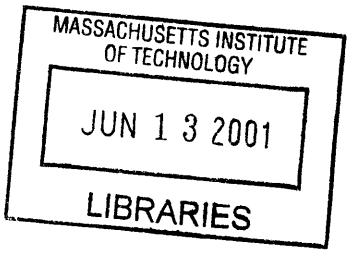
PlusShorts:
Using Punctuation as an Iconic System for Describing and Augmenting Video Structure

Aisling Geraldine Mary Kelliher

B.A. Communication Studies, Dublin City University
Dublin, Ireland (1996)

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning,

in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences



at the Massachusetts Institute of Technology
May, 2001

ROTCH

Signature of Author

.....

Program in Media Arts and Sciences
May 11, 2001

Certified by

.....

Glorianna Davenport
Principal Research Associate, Interactive Cinema Group
MIT Media Arts and Sciences
Thesis Supervisor

Accepted by

.....

Stephen A. Benton
Chair, Departmental Committee on Graduate Studies
Program in Media Arts and Sciences

PlusShorts:
Using Punctuation as an Iconic System for Describing and Augmenting Video Structure

by Aisling Geraldine Mary Kelliher

Submitted to the Program in Media Arts and Sciences
School of Architecture and Planning, on May 11, 2001
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

Affordable digital cameras, high bandwidth connectivity and large-scale video hosting websites are combining to offer an alternative mode of production and channel of distribution for independent filmmakers and home moviemakers. There is a growing need to develop systems that meaningfully support the desires of these filmmakers to communicate and collaborate effectively with others and to propel cinematic storytelling into new and dynamic realms.

This document proposes the development of a networked software application, called PlusShorts, that will allow a distributed group of users to contribute to and collaborate upon the creation of shared movie sequences. This system introduces an iconic language, consisting of punctuation symbols, for annotating, sharing and interpreting conceptual ideas about cinematic structure. The PlusShorts application presents individual movie sequences as elements within an evolving cinematic storyspace, where participants can explore, collaborate and share ideas.

Thesis Advisor: Glorianna Davenport
Title: Principal Research Associate, Interactive Cinema
Program in Media Arts and Science

The author gratefully thanks the Eircom Fellows program for its support

PlusShorts:
Using Punctuation as an Iconic System for Describing and Augmenting Video Structure

by Aisling Geraldine Mary Kelliher

The following people served as readers for this thesis:

Reader

Brian Smith
Assistant Professor
Explanation Architecture Group, MIT Media Laboratory

Reader

Maureen Thomas
Creative Director, Cambridge University Moving Image Studio
Senior Creative Research Fellow, Swedish Interactive Institute's Narrativity & Communications Studio

Dedication

To my Grandma, a truly amazing woman.

Acknowledgements

To my parents, Sheila and Jerry, who encouraged me to embark on this mad adventure and supported me throughout, even when my hair was pink. Thanks to my father for his daily sports reports, hilarious emails and all those 'research-related' articles. Thanks to my mother for the cards and the candles - you are my rock.

To my sister, Gráinne, whose courage and strength inspires me to try harder and be the best I can. You have a heart of gold.

To my advisor, Glorianna, who has steered me down this sometimes rocky road. For giving me the room to explore and for opening my eyes to unimagined possibilities, I am grateful.

To my UROP, Danny Talavera, who put in long hours coding the software. Thank you for your sterling work and good humor. Nice one.

To my homeboys, Jamie and Jonah. For introducing me to interferometers, $2 bets and Jay-Z. How can I ever thank you? To the rest of the unlockedgroove posse: Heemin, enough trance already; Dan, tek fu rock da mic; Alise, the best ex fat-hick vegan I know.

To the away team: Ellie, Aidan, Andrew, Alma, Karen and Julie. Thanks for watching out for me here and looking after me there. The best craic ever.

To the even further away team: Eleanor, Muireann (well, Cork is practically a foreign country), Tom, Carter, Emma, Rory and Paul. Thinking of you all.

To my IC cohorts: Barbara, for all your thoughtfulness; Roger, for being a calm and soothing officemate; James, for knowing more about movies than anyone I've ever met; Pengkai, for your cheerfulness; Paul, well, for being Paul; Cathy, for your persistence and dedication; Tony, for having the best laugh.

To my readers, Maureen and Brian, for your thoughtful comments and help in structuring this document. Thanks also to Sile, for her perceptive remarks.

To Linda Peterson, for your considerable patience and kindness over the last two years.

To Samuel Beckett, for making me laugh at nothing.

To Cameron, my partner in crime and one in a 10^{70} . For teaching me what that number means, archiving the waffles and being the delicious person you are. Love always. (ais + cam)!

Contents

List of Figures

List of Figures cont.

1.0 Introduction

1.1 Hypothesis

The growing convergence of streaming media technologies, content delivery mechanisms, extensive databases and smart collaborative filtering procedures has given rise to the development of large-scale video hosting websites, that support the storage, editing and delivery of user created content. Independent filmmakers, and home moviemakers alike have begun to use these websites to distribute and promote their work and to share video content with family and friends. However, there is an increasing desire to stretch the cinematic capabilities of the Internet beyond that of a movie distribution channel, towards the creation of an environment supporting large-scale collaborative productions that advocate and develop new cinematic forms.

This raises a number of important questions. How do we create an environment that facilitates the online production of many different types of movies by a distributed group of collaborators? How will these filmmakers interpret, evaluate and understand the cinematic contributions of other system users? And how will this environment affect not only our understanding of what it means to watch or produce a movie, but also our fundamental perception of what constitutes a cinematic artifact?

Effective communication and collaboration between people benefits from a structured framework accompanied by a common group understanding of the purpose and scope of the task at hand. For example, calendars provide a structured system for organizing units of time over prolonged intervals, to satisfy the needs and preoccupations of society.

They are a vital and indispensable part of everyday life, providing the basis for planning agricultural and migration cycles, for divination and prognostication, and for maintaining cycles of religious and civil events [Seidelmann 1992]. This structural framework provides a general vocabulary and syntax for the description of time, thus allowing people to easily plan and coordinate their lives with each other. Online, creative collaboration would greatly benefit from the development of such a structured framework, providing users with a common language for expressing, describing and sharing ideas about the creative process.

Movies are structured according to the nature of the story they wish to tell, the pace at which they choose to tell it, and the atmosphere they wish to evoke by it. The ability to clearly describe and share the conceptual ideas underlying the edited structure of a movie sequence is a useful, indeed if not necessary, tool for any collaborative moviemaking system. A structural markup language for the annotation and interpretation of movie structure would be a constructive way of implementing this feature.

The provision of an iconic visual language to describe these cinematic structures using a familiar symbol set is proposed. Punctuation symbols are a readily identifiable system of structural representation. Their semantic meaning is generally and consensually understood, yet flexible enough to allow for the comprehension of creative and alternative usage. The adaptation of punctuation notation as a representative system to describe, annotate and share ideas about movie structure is proposed in this research.

1.2 PlusShorts Overview

This thesis proposes the development of a software application that will allow a distributed group of users to collaborate upon the construction of shared movie sequences from a shared database of short video clips. This application, called PlusShorts, will allow users to both reveal and interpret the conceptual thinking underlying the editorial construction of these shared movie sequences. Using punctuation symbols as an iconic mark-up language, contributors to the system can annotate their edit decisions, while viewers can use the symbols to interpret the creator's intent and to generate debate and discussion within the PlusShorts community of users.

The movie sequences will be accessed and navigated through using a newsgroup-type threading technique, where each sequence will be represented as a singular video sequence 'post', forming part of a larger, shared video sequence 'thread'. Original movie sequences will be displayed as header threads, which can be expanded to reveal the hierarchically structured video sequence 'posts' submitted in response underneath. The conception of a video sequence as a variant instance within a larger story thread allows for a more playful, explorative approach towards moviemaking, where there is no definitive final cut, but rather a series of investigative story developments.

1.3 Overview of the Thesis

Chapter 2.0, "Extended Example", presents a scenario detailing how the PlusShorts system can be used to facilitate the collaborative construction of an open-ended evolving movie. The scenario indicates how filmmakers can share ideas about movie structure using punctuation symbols as a visual markup language to describe their edits.

Chapter 3.0, "Theory and Rationale", describes the research motivations and theoretical base guiding this research. This thesis is informed by exploratory work conducted in several disciplines including film studies, semiotics, annotation and knowledge representation. This chapter elucidates the key concepts from each field that have inspired the development of the PlusShorts application, and concludes with a description of previous relevant research.

Chapter 4.0, "Design and Implementation", outlines the design principles, structure and implementation of the PlusShorts software. Two iterations of the software are discussed: the local PlusShorts application designed for evaluating the central hypothesis in a workshop environment, and the online PlusShorts applet developed as a front-end application for the Shareable Media Project. The Shareable Media Project represents a large-scale research initiative within the Media Lab's Interactive Cinema Group, and will be discussed in this chapter in relation to the PlusShorts application.

Chapter 5.0, "Evaluation", presents the evaluation of the PlusShorts application. A one-day workshop for artists and filmmakers was conducted at the MediaLabEurope (Dublin), where participants used the PlusShorts software to collaboratively construct, annotate and interpret movie sequences. The design of this workshop will be discussed along with a detailed description and analysis of the day's proceedings.

Chapter 6, "Conclusion", describes the primary contributions of the PlusShorts application, summarizes the main conclusions derived from this research and posits suggestions for further research directions.

2.0 Extended Example

2.1 Collaborative Construction

Peter, a filmmaker from San Francisco, met Ellie while he was traveling in India. Ellie had just graduated from Art College in Dublin, and was backpacking around Europe and Asia. They were both using digital video cameras to document their adventures, and, by the time they encountered each other in Goa, had amassed a considerable amount of footage. They decided to travel together throughout Western India for a number of weeks before Peter returned to America. As they traveled, Peter talked about the different types of stories and movies he wanted to make using the footage he had shot. Ellie was interested in recounting her adventures a number of different ways: as a video diary, as a cultural study and perhaps as an art piece. They both became excited about collaborating together to tell the multifaceted story of their travels. Peter had previously used the PlusShorts system to make movies online with some friends of his in Boston and before he left, he suggested that this might be a good tool for them to use.

Shortly afterwards, Ellie returned to Ireland and began digitizing her footage into short clips. She uploaded the movie clips onto the Shareable Media website and began using the PlusShorts software to sequence the clips into short story segments. She put together a sequence detailing her travels chronologically up until her arrival in Goa. She used the punctuation symbols to define the structure of her sequence, for example, by using the colon icon before a clip that introduced a new character Ben. This clip was followed by a plus sign and a short movie clip of Paris, as Ben suggested she travel there, which in turn was followed by another plus icon and a short clip of Berlin, the next city visited on her journey.

Ellie used the plus icons repeatedly between clips to indicate the progress of her trip, which culminated with an equals sign placed before the final clip representing her arrival in Goa.

Peter, now back in San Francisco, also uploaded his captured footage to the Shareable Media site. Having looked at the movie sequence Ellie had constructed, he decided to replicate her edit structure to tell his own story of how he ended up in Goa. He posted a reply to Ellie's 'video thread', where he replaced all Ellie's footage with his own. He also added a new scene, represented as a 'video sentence' detailing a day trip they had made together to Chapora Fort. Peter used comma icons between various clips of the Fort, indicating that all of the shots presented formed part of the same segment, which he concluded with a period icon.

Meanwhile, Ellie had been collecting numerous news clips from the Internet describing real events that had been taken place not only in India, but also throughout the world, while she was away. She created another video thread that juxtaposed her adventures, trials and tribulations with those going on in the world at that time. She used the parenthesis icons to segment the different clip pairings, where each set of parenthesis contained a news clip, followed by an ellipses icon to indicate 'meanwhile', and a final clip denoting a concurrent event in Ellie's trip.

Peter thought it would be fun to use Ellie's latest video thread as the basis for a dramatic reenactment. He shot some footage of a friend of his pretending to be Ellie's concerned mother, watching the news as it described all the disastrous events taking place in the countries where her daughter was traveling. He put together a sequence using the news clips, footage of "Ellie's mother" reacting to it, and some shots of the blissfully unaware Ellie on her happy travels. Peter also used the parenthesis icons to segment the sequence into its requisite parts.

One such segment contained a shot of Ellie's fictional mother reading a postcard from Ellie in Paris, followed by a plus icon connecting that shot to a news clip of bombs falling on the Eiffel Tower. An equals icon led to the mothers' reaction shot, while outside the parenthesis, an ellipses icon was used to indicate that meanwhile, Ellie was having a great time, as demonstrated by the clip showing her visiting vineyards in Normandy.

Karla was teaching a class on video editing to a group of high school students. She was looking on the web one day for examples of video clips she could use in her class. She came across the Shareable Media site and launched the PlusShorts applet. She decided to use the system in her class to help students think about the structure of the movies they were trying to make, and why certain edit methodologies worked best with certain types of footage. Her students were impressed with the video threads created by Ellie and Peter and began to formulate diverse stories inspired by their footage. Some students used the more exotic clips as a dream sequence forming part of a story they were developing, whilst others incorporated additional footage from action adventure movies to make it look like Ellie and Peter were fugitives on the run. The students used the punctuation icons with each story treatment to describe the edit structure of their sequence. They examined, interpreted and discussed the marked up structures of the various sequences created by the class, identifying the most successful formats for each type of story and movie genre.

3.0 Theory and Rationale

This research describes the development of a tool that uses punctuation symbols as a visual markup language for annotating and sharing conceptual ideas about the structure of video sequences. The theoretical basis for the development of this practical application is derived from a number of disparate fields: notation, annotation, literature, art, film studies, philosophy, semiotics, information retrieval and knowledge representation. The purpose of this chapter is to extract and elucidate the factoring influences and relevant work from each of these fields that pertains to the development of the PlusShorts application.

3.1 Notation

> **no·ta·tion 1a :** the act, process, method, or an instance of representing by a system or set of marks, signs, figures, or characters **1b :** a system of characters, symbols, or abbreviated expressions used in an art or science or in mathematics or logic to express technical facts or quantities.
>
> [Merriam-Webster 1998]

A notational system can be defined as an ontological dimension reified by a set of tokens, which has a set of combination rules defining operations that may meaningfully be performed upon the tokens [Long 1996]. The power of these notational systems resides not with the symbols themselves, but rather with the abstractions that are represented by these symbols. The philosopher Max Black, in disputing theoretical attempts to "improve" symbolism, stated "tokens of any properties whatsoever can be used as the material for a complete language." [Black 1949: 160]. Symbol sets may be created from entirely new tokens or they may borrow from a preexisting symbol set. What is of importance is the ability of the chosen symbol set to represent all aspects of the ontological type it is seeking to describe.

Notation systems can represent a wide range of ontological dimensions, including entities, groups, values, forms, relationships, instruction and process. Examples of notation systems include the alphabet, mathematics, music, symbolic logic, chemical formulae, dance, and systems design. The meaning of notation systems can be governed by law (money or traffic signs), non-governmental professional groups (mathematical or engineering notations) or in some cases by community authority (for example, the alphabet is controlled by ongoing public consensus). Of specific relevance to this research is the use of instructional notation schemas relating to injunctive description and interpretation within the artistic and expressive domains, in particular within the fields of music and dance.

3.1.1 Musical Notation

Iconic notation schemata have been used to annotate, define and augment creative works for centuries, if not millennia. Some evidence exists to indicate that the Egyptians practiced the notation of music from the 3rd millennium BC [Britannica 1990]. Musical notation, the written, printed or visual representation of music, operates as a memory aid during composition and exists to preserve music for later performance and analysis. From the earliest neumatic system of squiggled shapes introduced in Carolingian Europe in the 8th and 9th centuries, musical notation matured over the centuries to reflect increasing precision in rhythm, the development of the staff and the concept of scale. By the early 16th century, Western musical notation had become standardized with the mandatory components of staff, clef, time signature and bar lines, as shown in FIGURE 3-1.

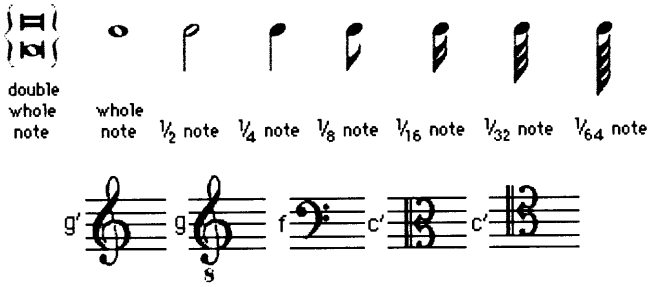


FIGURE 3-1: Standard Western Music Notation

However, the mid-20th century heralded a revising of this conventional system, leading to the alteration and indeed frequent abandonment of the standard scheme. Increasingly complex rhythmic values and pitches outside the normal chromatic scale have led to the development of new symbols. Some composers have chosen to discard the staff system entirely, appropriating instead the notation schemes for graphs, mathematics and diagrams in an attempt to represent their non-traditional compositions. FIGURE 3-2 shows a page from the score of Karlheinz Stockhausen's *Electronic Study No.2,* where the composer used a graph to represent the tunings of a sine oscillator, all of which were based on the 25th root of 5 (1.066494942).

