to a complexity that to our knowledge has not been dealt with before.

A. Functional Model of the Framework

The architecture of the framework is shown in Figure 1. Similar to the architecture of SEMAINE [12], it comprises components of the classes: *Analyzers*, *Interpreters*, *State*, *Proposers*, and *Realizers*. The different instantiations of each class of component reflect the multi-modality required for REVERIE, each instantiation dealing with a specific aspect of human behavior. To ensure modularity and flexibility in integrating modules we rely on standard markup languages. Inputs and outputs of modules are represented with the XML languages: EmotionML [13], FML [14], BML [15] and MPEG–4's FAPs and BAPs [16].

The following subsections describe the components of this framework.

⁶ In our work *autonomous* implies fully computer controlled and, logically, *semi-autonomous* is a mixture of user and computer controlled.

1) Analyzers

Analyzer modules receive input from external data capturing components and perform basic processing on the incoming data, ranging from transformation of data up to gesture recognition; note: it does not include interpretation. Output is a representation of a current state. In our framework we identify four Analyzer modules: *Gaze & Attention*, *Emotion & Affect*, *Keyword Spotting*, and *Facts from the World*.

a) Gaze & Attention

The Gaze and Attention module deals with gaze and attention identification and eye tracking of the students in the real world. It analyzes their point of interest to control their avatar's gaze and serves as an indicator of being attentive.

b) Emotion & Affect

The emotion and affect recognition module's purpose is twofold: the students will be able to control the expressive signals of their avatar's face via facial expressions while also utilizing the underlying emotional information to infer the affective quality of the perceived stimulus. Based on live video of a student, an emotion and core affect state description is produced.

c) Keyword Spotting

To identify input text from spoken words of the users, a phoneme-based keyword spotter is used. In our use case this module is used for the guide consulting the teacher.

d) Facts from the World

This module receives and analyzes dynamic data of the virtual world such as: position and geometry of the students and moving objects, lighting or atmospheric conditions and other facts that influence the experience of the students.

2) Interpreter Modules

The Interpreter modules are responsible for high-level interpretation of input produced by the analyzers in the context of the states (user, dialog, agent, group and world) that are relevant for the specific type of interpretation and subsequently update the appropriate states. In other words, the Interpreter modules add intelligence; they are responsible for how input is perceived subjectively, rather than objectively. The context makes a difference. The framework has four Interpreter modules: *Gaze & Attention Behavior Assessment*, *User Learning*, *Affect & Social Module*, and *Obstacle Avoidance & Path Finding*

a) Gaze & Attention Behavior Assessment

The gaze and attention behavior assessment module will decide on the target location at which to shift character gaze, based on the character's emotional state following perception. This behavioral cue does not only provide characters with the ability to observe their surrounding environments in order to perceive and localize possible threats and opportunities, but also offers an insight to human users who can identify, and possibly relate to the autonomous characters behaving the way they do because a specific stimulus caused them to feel a certain way. As an example, if some students in the back of the group are making jokes, disturbing the guided tour, others may turn their head and either frown or smile.

b) User Learning

This module observes avatars as they interact with their 3D virtual environment and other characters. The indefinite quantity and complexity of this data will be converted into a set of fuzzy logic rules that describes the expected behaviors of the autonomous agent rather than simply try to mimic or copy the human-controlled character in every way.

c) Affect & Social Module

This module computes the emotions of an autonomous agent elicited by events in the environment, considering the influence of the agent's personality and social relations. The modularity of the compute model enables the designer to choose which modules to activate. The designer can define the initial social attitude of the autonomous agent (being hostile or friendly) that impacts the mood, the elicited emotions and the actions of the agent in the environment. Users are more likely to trust an interface when the autonomous agents express appropriate social and emotional responses during interactions with virtual characters. For instance, the agent will avoid close distance if it wants to express dislike to a participant. Two dimensions will represent the social relations: dominance and liking. The social relations will be updated during the interaction given the non-verbal behavior of the participants. For instance, a lean-forward posture and sustained gaze could be interpreted as a dominant attitude.

d) Obstacle Avoidance & Path Finding

This component is responsible for providing collision free paths for autonomous agents and semi-autonomous avatars to allow them to navigate through the 3D virtual environment. It collects information from the World state, in particular the location, of all moving avatars, agents and objects.

