

A Two-time Model for Video Content Representation and Retrieval

Nastaran Fatemi

1) CLIPS, BP 53, 38041 Grenoble, France
Nastaran.Fatemi@imag.fr

2) EPFL-DI-LITH, Ecublens, 1015 Lausanne, Switzerland
fatemi@di.epfl.ch

Philippe Mulhem

1) CLIPS, BP 53, 38041 Grenoble, France
Philippe.Mulhem@imag.fr

2) IPAL, School of Computing, National University of Singapore
Lower Kent Ridge Road, Singapore 119260
mulhem@krdl.org.sg

Abstract

This paper presents a temporal model of video for video retrieval systems. Time is the dimension that plays an essential role in the perception of video document content. The modelling of the different temporal characteristics of the video content is however a complicated problem. Our approach to this problem has consisted of distinguishing two different dimensions of time for video documents: video time and story time. The semantic structure proposed based on these two temporal dimensions, provides the user with a new modality of content description which is close to the semantic perception of the video documents. We present here how using the conceptual graphs formalism, the details of the temporal model can be represented in query, document and matching function of the retrieval process.

1 Introduction

Video is a temporal media. Its temporality is essentially due to a certain speed that the video imposes to its users in the perception of its content information. Despite the importance of the temporal aspect of the video, it has not been sufficiently studied in the context of video Information Retrieval systems. The aim of this research work is to determine the temporal information of video content, to define its use in a video Information Retrieval system and finally to propose a temporal model of video which allows to explore these information in the retrieval process. We have especially emphasised the generality aspect of the model proposed: a model to be simply adaptable to different types of video and application contexts.

To construct the model, our approach has consisted of two preliminary phases: first, we studied the characteristics of time-based perception of video to discover the temporal information of the video content and their relation to other content information; second, we studied the main current video models to verify their adequacy of serving as a general base model which may be extended to represent temporal information. This later study revealed us the lack of such a “general model” for video retrieval purpose. We propose then in this paper a new semantic structure for video documents, which is based on our study of the characteristics of time-based perception of video.

In the following, we first describe in the section 2 the existing video document models and their limits. The section 3 focuses on the features of the time-based perception of video documents. These features lead to the principles of our model described in section 4. A formalisation of the model based on the conceptual graphs is presented in section 5. The section 6 explains shortly how we carried out the model on a specific context. The section 7 proposes some key points for the automatic indexing of the video using the actual feature extraction techniques. We conclude in section 8.

2 Related Works

The main idea of most of the existing video models may be classified into three categories: the hierarchical cinematographic model [1], [2] and [3], the stratification model [4] and [5], and models based on the objects present in the video [6], [7].

The hierarchical cinematographic model is based on the cinematographic structure of video documents,

which is the result of the editing of a video. In this structure, the whole document is composed of a set of scenes, each one containing a set of shots which are themselves composed of a succession of images. In such models, the whole video content is represented by describing the content of each part in the hierarchy. The existing techniques of shot cut detection and scene detection facilitate the reconstruction of this structure and so the indexing of video documents following this model. However, the utilisation of this structure is more adequate to the structured video documents (in which this syntactical cinematographic structure is highly related to the semantic structure of the video document) for example the television news. In the other cases, i.e. when such a semantic structure is not dominant, it would be restrictive to represent the content information through the cinematographical structure.

The stratification model is particularly designed to overcome this limitation. In this model, unlike the hierarchical cinematographic model, it is the description of the whole document which is partitioned into smaller pieces named "strata" (not related to scenes and shots of the video document) and then the first and last images of the video portion corresponding to each strata are determined. This model permits a large flexibility in the description of content; but as it has been mainly designed for the video editing and annotating systems, it does not provide a precise definition of what the content of each strata may be, how this content description is organised and how is it possible to retrieve a video document by describing its content through a set of strata.

The third category of models is based on the objects present in the video. These models represent for important objects of the video, the video portions in which they have been present and also a description of the object. Other models, e.g. [7], using spatio-temporal characteristics of the objects, allow to query on the spatio-temporal relations between objects. But, these models are eventually restricted to the representation of objects and they don't permit more elaborate description of the semantic content of the video documents which is obviously desired in a video retrieval system.

Despite the inadequacy of the existing video models to serve as a general video model appropriate to a video retrieval system, the study of characteristics of the time-based perception of video revealed us a semantic structure for video perception which permits a general framework for video content description in time. The result of this study has been explained in the next section.

