Changing Video Arrangement for Constructing Alternative Stories

A. Piacenza, F. Guerrini, N. Adami, R. Leonardi
DII, University of Brescia, Italy
{firstname.lastname}@ing.unibs.it

J. Teutenberg, J. Porteous, M. Cavazza
School Of Computing, Teesside University, UK
{J.Teutenberg,J.Porteous, M.O.Cavazza}@tees.ac.uk

ABSTRACT

Currently, automatic generation of filmic variants faces a number of key technical issues and thus it usually resorts to the shooting of multiple versions of alternative scenes. However, recent advancements in video analysis has made this objective feasible, though semantic consistency must be somehow preserved. This demo presents a video-based storytelling (VBS) system that successfully integrates video processing with narrative generation by means of a shared semantic description. The novel filmic variants are constructed through a flexible video recombination process that takes advantage of the polysemy of baseline video segments. The short output video clips shown in this demo prove how the generated narratives are semantically consistent while keeping generative power intact.

Categories and Subject Descriptors

I.2.10 [**Artificial Intelligence**]: Vision and Scene Understanding—*Video Analysis*

General Terms

Algorithms

Keywords

Interactive Storytelling, Narrative Modeling

1. INTRODUCTION

Current interactive movie systems rely on the shooting of multiple versions of alternative scenes due to the inability to automatically generate video content in real-time [1]. In contrast, a number of narrative generation systems have opted to use 3D computer graphics [2], whose visual quality however is not comparable to that of video imagery. A key technical challenge for the automatic generation of video content in real-time is the need to generate multiple instances of novel narrative actions from the same baseline video. The video-based storytelling (VBS) system presented in this demo exploits advances in content and semantic analysis of videos to generate multiple instances of new narrative actions from any baseline movie. In particular the system features a novel approach based on the conceptual integration of video processing and narrative generation.

Copyright is held by the author/owner(s).
MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.
ACM 978-1-4503-0616-4/11/11.

Our technical solution recombines video segments to generate short (by filmic standards), completely novel filmic variants of a movie, given an expected plot goal and characters set provided by the user. In this demo, we illustrate the VBS system by showing some alternative narratives of Michael Radford's *The Merchant of Venice* [4].

The paper is organized as follows: Section 2 glances briefly at the technical aspects of the VBS system, with particular emphasis on its inputs and outputs, Section 3 describes what will be actually shown during the demo and Section 4 concludes the paper.

2. SYSTEM DESCRIPTION

The VBS system is schematized in Fig. 1. The core of the VBS system integrates video analysis on the one hand and the generation of narrative variants on the other. The narrative generation module outputs story variants (sequences of narrative actions) in response to user input (see Section 3). When a new plot variant is constructed its consistency cannot be preserved by means of low-level video analysis alone, hence our approach features high level reasoning through a Planning AI.

The integration of video analysis and narrative generation is achieved via the use of a shared semantic representation to enable communication between low-level video content and the model of the narrative domain. In particular, a video is first segmented into shots. They are annotated by the video system with several semantic tags that describe intermediate level concepts as specified by an agreed vocabulary, and

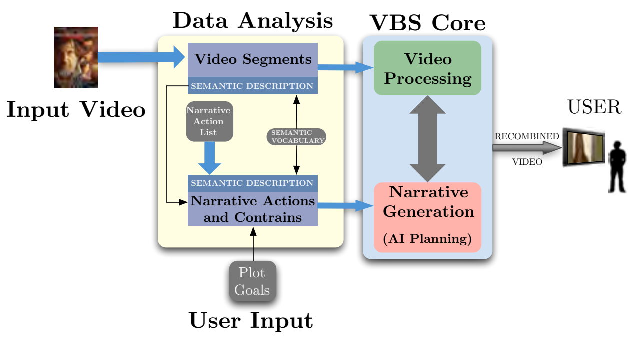

Figure 1: VBS system overview: at its core, it integrates video processing and narrative generation by allowing the exchange of semantic information drawn from a common vocabulary.

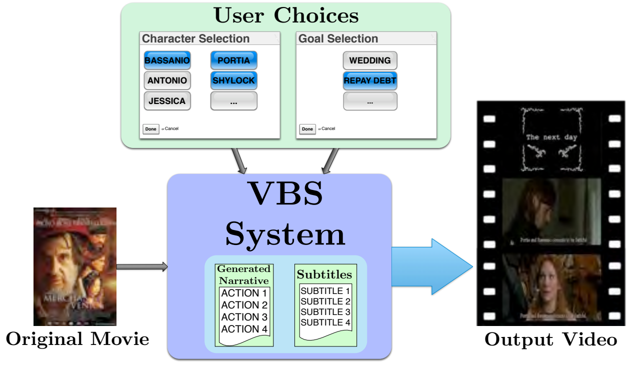

Figure 2: Typical demo workflow: the user inputs its narrative choices; the VBS system recombines the shots of the baseline movie and adds subtitles and panels to reflect those choices in the output video.

the planner provides the bridge between them and the narrative actions through an appropriate mapping as part of its domain model. Shot semantic tagging is commonly adopted for a variety of tasks, e.g. for emotional categorization. The tags extracted in this work are: the characters' name and mood (positive, negative or neutral), the field of vision of the camera (a ternary value) and a triplet of environmental binary descriptors: time of the day, crowd presence and location (indoor or outdoor). For more details, refer to [3].

