Fourth Annual Online Public Report
SALERO Deliverable 10.5.7
Fourth Annual Online Public Report

SALERO Deliverable D10.5.7

SALERO identifier: SALERO-D10.5.7-JRS-PublicAnnualReport-v06.doc
Deliverable number: D10.5.7
Author(s) and company: G. Kienast, G. Thallinger (JRS); R. Fach (GVG);
S. Miarnau Freixes (AM); O. Mayor (UPF);
T. Bürger (LFUI); S. Cheung (DTS); T. Stolt (TAIK);
R. Villa (UG); M. Romeo (FBM-UPF);
C. Goodman (PGP); M. Matthews (BLITZ);
J. Adell (URL)

Work package / task: WP10
Document status: Final
Confidentiality: Public

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Reason of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2009-10-29</td>
<td>Document created, call for input from partners</td>
</tr>
<tr>
<td>2</td>
<td>2009-11-09</td>
<td>First input collected from partners</td>
</tr>
<tr>
<td>3</td>
<td>2009-11-10</td>
<td>Input from other partners integrated</td>
</tr>
<tr>
<td>4</td>
<td>2009-11-13</td>
<td>Final call for input (BLITZ, UPF, AM, URL, PGP)</td>
</tr>
<tr>
<td>5</td>
<td>2009-11-23</td>
<td>Version for internal review</td>
</tr>
<tr>
<td>6</td>
<td>2009-11-24</td>
<td>Final version</td>
</tr>
</tbody>
</table>
The work presented in this document was partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th framework programme for R&D.

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

This document contains material, which is the copyright of certain SALERO consortium parties, and may not be reproduced or copied without permission. All SALERO consortium parties have agreed to full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the SALERO consortium as a whole, nor a certain party of the SALERO consortium warrant that the information contained in this document is capable of use, nor that use of the information is free from risk, and does not accept any liability for loss or damage suffered by any person using this information.
# Table of Contents

1 Project Overview .................................................................................................................. 1

2 Summary of Activities ......................................................................................................... 1

3 Third Phase Experimental Productions .............................................................................. 2
   3.1 Karaoke Revolution ......................................................................................................... 2
   3.2 Turing Machine ............................................................................................................... 2
   3.3 MyTinyPlanets.com ........................................................................................................ 3
   3.4 Interactive Videojockey (i-VJ) ......................................................................................... 4

4 Procedural Character Generation ......................................................................................... 5
   4.1 Character Generation ...................................................................................................... 5
   4.2 Facial Animation Pipeline ............................................................................................... 6

5 Speech Synthesis and Audio Processing ............................................................................ 7

6 Semantic Media Research .................................................................................................... 9

7 Integrated Applications ........................................................................................................ 9
   7.1 Automatic Program Generation Pipeline ..................................................................... 9
   7.2 Scalable Codec & Sound Virtualizer ............................................................................. 10
   7.3 Context-Based Search and Retrieval ........................................................................... 11
   7.4 Bones Dailies ............................................................................................................... 13
   7.5 Game Development Tools ............................................................................................ 14
   7.6 Audio Transformation Tools ....................................................................................... 14
      7.6.1 Voice Transformation ............................................................................................. 14
      7.6.2 Tempo & Pitch Transformation ............................................................................... 15
      7.6.3 Advanced Audio Equalizer .................................................................................... 17
   7.7 Vowel Builder & Animation Builder ............................................................................ 17

8 User Involvement, Promotion and Awareness ................................................................... 18
   8.1 User Group ................................................................................................................... 18
   8.2 Publications & Presentations ........................................................................................ 19
   8.3 Awards .......................................................................................................................... 19

9 Conclusions .......................................................................................................................... 19

10 Further Information ............................................................................................................ 20
   10.1 Useful Web Links .......................................................................................................... 20
   10.2 Project Partners ............................................................................................................ 20
   10.3 Contact ........................................................................................................................ 21
1 Project Overview

SALERO’s aim is to make cross media-production for games, movies and broadcast faster, better and more cost efficient by combining computer graphics, language technology, semantic web technologies as well as content based search and retrieval.

