# Variations on Multimedia Data Mining


Simeon J. Simoff
Faculty of Information Technology
University of Technology, Sydney
Broadway, NSW 2009, Australia
+61 2 9514 1838
simeon@socs.uts.edu.au



## ABSTRACT

Is multimedia data mining just a new combination of buzz-words or is it a new interdisciplinary field which not only incorporates methods and techniques from the relevant disciplines, but is also capable to produce new methodologies and influence related interdisciplinary fields. Rather than making an overview of existing methods and techniques, or presenting a particular technique, this paper aims to present some facets of multimedia data mining in the context of its potential to influence some relatively new interdisciplinary domains.



## Keywords

multimedia, digital media, data mining, knowledge discovery, knowledge representation, case-based reasoning, computer-supported collaborative work.


## 1. INTRODUCTION

Multimedia and digital media, and data mining are perhaps among the top ten most overused terms in the last decade. The field of multimedia and digital media is at the intersection of several major fields, including computing, telecommunications, desktop publishing, digital arts, the television/movie/game/broadcasting industry, audio-video electronics. The advent of Internet and low-cost digital audio/video sensors accelerated the development of distributed multimedia systems and on-line multimedia communication. The list of their application spans from distance-learning, digital libraries, and home entertainment to fine arts, fundamental and applied science and research. As a result there is some multiplicity of definitions and fluctuations in terminology [2]. In this paper digital (multi)media denotes computer-mediated and controlled integration of numeric, text, graphics and other geometry representations (CAD drawings, 3D models, virtual universes), images, animation, sound, video and any other type of information medium which can be represented, stored, processed and transmitted over the network in digital form.

Another result of the rapid progress in these fields is the number of challenges for computer systems research and development, including:

- enlarged data sets with variety of formats and structures;
- variety of models for integration of media elements and components;
- demands on computational efficiency of media analysis and retrieval algorithms;
- knowledge representation schemes;
- visualisation metaphors.

Multimedia (or digital media) representations comprise a collection of domain descriptions in "native" for the domain format. Figure 1 illustrates that variety in terms of the degree to which representation is structured and to what degree the specification of this representation complies with some formal models. The term "hypermedia" is added to stress the presence of links in the digital media under consideration. of knowledge representation and specification.

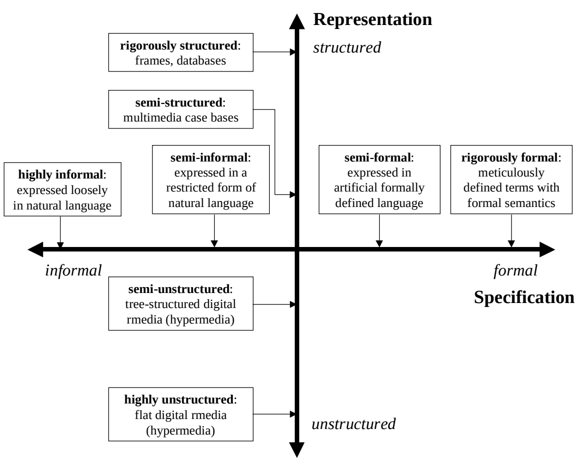



**Figure 1. Digital media and domain description (adapted from [8]).**

The specifics of the domain, where multimedia is used may influence the conceptual model that defines the representation of the multimedia content. The knowledge about the model can be invaluable in the development of multimedia data mining schemes, providing some initial assumptions and structure insights and assisting in the attribute identification.

Researchers in the database community are viewing multimedia data mining basically as an extension of the knowledge discovery



Proceedings of the International Workshop on Multimedia Data Mining (MDM/KDD'2000), in conjunction with ACM SIGKDD conference. Boston, USA, August 20, 2000
(O.R. Zaïane, S.J. Simoff, eds.)

in databases, for example, as "the mining of high-level multimedia information and knowledge from large multimedia databases," a "subfield of data mining that deals with the extraction of implicit knowledge, multimedia data relationships, or other patterns not explicitly stored in multimedia databases [14]. Consequently, the framework of KDD applied in multimedia database mining provides similar knowledge representation schemes - association, classification, characterisation and other types rule patterns. This approach, more extensively described in [15], is consistent with the overall KDD methodology.