Individual narrative actions are then described by a recombination of shots appropriately selected according to their semantic description. This is a powerful result since it provides a complete decoupling of the narrative model and the baseline video content, which is only described by its semantics. Another innovative feature of the system is that the output of video processing is able to constrain the narrative generation process by preventing the creation of narrative variants that feature individual actions that can't be presented given the content of the baseline video.

Last, to attach a different high-level meaning to a given narrative action, represented by a certain sequence of shots, the video is stripped of its audio portion and an appropriate subtitle is added for each action. Moreover, to further improve the overall narrative understanding by the user, a text panel detailing the change of context is interposed between a narrative action and the next when any of the environmental tags changes.

3. DEMO DESCRIPTION

This Section describes the functionality of the proposed VBS system demo, how the user can drive its flow and what output is expected. At the beginning of a standard demo session, the user is presented with a list of choices that allows him to influence the generation of the filmic variants. In particular, the user can select the characters who wants to be involved in the narration, chosen among the main characters of the original movie, and the goals of the new plot. Some plot goals may be unavailable for certain characters selection, e.g. for insufficient video content resources. The user input forces the narrative generation module to build a narrative path that satisfies the input requirements and then the user is presented with the actual output of the system, i.e. a recombination of video segments taken from the baseline video. The demo workflow is depicted in Fig. 2. Fig. 3 shows an excerpt of a recombined video output.

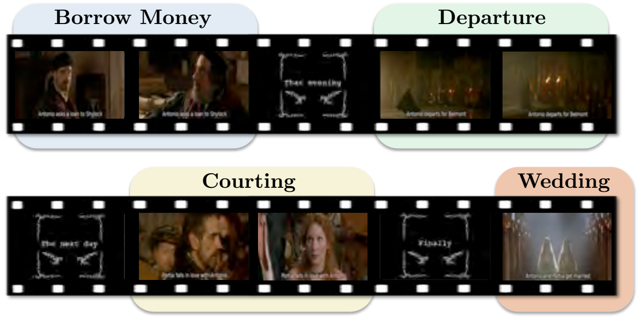

Figure 3: Excerpt of a recombined video output for the case of a wedding goal between Antonio and Portia: in this case, Antonio asks Shylock for a loan in order to make the trip and adequately court Portia and later they are eventually married.

Another feature of our VBS system is that, at any time during the playback, the user can revisit the choices that he has made, thus changing dynamically the evolution of the story. This can be performed at run-time, forcing the AI Planning to take into account what has already been shown in the playback and changing only the future actions. However, dynamic radical changes of the narrative or characters are not allowed. In those cases, as an alternative the planner can be rewinded and asked to construct a whole new narrative back from scratch.

4. CONCLUSIONS

This demo shows the functionalities of the video content-based storytelling system for the automatic generation of novel filmic variants. To better highlight the good performance in terms of both visual presentation and narrative generation power, actual video output is generated and then shown to the audience. As soon as the user provides a new set of input parameters, the system can generate and playback the requested narrative almost in real-time.

5. ACKNOWLEDGEMENTS

This work has been founded (in part) by the European Commission under grant agreement IRIS (FP7-ICT-231824).

6. REFERENCES

[1] P. Cesar, D. Bulterman, and L. Soares. Human-centered television: directions in interactive television research. *ACM Trans. on Multimedia Computing, Communication and Applications*, 4(4), 2008.

[2] M. Mateas and A. Stern. Structuring Content in the Façade Interactive Drama Architecture. In *Proc. 1st Conf on AI and Interactive Digital Entertainment (AIIDE)*, Marina Del Rey, USA, 2005.

[3] A. Piacenza, F. Guerrini, N. Adami, R. Leonardi, J. Porteous, J. Teutenberg, and M. Cavazza. Generating story variants with constrained video recombination. In *Proc. ACM Multimedia Conf.*, Scottsdale, AZ, USA, 2011.

[4] M. Radford. MGM Home Ent. (Europe) Ltd., 2004. The Merchant of Venice (film adaptation).