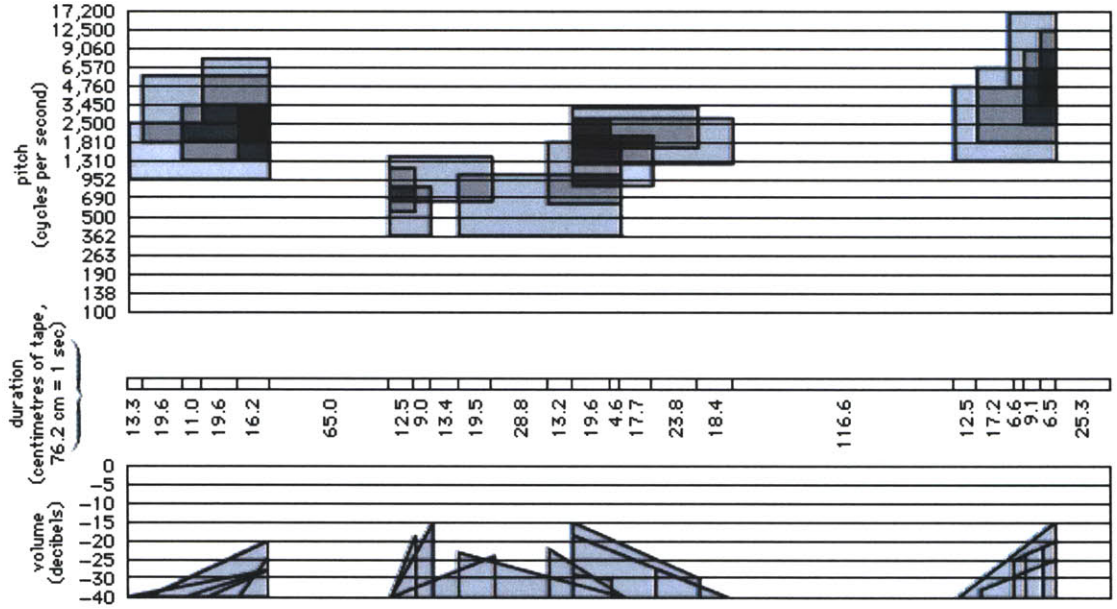
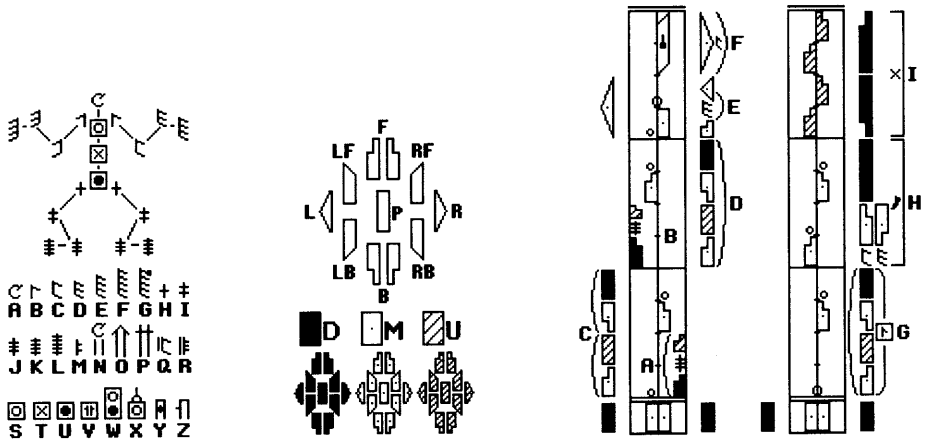


FIGURE 3-2: Page from the Score of Stockhausen's *Electronic Study No. 2*

3.1.2 Dance notation

The first written record of dance notation originated in the latter half of the 15th century in the form of letter-symbols found in manuscripts in the municipal archives of Cervera, Spain. However, it is the Hungarian, Rudolf Laban, who is credited with developing the first notation technique for describing precise movement and rhythm in dance. [Laban 1928].

His system, Labanotation, aims to record every aspect of human motion as precisely as possible and is not connected to a singular, specific style of dance. Similar to musical notation, Labanotation uses a staff, which is marked with lines, symbols and accents to indicate parts of the body, direction, gesture, center of weight and timing. FIGURE 3-3 shows the differentiated system used to label all parts of the body, the direction of a movement, and finally the use of vertical bows to demonstrate simultaneous actions and phrasing.



Body Parts **Direction Symbols** **Simultaneous Actions**

FIGURE 3-3: Examples of Labanotation Dance Notation

Both music and dance instructional notation systems have evolved their representation schema in accordance with the radical changes and developments occurring within their professed domain. It can take centuries for a new symbol set or representation methodology to become broadly accepted and understood by the populace, which leads in some cases to the appropriation of a previously defined notation scheme from another discipline, as with the incorporation of a graphical representation in the Stockhausen example. It is for this reason that I chose to adopt for my purposes, the symbol set from an already firmly established notation system - punctuation. The meanings of the symbols chosen are generally already clearly defined, yet have developed sufficiently to allow for their inventive adaptation and reinterpretation both within their original ontological domain and beyond.

3.2 Punctuation

"No iron can stab the heart with such force as a full stop put just at the right place."

Isaac Babel [Babel 1955]

The system of punctuation used today with English and other western European languages derives from the punctuation used in the classical period with Greek and Latin. Aristophanes, librarian at the Museum at Alexandria, devised many of the significant signs and accents incorporated into Greek texts [Britannica 1990]. These signs originated from the theory of rhetoric and so referred to the rhythmical units of speech. The earliest Roman inscriptions used points to separate words, and paragraphs were sporadically utilized to indicate a change of topic. The 7th and 8th centuries heralded the introduction of new punctuation symbols thought to have derived from the neumatic musical notation used at the time to notate Gregorian chants. These new marks, which include the question mark, indicated appropriate inflections of the voice. In the 16th century, Aldo Manuzio declared that the clarification of syntax was the main goal of punctuation, and with that he postulated the modern comma, semicolon, colon and period [Manuzio 1970]. The exclamation mark, quotations and the dash were in place by the end of the 17th century.

The invention of printing standardized the system of punctuation, disseminating grammatical and typographical standards and placing the emphasis squarely on structure, rather that on spoken sound.

"As he Sets on, he [the printer] considers
how to Point his Work,
viz. when to Set, where; where. Where to make ( ) where [ ]
and when to make a Break...
Joseph Moxon [Moxon (1683) 1962]

Punctuation today is still largely an intuitive practice, its rules governed by nothing more than public consensus. Structure may be the most prominent feature, but rhetoric and pace still exert some influence on a writer's choice of punctuation. These choices can stretch beyond the straightforward purpose of structuring a document, and move instead towards fulfilling creative commutative, interpretive and teaching functions, as the following examples demonstrate.

3.2.1 Punctuation and Literature

"Dante...Bruno. Vico.. Joyce"

Samuel Beckett [Beckett 1929]

Samuel Beckett explained his first published writing on Joyce, which explores the use of punctuation as a visual metaphor for time, as follows: "From Dante to Bruno is a jump of three centuries, from Bruno to Vico is about one and from Vico to Joyce about two."[Bair 1978: 80]. Throughout his writings, Beckett used punctuation marks to establish the "dynamic rhythms" of his "word-music", in a method that drew attention to the essential musicality of his work. [Whitelaw 1995: 78]. His writings are often compared to musical scores with his rhetorical marks functioning as musical notation. Steven Connor has identified four separate epochs of punctuation in Beckett's work. His early work can be identified as the classical phase where he utilized all normative punctuation resources. The novel Watt [Beckett 1970] ushered in a new era defined by "Beckett's discovery of the extraordinary capacities of the comma, to create a kind of counterpoint between the sheer going on of the sentence, with no awareness of its likely end, and the interruptions, resumptions and folding over that the comma gives."[Connor 1998]. The next phase saw an outright rejection of the use of punctuation, which developed finally to a resuscitation of syntax and a delight in formal spacing in his later writings.

e.e. cummings is another notable example of a writer who experimented radically with form, syntax, spelling and punctuation. His highly original style of poetic expression sought to abandon traditional techniques and structures in favor of idiosyncratic layout and inventive wordplay.

r-p-o-p-h-e-s-s-a-g-r

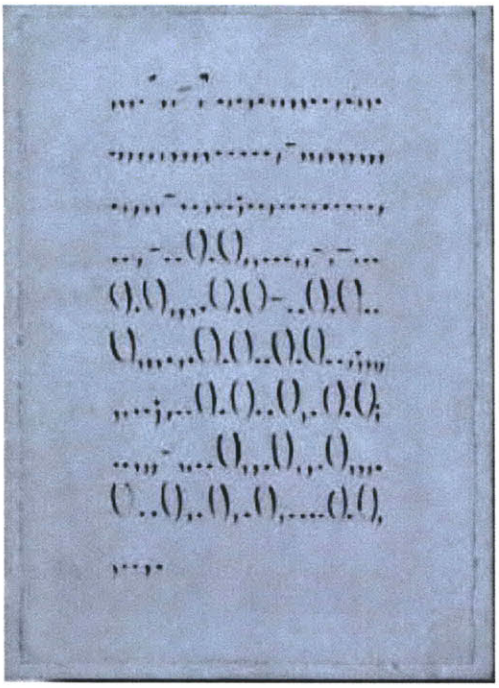
r-p-o-p-h-e-s-s-a-g-r
who
a)s w(e loo)k
upnowgath
PPEGORHRASS
eringint(o-
aThe):l
eA
!p:
S a
(r
rlvlnG .gRrEaPsPhOs)
to
rea(be)rran(com)gi(e)ngly
,grasshopper;

[Cummings 1923]

Gertrude Stein, one of the great modernist writers, abandoned traditional literary style altogether for a highly repetitive prose without conventional punctuation. Edith Sitwell, writing in 1935, describes how Stein is "bringing back life to our language by what appears, at first, to be an anarchic process. First she breaks down the predestined groups of words, their sleepy family habits; then she rebrightens them, examines their texture, and builds them into new and vital shapes." [Sitwell 1925]. Other illustrious writers and poets of the period include Hart Crane, Ezra Pound and the inimitable James Joyce. Each of these writers tortured and stretched the boundaries of the English language, using inventive spelling, wayward grammar and radical punctuation to commentate directly on the semantics of the words their literary compositions contained.

3.2.2 Punctuation and Art

Richard J. Galpin's *Punctuation Extraction* series is a powerful example of the ability of punctuation notation to convey an author's emotional tone and structural style. Galpin takes familiar pieces of writing (books, plays, declarations) and removes all the words contained within them, leaving only the inky punctuation symbols behind. These stark and simple black and white paintings immediately manage to convey the formal nature of the *Universal Declaration of Human Rights* [FIGURE 3-4], the humor and zaniness of *The Cat in the Hat* [FIGURE 3-5], the overtly conversational style of a Mills and Boon novel, *Private Pleasures* [FIGURE 3-6], and the idiosyncratic and aesthetic formalism of Samuel Beckett's *Endgame* [FIGURE 3-7].



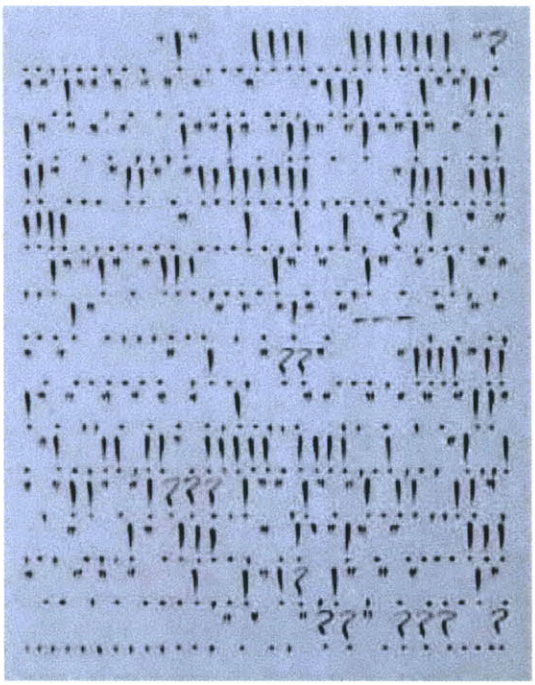


FIGURE 3-4: Declaration of Human Rights FIGURE 3-5: Cat in the Hat

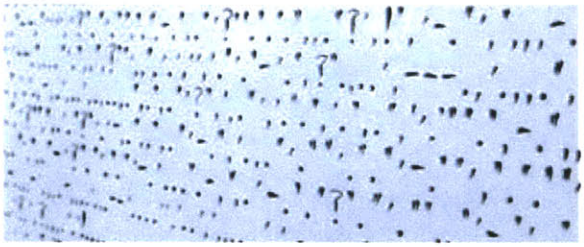


FIGURE 3-6: Private Pleasures (Detail)

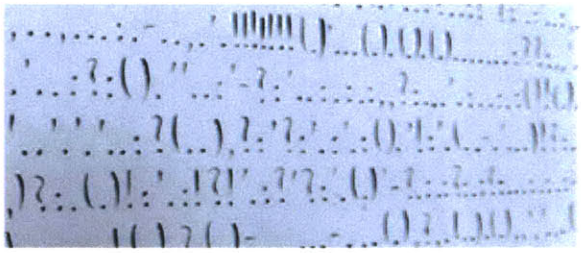


FIGURE 3-7: Endgame (Detail)

3.2.3 Punctuation and Dance

In "Writing Movement / Dancing Words: A Collaborative Pedagogy", Donna Davenport and Cheryl Forbes write about the relationship they have discovered between writing and dance and the pedagogical application of such a connection [Education 1997]. Davenport uses language metaphors to teach students dance technique and dance composition at the Hobart and William Smith College, where she describes the nuances of movement in terms of punctuation, vocabulary, and voice. Using punctuation to define movement phrases and transitions she draws on the following analogies:

- the exclamation point as an exciting declarative conclusion
- the semi-colon completes an idea and leads the dancer into the next statement
- the comma as a slight pause, that shapes continuity from one sequence of steps to the next
- the parentheses frame a new movement theme which digresses from the main text
- the period is the definitive end of a regularly metered phrase.

Forbes contrasts the structural use of space on a page with the spatial relationship between sequences in a dance, whilst also alluding to the use of punctuation in texts to affect a reader's perception of time as being analogous to how a dancer uses structured pauses and breaks to control the passage of time in a dance.

Punctuation has been used not only to structure works, but also to facilitate the interpretation of style and tone, to comment directly on the text or material it contains, and to act as a structural metaphor for the teaching of another expressive discipline. Paul Hirsch, editor of *Star Wars* (1977), *Blowout* (1981) and many of Brian DePalma's movies, in describing his editing methodology stated: "You can create punctuation by the timing of the cut. If you establish a certain pace, and then change the tempo at a certain point, you break the rhythm and that can act as a period on a sentence. If you want to start a whole new paragraph, sometimes it's important to use a dissolve which the audience interprets as a paragraph or a chapter."[Oldham 1992: 194]. With this research, I am most interested in investigating the capacity of this symbol set to operate within the cinematic realm as a form of metadata, to annotate and explain a sequence such that an additional layer of interpretive meaning is applied.

3.3 Annotation

an·no·ta·tion : a note added by way of comment or explanation

[Merriam-Webster 1998]

The annotation of a document generally involves the addition of a supplementary layer of purposeful data that can be used in order to obtain information about the structure and content of the described document. This metadata, data about data essentially, can be used to define element names, descriptions, representations, semantics and classifications. Tim Berners-Lee defines metadata as "machine understandable information about web resources or other things" [Berners-Lee 1997] Content-specific metadata is generally recognized as the only form of annotation to provide a sufficient degree of generality for the efficient indexing and retrieval of web objects, and in particular, multimedia objects. However, some content-specific approaches serve only to describe the structural and physical characteristics of objects (color, shape, motion patterns etc.).

Of interest to this research in particular, is the current effort to address this issue by moving towards a more comprehensive semantic model that uses conceptual annotation methodologies.

3.3.1 Conceptual Annotation

Elisa Bertino and her colleagues at the University of Milan have proposed the association of Natural Language captions with web objects. These captions would take the form of "short texts representing a general, neutral description of their informational content." [Bertino et al 1999] Such captions could then be converted into a conceptual annotation in NKRL (Narrative Knowledge Representation Language), as demonstrated in the WebLearning and CONCERTO projects [Zarri G. et al 1999]. Graziella Tonfoni has developed an iconic annotation system to enhance content and context visibility in documentation. CTML (Context Transport Mark Up Language) is a derivative language of CPP-TRS, which is a visual language that provides users with a set of interpretive clues that aim to facilitate a deeper understanding of a document. CTML aims to make precise context attachment possible by providing a variety of annotation signs indicating different document features such as communicative function, style, intention and turn-taking [Tonfoni 1999]

One of the goals of this research is to develop a similar iconic annotation system, which allows users to describe and share, conceptual ideas not about documents, but about the structure of edited movie sequences. Recent years have seen a significant surge in the proliferation and use of digital video information. Efficiently analyzing, annotating, indexing and retrieving vast amounts of rich content is an arduous and difficult task that has been the subject of significant academic and commercial research.

3.3.2 Video Structure and Video Annotation

Considerable research has been conducted into the automatic annotation of video structure: detection of low-level information such as color, texture, shape [Zhuang 1998], shot boundary detection [Arman 1993] and group based representations of similar clusters of frames [Zhang 1997][Zhong 1995]. However, these approaches although effective in detecting accurate data about the physical properties of audiovisual content, provide a somewhat impoverished model for deriving the semantic concepts of a video.

