4
Video-Based Interactive Storytelling

This thesis proposes a new approach to video-based interactive narratives that uses real-time video compositing techniques to dynamically create video sequences representing the story events generated by planning algorithms. The proposed approach consists of filming the actors representing the characters of the story in front of a green screen, which allows the system to remove the green background using the chroma key matting technique and dynamically compose the scenes of the narrative without being restricted by static video sequences. In addition, both actors and locations are filmed from different angles in order to provide the system with the freedom to dramatize scenes applying the basic cinematography concepts during the dramatization of the narrative. A total of 8 angles of the actors performing their actions are shot using a single or multiple cameras in front of a green screen with intervals of 45 degrees (forming a circle around the subject). Similarly, each location of the narrative is also shot from 8 angles with intervals of 45 degrees (forming a circle around the stage). In this way, the system can compose scenes from different angles, simulate camera movements and create more dynamic video sequences that cover all the important aspects of the cinematography theory.

The proposed video-based interactive storytelling model combines robust story generation algorithms, flexible multi-user interaction interfaces and cinematic story dramatizations using videos. It is based on the logical framework for story generation of the Logtell system, with the addition of new multi-user interaction techniques and algorithms for video-based story dramatization using cinematography principles.

This chapter discusses the related works and describes the main differences between the proposed system and previous work. It also presents an overview of the architecture of the video-based interactive storytelling system from a software engineering perspective.

PUC-Rio - Certificação Digital Nº 1021793/CA

4.1. Related Work

The idea of using videos as a form of visual representation of interactive narratives is not entirely new. The first attempts to use prerecorded video segments to represent some form of dynamic narrative date back to the 1960s (Činčera et al. 1967; Bejan 1992; Chua and Ruan 1995; Davenport and Murtaugh 1995; Ahanger and Little 1997) and several other interactive narrative experiences using videos have been developed through the years. The game industry was the first to explore the use of videos as a form of interactive content. During the early 1980s a new class of games, known as full motion video (FMV) based games or simply by interactive movies, emerged and became very popular. The main characteristic of these games is that their content was mainly based on pre-recorded video segments rather than sprites, vectors, or 3D models.

The first game to explore the use of full motion videos as the game content was Dragon's Lair (1983). Although the genre came to be associated with live-action video, its first occurrence is an animated interactive movie. In Dragon's Lair, the player has the role of a sword fighting hero who needs to win many fights and gather items to finally free a princess from a dragon. The gameplay consists of making decisions by using a joystick to give directions to the virtual character. If the player chooses the right action and its respective button is pressed at the right moment, the obstacle is overcome. If not, the character dies and the player loses a life. Space Ace (1984), another game from the same production team of Dragon's Lair, used a similar idea, but improved on its predecessor with an expanded storyline, with multiple branch points and selectable skill levels.

FMV-based games were considered the cutting edge technology at the beginning of the 1990s and were seen as the future of the game industry. However, as the consoles of that time evolved, the popularity of these games decreased drastically. Today, they are known as one of great failures of the game industry (Wolf 2007). The main problem was the lack of interactivity. The gameplay of most part of them was based on pressing a sequence of buttons in pre-determined moments to keep the narrative moving forward. The narratives also had a very limited branching factor, because every action, every movement, every success and every failure had to be either pre-filmed or pre-rendered.

PUC-Rio - Certificação Digital Nº 1021793/CA

Obviously it was expensive in terms of production, so the designers had to reduce the interaction options to reduce costs. At that time, FMV-based games failed in the attempt of creating a link between games and films.

In the same time, academic researchers begin to explore the capabilities of videos as an interactive content. Davenport and Murtaugh (1995) present a method to maintain temporal continuity between segments of videos by scoring metadata associated with all available scenes. In their application, users are able to navigate through a collection of documentary scenes describing theme, time and location. Terminal Time (Mateas et al. 2000) is another example of narrative system that uses videos to produce historical documentaries based on the audience's appreciation of ideological themes. It focuses on the automatic generation of narrative video sequences through a combination of knowledge-based reasoning, planning, natural language generation, and an indexed multimedia database. In their system, video clips are subsequently selected from the multimedia database according to keywords associated with the documentary events and annotated video clips. In a similar approach, Bocconi (2006) presents a system that generates video documentaries based on verbal annotations added to the audio channel of the video segments. A model of verbal relations is used to automatically generate video sequences for user-specified arguments. In another work, Chua and Ruan (2005) designed a system to support the process of video information management: segmenting, logging, retrieving, and sequencing. Their system semi-automatically detects and annotates shots for later retrieval. The retrieving system uses rules to retrieve shots for presentation within a specified time constraint.

Ahanger and Little (1997) present an automated system to compose and deliver segments of news videos. In the system, content-based metadata and structure-based metadata are used to compose news items. The composition process is based on knowledge about the structure of a news item (introduction, body, and end) and how various types of segments fit into the structure. Within restrictions imposed by the composition grammar, segments belonging to the body can be presented in any order if their creation times are within a small range. Related segments can be included or excluded to meet preference to time constraints without sacrificing continuity. The authors also present a set of metrics to evaluate the quality of news videos created by the automated editing process.

PUC-Rio - Certificação Digital Nº 1021793/CA

These metrics include thematic continuity, temporal continuity, structural continuity, period span coverage, and content progression (Ahanger and Little 1998). In a similar approach, but not focusing on news videos, Nack and Parkes (1997) present a method to establish continuity between segments of videos using rules based on the content of the segments. Their application is capable of automatically generating humorous video sequences from arbitrary video material. The content of the segments is described with information about the characters, actions, moods, locations, and position of the camera.

