Indexing and Searching Video Sequences Using Concept Lattices

Nada Mimouni[1] and Yahya Slimani[2]

[1,2] Department of Computer Science, University of Sciences, El ManarII, BP. 1002, TUNIS

Abstract. The explosion of multimedia content in numerical databases have created a world need to new paradigms and techniques allowing to express how to explore, seek and summarize media collections and exploiting their content. In this paper we present a FCA-based approach to classify and search relevant video sequences from a set of video sequences. It consists in building a lattice from the binary relation between video sequences and their annotations. These annotations are extracted from a set of XML files associated to sequences. An indexation technique based on MPEG7 norm and XML language is used to annotate video sequences. Then we transform an user request into concept and merge it in the lattice. As result we obtain a set of sequences belonging to the extents of the request formel concept subsumers in the resulting lattice.

1 Introduction

The exploration of complex data, mainly video sequences, implies many problems, in particular with regard to their structuring, their storage and their indexation on the one hand and their analysis on the other hand. The representation of complex data for their analysis (indexation, search, etc.) requires a modeling making it possible to take into account the various aspects of these data. The use of an unified format to describe the video format is made possible thanks to XML.

In this paper we propose a solution to a particular IR problem where video sequences instead of data are searched and indexated with metadata reflecting information about video sequences rather than on data extracted from documents. The rest of the paper is organized as follows. Section 2 presents related work on Data mining and Multimedia mining. Section 3 provides an overview about certain fields as well as works that combine FCA and IR. The proposed approach using XML and FCA is described in Section 4. Section 5 concludes with a discussion.

2 Related works

2.1 Multimedia data mining

Generally speaking, knowledge discovery or data mining is a nontrivial extraction of previously unknown and potentially useful information from data [1].

Multimedia data mining can be defined as a technique to find unknown knowledge and unique patterns in multimedia data by reducing the gap between the high-level semantics of media objects and the low-level features extracted from them [2].

Within the framework of these complex data, knowledge discovery is even more delicate. A complex object can be regarded as an aggregate of complex documents [3]. XML files can be well thought-out as means making it possible to manage this level of information, considered in this case as metadata.

2.2 Indexing multimedia data

The indexing operation describes and characterizes a document or a fragment of the document by representing concepts contained in this document. Three principal aspects of video indexing are described in what follows [4]. The first aspect is the segmentation; it allows the identification of the units of indexing in video contents. The second aspect is the representation and classification; once units of indexing determined, it is necessary to represent and to classify the concepts of interest by textual annotations or descriptors resulting from the analysis of the video sequence. The third aspect relates to the methods of research which must provide relevant answers in lowest time. Annotations are texts explaining video sequence contents (whole document, scene, plan, image or even video object). There are two distinct aspects which can be expressed by annotations: the description or the interpretation of video sequences.

3 Concept lattices related theory

3.1 Concept lattices for Information Retrieval (IR)

As it was mentioned by Messai in [5], the application of concept lattices in Information Retrieval was originally present at the beginning of FCA. Information Retrieval was then mentioned as one application field for concept lattices usage [6]. A lattice-based information retrieval approach is proposed in [7]. In both propositions [6] and [7], lattice-based information retrieval shows performances that are better than boolean information retrieval.

Several works deal with the problem of generating the set of concepts and the concept lattice of a given formal context [5]. Some of the proposed algorithms allow an incremental construction of concept lattice for a given formal context such as proposed in [8]. This aspect is particularly beneficial for the information retrieval applications in general and for our video analysis problem in particular for two reasons. First, user queries need to be merged into the set of concepts. Second, incremental lattice construction allows the insertion of new concepts. A detailed work based on this aspect was achieved by Messai [5] on bioinformatic data sources within a BioRegistry repository on the web, and the results show a big performance compared with works on this field.

4 Annotation system and concept lattices for classifying and searching video

4.1 Construction of video concept lattice

In this section, we show how the FCA framework can be applied to the formalisation indexed video sequences. More detailed FCA related definitions can be found in [5].

