Indexing and Mining Audiovisual Data

Pierre Morizet-Mahoudeaux

Department of Computer Science and Engineering

UMR-CNRS 6599 - Heudiasyc

University of Technology of Compiègne, France

pmorizet@hds.utc.fr

Outline

Introduction

Revisiting the notion of document

Indexing audio-visual documents

Multimedia indexing

Mpeg 7

Current approaches

1 – Introduction

- Motivation
- Outcomes of digitised audio-visual documents production
- Major issues

Motivation

- Archiving: Professional Archives INA, France
  - 1.1 million hours of radio and TV programmes;
  - 575 000 hours / Radio ;
  - 535 000 hours / TV
  - $\approx$ 70 000 hours a year of audio and audio-visual programs since 1995.
- Legal deposit:
  930 000 hours (430 000 hours of TV, 500 000 hours of radio).
- 2.5 millions documents covering 113-km of shelf space; 8-km / year
- 133 years for watching or listening all archives;
- Restoring
- Allowing access

Outcomes of digitised audio-visual documents production

- Dematerialisation (annihilation) of contents
- Integration of several media on the same medium
- Temporal objects digitisation

➥ What is a document ?

➥ What is indexing ?

➥ What are the numeric multimedia possibilities ?

2 – Document

- What is a document ?
- The document-medium articulation
- Paper-based textual documents / Temporal documents
- Conclusions of section 2

What is a document ?

Definition: a document is a *content* instituted by a *publication act*, written down a medium, which possesses a *spatial* and *temporal delimitation*.

- Content: a pattern written down a medium, which can be meaningfully interpreted by a person or a community.
- Publication act: a meaningful material pattern is a document whenever it becomes available to some public.
- Spatial delimitation: to know where the document begins and where it ends.
- Temporal delimitation: the meaningful material pattern of a document must remain the same, whatever the time of its consultation or reading. The modification of its content gives rise to a new publication.

The document-medium articulation-1/3

The document-medium articulation can be described along six characteristics

- The recording medium
- The recording pattern
- The appropriation medium
- The appropriation physical set
- The semiotic appropriation form
- The appropriation modality

The document-medium articulation-2/3

- The recording medium
  - paper for a book
  - videotape for an audio-visual document or silver film for a movie
  - accessible memory for a digitised document
- The recording pattern
  - typography of an alphabet for a book
  - magnetic signal for a video
  - binary code for numerical documents

The document-medium articulation-3/3

- The appropriation medium
  - paper
  - screen
  - speaker
- The appropriation physical set
  - magnetic signal (recording pattern) → video-tape (recording medium) → TV set (appropriation medium) → visual signal (colour pixels on the screen)
- The semiotic appropriation form
- The appropriation modality

Paper-based textual documents / Temporal documents

For paper-based documents:

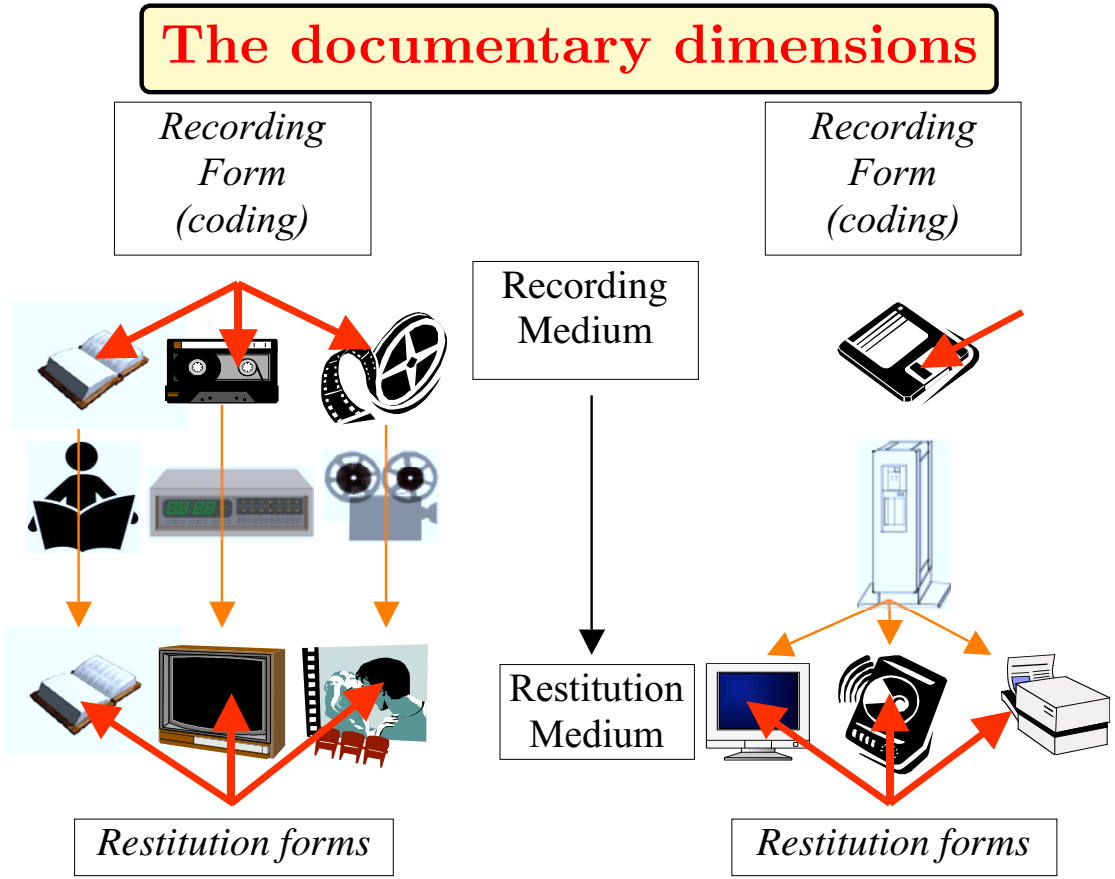
- The recording medium = the appropriation medium
- The recording form = the appropriation form

