Towards a Non-Linear Narrative Construction

Vidya Setlur, David A. Shamma, Kristian Hammond, Sanjay Sood
Intelligent Information Laboratory, Northwestern University
1890 Maple Avenue, Evanston IL 60201 USA
{vidya, ayman, hammond, sood}@cs.northwestern.edu

ABSTRACT

This article describes the implementation of a system that 'imagines' while a movie is being played by finding associations in the movie's content and presenting them to the viewer. This related information, in the form of images and movie clips, helps enhance the viewer's experience in a new immerse environment.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems

General Terms

Design, Human Factors

Keywords

Imagination, Narrative, Image Association, Movies

1. INTRODUCTION

Being part of an engaging activity, such as watching an artist paint, viewing a play or even listening to poetry initiates creative thought; in turn instigating imagination along various tangents. We chose to use movies as a form of such engagement because they provide a similar environment devoid of superfluous complexity and abstractness. The focus of our research is on developing a system to find related context as the movie is playing, and providing a platform for user controlled non-linear narrative.

2. FRAMEWORK

Movies are rich in imagery content and serve as excellent sources for video as well as audio information. We use dialogs from the movie as a basis for the system to retrieve related images and clips. They are pulled from the union of the DVD's subtitle track and closed captioning. Both tracks are important as closed captioning does not provide accurate line-by-line dialog but does provide audio cues (speaking Italian, band playing, etc.). The dialog events thus obtained, contain the dialogs and timestamps, which are relative to the running time of the movie.

The textual content present in these dialog events is used to extract URLs of images from the Internet. The text is initially filtered out by eliminating commonly used English stop words. In order to improve the coherency, terms are categorized. For instance, images of actors portrayed in different movie roles are pulled off the character names. The search could be further refined with the use of corresponding synonyms with each term during the query. We hence obtain images that associate with the terms, like a Brooklyn Bridge photo for the term 'New York'.

Clip to clip matching is done using simple term frequency analysis. Once a candidate list of term similar clips has been found, the successful candidate is chosen though several simple condition checks that take into account the run length of the current playing (source) scene, the run length of the candidate (target) scene, and the distance from the source to the target. We found that most clips will match neighboring clips, but more interestingly they match to the last scene of the movie. Likely this artifact is a sign of a good movie.

We chose a three by three array of monitors, a.k.a. 'The Wall', to present our environment. A simple set of rules is used to control how content is displayed on the neighboring screens as the movie plays in the center tile.

3. FUTURE WORK

The architecture provides a foundation for developing richer semantics with a deeper understanding of story boarding, character revelation and mood determination. When expanded across several movies, we plan on a user controlled centroid where the source clip can be altered on the fly by the neighboring content. This will allow the user to follow his own narrative in the imagination space. In turn, we hope that it will facilitate a richer environment for the task engaged viewer.

IUI'03, January 12–15, 2003, Miami, Florida, USA.
ACM 1-58113-586-6/03/0001.