Building Visual Knowledge in a 'Cinematic Hypertext' authoring environment

Clara Mancini

Knowledge Media Institute
The Open University
Milton Keynes
MK7 6AA
UK
E-mail: C.Mancini@open.ac.uk

INTRODUCTION

One of the main issues in hypertext is what we would like to call "discourse shaping". The possibility of providing or recognising discourse coherence in the interactive medium has a lot to do with the possibility that authoring environments offer to shape the hypertext discourse. The answer of spatial hypertext to this problem is 'visual' shape [11,12]. Its peculiarity and strength is that it allows the user to express relations between objects in a visual dimension, where shape is more immediately perceivable and can express in an implicit form what might be difficult to explicitly define. From painting, to photography, to cinema, all visual media have this peculiarity, due to the nature of their significant. Among them, the cinematic medium is the closest to hypertext and the analogy between cinema and hypertext in general[1] has been pointed out in more than one occasion [3,7,8]. In particular, though, cinematic language has a lot to do with spatial hypertext and would have contributions to make to its development, as a medium that creates discourse coherence through dynamic visual shape.

CINEMATIC LANGUAGE AND HYPERTEXT

Cinematic signification is based on the juxtaposition of shots, by which the film's discourse is generated. The cognitive connection of shots is conventionally based on a set of rhetorical patterns (the result of half a century of cinematic linguistic evolution), which provide coherence to the linear chain of shots, assisting viewers in recognizing the articulation of a discourse. In fact, the essence of the viewer's "reading" ability is in establishing connections between shots, as the meaning of any single shot dramatically changes depending on how it is connected to the others [6].

The parallel between the cinematic and hypertext medium is precisely based on the fact that the latter also consists of the articulation of rich semantic units whose connection, due to the activation of a link, effects and expresses a strong semantic relationship. Although hypertext units constitute self standing cores of content, the meaning of a single unit changes depending on how its connections with the others are activated. That is, the role of any hypertext unit depends on the navigational path it happens to be part of according to the user's choices.

---

[1] By *hypertext* we refer to the medium, the concrete form which can only work in an interactive medium. It can follow different registers or working modalities, that we distinguish according to two different parameters: the visibility on the contents' organisation, on the one hand, and the definition of connections between hypertext units, on the other hand. On the one hand, the reader can have access to the contents through a global, *aerial level* view (node maps) or through local, *ground level* discovery ('link by link paths'). On the other hand, the connections between content units can be *enunciated* (explicitly indicated and/or classified as in most semantic hypertexts), or their identification and classification can be *evoked* by visual/kinetic features (like in spatial hypertext). For instance, we see *ScholOnto* [1] as aerial with enunciated connections; *VKB* [12] as aerial with evoked connections; the *Assembly 3/1* hypertext [2] as both ground level and aerial, with indicated but unclassified connections; and *HyperCafé* [10] as ground level with enunciated but unclassified connections.

CINEMATIC LANGUAGE AND SPATIAL HYPERTEXT

Its iconic nature makes cinematic language very close to the concept of spatial hypertext itself and to the ideas it is based on and which distinguish it from other forms of hypertext entirely based on written text.

Although cinematic language includes different codes like oral speech, written text, music and sound, it is essentially iconic, being based on moving images. Hypertext, like cinema, is a visual medium, as navigation develops on a computer screen and its discourse can make use of images, as well as of sounds. Nevertheless, in the majority of cases, written text prevails, if it is not the only code to be used, as it lends itself to the explicit expression of abstract concepts.

The minimal linguistic unit of natural language is the phoneme, a *symbolic* non–signifying differential element, whose combination generates morphemes successively articulated to generate the enunciation. The cinematic minimal linguistic unit is the *shot,* an *iconic* and *indexical* semantically rich element, which, in semiotic terms, is the equivalent of a linguistic enunciation. Because in natural language the nature of the signifier has nothing to do with the nature of the signified, the correspondence between the signifier and the signified is more convention-based. Therefore, concepts can be expressed more explicitly, the user being required to build a mental representation of concrete elements. In contrast, because in cinematic language the nature of the signifier does have something to do with the nature of the signified, the correspondence between signifier and signified is less convention-based, so that cinematic language can suggest concepts more implicitly, through the re-presentation of concrete elements and events.