SALERO defined and developed ‘Intelligent Content’ for media production, consisting of multimedia objects with context-aware behaviours for self-adaptive use and delivery across different platforms. ‘Intelligent Content’ should enable the creation and re-use of complex, compelling media by artists who need to know little about the technical aspects of the tools that they use.

Based on research into methodologies for describing, creating and finding intelligent content, SALERO developed toolsets to create, manage, edit, retrieve and deliver content objects, addressing characters, objects, sounds, language sets, and behaviours. The toolsets developed and the concept of intelligent content have been be verified by experimental productions.


2 Summary of Activities

In the fourth year of the project major progress has been made in all areas. Examples are the development of advanced algorithms for facial animation, expressive speech synthesis, improved tools for game development, establishment of an ontology-based annotation and search tool taking content based and textual information also into account and the implementation and evaluation of the third phase experimental productions.

Research in facial animation yielded techniques for emotional animation based on the Maskle approach and using the activation/evaluation model to specify emotions. This is combined with work on highly realistic modelling and animation of the appearance of persons.

In respect to expressive speech synthesis a limited domain engine has been implemented. The work in respect to audio transformation has yielded different algorithms to alter the quality of a voice (e.g. converting a female voice to a male one, changing the perceived age of the speaker, pitch transformation).

With i-VJ, Karaoke Revolution, MyTinyPlanets.com and Turing Machine the third phase experimental productions were implemented with the tools developed within the SALERO project and subsequently evaluated.

Results of the project have been presented at IBC 2009 and a showcase of tools and experimental productions has been made available on the project website.
3 Third Phase Experimental Productions

The experimental productions of SALERO have been designed to highlight how an integrated approach to intelligent content can benefit multimedia productions. There have been three phases of experimental production, each allowing partners to integrate their work and research. The third phase experimental productions were designed to integrate as much of the SALERO technology as possible.

3.1 Karaoke Revolution


The singing game was developed for the PlayStation 3, XBOX 360 and Wii and, with the Character Creator, offers almost infinite choices for creating an avatar with different ethnicities, ages and weights as well as facial features, hair styles and clothes. The techniques used worked equally well at the high fidelity level required for the PS3 and XBOX 360 and with the more stylised approach required for the Wii. The game features single and multiplayer modes, 16 venues to perform in and 50 original master tracks. PS3 and XBOX 360 offer online play and further tracks to download.

Figure 2: Screenshots from Karaoke Revolution

3.2 Turing Machine

The Turing Machine Cross Media Experimental Production by TAIK integrated SALERO tools into a research project that investigates associational storytelling and interaction structures, while making the patterns of human-machine communication more visible. It is made up of the physical space installation Alan01, Turing Impact digital High Definition short films and the online service Alan Online.

Figure 3: Alan01 Installation at Media Centre Lume
The *Alan01 interactive installation* engages its audience in a symbolic dialogue with the consciousness of a fictional Alan Turing, using a touch interface developed for this purpose. *Turing Impact HD short films* were produced in a *High Definition Professional Training workshop* organized by TAIK and GVG at Media Centre Lume in Helsinki. These documentaries/fiction films explore the impact Alan Turing and the digital computer has made on contemporary culture. *Alan Online* is the non-material counterpart of the Alan01 installation. It offers a second interface and another aspect to the media content that is shared with the physical installation.

![Figure 4: Screenshot of Alan Online](image)

### 3.3 MyTinyPlanets.com

In the third experimental production phase of SALERO, Peppers Ghost Productions (PGP) produced animation tests for the *My Tiny Planets* browser-based virtual world using both methodologies devised by Dublin Institute of Technology (DIT) and traditional toolsets, and evaluated the integration of these into a variety of on-line delivery mechanisms.

Having established an initial test group of several thousand users, PGP was better able to understand behaviours and requirements of players in virtual worlds, and revise development plans accordingly. It became apparent that animated sequences as a support for game-play, narrative and immersion represent a significant commercial advantage, and will be pursued aggressively during the next phase of development.

DIT’s approaches for both lip-synch (driven primarily by audio amplitude) and activation of animation proved to fit well as a layer over the top of more usual approaches such as layering these over the top of ‘ambient’ motion capture files to prevent stilted movements.