More recent research has begun to focus on exploring video structure at a coarser granulation, in accordance with the dictum that people understand movies according to events or semantic scenes rather than as a series of shots or fleeting keyframes. The automatic construction and elucidation of "scenes"(collection of semantically related and temporally adjacent shots) in an audiovisual stream is becoming of increasing importance in many video applications, where subsequent browsing and retrieval of the video data is to be supported. [Yeung 1996] [Rui 1998] Specific application models have been developed for dealing with the parsing of particular types of video content, for example, news. [Zhang 1995]

Constructing a model for the description and representation of video content at this level of granularity requires a clear understanding of video and cinematic structure. Research conducted into investigating the annotation of audiovisual content using story and narrative structural constructs is discussed in Section 3.6.1., while the cinematic theory grounding this work will be discussed in the next section.

3.4 Cinema, Editing and Structure

"Editing is choosing what's to be in the film and in what order."

Paul Hirsch, Film Editor [Oldham 1992]

Editing plays a vital role in the overall stylistic system of a film, strongly shaping the experience for the audience, although they are often unaware of this. Shots are spliced together in a movie using straightforward instant cuts, or more gradual dissolves, fade-in and fade-outs, wipes and other optical effects. The relationship between adjacent shots can be graphic (visual dimensions of shots), rhythmic (duration of shots), spatial (location of shots) or temporal (time of shots).

3.4.1 Classic Narrative Continuity

Edwin Porter was the first filmmaker to arrange shots specifically to present a narrative continuity, but it was D.W. Griffith who introduced such major tenants of classical editing as the extreme long shot, the close-up, the cutaway, the tracking shot, parallel editing and the variation of shots for impact and pace. Films such as *The Birth of A Nation* (1915), *Intolerance* (1916) and *Broken Blossoms* (1919) epitomize Griffith's gift for enunciating the emotional and dramatic impact of a story through the use of skilled and imaginative editing.

The classic continuity style of editing that has developed from this era is largely concerned with presenting narrative action, through the handling of time and space. The space of a scene is constructed along what is called the "axis of action" or the "180 degree line", which determines a half circle where the camera can be placed to present the action. [Bordwell 1997: 285] Temporal continuity is maintained through the manipulation of time, which can be further distinguished by temporal order, frequency, and duration.

Temporal order relates to the presentation of the story events in a linear fashion and frequency alludes to the fact that most events in a movie are presented only once (flashback obviously violates these norms). Temporal ellipses, represented by dissolve, wipe or fade transitions are often used to indicate to the viewer that time has passed. John Huston's *The Maltese Falcon* (1941) and William Wyler's *The Best Years of Our Lives* (1946) are pertinent examples of movies produced using the classical Hollywood conventional style of continuity editing.

3.4.2 Discontinuity and Montage

An alternative form to the classical style of continuity editing was proposed by the Soviet filmmakers of the 1920's, in particular, Vsevolod Pudovkin and Sergei Eisenstein. They modified Griffith's methodology to include a more formalized process that sought to not only relate stories, but also to draw intellectual conclusions from them [Reisz 1968: 27]. Eisenstein formulated the idea of montage: the radical juxtaposition of shots to create a new and startling idea or emotion, what he called, a collision of attractions. Spatial and temporal discontinuities were encouraged as evidenced in such films as *Strike* (1924), *Oktober*(1928) and *Battleship Potemkin* (1925). Eisenstein used spatial, graphical and temporal editing to construct analogies and powerful comparisons to help interpret the story events. Luis Buñeul was similarly interested in the principle of cinematic counterpoint, but from a decidingly surrealist perspective. *Un Chien d'Andalou* (1929) and *L'Age d'Or* (1930), both produced in collaboration with Salvador Dali, contain shocking, surreal shots sequenced to deliberately create a radical visual asymmetry.

The films of the New Wave, beginning with Francois Truffaut's *400 Blows*(1959) and Jean-Luc Godard's *Breathless*(1959), introduced mainstream cinema audiences to the disruptive jump cut. This had the effect of immediately drawing the viewer's attention to the fact that they were engaged in the process of watching a film.

This invocation of self-reflexivity was an important element in Godard's cinematic counterstyle, which sought to negate the so-called manipulative character of narrative storytelling.

3.4.3 Influences on Editing

The development of cinematic editing has been informed not only by the two contrasting approaches outlined above, but also by influences within the popular arts, the emergence of documentary and the invention of television. Vaudeville-influenced film offered audiences a different cinematic experience from Griffith's dramatic epics or the intellectual polemics of the Soviets. The films of Charlie Chaplin and Buster Keaton epitomize many of the characteristics of the burlesque (the victim, the routine, the performance, the set pieces) and are specifically edited to centrally situate the star's persona along with clearly articulating and defining the comic set pieces.

John Grierson, affected by the powerful editing of the Soviet filmmakers in the 20's, sought to take their principle of social and political purpose and marry it with a strong visual aesthetic. The development of sound in the 1930's greatly aided the emergence of documentary as an instrument of social policy in Europe and America [Dancyger 1997: 54]. Grierson's *Night Mail* (1936), Flaherty's *Man of Aran* (1934) and Lorentz' *The Plow that Broke the Plains* (1936) produced tightly woven narratives of sound and image that engaged critically and imaginatively with social, economic and cultural issues.

The introduction of television, with its tremendous sense of immediacy, its rapid-fire commercial presentations and its general fast pace, had a profound effect on film editing. The docudrama approach - a mixture of cinema verite, jump cut editing, and narration - was adopted to make films seem more real and more immediate (*The Manchurian Candidate* (1962), *All The President's Men* (1976)).

The filmmaker Richard Lester was one of the first to wholeheartedly adopt many of the techniques of television editing, as exemplified in his two Beatles films, *A Hard Day's Night* (1964) and *Help!* (1965). Using a combination of cinema verite, with an offbeat narrative, along with jump cuts, extreme close-ups, crazy zooms and ridiculous cutaways, Lester infused the movies with an exuberant energy that pushed the narrative (or semblance of) along at an accelerated pace.

The readiness of audiences to accept this quicker pace meant that movies began not only to accelerate the narrative, but also the editing. A new visual sensibility involving super fast cutting, non-linear narratives and the elevation of style over substance became the staple of movies in the 60's, including *Bonnie and Clyde* (1967), *The Graduate* (1967), *Easy Rider* (1969), and *Woodstock* (1970). This development was simultaneously echoed in the related audiovisual domains of television commercials and music videos.

"No art form is as schizophrenic as the music video. In part a commercial and in part a short film, it has flaunted the line between art and commerce, undermined narrative and character development, and shortened an entire generation's attention span."

Steve Reiss, Author and Music Video Producer [Reiss 2000: 10]

The shaping device of the music video and the source of its energy is the pace of the music itself. Narrative is relegated to the sidelines in favor of the creation of a "feeling state". Close-ups, jump cuts, extreme foregrounding and the overuse of pace all function to obliterate any semblance of definitive time or place. FIGURE 3-8 shows selected shots from the music video "*Big Time Sensuality*", which features multiple jumpcuts between close-up, strongly foregrounded shots of the singer Björk.

FIGURE 3.8: Shots from Bjork's "Big Time Sensuality" Video, dir. Stephane Sednaoui

Music videos are necessarily densely textured as they most hold up for repeated viewing. They can be performance driven (Aerosmith and Run D.M.C's "*Walk this Way*"), narrative (Beastie Boys "*Sabotage*"), animated (Radiohead's "*Paranoid Android*", A-Ha's "*Take on Me*"), classic portraiture (Sinead O'Connor's "*Nothing Compares 2 U*") to name but a few genres.

David Bowie described the music video as "the logical fulfillment of art and technological destiny…I can visualize the day when the interface of music and video will create an entirely new kind of artist."[Reiss 2000: 18] As video makers begin to cross over to feature films – witness Spike Jonze, Julian Temple, Russel Mulcahy and Dominic Sena - the influence of their 'breeding ground' on these hybrid artists is evident. The distributed environment of the Internet provides a potential new home for the music video, where music and images made for the small screen can be sampled, combined and mixed in real time by a dispersed group of online collaborators.

3.4.4 Editing and Genre

The style of editing used in a cinematic scene or even throughout an entire film is often dictated by the genre category of the movie itself. Action scenes, for example, frequently represent climactic or turning point scenes in a movie. The editing in these dynamic, fast-moving scenes centers around a number of issues: identification, excitation, conflict and intensification. Close-ups and point-of view shots encourage identification. Movement within shots (pans, tilts, zooms, tracking, dolly, hand-held), movement between shots and variation in the lengths of shots enhance the element of excitement in a scene. Conflict is introduced by using crosscutting, while varying the lengths of shots, from short to long to set pattern helps to intensify the dramatic intent of a scene.

The editing in a dialogue scene is determined at a general level by the genre of the movie and at a specific level, by the actual meaning of what is being said. For example, if a piece of dialogue is important for moving plot forward, then maybe a close-up or some shift in the pattern of shots is required. In action movies, dialogue is often used to fill in the details, whereas in more character driven movies, such as a Woody Allen production, dialogue is used to establish character and is indeed the fulcrum around which the film plot revolves. In *The Graduate,* dialogue fulfills multiple functions, where it is used to advance the plot, to reveal a character's nature and to provide comic relief.

The editing of a comedy film relies on an interdependent relationship between the writer, the director, the actor(s) and the editor. The key thing to understand is the material itself: what is its narrative intention, its source of humor (character, situation, satire or farce) and the intended target of the humor.

The documentary film is concerned with the exposition of a theme, which is found and shaped in the editing. Imaginative documentary, such as *The Thin Blue Line* (1988), offers the editor many options for creating a new interpretation of reality using sound, juxtaposition of sequences, use of different types of shot, close-ups and pace.

3.4.5 Sound and Cinematic Structure

"The eye solicited alone makes the ear impatient, the ear solicited alone makes the eye impatient. Use these impatiences."

Robert Bresson [Bresson 1986]

Michel Chion describes the relationship of sound (dialogue, music, sound effects) to image in the movies as "added value", where the added value "engages the very structuring of the vision – by rigorously framing it." [Chion 1994: 7]. Music can be used in movies to both reinforce the emotional feeling of the scene and to radically juxtapose the depicted situation by remaining blatantly indifferent to the images onscreen, thus creating what Chion calls an "anempathetic" effect. An example of this latter effect would be in Hitchcock's *Psycho* (1960), when the noise of a shower running continues deferentially after, and in spite of, Marion's violent death. The main definitive function of film sound, according to Chion, is to unify the flow of images by bridging the visual edits and by establishing atmosphere.

3.4.6 Sound and Rhythm

Animated films demonstrate a particularly close relationship between onscreen movement and sound. As the sound track is generally recorded prior to any actual drawing, animated characters, particularly those in Disney films, tend to move in exact synchronization with the music. In feature films, such a close synchronization between sound and image is called "mickey-mousing", a phrase coined by David O. Selznick to describe Max Steiner's music scores, which so directly referenced and alluded to the activity on screen [Weiss 1985: 408].

Contemporary cinematic techniques have moved away from "catching action" with sound, preferring instead to elucidate the implicit values in a scene by, for example, cutting against rhythm to emphasize dialogue. A powerful example is Chris Marker's *La Jetee* (1962), the central conceit of which is the contrasting rhythms between the still images and the rapid, constantly changing and evolving soundtrack.

3.4.7 Sound and Punctuation

Early silent cinema borrowed heavily from the narrative techniques of theatre and opera, both of which used breathing, gesture, pauses and intonation to punctuate their texts. Meaning was derived and inflected in silent cinema using gestural, visual, textual and rhythmical modes. The advent of synchronous sound thus served to accentuate the subtleties of movie punctuation, whereby the bark of a dog off-screen, the toll of a bell or the tinkling of a piano could be used to emphasize a word, dramatically close a scene or indicate a change of location. Such sounds served to define a film's space, establish the framework of the action and help create and transfigure the movie's overall rhythm. A distinct type of sound/image synchronicity can be witnessed in martial arts and fight films, which rely heavily on the dramatic stylization of time and the use of slow motion. The exact moment of a contacting punch is often bracketed by episodes of temporal elasticity, as exemplified in the brutal fight scenes in Scorsese's *Raging Bull* (1980).

Chion cites John Ford's *Informer* (1935) as a pertinent example of a movie using music as symbolic punctuation. Max Steiner's score illustrates the principle of the leitmotif, where a musical theme is assigned to each of the main characters, the key thematic ideas and particular settings. Such musical themes are heard whenever the corresponding character appears, and they react and change accordingly to reflect the changing fortunes of that depicted character. Akira Kurosawa also uses sound to both punctuate and inform the narrative in the final battle scene in *Seven Samurai* (1954).

Incorporating "sound perspectives" to spatially orient the shifting areas of activity, this incredibly dense sound track shifts and swirls as new narrative elements are constantly introduced (hooves, battle cries), abruptly and piercingly punctuated (women's screams, gunfire) and gently modulated (rain falling).

Sound helps to direct and structure the audience's response to a movie, by framing the narrative, highlighting what is most important at certain times, and supporting the emotional resonance of the film. It can be used to refer directly to the action on screen, mimicking events and emphasizing moments of high drama. This research seeks to use a symbolic notation system to help add a similar layer of interpretive meaning to a movie. This representation scheme can be used to directly reference and draw our attention to a particular event on the image or soundtrack. It can be attached as in Steiner's leitmotif, as specific annotation describing a particular character, remaining with that character throughout the sequence. It can be also utilized to accentuate the evolving importance of events, as in *Seven Samurai*, where particular notation can always be solely associated with the focus of attention.

3.5 Film and Semiotics

In "Signs and Meaning in The Cinema", Peter Wollen, in arguing why the field of semiotics is pertinent to the study of cinema, declares that criticism depends upon knowing what a text means and understanding the code or mode of expression which permits cinematic meaning to exist. [Wollen 1972: 17]. The essential principle of semiotics is that the whole of human experience can be understood as an interpretive structure mediated and sustained by signs. [Deely 1986: xi]. This research seeks to provide a framework and vocabulary for filmmakers to elucidate the meaning and mode of expression of their cinematic constructions. The development of this framework has been guided by the semiotic, cine-semiotic and cognitive semiotic theory described in the following sections.

3.5.1 Semiotics

Contemporary semiotics, as espoused by Charles Pierce and Ferdinand de Saussure, proposed the inauguration of a new and comprehensive discipline based on linguistic methods. Saussure, using the term semiology, defined it as a science that would "show what constitutes signs" and "what laws govern them" [Saussure 1966: 16]. He defined the sign as the union of a form which signifies – the signifier and an idea signified – the signified. Saussure described signs as entering into two fundamental types of relationship: paradigmatic (involving choosing between signs) and syntagmatic (involving combining signs). For his part, Pierce famously defined a sign as "something which stands to somebody for something in some respect or capacity."[Pierce 1931].

3.5.2 Language and Film: Cine Semiology

The Russian Formalists, who understood art as a system of signs and conventions, were the first to explore the analogy between language and film. Shklovsky, Tynianov, and Eikhenbaum described in the 1927 anthology Poetic Kino, how the "poetic" use of film was analogous to the "literary" use of language, thus championing the anti-normative, the deviant and the avant-garde. Eikhenbaum described cinema as a "particular system of figurative language", the stylistics of which would treat filmic "syntax", the linkage of shots into "phrases" and "sentences" [Eikhenbaum 1982]. Vladimir Propp extended the Formalist approach to narratology, the study of narrative structure. In The Morphology of the Russian Folk Tale, Propp outlined a typology of narrative structures based on his analysis of the "narratemes" or smallest narrative units of traditional Russian folk tales. [Propp 1968]. This seminal work helped lay the foundations for French Structuralism, and served to influence in particular the work of Roland Barthes.

While Formalism emphasized the message and poetic function of art, structuralism emphasized the code and the metalinguistic purpose of art. The advent of structuralism in the 1960's, defined by Roland Barthes as a "mode of analysis of cultural artifacts which originated in the methods of contemporary linguistics" inspired a deeper exploration of the cine-language concept. [Barthes 1967: 897]. As Barthes was examining the paradigmatic and syntagmatic relationships between food items on a menu (*choose* one type of soup and then *combine* it with an entrée and dessert), Godard was exploring the same operations in the cinematic domain. His whispered voiceover in *Two or Three Things I Know about Her* (1967) – "Should I focus on the leaves, or the sign?...Am I too close?...Is my voice too loud?" - indicate a self-reflexive, metacinematic consideration of the choosing and combining functions of film.

Film semiotics emerged in the early 1960's as part of the general academic movement emphasizing the applicability of structural linguistic to multiple disciplines. The pioneering film semiologists of the time, including Pier Paolo Pasolini, Umberto Eco and Christian Metz, expended considerable effort in investigating the literal equivalences between the signs and structure of natural language and those of cinema. Much of this formative debate centered on the issue of the nature of the minimal semiotic unit, as defined within both disciplines.

The French linguist, Andre Martinet, had developed the concept of double articulation to refer to the two structural levels on which natural languages are organized. Speech could be analyzed into morphemes (significant units), which constitute the first articulation and phonemes (meaningless sound units), considered as the second articulation. Christian Metz, whose self-declared purpose was to "get to the bottom of the linguistic metaphor", argued against the notion of cinematic double articulation. He claimed that there was no corresponding phoneme in film, as even a single frame, possibly the smallest unit in cinema, signifies in and of itself.