When we come to temporal documents, we must distinguish:

- Static and spatial appropriation forms
- Temporal and dynamic appropriation forms

In the second case:

- The recording form must be spatial, and the semiotic form must be temporal
- The recording medium $\neq$ the appropriation medium
- The recording form $\neq$ the appropriation form



- A mechanical means is mandatory for re-building a temporal document
- A *player* makes it possible to decode the recording form
- It makes accessible the semiotic form of the document

Conclusions of section 2

Where lies the true nature of a document ?

- A temporal document exists only during its temporal progress
- The recording medium is not a document: it is a resource
- The viewed document is always the result of a (computed) transformation
- The transformation process uses meta data of the document

➢ Meta data makes it possible to use and exploit information. They are a generalisation of index and information retrieval concepts

➢ Meta data must be created simultaneously with the document, and not afterwards like index in general

3 – Indexing

- What is indexing in general ?
- Acuteness and granularity of indexing
- Digitised documents indexing
- Content-based indexing
- The three steps of indexing

What is indexing in general ?

According to the AFNOR standard indexing is:

- Describing and characterising a document by the representation of the concepts contained in the document
  - the document is analysed for concept extraction
  - concepts are translated into a documentary language with tools such as thesaurus, classification, etc.
- Recording the concepts in some organised file, which can be easily accessed
  - files are organised as paper or digitised sheets, indices, tables, catalogues, etc.
  - queries are handled by these files

This definition concerns textual documents, it must be generalised:

➥ We will introduce later the notion of descriptors

Acuteness and granularity

Acuteness depends on the wanted faithfulness of the research results. It is defined by:

- The concepts repository
- The structure of indices

Granularity depends on the wanted *manipulation unit*.

- It is *the book* or *the video-tape* for librarian or archivist,
- It is *the page* or *the scene* for the reader.

Outcomes of digitising: Granularity and digitising

- Digitising the physical support brakes the links between the digital unit and the document manipulation unit. Ex.: the pixel
- Digitising the semiotic support means that a link between each unity and the content interpretation has been defined
  - ex.: the alphabet characters
- The case of audiovisual documents: the MPEG standards
  - MPEG-1 and MPEG-2 specify compression and digitising standards based on the pixel unit. They have no link with the interpretation of the documents
  - MPEG-4 analyses the audiovisual flow as a set of objects linked by spatio-temporal relations. Example: the television screen background.

➥ Indices can be based on markup languages for digitised documents

Content-based indexing

- For textual documents
  - Queries are built with words (strings built with alphabetic characters).
  - The system performs a comparison with patterns (strings, words, ...) in the text, separated by spaces or punctuation characters.
- For audiovisual documents we must make the assumption that
  - Queries built with sounds or images make it possible to find similar documents based on a distance measurements between images or sounds.
  - We can index a document content with image-indices or sound-indices.
- The second assumption does not hold.
- The first assumption does not refer to the exact notion of index, but rather to descriptors.

Descriptors / Indices

- A descriptor is a piece of information extracted from a document by automatic means. It is close to the physical content of the document.
- An index is a position or a value given in the scope of an interpretative system associated to the management of a set of documents.

Descriptors / Indices (continued)

- Two descriptors containing different values of the same kind of information can refer to the same interpretation of documents. Ex. Two different colour histograms do not mean that the images from which they have been extracted are significantly different.
- Two different indices or concepts refer to two different interpretations of the contents.
- Indices can by defined based on languages, because they rely on the concepts hold by words.
- We have no such generic and stable concepts for images or sounds

Interpretation of image descriptors

What does show this image: a storm, a criminal attempt, a riot ?