In the information system approach, the methods of multimedia data mining are better known as multimedia information analysis and retrieval. The research in the field includes a collection of works in content-based image and video search, fusion of pictorial and other media and efficient storage organisation for multimedia data [3].

Is multimedia data mining just a new combination of buzz-words or is it a new interdisciplinary field which not only incorporates methods and techniques from the relevant disciplines, but is capable to produce new methodologies and influence related interdisciplinary fields? The aim of this paper is to present some facets of multimedia data mining in the context of case-based reasoning and computer-supported collaborative work environments.

## 2. MULTIMEDIA DATA MINING IN HYPERMEDIA CASE BASES - AN EXAMPLE OF EXTENDING AI SCHEMES

Methods developed in multimedia data mining can have significant impact on related fields from data analysis and artificial intelligence. The potential is illustrated on the use of multimedia data mining approach in hypermedia case bases for automating case-based reasoning with unstructured case data.

### 2.1 Knowledge representation

Case models based on hypermedia representations are becoming popular alternatives to the strict format of object-oriented and attribute-value representations. What exactly is denoted by a case and how it is represented are major structural issues in CBR. When in financial and business applications cases are usually well-structured object-oriented or relational attribute-value representations, in interdisciplinary domains like design, digital media production and visual reasoning, cases are represented in more informal way (see Figure 1). Among the reasons for such diversity, perhaps the major one is the limited expressive power of formal representations. Hypermedia case representation is suitable for domains where it is difficult to fit domain knowledge into structured knowledge representation schemes (see Figure 1). Hypermedia case models offer richer semantics which may be considered both as alternative and extension to the strict format of object-oriented and attribute-value representations. The hypermedia representations comprise a collection of case descriptions represented as text in free or table format and other multimedia data, such as CAD drawings, images, video, sound, etc. Another characteristic of hypermedia is the use of links, where the links can connect information within a case, between different cases, or links to data that lies outside the case library.

Usually, the representation of the cases in the library is organised according to some conceptual model of the domain. This approach towards complexity is based on the following assumptions:

- a case is a hierarchy of concepts, or "subcases";
- a case is represented by different views.

This supports case-based reasoning paradigm because subdividing a case in this way allows reasoning to focus only on the relevant parts of that case. By processing only some of the knowledge associated with a case, reasoning can become more efficient. The development of a case-base that has a hierarchical structure usually requires defining a typical decomposition of domain experience. Figure 2 presents an example of domain decomposition - the decomposition of the building design domain according to a structural engineering view of the building. Figure 3 illustrates the use of the ontology in Figure 2 for organising the access to case elements (subcases).

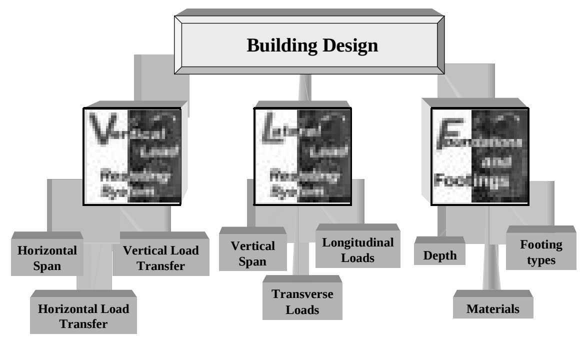



**Figure 2. Case as a hierarchy of concepts.**


**Figure 3. Example of hypermedia case organisation.**

The use of different views of a domain case recognises that the experience in a domain can be understood from different perspectives. An example of this approach is presented in [12]. In this approach, a single, complex design project is represented as multiple cases. The use of multimedia can make it easier to understand complex systems - icons, images, sketches, etc. can highlight and illustrate corresponding text or tabular information.

