UniS

Navigating Stories in Films

Andrew Salway and Yan Xu

School of Electronics and Physical Sciences

Department of Computing

CS-05-04

This work was supported by EPSRC grant GR/R67194/01 – TIWO: Television in Words.

THE QUEEN'S ANNIVERSARY PRIZES
FOR HIGHER AND FURTHER EDUCATION
2002

UniS
University of Surrey

Navigating Stories in Films

Abstract

This report describes the transformation of feature films into hypervideo by representing their story structures using plot unties. Plot units represent cause-effect relationships between characters' affect states and the events in a story. We use plot units to structure hypervideo links between intervals of video data. We have manually analysed two full-length feature films in terms of plot units. A system was developed to store and edit data about plot units and to navigate films by following hypervideo links based on the structure of the plot units. The effect of plot units was evaluated by having users complete question-answering tasks using the system. Results suggest that when using the links subjects gave better answers to questions about the film. In questionnaire feedback, users supported using this kind of hypervideo for watching and re-watching films on future domestic video players.

1. Introduction

Hypervideo offers new ways to watch and to interact with video data, but little research has been done into how hypervideo can be used to watch and re-watch feature films. We are interested in this because we see hypervideo as a next step in the sequence of interactivity from cinemas, to television, to video players, and to DVDs. Recent developments in DVD and interactive television technology, such as Blu-ray [1], mean that mass delivery of hypervideo is becoming feasible. Film distributors should be interested in hypervideo as a way of adding-value to film purchases, to encourage multiple viewing of a film, whether on DVD or via video-on-demand. To be engaging and to facilitate intuitive interaction, it is important that hypervideo structures reflect story structures.

Ours is the first research to transform complete full-length feature films into hypervideo based on detailed story structures. Other important features of the work are the use of typed hypervideo links to describe different kinds of relationships between affect states in plot units, and evaluation with question answering tasks. Section 2 reviews previous research in hypervideo. Section 3 introduces the theory of plot units and describes the analysis of two feature films. Section 4 presents the NAFI system which implements hypervideo navigation by plot units. Section 5 reports our evaluation of this approach and Section 6 closes the paper with a brief discussion.

2. Hypervideo

Hypermedia consists of nodes (pieces of media) and links between nodes that users follow to navigate the media. It is intended to be an intuitive way to create, share and access information. Hypermedia structures are said to reflect cognitive structures. Typed links enhance hypermedia structures by giving semantics to the links [2]. Typed links help a user to understand what to expect if they follow a link, and can be exploited by machines for analysing and generating hypermedia. Whilst the notion of

hyper*text* is quite well understood, the idea of *hypervideo* is still developing and has been interpreted in various ways.

An early proposal for hypervideo considered it to be a new kind of cinematic experience and the researchers discussed its aesthetic properties [3]. In their system a filmmaker authors a set of possible narrative sequences in hypervideo material and the viewer chooses which sequences to watch. In [4] a generic data model for hypervideo represented semantic associations between video entities, i.e. regions in consecutive video frames, and other logical video abstractions. The focus was on semantic associations between entities, e.g. X is-a Y, rather than on story structures. The use of hypervideo for interactive training has been demonstrated in [5]. Recently a system to support object-based hypervideo authoring has been proposed [6], and issues to do with hypervideo transmitted via interactive television have been discussed [7]. Without mentioning hypervideo explicitly, previous work has used conceptual representations of story structures to facilitate access to digitised films. Semantic networks were used to support browsing between scenes of a film, but the emphasis was on objects and class hierarchies and only 5 scenes of one movie were analysed [8]. In [9] film stories were analysed with plot units and story threads, but the analysis was at a relatively low-level of detail appropriate for the visualisation of story structure for video navigation.

3. Representing a story with plot units

Plot units are a conceptual structure to represent stories and were used for narrative summarization in computer simulations and psychological experiments [10]. The kinds of narrative worked with were short written fictive stories. From this perspective, the central parts of a story are the characters' *affect states* (emotional reactions to events, i.e. positive and negative reactions, and unspecified mental states) and links. Links, including causal links, are between affect states. Previously, these structures were found to be effective for story summarization and also highlighted repeating structures that were common between stories.