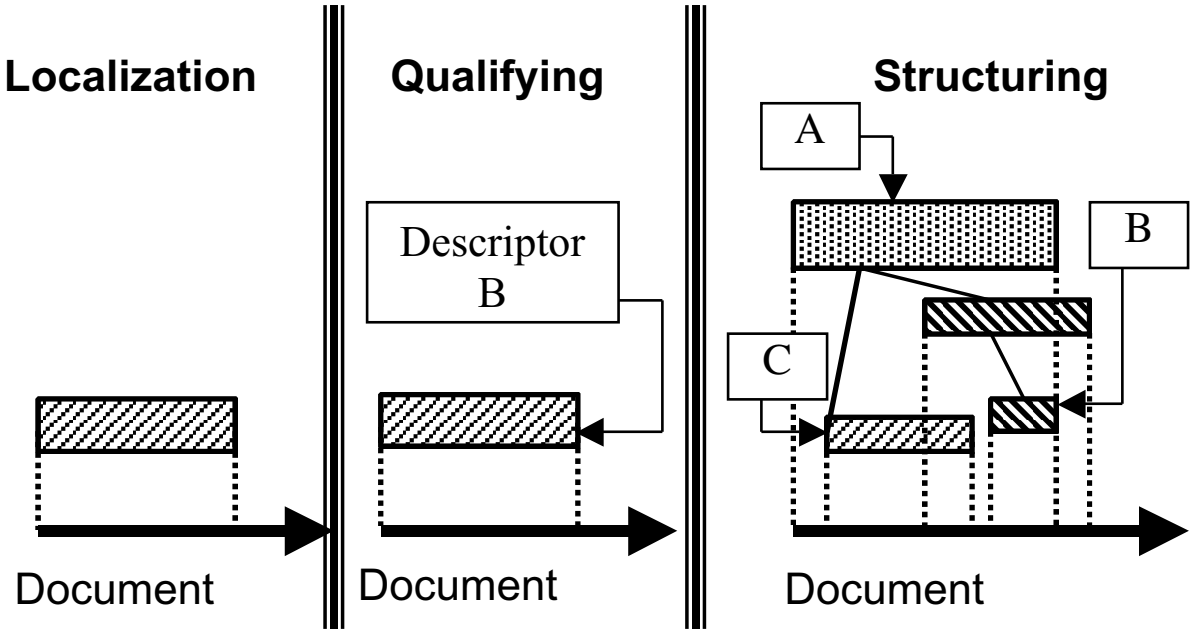
Audiovisual descriptors

- Iconic descriptors.
- Key-frames.
- Descriptors of the digital signal (texture, histograms, ...)
- Descriptors computed from descriptors of the digital signal

Three types of indexing

- Conceptual indexing
  The content is qualified by a concept, which describes what it is about in the document.
- Structural indexing
  These indices describe how two indexed segments of the document are linked.
- Content-based indexing
  Rather than indexing it is content-based description. It concerns the extraction of characteristics. It makes it possible to link a document content, according to some distance measurement.

The three steps of indexing



- Localisation: where the descriptors are,
- Qualifying: what do they concern,
- Structuring; how they are each other linked.

4 – Multimedia indexing

- Temporal objects
- Computer information systems
- Indexing a temporal object

Temporal objects

They are contents.

- Their semiotic restitution form is dynamic or temporal
- The order and the rhythm of reading is imposed
- They have a duration
- They are not randomly meaningfully readable
  - In an analog context: they need dedicated, specialised reading tools.
  - In a numerical context:
    * Digitising provides tools for signal processing (compression, scripting, watermarking).
    * Temporal contents can be integrated in a computer information system for exchange and management.

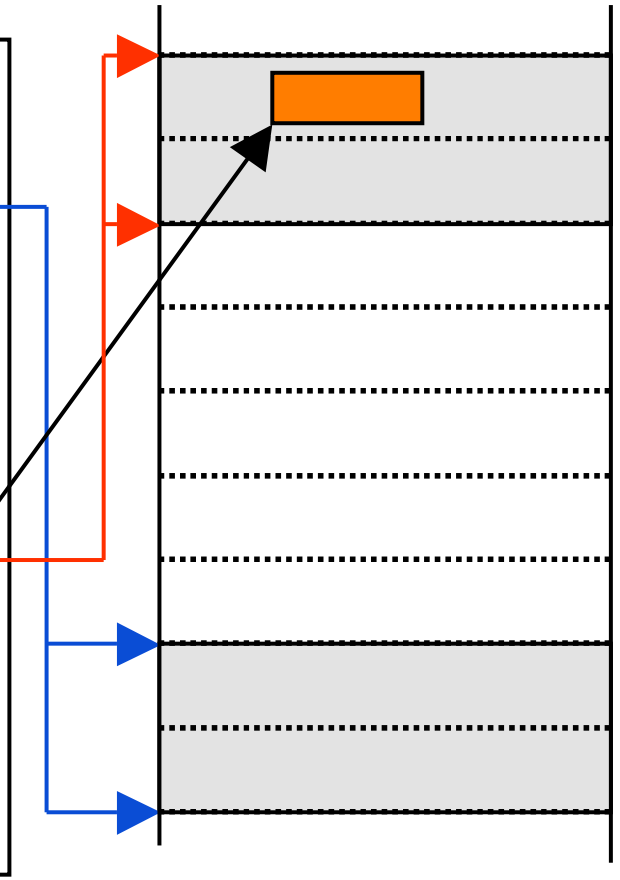
Computer information systems

- The documentation and the documents line:
  - production (scripts, story-board, ...)
  - broadcasting (TV magazines, TV conductors, ...)
  - archiving (notes, annexes, ...)
- Document and documentation link.
- Integration of the document and the documentation.