3 The Time-based Perception of Video

A video document is a sequence of images played at an accurate frequency to create the illusion of animation. All information we receive from this media, either temporal or non-temporal, is the result of this succession of images in time. But, as we will see later, this time is not the only time dimension we can perceive when watching a video document. To show the different aspects of the time-based perception of video and so to define the temporal information of their content which may be used in the retrieval process, we explain, in the following, the different levels of the perception of the video content and then the different time dimensions related to these levels.

3.1 Two Levels of Perception of Video Content

We define two levels of perception of video content: visual perception and semantic perception.

3.1.1 Visual Perception

Visual perception corresponds to the elementary information we receive when watching a video independently of any abstraction we may attach to these information. These information are: the presence of pictorial elements¹ and the change of spatio-temporal characteristics of these elements. A simple example for this level of perception is given in the figure 1. At the visual perception level, the information considered to be transferred by the images shown in this figure is the presence of the circle in the image suit and its moving up in the first three images and its moving down in the last two images.

¹ A pictorial element is defined as a set of pixels that verify common criteria in a sequence of frames. Such criterion may be for example the form or texture created by the set of pixels, their color, etc.

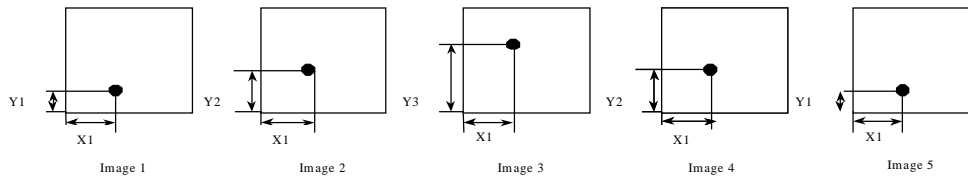


Figure. 1: An Example of Visual Perception of Video Content

3.1.2 Semantic Perception

Once pictorial elements are observed, abstractions are made through these elements towards different concepts². Such abstractions depend mainly on the context of the presentation of video. For example the circle of Figure 1 can be perceived in a particular context as the sun and its going up and coming down as the sunrise and the sunset. This abstraction process may continue in a hierarchical way: new concepts can be created as abstraction of one or several other concepts. Then, in the last example, the sunset followed by the sunrise may be considered as a “day” and so on.

The abstraction of the content information explained above does not end with the construction of concepts (and relations). The understanding of the video content leads finally to the perception of a *story*. In fact the story is first created by the maker of the video document by the use of the techniques and the art of cinematography (the cutting of video into shots and scenes is part of this artwork). Then, the understanding of the cinematographical language by the spectator reshapes the story. The notion of story has a primordial importance in the semantic content information perceived in a video as it represents the conceptual description of the video content for the spectator.

Each story contains a set of *events*, which are the dominant facts happenings during the story. An event is composed of a set of concepts and relations abstracted in the semantic level. According to the context of video, an event may represent the presence of certain objects or persons, an action realised by/on some objects or persons, etc. The events have temporal continuity in the story and so a description of the whole story may be formed by describing the events in time. In the next section, we precise the temporal descriptions of videos by presenting two time dimensions of the video documents.

3.2 Two Dimensions of Time for Video

The notion of story creates a new time dimension perceived by the user namely the *story time*. The story time is the time dimension during which the story of video “takes place”. This time dimension is different from the real time dimension of the video called the *video time* during which the story is “shown”. Events are then present in the story and also in the video. But the temporal characteristics of the events, i.e. the time intervals and the temporal relations of events according to the story and video, are different. For example, an event shown in video during a few minutes may create the illusion of its happening during one day in the story. Another example is two concurrent events of story which happen in different places. To show this, the events are cut into a few smaller parts and shown alternatively. The temporal relation between the intervals of the events in video and story time are different (As shown with two events E1 and E2 in figure 2). The flashback is another example where an event A happens before an event B in the story time but after in the video time.

² For the sake of simplicity, we use here the term of concept for the result of an abstraction process. We will see later that in fact there exist concepts and also relations between concepts.

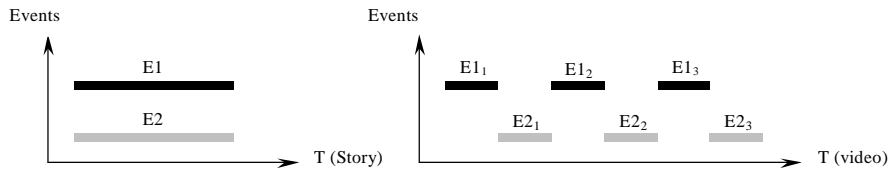


Figure 2: Two Concurrent Events in Story. The Demonstration of Their Temporal Relations in Story and Video Time.