![Figure 5: MyTinyPlanets Games and Demos](image)
3.4 Interactive Videojockey (i-VJ)

This experimental production is a digital television broadcast over Internet TV. Internet TV can be described as a set of multimedia services that are distributed on the “open” internet that is using the already existing network structures. Internet TV is in line with two growing trends in digital content, making the TV experience more interactive and allowing for customization.

![Screenshot from i-VJ](image)

**Figure 6: Screenshot from i-VJ**

*i-VJ* (interactive Video Jockey) is an Internet TV music channel presented by an interactive virtual character and based on a 24/7 automatic broadcast and VoD (Video-on-Demand) services. The content of some sections can be customized by users – I-VJ is an interactive service, hence the structure is not linear. The main services are as follows:

- **Personalization:**
  - My i-VJ: allows users to decide the virtual VJ's appearance.
  - My Playlists: users can create their own playlists.
  - My News: users can activate music news headlines.
  - My Account: users can set up and manage their account.

- **Control:** allows users to have the control of the video with the trick mode options: play/pause, go to previous / next video, repeat video, Random play similar videos, go to the list of songs

- **Add this video:** allows users to insert the music video they are watching into a playlist.

- **Info:** users can view additional information of the video.

- **Search:** users can search music videos with a virtual keyboard.

There are currently two levels of content customization. Users can:

- prepare their own music video playlist and see the videos they want. They can also choose the music headlines they wish to see.
- customize the virtual Video Jockey's appearance by choosing gender, hairstyle, clothes, accessories, etc.
4 Procedural Character Generation

4.1 Character Generation

Creating a new game character can take up to one month of artist and animator time. The aim of procedural generation is to ‘automatically’ and instantly generate a new character using a code driven combination of pre-prepared base assets, meshes, animation rigs, textures and skin shaders.

Instead of a standard linear blend between morphs, which usually results in flat averaging of the two (or more) base meshes, the work in SALERO has led to techniques to additively blend morph offsets on top of existing meshes. This means that we have the ability to blend from ONE base mesh to up to three ethnic variations of this mesh while retaining features from all of the morphs.

The same principles have been applied to full body meshes, enabling a realistic representation of the full range of human shapes and sizes.
The system was designed after extensive research into the current state of the art and is the first to be platform and game title agnostic as well as incorporating a full range of ages, weights and ethnicities.

### 4.2 Facial Animation Pipeline

Following the drive to save modelling and animation time, a new pipeline for facial character animation has been developed which uses the concepts of "activation", controlling the strength of the expression; and "evaluation", controlling the positive-negative balance of the expression, to animate a face. Together with a third variable controlling the asymmetry of the face, this system allows the creation of a video variety of expressions. This model is applied using a new tool called the Maskle, which allows artists and modellers to rapidly weight a face for animation, and use the activation/evaluation system.

![Figure 11: Emotional wheel for Activation (vertical axis) and Evaluation (horizontal axis)](image)

The time required use the facial animation pipeline is a fraction of the time that it would take a computer animator to achieve the same number of expressions using usual methods.

![Figure 12: Screenshot of interface as a Maya (left) and Max (right) plug-ins](image)
Speech synthesis research by Ramon Llull University was oriented to improve the quality when the text input does not belong to a limited domain. In this sense, a new speech database in English was used. The database consists on 10 hours of speech recorded for speech synthesis purposes. The results of the speech synthesis were used together with voice conversion tools in three of the experimental productions: MyTinyPlanets, i-VJ and Alan01. The new voice outperforms the quality of previous created voices.

The UPF voice transformation tool, mentioned above, takes the output synthetic voice generated by the TTS and transforms it via a voice transformation web-service, where the user can select between several voice presets to create multiple synthetic voices (male, female, kid, old, robot, alien, clown, monster) from one single database (Figure 14).
The new version of the TTS system incorporates some improvements related to the research topics we are currently working on:

- Analysis and synthesis of Voice Quality parameters
- Prosody estimation of expressive speech
- Unit Selection Algorithms.