The content and its documents are integrated on the same support (ex. DVD). Markup are not in the document by itself, but in a separated document (notes, review). The indexed document and its indices must be linked. The description order may be different from the order of the segments in the document.

Indexing a temporal object

<title> the title <\title>
<anthology-extracts>
<loctime> 14:00 -- 18:00 <\loctime>
<title> shower scene <\titre>
<\anthology-extracts>
<autthor> Hitchcock <\author>
<scene> this is a scene
<loctime> 12: 43 -- 13:50 <\loctime>
<object-of-interest> an object
<locspace> x+d1:y+d2 <\locspace>
<\object-of-interest>
<\scene>



The document information is not in the document, and its organisation is not that of the described document.

5 – MPEG-7: Multimedia Content Description Interface

- General Description of MPEG-7
- Structural information of MPEG-7
- MPEG-7 basic descriptors and schemata
- Example

General Description of MPEG-7

MPEG-7 is a new standard for the description of multimedia content, designed to assist content-based access to multimedia data.
The standard defines three kinds of information that comprise the description:

- Creation and usage information:
  It concerns mostly textual information (metadata).
- Structural information
  It concerns low-level and machine-oriented description.
- Semantic information
  It concerns the conceptual aspects to express high-level and human oriented information. They deal with semantic entities, such as objects, events and concepts.

Structural information of MPEG-7

Several information structures are available

- Features: they concern all what we need to express for describing the document (authors, scenes, segments, ...)
- Descriptors: they are the formal representation of a feature,
- Schemata description: they correspond to structures linking two different descriptors or other schemata descriptors. To a certain extent, they are an adaptation to MPEG-7 of the notion of DTD of XML, or of Schema of XML-Schema
- The *Description Definition Language* (DLL), to create descriptors and description schemata. It is an extension of XML-Schema

MPEG-7 basic descriptors and schemata

MPEG-7 proposes Three standards:

- sound schemata,
- video schemata,
- multimedia schemata.

Sound schemata

They are designed to allow the expression of information about the sound signal, which can be extracted with tools available today.

- sounds effects
- the quality in tone,
- the speech,
- the melodies,
- low-level descriptions (temporal envelope, spectral representation, harmonics, . . .).

A ”silence” descriptor permits to describe a silent content

Video schemata

- Basic descriptors to localise, with a variable accuracy, a part of a visual or video content,
- Colour descriptors: quantisation, colour space, prevailing colour, ...
- Texture: partition of homogeneous texture area, histograms, ...,
- Patterns: area analysis, edges representation, three dimensions patterns,
- Movements: camera movement, trajectory, movement analysis, ...
- Localisation: spatio-temporal localisation of a video sequence as a "pipe" or spatio-temporal region,
- Face recognition.

Multimedia schemata

- Content management: content creation, production, encoding, coding and files formats,
- Content description: it concerns the structural and semantical aspects. The description of the semantical content uses conceptual description, based on the "real world", which is close to human understanding of multimedia information.
- Browsing and content accessing: summary structures, as well as partitions or decompositions are available. "Variations" are the description of different resources, which represent variants of the same content depending of the context. For example, the linguistic variants of the same audio-visual program.

Multimedia schemata (continued)

- Content organisation: for the description of the organisation of collections of objects, events or content segments.
- User interaction: it concerns user preferences or profile for content reading. They can be used with the "variation" schemata to propose user-adapted content (the user language for example).

Example: description schema of the video mounting transitions

```
<!-- ################################################-->
<!-- Definition of Transition DS -->
<!-- ################################################-->
<complexType name="TransitionType">
  <complexType>
    <extension base="mpeg7:VideoSegmentType">
      <sequence>
        <element type="GradualEvolution"
                 type="mpeg7:GradualEvolutionType"
                 minOccurs="0"/>
        <element name="SpatioTemporalLocator"
                 type="mpeg7:SpatioTemporalLocatorType"
                 minOccurs="0"/>
      </sequence>
      <attribute name="editingLevel" use="optional">
        <simpleType base="string">
          <enumeration value="global"/> <!--Or InterShot-->
          <enumeration value="composition"/>
```

Example: description schema of the video mounting transitions

TransitionType	Describes a transition realised between two edited video segments during an editing process. Three different types of transitions are distinguished based on the value of the attribute editingLevel: global transitions, composition transitions, and internal transitions. Three different types of transitions are also distinguished based on the value of the attribute evolution: cuts, analog cuts, and gradual transitions.
GradualEvolution	Describes the transition when it is a gradual (optional). It does not apply to transitions that are cuts and analog residual cuts.