3.3 Generic and Specific Features in Time-based Perception of Video

The overview of the description of time-based perception of video presented above, lets us distinguish the characteristics which are *generic* for different video types and applications from those which are considered as *specific* for a given context. This early distinction clarifies the key points in defining the principles of the temporal model.

Generic characteristics consist of visual perception, the process of abstraction of concepts and relations (from visual to semantic perception), the formation of the events, and the perception of the whole content in the form of a set of events which occur in time (video and story).

In the other hand, specific characteristics consist of the type of concepts and relations abstracted and also the internal structure of events (possible bindings between concepts and relations). The choice of the time dimension (video, story or both) to be represented may also depend on the type of the video and the application and may thus be considered as a specific characteristic.

4 The Principles of the Temporal Model

Following the results of our study of time-based perception explained earlier, we now describe the principles of the temporal video model. Before going into such description, we present an intuitive vision of the kinds of queries that we would expect the model to handle.

Examples of temporal queries

Following are the examples of the queries which are expected to be handled by the temporal model.

- Q1: The videos in which we see Alfred Hitchcock for at least 1 minute.
- Q2: The videos in which we see a car falling to a valley after a car pursuit.
- Q3: The videos in which an explosion happens at the same time that two persons are talking together.
- Q4: The videos in which we see a scene of revolution in 1980.
- Q5: The videos in which we see soldiers the day after the war.

A simple analyse of the queries presented above results that the temporal model should permit:

- The description of the events of a video (e.g. a car falling down to a valley), and
- For each event, the description of its video time interval (e.g. Q1) and its story time interval (e.g. Q4) and also the video temporal relations (e.g. Q2) and story temporal relations (e.g. Q3 and Q5) it has with other events.

The next two sections describe the way these two principles are taken into account in the temporal model.

4.1 Event Content Description

Reminding the definition of events in section 3.1.2, events are the description base units composed of a set of concepts and relations. As events have temporal continuity, there may be no temporal relations inside an event.

Such definition is general and postulates the definition of the set of possible concepts and relations and also their various bindings. The specification of these is context dependent is considered as a specific characteristic of the model. For example, in a TV news archive the set of concepts may be defined as the important *persons*, *objects*, *places*, *actions*, etc. The relations may be the relation *is_agent_of* which signifies that a person (or an object) is the agent of an action, the relation *has_location* which relates a person, an object or an action to the place they are present, etc. Then an event describing the fact that “Bill Clinton shakes the hand of Jacques Chirac at the White House” would be represented by the concepts *Person*: Bill Clinton, *Action*: Shaking hands, *Person*: Jacques Chirac, and *Place*: White House, and the relations *is_agent_of* (relating the two persons with the action) and *has_location* (relating the action to the place). Whatever are the elements that compose events in a particular context, a fundamental point to underline is that the temporal features of the events are domain independent. These features are detailed in the next section.

4.2 Temporal Description of Events

As mentioned before, in our model the temporal description of the video content is allowed by the temporal description of the events in video and/or story time. In the following, we focus on the principal temporal features to be represented for events. These features exist at the video time level as well as the story time level in the model.

- The time interval of an event e , called *Intvl*, which is a couple consisting of the beginning and the end points of the time interval: $\text{Intvl}(e) = (\text{Beg}(e), \text{End}(e))$
- The time duration of an event e , called *Duration* (e).
- The temporal relations between the time intervals of the events. These relations come from the 13 well-known Allen Relations [8] which permit the specification of all possible temporal relations between two time intervals. To permit also the representation of the quantification of these relations, we use Moulin Relations [9]. Moulin proposes the representation of all Allen Relations by only two relations *Before* and *During* and the parameter *Lap* bound to *Before* and *DB* and *DE* bound to *During*. The attribution of the negative, zero, or positive values to these parameters permit at the same time, the distinction of Allen relations and also the quantification of these relations. Moulin relations are as following:

$$\begin{aligned} \text{Before}(\text{Intvl}(e1), \text{Intvl}(e2), \text{Lap}) & \quad \text{where } \text{Lap} = \text{End}(e2) - \text{Beg}(e1) \\ \text{During}(\text{Intvl}(e1), \text{Intvl}(e2), \text{DB}, \text{DE}) & \quad \text{where } \text{DB} = \text{Beg}(e1) - \text{Beg}(e2) \\ & \quad \text{and } \text{DE} = \text{End}(e2) - \text{End}(e1) \end{aligned}$$

Here, we may notice that having the time interval, other temporal information like the duration and temporal relations may be calculated; however it should be noticed that if the time dimension to be represented is the story time, we cannot be always sure to have a precise information about time interval of a given event (whereas the temporal relations are easier to obtain), so we cannot expect the deduction of other information like the duration of the event and the temporal relations between the events. In the temporal model we present, we will represent all these information. However, certain implementations of the model may let part of these information to be calculated from others and so decide to store only the necessary part.