Because in natural language, connections can be explicitly enunciated, their correctness is constantly verified against the content of the units that are connected and a connection can in fact be right or wrong (always implying a certain degree of commitment). In contrast, in cinema, because connections are performed without being enunciated, any connection tends to be seen as coherent and sensible in one way or another, as viewers seek constantly to make sense of transitions. As Miles underlines [7], in cinema a connection is not in principle right or wrong, but good or bad, effectual or ineffectual, and a cinematic sequence has to be deciphered as a structural whole, the single element (unit or transition) having no specific meaning in itself. This is why cinematic language, unlike natural language, cannot be considered as a grammar, but as a rhetoric [6]. And this is where cinematic language has so much to do with the concept of spatial hypertext, where connections that would be difficult or committing to express in natural language can be implicitly suggested by visual and spatial features rather than being explicitly defined.

Although a rhetoric is much more flexible than a grammar, one thing that it certainly needs in order to be expressive, that is, in order to be able to generate ‘shape’, is consistency. In other words, the way visual elements are used in cinema to express concepts or to represent events has to be consistent, especially locally, that is, in the single film. What usually happens is either that a particular director (read: author) develops a personal style that constitutes his personal cinematic language through all his works, or that a particular film displays a peculiar style or set of representational conventions that overlap with the most basic ones generally followed by everyone. Anyway, no matter how loose a style is, the essential is that it is consistent. Without consistency not only there would be no style, but there would not be possible to generate shape at all.

Rosenberg extensively discuss how spatial hypertext would benefit from an articulated use of visual features, exploring a number of possibilities [9], and VKB seems to constitute a development of VIKI in that direction [12]. In this respect, a contribution could also come from the way cinematic language shapes discourse, according to the essential criterium of representational consistency.

COGNITIVE COHERENCE RELATIONS AND SPATIAL HYPERTEXT CONNECTIONS

To provide consistency to the use of visual features, there would be the need for criteria to define hypertext connections. These criteria would not need to be specific and in fact, the more general they were, the more generally applicable and the more flexible they would be.

As we describe else where [4], computational and psycholinguistic research on text coherence relations, reports evidence that the articulation of complex text structures is based on a small set of elementary cognitive relations, from whose combination other coherence relations can be derived. These coherence relations describe the cognitive connection that the reader assumes to hold between the content units of two spans of text and they are hypothesised to be "universal" in the sense that they are *cognitive,* not specific to a particular domain or genre of text. In brief, the relational categories identified belong to basic relation types like *causal* (A causes B), *disjunctive* (A is alternative to B), *conjunctive* (A coexists with B), *conditional* (A presupposes B) and *sequential* (A follows B). These relation types imply relational rules, which may either succeed or fail. If the rule succeeds, the relation is considered positive (e.g. A *does* follow B), if the rule fails the relation is considered negative (A *does not* follow B). Apart from *logical* connections, the Cognitive Coherence Relations theory also offers a base to account for *analogical* connections, namely with the *comparison* relation (or its opposite: *contrast*), implying a underlying *similarity* relational rule.

Because CCR categories are cognitive and very basic, they may constitute quite intuitive distinctions in the way objects, like hypertext chunks, can be related. Without being committing or too constraining, in the phase of information *triage*, CCR could support spatial hypertext users in consistently using visual and spatial features to shape their collection of information chunks. While, in the hypertext presentation phase, CCR could support users in the discovery and interpretation of any shape before their eyes.