6 – Current approaches

- Text mining
- XML Schema
- Ontology-based mining

Text mining: Statistical analysis approaches

- Let N the total number of documents and idf_i the number of documents where the string i appears.

$$tr_i = \log \frac{N - idf_i}{idf_i}$$

  Denotes the level of discrimination of the document by this string.

- It is preferable to take into account the distribution of words in one document together with the distribution of words in the collection of documents.

  Let N the total number of documents, $tf_{i,j}$, the number of occurrences of the word w_i in the document D_j, and n the total number of different words in D_j. We compute the combination of two parameters :

$$P_{1_{i,j}} = (1 + \log tf_{i,j}) \frac{1}{\sqrt{(1 + \log tf_{1,j})^2 \ldots (1 + \log tf_{n,j})^2}}$$

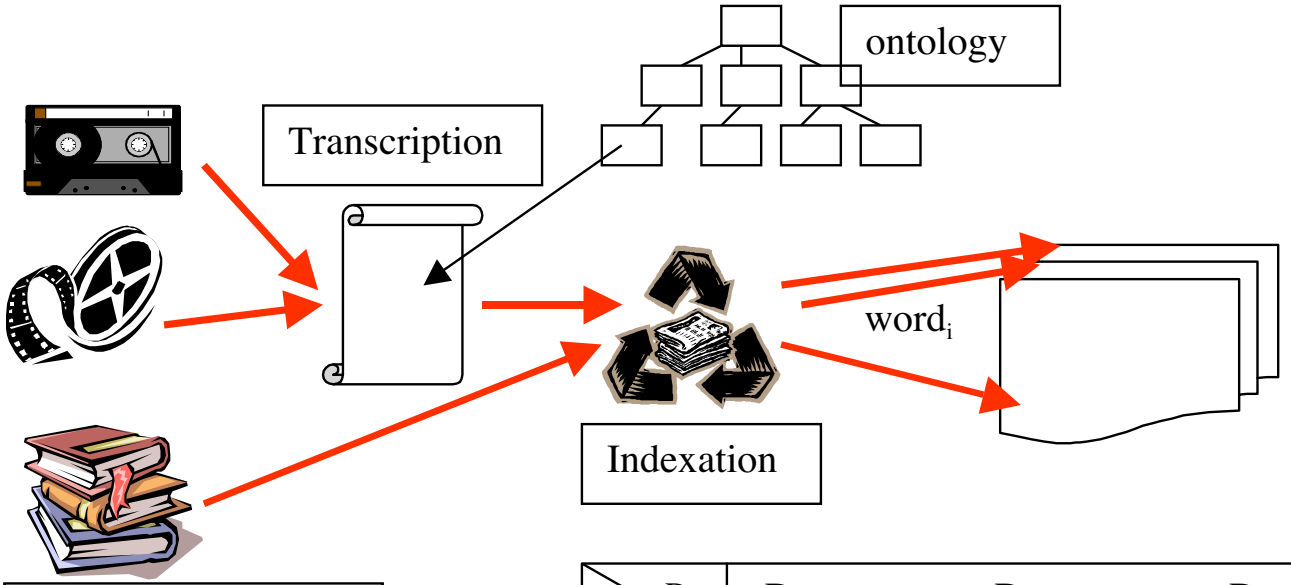
$$P_{2_{i,j}} = (1 + \log tf_{i,j}) \log \frac{N}{df_i}$$

where, df_i is the number of documents, where the word w_i appears.

This basic notion is completed by the statistical disribution of strings in the documents or linguistic approaches to compute the weight of strings.