In addition, a website has been developed to be used by mobile phones such as the iPhone, so the TTS is as accessible as possible (Figure 15).

Figure 15 Mobile version of the TTS web service

In order to continuously improve the system, an evaluation platform has been developed that can be used to evaluate any multimedia system. It is based on perceptual evaluations of selected stimulus. TRUE (Testing platfoRm for mUltimedia Evaluation) is an online platform developed to create and perform subjective tests oriented to the evaluation of stimuli of different nature such as audio, video, graphics and text (Figure 16). Due to the high flexibility that the platform offers to researchers different kinds of tests can be carried out, such as emotion identification or quality assessment of synthesis systems, among others. The results can be used for different purposes depending on the research goals, e.g. to validate the emotional content of multimedia data of a corpus.

Figure 16: TRUE allows evaluation by radio buttons and by means of a variety of plugins

In the latest International Conference of the International Speech Communication Association, URL presented work done on automatic recognition of affective information. URL participated in the “Emotion Challenge” which consisted of detecting the emotion elicited by a group of children while they were
playing with the Aibo robot (Sony). The children’s speech was recorded and processed while they were interacting with the robot and expressing a set of different emotions. The greatest challenge was that the speech was spontaneous and recorded in a real environment. The results were been promising and URL finished in first and second position in both modalities of the challenge.

6 Semantic Media Research

In the final year of the project work on the semantic workbench continued, which is used in SALERO to manage ontologies and which provides a set of services that can be used in other applications to use ontologies for annotation or for retrieval of resources. Most notably, the final iteration of the workbench was extended in the search result interface to interfere with a fusion service implemented to centralize the access to search services in the overall project.

On top of the services, a tool for annotation and search has been implemented and evaluated. This tool, called the Intelligent Media Annotation and Search (IMAS) tool, was tailored to the need of untrained users and for that was equipped with a fast and easy to use user interface. Novel services provided by the workbench such as statement-based querying or annotation-based ranking have been implemented there as well. Additionally further ontologies have been built to both capture the annotation model used in the system as well as relationships between resources. All these ontologies are used by the annotation tool.

7 Integrated Applications

7.1 Automatic Program Generation Pipeline

The tools and technology developed during the SALERO project for automatic programme generation can be broadly split into two aspects. The first is introduction of the multimedia production scripting framework which allows fine control of the creation of an animated production or clip. The principal
contribution of this work is enabling aspects of animated production to be automated and re-used according to external data. The framework also allows multiple reuse of assets and semi-automatic programme creation, and the complete separation of content from rendering.

The second aspect deals with the creation of an advanced editing application, the ‘Programme Editor’ for rapid prototyping and programme setup, and designed specifically for use with the framework. The Programme Editor is a Windows application that provides an intuitive drag-and-drop style interface to the placement of programme components in the relation to both themselves and the timeline, and is capable of rapidly adjusting different aspects/sections of a programme (for example, camera type, actor locations, audio, and provide on-the-fly previewing of the results. While essentially acting as a front end for the programme scripting and generation aspects of the framework, the Programme Editor is nevertheless a powerful application in its own right, using a custom 3D graphics engine to display results in real time (see Figure 18). The Framework and Programme received a good reception at the ACM Conference in Advances in Computerised Entertainment (October 2009), where they were presented as a full paper.

![Figure 18: The Program Editor and real-time preview window](image)

### 7.2 Scalable Codec & Sound Virtualizer

The Scalable Audio Codec partly developed within SALERO represents a step forward to offer new services and better user experiences in a world of increasing mobility and connectivity. Often, the network and connectivity is based on wireless technologies. Currently, user experience of media access, for example server-based audio listening through wireless, frequently is not a continuous and glitch-free one, be that in a home wireless network environment or in a travelling automobile. Interruptions occur and services become unavailable due to fluctuations in data bandwidth. With the new codec, the designed scalability embedded in the generated data streams helps to greatly alleviate or even eliminate such problems so that continuous services can be offered.