Hypervideo, or hyperlinked video, is another form of media that explores interactivity by including embedded user-clickable anchors into videos, allowing the user to navigate between video and other hypermedia elements. HyperCafe (Sawhney et al. 1996) is one of the first hypervideo examples that were primarily designed as a cinematic experience of hyper-linked video scenes. Currently, hypervideo research is mainly focused on the efficient definition of interactive regions in videos. VideoClix (2014) and ADIVI (2014) are examples of authoring tools for defining flexible hyperlinks and actions in a video. However, they do not directly support the generation of interactive narratives.

Another research problem closely related to automatic video editing is video summarization, which refers to the process of creating a summary of a digital video. This summary, which must contain only high priority entities and events from the video, should exhibit reasonable degrees of continuity and should be free of repetition. A classical approach to video summarization is presented by Ma et al. (2002). The authors present a method to measure the viewer's attention without fully semantic understanding of the video content. As result, the system could select the high priority video events based on the evoked attention.

The idea of a generic framework for the production of interactive narratives is explored by Ursu et al. (2008). The authors present the ShapeShifting Media, a system designed for the production and delivery of interactive screen-media narratives. The productions are mainly made with prerecorded video segments. The variations are achieved by the automatic selection and rearrangement of atomic elements of content into individual narrations. The system does not incorporate any mechanism for the automatic generation of stories. Essentially, their approach is to empower the human-centered authoring of interactive narratives rather than attempting to build systems that generate narratives by

PUC-Rio - Certificação Digital Nº 1021793/CA

themselves. The applications developed with the ShapeShifting Media system include My News & Sports My Way, in which the content of a continuous presentation of news is combined in accordance with users' interest, and the romantic comedy Accidental Lovers, in which users can watch and influence a couple's relationship. In Accidental Lovers, viewers are able to interact with the ongoing story by sending mobile text messages to the broadcast channel. Changes in the emotional state of the characters and their relationships depend on the existence of some specific keywords found in the viewer's messages. Accidental Lovers was broadcasted several times on Finnish television in late December 2006 and early January 2007 (Williams et al. 2006). Another example of system for the production of interactive narratives is presented by Shen et al. (2009). Their system helps users to compose sequences of scenes to tell stories by selecting video segments from a corpus of annotated clips.

Another example of interactive narrative automatically edited and broadcasted by a TV channels is Akvaario (Pellinen 2000). Similarly to Accidental Lovers, in Akvaario viewers also can influence the mood of the protagonists through mobile text messages. The system uses a large database of clips (approximately 5000), and relies on many features of the database organization to choose the adequate video segments based on the content of the viewer's messages (Manovich 2001).

There are also some examples of video-based interactive narratives for cinema. *Kinoautomat* (Činčera et al. 1967) is one of the first interactive films produced for cinema (Hales 2005). The film comprises nine interaction points, where a human moderator appears on stage and asks the audience to choose between two scenes. Then, the public votes on their desired option by pressing colored buttons installed on the theater seats. Based on the audience votes, the lens cap of two projectors is manually switched to project only the selected scene. Kinoautomat was exhibited for six months and attracted a public of more than 67 thousands of viewers. Following a similar approach, the short interactive film *I'm Your Man* (Bejan 1992) was also exhibited on movie theaters and allowed the audience to interact in six points of the story by choosing between three different options. A more recent interactive experience is Last Call (Jung von Matt 2010), which is an interactive advert for the 13th Street TV Channel exhibited experimentally in movie theaters. In Last Call, the audience interacts with the

PUC-Rio - Certificação Digital Nº 1021793/CA

actress talking to her via cell phone. Based on the audience voice commands, the system selects the sequence of videos to be presented according to a fixed tree of prerecorded video segments.

In a recent research work, Porteous et al. (2010) present a video-based storytelling system that generates multiple story variants from a baseline video. The video content is generated by an adaptation of video summarization techniques that decompose the baseline video into sequences of interconnected shots sharing a common semantic thread. The video sequences are associated with story events and alternative storylines are generated by the use of AI planning techniques. Piacenza et al. (2011) present some improvements to these techniques using a shared semantic representation to facilitate the conceptual integration of video processing and narrative generation. However, continuity issues are not tackled by their approach. As these video segments can be joined in different orders, several continuity failures may occur, in particular because their system uses video segments extracted from linear films. The planning algorithm ensures only the logical continuity of the narrative, but not the visual continuity of the film.

Another recent research that explores the use of videos in interactive storytelling is presented by Müller et al. (2013). Those authors describe a system for the production and delivery of interactive narratives, whose web-based client interface represents stories using short video snippets. However, as other previous works, their system relies only on static video segments and cinematography principles are not applied to ensure the consistency of presented video stories.

In general, the previous works surveyed here focus basically on the creation of stories by ordering pre-recorded video, without using cinematography concepts. The interactive narratives broadcasted by TV channels and exhibited in theaters are entirely based on predefined branching narrative structures. Moreover, previous works adopt only immutable pre-recorded videos, which reduce interactivity, story diversity, and increase the productions costs. None of the previous works uses video compositing techniques to generate video-based interactive narratives in real-time. The proposed thesis differs from the aforementioned works because it proposes a general model for video-based interactive storytelling based on planning and cinematography theory. The proposed approach uses video compositing techniques in order to create video

PUC-Rio - Certificação Digital Nº 1021793/CA

sequences representing story events generated by a planning algorithm in real-time.