3) State Modules

State modules represent the current state of the interpreted avatar and agent related entities. The framework has five State modules: *User State*, *Agent State*, *Group State*, *Dialogue State*, and *World State*.

a) User State

The user state is a set of elements describing the current state of the user, including gaze and attention and emotion and affect, as has been observed by different input modalities such as camera, microphone, keyboard, etc. The framework serves to control an avatar that represents the user in the virtual world, but the avatar state is not necessarily an exact copy of the user state.

b) Agent State

The agent state is composed of several elements describing the internal state of the autonomous agent or avatar. Connecting to the emotional and social module, the agent state integrates the current emotions of the agent, its social attitude, and its social relations with other virtual characters. Moreover, in the state of the agent its social personal goals (for instance "assume a dominant position over other participants") and its social role in the environment are defined. Such elements are used in the emotional and social proposer to select the appropriate non-verbal behavior. The role of this module as a link between the User Learning and Adaptive Behavior

Proposer modules is to relay the set of fuzzy logic rules which describe the way in which the avatar interacts with its environment and with other virtual characters. These fuzzy rules will describe the physical actions and interactions of the (semi-autonomous) avatar as it interacts with the virtual world and the other virtual characters.

c) Group State

The group state enables the virtual character to be aware of the group and its members. Each group in which the autonomous agent is involved is characterized by: an individual level that defines what the autonomous agent knows about the individual characteristics of each member of the group (name, social relation, characteristics, etc.); a group level that characterizes the structure of the group (members, members' roles, etc.); an interaction level that categorizes the possible actions in the group [17].

d) Dialogue State

The dialog state is composed of: the task stack, containing the tasks of the Hierarchical Task Networks (HTNs) being executed; the current dialogue tree and the list of the past utterances. With these elements, the dialog system is able to track down the current point in the tree, and compute which sentences could be proposed to the user, with respect to the current task and dialogue.

e) World State

The World State relates to the state of the *virtual* world created in the computer as distinct from the real world state where the users reside. The virtual world state can be defined as all the information existing within the virtual environment at any moment in time. Key information within the virtual world is the position and orientation of the human-driven avatars, the semi-autonomous avatars and autonomous agents. Orientation in this context is a generic term encompassing the gaze and attention behavior of the avatars. One of the major contributing factors to the AI (Artificial Intelligence) decisions that dictate the autonomous agent behavior will be based on the individual and collective positioning of the human-driven avatars in the virtual world. These AI decisions will inform both autonomous agent navigation throughout the virtual environment, i.e. avoidance and path finding, and autonomous agent behavior, gaze and attention decisions.

4) Action Proposer Modules

The Action Proposer modules produce the candidate actions based on the current state of relevant aspects of the avatar and its context. Actions produced are indicated as being *candidate* actions because different modules may propose conflicting actions that have to be resolved in the next step of the framework. The framework has five Action Proposer modules: *Dialogue Proposer*, *Gaze & Attention Behavior Proposer*, *Adaptive Behavior Proposer*, *Emotion & Social Behavior Proposer*, and *Navigation Proposer*.

Figure 2.   Same input, but different social relation leads to different gestures.

a) Dialogue Proposer

The dialog system, Disco4games [18], is a task-based dialogue manager, which is based on collaborative discourse theory. It combines traditional dialogue trees and HTNs to partially automate dialogue authoring, and to improve dialogue structure reuse. HTNs are used to represent the high-level task structure of the dialogue, while sub-dialogue trees are used for more application-specific dialogues. Formatting rules can be used to add color to the automatically generated utterances.

b) Gaze & Attention Behavior Proposer

This module suggests something for the character to look at. Internally it needs to assess possible points of interest and then choose which one it should look at. As part of this decision the module takes into account the current emotional state, the influence of the individual points of interest on that state, and also what points of interest the character has been looking at in the recent past, and their relative importance.

c) Adaptive Behavior Proposer

This module processes data based on a set of the previous behavioral decisions over a certain time period that the (semi-autonomous) avatar performed to determine a current behavior for the autonomous agent. The learning algorithm will also be presented with information about the world and navigation via the World State module. The output will be a selected behavior for the autonomous agent to implement, i.e. move towards and interact with another virtual character; remain distant from another character to be avoided; move towards and inspect an object within the world; take instruction from and listen to the virtual teacher, etc.

