New Ways to Find and Watch Films

Dr Andrew Salway
a.salway@surrey.ac.uk
Department of Computing, University of Surrey, United Kingdom

2nd Congress of the Technological Age
Morelia, 1st April 2006
A (Very) Brief History of Film

[Main source www.wikipedia.org]

- **1890’s**
  - Edison’s Kinetograph
  - Lumières’ Cinematograph

- **1900’s**
  - 10,000 movie theatres in US by 1908
  - Use of films to tell stories

- **1920’s**
  - Added sound – speech, music, sound effects
  - Well-developed ‘language of film’
A (Very) Brief History of Film

[Main source www.wikipedia.org]

• **1940’s-1960’s**
  - Colour films
  - Television

• **1970’s**
  - Video Cassette Recorders
  - Subtitling for the deaf and hard of hearing

• **1990’s**
  - MPEG Standards and other video coding formats
    ➔ Digital video on PC’s
    ➔ DVD Players
A (Very) Brief History of Film

MPEG STANDARDS

Moving Pictures Expert Group

- **MPEG-1 (1992):** compression of video data to 1.5MBits/s; store on CD-ROM; Internet delivery
- **MPEG-2 (1996):** high quality; “efficient transmission over error-prone delivery systems” (digital TV; DVDs)
- **MPEG-4 (1999):** further compression (mobile phones); change from video as sequence of frames to video as a composite of audio-visual objects
- **MPEG-7 (2001):** “Multimedia Content Description Interface”; Metadata – ‘data about data’
These Days...

Now Playing on TiVo

- The Apprentice Thu 3/25
- Alias Sun 3/28
- Survivor Mon 3/29
- Stanford vs. Arizona Tue 3/30
- The Real World Tue 3/30
- American Idol (2) Wed 3/31
- Six Feet Under Sun 4/4
- Emeril Live (4) Wed 4/7

http://tivo.com/resources/images/tv_tivo_fulldpi.jpg
These Days...

http://moto-rikup.seesaa.net/image/m_ipod.jpg
Hollywood to offer Internet movie downloads
by Derek Sooman on Sun 26 Mar 2006, 03:00 AM

In seeking to combat piracy, Hollywood studios are to offer Internet movie downloads at the same time that major blockbusters enter retail stores in DVD form. Universal Pictures International plans to introduce a "download to own" service in Britain in partnership with LoveFilm, an online video rental company. The industry, which has been burned by illegal copying of DVDs and trading of movies on the Internet, is being forced to embrace digital distribution.

"Download-to-own has the potential to completely revolutionize the way people watch movies," said Peter Smith, president of Universal Pictures International, part of the NBC Universal division of General Electric. "The entertainment industry is changing rapidly, with the introduction of new delivery channels to consumers and an emphasis on instant access."

These Days...
Audio Description

• Subtitles allow the deaf and hard-of-hearing to read what they cannot hear: **audio description** allows the blind and visually-impaired to hear descriptions of what they cannot see.

• In between dialogue the describer gives essential details about on-screen scenes and events, and about characters’ actions, appearances, gestures and expressions.

• Audio description is increasingly available internationally on television, in cinema and on VHS/DVD releases.

• In UK 3-4 major film releases a week are described (with a back catalogue approaching 1000 films) – 200+ cinemas provide audio description; RNIB campaigning for 50% of broadcast television to be described.
[11.55] Laughing, Jan falls back into her seat as the jeep overtakes the line of the lorries.
[12.08] The jeep has hit a mine.
[12.09] Hanna jumps from the lorry.
[12.20] Desperately she runs towards the mangled jeep.
[12.27] Soldiers try to stop her.
[12.31] She struggles with the soldier who grabs hold of her firmly.
[12.35] He lifts her bodily from the ground, holding her tightly in his arms.
These Days...
These Days...

http://www.imdb.com
These Days...
These Days...

http://medialit.med.sc.edu/chapterlist.jpg
What Has Changed?

• Affordances for Storytelling in Films

• Accessibility of Films

• Interactivity with Films
Ideas for a Future Video Player

• “Find me scenes showing X and Y”
• “Make me a summary of the film”
• “Find me films with a similar story to this”
• “Why did that happen?”
• “Give me background information about…”
• “How is this film genre different from that one”
VIDEO DATA
Frames made of pixel, 25fps
Digitised soundwave

“Semantic Gap”
To put it another way...