Figure 4 illustrates a typical case page, which includes text description, images, CAD drawings and videos. In general, page layout and format are not restricted. Different people develop the actual page content of the different cases over time. The developments in web and multimedia technologies may influence the variety of elements, included in these loosely structured case descriptions.

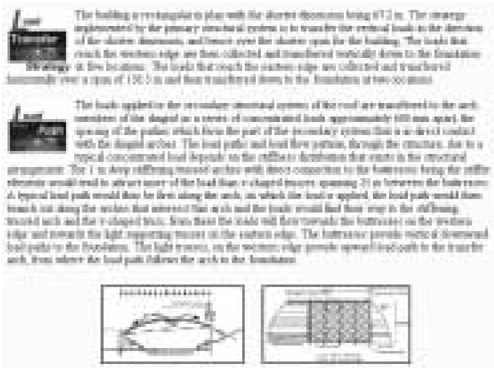
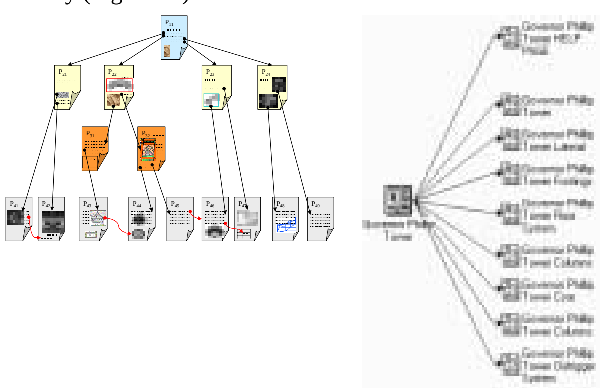


**Figure 4. An example of a case page.**

Figure 5 illustrates the idea of shaping the case structure according to a conceptual hierarchy. The semantics of the links depends on the relations between concepts that constitute the representation (Figure 5a). The tree-like structure (Figure 5b) may include some links between pages within a same level. Such links usually appear at a lower level in the hierarchy, where a concept may be related to concepts that belong to different branches in the hierarchy (Figure 5a).



a. b.

**Figure 5. Hierarchy of domain concepts, reflected in the case representation.**

The use of hypermedia to develop case base systems, however, does not solve the problems in building automated reasoning algorithms, which benefit directly from the information in the library. Hypermedia representation as it is supports human reasoning rather than automated reasoning. When this is a reasonable compromise for educational purposes, there is not much use of this approach in the research and industrial case base systems – the system remains simply a structured hypermedia handbook. A common solution is to build an additional structured representation layer, a vector of case attributes. As a rule, the methodology for building such additional layers employs some knowledge engineering techniques for identifying the attributes that represent the domain case and involves domain expert(s). The attribute-value representation of the case is linked to the "entry" page of the corresponding multimedia case representation (an example of entry case page is shown in Figure 3), as shown in Figure 6. The reasoning algorithms operate over the values of the attributes, without utilising the advantages of the multimedia information available in the case.

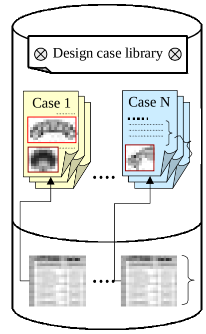



**Figure 6. Hypermedia case base with additional attribute-value representation**

## 2.2 A framework for multimedia data mining in hypermedia case libraries

Discovering implicit knowledge in hypermedia case bases is substantially different to data mining in databases. The data organisation units in database data mining are the data tables, in particular their columns or rows. Inside the hypermedia case base the organisational unit is the case or subcase, which merely consists of one or more multimedia pages. The pages comprise a variety of data formats. Knowledge discovery then, in our use of the term, involves finding patterns in primarily unstructured data. More formally the knowledge discovery process in multimedia case representations can be viewed as *machine learning where a case library replaces the training set*. The approach is illustrated in Figure 7.