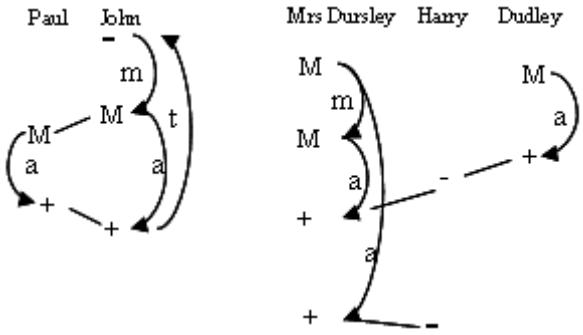


Figure 1. Example Plot Units

A primitive plot unit comprises two affect states and a causal link between them. Primitive plot units are combined to represent sequences of reactions and mental states, including interactions between characters with cross-links. Consider a short story based on an example from [10]. John tries to start his car but it is broken so he asks Paul for help; Paul fixes the car; John thanks Paul and drives to work. The plot unit representation for this story, Figure 1, shows John's negative reaction (-) to his car being broken, and a motivation link to a mental state that leads him to ask

Paul for help (m→ M). Paul agrees to help (M) and there is an actualisation link to the event of fixing the car which is positive for both characters (a → +). The termination link (t) indicates that the car fixing event resolves John's problem. Figure 1 also shows part of one scene from *Harry Potter and the Philosopher's Stone* in which Mrs Dursley and her son Dudley wake up Harry against his wishes and force him to cook their breakfast. Note that links can hold between affect states depicted in different scenes: these links are probably more relevant for hypervideo.

We analysed the plot units of two feature films, including three affect states (+ , –, M), four types of causal link between affect states (actualization, motivation, termination, equivalence) and cross-links between characters. This manual task took about 35 person hours per film. In general *The Pelican Brief* (a crime thriller) has more primitive plot units, and in particular, more m-links and cross links, than *The Matrix Revolutions* (a sci-fi action thriller), Table 1.

Table 1. Data about plot units in films

	The Matrix Revolutions (129 minutes)	The Pelican Brief (141 minutes)
plot units	98	137
affect states	151	201
a-links	54	65
m-links	19	38
t-links	16	25
e-links	9	9
cross-links	67	108

4. NAFI: A system for navigating films

We developed the NAvigating FIlms system (NAFI) to investigate hypervideo based on the story structures of films. It lets researchers enter data about story structures, e.g. plot units, to structure films as hypervideo. Users navigate the film with hypervideo links and the system records their actions for subsequent analysis by researchers. Figure 2 shows a simplified version of the NAFI class diagram. NAFI stores data about films, scenes, the events in scenes and the connections between events (both within a scene and across scenes). A plot unit is associated with 1 or 2 scenes (for intra-scene links and inter-scene links); a scene may have many plot units. A plot unit is made of 2 affect states and a causal link: there are 3 types of affect state and 4 types of link. Affect states and plot units belong to a character and cross-links join affect states belonging to different characters. Plot units are associated with video data by the start and end time of scenes, and the timepoints of affect states.

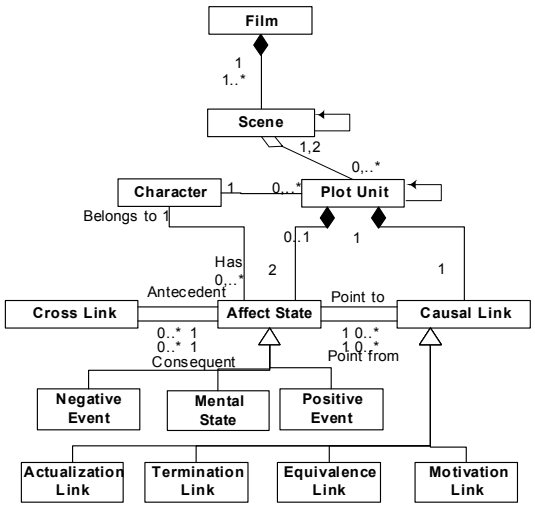


Figure 2. Simplified NAFI Class Diagram

When the plot units are associated with video data then the links between affect states become, in effect, typed hypervideo links. To enable hypervideo browsing we need to show users what links are available from the current video interval, and we need to