$$index_i : d_k(weight_{i_k}) : d_l(weight_{i_l}) : d_m weight_{i_m})$$

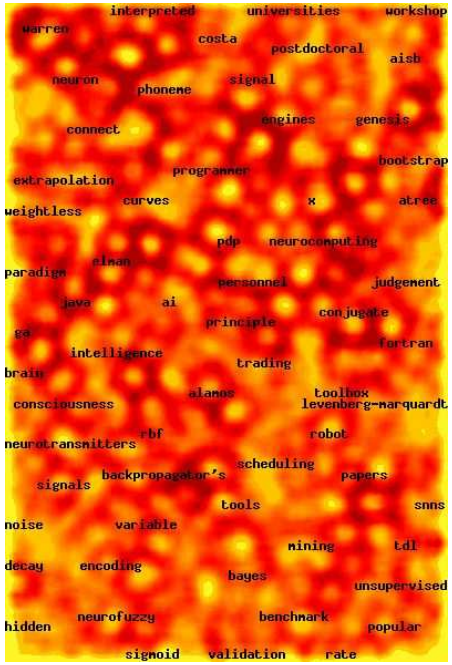
Text mining with meta-knowledge



Index + context :
- arounding words
- doc
- temp. pos. in doc.
- ontology

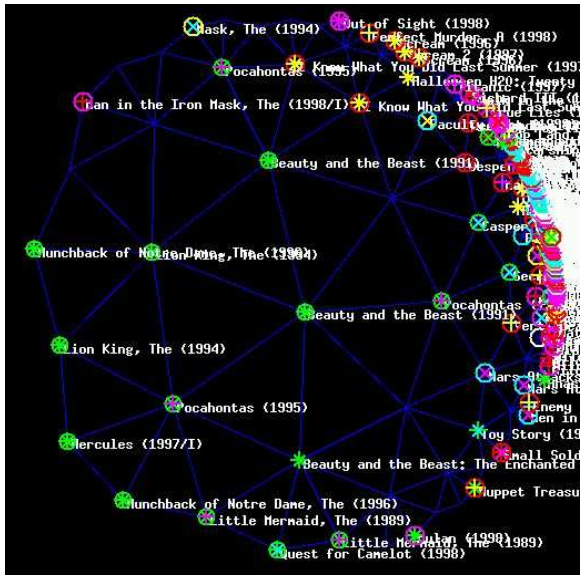
Index \ Doc	D_1 ……. D_j …….. D_n
$word_i$	…….. $f_{i,j}$

Text mining with SOM (Self organisation maps: Kohonen)



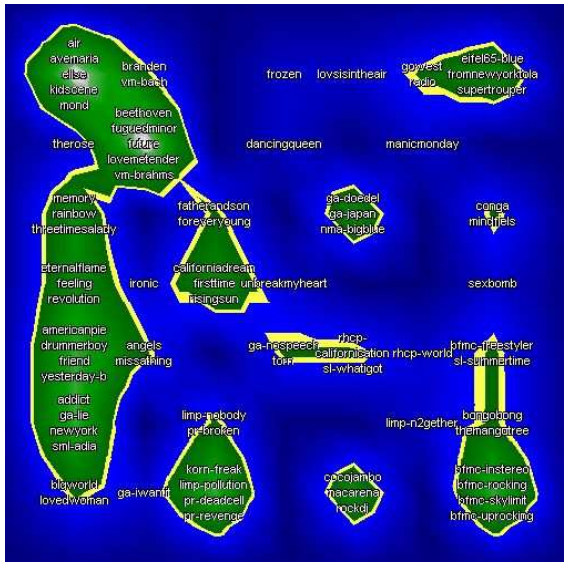
WEBSOM

Kohonen, 1997



HSOM

Ritter, 1999



Islands of Music

Pampalk 2001

XML-Schema and databases (I. Varlamis, M. Vazirgiannis)

Define a XML-Schema for describing documents, objects and relations. The XML-Schema file contains five different types of subelements.

- xsd:element: They define the name and type of each XML element.
- xsd:attribute: They describe the name and type of attributes of an XML element. They can be of simple (integer, float, string, etc.) or complex type (i.e. enumerations, numerical range etc.)
- xsd:simpleType: They define new datatypes that can be used for XML attributes.
- xsd:attributeGroup: They group xsd:attribute definitions that are used by many XML elements.
- xsd:complexType: They represent the various entities of the metadata model. They contain:
  - one ore more <xsd:attribute> tags
  - one ore more<xsd:element> tags

Example of a XML element: Audio-visual

```
<xsd:element name="AudioVisual type=" AudioVisualDS " />
<xsd:complexType name="AudioVisualDS">
     <xsd:attribute name="id" type=" ID use="required />
     <xsd:attribute name="AVType" type="AVTypeD" use="required"/>
     <xsd:sequence>
<xsd:element maxOccurs="1" minOccurs="0"name="Syntactic" type="SyntacticDS" />
<xsd:element maxOccurs="1" minOccurs="0"name="Semantic" type="SemanticDS" />
<xsd:element maxOccurs="1" minOccurs="0" ref="MetaInfoDS" />
     </xsd:sequence> </xsd:complexType>
<xsd:simpleType name="AVTypeD">
     <xsd:restriction base="string">
          <xsd:enumeration value="Movie" />
          <xsd:enumeration value="Picture" />
          <xsd:enumeration value="Document" />
     </xsd:restriction> </xsd:simpleType>
```

Structure of the XML Schema file

It contains the definition of the various entities and some supportive structure, and the description of the structure of the database commands.
The extension elements are:

- DBCommand
- DBInsert, DBUpdate
- DBDelete
- DBReply
- DBSelect

Mapping XML-SCHEMA into a Relational schema

- the XML-Schema file that contains information about structure of the exchanged XML documents is used to generate the relational database structures
- XML documents are parsed to contruct the appropriate SQL commands

Audio-visual Event Description Interface (INA))

- It is a representation model for meta-data.
- It permits the expression of description schemata, which defines the n dimensional structure of an audio-visual document.
- The meta-data, which are produced by the AEDI model are linked to the content of the audio-visual documents.
- It has been implemented in the MIPE environment (*Multimedia Indexing Publishing Environment*), which is an editor/composer/publisher environment.