- The bit stream that is the text data, image data, video data, audio data describes the expression of the media, but not the content

“every symbolic object consists of two interacting components, one of which is situated on the plane of expression while the other is situated on the plane of content”

“difference between what constitutes the objects themselves (media objects) and what is actually being communicated”

Smoliar and Wilcox (1997)
What happens in the mind/brain of a viewer when watching a film?
Film and Narrative

Film

- moving images - *filmed and special effects*
- sound - *speech, music, sound effects*
- camera techniques – *pan, zoom*
- editing techniques – *montage*

*Films often tell stories...*
Story

Stories involve chains of events in cause-effect relationships occurring in space and time, where the agents of cause-effect are characters with goals, beliefs and emotions.

(Bordwell and Thompson 1997)
“More than reconstructed timelines and inventories of existents, … *interpreters are called upon to live out complex blends of cognitive and imaginative response, encompassing sympathy, the drawing of causal inference, identification, evaluation, suspense*”

(Herman 2002)
VIDEO DATA
Frames made of pixel, 25fps
Digitised soundwave

"Semantic Gap"

 Need to store metadata - "data about data"
First Question

How do we model films, in particular the stories told by films: in other words, what metadata do we need for each film?
### Simple Video Data Model

<table>
<thead>
<tr>
<th>FILM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title:</td>
</tr>
<tr>
<td>Director:</td>
</tr>
<tr>
<td>Year:</td>
</tr>
<tr>
<td>Genre:</td>
</tr>
<tr>
<td>Actors:</td>
</tr>
</tbody>
</table>
Generic Video Data Models

Attributes include descriptions of events and objects, e.g. kick(Andrew, ball)
Figure 1. The conceptual schema graph modeling movie informative structure
Roth (1999)

Figure 3: An excerpt of the semantic network representation of scenes from the motion picture “True Lies”
TIWO Project

• **AIM:** to develop a computational understanding of narrative in multimedia systems

• **Research Scenario:** audio description for television and film → technology to assist the production of audio description, and technology that reuses audio description for video metadata

TIWO (Television in Words) was funded by UK government 2002-2005. The work continues with 8 students (BSc, MSc and PhD) working with me.

**Partners:** BBC (British Broadcasting Corporation), RNIB (Royal National Institute of the Blind), ITFC and Softel
Plot Units to Represent Stories

- As part of TIWO Xu (2006) investigated the use of Plot Units (Lehnert 1978) to represent the story structures of full-length feature films

- Plot units describe characters’ affect states (emotional reactions to events and goals) and relate them to events and other characters’ affect states

- Xu (2006) developed and evaluated the NAFI hypervideo browsing system where film browsing is based on plot unit structure
Plot unit representation for one scene from *Harry Potter and the Philosopher’s Stone*. Mrs Dursley and her son Dudley wake Harry up because she wants him to cook breakfast.
NAFI: Navigating Films

<table>
<thead>
<tr>
<th>Highlight the related scenes and list related events in each scene:</th>
<th>by 'What' link</th>
<th>by 'Why' link</th>
</tr>
</thead>
<tbody>
<tr>
<td>by 'conclusion' link.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>by 'coincidental' link.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>by 'characters interaction' link.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>by 'character's goal' link.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Highlight Scenes</th>
<th>Reset Key Frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward 5 sec.</td>
<td>SCENE No.: 4</td>
</tr>
<tr>
<td>Backward 5 sec.</td>
<td>Time: 00:10:14</td>
</tr>
</tbody>
</table>
NAFI: Navigating Films

Highlight the related scenes and list related events in each scene:
by 'What' link
by 'Why' link.

1st Event 6, Time: 00:06:55, Scene 2, E
1st Event 5, Time: 00:06:18, Scene 2, E

by 'conclusion' link.
by 'coincidental' link

by 'characters interaction' link.
by 'character's goal' link

Highlight Scenes
Forward 5 sec.
Backward 5 sec.
Reset Key Frames

Event 6, Time: 00:06:55, Scene 2, E
Event 5, Time: 00:06:18, Scene 2, E

SCENE No.: 2
Time: 00:07:15
Plot Units to Represent Stories

• A hand-made plot unit description for two full-length feature films showed that important aspects of a film’s story could be described methodically

• Evaluation of NAFI suggested that users enjoyed the hypervideo feature, and it helped them to understand the story

• Could be part of future Blu-ray / HD DVD discs?