d) Emotion & Social Behavior Proposer

The emotional and social proposer module will allow the autonomous agent to choose the non-verbal behavior to display during its interaction with other virtual characters. The non-verbal behavior will be selected based on the emotions of the autonomous agent, its social relation and the social relations it thinks the other has toward itself. The role of the agent (for instance "supporter" or "guide") will determine the initial social relation of the autonomous agent and its personal goals. A dominant agent will for instance do wider gestures. Figure 2. illustrates an example of the different gestures generated by the model. This social behavior model is the one described in [19] and implemented in the virtual agent Greta [20]. The same input was sent to the model, an intention to inform, but the social relation was different. In the left image the agent is more dominant, while in the right image, the agent is more

submissive. The emotion proposition part of the Emotion and Social Behavior Proposer therefore will also be to choose the parameters and assets that are to be used by the animation-blending module to be able to correctly render the character. Given that emotions can be expressed through a number of factors including facial expressions, plus body posture, the sorts of assets and parameters to be passed along are likely to be: desired facial animation state; desired body pose animation state; head and eye movement properties; desired skin surface properties. For example, if the input indicates that the avatar should react sadly to an event, the emotional proposer may propose depicting sadness by proposing that the eyes and head look down, corners of the mouth lower to form a frown, and shoulders and body posture drop.

e) Navigation Proposer

This component is responsible for proposing collision free paths for autonomous agents and semi-autonomous avatars to allow them to navigate through the 3D virtual environment. Group behavior and emotional state are new elements that given time may be introduced as innovative aspects in the planning algorithms.

5) Action Realization Modules

Action Realization modules are responsible for checking consistency and have to resolve conflicts in proposed actions. Preferred actions are selected and turned into a behavior plan. The Navigation, Animation Blending and Speech components produce the data needed for rendering. The framework has four Action Realization modules: *Behavior Selection Planner and Realizer*, *Navigation*, *Animation Blending*, and *Speech*.

a) Behavior Selection Planner and Realizer

The Behavior Selection Planner and Realizer module takes the proposed candidate behaviors and decides which of them to action; ensuring that the sequence is in order will make the character more believable. The Animation Blending module has an additional complication in that it is dependent on other selection and realizer modules. The output of the selection planner goes to the renderers via the Navigation, Animation Blending and Speech modules. As this module will receive input from all of the proposer modules, i.e., Dialogue Proposer, Gaze and Attention Behavior Proposer, Adaptive Behavior Proposer, Emotion and Social Behavior Proposer and Navigation Proposer, a form of conflict resolution will have to occur. This will assign different levels of priority to the different proposers, e.g. before a virtual character begins conversing with another character, they should first gaze at that other character; therefore the gaze proposer would have a higher level of temporal priority to the dialogue proposer in that instance. Another simple example of conflict resolution could involve the selection between the avatar running and speaking, as both probably wouldn't occur at the same time.

b) Navigation

Based on the current location of the avatar and the proposed gaze, the Navigation module provides the field of view parameters needed to produce a view on the environment from the avatars point of view. It is up to the rendering modules to select between this option and possible other camera settings.

c) Animation Blending

We see the final animation being a blend of various key-framed animations, static poses, procedural animation, IK (Inverse Kinematics) solvers to join the procedural to the key-framed animation, lip sync, etc. There are two aspects to this work: the purely technical aspect of making these systems work, and the more subjective aspect of making them work together in a way that sells the emotional state to the viewer.

d) Speech

In order to enable the autonomous agent to speak, an open-source, multilingual Text-to-Speech Synthesis system called MARY will be used. OpenMary [7] contains a number of male and female voices that can be used. Prosody is also modeled: the location of pitch accents and prosodic phrase boundaries is determined by corpus-derived rules, and the actual tones are then adapted depending on the sentence type (declarative, exclamatory, etc.).

V. CONCLUSIONS AND FUTURE WORK

We have seen in the use case (Section II) how REVERIE allows a teacher to take his class on a virtual trip to a virtual environment. In this paper we were not concerned with the accuracy of the rendering of avatars and agents; the framework serves to bring in accuracy of their behavior. As was mentioned in Section III, studies have shown that this is the most essential element of feeling immersed. In that sense, having a discussion with virtual fellow-students that look you in the eye outperforms the experience of using tele-conferencing systems that display participants with their eyes fixed on a screen.

Future work will expand and outline the proficiencies of the individual modules of the framework, e.g. the capability of the system to learn from user behavior can signal uncommon or undesirable behavior of the real user, and serves to have an agent replace the user when he is unavailable for a while. The REVERIE system will be capable of high quality full-body capturing of users. Their representation, according to the definition of an avatar, is a one-to one replication of the user. The system is designed to be able to blend these *replicants* seamlessly into the virtual environment populated with agents and avatars.