5 Formalisation

To permit the precise specification of the model, we present here a formalised instance of it using the *conceptual graph* knowledge representation formalism defined by Sowa [10]. The most important advantages of this formalism may be named as following:

- There exist algebraic operators that are in accordance with the logical interpretation of conceptual graphs. This leads to a strong theoretical validation of the formalism synonym to the strong validation theory [11].
- The algebraic interpretation gives a basis to achieve query processing with polynomial complexity when

adequate pre-processing is performed [11].

- The graphical representation of the conceptual graphs allows an easy exchange between the user of the model and its provider.

5.1 The Conceptual Graphs

Before presenting the formalisation of our temporal model we give here a brief introductory to the conceptual graphs and their relationship to Information Retrieval.

5.1.1 An Introduction to Conceptual Graphs Formalism

Conceptual graphs are a knowledge representation formalism based on the linguistics, psychology and philosophy [10]. A conceptual graph represents information as a finite, connected, oriented, bipartite graph having two types of nodes: *concepts* and *relations*.

Concepts are discrete (atomic) entities which may correspond to the human mental images (concrete concepts) or not (discrete concepts). A concept, represented graphically by a box, has a *type* (which corresponds to a semantic class) and possibly a referent (which corresponds to an instantiation to an individual of the class). In a concept, the referent and the type are related by a *conformity relation* which permits to verify whether the association of a label of type *t* to a referent of type *r* is meaningful: the concept is so well-formed. There exist two categories of referents: individual referents each of which designates a particular individual and generic referent noted by * which designates any individual referent that conforms to the type of the concept.

Conceptual relations are represented graphically by an oval. These specify the relation which exists between the concepts of the graph. The relations are identified by a type and they give a direction to the conceptual graph containing them.

Figure 3 shows two types of representations for conceptual graphs: graphical and linear representations.

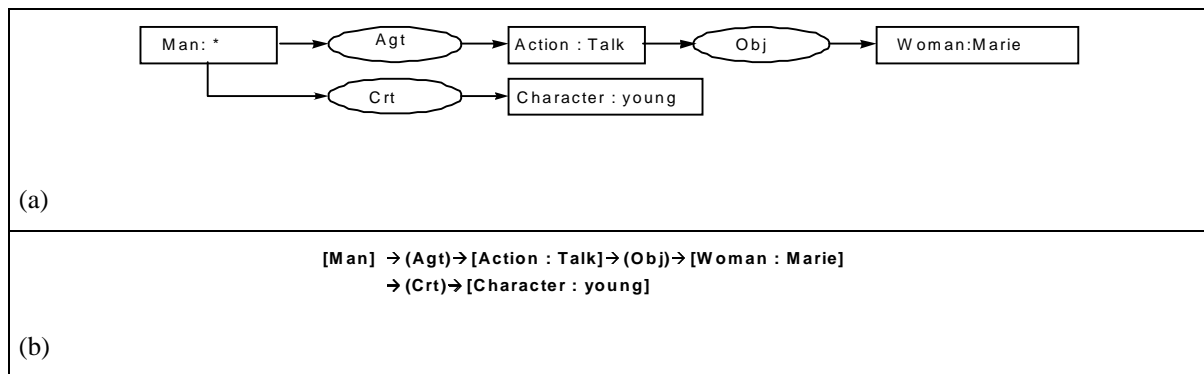


Figure 3: Graphical (a) and Linear (b) Representation of the Conceptual Graph for “A Young Man Talks to Mary”.

The formalism defines a knowledge base that contains a concept and a relation lattice. The lattice T_c is a set of *concept types*. T_c is provided with a partial ordering relation \leq_c . The lattice T_r is a set of *relation types*. T_r is provided with a partial ordering relation \leq_δ (The partial ordering relations \leq_c and \leq_δ represent the notion of Generalisation/Specialisation). The set of concepts and relations in T_c and T_r are relatively restricted between $Top_c, Bottom_c$ and $Top_r, Bottom_r$. The figure 4 is a simple example of the concept lattice (The relation lattice has the same presentation).