AEDI Modules

- The *ontology* defines the semantics of concepts used for describing the audio-visual contents.
- The *description schema* defines the classes of objects available for the description of documents, and the structural relations between classes.
- The *AEDI description* defines the documents segments, which have been indexed and builds the descriptors instances corresponding to the classes defined in the description schema.
- The *covering texts* (notices, transcriptions, scripts, copyrights, ...), are XML encoded and linked to some AEDI description by Hypertext links. For example the text/sound can be aligned according to a transcription.
- The *audio-visual document* is a specific temporal selection in one or several video or audio files.

Conclusion

- the main outcome of audio-visual document digitizing is the integration of all documents linked to a video production line on the same support.
- the main consequence is a revisiting of the definition of indexing, which generalizes the text indexing definition.
- breaking the document integrity and putting all documents on the same support makes it possible to define indices linked to the semantical objects, which compose an audio-visual document. The most important property of these objects is that they can be temporal objects.
- Marked-up languages has open the possibility of developing indexing based on a semantical document description. It also made it possible to propose effective computer-based systems.

7 – References

- Gwendal Auffret and Jean Carrive and Olivier Chevet and Thomas Dechilly and Remi Ronfard and Bruno Bachimont, Audiovisual-based Hypermedia Authoring - using structured representations for efficient access to AV documents, in ACM Hypertext, 1999
- G. Akrivas, S. Ioannou, E. Karakoulakis, K. Karpouzis,Y. Avrithis, A.Delopoulos, S. Kollias, M. Vazirgiannis, I.Varlamis. An Intelligent System for Retrieval and Mining of Audiovisual Material Based on the MPEG-7 Description Schemes. European Symposium on Intelligent Technologies, Hybrid Systems and their implementation on Smart Adaptive Systems (EUNITE), Spain, 2001.
- Bachimont B. et Dechilly T., Ontologies pour l'indexation conceptuelle et structurelle de l'audiovisuel, Eyrolles, dans Ingénierie des connaissances 1999-2001.
- Bachimont B., Isaac A., Troncy R., Semantic Commitment for Designing Ontologies : a Proposal. 13th International Conference EKAW'02, 1-4 octobre 2002, Siguenza, Espagne

References (continued)

- Bachimont B. Indexation multimédia, in Assistance intelligente à la recherche d'information, édité par Eric Gaussier et Marie-Hélène Stefanini,, Hermès, 2002
- Bachimont B. "audiovisual indexing and automatic analysis: problems and perspectives from an archiving point of view" Imagina 2002, Monaco, 12 february 2002
- Bruno Bachimont. Le document audiovisuel, le numérique, ses usages et son archivage. Le document multimédia en sciences du traitement de l'information. Ecole thématique du CNRSGDR I3, Documents et évolution. Cépaduès-Editions, tome 1, p. 111-128,2000.
- Bachimont B., (1998), MPEG-7 and Ontologies: an editorial perspective, Virtual Systems and Multimedia 98, Gifu, Japan, Oct. 1998.
- Carrive, J., Pachet, F., Ronfard, R. (1998). Using Description Logics for Indexing Audiovisual Documents. International Workshop on Description Logics (DL'98), Trento, Italy, ITC-Irst pp.116-120, 1998.

References (continued)

- Carrive J., Pachet F., Ronfard R. (2000b), Logiques de descriptions pour l'analyse structurelle de film, in Ingéniérie des Connaissances, évolutions récentes et nouveaux défis, Charlet J., Zacklad M., Kassel G., Bourigault D. (ed.), Eyrolles, pp. 423-438.
- Carrive J., Pachet F., Ronfard R. (2000a), Clavis: a temporal reasoning system for classification of audiovisual sequences, in proceedings of the Content-Based Multimedia Access (RIAO'2000), Paris, France, April 12-14, pp. 1400-1415.
- Gwenaël Durand, P. Faudemay, Cross-indexing and access to mixed-media contents, Proc. CBMI01 International Workshop on Content-Based Multimedia Indexing, Brescia, Italy, September 2001.
- M. Erwig, R. H. Gueting, M. Schneider, M. Vazirgiannis. Spatio-Temporal Data Types: An Approach to Modeling and Querying Moving Objects in Databases, GeoInformatica Journal 3(3), Kluwer Publishers, 1999.

