9
Conclusion

This thesis proposed a new approach to video-based interactive narratives that uses video compositing techniques to dynamically create video sequences representing the story events – rather than using only prerecorded scenes. In our method, actors are filmed by several cameras in front of a green screen in a variety of emotional states and situations that are compatible with the logical structure of narratives. Afterwards, an automatic process controls the cinematography language, compositing scenarios and choosing view parameters (zoom, angle and camera movements). This approach allows the generation of more diversified stories, increases interactivity, and reduces production costs. This chapter presents the conclusions remarks, summarizes the contributions, points some limitation of our approach and suggests topics for future research work.

9.1.
Concluding Remarks

This thesis explored video-based interactive narratives from three different points of view: authors, developers and users. For authors, we presented a general guide on how to write and film interactive stories, and also developed some computational tools to help them in the production process. For developers, we proposed an architecture for video-based interactive storytelling systems and presented the technical details about the implementation of the real-time video compositing and editing algorithms. For users, we designed attractive and engaging interaction mechanisms.

In any form of storytelling, the author is the key component for a successful story. However, authoring for interactive storytelling is a difficult task. It involves the process of logically specifying the context of the story, thinking about possible events in terms of parameters, preconditions and effects, which are tasks that typical story writers are not familiar with. In addition, the production team (e.g.

artists, cinematographers, actors) must produce the visual content (e.g. 3D models, videos, 2D animations) to represent the narrative in accordance with the logical specification of the story and considering all possible storylines that may be created to comply with the user's desires. Our initial thought about the use of videos to dramatize interactive narratives was that it would simplify and reduce the production work in comparison with the artistic efforts necessary to produce 3D models and animations for a 3D dramatization. However, along the development of this work, we have realized that filming and editing the video material necessary for a video-based dramatization require as much authorial work as producing a 3D/2D interactive narrative. Obviously it involves different professionals – while a 3D interactive narrative requires designers and artists, a video-based dramatization requires filmmakers and actors.

From the technical point of view, developing a video-based dramatization system required completely new algorithms and techniques for video compositing and virtual cinematography. Traditional 3D/2D dramatization systems adopt many of the techniques used in games and simulations to create the story worlds and control the virtual characters, animations and cameras. On the other hand, video-based systems cannot make use of most of those techniques due to limitations imposed by the video resources (e.g. lack of freedom to show characters from any angle, non-parameterized actions and movements, immutable video sequences). In order to overcome some of those limitations, we proposed filming the actors and locations from multiple angles in front of a green screen, which allowed us to create specialized algorithms for positioning actors and cameras, selecting the best shots to film the scenes, simulating camera movements, and compositing all scene elements into a single piece of motion picture in real-time. The results of the technical evaluation tests demonstrated the efficiency and applicability of the proposed video compositing and editing algorithms for video-based interactive storytelling.

Since the beginning of this research, television and cinema were the main target mediums of the proposed video-based interactive storytelling system. The advance of interactive storytelling technology to these new mediums required new interaction mechanisms to support the multi-user characteristic of these platforms. The proposed user interaction interfaces (social networks and mobile devices) provided the basic multi-user setting required for both mediums. In addition, they

provided an engaging way for users to interfere in the narratives through natural language. Although few experiments had been conducted on applying these mechanisms on real environments of television and cinema, the small scale experiments confirmed their applicability and efficacy in providing an engaging user interaction interface that support multi-user interactions. We believe that video-based interactive storytelling may be the first step towards the emergence of real interactive films, which can expand the boundaries of traditional branching interactive films towards a new form of digital entertainment.

The next sections present more details about the contribution of this thesis, its limitations and some directions for future research.

9.2. Contributions

The main contributions of this thesis are:

- *Video-based interactive storytelling using video compositing techniques*. We proposed a new method to represent video-based interactive narratives using real-time video compositing and editing techniques. Previous works on video-based interactive storytelling are all based on static and immutable pre-recorded video sequences that are rearranged during presentation, which reduce interactivity, story diversity, and increase the productions costs. The proposed method dynamically generates video sequences representing the story events in real-time, which provides the system with the possibility of generating more diversified stories without increasing production costs.
- *Interactive video narratives based on cinematography principles and techniques*. The proposed techniques for the generation of video-based interactive narratives follow cinematography principles and rules to create attractive and engaging visual representations for the story events. The architecture of our system is composed of a set of cinematography-based autonomous agents that share the responsibility for creating dynamic video sequences respecting cinematography rules. Previous works on video-based interactive narratives focus mainly on the creation of stories by

ordering video segments, without taking into account cinematography concepts.