In the graph conceptual formalism, *canonical graphs* represent the possible situations of the real world. These graphs express the valid combinations of the concepts and relations. There exists a set, namely the *base*, of canonical graphs that are defined *a priori* and which express the elementary semantic constraints of the represented domain. Other canonical graphs are *derived* from the canonical base by the *canonical operations* of

copy, join, restriction, and simplification proposed in the conceptual graph formalism.

If a graph g_2 is derived from a graph g_1 , g_2 is a specialisation of g_1 and noted as $g_2 \leq g_1$.

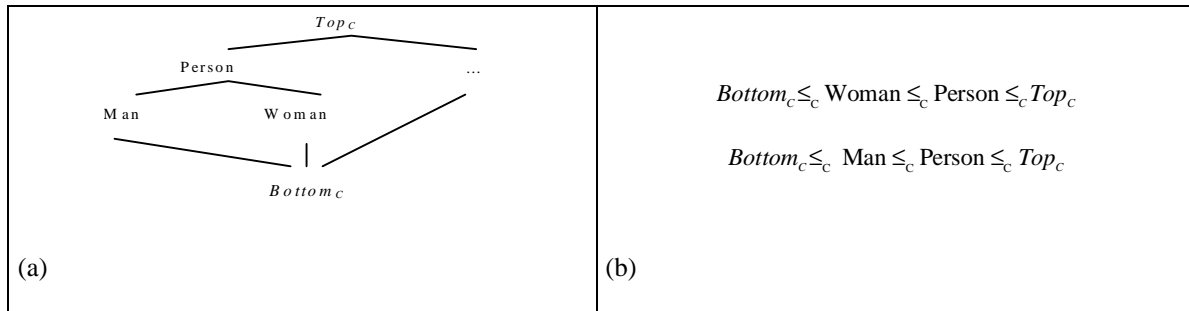


Figure 4: Graphical (a) and linear representation (b) of the concept lattice.

5.1.2 Conceptual Graphs and Information Retrieval

The important advantage of the conceptual graphs, besides their flexibility and expressive power is specially their connection with the logical model of Information Retrieval [12] as discussed in [13] and [14]. This connection is due to the explicit relation of the formalism with the first order logic. Sowa [10] defined an operator Φ which permits to associate to each graph u a formula $\Phi(u)$ expressed in first order logic. This operator has the property of conserving an order on the graphs: if $u \leq v$ (u is a specialisation of v) then the associated logical formulas verify the following implication: $\Phi(u) \supset \Phi(v)$.

The relationship between this formalism and a retrieval process is that: a document indexed by a conceptual graph D is relevant to a query represented by the graph Q if D is a specialisation of Q , i.e., $\Phi(D) \supset \Phi(Q)$.

The example of the figure 5 illustrates the projection to be done when the user wants to know if there exists somebody talking to somebody else.

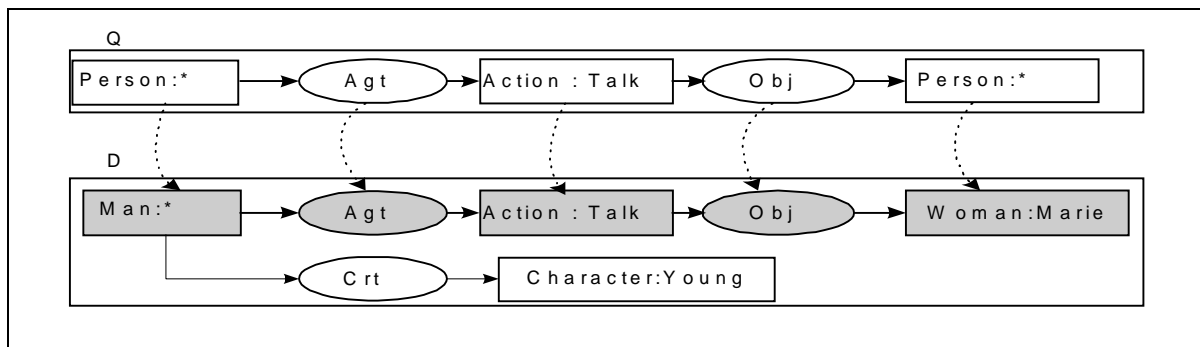


Figure 5: The Projection of the Document D on the Query Q .

The realisation of such implication is achieved on conceptual graphs using the *projection operation*: if $u \leq v$ then there exists a projection of v on u . So using the projection operation of conceptual graphs we can affirm that $D \leq Q$ and hence $\Phi(D) \supset \Phi(Q)$.