3.5.3 Christian Metz

Metz is a seminal figure in the field of film semiotics, who sought to extend and integrate the insights of Saussurean linguistics with the poetics of the Formalists. His early work was predicated around determining whether cinema was a language or a language system. He concluded that cinema was a language without a language system, but his reasoning deserves close examination. He rejected cinema as a language system primarily because "it lacks the equivalent of the arbitrary linguistic sign" [Metz 1974]. He also argued that cinema could not be considered a language system as it allowed only for deferred communication: that which is one-way and involves a considerable time-lapse between production and viewing, and reception and critique. The world of cinematic entertainment has changed considerably since Metz first propounded his theories, and this latter argument in particular is somewhat challenged by the immediacy of today's Internet distribution channels and the evolving relationship of the consumer as producer.

In Language and Cinema, Metz further argued that the analogy of the shot as a word and the scene as a sentence was deeply problematic [Metz 1974]. He maintained that shots are the creation of the filmmaker and infinite in number, unlike words, which are from a lexicon that is finite in principle. He also stated that shots do not gain meaning by paradigmatic contrast with other shots that might have occurred in the same place on the syntagmatic chain. Metz declared that to speak a language is simply to use it, while to speak "cinematic language" is always to a certain extent to invent it. By this he meant that, unlike the straightforward process of producing sentences, to produce so-called filmic utterances required talent, training and access.

Metz's most influential contribution to film semiotics was his Grand Syntagmatique system, where he attempted to isolate the main syntagmatic figures of narrative cinema and, in so doing, provide a precise terminology for the description of film.

The Grand Syntagmatique describes eight diverse ways in which time and space can be ordered through editing within the segments of a narrative film, including the autonomous shot, alternating motifs, bracket sequences, scenes and episodic sequences. Responding to criticism about the narrow sector of movies that the system could only be applied to (mainstream Hollywood narrative), Metz later modified his claim about the system to redefine it as one of many cinematic codes, which could "serve as an attention-focusing device, of interest even when only partially applicable."[Stam, 1992: 48].

3.5.4 Cognitive Semiotics of Film

Over the past thirty years or so, Metz's film semiotics have been rejected, reexamined and reworked by means of theories of post-structuralism (Stephen Heath, Colin MacCabe), pragmatics (Roger Odin), cognitive science (David Bordwell, Edward Branigan) and transformational generative grammar (Michel Colin, Dominique Chateau). Of these developments, cognitive film semiotics can be seen to represent the next stage of semiotic film theory. It aims to model the actual mental activities/intuitive knowledge/competence involved in the making and understanding of filmic texts. This approach positions itself somewhere between the linguistic determinism of Metz and the rational autonomy that cognitivists confer upon film spectators.

Of interest to this research in particular is Roger Odin's contribution to film theory, which is based on an analysis of the relationship between semiotics, pragmatics and cognitivism. His approach, which can be called semio-pragmatic, deals with the contention that the ontology of a film is not automatically fixed in advance, but is negotiated and determined through the process of watching/reading a film [Buckland 2000: 78]. The semio-pragmatic approach to filmic texts is chiefly concerned with the production of meaning from the engagement of the spectator with the text. Odin's pragmatic framework is based on the immediate discursive nature of film, drawing on Metz's notion of cinema as 'langage sans langue'- a language without a language system.

Odin alludes to Metz's comparison of the analysis of a film by a semiotician and the reading of a film by a spectator. "The path that the semiotician follows is (ideally) parallel to that of the film viewer. It is the path of 'reading', not of 'composition'. But the semiotician forces himself to make explicit this procedure, step by step, while the viewer practices it directly and implicitly, wanting above all to understand how a film is understood." [Metz 1974: 73]. Odin concludes that what Metz was ultimately trying to get at was to understand how film is understood.

Odin is interested in understanding the function of audio-visual productions in a given social space. To frame his argument, he talks of meaning being constrained by "Institutions", which he assigns as being both external (commercial cinemas, schools, art house cinemas) and internal (discursive competence comprising modes and operations). These modes and operation constitute the filmic dimension of the spectator's competence. Modes of filmic meaning that a spectator might infer from an audio-visual production include: spectacle, fictional, dynamic, home movie, documentary, instructional, artistic and aesthetic.

This research aims to examine Odin's framework within the context of new developments in the domain of cinematic production and distribution. Audiovisual productions can now be both produced and viewed over the Internet by a distributed group of collaborators and viewers. One of the functions of this work is to examine the changes, if any, that these new institutional parameters bring to the reception and comprehension of audiovisual content.

3.6 Related Work

3.6.1 Video Annotation, Retrieval and Construction

Interactive Cinema

The Interactive Cinema Group in the Media Lab has always engaged in expanding and diversifying the role and function of, and the relationship between, the storymaker, the storyteller and the story audience. The research described in this thesis seeks in particular to build upon the group's explorations into video annotation, retrieval and construction.

Amy Bruckman's *Electronic Scrapbook* is a system designed for annotating home-video footage, where the attached descriptions can subsequently be used in a story template system that helps automate the editing process [Bruckman 1991]. Electronic Scrapbook utilizes an inference mechanism to suggest potential annotations to the moviemaker for describing new footage. The system also uses an adapted form of case-based reasoning to help the moviemaker construct personalized movies. Borrowing from Seymour Papert's constructionist theories of "objects to think with", Bruckman describes segments of video clips as objects that can be put somewhere, described, copied or thrown away.

Thomas Aguierre Smith's *Stratagraph* enables ethnographic filmmakers to attach a richly layered set of contextual annotations to a video stream. Implemented as the *Anthropologist's Video Notebook*, this system deals with the joint issues of "how to represent dynamically changing content information in a video database system" and "how to describe video in such a way that these descriptions can serve to generate new content." [Smith 1992]. Users can assign descriptions to contiguous sequences of video frames, called "stratums", that can overlap or even totally encompass each other. FIGURE 3-9 shows a graphical representation of the streams of annotation attributed to a movie sequence using Smith's method of Stratification.

Instead of creating a database of video objects or chunks of frames, a database of descriptions which have specific frame addresses is constructed. The video streams can in turn be edited, and so the list of descriptive annotations for any particular section of video will develop to reflect not only the context of the source material itself, but also the context of how it has been reused in multiple edited sequences.

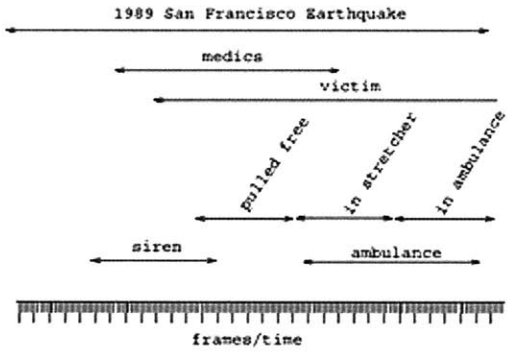


FIGURE 3-9: Annotation of Video Footage Using Stratification

Ryan Evans created two interrelating tools, *Logboy* and *FilterGirl,* designed for creating 'multivariant' movies or movies that play out differently each time they are presented [Evans 1994]. Video clips are annotated using descriptive categorizations such as Scene, Setting, Character, Point of View and Framing in the LogBoy video database tool. FilterGirl is a story modeling tool that allows the moviemaker to construct story constraints or filters that govern the selection and playback process of the annotated clips. The filters created can be combined in a number of ways:

- Temporally - directing story structures according to time
- Logically - allowing for Boolean combinations or descriptions
- Contextually – check annotations for continuity matching
- Viewer Preference – provides for viewer interaction

As Evans points out, neither of these tools can actually understand stories, but rather provide filmmakers "with a way to think about and construct multivariant stories from descriptions and constraints."

The PlusShorts system seeks to provide an environment that allows filmmakers to think and share conceptual ideas about movie structure. An iconic symbol set of punctuation icons is provided for annotating and marking up the structure of movie sequences, constructed from the video 'objects' in the movie database. Similar to Smith's notion of annotation description's 'thickening' over time as movie segments are reused, a user's understanding of a movie clip or segment in the PlusShorts system deepens as she observes not only how it has been reused in other sequences but also how the author of the sequence has chose to describe it. Users of the PlusShorts system can also view the punctuation markup as resembling the story templates of Bruckman's Electronic Scrapbook or the story filters of Evan's FilterGirl. The PlusShorts system allows users to easily replicate the sequence structure of another contributor by replacing the content of the original author's with their own.

Media Streams

Marc Davis' *Media Streams* introduces an iconic visual language for video content representation. This system allows users to create multi-layered annotations of streams of video data using a library of over 3,000 hierarchically structured, composable, and searchable iconic primitives. FIGURE 3-10 shows a selection of pictorial icons from the Media Streams lexicon.

FIGURE 3-10: Selection of icons from the Media Streams lexicon

It attempts to create both a standardized global human, and, computer readable and writable visual language for representing the consensual interpretations of video content. [Davis 1993].

Annotations are created by dragging icon descriptors onto a Media Timeline, where they are used to temporally index the video stream. FIGURE 3-11 shows a screengrab of the Media Streams Timeline, complete with iconic annotations arranged along the linear streamlines. Icons can be organized across multiple axes corresponding to various aspects of audio-visual content such as characters, location, actions and transitions. Davis describes the system as possessing a syntax for the composition of iconic sentences and a means for extending the visual language.

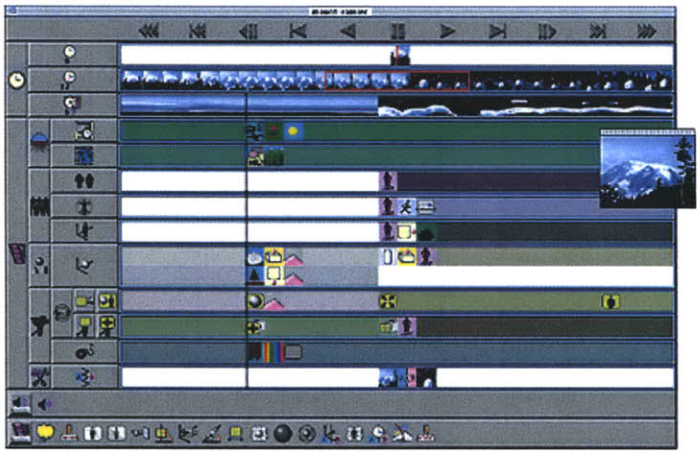


FIGURE 3-11: Screenshot of Media Streams Timeline

Unlike the continuous video streams that the Media Streams system attempts to annotate, the PlusShorts systems aims to describe and annotate the structure of edited sequences composed of segmented clips that have already been separated from their original context of origin. The purpose of PlusShorts is not to describe the actual content of the video clips, but rather denote the conceptual thinking underlying the editing choices of the author. The limited set of visual icons used in PlusShorts, are also already familiar, both visually and semantically, to most users of the system. However, the *exact* meaning of these icons in the context of their use in describing video structure is open to interpretation, discussion and agreement between users of the system.

Algebraic Video for Composition

The Programming Systems Research Group at MIT's Laboratory for Computer Science proposed an algebraic video data model that "provides operations for the composition, search, navigation and playback of digital video presentations." [Weiss et al 1994]. These operations function both to attach attributes to, and temporally and spatially combine, video segments. Their model supports the hierarchical nesting of video expressions and the inheritance of attributes, which can also be used as search criteria to discover video streams of interest and explore a presentation's context. For example, a user might come across an interesting algebraic expression and then choose to examine the encompassing video segments. Expressions can be nested hierarchically, where overlapping segments allow for multiple coexisting annotations and views of the same data, thus enabling the user to assign and interpret numerous meanings from the same audiovisual content. New video expressions are created from a database of video clips by combining desired video segments using algebraic combinations. FIGURE 3-12 shows two screenshots of the Algebraic Video Browser, containing two different 'video expressions' of varying spatial layout.

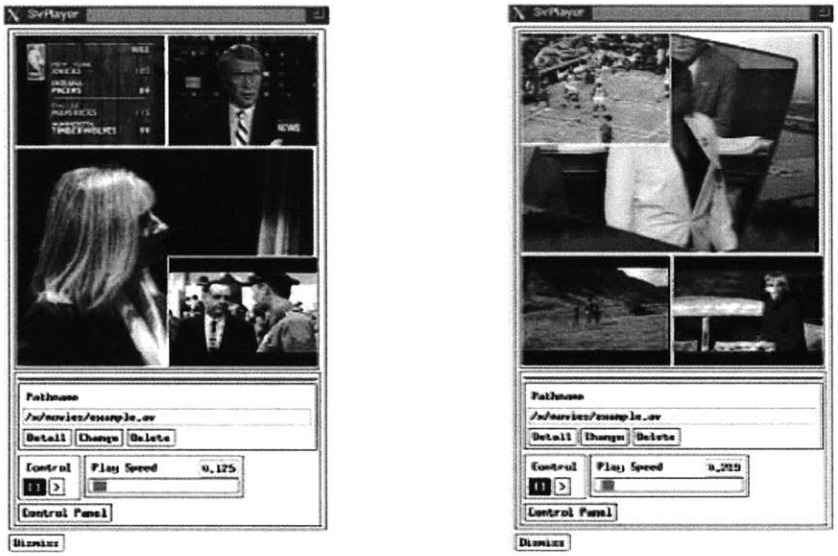


FIGURE 3-12: Algebraic Video Browser Snapshots

3.6.2 Online Collaborative Tools

Structured Online Collaboration of Graphic (Re)Interpretation

The Versus Project concept is described on its website as: "Two designers take turns providing their own interpretations of each other's work. Beginning with an original image and cycling back-and-forth between each other with the product. Leaving it entirely to the other designer to modify it as they wish under the time constraint of one hour. The process is repeated until one designer tires or gives up."[http://www.bloop.org/choco/versus/main.htm]

Each designer can also submit an additional piece of text with every image outlining the reasoning, inspiration and motivation behind the changes he/she has made. This commentary is then displayed underneath or alongside the presented image. FIGURE 3-13 shows ten successive screenshots of the images and text alternately submitted by designers Mike Cina and Harsh Patel who formed one face-off pair in the Versus Project. The designers used the text area to describe anything from how they were feeling that day, to what music they were listening to, to what it was they had found inspiring in the other's work.

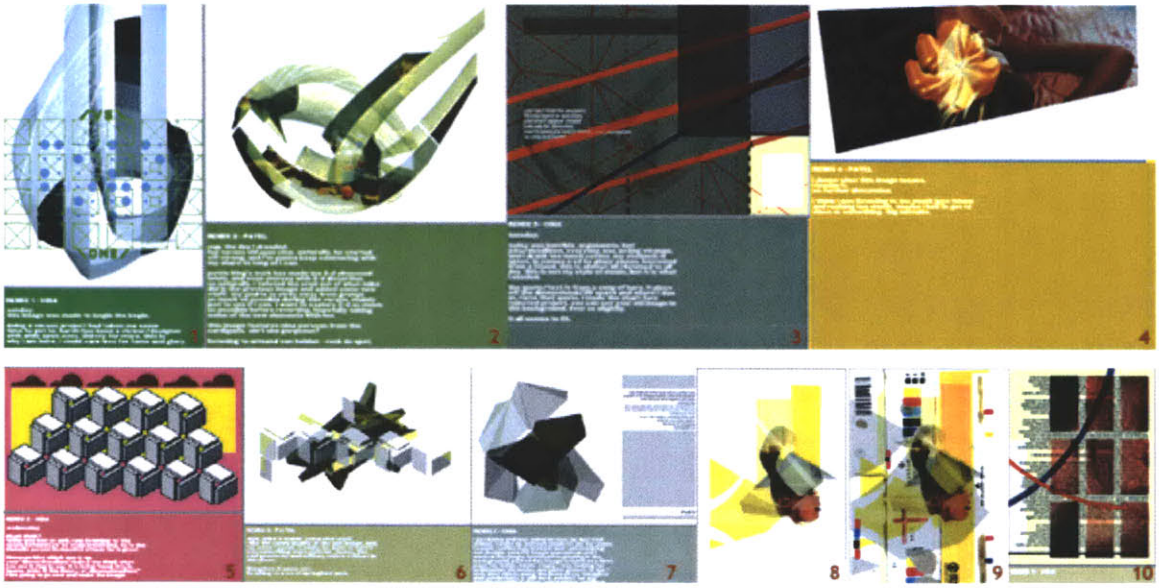


FIGURE 3-13: Successive Screenshots from the Versus Project (Cina vs Patel)

Visitors to the Versus Project website can follow the progress of the graphic challenge by clicking through successive web pages, each of which displays on one page both the submitted artifact and the description of the reasoning behind it. With the PlusShorts application I also want to create a structured, albeit asynchronous and non-time dependent environment, where artists and filmmakers can collaborate together and develop evolving and revealing works. I am interested in replicating the Versus Project method of displaying in parallel, both the artifact, and the conceptual thinking behind it. It is my intention to integrate the author/filmmaker's commentary into the PlusShorts interface in such a way that it can be observed as the movie sequence plays out.

Online Collaborative Tools for Moviemaking

There are currently few tools available that allow users to collaborate together in an online environment on creating movies using self-generated content. The Mimicry system comprises a Java applet that allows users to dynamically create links to and from temporal media on the Web regardless of the users' ownership of the Web pages or media clips involved [Bouvin 1999]. Websites such as VideoFarm.com, Earthnoise.com and Fjallfil.com provide online editing tools that allow users to modify and share their own material exclusively. GetMusic's *VideoLab* enables users to create Flash based music videos online. Users can upload their own still images to combine with the server provided images, videos, sound effects and stills [FIGURE 3-14], using a real-time editing interface with basic transition effects. [FIGURE 3-15]

FIGURE 3-14: Screenshot of VideoLab's Content Selection Interface

FIGURE 3-15: Screenshot of Videolab's Real Time Editor Interface

I-Views and the Shareable Media Project

Pengkai Pan's *I-Views* is an online asynchronous story sharing system that allows individuals to use communally owned media clips to create narrative sequences that are subsequently compared and used to help foster the development of communities of similar interests. [Pan, 1999]. The I-Views system incorporated and interweaved two distinct tool types: the first a video editing tool and the second a virtual community building tool. The editor allowed users to view, edit, compare, evaluate and discuss streaming video content, while the community focused section allowed users to initiate dialogue with other participants with similarly matched interests. This system of shared authorship, tools and virtual environments demonstrated a new story form: "Shareable Documentary."