To further enhance user experience on 2-channel play-back systems (e.g. headphones, and stereo speaker system) a multi-channel Surround Sound Virtualizer is also included. It employs the latest research results to bring a new level of 3D realism and clarity, not only for media entertainment experiences but also possibly for communication when used with the Scalable Audio Codec.
Figure 19: Surround sound virtualization

7.3 Context-Based Search and Retrieval

Three user interfaces have been developed and are currently deployed and ready for use by end users: the first is the AspectBrowser (shown in Figure 20), which allows a user to search and organise text, video and image databases. The second is the Indexing interface which allows new image and video collections to be created. The third user interface is called ViGOR (see Figure 21 and Figure 22) and allows video searching. All three systems share a common backend search and indexing service, which enables different interfaces to be “plugged-in” as required.

Using the indexing interface and a secure FTP client, a new data collection can be created, image and videos can be uploaded onto the server, and the automatic indexing of the uploaded content can be started. Once a collection has been indexed, it appears automatically on the AspectBrowser interface, where it can then be searched using content-based techniques. The AspectBrowser interface allows the user to create multiple search “aspects” which can be organised to aid the user searching the multiple collections; the overall sequence of aspects can be used to provide context to individual aspects.

Figure 20: The AspectBrowser interface, which allows the searching of text, video and image collections

ViGOR is a video retrieval system that employs grouping and recommendation techniques to facilitate online video retrieval tasks. The main objective of this work is to aid users while carrying out exploratory
video retrieval tasks; these tasks can be often ambiguous and multi-faceted. ViGOR (see Figure 21) comprises of a search panel (A), results display area (B), workspace (C) and playback panel (D). These facilities enable the user to both search and organise results effectively. The users enter a text based query in the search panel to begin their search. The result panel is where users can view the search results (a). Additional information about each video shot can be easily retrieved. Placing the mouse cursor over a video keyframe for longer than 3 seconds will result in any text associated with that video being displayed to the user (we will hence forth refer to this action as tooltip) (e). If a user clicks on the play button the highlighted video shot will play in the playback panel. Users can play, pause, stop and navigate through the video as they can on a normal media player.

The workspace is designed as a potentially infinite space to accommodate a large number of groups. Each group can also be used as a starting point for further search queries. Users can select particular videos in the group’s panel (b) and can choose to view an expansion of the group that contains similar videos based a number of different features (d). We will call this functionality local expansion. As the ViGOR system uses YouTube as a backend, the features available to perform a local expansion of the group are mainly standard YouTube features. The interface offers three expansion options (c): 1) text expansion, which is the result of a new search using text extracted from the selected videos; 2) related videos; and 3) videos from the same user. All of the videos returned by these expansion options are retrieved using the YouTube API.

![Figure 21: Screenshot of ViGOR](image)

In order to help facilitate these exploratory video search tasks, two types of recommendation based on past usage history are presented. The users are presented with recommendations of video shots that might match their search criteria based on their interactions (E); these are global recommendations that incorporate the entire search session. These global recommendations are updated whenever the users plays a video, issues a new text query or moves a video to a group panel. One new option for a local expansion is added: the local recommendation (f), with which users may also retrieve recommended videos within a group. These recommendations are localised to each group and are based on the interactions of previous users with videos that the current user has selected.
7.4 Bones Dailies

The Bones Dailies platform is a Linux-based system designed to double the throughput of current dailies solution. This fully scalable turnkey system manages the entire Dailies production process, from ingest to the creation of a colour-graded dailies master.

The system is able to work with SD, HD, 2K or 4K material, to ingest content from film scanners, or other devices such as video tape machines, or digital files (e.g. from digital acquisition cameras). Bones Dailies offers a non-linear approach that increases the efficiency of the entire Dailies process.

It features a template-driven architecture for processes and logical rules, which helps minimize mistakes. It also features built-in real-time colour correction capabilities, which remove the need for expensive colour correction hardware. Bones Dailies also supports ASC based colour grading for interchange with Digital Intermediate systems.

Figure 23: Bones Dailies Production Environment

The Bones Dailies system enhances the efficiency of almost any workflow, saves time and eliminates mundane tasks, automates a variety of processes, and improves the image quality of the Dailies...
material being viewed by clients. It enables collaborative and parallel workflows by sub-dividing the Dailies production process into individual tasks. In a multiple seat configuration all tasks can be shared and performed in parallel: Telecine or tape transfers, audio preparation and sound synchronization, colour correction and playout.