4.2. System Requirements

Based on the cinematography and interactive narrative theories, we established some basic requisites for a video-based interactive storytelling system:

1. **Interactivity:** Interactivity is the key element of interactive narratives. It differentiates interactive narratives from simple linear stories. However, the level of interaction must be carefully planned. The audience must keep the attention on the narrative content without being distracted by the interaction interface. A video-based interactive narrative must handle user interactions and present the results of the user interventions without breaking the continuity of the narrative. In addition, the interaction interface must support multi-user interactions and be unobtrusive to users that just want to watch the narrative without interactions.
2. **Flexibility:** One of the main challenges when developing the dramatization module of an interactive storytelling system is how to make it generic, flexible and adaptable for the presentation of different story domains. A video-based dramatization system must be flexible and independent of story domain.
3. **Autonomy:** In interactive storytelling, stories are usually generated in real-time and the system must be capable of representing all the stories without human intervention. A video-based dramatization system must be capable of:
    - **a.** Automatically compose the scenes to represent the story events;
    - **b.** Autonomously control the behavior of the characters participating in the action;
    - **c.** Automatically select the best shots during the compositing process;

PUC-Rio - Certificação Digital Nº 1021793/CA

   d. Automatically select the best music and illumination to express the emotions of the scenes;

4. **Real-Time Performance:** The ability of generating and presenting narratives in real-time is crucial to any interactive storytelling system. In a video-based interactive narrative, the system must be capable of composing video sequences to represent the story events in real-time, without noticeable delays and keeping the visual continuity of the film.
5. **Continuity:** In a film, continuity means keeping the narrative moving forward logically and smoothly, without disruptions in space or time. When the audience becomes aware of continuity errors, they simultaneously become conscious that they are watching a movie, which breaks the storytelling illusion. A video-based interactive storytelling system must be capable of keeping the visual and temporal continuity of the narrative.
6. **Expressing emotions:** Expressing and evoking emotions is a key factor to engage the audience in a narrative. The cinematography theory describes several ways to emphasize emotions by using specific camera shots, camera movements, light and music. A video-based interactive narrative must emphasize the dramatic content of the story by correctly employing cinematography principles according to the emotional content of the narrative to create an attractive and engaging visual representation of the story.

PUC-Rio - Certificação Digital Nº 1021793/CA

4.3. Operating Cycle and System Modules

Similarly to previous interactive storytelling systems, the main operating cycle of a video-based interactive storytelling system can be divided in three main processes: *story generation*, *user interaction* and *story dramatization*:

1. The *story generation* phase makes use of planning algorithms to create and update the story plot;

2. The *user interaction* phase allows users to intervene in the narrative in a direct or indirect way;
3. The *story dramatization* phase represents the events of the story plot using videos.

The main difference between previous systems and a video-based interactive storytelling system lies in the *story dramatization* phase, which uses videos with living actors to present the story events instead of computer generated 2D or 3D animations.

Each phase of the operating cycle is implemented in a different module. The proposed system is composed of three main modules (Figure 4.1): Story Generator, User Interaction and Story Dramatization, which implement their respective phases in the operating cycle (*story generation*, *user interaction* and *story dramatization*). Each module integrates a dedicated controller in charge of handling the network communication between the components: a Planner Controller for the Story Generator, a Drama and an Interaction Controller for the Story Dramatization, and a Global and a Local Interaction Controller for the user Interaction module. Each controller is responsible for interpreting and managing the messages received from other modules.

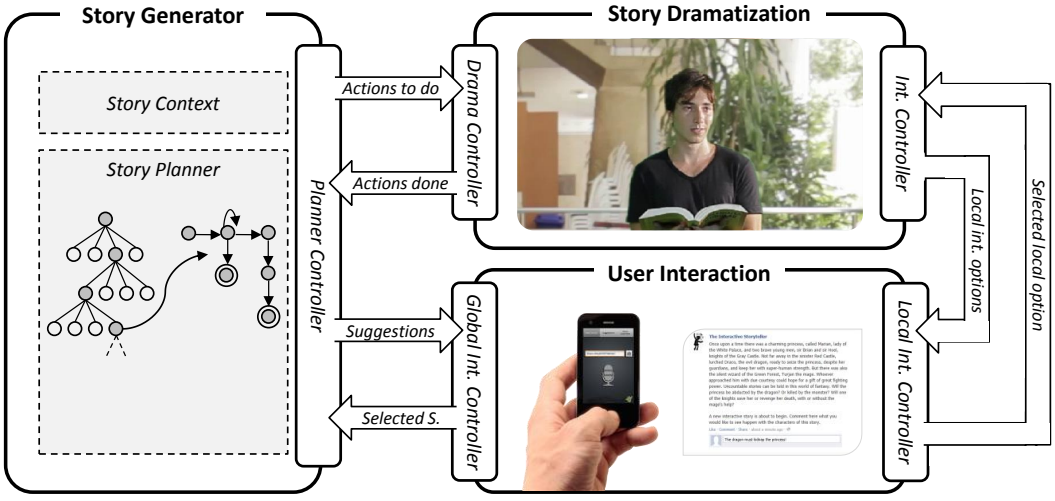

Figure 4.1: System modules.

The three modules were designed to work independently and to run on separate computers, which reduces the computational overhead of running a

PUC-Rio - Certificação Digital Nº 1021793/CA

complex planning task together with time consuming image processing algorithms for video dramatization. The system adopts a client/server architecture, where the story generator and user interaction modules are both servers, and the story dramatization module is the client interface. This architecture allows several instances of the story dramatization module to be connected with the story generator and the user interaction servers, allowing several users to watch and interact with the same or different stories. The communication between the modules is done through a TCP/IP network connection.