The result of Pan's initial exploration has been the *Shareable Media Project*, a research initiative in the Media Lab's Interactive Cinema Group involving four graduate students and a number of undergraduate researchers. This project represents an attempt to provide a coherent structure that will facilitate distributed collaboration and communication among filmmakers, storytellers, artists and audiences.[Kelliher et al, 2000] The aims of the project are threefold:

- Infrastructure: create a scaleable, extensible database and network
- Interface: create dynamic, novel visualization tools for video playback
- Community: facilitate in the construction and maintenance of online video communities

The Shareable Media Architecture developed to support this project has the capability to deploy multiple applications on wired and wireless devices through both narrow and broadband connections [Lin et al, 2000] Applications are developed in accordance with the published API for the system, called the Shareable Media Framework.

The online version of the PlusShorts software was developed to operate as part of the Shareable Media System. In addressing the overarching goals of the Shareable Project, it seeks to provide a novel, compelling and easy-to use tool for both the effective visualization and construction of streaming video content. Designed to foster interaction and collaboration between filmmakers online, the PlusShorts application draws on the endeavors of the I-Views system to encourage the development and maintenance of virtual communities of users.

3.7 Summary

In this chapter I have described the development and evolution of instructional notation systems relating to the expression and description of creative works. I discussed the historical evolution of punctuation notation, chosen as the representative symbol set in this research owing to its well-established familiarity, communally agreed upon meaning, and openness to application and reinterpretation in other non-writerly domains. I examined the creative application of this notation scheme in the fields of literature, art and dance, where it has been used not only as a structuring tool, but also to interpret and convey meaning.

The broader field of annotation was investigated, and the use of annotation to describe a document's conceptual meaning and to represent video was explained. The structuring of cinematic and audiovisual content by editing techniques was examined in relation to historical development, genre and sound. I then described the relationship between cinema and language from semiotic, cine-semiotic and cognitive semiotic perspectives.

This investigation informs the theoretical assumptions made concerning how filmmakers and audiences construct and subsequently negotiate the meaning of a filmic text. Of interest to this research is to investigate how these assumptions are challenged or supported by the delivery of cinematic content over the Internet, the distributed collaborative process of making an online movie, and the nature of the cinematic artifact itself. In concluding this chapter, I reviewed practical applications and prior research conducted within the fields of video annotation/ retrieval and distributed story-sharing networks.

4.0 Design and Implementation

4.1 Motivation

4.1.1 Found Footage Films and Collaborative Co-Construction

The movies of found footage filmmaker Jay Rosenblatt have been described as "fusing optical printing, hand processing and an abundance of found footage – public service films, lush Technicolor pics and garish porn - in the service of a kind of cultural dream analysis, a place where "meaning" is found at the intersection of the mundane and the bizarre." [Golembiewski 2000]. The films of Rosenblatt, along with those of Craig Baldwin and Bruce Connor amongst others, are wonderfully compiled cinematic collages using material appropriated from a wide range of diverse sources.

The Shareable Media system, described in Section 3.6.2, provides users with a large database of video clips (found footage as such) that can be appropriated and reused to create new sequences that divulge alternative meanings and interpretations of the content. With the PlusShorts application, I want filmmakers to be able not only to engage in radical rereadings of the shared content, but also to reveal the conceptual thinking underlying their work and the methodology governing it. By introducing an iconic mark up language, I want to engage filmmakers in beginning a dialogue with other producers/artists as they strive to interpret and reach concordance or create dissonance around the meanings of the symbols and the semantic structures they seek to define.

4.1.2 Cinema and Language in a Networked Environment

This work also seeks to reexamine and reassess the relationship between cinema and language, in light of recent developments in networked communication, the proliferation of cheap digital video cameras and the emergence of distributed moviemaking/viewing. The Shareable Media System aims to challenge the binary notion of the producer versus the consumer and promote instead a more fluid and complementary understanding of the relationship between these two roles. This system allows for the development of applications supporting collaborative activities. These applications ultimately aspire to move towards an audiovisual experience that is multi-user, multi-directional and open to all.

The PlusShorts application will support the construction of many types of 'filmic utterances', in a straightforward process that questions the narrow definition of what constitutes a correct cinematic construction or a 'proper filmmaker', as espoused by Christian Metz [Section 3.5.3.]. With the PlusShorts application, I want to create an environment that exploits the notion of cinematic possibility – an environment that supports the paradigmatic contrast of shot use in a sequence - in deference to Metz' claim that this was simply not possible with movies [Section 3.5.3]. Reexamining and revisiting theories of film semiotics and cinematic language, originally derived from and applied to the traditional cinematic domain, in this emerging, open-ended environment is another primary goal of this project.

4.1.3 Teaching Tool for Moviemaking

In teaching video editing and production to students previously, I have always accompanied the lessons with screenings showing segments from movies skillfully illustrating whatever cinematic principle is being studied in the class, such as lighting, shooting dialogue or editing a sequence to create suspense.

By observing, analyzing and interpreting the production methodologies of others, students develop not only a language and terminology for discussing movies, but also a keen understanding of a director's style, intent and purpose. It is my aim with the PlusShorts software to develop an application that allows filmmakers to share ideas about their work, analyze the contributions of others, and engage in collaborative exercises exploring the process of filmmaking.

4.2 Scope

The PlusShorts Java application functions in a standalone environment or on the WWW as part of the Shareable Media System. The application allows a distributed group of users to contribute to, and collaborate upon, the creation of shared movie sequences constructed from user submitted movie clips stored in the central database/server. This application additionally provides a visual markup language, in the form of punctuation symbols that describe the conceptual thinking underlying the structure of the movie sequence. The application, although open to use by anyone, has been specifically designed to be used and evaluated by filmmakers collaborating together on the co-construction of movie sequences.

4.3 System Overview

The PlusShorts application, in its design and functionality approaches a new type of story form, where the navigational tools for finding content, the construction area used for sequencing content, and the final playback mechanism, operate seamlessly together to inform and evidence the meanings produced by each component [Kelliher et al 2000]. One of the main design goals I retained from my first rough sketches of possible interfaces for the PlusShorts application was the encapsulation of all the functional elements in one screen.

I wanted the user to be able to find sequences, search for clips, edit sequences and playback movies all within the one screen. I imagined a scenario in which a user could be focusing on watching a movie playback, while also being peripherally aware not only of the author's description of the structure of the sequence, but also of the sequence's relationship with other sequences in the system. I wanted to create an all-inclusive interface that would be straightforward to use and would integrate all components of the task in one screen. The final PlusShorts application interface, in both its design and functionality, aims to facilitate and fulfill this particular design goal.

4.4 Interface

The PlusShorts interface contains three main activity areas: the sequence navigation and selection interface, the sequence construction interface and the sequence playback interface. FIGURE 4-1 shows a screen shot of the PlusShorts application, with each specific activity area clearly marked.

FIGURE 4-1: PlusShorts Application Interface

4.4.1 Sequence Navigation and Selection Design

This element of the application resembles a newsgroup interface, as shown in FIGURE 4-2. Newsgroups, among the oldest forms of communication on the Internet, are designed to support discussions of particular topics among people connected to the Internet. They utilize a technique called threading, whereby topics posted, that have replies subsequently attached to them, are displayed as header threads, which can be expanded to reveal the hierarchically structured replies underneath.

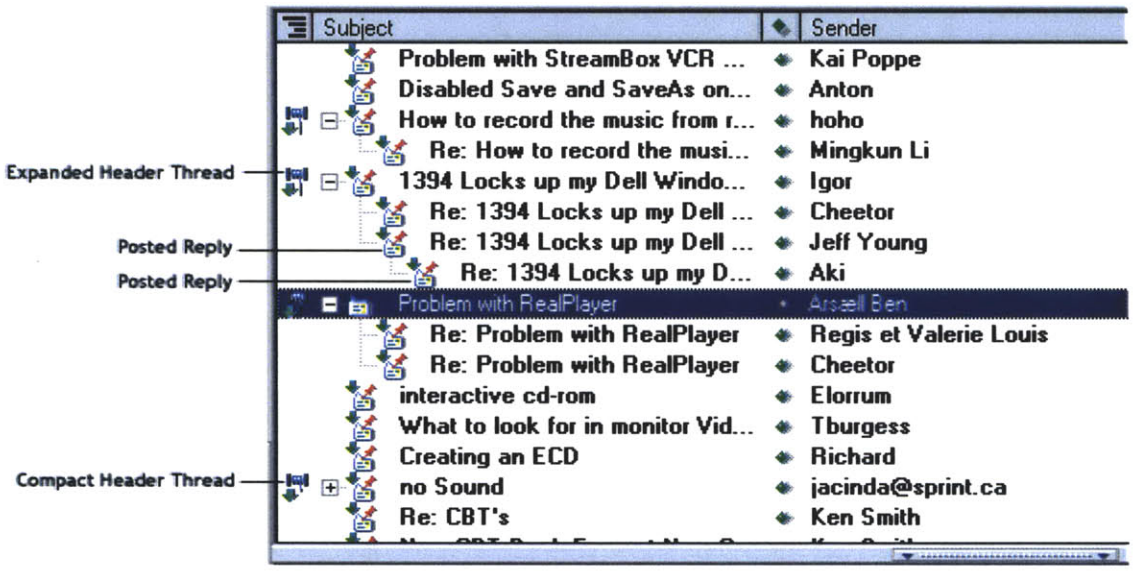


FIGURE 4-2: Newsgroup Interface

In the PlusShorts Application, original movie sequences are conceived as "movie threads" and sequences submitted in response to these threads are considered as "movie posts". These movie threads and posts are displayed and navigated the same way that a newsgroup thread is displayed and navigated.

Function

Similar to threads in a Newsgroup, movie threads can be expanded to display associated movie posts in a hierarchical structure that denotes the development and evolution of the movie thread over time. The threads can be sorted and displayed chronologically by date of thread submission, or according to author, whereby all sequences (threads and posts) created by an individual author are accessible by selecting the author's name. This area of the interface also contains a search component, whereby users can utilize keyword searching to query both at a sequence name and a clip name level. Once a user selects a particular thread or post, the corresponding sequence is loaded into the sequence construction area. FIGURE 4-3 shows a screenshot of the navigation and selection interface in the PlusShorts application.

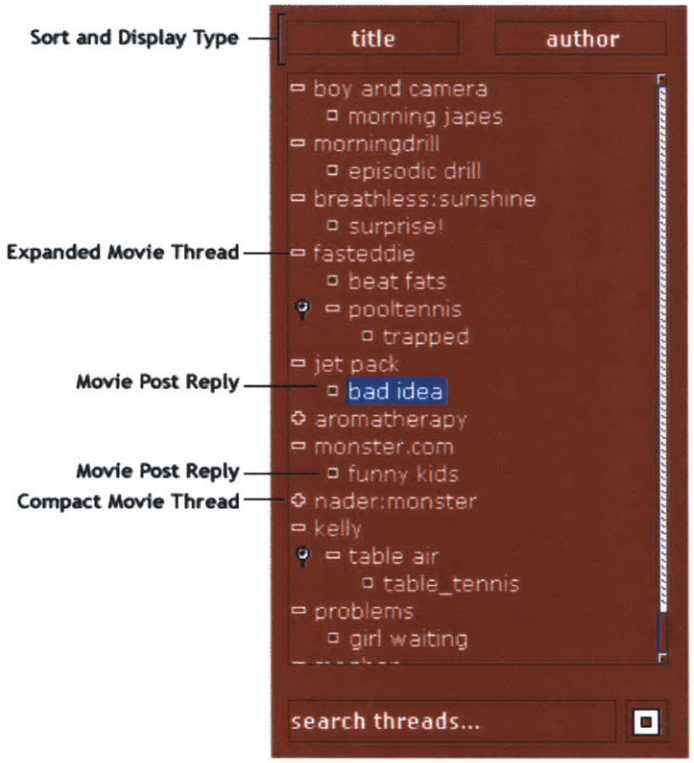


FIGURE 4-3: PlusShorts Sequence Navigation and Selection Interface

The user can use the navigation interface to trace the evolution of a movie thread, from its initial submission, through its multifarious iterations, by selecting and viewing the sequences constructed by all contributors in that thread.

4.4.2 Sequence Construction

This area of the application interface encompasses three components: punctuation icons, sequence construction grid and search area.

Punctuation Icons

Design

Punctuation notation has been used to structure texts, to comment directly on the content of the text it is containing, and, as a metaphorical tool for learning [Section 3.2]. In section 3.1, I described the importance of the ability of the selected symbol set to represent all aspects of the ontological type it is seeking to describe.

The punctuation symbols selected for use in the PlusShorts application, (as shown in FIGURE 4-4), were chosen because of the applicability of their original function and purpose to metaphorically describe both spatial and temporal relationships between shots in a movie sequence.

FIGURE 4-4: Punctuation Symbols used in the PlusShorts Application

For example, David Bordwell describes the presentation of action in such a way that it occupies less time on the screen than in the story, as elliptical editing [Bordwell 1997: 283]. Film directors can use optical effects such as the dissolve, wipe or fade to indicate that time has been omitted, or they can perhaps use a conventional shot change, such as the cutaway, which shows an event occurring elsewhere that does not last as long as the omitted action. The ellipses punctuation icon has an obvious correlation with this aspect of movie structuring. Similarly, Christian Metz described the bracket sequence structure in his Grand Syntagmatique system as pertaining to brief, conceptually driven scenes in a movie [Section 3.5.3]. There is a direct iconic translation of this cinematic conceit in the form of the parentheses symbols.

Temporal editing techniques used to manipulate and control events in the story time can be represented using symbols 'borrowed' from numeric notation. For example, Errol Morris's *The Thin Blue Line* (1988) intersperses present-day footage of suspects and witnesses, with flashback scenes depicting reenactments of the crime committed. *They Shoot Horses, Don't They?* (1969) uses flashforward shots throughout the movie, which hint and allude to the film's final ending. Both of these structuring methods for controlling the timing of depicted events in a movie could be represented using the 'minus' (flashback) and 'plus' (flashforward) icons.

In Section 3.1, I discussed how the meaning of notation systems can be governed, as in the case of the alphabet, by public consensus. In conducting this research, I am interested in observing how people use the symbols, attempt to interpret what others have done, engage in dialogue and discussion, and, ultimately come to an agreement about what they mean within the cinematic domain. For my own purposes in using the application, I imagined the symbols could be used as follows:

:	follows an establishing shot
,	separates shots in a scene
.	denotes the end of a scene
"	denotes a shot containing dialogue
...	denotes shots omitted – spatial or temporal ellipse
( )	to separate significant sections (shots/segments/scenes)
+	joins sections together or indicates a flashforward
-	indicates a flashback
=	placed before a concluding or resolution shot/sequence
?	direct reference to author's intent – query of edit decision or emphasis of certain section
!	expression of author's feeling about a shot/sequence

Function

The punctuation symbols can be dragged into the sequence construction grid and placed between the thumbnail images representing the movie clips. The symbols do not affect the physical playback of the movie sequences, but exist rather as an additional layer of interpretive information to be referenced as the sequence plays out.

Sequence Construction Grid

Design

The idea of using a grid layout for the construction area was inspired by the storyboard layouts used in the planning stages of animation and movie productions. Before filming has even begun, the director, the cinematographer and the entire crew can have a clear and cogent understanding of how the film is ultimately going to look. Alfred Hitchcock, a major proponent of this pre-production method, explained to François Truffaut that once the storyboards were completed, he never had to look through the camera during production, because he already knew what he would see [Universal Studio Archives 2001]. FIGURE 4-5 shows the original storyboard from the opening Rosebud scene in *Citizen Kane* (1941), which illustrates a clear sense of camera movement, mise-en-scene and the sequential order of the scene.

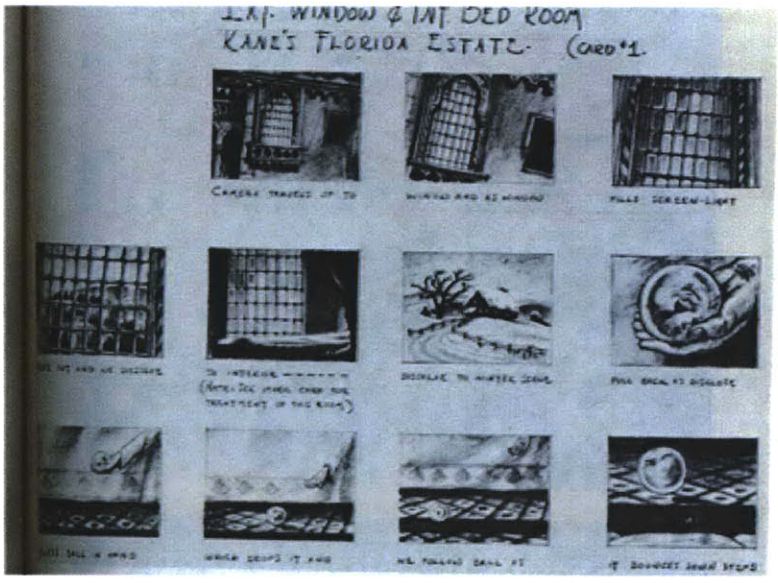


FIGURE 4-5: Storyboard from Rosebud Scene in *Citizen Kane* (1941)

Animation storyboards in particular lend themselves to being constantly rearranged and reorganized, as the narrative direction of the story is negotiated by moving individual drawings around on a pin board.

