The Evolution of Film Editing Technique and its Implications to the Parsing and Summarization of Motion Pictures

John Mateer
Dept of Electronics
Dept of Theatre, Film & TV
(011 44) 1904 433 245
jwm10@ohm.york.ac.uk
Overview

- Overview of Shooting Related to Editing
- Evolution of Grammar and Techniques
- Relevance to Automated Parsing
- Benefits to Summarisation
- Issues and Obstacles
- Conclusions
Director controls what is shot

‘Surrogate audience member’

Composition attributes affecting editing

Subject Position

Framing

Lead space

Eye line
Subject Position within Frame

1/3rd

2/3rds

Camera Left

Camera Right

Saving Private Ryan ©1998 Paramount
Lead Space

Saving Private Ryan ©1998 Paramount
Eye Line
Shooting and Editing Scenes

- Scene Structure
  - A clear *beginning, middle and end*
  - Not a resolution of story but of moment

- Coverage
  - All shots taken to record the scene
  - Each new shot should advance story
    - *What does the audience need to see?*

- Processed raw footage called “Rushes”
Choosing Shots

- Variation in composition and angles will determine how shots “cut”
- Continuity
  - In Setting and Space
  - In Action
  - In Camera motion
  - The “Line”
Thomas Crown Affair (1968)

Clip from the bank manager scene
Relationship of Time to Editing

- **Time**
  - Running time – physical duration of program
  - Audience time – perceived duration
  - Story time – local duration within program
  - Action time – event duration

- **Flow of time controlled through “beats”**
  - In action, camera movement or transitions
  - Enable breathing space, emphasis
Evolution of Film Grammar

- Pre-1900: Birth of Cinema
- 1900-1910: Evolution of Continuity
- 1910-1920: Beginning of Common Language
- 1920-1930: Conceptual Languages Emerge
- 1930-onward: Expansion of established grammars, development of new ones
Pre-1900: Birth of Cinema

- Camera not mobile
- Requirement of light very high
- Theatrical Presentation
  - Fourth Wall model
  - Audience kept distant
- No transitions
- Simply to record a constrained event
1900-1910: Notion of Time

- Long takes
- Straight cut transitions emerge
- No line but evolution of continuity
  - entrance and exits from correct side
  - action consistent between adjacent shots
- Initially real time (no ellipsis)
- Location work just starting
Yorkshire Egg Collectors (1908)

Clip of ‘climmers’ gathering edge from cliff faces in Flamborough

From the Yorkshire Film Archives
1900-1910: Continuity Evolves

- Linear time but in multiple locales
  - very beginning of ellipsis
- Variation in framing dictated by location
- Audiences began to understand nature of film and basic continuity
- Novelty drives filmmaking
The Thieving Hand (1908)

Clip of the hand being pawned, stealing from the pawn shop and the police being summoned
1900-1910: Continuity Evolves

- Early notion of ‘Line’ through Eye Line
  - Lead space used to guide viewer’s attention
  - Continuity of space still inconsistent
  - Still ‘fourth wall’ model

- Technical innovations enhance technique
  - Stop motion photography, Multiple exposure
  - Variation in recording frame rates

- Audience sophistication drives innovation
Princess Nicotine (1908)

Clip of the stop-motion animation of cigar being created, man in living room smoking and tormenting the Princess
1910-1920: Common Language

- Master-Scene language established
- Start with Establishing Shot
- Closer shots to see subjects, action
- Close-ups to see emotion, reaction
- Wider shots for context, reinforcement
- Return to Closer shots to reinforce reaction
- Often Wide end shots to show resolution, exit
- ‘The Line’ gains acceptance as standard
1910-1920: Languages Develop

- **Parallel Language established**
  - Action in multiple locations occurring simultaneously – ‘cross-cutting’

- **Experimentation with convention**
  - Colour used to indicate time, emotion

- **New technology enables moving cameras**

- **Optical transitions become common**
The Lonedale Operator (1911)

Clip of the cross-cut sequence of the woman telegraphing for help as the robbers try to break in
1920-1930: Conceptual Language

- Pudovkin tests ‘Constructive’ model
  - Meaning of film can be broader than content of individual shots
  - Enables conceptual communication
  - Creates tension, enhances action

- Eisenstein expands this through ‘Montage’
  - Graphic, Rhythmic, Picture-Sound and Ideological juxtaposition
Battleship Potemkin (1925)

Clip of the Odessa Steps sequence
1930-onward: Grammar Explored

- **Visual Discontinuity**
  - Buñuel, surrealists challenge perception and expectation
  - Exploit audience understanding of continuity
  - Continuity, Lead Space, Eye Line ‘abused’
  - Development of *Graphic Match* cuts
  - Advent of *Jump Cuts* to jar viewer