The system adopts the story generation algorithms of Logtell, and consequently follows its approach of generating stories in chapters, which are represented as contingency trees, where the nodes are nondeterministic events and the edges correspond to conditions that enable the execution of the next event. As illustrated in Figure 4.2, a nondeterministic event e_i is executed by a nondeterministic automaton composed by basic actions a_i. The basic actions correspond to the primitive actions that can be performed by the virtual characters during dramatization.

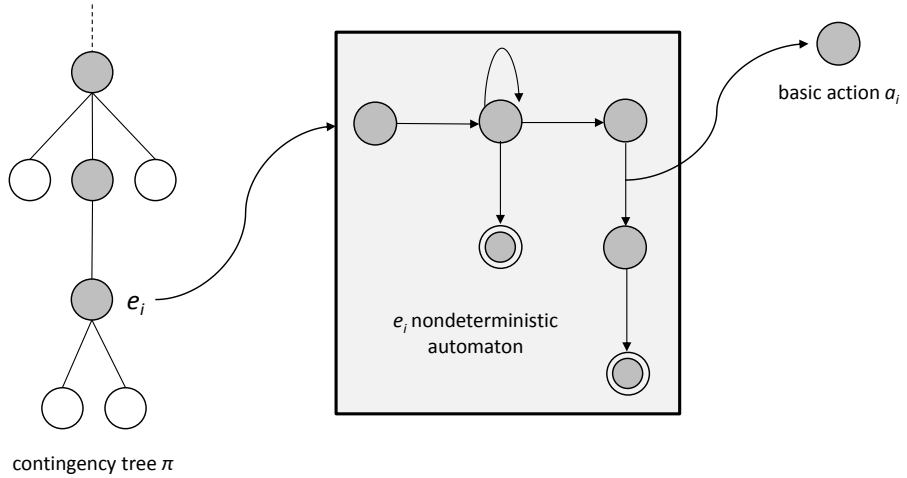

Figure 4.2: Overview of the story generation process.

The system offers two types of user interactions: global and local. In global user interactions, users are able to suggest events to next story chapters, directly interfering in the generation of the contingency trees for the chapters. Such interactions do not provide immediate feedback, but can directly affect the narrative plot. Local user interactions occur during the execution of the nondeterministic automaton and are usually more direct interventions, where users

PUC-Rio - Certificação Digital Nº 1021793/CA

have to choose between the available options in a limited time. In this type of intervention, users can observe the results of their choices immediately, but such interventions only affect the story plot when the decision leads the execution of the nondeterministic automaton to a different final state.

The system has a dynamic behavior with several tasks running in parallel. Figure 4.3 presents an overview of the behavior of the whole system through an activity diagram, where thick black bars indicate parallel activities. Initially, the story generator module creates the first chapter of the story, according to the initial state of the world, while the dramatization module exhibits an overture. In parallel with the dramatization process, the user interaction module is continuously collecting all the suggestions sent by the users (G facts) and combining them with the facts added (F^+) and removed (F^-) from the current state of the world by the story planner. When the end of a chapter is reached, the facts that are more frequently mentioned by the users and that are not inconsistent with the ongoing story are then incorporated into the story plot. During dramatization, if a local decision point is reached, the user interaction module collects opinions of users to decide the course of the dramatization.

The next sub-sections describe in more details the main operating cycle of the three modules of the system and their respective tasks.

4.3.1. Story Generation

The Story Generator module is in charge of creating and updating the story plot according to user interactions. In every operating cycle, a new chapter of the story is generated. The story generation phase main cycle is composed of six steps: (1) Request Reception; (2) Suggestions Retrievement; (3) Chapter Generation; (4) Automaton Transmission; (5) Suggestions Generation; and (6) Suggestions Transmission.

The first step of the story generation phase is triggered by the reception of a request from the story dramatization module, which can be a request for: (1) the first chapter of a new story; (2) the next chapter of an ongoing narrative; or (3) the next basic event of an ongoing chapter. In the case of a request for the first or the next chapter, the story generator module retrieves all the suggestions given by

PUC-Rio - Certificação Digital Nº 1021793/CA

users and starts a new planning process in order to generate the story events for the next chapter using the users' suggestions to guide the development of the narrative. Once the chapter has been generated, a new message containing the nondeterministic automaton for the first basic event of the contingency tree of the chapter is constructed and sent back to the story dramatization module. Then, a new set of possible suggestions, based on the possible outcomes of the story, is created and sent to the user interaction module. Otherwise, if the story generator receives a request for the next basic event of an ongoing chapter, the module only creates and sends back a new message containing the nondeterministic automaton for the next basic event of the contingency tree of the current chapter.