FIGURE 4-6 depicts two storyboard segments from Pixar's *Toy Story 2* (1999) animated film. These rough sketches allow the story developers to gain an initial sense of the character development and narrative progress of the film, before the arduous task of actually creating the final rendered images begins.

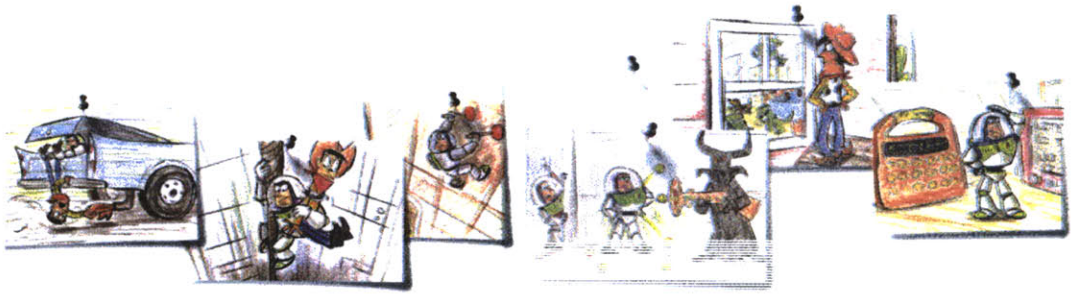


FIGURE 4-6: Storyboard segments from *ToyStory 2* (1999)

In the PlusShorts application, the sequence construction grid consists of a 5*5 grid of panels. There are in fact three identical grid "pages" which can be flipped through using the numbers situated above the grid, making 75 panels altogether. Given the network capabilities of most users and buffering issues associated with streaming content, it was decided to limit the length of the sequences to a maximum of 75 clips. This could be easily extended, if initial evaluation results expressed a user desire for such an expansion. Within the grid, users can manipulate the thumbnail images that represent each video clip stored in the database. These thumbnails are automatically generated when a clip is first uploaded to the website/added to the database. FIGURE 4-7 shows a screenshot of the PlusShorts construction grid interface, with a sequence denoting a popular television advertisement loaded within. The punctuation symbols have been dragged from the punctuation grid and placed between the clip thumbnails to annotate the structure of the advertising sequence.

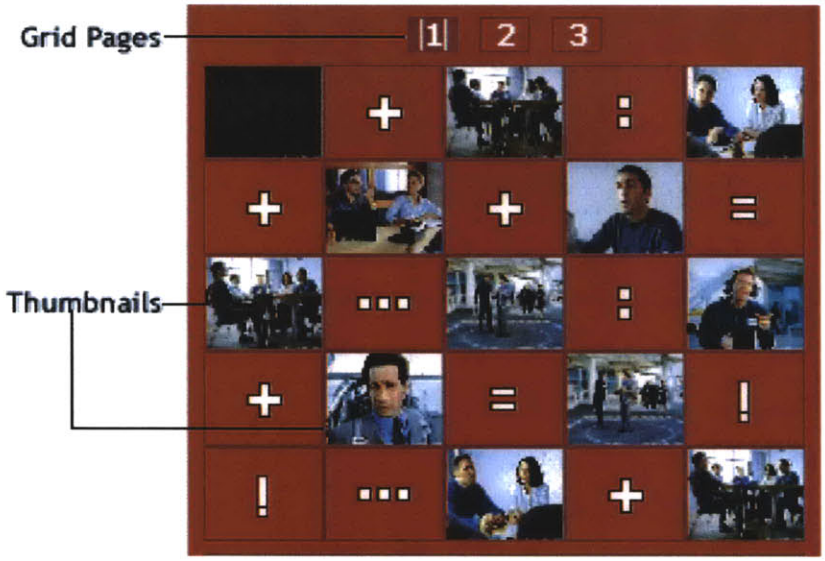


FIGURE 4-7: Sequence Construction Grid in the PlusShorts Application

Function

These representational thumbnails can be moved around, copied, cut, pasted and deleted to form new sequence layouts and alternative interpretations of video threads. Thumbnails and/or groups of thumbnails can be copied from one thread and pasted into another. Thumbnails can also be selected and used as search criteria, where the returned results, displayed in the selection and navigation area, are other threads containing that thumbnail/group of thumbnails. This facility allows a user to investigate not only how a particular clip(s) has been appropriated and reused in other sequences, but also to examine how different users have described the relationship between that clip and other clips in the sequence using the punctuation icons. Like Smith's Stratification annotation scheme, which allowed for multiple layers of description to be attached to video data, clips in the PlusShorts system can also have multiple representations attached to them, creating a 'thick description' [Smith 1992]. The grid can be entirely cleared to begin a new thread, and new clips can be added to the construction area using the search clip component situated underneath the main grid area. The movie sequence contained in the grid will play back according to a left to right, top down rule, similar to reading a Western comic strip.

Search Clip Component

Design

The search results display area replicates the grid structure of the construction area and is situated directly beneath it to facilitate the quick and easy dragging of desired clips. FIGURE 4-8 shows a screenshot of the PlusShorts application search and search results display, which contains five thumbnails returned from the keyword query "kelly".

FIGURE 4-8: Clip Search Component in the PlusShorts Application

Function

Users can utilize keyword searching to find clips in the database that they would like to add to their sequences. Results are returned in the display grid above the search text box, and can be dragged into the construction grid. The display grid shows five of up to fifteen returned results at a time and the user can cycle through the returned clips by clicking on the panels representing each set of five results, located above the text area.

4.4.3 Sequence Playback Interface

Design

This element of the PlusShorts interface is contained within the application itself in the offline version, but is outside of the java applet in the web version and embedded instead in an external HTML page. However, the location and look of the player is identical in each version.

Function

Users can play back the entire movie sequence by clicking on the play button located to the right of the punctuation grid. To play just a segment of the movie or a random selection of clips, a user can highlight just the clips she wants to see in the construction grid before selecting the play button. Having viewed the sequence, if the user is happy with her construction, she can name the sequence and then upload it to the database by selecting the submit icon, located to the left of the punctuation grid. FIGURE 4-9 shows a screenshot of the playback interface component in the PlusShorts application. The grid construction area contains a sequence depicting a television advertisement. The user has highlighted the first thirteen panels to indicate that she only wants to play the sequence as far as the thirteenth panel. The movie sequence plays out in the QuickTime Player located to the right of the construction grid.

FIGURE 4-9: Screengrab of selected clips playing out in the QT Player

4.5 Design Evolution

There are two major intertwining strands of software development accompanying this research. One has been the development of an application to evaluate the contention regarding the iconic description and interpretation of cinematic structure. The other has been the development of an online version of this application that is integrated with the Shareable Media Project. Both of these endeavors were carried out simultaneously, and both served to influence and inform the design and evolution of the other.

4.5.1 PlusShorts Application

The initial prototype of the PlusShorts application software was developed as a standalone Java application that utilized the native QuickTime, Java and QuickTime for Java API's. The QuickTime for Java API provides both an object model for the QuickTime API and a logical translation or binding of QuickTime's native calls into Java method calls. In this initial prototype, in addition to the standard java classes dealing with GUI layout and functionality, a QuickTime Sequence layer was developed to interface between the main PlusShorts java classes and the QuickTime player embedded within the application.

All media content (movies and thumbnails) was stored in a flat directory structure and passed to the QuickTime player when called by the QuickTime sequence layer. This version of the software allowed users to create, edit, markup and save sequences using a limited content set. Sequence threads could be accessed using a list structure and basic keyword searching of the clips was supported. This prototype was demonstrated to numerous colleagues, friends and lab sponsors and their comments and feedback proved instrumental in influencing the future designs of the application. FIGURE 4-10 shows a screenshot of the interface for the first PlusShorts application prototype.

FIGURE 4-10: Early PlusShorts Application Prototype

The final version of the local PlusShorts application was tested and evaluated at a workshop for filmmakers conducted at the MediaLabEurope, Dublin. For the purposes of the workshop, the PlusShorts application ran over the local network in the Dublin lab. All media content, movie clip keyword and movie sequence information was stored in a shared directory on one designated "central" computer, and each individual workstation running the PlusShorts application accessed the movies and thumbnails from there.

Influenced by concurrent developments with the online version of the software, the PlusShorts application used in the workshop featured a lightweight front-end dealing with GUI functionality, separate utility classes responsible for calling and retrieving content related data and the use of SMIL files to send reference data to the QuickTime player [see section 4.5.2 for more information on the use of SMIL in the PlusShorts software]. The hierarchical display, searching and sorting of video threads/posts was also fully supported. The interface for this application of the software is shown in FIGURE 4-1.

PlusShorts Applet and the Shareable Media Project

In November 2000, the first version of the Shareable Media System, developed by the Interactive Cinema Group, came online. The system is organized according to a three-tier model, comprising a low-level Data Tier (media content), middle-level Business Logic Tier (Shareable Media Architecture) and high-level Presentation Tier (containing front-end applications such as PlusShorts). The Shareable Media Architecture (SMA) itself consists of three modules: Shareable Media Framework, Application Manager and Storage Manager. Software developers can use the API provided by the Shareable Media Framework to develop front-end applications that can be registered, tested and released using the Application Manager. The Application Manager and the Shareable Media Framework can access, change and update content through the Storage Manager that functions as the only gateway to the content data [Lin et al, 2000]. FIGURE 4-11 shows an outline diagram of the Shareable Media Architecture, depicting all three layers of the system and their requisite components.

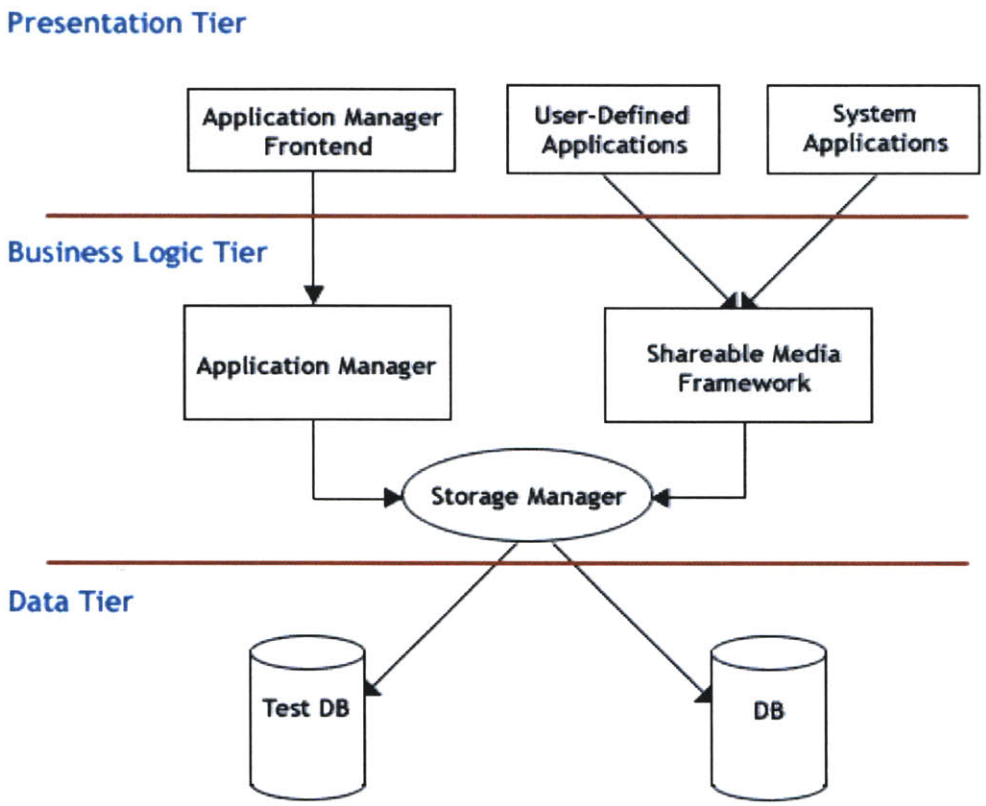


FIGURE 4-11: Diagram of the Shareable Media 3-Tier Architecture

The first online version of the PlusShorts application was prepared in accordance with the Shareable Media Framework and testing began on the PlusShorts Java applet shortly before Christmas. A set of utility classes was developed to deal with all functional access queries to the database (save sequence, search, find clip etc.), while the main PlusShorts classes once again handled all GUI functionality. The QuickTime player initially remained integrated within the PlusShorts applet, similar to the standalone application. However, as this caused the applet to function very slowly and ultimately freeze after a number of sequences had played out, it was decided to separate the QuickTime player from the applet and integrate it into the web browser instead.

This refinement resulted in a version of the applet that was considerably more lightweight than its predecessor. Movie sequences would now be sent to the QuickTime player from the Shareable Media server as SMIL files. SMIL (Synchronized Multimedia Integration Language) is a markup language designed to facilitate the positioning and coordinate the sequencing of media elements on the WWW. A SMIL servlet was developed to handle all calls from the PlusShorts applet to the database. This servlet returned a SMIL file containing references to the desired movie clips and the browser subsequently retrieved the movie files and played them out in sequential form in the QuickTime player. This version of the software worked much faster than previous versions and also saw the development of greater GUI functionality, such as the implementation of the hierarchical navigation structure.

The Shareable Media Project is a long-term initiative in the Interactive Cinema Group. To further facilitate the development of future applications and the upgrading and maintenance of current applications, the Shareable team decided in April 2001 to upgrade the application server module of the system from the J2EE model to the more robust Borland Application Server. This development necessitated the reconfiguring of several aspects of the Shareable Media Architecture, and consequently, the reconfiguring of the application software it was supporting. At the time of writing this document, the PlusShorts software is being integrated with this latest iteration of the Shareable Media system. This applet will contain utility and servlet files specific to the PlusShorts applet and will be formally registered with the new Borland application server.

5.0 Evaluation

5.1 Design

A daylong workshop with a small number of participants was organized to test and evaluate the PlusShorts application. The participants had varying degrees of experience with making and producing movies, using non-linear editing tools, networked environments and computer systems in general. The workshop activities were fourfold:

5.1.1 Introduction and Familiarization with the PlusShorts Application

Here the participants used the PlusShorts system to examine, evaluate and experiment with the mark-up assigned to describe the structure of familiar professionally edited audiovisual sequences (movies, advertisements and music videos). The movies demonstrated disparate editing styles and methodologies and derived from various cinematic eras such as Classical Hollywood, Early Russian Montage and French New-Wave. For the purpose of the workshop, movie segments from *Battleship Potemkin* (1925), *The Man with A Movie Camera* (1929), *Breathless* (1959), *The Hustler* (1961), and *The Untouchables*(1987) were used. The television advertisements selected for analysis included those forming part of a running series of similarly constructed and/or themed advertisements for a particular product *(Budweiser "Wazzup", Budget "Marketing Executives", MasterCard "Priceless").* Other selected advertisements included those mimicking or performing parodies of other advertisements, including advertisements from Ralph Nader's presidential campaign, which allude to both monster.com and MasterCard ads. A compilation of music videos comprised the final selection of audiovisual segments to examine and markup.

The purpose of this exercise was to familiarize participants with the operation and function of the PlusShorts system. They learned how to reorder sequences, cut/copy/ paste clips, start new video threads and search for clips in the database. Using familiar material that had been edited and constructed according to a wide variety of styles and purposes helped focus attention on how the punctuation symbols could be used to describe disparate audio-visual sequences. Participants could begin to evaluate the relationship between the structure of the video sequences and the punctuation symbols used to describe that structure.

5.1.2 Sequence Construction and Markup using Seed Content

Here the participants used the PlusShorts system to construct and mark-up sequences using video material specifically shot and produced in short, modular units that lend themselves to being used and appropriated in a multitude of ways. This activity involved the construction of sequence threads using footage in the database that had been deliberately produced for collaborative construction, display and playback within an open-ended non-linear system. The footage consisted of approximately 200 short movie clips that had been produced for the Interactive Cinema Group's recent *Flights of Fantasy* Installation at the DeCordova Museum. Participants worked on creating both original sequences from the material and modifying the sequences of others in a collaborative exercise. Participants could also choose to add new material or reuse material used in exercise one if they wished. Participants were encouraged to discuss their edit decisions, their choice of markup and their interpretation of the constructions of other participants

5.1.3 Production of Original Content and Sequence Construction using Original Content

Here, participants used the PlusShorts system as a guiding framework for the production of original material. This material was used to collaboratively edit and construct sequences that were then marked-up using the descriptive punctuation symbols. Participants, working on their own and together, produced some short digital video clips for use within the PlusShorts system. This process involved pre-production (planning, writing, location scouting), production (shooting) and post-production (capturing, rudimentary editing, keywording and compression). The clips were ultimately added to the system where they were used by the participants to create and construct movie sequences.

5.1.4 Discussion and Evaluation

Throughout the workshop, participants were asked to articulate and describe their progress through the various activities. They were asked periodically to examine, interpret and reflect upon the contributions of the other participants. Several structured group discussions were held during the day to discuss both the practical and the conceptual underpinnings of the PlusShorts system. Some of the issues discussed included:

- Did they find it easy/difficult to use?
- Did they find any elements of it confusing?
- What did they think of the use of punctuation as a symbol system for describing movie structure?
- Did they find PlusShorts a helpful tool for interpreting the conceptual ideas of others?
- What would they like to add to/change about the software?
- How did the final production activity differ/compare to their usual approach to the movie production process?
- Would they use this system in the future?