Bones Dailies excels where traditional linear Telecine-based Dailies production workflows fail e.g. traditional workflows cannot keep track of new requirements such as multiple versions, arbitrary tape content, scene and take order.

### 7.5 Game Development Tools

Blitz Games Studios has developed a suite of tools to improve the production pipeline and empower artists, animators and designers to realise their creative vision. The game creation tools are visually based to ensure ease of use and give instant feedback.

The animation visual editor allows game teams to blend many animations seamlessly per character, giving a smooth transition and allowing for subtleties not previously achievable without substantial hand keyed animation effort. The editor replaces manual set ups by programmers and also provides debugging feedback in real-time.

The visual shader editor gives artists the opportunity to create their own visual effects, whether that's improved skin tones, light and shade or post production effects. Previously programmers were needed to implement these in game and the artist's vision was often unrealised.

The visual editor for game scripting replaces the need for programme code in level production and enables animation sequencing, scripted sequences, character behaviours and AI to be placed in a scene, previewed in real time and then iterated to improve the design if needed. This approach supports designers to iterate game play mechanics and level design many more times than was previously possible.

![Figure 24: Visual Animation Editor Screenshot](image)

The Distributed Asset Processing system also supports designers to iterate their work by reducing the time taken to process assets from 90 minutes to under 10 minutes. It employs cluster processing over a LAN and is optimised for game assets. The system is invisible to users and is fully integrated into the BlitzTech SDK.

### 7.6 Audio Transformation Tools

The Audio Transformation Authoring Tools are delivered as a set of online java applets and VST plug-ins, an audio plug-in standard created by Steinberg to allow any third party developers to create VST plug-ins for use within VST host applications.

#### 7.6.1 Voice Transformation

Several singing and speech voice transformations are included in a real-time VST plug-in effect. The system uses spectral techniques for modifying the character of the voice. Transformations include: transposition, quantization, vibrato, roughness, breathiness, whisper, timbre mapping and other spectral
transformations. This tool is targeted to recording or post-production studios that require voice processing.

Figure 25: Voice Transformation VST Plug-in

The Java applet version of the Voice Transformation tool, allows any user to run the voice transformation tool within the web browser. The applet sends the recorded user's voice to a server that transforms the voice based on the transformation parameters chosen by the user from eight different presets and the server sends back the transformation audio file to the client.

Figure 26: Voice Transformation Java applet Web application

1. **Set up your microphone and adjust sound level** (you should see the transform heading to your input voice).
2. **Record** your voice by pressing the record button (recording is limited to 15 seconds, then you can listen to your recorded voice pressing the play button).
3. **Apply the desired transformation** pressing the browse buttons (this process can take some time depending on your connection speed).
4. **Listen** to your voice transformed by playing each of the play buttons below the transformation icons.

7.6.2 Tempo & Pitch Transformation

The Tempo Transformation VST plug-in combines automatic rhythm analysis and time-scaling in order to transform the rhythm of an audio mix. An input audio is first pre-analyzed in order to extract a musical rhythm description. After this process, the user can vary the tempo or add/subtract swing by using the GUI controls during the playback.
The java applet version of the Tempo Transformation tool incorporates also Pitch Transformation for music. This allows transforming tempo and pitch for any mp3 or wave file. The user just drags a sound file to the applet running on the browser and then modifies each of the envelope curves (allows variable pitch and tempo transformations), then pressing the transformation button, the server transforms the file and returns back the transformed file to the client.

Figure 27: Tempo transformation VST Plug-in

Today's online Java applet version of Tempo Transformation tool incorporates also Pitch Transformation for music. This allows transforming tempo and pitch for any mp3 or wave file. The user just drags a sound file to the applet running on the browser and then modifies each of the envelope curves (allows variable pitch and tempo transformations), then pressing the transformation button, the server transforms the file and returns back the transformed file to the client.