References (continued)

- T. Kohonen, S. Kaski, K. Lagus, J. Salojärvi, J. Honkela, V. Paatero, and A. Saarela. (2000) Self Organization of a Massive Document Collection. IEEE Transactions on Neural Networks, Special Issue on Neural Networks for Data Mining and Knowledge Discovery, volume 11, number 3, pages 574-585. May 2000.
- Lagus, K. (2000) Text Mining with the WEBSOM. Acta Polytechnica Scandinavica, Mathematics and Computing Series no. 110, Espoo 2000, 54 pp. D.Sc.(Tech) Thesis, Helsinki University of Technology, Finland. URL http://www.cis.hut.fi/krista/thesis/
- Lespinasse K. and Bachimont B. (2001). Is Peritext a Key for Audiovisual Documents ? The Use of Texts Describing Television Programs to Assist indexing. CICLing2001, Mexico, Springer-Verlag. conference: CICLing2001.
- T. Markousis, D. Tsirikos, M. Vazirgiannis, G. Stavrakas. A Client-Server Design for Interactive Multimedia Documents based on Java, in Elsevier - Computer Communications Journal, 2000.

References (continued)

- I. Mirbel B. Pernici T. Sellis S. Tserkezoglou, M. Vazirgiannis. Checking Temporal Integrity of Interactive Multimedia Documents, in Very Large Data Bases journal,9(2): 111-130, 2000.
- I. Mirbel, B. Pernici, M. Vazirgiannis. Temporal Integrity Constraints in Interactive Multimedia Documents, in the proceedings of IEEE-International Conference on Multimedia Computing and Systems (ICMCS '99), Florence, Italy, June 1999.
- J. Ontrup and H. Ritter, Text Categorization and Semantic Browsing with Self-Organizing Maps on non-euclidean Spaces, Proceedings of PKDD-01 5th European Conference on Principles and Practice of Knowledge Discovery in Databases, pp. 338-349, ed. L. De Raedt and A. Siebes, publ. Springer LNAI 2168, 2001
- J. Ontrup and H. Ritter, Hyperbolic Self-Organizing Maps for Semantic Navigation, Advances in Neural Information Processing Systems 14, 2001

References (continued)

- I. Varlamis, M. Vazirgiannis, P. Poulos. Using XML as a medium for describing, modifying and querying audiovisual content stored in relational database systems, International Workshop on Very Low Bitrate Video Coding (VLBV). Athens 2001.
- E. Pampalk, A. Rauber, and D. Merkl, Proceedings of the ACM Multimedia 2002, pp 570-579, Juan les Pins, France, December 1-6, 2002, ACM.
- E. Pampalk, A. Rauber, and D. Merkl. Proceedings of the International Conference on Artificial Neural Networks (ICANN'02), Springer Lecture Notes in Computer Science, Madrid, Spain, August 27-30, 2002, Springer.
- I. Varlamis, M. Vazirgiannis. Bridging XML-Schema and relational databases. A system for generating and manipulating relational databases using valid XML documents, in the proceedings of ACM Symposium on Document Engineering, Nov. 2001, Atlanta, USA
- I. Varlamis, M. Vazirgiannis. Web document searching using enhanced hyperlink semantics based on XML, proceedings of IDEAS '01, Grenoble, France.

References (continued)

- M. Vazirgiannis D. Tsirikos, Th.Markousis, M. Trafalis, Y. Stamati, M. Hatzopoulos, T. Sellis. Interactive Multimedia Documents: a Modeling, Authoring and Rendering approach, in Multimedia Tools and Applications Journal (Kluwer Academic Publishers), 2000.
- M. Vazirgiannis, Y. Theodoridis, T. Sellis. Spatio-Temporal Composition and Indexing for Large Multimedia Applications, in ACM/Springer-Verlag Multimedia Systems Journal, vol. 6(4), 1998.
- M. Vazirgiannis, C. Mourlas. An object Oriented Model for Interactive Multimedia Applications, The Computer Journal, British Computer Society, vol. 36(1), 1/1993.
- M. Vazirgiannis, T. Sellis. Multimedia Data Base Management Systems: Modeling and Querying Interactive Multimedia Documents, to appear in Advanced Databases (Editors, M. Piattini, O. Diaz), Artech House.

References (continued)

- M. Vazirgiannis, T. Sellis. Multimedia Database Management Systems: Modeling and Querying Interactive Multimedia Documents, in Advanced Databases (Editors, M. Piattini, O. Diaz), Artech House.
- M. Vazirgiannis. Multimedia Data Base Object and Application Modeling Issues and an Object Oriented Model, in Multimedia Database Systems: Design and Implementation Strategies, Kluwer Academic Publishers, 1996, Pages: 208-250.
- E.Veneau, R.Ronfard, P.Bouthemy, From Video Shot Clustering to Sequence Segmentation, Fifteenth International Conference on Pattern Recognition (ICPR'2000), Barcelona, september 2000.
- D. Vodislav, M. Vazirgiannis. Structured Interactive Animation for Multimedia Documents, Proceedings of IEEE Visual Languages Symposium, Seattle, USA, September 2000.