5.2 Results

The workshop took place at the MediaLabEurope in Dublin, Ireland. The local version of the PlusShorts system application was installed and configured to use the local network to share media files, keyword and sequence information. The PlusShorts application used in the workshop allowed participants on different workstations to see each other's work by periodically restarting the application to refresh the sequence/thread information. I had seeded the application with a considerable number of unmarked (no punctuation symbols) "starter" sequence threads using the movie, advertisement, music video and Flights of Fantasy content.

In addition, I added my own interpretation of the content using the punctuation symbol set to mark up the original sequences and posted these as secondary "reply" threads underneath each of the parent threads. Participants in the workshop could use these example sequences as initial indicators of what the punctuation symbols might mean and generate some immediate discussion as to how to interpret both the original sequences and the marked up secondary threads. Participants were asked periodically throughout the day to verbalize their actions and describe their thought processes in relation to the markup language. These conversations were recorded, as were the more structured group discussions also held at regular intervals throughout the day.

All three participants in the workshop had considerable prior experience with non-linear video editing systems (Premiere, Avid, Lightworks, After Effects, Media 100) and each was computer literate and Internet-savvy.

Paul Smyth holds a B.A. in Philosophy and Politics and a Masters in Film. Since 1994 he has worked part time for Fashion Television in New York, Milan, Paris and London as a Field Assistant.

He also worked for Entertainment Tonight on several occasions as Satellite Coordinator at events including the MTV Europe Music Awards (Dublin), the World Music Awards (Monaco) and the Cannes Film Festival.

Michelle Richards holds a BA in Fine Art and Sculpture and a specialist diploma in Art and Design teaching. She worked in arts administration before becoming involved in third level education. She also continues to work on her own art practice and has exhibited her work throughout Ireland as well as in New York, Yugoslavia, Poland and France.

Simone Poster graduated with a BA in Fine Art and Decoration and she also holds a MSc. in Multimedia Systems. She has worked as an arts journalist, while continuing to develop her own artwork. She is currently completing here PhD. studies, where she is designing responsive environments in public spaces to create interactive art installations.

5.3 Observations

5.3.1 Sequencing

Paul and Michelle began by concentrating solely on reediting and reorganizing individual sequence threads. They spent considerable time getting familiar with the application itself, examining the video content in the database and having fun creating strange and wonderful reconstructions of television advertisements. FIGURE 5-1 shows a screenshot of a sequence created by Michelle that demonstrates her inclination to construct sequences using shots that were visually similar. She had searched through all the threads in the database looking for clips of faces shot in closeup and in black and white, to put together.

FIGURE 5-1: Video thread created by Michelle using visually similar images

FIGURE 5-2 shows a screenshot of a sequence created by Paul, where he used footage from the famous Odessa steps scene in Eisensteins' *Battleship Potemkin*, depicting crowds of people running in panic from the murderous army. He incorporated this footage with some clips from the *Budweiser "Wazzup"* commercial, to create the impression that the crowds of people were fleeing not from the army, but from the boorish, screaming Budweiser guys.

FIGURE 5-2:Video thread created by Paul using clips from different threads

Gradually, the participants began to create sequences using content from multiple movie threads, and, once at ease with the workings of the application and its various editing features, they began to markup the video threads they were constructing with the symbolic language provided.

Simone became familiar with the workings of the application very quickly and began using the punctuation symbols shortly after the first activity period began. She had been previously introduced to the application prior to the workshop and so was somewhat more aware of its purpose and machinations.

The participants spent more time examining the content from the first exercise and seemed to incorporate that content more into the sequences they chose to construct. The television advertisements and the music videos with strong narrative elements seemed to appeal particularly to the workshop participants. The humor and visual strength of these clips lent themselves to numerous interpretations. Towards the end of the final session, the participants, and in particular Paul, began to integrate material from the Flights of Fantasy content with footage they themselves had shot.

5.3.2 Use of the Punctuation Symbols

All three participants used and interpreted the punctuation symbols in different ways. Simone wanted to use them to offer visual clues to the viewer as to the different playback pathways that the movie thread could take. For example, FIGURE 5-3 shows a screenshot of a sequence she constructed where she suggested using the question mark symbol before different segments in a thread to denote one of a number of possible playback paths

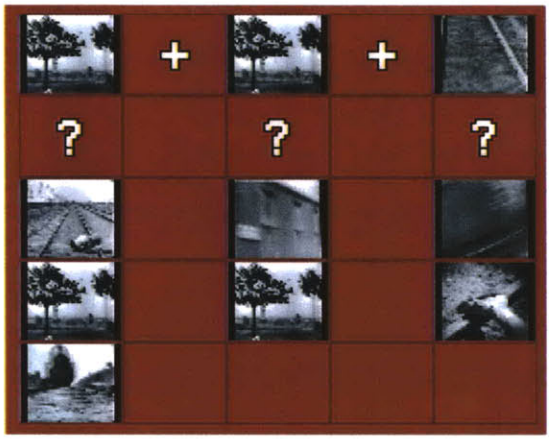


FIGURE 5-3: Video thread using annotation to indicate playback possibilities

Michelle was less interested in using the symbols themselves to denote the structure of the sequences she was putting together. She used them primarily during the first activity, but when it came to the footage that she had shot herself, she felt the nature of the content she had shot did not necessarily lend itself to being easily described using the symbols provided. Most of her footage denoted clips of smoke, paving stones, water and building interiors shot using abstract angles and unconventional framing. Paul utilized the symbols throughout all the workshop activities, where he used them to enhance the narrative of the sequences he was putting together. FIGURE 5-4 shows a screenshot of a sequence he created using footage from an *AIR* music video. He used the exclamation mark symbol to reinforce the dramatic nature of certain clips and used the ellipse icon to indicate **"that things are just continuing in the same vein."**

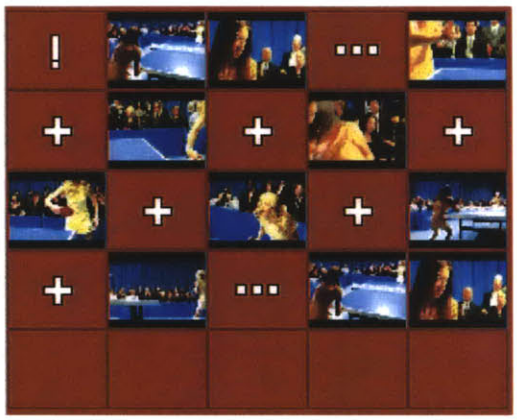


FIGURE 5-4: Video thread using reedited and annotated music videos

FIGURE 5-5 shows a screenshot of another sequence that Paul created using clips from *The Hustler, The Man with A Movie Camera* and some video clips from the Flights of Fantasy footage. Here he used the plus symbol to denote time was moving on and that there were active changes between the edits. He also used the parentheses to draw attention to certain segments in the final video thread where he was trying to build up tension.

FIGURE 5-5: Annotated video thread created by Paul

Each of the participants agreed on the necessity of group concurrence on the meaning of the individual symbols if working on a collaborative group project. Michelle also pointed out that a user's background experience or motivation for using the application might affect their approach to using the symbol system.

"It might have something to do with your background and where you come from. If you're more a scriptwriter you might be more inclined to use it."

5.3.3 Group Discussions

Prior to going out to shoot content for the third activity of the day, the group discussed how they were going to tackle the shooting project, given their understanding of the manner in which their content would be later edited, reconstructed and displayed within the PlusShorts application. Paul was the most experienced filmmaker of the trio, having worked on the production of a number of shorts in various capacities. He was most interested in gathering material that could be reused in a number of different contexts and that would be useful to all the workshop participants:

"I'm thinking that definite and defined movements would be good. Various kinds of cutaway shots or generic shots that could be used in a multiplicity of ways and work in a variety of contexts. My idea would be to make things that could be usable in the most possible contexts in the database. Clips that would be useful for more people than just yourself. "

Michelle, whose background is as a fine artist, declared a very different intent:

"The kind of stuff I would normally shoot would be quite abstract stuff like bubbles or patterns. An abstract reaction to an action. Surreal stuff. Anything that has a pattern. The kind of way that I would work would not be story based at all. So, all these clips in the database seem to be all about some type of event. Whereas I would look at how the clips look and how that clip would meet up with the next clip. Make a pattern out of it. I would use the system in a different way to Paul."

Paul and Michelle agreed to accompany each other on the shoot and to attempt to use each other's footage in their respective constructions later on. Towards the end of the workshop, Paul reflected further on the footage he had shot:

"I definitely wanted to keep things short and concise and try and have in and out points very quickly. In this case you had to think of taking short, snappy, quick, reworkable shots and maybe take things you could use in a variety of circumstances as you're aware that they will go into a database and you will draw on them regularly as stock shots almost. So you end up trying to shoot stock things, door banging or a person walking away or any sort of thing that's used regularly in movies. Then you might shoot more dramatic, specific shots to mix with the stock footage, so you have two types. In the time we had, I don't think I did very well though!"

5.3.4 Suggested refinements and improvements to the PlusShorts application

All three participants declared that the application was straightforward and easy to use. The one problem that they collectively agreed upon, was the fact that you couldn't temporarily store your work while you searched through the database for clips to add to your construction or navigated through other threads looking for content. All three participants recalled losing work at one or several times during the workshop, when they loaded up another thread inadvertently without having previously submitted their work. Simone suggested adding a text input component to the interface, whereby users could attach text comments to particular clips or threads, to further explain what they were doing. Paul disagreed, stating:

"I feel like if a group of people were collaborating on something and they understood what each punctuation symbol meant, there would be no need for textual comments. A user or a group of users could just define what they meant by the symbols and that would be less laborious than adding written text."

Simone drew attention to the problem with audio and its binding relationship with the video track. She wanted to be able to take the original audio off the clips and add her own soundtrack which she could insert into the video thread using a visual track guideline that would be displayed underneath the video clips:

"It would be good also to have a separate place for audio clips. When I was playing this particular thread, I would immediately look for a clip of dramatic music to fit the part where the guy gets run over by the train. And I would just add it underneath like a separate track."

FIGURE 5-6 shows a mock-up of the PlusShorts interface demonstrating Simone's idea of displaying the added audio track as a graphic line, indicated here in blue.

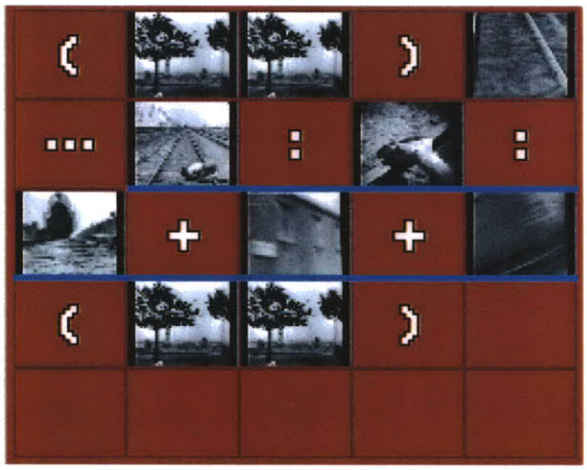


FIGURE 5-6: Prototype rendition for visually representing the audio track

Paul suggested developing the search engine to allow for a wider variety of search criteria. He thought it would be useful if you could visually search for clips by color, or tone or any other type of visual descriptor.

Michelle wanted to be able to exercise greater control over the visual playback of the video sequence. She wanted to be able to use the symbols to determine the visual layout of multiple video streams. For example, she imagined using the plus symbol to divide the screen into four, and then using another plus symbol to divide the screen into eight etc. Initially all participants advocated developing more video effects capabilities for the applications. Michelle thought this would help compensate for the jarring implications of constant jump cuts. She also thought some form of image processing would be useful to make the edits from disparate clips a little less obvious.

She suggested that adding a function averaging out the visual discrepancies between clips would make the final playback more pleasurable:

"When you're putting pieces together like in a collage, it would be really useful to make them merge in some way. Like here's some city stuff and I want to find something similar, but it's a very different color. If you could affect it in some way to blend it together better, that would be useful"

Paul thought the addition of a fade out/in capability and the capacity to add black frames would enhance the final playback and make the sequence look more professional. After considering the implications of adding all these additional features to the interface, Paul reconsidered his original position and stated instead that the beauty of the PlusShorts system was its simplicity and ease of use. He believed that adding numerous features and effects might only detract and needlessly complicate the application:

"It would definitely be interesting, apart from it being just an assembly of clips if you could affect them it would be great. It would give it an extra layer. Obviously the more functions it has the more uses you can think for it. I'm not sure though. It might complicate things needlessly. The fact is it is simple to use and does simple things and if you complicate it more it might make it too much like other stuff that's out there that's too complicated to use and less user-friendly."

All three participants also agreed that it would be beneficial if you could refine the in and out points on individual video clips. Michelle also pointed out a somewhat obvious interface oversight – there was no stop button to control the video playback. Although the video stream could be stopped by simply clicking on the video image itself, this was not initially readily obvious to the participants.

5.3.5 Future Use of the PlusShorts application

Paul and Michelle both agreed that they would use the system to collaborate with others remotely. Paul thought that the system would provoke a user to think more deeply about the editing process and would encourage users to experiment more with different sequencing arrangements while making a movie:

"I feel like I'd enjoy using it for just making you think about editing and make you think about sequences and playing with content. I feel like I could definitely use it if I was working with someone remotely. Say I was collaborating with someone and I could say go look at this thing I did and here's the name of it and they could go and look at it and decide if they agreed with it or wanted to change it and they could send it back and then I could look at it too. I'd especially use it as you could provide your own content and upload your stuff and put it together and people could see it and comment on how you put it together."

Paul also talked about using the application as a storyboarding tool:

"It could be very good for storyboarding. Even just using the stills you could probably sequence your whole project. The fact that there are three pages of storyspace means you are able to have different versions of it on each page and layout the structure of your story as a big storyboard. Without even playing it you could just use the stills to tell the story."

Paul drew attention to the fact that you could use the thumbnail images in the construction grid as a useful guide to maintaining certain editing rules of thumb, namely matching eyeline and movement continuity between clips. FIGURE 5-7 shows a screenshot of a sequence that Paul constructed using the thumbnails as visual clues for maintaining the 180 degree line or "axis of action", as described in Section 3.4.1.

FIGURE 5-7: Video thread created by Paul emphasizing the "axis of action"

"I like it because there are a few different eyelines in this piece – seven stills or whatever - and to match the eyelines is really easy because you can just drag them around and check quickly. Any time I've ever made a short you put so much time into checking the eyeline and cheating when you're making it, but this would totally help you piece it together instantly afterwards. You might be able to pay less attention when you were making it even."

Michelle also remarked on the advantage of being able to gauge a rudimentary idea of the visual composition of a sequence from the thumbnail images displayed in the construction area:

"I like the way you can see all the clips and see the color so without playing it back you can see the color shifts and get an idea of what it looks like."

FIGURE 5-8 shows a screengrab of a sequence that Michelle made using the color of the thumbnails as a visual reference for constructing a strongly patterned sequence.

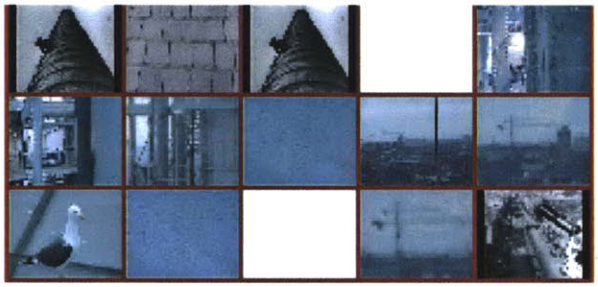


FIGURE 5-8: Video thread created by Michelle using visually similar clips

5.4 Comments

The background experience and interests of the participants very much dictated how they chose to interact with the application. Michelle, as an artist with a particular interest in visual patterning, concentrated on putting together sequences using clips that looked visually similar both within the construction grid and during playback. Paul, with a background in television and experience making fiction shorts was more interested in constructing actual stories and developing narratives from the content in the database. Prior to the workshop, Simone had already shot footage for an interactive narrative piece that had not yet been edited. With this in mind for use for the final exercise of the day, she explored the possibilities of multiple playbacks and forking narratives from the very first exercise. She was the first to begin using the punctuation symbols to both comment on the structure of the video threads and to give visual indications in the construction grid of possible playback paths.

The nature of the discussion both within the arranged group meetings and casually throughout the day was open and honest. Participants worked together in discovering how the application worked and in particular, how to find clips in the database and use the search function most effectively. Participants showed interest and enthusiasm in engaging and interpreting each other's work and unprompted discussions about the activities and the work being conducted took place between participants throughout the day. Michelle and Paul decided to shoot together in the same location but with the intention of capturing very different footage and ultimately try and integrate the two content sets later on. Time constraints meant that only three sequences were constructed from their merged content. In one such sequence, Paul used Michelle's footage that she had shot traveling through a building in a glass elevator, which stopped at each floor. He combined this with his footage of people doing various activities (eating, talking, drinking, walking) making it seem as if each activity was occurring on a separate floor in the building.

The type of footage shot by the workshop participants was readily influenced by the workings of the PlusShorts application itself (apart from Simone's footage, which was shot prior to the workshop but with the overarching idea of constructing an interactive narrative with the content). Both Paul and Michelle described their awareness when they were out with the cameras of shooting short, modular clips with definite in and out points. Michelle was interested in building up a corpus of visually arresting and relational clips, while Paul was specifically targeting stock action/reaction shots of short duration, which would be useful to have in the database in general.