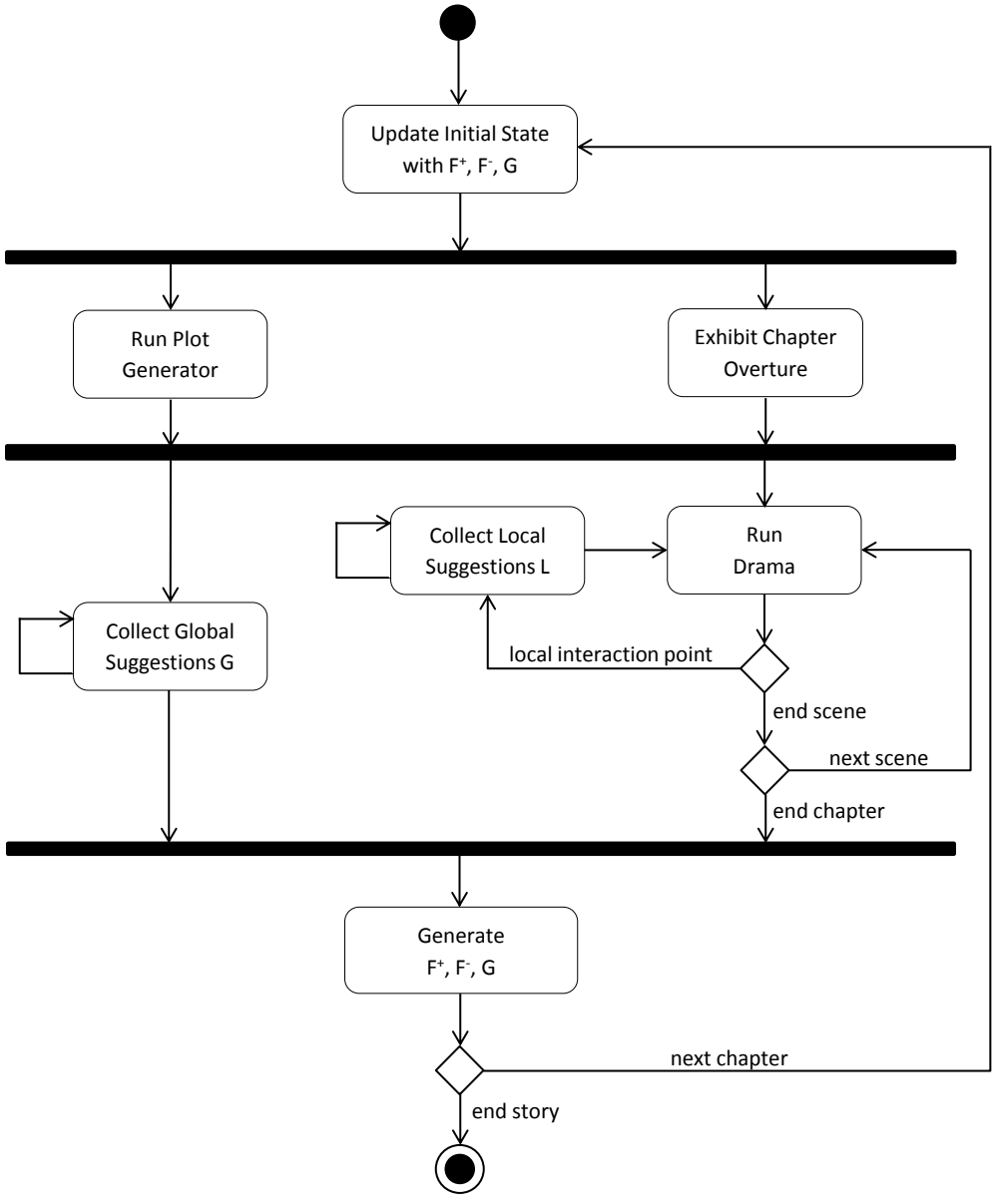

Figure 4.3: Activity diagram of the proposed system.

PUC-Rio - Certificação Digital Nº 1021793/CA

When a new chapter is requested, the story planner must check the coherence of the user suggestions and compute the story events for the next chapter considering the possible consequences of the user interventions in the rest of the story. However, this is not a trivial task and may become excessively time-consuming. In order to synchronize the process of generation and dramatization, stories are strategically divided into chapters. While a chapter is being dramatized, the story planner can already start generating the future chapters. When user interventions are coherent, they are incorporated in the next chapters. In this way, the system keeps the plot generation some steps ahead of the dramatization, so that chapters are continuously generated and dramatized. While the story is being dramatized, the system tries to anticipate the effects of possible user interventions, so that future chapters will be ready when necessary (Camanho et al. 2009). If the system detects that more time is needed for generating the next chapter, a message is sent to the dramatization module in order to extend the duration of the remaining events in the current chapter, as detailed by Doria et al. (2008).

PUC-Rio - Certificação Digital Nº 1021793/CA

4.3.2. User Interaction

The user interaction module is in charge of handling and managing multi-user interactions. The user interaction phase cycle is composed of three steps: (1) Suggestions Reception; (2) Vote Collection; and (3) Selected Suggestion Transmission.

The first step of the user interaction phase is triggered by the reception of interaction suggestions, which can be global suggestions generated by the story generator module, or local interaction options received from the story dramatization module. After parsing the suggestions, the process of collecting votes from users starts. Although there is a set of valid global suggestions, users are free to suggest any event for the story. The user interaction module maintains a list of user's desires, which contains the number of votes for each suggestion, even if it is not in the current set of valid suggestions. In this case, if it appears in the set of valid suggestions during a future chapter, it will already have the amount votes previously accumulated.

Global suggestions are continuously collected by the system. When the story generator module requests the results of user interactions, a new message containing the most voted current global suggestion is created and sent back to the story generator module. Local user interventions occur in parallel with the global user interaction. When the system receives local interaction options from the story dramatization module, it shows and collects user votes for the local decision point. In this type of intervention, users are more restricted and have to choose between the available options in a limited time. When the dramatization module requests the results of the local intervention, a new message containing the most voted current local option is created and sent back to the dramatization module. Meanwhile, the system is still collecting global suggestions for the next chapters.

4.3.3. Story Dramatization

PUC-Rio - Certificação Digital Nº 1021793/CA

The story dramatization is the third process in the main cycle, and is handled by the Story Dramatization module. The dramatization phase cycle is composed of three steps: (1) Automaton Reception; (2) Automaton Execution; and (3) Confirmation Transmission.