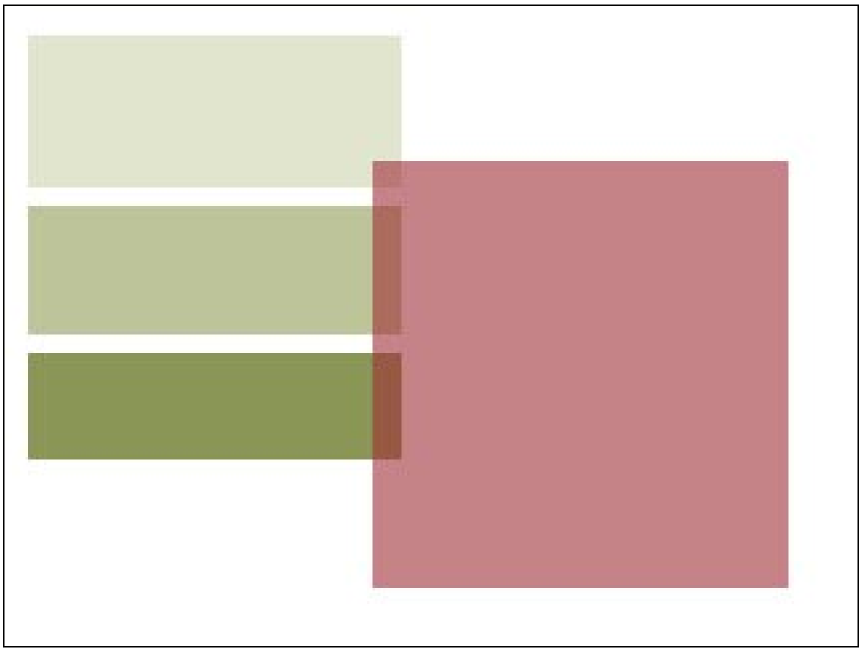
COGNITIVE COHERENCE RELATIONS, CINEMATIC LANGUAGE AND SPATIAL HYPERTEXT

In synthesis, Cognitive Coherence Relations could be used to provide consistency and to support the coherent development of hypertext relational structures. Even in a first attempt to give shape to a chaotic collection of information chunks, it is likely that people would be able or willing to distinguish between *causal, conjunctive, disjunctive, conditional, sequential* or *similarity* relations. For instance, there are many types of causal relations and they might be difficult to identify or committing to define at an early stage of information *triage*. However, assuming or "having the feeling" that a causal relation holds between the content of two information chunks is much easier and much less committing. An incremental formalisation of the hypertext connections, that is, the articulation and distinction of a general causal relation into more specific ones might or might not take place in further phases of elaboration.

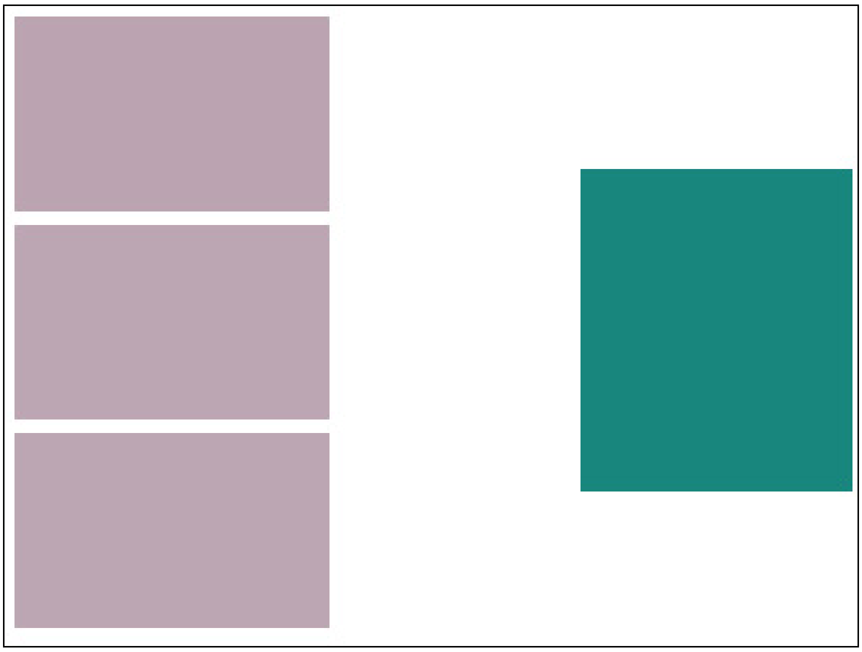
On the other hand, visual features (like shape, colour, font, spatial position/composition and juxtaposition of hypertext units) and maybe dynamic features (like transitions' duration and synchronisation of hypertext units), which characterise cinematic language, could be used both to express hypertext structures in the information *triage* phase, and to *present* already existing hypertext structures. Authors could be allowed to define their "visual style" for the connection types that they are going to use, thus establishing a local, consistent, visual language, just as film directors do. The author might decide whether the hypertext connections should be explicitly indicated and classified, or be left implicit and only used by the system to render the hypertext's connections and structures, which the end user would infer on the basis of visual features (in this respect such environment would be different from systems like Aquanet [5]).