Capturing, editing, compressing and keywording the video clips shot on the day of the workshop incorporated a large proportion of time in the afternoon, and all participants expressed regret about the briefness of the preparation and pre-production stage. They felt that they would have produced better and more appropriate content if they had had a longer time to prepare. However, all participants readily agreed that they would like to use the system again, given a time structure that was somewhat less confining. They also concurred that the application could be best utilized if users were collaborating together on a specific, predefined task and where the group shared a common understanding of the meaning of the punctuation symbols.

Of the refinements suggested for further iterations of the application, I think that dealing with the temporary storage of "working files" is of paramount importance. The frustration experienced with losing work, and the redundancy of having to repeatedly submit semi-completed threads could be alleviated by including a temporary buffer in the application. A more long-term problem pertains to the soundtrack issue highlighted by Simone. The ability to separate, mute or discard the audio channel associated with a clip and replace the audio on either an individual clip or throughout a sequence thread is important to the future development not only of the PlusShorts application, but to other applications currently being developed for the Shareable Media system.

Careful thought must be given as to how the audio track should be represented in the application interface, what functionality should be added to deal with the addition and subtraction of audio and what overhead this will bring to the data transfer rates between the client and the server in the online version of PlusShorts.

6.0 Conclusions

6.1 Summary

This research has introduced PlusShorts, a software application utilizing punctuation symbols as a visual markup language to describe and share conceptual ideas about movie structure. This application provides a vocabulary and syntax for analyzing, describing and interpreting the multifarious editing methodologies used in the construction of movie sequences. By using the punctuation symbols to describe and annotate the structure of movie sequences, a distributed group of collaborators can share and interpret the conceptual and aesthetic ideas governing the moviemaker's construction process.

6.2 Contribution

The PlusShorts application, to paraphrase Edith Sitwell's description of the modern novel [see Section 3.2.1], allows users to break down predestined groups of shots, examine their texture and then rebrighten them by building them into new and vital shapes. This application presents a visual language for describing the structural breakdown of video sequences within an explorative environment that serves, in accordance with Metz's definition, as an attention-focusing device. This research thus attempts not only to understand film, but, in supporting the collaborative construction and interpretation of movies by a distributed group of participants, seeks also to understand how film is understood.

The initial hypothesis that punctuation symbols could be used to share conceptual ideas about movie structure was answered by the evaluation. In Section 3.1 we ascertained that the power of symbolic tokens lies with the abstractions they represent, and their ability to embody all aspects of the ontological domain they are defining. The punctuation symbols selected to represent cinematic structure in the PlusShorts system were successful in defining the editing constructs of audiovisual content derived from a variety of genres. The symbols were used to describe the structure of movies exemplifying radically different editing styles: traditional Hollywood continuity, Russian montage, New Wave discontinuity and contemporary action. They also supported the structural description of additional forms of audiovisual content such as television advertisements and music videos.

Users of the PlusShorts application utilized the symbols in different and inventive ways [see Section 5.3.2], adapting their meaning to correspond with their own understanding of cinematic content. From this we can determine that the background and creative aspirations of the moviemaker has a profound effect not only on how they use the tool but also, on how they choose to interpret the filmmaking process and assign meaning. However, despite their different approaches to describing video, participants in the evaluation workshop were able, occasionally after some discussion, to understand and interpret the annotation of others. The workshop participants strongly agreed that consensus on the meaning of the punctuation icons, even when used in different ways, can be reached by a group of collaborators working on a common project.

Displaying the movie sequences as elements of an evolving movie thread in a newsgroup style interface proved successful in allowing users to easily navigate through the database of sequences. This presentation of audiovisual content also provokes a new understanding of what constitutes a 'cinematic utterance'. A movie thread can begin as a traditional narrative, dissolve into an abstract visual interpretation and resolve as gothic horror.

Movie utterances are no longer the product of one directional voice aimed towards a receptive audience, but rather the amalgam of collective interpretations from a community of viewer-contributors.

The display of the sequence thumbnails and punctuation annotations in the construction grid allowed users to quickly establish the visual look and feel of the movie sequence, along with facilitating the practice of continuity editing, regarding eyeline and movement matching [5.3.5]. This method of displaying the sequences also allows for the paradigmatic comparison and contrast of shots in a sequence as users can upload multiple versions of a sequence with different shots selected to play out at a particular point.

6.3 Improvements and Future Work

The addition of a component to deal with the separation, addition and general manipulation of the audiotrack would strengthen the ability of the PlusShorts application to both provide a greater lever of editorial flexibility to the filmmaker, and afford the viewer a more fluid and enjoyable cinematic experience.

The development of the search engine to allow searching of the sequence annotation was suggested for the PlusShorts application but unfortunately, never implemented. The ability to use the punctuation symbols themselves as search criteria would provide a novel means of searching the database at a purely structural level. For example, the submission of the query **+++++=!!!!** might return an action sequence or **“”...””** might return a sequence containing two dialogue scenes that take place in different locations.

The imminent release of SMIL 2.0(scheduled May 2001), with its full complement of transition effects, provides an opportunity to integrate an effects component into the PlusShorts application, as suggested by some participants at the workshop. These effects could be integrated with the punctuation icons and an option to use the application in either descriptive (symbols used only for annotation: no effect on playback) or prescriptive (symbols used as transition effects: effect playback) mode.

The online version of the PlusShorts software is currently being reintegrated with the Shareable Media System. This will allow for further user testing regarding the annotation and interpretation of the movie sequences using the punctuation symbols. This testing will involve a wider demographic of distributed contributors, interested in using the application to teach video editing, to prototype and develop ideas for a traditional film production, to collaborate with other moviemakers in exploring new means of cinematic expression, and most of all, to have fun.

One audiovisual format that lends itself particularly to further exploration in the online, collaborative domain is the music video. The sublimation of narrative and character development, the use of multiple types of footage and the densely textured structure combine to provide an audiovisual experience that is visually compelling, fleeting, yet memorable. The development of a compelling interface that facilitates a comprehensive engagement between the sound and image tracks, where the structure, rhythm and pace of each affects and informs the final output, is of importance here. One can imagine the emergence of an endless, open-ended music video, composed of video and audio files contributed by a distributed group of users. The music video playback could be controlled and directed in real time by these users, where the video images selected could influence the choice of music, its tempo, volume and pitch, and it, in turn, could affect the pacing, transitions and visual effects of the video image.

The ability to share and interpret conceptual ideas about the process of creating is an important one in collaborative projects. This is particularly pertinent in online environments where not all members of the community know each other and where there are multiple cultures and languages. The development of more sophisticated representative systems for the sharing of abstract and subjective understandings of digital information is an important goal in supporting not only human interactions on the web, but also the endeavor to provide machine readable semantic annotations of documents and digital artifacts created by humans. The web is slowly evolving and extending to accommodate semantically described data, in conjunction with the development of comprehensive ontologies relating conceptual information with logical rules. The research described in this thesis represents a preliminary initiative to provide users with a simple, familiar annotation scheme to describe and share with others, how they understand their creative output. Future research promises the development of broad ontologies and rich descriptive frameworks for audiovisual content, in a world where online environments will support not only the creative contributions of humans, but also the artistic output of computational agents.

7.0 References

Aigrain, Philippe; Joly, Phillipe and Longueville, Veronique. (1995) Medium knowledge-based macro-segmentation of video into sequences. *In IJCAI Workshop on Intelligent Multimedia Information Retrieval, p 5 – 14*

Arman, Fashid; Hau, Adringand Chiu and Ming-Yee. (1993) Feature management for large video databases. *In Proc. SPIE Storage and Retrieval for Image and Video Databases*

Babel, Isaac and Morison, W. (1955) Collected stories. New York: Criterion Books

Bair, Deirdre. (1978) Samuel Beckett: a biography London: Vintage Books

Barthes, Roland. (1967) Science versus Literature *Times Literary Supplement, 28 September*

Beckett, Samuel. (1929) "Dante...Bruno. Vico..Joyce" in Our Exagmination Round His Jactification For Incamination of Work in Progress London: Faber and Faber

Beckett, Samuel. (1970) Watt. New York: Grove Press.

Bertino, Elisa; Catania, Barbara and Zarri, Gian. (1999) A Conceptual Annotation Approach to Indexing in a Web-Based Information System. *Workshop on Advanced Issues of E-Commerce and Web/based Information*

Black, Max. (1949) Language and Philosophy Westport CT: Greenwood Press

Bordwell, David and Thomson, Kristin. (1997) Film Art: An Introduction New York: McGraw-Hill Fifth Edition.

Bresson, Robert. (1986) Notes on the Cinematographer London: Quartet Books

Britanica. (1990) The New Encyclopædia Britannica. Chicago: Encyclopædia Britannica.

Bruckman, Amy. (1991) The Electronic Scrapbook: Towards an Intelligent Home-Video Editing System. *MIT MS Thesis.*

Buckland, Warren. (2000) The Cognitive Semiotics of Film Cambridge: Cambridge University Press

Chion, Michel. (1994)Audio-Vision: Sound on Screen New York: Columbia University Press

Connor, Stephen. (1998) Slow Going *Presented at the Critical Beckett Conference, University of Birmingham, 26th September.* [http://www.bbk.ac.uk/eh/eng/skc/slow.htm]

Cummings, E. E. and Firmage, G. J. (1923) Complete Poems: 1904 – 1962 New York: Liveright

Dancyger, Ken. (1997) The Technique of Film and Video Editing: Theory and Practice Boston: Focal Press

Davenport, Donna and Forbes, Cheryl. (1997) Writing Movement / Dancing Words: A Collaborative Pedagogy *EDUCATION, Winter*

Davis, Marc. (1995) Media Streams: Representing Video for Retrieval and Repurposing. *Ph.D. Thesis, MIT*

Deely, John; Williams, Brook and Kruse, Felicia. (eds) (1986) Frontiers in Semiotics Bloomington: Indiana University Press

Eikhenbaum, Boris. (ed) (1982) The Poetics of Cinema, in Russian Poetics in Translation 9, trans. Richard Taylor. Oxford: RPT Publications

Evans, Ryan. (1994) LogBoy Meets FilterGirl: A Toolkit for Multivariant Movies. *MIT MS Thesis*

Golembiewski, Emily. (2000) Lost and Found – The philosophical appropriations of Jay Rosenblatt in *Filmmaker Summer Vol. 8 #4*

Kelliher, Aisling; Seo, James Jung-Hoon; Pan, Pengkai; Lin, Cathy and Davenport, Glorianna. (2000) Visual Interfaces for Shareable Media in *Proc. Int. Symposium on Electronic Art, Paris*

Laban, Rudolf. (1928). Script dancing. La danse écrite. Wien: Universal Edition.

Long, Jeffrey. (1996) A Study of Notation: The Structure of Perceptual Revolutions [http://www.cs.vu.nl/~tbr/repository/nel/ason.html]

Manuzio, Aldo. (1970) Phrases linguae latinae, 1579. Menston: Scolar Press.

Merriam-Webster. (1998) Webster's tenth new collegiate dictionary. A Merriam-Webster. Springfield, Mass. G. & C. Merriam Co.

Metz, Christian. (1974) Film Language: A Semiotics of the Cinema New York: Oxford University Press

Metz, Christian. (1974) Language and Cinema trans. Donna-Jean Umiker-Debeok The Hague: Mouton

Moxon, J. (1962) Mechanick exercises on the whole art of printing (1683-4). London: Oxford University Press.

Oldham, Gabriella. (1995) First Cut: Conversations with Film Editors Berkeley: University of California Press

Pan, Penkai. (1999) I-Views, a Storymaking Community of, by and for the Audience*MIT MS Thesis*

Pierce, Charles Sanders. (1931) Collected Papers, vols I – VIII, ed. Charles Hartshorne and Paul Weiss. Cambridge: Harvard University Press

Propp, Vladimir Iakovlevich. (1968) Morphology of the Folk Tale Austin: University of Texas Press

Reiss, Steve and Feineman, Neil. (2000) Thirty Frames Per Second : The Visionary Art of the Music Video New York: Abrams

Reisz, Karel and Millar, Gavin. (1968) The Technique of Film Editing Boston: Focal Press

Rui, Yong; Huang Thomas S. and Mehrotra Sharad. (1998) Exploring video structures beyond the shots. *In Proc. of IEEE conf. Multimedia Computing and Systems*

Rui, Yong; Huang Thomas S. and Mehrotra Sharad. (1998) Constructing Table of Contents for Videos. *In ACM J. of Multimedia Systems*

Saussure, Ferdinand de. (1966) Course in General Linguistics, trans. Wade Baskin, New York: McGraw Hill

Seidelmann, P. K. United States Naval Observatory. Nautical Almanac Office., et al. (1992). Explanatory supplement to the Astronomical almanac. Mill Valley Calif.:University Science Books.

Sitwell, Edith. (1925) Poetry and Criticism New York:Vanguard Press

Smith, Thomas Aguierre. (1992) If You Could See What I Mean... Descriptions of Video in an Anthropologist's Video Notebook. *MIT MS Thesis*

Stam, R., R. Burgoyne, et al. (1992). New vocabularies in film semiotics : structuralism, post-structuralism, and beyond. New York: Routledge.

Stockhausen, Karlheinz. (1993) Electronic Study No. 2. Score

Tonfoni, Graziella. (1996) Communication Patterns and Textual Forms. Exeter: Intellect

Tonfoni, Graziella. (1999) On augmenting documentation reliability through communicative context transport. *Proc. of the 1999 Symposium on Document Image Understanding Technology", p 283 – 286, Annapolis, MD*

Universal Studio Archives
http://www.universalstudios.com/archives/hitchcock/05_text.html

Weis, Elisabeth and Belton, John. (1985) Film Sound: Theory and Practice New York: Columbia University Press.

Whitelaw, Billie. (1995) Billie Whitelaw…:Who He? London: Hodder & Stoughton

Wollen, Peter. (1972) Signs and Meaning in the Cinema Bloomington: Indiana University Press

Yeung, Minerva; Yro, Boon-Lock and Liu, Bede. (1996) Extracting story units from long programs for video browsing and navigation. *In Proc. IEEE Conf. On Multimedia Computing and Systems*

Zarri, G. and Gilardoni, L. (1996) Structuring and Retrieval of the Complex Predicate Arguments Proper to the NKRL Conceptual Language. *In LNCS 1079; Proc of the Nineth Int. Symp. On Methodilogies for Intelligent Systems, p 398 – 417, Zakipane, Poland*

Zhong, Di; Zhang HonJiang and Chang Shih-F. (1997). Clustering methods for video browsing and annotation. *Technical report, Columbia University*

Zhang, H; Gong, Yihong; Smoliar, S.W. and Tan, Shuang Yeo. (1995) Automatic parsing of news video. *In Proc. IEEE Conf. On Multimedia Computing and Networking*

Zhuang, Yueting; Rui, Yong; Huang, Thomas S. and Mehrotra, Sharad. (1998) Adaptive key frame extraction using unsupervised clustering. *In Proc. IEEE Int. Conf. On Image Processing*

8.0 Filmography

400 Blows, The	(1959)	Francois Truffaut, France
A Hard Day's Night	(1964)	Richard Lester, Great Britain
All The President's Men	(1976)	Alan Pakula, United States
Battleship Potemkin	(1925)	Sergei Eisenstein, USSR
Best Years of Our Lives, The	(1946)	William Wyler, United States
Birth of A Nation, The	(1915)	D.W. Griffith, United States
Blowout	(1981)	Brian De Palma, US
Bonnie and Clyde	(1967)	Arthur Penn, United States
Breathless	(1959)	Jean-Luc Godard, France
Broken Blossoms	(1919)	D.W. Griffith, United States
Citizen Kane	(1941)	Orson Welles, United States
Easy Rider	(1969)	Dennis Hopper, United States
Graduate, The	(1967)	Mike Nichols, United States
Help!	(1965)	Richard Lester, Great Britain
Hustler, The	(1961)	Robert Rossen, United States
Informer	(1935)	John Ford, United States
Intolerance	(1916)	D.W. Griffith, United States
L'Age d'Or	(1930)	Luis Buñuel, France
La Jetee	(1962)	Chris Marker, France
Maltese Falcon, The	(1941)	John Huston
Man of Aran	(1934)	Robert Flaherty, GB
Man with A Movie Camera, The	(1929)	Dziga Vertov, USSR
Manchurian Candidate, The	(1962)	John Frankenheimer, US
Night Mail	(1936)	Basil Wright, Great Britain
Oktober	(1928)	Sergei Eisenstein, USSR
Plow that Broke the Plains, The	(1936)	Pare Lorentz, United States
Psycho	(1960)	Alfred Hitchcock, US
Raging Bull	(1980)	Martin Scorsese, US
Seven Samurai, The	(1954)	Akira Kurosawa, Japan
Star Wars	(1977)	George Lucas, United States
Strike	(1924)	Sergei Eisenstein, USSR
They Shoot Horses, Don't They?	(1969)	Sidney Pollack, United States
Thin Blue Line, The	(1988)	Errol Morris, United States
ToyStory 2	(1999)	John Lasseter, US
Two or Three Things I Know about Her	(1967)	Jean-Luc Godard, France
Un Chien d'Andalou	(1929)	Luis Buñuel, France
Untouchables, The	(1987)	Brian De Palma, US
Woodstock	(1970)	Michael Wadleigh, US