The dramatization process is initiated by the story dramatization module after receiving a new automaton with basic actions to perform. The nondeterministic automaton is executed starting from the initial state until it reaches a final state. As previously detailed in Section 2.2, in each automaton, states are described by situations observed in the world, and the transitions between states are associated with basic actions that virtual actors can perform. The basic actions are parsed during the execution of the automaton and delegated to their respective actors. The execution of the automaton progresses to the next state when an actor finishes its performance of a basic action.

In general, there is always a set of states that can be reached after the execution of a basic action and the selection of which transition must occur is based on local user interaction. When starting the dramatization of an action that leads to a decision point, the dramatization module creates a new message containing the local interaction options and sends it to user interaction module. After finishing the execution of the action, the dramatization module retrieves the

most voted option and selects the next action to continue the execution of the automaton. When the execution of the automaton reaches a final state and all the basic actions have been successfully executed, a confirmation message is sent back to the planner controller indicating the end of the dramatization of the current automaton, and requesting a new automaton to continue the narrative based on the final state reached during the dramatization of the current automaton.

4.4. Architecture

The architecture of the proposed video-based interactive storytelling system comprises three main modules: story generator, user interaction and story dramatization.

4.4.1. Story Generator

The story generator module is based on the third version of Logtell, which incorporates the basic temporal modal logic of the first version (Pozzer 2005; 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 were introduced in the third version of the Logtell system. Only a few relevant modifications were made in the original architecture and implementation of Logtell story generation module. The main modification is the introduction of a new module called Planner Controller to manage and centralize the communication of the story generator module with the other modules of the system.

Figure 4.4 shows the architecture of the story generator server. In the architecture, story contexts are stored in a database of contexts (Context Database), where each context contains a description of the genre according to which stories are to be generated, and also the intended initial state specifying characters and the environment at the beginning of the story. The Context Control Module stores and provides real-time access to all data of the Context Database. The Simulation Controller is responsible for informing the dramatization module,

PUC-Rio - Certificação Digital Nº 1021793/CA

at the client side, the next events to be dramatized; receiving interaction requests and incorporating them in the story; selecting viable and hopefully interesting suggestions for users who are intent on performing global interactions; and controlling a number of instances of the Nondeterministic Interactive Plot Generator (NDet-IPG), which is responsible for the generation of the plan to be used as input to the dramatization process. The Chapter Controller is responsible for generating the plot, chapter by chapter, including the treatment of nondeterminism and the parallel processing of multiple valid alternatives at any given moment. The Interface Controller controls the user interventions and centralizes the suggestions made by the users.

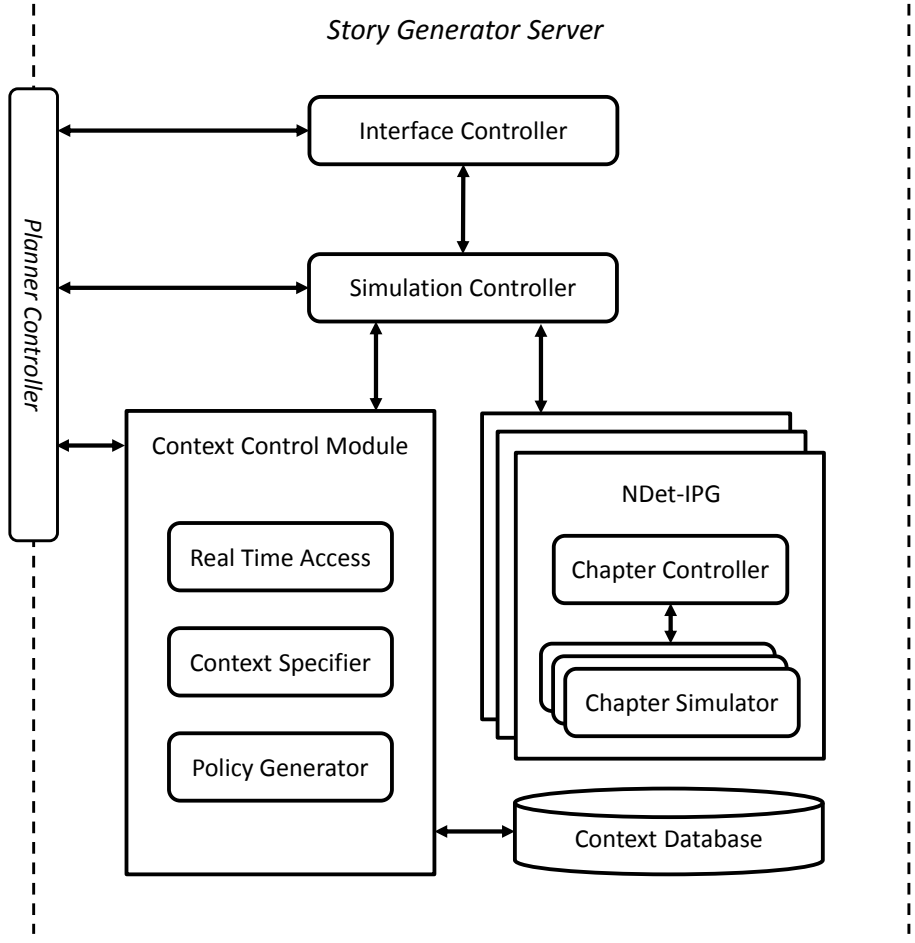

Figure 4.4: The new architecture of the story generator server of Logtell.

More details about the architecture of the Logtell are available in (Pozzer 2005; Ciarlini et al. 2005; Camanho et al. 2009; Silva 2010).