**BUT** – it took about 50 hours to describe plot units for one film
Second Question

How do we analyse video data in order to extract representations of stories: in other words, how do we create this metadata automatically?
Analysing Video Content

• Recall, video data comprises:
  • Moving image – pixels
  • Speech, music, sound effects – digitised soundwaves

• During the last five years, a lot of research has been done on automatic video content analysis: while this has proved successful for some applications we are a long way from being able to extract representations of films’ stories from video data

• To get an idea of what can be achieved with automatic analysis of video data, see the TRECVID conference and systems such as CMU’s Informedia and DCU’s Fishclar.
Use of Collateral Text

- In TIWO, so far we have concentrated on extracting information about films’ stories from collateral texts:
  - audio description
  - film scripts
  - plot summaries
  - subtitles/closed captions
[11.55] Laughing, Jan falls back into her seat as the jeep overtakes the line of the lorries.
[12.08] The jeep has hit a mine.
[12.09] Hanna jumps from the lorry.
[12.20] Desperately she runs towards the mangled jeep.
[12.27] Soldiers try to stop her.
[12.31] She struggles with the soldier who grabs hold of her firmly.
[12.35] He lifts her bodily from the ground, holding her tightly in his arms.
Extracting Information about Emotions

METHOD

• Used 22 ‘emotion types’ (Ortony et al. 1988)
• Created lists of 600+ ‘emotion tokens’ for each type; used WordNet on grammatical variants of type names

  **FEAR:** fearful, nervously, desperately, terrified…
  **JOY:** joy, contented, cheerful, delighted…
  **HOPE:** hope, expectation…

• Map emotion tokens \(\rightarrow\) emotion types

  Salway and Graham (2003)
Emotions Extracted for *Captain Correlli’s Mandolin*

52 tokens of 8 emotion types
Emotions Extracted for *Captain Correlli’s Mandolin*

**52 tokens of 8 emotion types**

15-20 minutes: Pelagria’s betrothal to Madras
20-30 minutes: invasion of the island
68-74 minutes: Pelagria and Correlli’s growing relationship
92-95 minutes: German soldiers disarm Italians
Emotions Extracted for *The Postman*
Some Common Events in Films

• Vassiliou (2006) analysed a corpus of audio description (45 films, 400,000 words) and a corpus of film scripts (75 films, 2,000,000 words)

• Results show some common phrases which tend to occur in descriptions of all films, e.g.:
  ➢ “… looks at …”
  ➢ “… turns to …”
  ➢ “… opens the door …”
  ➢ “… smiles at …”

• This suggests some important film events to extract information about, and useful patterns to use for information extraction
Plot summary
A young, shell-shocked war nurse (Hana) remains behind to tend her doomed patient.

Audio description
[23:54] Hana makes her patient comfortable

Audio description
[45:09] Gently she washes the tender skin on the patient's chest.
Further work in the TIWO Project

- Videsh Lingabavan, an MSc student at Surrey, is investigating the extraction of information from film dialogue
  - “I don’t know…”
  - “I don’t wanna…”

- Matthew Knight, an MSc student, is prototyping a system to extract story structures (such as plot units) from data about common film events

- Jamie Lakritz, a BSc student, has developed a system to convert a film screenplay into a first draft audio description
Some Important Related Work

- **Automatic Scene Segmentation**
  (Sundaram and Chang, Columbia University)

- **Analysis of Film Tempo**
  (Adams, Dorai and Venkatesh)

- **Analysis of Emotion in Film Audio**
  (Jones et al. at CDVP in Dublin City University)

→ How to integrate information about films’ extracted from multiple media streams?
Summary and Discussion Points

• Can we describe stories formally?

• Can machines understand stories?

• If so, how does this benefit us?

• How can we develop interdisciplinary approaches to answer these questions?
New Ways to Find and Watch Films

Dr Andrew Salway
a.salway@surrey.ac.uk
Department of Computing, University of Surrey, United Kingdom

2nd Technological Age Congress
Morelia, 1st April 2006