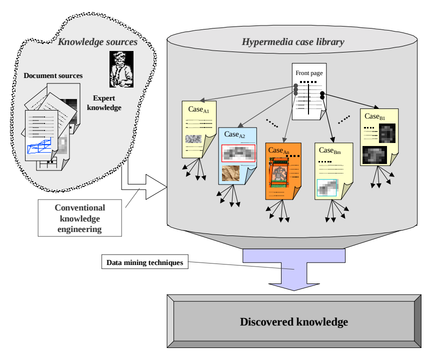



**Figure 7. Knowledge discovery in hypermedia case libraries.**

As illustrated in Figure 8, the information model of the case library constitutes the initial basis for the multimedia data mining [10]. During the data segmentation multimedia data are divided into logical interconnected segments. For example, in a hypermedia case library each segment can include one or more pages. Within particular media type a segment may have different meaning. For example, within a text a segment could be a paragraph, a sentence. The actual mining and analysis procedures are expected to reveal some relations in the segments and between the segments at different levels. For instance, the text analysis can identify relations between concepts presented in the text and the CAD drawings presented on that page (or vice versa, find that the actual CAD drawings are not connected with the content of the text description). The analysis within text segments can identify word concordances that denote complex terms, not explicitly defined in the case base. Extracted patterns are incorporated and linked under the framework of the information model. As a result there can be additional attributes, change in links, revision of identified attributes, changes in some attribute values. Some paragraphs, images or other media segments could become insignificant. Consequently, the information model should be able to accommodate changes in the structure and media content.

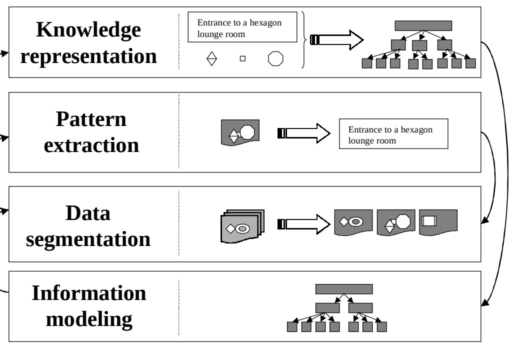



**Figure 8. A model of multimedia data mining**

The model is expanded in Figure 9. In case-based reasoning systems the specific interest can be focused in finding patterns in the cases that can assist with indexing and adapting cases as a way of improving the retrieval of related previous experience and indication when an adaptation lies outside some reasonable constraints, based on the experience in the case base. Patterns in the form dynamic thematic paths[1] can assist with the navigation in retrieved cases. The framework combines two consecutive complementary strategies - data- and hypothesis-driven exploration, discussed in more details in [12].

The above described data mining schema can be integrated in the learning loop of the cased-based reasoning. The overall enhanced model of case-based reasoning with knowledge discovery back-end, as shown in Figure 10, illustrates the dynamics of case manipulation, analysis, mining, knowledge formulation and case update. The visual "symmetry" in Figure 10 reflects in some sense the mutual benefit from the amalgamation of these computing approaches. On the one hand, multimedia data mining has the potential to improve the case-based system. On the other hand, the case-based paradigm provides mechanism for incorporation and management of discovered knowledge. On the indexing side potential advantages in the extended case-based reasoning model include:

- generation of term indexing schemes, based on the words used in the text representation [11];
- generation of term indexing schemes, based on relating terms to regularities discovered in other media types;
- generation of alternative indexing schemes (for example, a graph structure indexing scheme), based on structural patterns discovered in graphics, image, audio and video media in the case
- association of multiple indexing schemes to one case library;
- dynamic generation of the above listed indexing schemes.

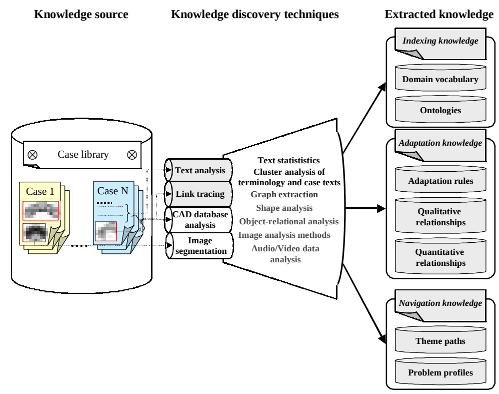



**Figure 9. The process for multimedia data mining in case libraries**

On the retrieval side the major advantage comes from the terminological flexibility in formulating queries due to the ontology-guided semantic transformation of the initial query and its match against the ontology-based indexing scheme [7].