- *Video-based interactive storytelling and robust story generation algorithms*. The proposed video-based interactive storytelling system is integrated with a robust story generation system based on planning under nondeterminism and capable of generating complex and diversified interactive story plots. Most of the previous works on video-based interactive storytelling, especially the interactive films produced for TV and Cinema, are based on rudimentary branching narrative structures, which simplify the use of videos for the representation of the story (all possible events are predefined and can be pre-recorded), but compromise the diversity of stories and the user's sense of agency. By integrating real-time video-compositing techniques with robust automated story generation algorithms, these limitations can be overcome and video-based interactive narratives with real interactive and dynamic plots can be created.
- *Real-time video compositing algorithm*. We proposed a parallel frame compositing algorithm capable of managing and compositing multiple video frames simultaneously to guarantee real-time performance. The algorithm was evaluated through a performance test, which demonstrates that compositing a frame becomes more expensive as more scene elements are added to the frame. However, the parallel architecture of the proposed algorithm can compensate the expensiveness of the frame compositing task by dividing the work among multiple CPU cores. The algorithm can also be used in other applications that require some form of automated video compositing process.
- *Automated method for shot selection using expert cinematography knowledge*. We proposed a method to select the best camera shots to show the generated scenes of video-based interactive narratives. Our approach consists of representing the knowledge of a real film editor using several artificial neural networks trained to solve cinematography problems involving camera shot selection. The proposed technique is capable of learning the personal editing style of human editors and replicating it during the video compositing process, which keeps the signature of the human artist in the computer generated content. This method is based on

our previous work that used support vector machines (SVM) to select the camera angles in a 3D environment (Lima et al. 2010).

- *Automated method for the selection of scene transitions based on cinematography theory*. We proposed a method to guarantee the temporal and spatial continuity of video-based interactive narratives by avoiding jump cuts and selecting the most adequate shot transition for the narrative scenes. Our approach consists in translating cinematography principles and practices directly into logical rules. The method was evaluated by comparing the results of the proposed method with the decisions made by human editors of well-known movies. The results indicate that our method is capable of selecting scene transitions as professional human editors in most of the cases.
- *Multi-user natural language interface for interactive storytelling using mobile devices*. We proposed a new multi-user interface that allows users to freely interact with virtual characters by text or speech using mobile devices. The interaction mechanism was designed to support both global and local user interactions. By using the proposed method, users are able to write or speak what they want to happen in the story, or easily select the desired outcome for local decision points. Most previous works on interaction methods for interactive storytelling focus mainly on single-user interactions.
- *Multi-user interaction though social networks*. We explored the use of social networks as a multi-user interface. We presented and evaluated an interaction method that allows users to interact and change stories through social networks (such as Facebook, Twitter and Google+). This method allows users to collaborate with the development of interactive stories in a social environment through their own social network clients, using smartphones, tablets, or personal computers without having to install any additional software. The activity that results from the user interactions in the social network may attract more viewers to the broadcasting channel (increasing the audience). In addition, viewers can make new friends through the interaction in the social network. As far as we are aware, this is the first time this form of interaction is explored in an interactive narrative.

- *General guide and computational tools for the production of video-based interactive narratives*. We presented a general guide on how to write and film interactive narratives, and proposed some computational tools developed to assist the production of video-based interactive narratives. The author is one of the most important components in any form of storytelling, especially in video-based interactive storytelling, where he/she has to specify the logical context of the story and cinematographers have to film and edit the videos resources necessary for the dramatization of the interactive story. As far as we are aware, this is the first research work to explore the concepts of authoring in video-based interactive storytelling.

9.3. Publications and Awards

The results of the research on video-based interactive storytelling were published in leading conferences in the field of multimedia and interactive storytelling. The real-time video editing method that automatically generates the most adequate shot transitions, avoids jump cuts, and creates looping scenes, was published in the International Conference on Multimedia and Expo (Lima et al. 2012A). In addition, more papers on this matter are being prepared for submission to journals of multimedia and computer entertainment.

The research on user interaction methods also has originated some publications: the social interaction method for interactive storytelling was published in the International Conference on Entertainment Computing (Lima et al. 2012B); the multi-user natural language interface using mobile devices was published in the Brazilian Symposium on Computer Games and Digital Entertainment (Lima et al. 2012C); and another paper describing a study on multimodal, multi-user and adaptive interaction methods was also published in the Brazilian Symposium on Computer Games and Digital Entertainment (Lima et al. 2011B). The results of the research on story dramatization also have originated some publications: a paper exploring the use of an augmented reality visualization interface combined with a sketch-based interaction interface was published in the International Conference on Entertainment Computing (Lima et al. 2011A), and in

the journal of Entertainment Computing (Lima et al. 2014A). In addition, another paper presenting a system capable of generating dynamic interactive narratives in the format of comic books was published in the International Conference on Advances in Computer Entertainment Technology (Lima et al. 2013).