help them understand about the different kinds of link. The main NAFI interface, Figure 3, includes a video-player (top-left), keyframes for all scenes (bottom-half) and information about available links (top-right). The user clicks the "Highlight Scenes" button to highlight the keyframes of all scenes containing affect states related to the affect states of the current scene. By choosing a related scene the user can see details of the links and choose an affect state to view – the video is then moved to the timepoint of that affect state. In Figure 3 the current scene depicts a murder victim and the highlighted scenes depict other characters' affect states that explain why and how the murder took place.

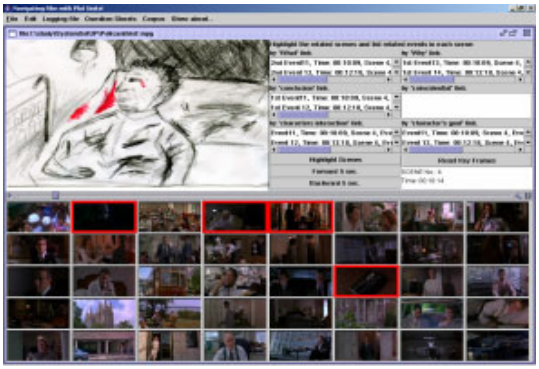


Figure 3. Interface for Hypervideo Navigation

5. Evaluation

We wanted to test the hypothesis that structuring films with hypervideo based on plot units helps users to find answers to questions about film content. Subjects used NAFI to find answers to questions about the story of a film with/without hypervideo links enabled. We chose *The Pelican Brief* because it has a lot of 'cause-effect' links between human motivations and physical events. In a previous exercise 6 volunteers were asked to watch the film and think of questions they had as they watched. Ten of these questions were chosen for the evaluation, Table 2.

Table 2: Questions used in the evaluation

WHO Questions
Q1. "Who killed the ill old man in Scene 4?"
Q2. "Who are implicated as suspects of the assassinations in the Pelican Brief?"
Q3. "Who was the car bomb in Scene 10 intended to kill?"
Q4. "Who wanted to assassinate Darby Shaw and Gray Grantham?"
WHY Questions
Q5. "Why were the people killed in Scene 4?"
Q6. "Why is it called the Pelican Brief?"
HOW Questions
Q7. "How is the President related to the assassination of the two judges?"
Q8. "How was Curtis Morgan connected to the assassinations?"
WHAT Questions
Q9. "What do the two dead judges have in common?"
Q10. "What was in Morgan's safe deposit box?"

Subjects were final year undergraduate students and PhD students in Computing; 25% had seen the film before (but some years ago). Each subject was given a 20-minute tutorial in using NAFI. 13 subjects answered 3 questions with hypervideo enabled, and 3 questions without hypervideo. Another 7 subjects answered the same questions, but with the use of hypervideo reversed. Subjects were given a scene and a time-point for each question to start with, and a time limit of 5 minutes to browse for an answer. The subjects' actions and the timing of their actions were recorded in a system log. The system log recorded when they started and finished answering each question, and how many and what kinds of links they followed, and when they used the keyframes to browse directly to a scene.

Our analysis of the subjects' behaviour has so far focused on the length of time it took them to answer questions, and on the quality of their answers

which we scored from 0-10. The summary of the results shows little difference in the time taken to answer questions but suggests that subjects did give better answers when using hypervideo, Table 3a. The results are also categorised by question type, Table 3b, and by individual question, Table 3c. Considering the different types of questions, it seems the biggest effect was for 'Why?' questions. Question 2, 5, 8 and 10 seem to stand out as being answered better with the use of hypervideo. More analysis is required to properly test these observations, and also to consider different patterns of navigation behaviour in subjects. Separate questionnaire feedback included encouraging comments from subjects who would like this kind of hypervideo in domestic video-players.