In the following, the formalisation of video sequences database is given by a formal context $\mathcal{K}_{vid} = (V, T, I)$, where V is a set of video sequences (e.g. President Clinton, Ice Age, ...), T is a set of annotations (e.g. sun, car, ...) and I is a binary relation between V and T called the incidence of $\mathcal{K}_{vid}$ and verifying: $I \subseteq V \times T$ and $(v, t) \in I$ or (vIt), where v, t are such that $v \in V$ and $t \in T$, means that the video sequence v is indexed by the textual annotation t. An example of formal context is given in table 1.1 with video sequences and annotations full names in table 1.2 (symbols and abbreviations are used for a better visibility in the lattice). Consider $A \in G$ a set of video sequences, then the set of annotations common to all sequences in A is $A' = \{t \in T \mid \forall\, v \in A,\ vIt\}$.

Table 1. 1.Example of video sequences formal context ($\mathcal{K} = (V, T, I)$). 2.Complete names of video sequences and their annotations.

Sequences\Key words	SN	HN	CR	WR	AL
Zero One Zero	1	1	0	1	0
President Clinton	0	1	1	1	0
Salmon	1	1	0	0	0
Leave Shop	0	0	1	1	0
Old Man Fish	0	1	0	0	1
Ice Age	0	1	0	1	0

Abbreviation	Sequence
S1	Zero One Zero
S2	President Clinton
S3	Salmon
S4	Leave Shop
S5	Old Man Fish
S6	Ice Age

Abbreviation	Key word
SN	Sun
HN	Human
CR	Car
WR	Water
AL	Animal

$(\mathcal{C}, \sqsubseteq)$ is a complete lattice called the concept lattice corresponding to the context $\mathcal{K}_{vid}$. In the following, $(\mathcal{C}, \sqsubseteq)$ will be denoted by $\mathcal{L}(\mathcal{C})$. Figure 1 shows the concept lattice $\mathcal{L}(\mathcal{C})$ corresponding to indexed video sequences formal context $\mathcal{K}_{vid}$ given by table 1.

One important characteristic of $\mathcal{K}_{vid}$ is that the set T of annotations is carefully structured during the construction of XML indexing files so that its cardinality remains small [1]. This particularity led us to choose the Godin algorithm [8] to generate the corresponding concept lattice since the context is small and sparse. In addition, as mentioned in section 3.1, this algorithm allows the addition of new concepts to an existing lattice. This aspect is useful for the querying method described in section 4.2.

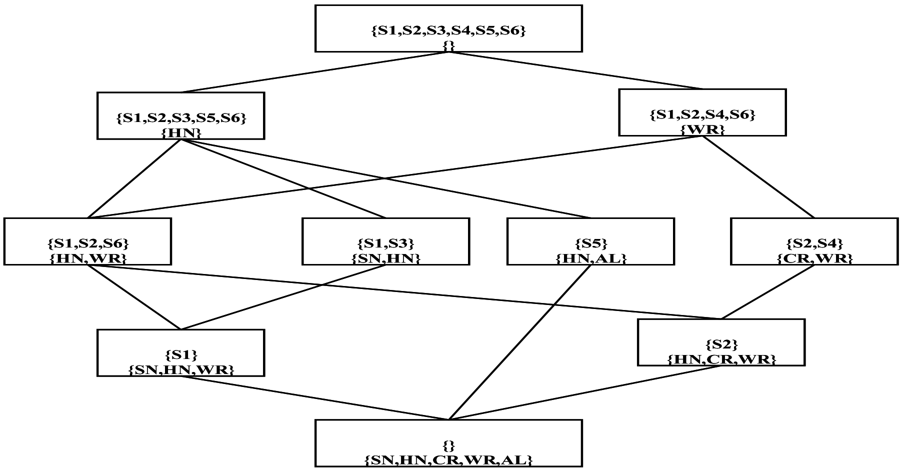

Fig. 1. Concept lattice $\mathcal{L}(\mathcal{C})$ corresponding to formal context $\mathcal{K}_{vid}$

4.2 Querying lattice of video sequences

Relevant video sequences retrieval We define a query as a formal concept $Q = (Q_A, Q_B)$, where $Q_A = Query$, i.e. a name for the extent to be formed and Q_B is the set of annotations to be used during the search. As an example, consider a query that searches video sequences with the annotations *Sun* and *Car*. Using the abbreviations given in table**??**, the query is represented by $Q = (\{Query\}, \{SN, CR\})$.