REFERENCES

[1] Vinayagamoorthy, V., Gillies, M., Steed, A., Tanguy, E., Pan, X., Loscos, C. and Slater, M. *Building expression into virtual characters*. In: Eurographics Conference State of the Art Report, Vienna, 2006

[2] Bailenson, J.N., Blascovich, J., Beall, A.C. and Loomis, J.M. *Interpersonal distance in immersive virtual environments*. Personality and Social Psychology Bulletin, 2003. 29(7): p. 819-833

[3] Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., Barker, C., Pistrang, N. and Sanchez-Vives, M.V. *A virtual reprise of the Stanley Milgram obedience experiments*. PLoS One, 2006. 1(1): p. 1-10

[4] Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A. and Sasse, M.A. *The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment*. Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 529-536. ACM, 2003

[7] http://mary.dfki.de/

[5] Vinayagamoorthy, V., Garau, M., Steed, A. and Slater, M. *An eye gaze model for dyadic interaction in an immersive virtual environment: Practice and experience*. In Computer Graphics Forum, vol. 23, no. 1, pp. 1-11. Blackwell Publishing Ltd., 2004

[6] Mori, M., *The uncanny valley* (Originally in Japanese: *Bukimi no tani*.) Energy, 1970. 7(4): p. 33-35

[7] Vinayagamoorthy, V., Steed, A. and Slater, M. *Building characters: Lessons drawn from virtual environments*. Proceedings of Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop, pp. 119-126. 2005

[8] Tinwell, A., Grimshaw, M., Nabi, D.A. and Williams, A. *Facial expression of emotion and perception of the Uncanny Valley in virtual characters*. Computers in Human Behaviour, 2011. 27(2): p. 741-749

[9] Gratch, J. and Marsella, S., *A domain-independent framework for modeling emotion*. Journal of Cognitive Systems Research, 5(4):269–306, 2004

[10] Thomas, F., Johnston, O. and Rawls, W. Disney animation: *The illusion of life*. Vol. 4. 1981: Abbeville Press New York

[11] Lasseter, J., *Principles of traditional animation applied to 3D computer animation*. ACM SIGGRAPH Computer Graphics, 1987. 21(4): p. 35-44

[12] Schröder, M. *"The SEMAINE API: Towards a Standards-Based Framework for Building Emotion-Oriented Systems"*. Advances in Human-Computer Interaction, 2010

[13] Schröder, M., Baggia, P., Burkhardt, F., Pelachaud, C., Peter, C., and Zovato, E. *EmotionML - an upcoming standard for representing emotions and related states*. Proc. Affective Computing and Intelligent Interaction. Memphis, TN, USA, 2011

[14] Heylen, D., Kopp, S., Marsella, S., Pelachaud, C., Vilhjalmsson, H. *The Next Step Towards a Functional Markup Language*, Intelligent Virtual Agents, IVA'08, Tokyo, p. 270-280, September 2008

[15] Vilhjalmsson, H., Cantelmo, N., Cassell, J., Chafai, N.E., Kipp, M., Kopp, S., Mancini, M., Marsella, S., Marshall, A.N., Pelachaud, C., Ruttkay, Z., Thorisson, K.R., van Welbergen, H. and van der Werf, R. *The Behaviour Markup Language: Recent Developments and Challenges*, In: Intelligent Virtual Agents, IVA'07, Paris, p. 99-111, September 2007

[16] ISO/IEC 14496 *MPEG-4 International Standard*, Moving Picture Experts Group, www.cselt.it/mpeg

[17] Prada, R. and Paiva, A. *Believable groups of synthetic character*s. Proceedings of the Fourth International Joint Conference on Autonomous agents and multiagent systems, AAMAS '05, p. 37-43, 2005

[18] Rich, C. and Sidner, C.L. *Using Collaborative Discourse Theory to Partially Automate Dialogue Tree Authoring*. In Intelligent Virtual Agents, pp. 327-340. Springer Berlin Heidelberg, 2012

[19] Ravenet, B., Ochs, M. and Pelachaud, C. *A computational model of social attitude effects on the nonverbal behavior for a relational agent*, In Proceedings of Workshop Affect Compagnon Artificiel Interaction (WACAI), November 2012

[20] Niewiadomski, R., Bevacqua, E., Mancini, M. and Pelachaud, C. *Greta: an interactive expressive eca system*. In Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems- Volume 2, International Foundation for Autonomous Agents and Multiagent Systems, AAMAS '09, pages 1399–1400, Richland, SC, 2009