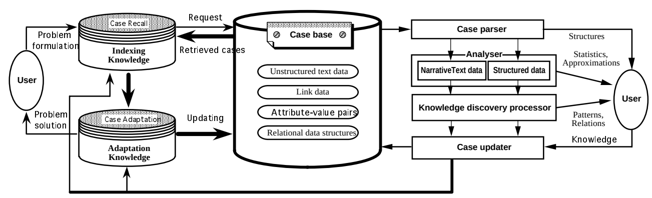



**Figure 10. Enhanced case-based reasoning model based on the symbiosis between MDM and case-based reasoning.**

The advantage for the adaptation in the enhanced model is the possibility to modify the new case description based on the information discovered from the multimedia analysis of retrieved cases in the context of the requirements.

## 3. MULTIMEDIA DATA MINING IN CSCW[2] ENVIRONMENTS

There are numerous approaches and techniques for setting up a computer-mediated environment for collaborative work [8]. The most common approach is to extend the desktop environment to include tools for meeting and sharing files. This approach takes the individual work environment and adds tools for communicating with others. An alternative approach is to create a virtual world environment in which the collaborators meet, work, and organise their projects. This approach differs conceptually because it creates a sense of place that is unique to the project, sort of a shared office space. A variation on this approach is to create a virtual world that is the model of the product or system being designed or developed.

The major feature of this kind of collaborative environment is the development of the project within the collaborative, multi-user

[1] Thematic path is a set of multimedia pages, each of which is part of the case library, that are relevant to the explanation and illustration of particular concept and have to be visited in particular sequence.

[2] Computer Supported Collaborative Work

environment. Project participants can work alone or collaboratively building the model and discussing the product as they view the model. There is only one representation of the model so there isn't a problem with simultaneous changes to different versions. There is a continuum of the process – a person does not shift environments when designing alone or collaboratively, and there is a continuum of the workspace during the design session - all working information about the product is accessed and shared through the same environment. An example of a project scenario in such environment based on Active Worlds, Inc. virtual world support is shown in Figure 11.

Such environment is a repository of multimedia data. Multimedia data mining in such environment can be used to enhance the functionality of computer support to project participants. The data includes 3D geometry of the product, data about allocation and behavior of participants in such environments, web multimedia data used in project documentation, presentation of ideas in collaborative sessions, communication transcripts, audio and video records. The example, from on-line analysis of bulletin board records, illustrates the potential for multimedia data mining and support in this field.

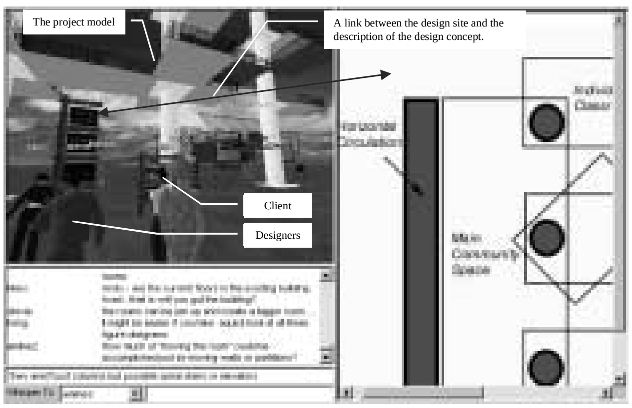



**Figure 11. An example of a collaborative project in 3D/2D virtual world**

Transcripts from online sessions, audio and video files can be used in the CSCW research or in education to form part of the student's assessment by including the amount and content of the student's participation. Text-based virtual worlds provide in explicit form a descriptive record of all activities inside the world. 3D virtual worlds provide transcripts from synchronous communications. Personal contribution to a collaborative session can be evaluated using text analysis of seminar transcripts [13] and multimedia analysis of related web pages.