The conceptual graph formalism has been used by a number of systems which have retrieval capabilities. We may cite the system KALIPSOS of IBM [14] which is a text understanding system for French language or the system MENELAS [15] which is medical Information Retrieval system. However these applications of the formalism are restricted to a use of conceptual graphs as knowledge representation formalism and not in the

correspondence procedure. The ELEN system [13] proposed for the first time the adaptation of this model for the Information Retrieval systems. Conceptual graphs have gained attention in Information Retrieval especially in the domains of new applications of IR such as digital libraries and hypermedia systems. The use of the conceptual graphs in the multimedia Information Retrieval systems is particularly interesting because of their rich power of expression which is also a matter highly desired in these types of applications. For example EMIR² [16] is an image retrieval system based on the conceptual graph formalism. Despite the benefit of using the conceptual graphs from the expressiveness point of view, the evaluations of the systems like EMIR² prove the complication and timeliness of the projection operation used in the matching phase: the relative algorithm being of exponential complexity. This problem was resolved by the proposition of the indexing model based on inverted files [11] resulting a correspondence algorithm of polynomial complexity. The image retrieval system RELIEF [11] which is an application of this model, shows the improvement in response time of queries and at the same time the improvement of the precision and recall measures.

5.2 Temporal Model Formalised with the Conceptual Graphs

In this section we present the formalised temporal model of video. To describe the document model by conceptual graph formalism, we shall present the canonical base C , the concept lattice T_c and the relation lattice T_r .

Each video document is composed of a set of *events*. The content of each event is described by a graph. The initial graph of C is thus the following:

[Video] → (Is_Composed_Of) → [Event] → (Has_Content) → [Content]

The content of the graph *Content* is, as we explained before, dependent on the characteristics of each domain. To distinguish clearly between the generic and specific aspects of the model, we present the description of the *content* graph by providing separate specific canonical base: C_{sp} , specific concept lattice: T_{CSP} , and specific relation lattice: T_{RSP} .

A very simple example of a canonical base used to represent persons, objects, actions and the places is:

[Entity] → (Is_Agent_Of) → [Action]
 [Action] → (Has_Object) → [Entity]
 [Action] → (Has_Location) → [Place]
 [Entity] → (Has_Character) → [Character]

The related T_{CSP} and T_{RSP} are presented in figure 6.

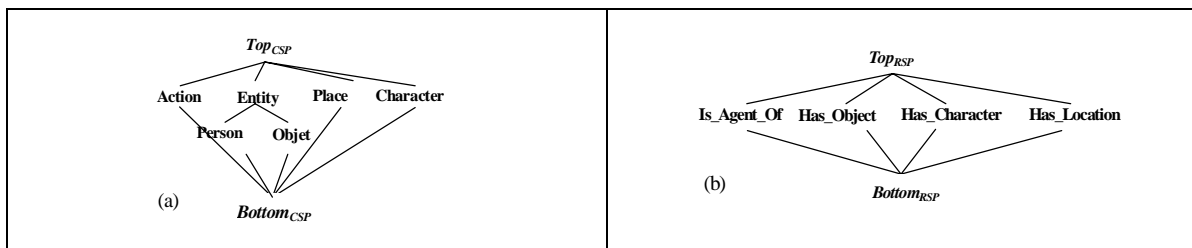


Figure 6: An Example of (a) Specific Concept Lattice and (b) Specific Relation Lattice.

Besides the content graph related to each event, there exist important temporal information representing the time interval of the event and its temporal relations with other events. These information are part of the generic characteristics of the model. According to the explanation given in section 4.2, these temporal information are represented by the following graphs in the generic canonical base, C :

[Event] → (Has_Beg) → [Time]	[Event] → (Before) → [Event]
[Event] → (Has_End) → [Time]	→ [Lap]
[Event] → (Has_Duration) → [Duration]	[Event] → (During) → [Event]
	→ [DB]
	→ [DE]