The research that led to this thesis also received two international awards from the International Telecommunication Union (ITU).[7] The first award is an honorable mention on "*Innovation*" in the "*1nd ITU IPTV Application Challenge*" competition (2011), with the video-based interactive narrative called "*The Princess Kidnapping*"; and the second award is an honorable mention on "*Interactivity*" in the "*2nd ITU IPTV Application Challenge*" competition (2012), with the comic-based interactive narrative called "*Little Gray Planet*". Both interactive narratives were designed for interactive TV. The ITU is the United Nations specialized agency for information and communication technologies.

9.4. Limitations and Directions for Future Research

Although the proposed approach to create video-based interactive narratives has achieved the primary objectives of this thesis, we also identified some limitations and directions for future research, which can be categorized into three main topics: image quality, authoring process, and evaluation experiments.

The image quality of the results produced by our system is still far from the excellent visual quality of feature films. Image quality depends on real-time techniques for realistic lighting, which relight actors with the proper illumination of the environment, and consider cast shadows and interreflection. Interactive real-time video rendering with complex illumination and materials is still an open issue even in the multimedia research area. A future work would be to explore the existent dynamic lighting techniques and verify the possibility of applying them in the real-time video compositing process for interactive storytelling. Examples of promising approaches include the use of techniques for capturing the actor's live-action performance illuminating him with a sequence of time-multiplexed basis lighting conditions (Wenger et al. 2005; Chabert et al. 2006), and the use of

[7] ITU - http://www.itu.int/

interactive ray tracing techniques in the compositing process (Pomi and Slusallek 2005).

The second main limitation of our approach is related with the amount of authorial work during the production and post-production phases. The proposed method to generate video-based interactive narratives is entirely based on the use of video resources filmed from different angles, which gives to the system the freedom to dramatize scenes applying the basic cinematography concepts during the dramatization of the narrative. However, filming the actors performing their actions from 8 different angles generates a huge number of video files, which grows according to the number of actions the characters can perform during the narrative. The process of editing and removing the background of all these videos using a traditional chroma key matting technique requires a huge amount of work in the post-production phase, which increases the production costs. An alternative to overcome this limitation could be the adoption of a more efficient and automated matting technique, such as a hardware-based solution (Joshi et al. 2006; Sun et al. 2006). These solutions may also improve the current visual quality of the compositing results, which suffers from color spills produced by the green screen background in the actors.

Another factor that increases the amount of work during the production and post-production phases is the existence of replaceable accessories or clothes in the characters, which will require the same actions to be filmed several times varying the accessories/clothes. A possible solution to this problem would be the inclusion of the dynamic and replaceable objects in the scenes during the compositing process using a tracking procedure to correctly sync the object movements with the actor movements. For example, if a character needs to hold different weapons during the narrative, he could be filmed holding a generic object with distinct tracking markers that would be tracked during the compositing process to identify the correct position to place any weapon in the character hands.

Another limitation of this thesis is the lack of large-scale user experiments to validate the usability of the proposed video-based system from a Human-Computer Interaction (HCI) perspective. An interesting experiment would be a comparative study between a video-based interactive narrative and a 3D/2D version of the same story. In this direction, the IRIS Evaluation Toolkit (Klimmt

et al. 2010; Roth et al. 2009) provides a good methodology to evaluate and compare the general users' experience provided by both dramatization modalities.

The visual quality of the video sequences produced by the proposed video compositing algorithms also has to be evaluated in more precise studies. An interesting experiment to complement the visual evaluation presented in this thesis would be a Turing Test applied to the generated video sequences in order to evaluate if human subjects are able to differentiate the scenes created by the compositing algorithms and the scenes created by the filmmaking professionals.

The interactive film production process described in this thesis and the authoring tasks for video-based interactive storytelling also need to be better evaluated. An interesting future work would be a deeper study about this process from the perspective of the people involved in the authoring tasks, which may provide a more detailed feedback about the problems and possible solutions to improve the process and the video-based interactive storytelling system in general. In addition, a more precise evaluation of the production costs is also necessary. In this direction, another interesting future work would be a comparative study of the costs for producing 3D interactive narratives, video-based interactive narratives using video compositing techniques, and video-based interactive narratives using only static video segments.

The present system was built based on the third version of the Logtell system, which incorporates the basic temporal modal logic of the first version (Ciarlini et al. 2005), the client/server architecture of the second version (Camanho et al. 2009), and planning under nondeterminism (Silva et al. 2010) combined with the use of nondeterministic automata to control the dramatization of events (Doria et al. 2008) that where introduced in the third version of the Logtell. Much work remains to be done towards the integration of the proposed dramatization system and user interaction interfaces with the recent advances in the Logtell Project, such as the stream-based architecture for delivering interactive narratives in multiple platforms (Camanho et al. 2013), the new non-deterministic planning model using dramatic properties of the story events (Gottin 2013; Ferreira 2013), and the incorporation of information-gathering events in the story plots (Silva et al. 2012).