More in concrete, a series of hypertext units connected by sequential relations might for instance be represented as a series of rectangles positioned one below the other and having the same width and different gradients of the same colour, according to their position into the sequence. To express conjunctive relations, for instance, the same features might be given to contiguous hypertext units, while different shape, colours, fonts, and opposite positions on the screen might express disjunctive relations. Causal or conditional relations might, for instance, be represented by inheritance/sum of features from some of the hypertext units to the others. Finally, similarity relations might, for instance, be rendered either through comparable or contrasting hypertext units' colours, shape, fonts, and so on.

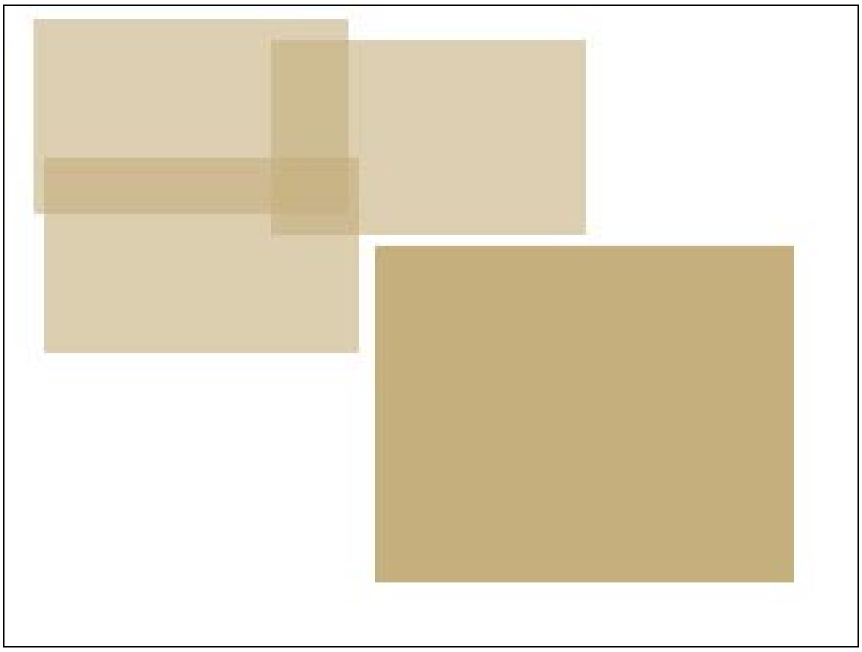
The expressive possibilities can be countless as countless can be the combinations of visual features. The figures below illustrate few combined examples.



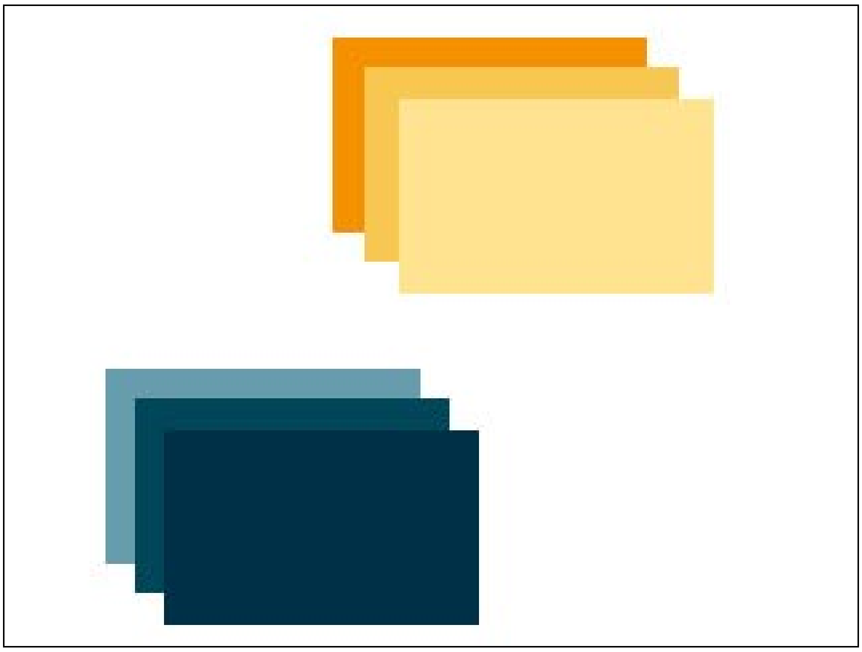
1.The figure above might represent a group of hypertext units connected by a sequential relation (green) and another hypertext unit related to the others by contrast or other analogical relation.