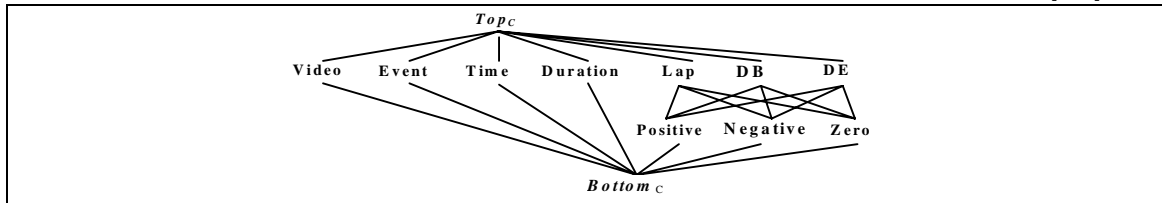


Figure 7: Generic Concept Lattice

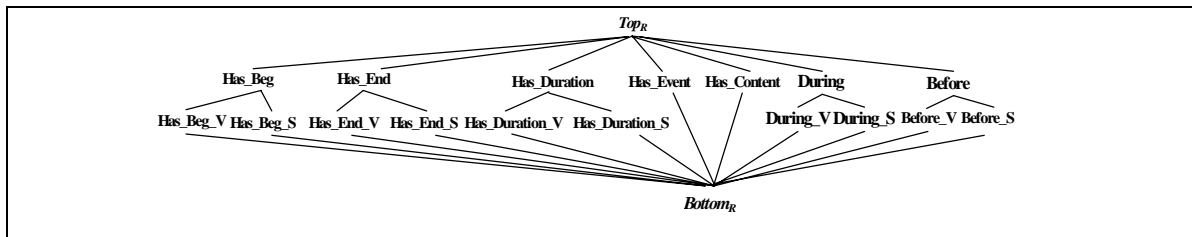


Figure 8: Generic relation lattice

The concept and relation lattices are presented in figures 7 and 8. In figure 7, the elements Lap, DB and DE are related to the temporal relations of Moulin seen in section 4.2. In figure 8, the specification of the relations used in the canonical graphs to their corresponding video and story relations, designated by “_V” and “_S”, permits the representation of video and story time.

In this model, the query is also a conceptual graph constructed based on exactly the same canonical base, relation and concept lattices provided for the document model. The matching between the query and document is allowed by the projection operation of conceptual graphs explained in the last section. An example of the representation of a temporal query based on the proposed model is given in the following.

Q: “A video portion in which two persons are talking at the same time that an explosion happens”:

```
[Event:e1]→(Has_Content)→[[Action:Exploding]→(Has_Object)→[Object:*]]
[Event:e2]→(Has_Content)→[[Person:*]→(Is_Agent_Of)→[Action:Talking]→
(Has_Object)→[Person:*]]
[Event:e1]→(During)→[Event:e2]
→[Positive:*]
→[Positive:*]
```

The use of During in the above query permits the expression of the concurrency of the two events without precisising whether the video or story time is intended. In the case where it is necessary to distinguish these times the corresponding relations, During_V or During_S, are used.

6 Experimentation

A prototype based on our temporal model has been realised and tested on a video corpus of 10 minutes. The corpus is an excerpt of a television series. This video was annotated following the proposed model: a set of 30 events were distinguished and then described following a predefined set of canonical graphs. The events were then described by their temporal time features and relations in video and story time. The realisation of the conceptual graph query and document representation and also the matching function was done using Prolog.

The present prototype permits the retrieval of video portions matching to the temporal description present in the query.

At present, the extraction of the events and also their temporal features is done manually. To make the model of practical use for large collections of video data, we need to provide procedures which will help to automate as much as possible the extraction phase. The next section describes shortly our intention of using the existing feature extraction techniques in this direction.

7 Towards an Automatic Indexing

To construct the document model, we need the extraction of events and their video and story temporal features. As defined previously, video events are described by concepts and relations. The indexing task aims at extracting from the video documents the concepts (like persons, objects and actions) and then establishing the relations between them. Face detection and speaker recognition techniques can help to extract persons. Features like color, shape and texture can be used to detect objects. The identification of actions should use image and sound feature extraction techniques, especially when the action produces sounds, as talking, laughing, clacking, exploding, etc. To relate the concepts inside the events many different trends should be used. A first idea is to consider the relations between the concepts which are close together in space and time. The concurrency of the image and sound features may also help to establish the relations as for instance when indexing a dialogue event. Moreover, domain dependent knowledge bases may be used to provide i) a restricted set of possible relations between the concepts and ii) the patterns of image and sound features associated with the relations. The video temporal features of events (begin and end time of events, their duration and the temporal relations between them) can be retrieved automatically once the events are determined. The extraction of story temporal features is a more elaborate task which should consider the inference of story temporal features from video temporal features. This is the subject of a study we are actually working on.

The usability of the feature extraction techniques mentioned above is dependent on their accuracy. Besides, we do not consider the fully automatic indexing of video as a realistic issue according to the important semantic throughput of the video content information.