PUC-Rio - Certificação Digital Nº 1021793/CA

4.4.2. User Interaction

The user interaction module of the proposed system is the result of several studies on user interaction mechanisms for interactive storytelling that were conducted during the development of this thesis (Lima et al. 2011B; Lima et al. 2012B; Lima et al. 2012C). The user interaction module works as a multimodal and multi-user interaction server that supports the integration of several interaction mechanisms based on suggestions (Figure 4.5). In this architecture, the Suggestion Manager is the main module that controls the interaction mechanisms, centralizes the users' suggestions and translates them into valid story suggestions. Each interaction mechanism acts as a multi-user server that has its own client interface, allowing several users to be connected in the same interaction network.

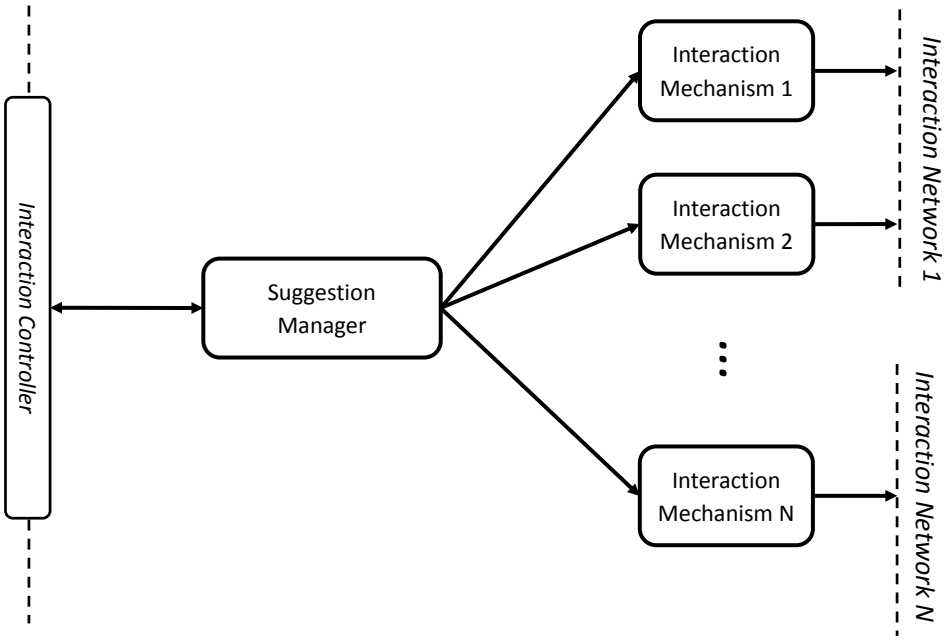

Figure 4.5: Multimodal interaction architecture.

The architecture of the user interaction module integrates two interaction mechanisms: social networks and mobile devices. The first method is based on the idea of using social networks (such as Facebook, Twitter and Google+) as an interaction interface. Three basic ways of interacting with the stories using social networks are: (1) interaction by comments – where users explicitly express their desires through comments in natural language; (2) interaction by preferences – where users express satisfaction or state preferences; and (3) interaction by poll –

PUC-Rio - Certificação Digital Nº 1021793/CA

where a poll is created and users vote in what they want. In the proposed architecture (Figure 4.6), the modules Interaction by Comments, Interaction by Preferences and Interaction by Poll implement their respective methods of social interaction and are responsible for accessing the social networks looking for user interactions and informing the Suggestion Manager about the user's choices. The second interaction mechanism combines the use of mobile devices (such as smartphones and tablets) with natural language to allow users to freely interact with virtual characters by text or speech. In the architecture, the Mobile Interaction module is responsible for receiving and translating user's advices into valid story suggestions, and informing the Suggestion Manager about the user's interventions.

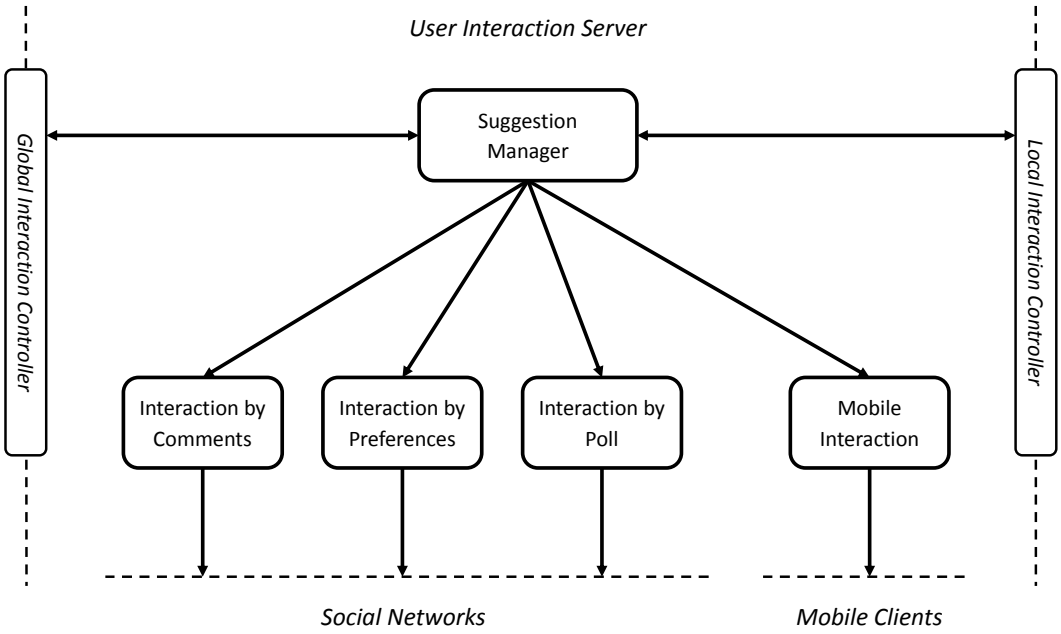

Figure 4.6: The architecture of the user interaction server.