Once Q is given, it has to be classified in the concept lattice $\mathcal{L}(\mathcal{C})$ using the incremental classification algorithm of Godin et al. [8]. The resulting concept lattice is noted $(C \oplus Q, \sqsubseteq)$ where $C \oplus Q$ represents the new set of concepts once the query has been added. In the following, the concept lattice $(C \oplus Q, \sqsubseteq)$ will be denoted by $\mathcal{L}(\mathcal{C} \oplus Q)$. For the given example, the modified concept lattice $\mathcal{L}(\mathcal{C}\oplus Q)$ is illustrated by figure 2. Numbered rectangles point out new or modified concepts due to the insertion of the query. Only these concepts share properties with the query and could thus be interesting for the user. The query concept is denoted by Q either in $\mathcal{L}(\mathcal{C})$ or in $\mathcal{L}(\mathcal{C} \oplus Q)$. If there exists in the lattice $\mathcal{L}(\mathcal{C})$ a concept of the form $(A, Q_B \cup B)$, then the classification of Q in $\mathcal{L}(\mathcal{C})$ will produce a subsumer concept of the form $(\{Query, A\}, Q_B)$ that will be the new query concept to be considered. For the sake of simplicity, we continue to denote by Q the query concept in $\mathcal{L}(\mathcal{C} \oplus Q)$.

Definition 1. *A video sequence is relevant for a given query if and only if it shares at least one annotation mentioned in the query. The degree of relevance is given by the number of annotations shared with the query and by the stage during which the video sequence is added to the result.*

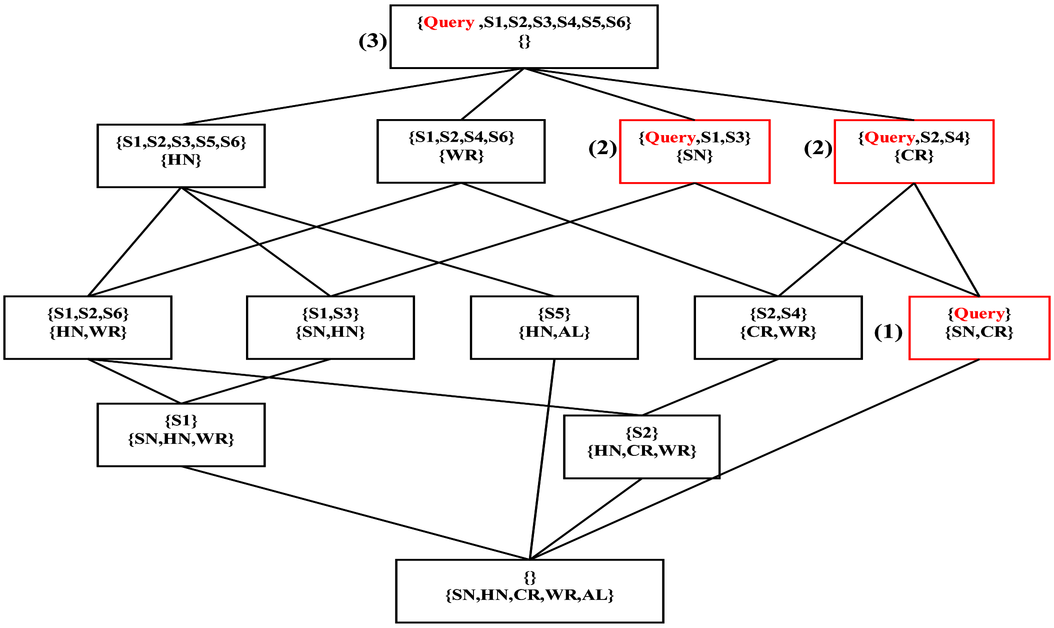

Fig. 2. The modified concept lattice $\mathcal{L}(\mathcal{C} \oplus Q)$

This definition of relevance is the basis of the retrieval process in the lattice and differs from the neighbourhood notion used in [7]. The latter can lead to retrieved documents lacking any query term which is acceptable in document retrieval but not suitable to our needs. The above definition of relevance is sufficient to explain the retrieval algorithm detailed hereafter.