8 Conclusion and Future Works

In this paper, we presented a temporal model of video to be used in a video retrieval system. Basing our approach on the study of the characteristics of time-based perception of the video, we defined event as the base unit of video content description in time. An important advantage of the definition of event is its flexibility to be adapted to different contexts and applications and so permitting the model to be general. We also defined the notion of story time which plays an important role in the perception of the video semantic content. The temporal description of the content based on the story time permits a second modality of content description besides the description based on the video time. Finally the formalisation of the temporal model using the conceptual graphs presents a solid logical basis to describe the details of the model concerning the query, document and matching function. The evaluation of the proposed model will consist of exploring the influence of the two-time modelling on the quality of the system results: the idea is to find out how the extracted temporal relations can enhance the precision of the system. The evaluation will be conducted in the following way: videos will be shown to the users, and they will be asked afterwards to use the system to find some parts of the videos. Users will be separated into two groups: one group will use only the video event descriptions and the second group will use also the temporal features and relations of events. Then we will compare the results of the system according to the users' queries. We will also be able to compare the enhancement of the results when considering different index and different temporal features and relations of the video events.

Our emphasis in the present model has been mainly on the exploitation of time during the document and query description. We consider further extensions of the model in various axes. At query level, studies should be done to determine the different modalities of temporal description in natural language and their correspondence to the temporal description using the well-known Allen Relations. This study will permit the definition of a simple and natural query interface and also the principles of processing of such queries. To extend the matching

function, we consider the study of the possibilities of involving temporal characteristics during the matching process. This may provide us useful measures permitting the determination of different levels of relevance replacing the exact matching. We are also working on the extension of the present model to provide different levels of content description, for example to permit another level concerning the description of physical attributes of video documents. Further extensions are also considered to integrate the features needed in editing and composing applications of video in order to provide a unique model for editing and retrieval tasks of the video documents.

9 References

1. Zhang H. J, Gong Y, Smoliar S.W, Tan S.Y. Automatic parsing and indexing of news video, In Proceedings of IEEE conference on multimedia computing systems, Boston, Massachusetts, 1994, pp 45-54
2. Swanberg D, Shu C. F, Jain R. Architecture of a multimedia information system for content-based retrieval. In Proceedings of 2nd international workshop on network and operating system support for digital video, Heidelberg, 1993 pp 387-392
3. Yeo. B. L Yeung M. M. The front pages of video database: Retrieving and visualising video content, IBM research division, RC 20901; 1997
4. Davenport G, Smith T. A, Pincever N. Cinematic primitives for Multimedia, In IEEE computer graphics and applications on multimedia, 1991, pp 67-74
5. Weiss R, Duda A, Gifford D. K. Composition and search with a video algebra, In IEEE multimedia, 1995, pp 12-24
6. Nagasak A, Tanaka Y. Automatic video indexing and full-video search for object appearances, In Proceedings IFIP TC2/WG 2nd Workshop conference on database systems, Budapest, pp 120-133; 1991
7. Aritsugi M, Tagashira T, Amagasa T, Kanamori Y. An approach to spatio-temporal queries: Interval-based contents representation of images, In Proceedings of 8th international conference on database and expert systems applications (DEXA'97), Toulouse, France, 1997, pp 202-213
8. Allen J. F. Maintaining knowledge about temporal intervals, Communications of ACM, vol. 26, num.11, 1983, pp 832-843.
9. Moulin B. Conceptual graph approach for the representation of temporal information in discourse, Knowledge based systems, vol. 5, num. 3, 1992, pp 183-192
10. Sowa J. F. Conceptual Structures: Information Processing in Mind and Machine, Addison-Weseley Publishing Company, 1984
11. Ounis I. Un modèle d'indexation relationnel pour les graphes conceptuels fondé sur une interprétation logique, PhD thesis, Grenoble: University of Joseph Fourier; 1998
12. Rijsbergen C. J. v. A new theoretical framework for Information Retrieval. Proceedings of ACM conference on research and development in Information Retrieval, Pisa, 1986, pp 194-200
13. Chevalet. J. P. Un modèle logique de Recherche d'Informations appliqué au formalisme des Graphes Conceptuels: Le prototype ELEN et son expérimentation sur un corpus de composant logiciels. PhD thesis, Grenoble: University of Joseph Fourier; 1992
14. Berard-Dugourd. A, Fargues J, Landau, M. C. Natural language analysis using conceptual graphs, In Proceedings of the International Computer Science Conference, ICSC'88, 1988
15. Zweigenbaum P et al. Linguistic and medical knowledge bases: An access system for medical records using natural language, Technical report, MENELAS: deliverable 9, AIM Project A2023, 1993
16. Mechkour M. EMIR², Un modèle étendu de représentation et de correspondance d'images pour la recherche d'information. Application à un corpus d'images historiques, PhD thesis, Grenoble: University of Joseph Fourier; 1995