Figure 28: Tempo/Pitch Transformation Java applet Web application
7.6.3 Advanced Audio Equalizer

The Advanced Audio Equalizer VST plug-in is aimed for the advanced equalization of commercial music productions. It is a powerful tool able to equalize the audio using the energy histogram based on different criteria. It is not limited to the typical energy distribution versus frequency, but it allows working with other useful energy distributions criteria such as panning, inter-channel phase difference, or energy variance. Some of its applications include:

![Advanced Audio Equalizer VST Plug-in](image)

**Figure 29: Advanced Audio Equalizer VST Plug-in**

Remixing a piece by changing the volume of each instrument independently or removing instruments from the mix and then re-adding them processed by some effects or isolate one or several instruments to perform a musical analysis of a piece.

7.7 Vowel Builder & Animation Builder

The Vowel Builder application uses information from the vowel stress tagging framework to assign animation events to an input speech signal. Vowel Builder implements characters with embedded movements and gestures that are completely independent of the authoring stage. In this manner, fully reusable character animation is possible across characters and languages. An authored asset can be re-used as often as desired, with the character concerned also being completely configurable at runtime.

![Dialogue animation authoring in Vowel Builder](image)

**Figure 30: Dialogue animation authoring in Vowel Builder.**
The tool loads interactive, reusable avatars (robots from My Tiny Planets in the prototype) into an authoring window that allows gestures and movements to be allocated to stressed vowels in the speech file. The application will only allow movements to be allocated to the current characters vowels. Each speech track is given a separate colour, with future work considering the possibilities of group animation authoring.

The Vowel Player is designed to provide online animated characters that can be triggered by any animation sequence authored using the Vowel Builder tool. Characters can be loaded in and out at runtime, allowing any number of reusable characters to be built for use by the same application. The player takes an authored asset for either monologue or dialogue speech and can deliver content interactively in any language:

![Interactive animation in Vowel Player.](image)

Figure 31: Interactive animation in Vowel Player.

The application will deliver speech in any language, in the above case the list of body parts will trigger different interactive responses from the avatar based on user selections.

8 User Involvement, Promotion and Awareness

8.1 User Group

The user group has been revised again in 2009 and an active participation schedule has been set up for the group members’ involvement with the project. In June, a User Group was held coinciding with the International Animation Festival Market (MIFA) in Annecy, France. This was an opportunity to give real demos as well as a hands-on experience session to the User Group Members to test the SALERO tools as well as watch the integration of the tools in the Experimental productions. Moreover, Peter Stansfield from Wavecrest Systems contacted and had individual interviews to introduce the SALERO project to VIDM (Very Important Decision Makers) from the audiovisual industry and get feedback from them.

In addition, results of the project have been disseminated during the IBC exhibition in September in Amsterdam where the project SALERO had a booth in the New Technology Campus to give live demos and presentations and have the possibility to meet with industrial representatives attending the IBC conference. Furthermore, the project was presented in one of the IBC conference sessions. For December a demonstration of SALERO results is planned during the Industry Day session of the SAMT 2009 conference in Graz, Austria.
8.2 Publications & Presentations

Results from SALERO project have been showcased at the following events:

- IBC 2009, Amsterdam, The Netherlands (September 2009): The SALERO tools and experimental productions have been presented on a booth in the New Technology Campus of the IBC exhibition.
- In April 2009 TAIK organized a “Training for Professionals Workshop” in Helsinki as a collaboration between Digital Film Technology (DFT) and the Media Lab (TAIK). Considerable assistance and resources were supplied by LUME at the University of Art and Design Helsinki.

SALERO’s research results were presented at the following conferences:

- 1st Young Researchers Workshop in Speech Technology (April 2009, Dublin, Ireland)
- Reasoning Web 2009 Summer School (August 2009, Brixen, Italy)
- International Conference on Semantic Systems – I-Semantics 2009 (September 2009, Graz, Austria)
- 10th Annual Conference of the International Speech Communication Association – INTERSPEECH 2009 (September 2009, Brighton, United Kingdom)
- Physicality Workshop at 23rd Conference on Computer Human Interaction – HCI 2009 (September 2009, Cambridge, United Kingdom)
- 9th Annual Conference on Information Technology and Telecommunication (October 2009, Dublin, Ireland)
- ACM International Conference on Multimedia (October 2009, Beijing, China)
- 4th International Conference on Semantic and Digital Media Technologies – SAMT 2009 (December 2009, Graz, Austria)
- ACM International Conference on Advances in Computerised Entertainment Technology – ACE 2009 (October 2009, Athens, Greece)