2. The figure above might represent a group of hypertext units connected by a conjunctive relation (purple) and another hypertext unit related to the others by contrast or disjunctive relation.



3. The figure above might represent a group of hypertext units connected by a conjunctive relation (transparent brown) and the descending hypertext unit related to the others by causal relation.



4. The figure above might represent two groups of hypertext units connected by conditional relations (overlapping series) and related to each other by contrast or comparison relation.

CONCLUSIONS

The process of knowing is a process of shaping, but there is no shape without consistency and no consistency. If cinematic language may suggest how to create discourse coherence through the consistent use visual features, Cognitive Coherence Relations might provide basic, intuitive relation types in order to consistently use visual features to support coherence in spatial hypertext, reinforcing its peculiarity, which is expressing structure through visual shape.

ACKNOWLEDGMENT

I would like to thank Simon Buckingham Shum for the insightful discussions and his helpful comments.

REFERENCES

1. Buckingham Shum, S., Motta, E. and Domingue, J. ScholOnto: An Ontology-Based Digital Library Server for Research Documents and Discourse. *Int. Jnl. on Digital Libraries*, 3, 3, 2000, pp. 237-248 [www.kmi.open.ac.uk/projects/scholonto].
2. Kolb, D.: *Assembly 3/1* version of the *Sprawling Places* hypertext project, (in press)
3. Mancini, C. From Cinematographic to Hypertext Narrative. In *Proc. Hypertext 2000*, San Antonio, 2000, ACM Press: New York, pp. 236-237
4. Mancini, C., Simon Buckingham Shum, S. Cognitive Coherence Relations and Hypertext: From Cinematic Patterns to Scholarly Discourse. In *Proc. Hypertext 2001*, Aarhus, Denmark, Aug 14-18, 2001, ACM Press: New York
5. Marshall, C.C., Halasz, F.C., Rogers, R.A., Janssen, W.C., Aquanet: A Hypertext Tool to Hold Your Knowledge in Place. In Proc. Hypertext 1991,
6. Metz, C. *Essai sur la Signification au Cinéma*. Editions Klincksieck: Paris, 1968
7. Miles, A. Cinematic Paradigms of Hipertext. *Continuum: Journal of Media & Cultural Studies*, 13.2, 1999, pp. 217-226
8. Miles, A. Hypertext Structure as the Event of Connection. In *Proc. Hypertext 2001*, Aarhus, Denmark, Aug 14-18, 2001, ACM Press: New York
9. Rosenberg, J. And And: Conjunctive Hypertext and the Structure Acteme Juncture. In *Proc. Hypertext 2001*, Aarhus, Denmark, Aug 14-18, 2001, ACM Press: New York
10. Sawhney, N., Balcom, D. and Smith, I. HyperCafé: Narrative and Aesthetic Properties of Hypervideo. In *Proc. Hypertext'96*, Bethesda, MD, 1996, ACM Press: New York, pp. 1-10
11. Shipman, F.M., Hsieh, H., Maloor, P. and Moore, J.M. The Visual Knowledge Builder: A Second Generation Spatial Hypertext. In *Proc. Hypertext 2001*, Aarhus, Denmark, Aug 14-18, 2001, ACM Press: New York
12. Shipman, F.M. and Marshall, C.C. Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems. *Computer Supported Cooperative Work*, 8, 4, 1999, pp. 333-352