More details about the implementation of the proposed interaction mechanisms are presented in Chapter 7.

4.4.3. Story Dramatization

The architecture of the proposed video-based dramatization module is inspired by the theory of cinematography, and the tasks of the system are assigned

PUC-Rio - Certificação Digital Nº 1021793/CA

to agents that perform the same roles played by the corresponding filmmaking professionals. This approach has been previously used in the 3D dramatization module of the second and the third version of the Logtell system (Lima 2010), and it has proved to be a good strategy to organize and maintain the modules of the dramatization system.

The video-based base dramatization architecture is composed of several cinematography-based agents (Figure 4.7). The agents share the responsibility of interpreting and presenting the narrative events using videos with living actors. In the proposed architecture, the Scriptwriter is the agent responsible for receiving and interpreting the automata of story events generated by story planner. The Director is responsible for controlling the execution of the nondeterministic automata and the dramatization of the basic actions, and for defining the location of the scenes, actors and their roles. The Scene Composer, using real-time compositing techniques, is responsible for combining the visual elements (video sources) that compose the scenes. The Cameraman is responsible for controlling a virtual camera and suggesting the possible shots (e.g. close-up, medium shot, long shot) for the scenes. The Editor agent, using cinematography knowledge of video editing, is responsible for selecting the best shot for the scenes and keeping the temporal and spatial continuity of the film. The Director of Photography is responsible for defining the visual aspect of the narrative, manipulating the illumination and applying lens filters to improve and create the emotional atmosphere of the scenes. A similar task is performed by the Music Director, which is responsible for creating and manipulating the soundtracks of the film to create the adequate mood and atmosphere of each scene. The communication between the Story Dramatization Client and the other modules of the system is handled by the Drama Controller and the Interaction Controller.

The video-based story dramatization module was specially designed to employ video compositing techniques to generate a visual presentation for the story events. However, it also supports the use of static video sequences to represent scenes. Thus, the system can dramatize both prerecorded and dynamically composed video-based interactive narratives. It can also blend both modalities and use static videos to represent complex scenes that cannot be dynamically composed by the system.

PUC-Rio - Certificação Digital Nº 1021793/CA

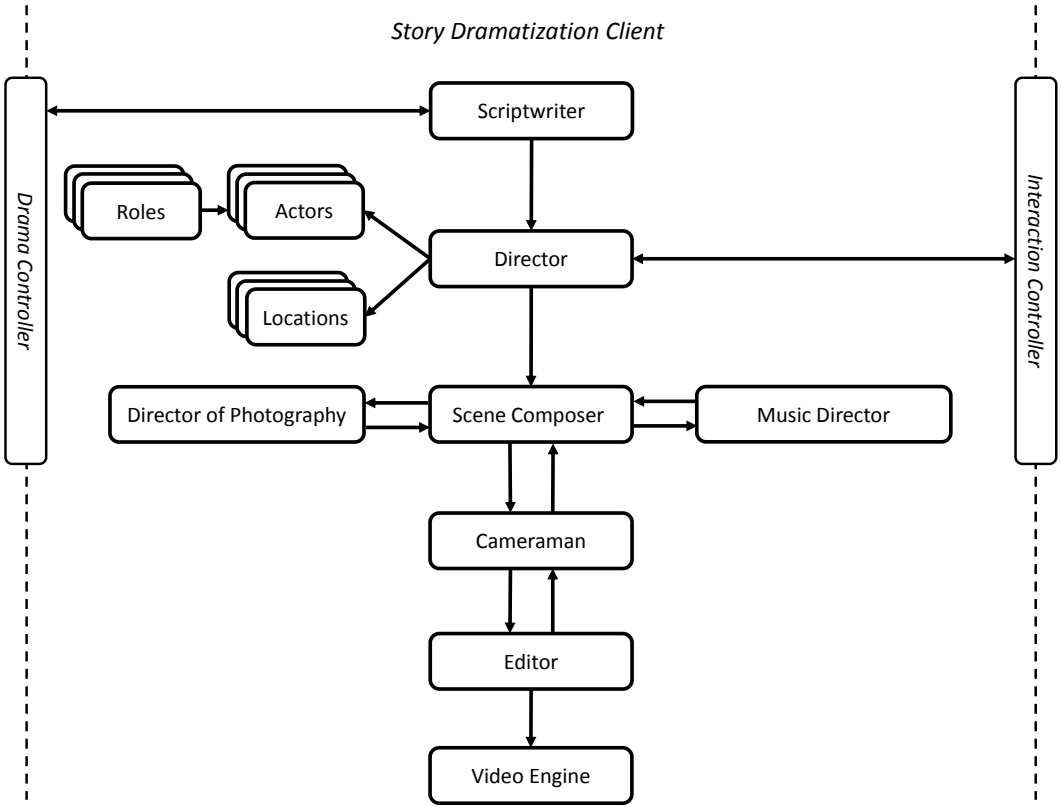

Figure 4.7: The architecture of the video-based story dramatization system.

More details about the implementation of the video-based dramatization system are presented in Chapter 6.

4.5. Conclusion

This chapter presented the general idea of the proposed approach to video-based interactive storytelling, related works, and the architecture of the proposed system together with a high level description of its components and operating cycles. The next chapters present the process for the production of video-based interactive narratives and describe in detail each of the system components.

PUC-Rio - Certificação Digital Nº 1021793/CA