8.3 Awards

The Turing Machine production received the EUROPRIX Quality Seal. The project was evaluated in two rounds by two different panels of jurors. By awarding the quality seal they wish to recommend the project to consumers and users. The EUROPRIX Quality Seal is proof that the project is highly innovative and creative.

9 Conclusions

With the results reached during the last year of the project a solid base is laid for the exploitation of the project results. The results will enable the different partners to either provide new software solutions to the market or provide new and enhanced services based on them. For the research partners on the one side it was possible to bring new insight to their fields of expertise and on the other side they started to explore new fields based on the collaboration with all partners – research as well as industry.
10 Further Information

10.1 Useful Web Links

- [http://www.salero.eu](http://www.salero.eu): SALERO project website including a showcase of SALERO’s results.
- [http://www.youtube.com/intelligentcontent](http://www.youtube.com/intelligentcontent): SALERO’s YouTube channel “Intelligent Content” publishing results from the projects research and the experimental productions.
- [http://ninoscompetence.wordpress.com](http://ninoscompetence.wordpress.com): Introduction to the Program Editor.
- [http://www.salero.eu/imas](http://www.salero.eu/imas): The Intelligent Media Annotation and Search tool
- [http://mir.dcs.gla.ac.uk:8080/proto/ExploratoryInterface.html](http://mir.dcs.gla.ac.uk:8080/proto/ExploratoryInterface.html): The “Facet Browser”.
- [http://mir.dcs.gla.ac.uk:8080/AspectBrowser/AspectBrowser.html](http://mir.dcs.gla.ac.uk:8080/AspectBrowser/AspectBrowser.html): The “Aspect Browser”.
- [http://mtg.upf.edu/project/kaleivoicecope](http://mtg.upf.edu/project/kaleivoicecope): MTG-UPF Voice Transformation Technology
- [http://mlab.taik.fi/alanonline](http://mlab.taik.fi/alanonline): The Alan Online website describes the Turing Machine Cross Media Production and includes the non-material counterpart of the Alan01 installation as an interactive interface
- [http://www.trusim.com](http://www.trusim.com): The character creator
- [http://www.blitzgamesstudios.com](http://www.blitzgamesstudios.com): Karaoke Revolution

10.2 Project Partners

- Blitz Games Studios Ltd, [http://www.blitzgamesstudios.com](http://www.blitzgamesstudios.com)
- Digital Film Technologies, [http://www.dft-film.com](http://www.dft-film.com)
- DLLNI Ltd, [http://www.dts.com](http://www.dts.com)
- Dublin Institute of Technology, [http://www.dit.ie](http://www.dit.ie)
- Fundació Barcelona Media Universitat Pompeu Fabra, [http://www.iua.upf.es](http://www.iua.upf.es)
- Leopold-Franzens Universität Innsbruck, [http://www.sti-innsbruck.at](http://www.sti-innsbruck.at)
- Pepper’s Ghost Productions Ltd, [http://www.peppersghost.com](http://www.peppersghost.com)
- Taideteollinen Korkeakoulu, [http://www.taik.fi](http://www.taik.fi)
- Universitat Ramon Llull, [http://www.salleurl.edu](http://www.salleurl.edu)
- Universitat Pompeu Fabra, [http://mtg.upf.edu](http://mtg.upf.edu)
- University of Glasgow, [http://www.dcs.gla.ac.uk](http://www.dcs.gla.ac.uk)
10.3 Contact

Georg Thallinger
JOANNEUM RESEARCH Forschungsgesellschaft mbH
Institute of Information Systems & Information Management
Steyrergasse 17
A-8010 Graz
AUSTRIA
Email: georg.thallinger@joanneum.at
phone: +43-316-876-1240
fax: +43-316-876-1191