6. Closing remarks

We have begun to investigate the transformation of feature films into hypervideo. Plot units provide a theoretically-motivated way to structure hypervideo according to story structure. Preliminary evaluation suggests that hypervideo links help users to understand some aspects of a film better, and informal feedback suggests that they found the navigation experience to be engaging and enjoyable. The analysis of video data into plot units is currently a manual task, but even as such there may be a business case to be made when you consider the cost relative to the budget of major films. In our ongoing work we have more data to analyse about subjects' navigation patterns, and more factors to consider in testing our hypothesis. Also, we are considering ways to partially automate the generation of plot unit data, and are encouraged by recent automatic analyses of story structure [11, 12].

Table 3a: Summary of results

	Average time to answer	Average quality of answer (0-10)
With Hypervideo	171	5.09
Without Hypervideo	168	4.55

Table 3b: Results by question type

Question Type	Average time to answer		Average Score	
	With H'video	Without H'video	With H'video	Without H'video
WHO?	160	153	5.8	5.2
WHY?	195	191	3.6	2.5
HOW?	159	188	5.8	5.1
WHAT?	179	139	4.7	5.3

Table 3c: Results by individual question

Question	Average time to answer		Average Score	
	With H'video	Without H'video	With H'video	Without H'video
Q1	109	164	4.0	6.0
Q2	172	191	6.0	2.8
Q3	146	145	7.0	6.3
Q4	211	112	6.2	5.5
Q5	178	171	4.9	2.5
Q6	231	241	0.8	2.5
Q7	135	198	5.2	6.1
Q8	225	167	7.5	3.0
Q9	185	159	2.0	4.8
Q10	171	124	8.0	5.8

References

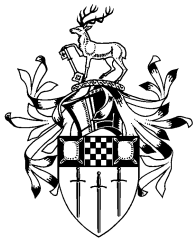
1. www.blu-ray.com/info/, accessed on 21 Dec 2004.
2. K.M. Hansen, C. Yndigegn, and K. Gronbaek, "Dynamic Use of Digital Library Material – Supporting Users with Types Links in Open Hypermedia", *Lecture Notes in Computer Science*, LNCS 1696, 1999, pp. 254-276.
3. N. Sawhney, D. Balcom, and I. Smith, "Authoring and Navigating Video in Space and Time", *IEEE Multimedia* Oct-Dec 1997, pp. 30-39.
4. H.T. Jiang, and A.K. Elmagarmid, "Spatial and Temporal Content-Based Access to Hypervideo Databases", *VLDB Journal* 7 (4), 1998, pp. 226-238.
5. J. Doherty, A. Girgensohn, J. Helfman, F. Shipman, and L. Wilcox, "Detail-on-Demand Hypervideo", *Procs. ACM Multimedia 2003*, pp. 600-601.

6. H.-B. Chang, H.-H. Hsu, Y.-C. Liao, T.K. Shih, and C.-T. Tng, “An Object-Based Hypervideo Authoring System”, *Procs. IEEE Int. Conf. on Multimedia and Expo, ICME 2004*.
7. M. Finke, and D. Balfanz, “A Reference Architecture Supporting Hypervideo Content for ITV and the Internet Domain”, *Computers and Graphics* 28 (2), 2004, pp. 179-91.
8. V. Roth, “Content-based retrieval from digital video”, *Image and Vision Computing* 17, 1999, pp. 531-540.
9. R. Allen, and J. Acheson, “Browsing the Structure of Multimedia Stories”, *Procs 5th ACM Int. Conf. Digital Libraries*, 2000, pp. 11-18.
10. W.G. Lehnert, “Plot Units and Narrative Summarization”, *Cognitive Science* 4, 1981, pp. 293-331.
11. B. Adams, C. Dorai, and S. Venkatesh, “Towards Automatic Extraction of Expressive Elements for Motion Pictures: Tempo”, *IEEE Trans. Multimedia* 4 (4), 2002, pp. 472-481.
12. A. Salway, and M. Graham, “Extracting Information about Emotions in Films”, *Procs. ACM Multimedia 2003*, pp. 299-302.

Department of Computing
University of Surrey
Guildford, Surrey
GU2 7XH UK

Tel: +44 (0) 1483 683133
Fax: +44 (0) 1483 686051
E-mail: a.salway@surrey.ac.uk
www.surrey.ac.uk

UniS
University of Surrey



UniS

University of Surrey