- *If squeamish, look away now!*
Un Chien Andalou (1929)

Clip of the graphic match between the cloud crossing the moon and the lady’s eye being sliced
1930-onward: Grammar Explored

Transition Driven

- Focus audience attention explicitly
- No attempt to hide technique
- Most common in newsreels
- Continues in ‘hard sell’ advertising
‘News on the March’ (1941)

Clip of the ‘news reel’ from Citizen Kane
1930-onward: Grammar Explored

*Dissection of Action (Vorkapich)*

- Similar to montage but salient shots shown to convey a linear process
- Focuses audience on steps involved, detail
- Widely used today in conjunction with other techniques

Clip of the opening sequence where Luke cuts the heads off parking meters
1930-onward: Evolution of Sound

- Initially based on Theatre model
  - Music accompaniment to enhance drama
- Evolved to basic recording of event
  - Dialogue, action but all live
- Multi-track recording, editing enabled control
  - Foley, looping, better synchronisation
- Sound design evolves to control how picture is interpreted
Modern Filmmaking

- *Draws on all previous languages*
  - Exploits audience knowledge and expectation of grammar and genre convention

- *There is little that is new…*
  - Re-invention, re-discovery of technique

- *… except what technology enables*
  - Snorkel lens, virtual cameras, etc.
Raging Bull (1980)

Clip of the Sugar Ray Robinson title fight that Lamotta loses
Automated Parsing

Can technique and language be of use?

- Picture and Sound analysis can reveal patterns to potentially classify language
- Parsing certain elements could enable higher-level semantic concepts to be identified
- Determination of grammar could better define search ranges
- The goal is reduce the scale of the domain, splitting it into manageable data sets
Picture Analysis

- Basic technical attributes can be readily parsed through image processing
  - Pans, Tilts, Zooms and transitions can be found using disparity measures

- Camera movement can be classified based on 3D extraction from images
  - Dolly moves, Crane shots, Steadicam, etc
Picture Analysis

- Relative motion and subject tracking can be extracted via segmentation

- Face Detection methods can facilitate the extraction of Framings and Groupings
  - Noting orientation of respective head positions can indicate Eye Line

- Face Recognition can aid with context

- Text recognition through OCR to add detail
Picture Analysis

- Object pattern matching techniques can be applied in stages
  - Location or Setting can be established based on landmark detection (Eiffel Tower = Paris)
  - More abstract settings can be detected by salient features (traffic lights = street)
- The goal is to build semantic database to enhance chances of correct parsing
Sound Analysis

- Segmentation of soundtrack into music, dialogue and effects for individual parsing
- Classification based on sound signatures noting time to correlate with picture
  - Give clues as to what to search for visually
  - Enhance accuracy of parsing through reinforcement
Film Language Analysis

Patterns of picture and sound characteristics in shots can provide insight into language

- Framings and durations can create signature
- Types of Continuity can be classified

- Could be used to determine genre
- Could be used to assist dating of programs
- Can provide additional detail for summarisation
Summarisation

- Rich Metadata and prosaic textual descriptions can be generated
- Potential cross links with related programs to confirm/enhance extracted data
- Benefit to visual summarisation unclear
  - Dependent on metaphor used
  - Static depictions can nullify effect of editing technique
Summarisation

Reverse Storyboarding Metaphor

- Based on film preproduction visualisation methods
- Draws on extracted technical attributes of pictures
- Automatically adds visual cues in keeping with storyboarding technique
  - onion skins, streaks, trail lines, arrows, field cuts, etc.
- Annotated, mosaic or combination form
Reverse Storyboarding

Tilt Up, Pan Left in Arrow Form

Tilt Up Reference
Reverse Storyboarding

Zoom Out in Mosaic Form

Zoom In Reference
Issues and Obstacles

- Film language very rich and not uniform
- Patterns of editing not always consistent
- Significant knowledge of film history required to create effective model
- Accurate training and testing difficult given diversity of programs in existence
- Computationally very expensive
Conclusions

- Film languages have evolved into a sophisticated and multi-faceted grammar
- Techniques that facilitate these languages are constantly growing and changing
- Understanding language and technique could enable better semantic extraction
- Textual Summarisation could benefit but relevance to visual summaries unclear
Useful References

*Technique of Film and Video Editing*

*A Director’s Method for Film and TV*
Ron Richards, Focal Press, 1992

*Film Directing Shot by Shot*
Steven Katz, Michael Wiese Prods, 1991
Thank You!

John Mateer
Dept of Electronics
Dept of Theatre, Film & TV
(011 44) 1904 433 245
jwm10@ohm.york.ac.uk