Considering the above definition, all the relevant video sequences are in the extent of Q and its subsumers in the concept lattice (indicated by numbered rectangles other than the top concept in Figure 2) since the intent of each one of these concepts is a subset of Q_B (the intent of the query concept). In the following, we will denote by $\mathcal{R}_{videos}$ the set of relevant video sequences for the considered query. It is important to mention here that all the video sequences in $\mathcal{R}_{videos}$ do not have the same relevance. In fact, they are ranked according to the number of shared annotations with the query and according to the stage during which they have been added to $\mathcal{R}_{videos}$.

Intuitively, the relevant video sequences retrieval algorithm consists first in classifying the query concept in the lattice. This operation instantiates the extension $\{Query\}$ (actually, Query could be considered as a variable to be instantiated). Then, the set of video sequences that are inherited from the subsumers of the query concept Q in the lattice are gathered in the result $\mathcal{R}_{videos}$. The rank of the returned video sequences may be memorized according to the distance of the sequences to the query concept. Consider C_1 the direct (most specific) subsumers of Q in the concept lattice. The set of video sequences in the extents of the concepts in C_1 and not already in the result are added to the result. The next step consists in considering the direct subsumers of concepts in C_1 (subsumers of distance two of Q) and adding new emerging video sequences to the result set $\mathcal{R}_{videos}$. Then we continue in the same way for C_2, C_3 etc until we reach an empty set C_n. In Figure 2, the numbers near the concepts show the iterations of the explained algorithm.

We have developed a system that allows the indexing, classification, research and visualization of video sequences for an user query.

5 Discussion and future work

In this paper, we have presented an approach combining formal concept analysis and multimedia mining for an information retrieval problem in video sequences. The approach is intended for the problem of relevant video sequences choice for a further interrogation. Such a choice rests on some data quality information expected by the user.

Acknowledgment

The author would like to thank Nizar Messai for his contribution in most parts of this work and for his valuable help.

References

1. N. Mimouni, *Indexation et recherche des séquences vidéo en utilisant les treillis de Galois*, Master Dissertation, Departement of Computer Science, University of Tunis, Jul. 2005.
2. K. Egiazarian and A. E. Hassanien, *Multimedia Mining in Soft Computing*, http://ai.ijs.si/informatica/.
3. J. Clech, *Contribution Méthodologique a la fouille de Données Complexes*, Ph.D. dissertation, Universite Lumiere Lyon 2, Mar. 2004.
4. M. Scuturici, *Contribution aux techniques orientées objet de gestion des séquences vidéo pour les serveurs Web*, Ph.D. dissertation, Laboratoire ERIC, Institut National des Sciences Appliquées de Lyon, 2002.
5. N. Messai, M. D. Devignes, A. Napoli and M. Smal-Tabbone, "Querying a bioinformatic data sources registry with concept lattices," *13th International Conference on Conceptual Structures (ICCS '05) Conceptual Structures: Common Semantics for Sharing Knowledge*, Kassel, Germany, July 2005.
6. R. Godin, G. W. Mineau, and R. Missaoui, *Méthodes de classification conceptuelle basées sur les treillis de Galois et applications*, Revue d'intelligence artificielle, 9(2) pp. 105–137, 1995.
7. C. Carpineto and G. Romano, *Order-theoretical ranking*, Journal of the American Society for Information Science, 51(7), pp. 87-601, May 2000.
8. R. Godin, R. Missaoui, and H. Alaoui, *Incremental Concept Formation Algorithms Based on Galois (Concept) Lattices*, Computational Intelligence, 11, pp 246-267, 1995.