Multimedia bulletin boards preserve the threads and the content of each message. Thus, the analysis of these data sets can be used to evaluate team collaboration. Below is an example of using a visualisation technique, which can provide quick feedback for monitoring collaborative projects.

Figure 12 presents a fragment from a team bulletin board. The messages on the board are grouped in threads.

A threefold split of the thread structure of e-mail messages in discussion archives in order to explore the interactive threads was proposed in [1, 9]. It included (i) reference-depth: how many references were found in a sequence before this message; (ii) reference-width: how many references were found, which referred to this message; and (iii) reference-height: how many references were found in a sequence after this message. The threefold split was extended in [4] to include the time variable explicitly. This model, expressed graphically as tree, allows the comparison of the structure of discussion threads both in a static mode (for example, their length and width at corresponding levels) and in a dynamic mode (for example, detecting moments of time when one thread dominates another in multi-thread discussions).

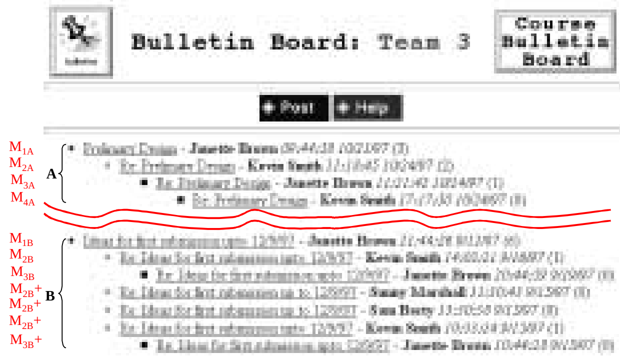
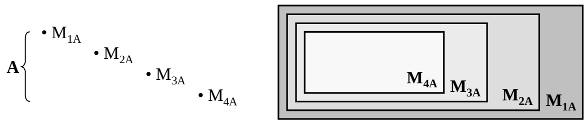


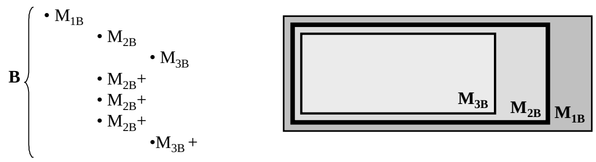

**Figure 12. Fragments from an asynchronous communication in a virtual world bulletin board.**

Visualisation techniques based on this model are modified versions of the nested set visualisation of tree structures [5]. Figure 13 shows an example of such visualisation applied to threads "A" and "B" from Figure 12. Each first message in a level is represented by a corresponding rectangle, labeled in this example to illustrate the message correspondence. Thus, there are four nested rectangles in Figure 13a. When messages are at the same level the thickness of the line is estimated based on the content-analysis of the message, including the text, included graphics and images. Each of the relevant messages on the same level is represented as additional 0.5 pt to the baseline thickness. In Figure 13b the base line thickness is 1 pt, thus rectangle "$M_{2B}$" has thickness 2.5 pt.

Figure 14 illustrates the application of the technique for monitoring collaborative design teams. Collaboration on a shared design and development task can be considered at different levels of abstraction and "degrees" of task sharing. Two extreme approaches to sharing design tasks during collaboration are identified in [6]: single task collaboration and multiple task collaboration. During single task collaboration the resultant design (or project development) is a product of a continued attempt to construct and maintain a shared conception of the design task. In other words each of the participants has his/her own view over the whole design problem and the shared conception is developed during intensive discussions. An example, of the visual pattern of such type of collaboration is presented in Figure 14b. It is characterised with relatively large amount of nested rectangles, usually indicating also several messages in respond to particular message. During multiple task collaboration the design problem is divided among the participants so that each person is responsible for a particular portion of the design. Thus, multiple task collaborative design does not necessarily require the creation of a single shared design conception, thus messages are usually related to the project management. Isolated messages and short threads dominate this collaboration style, as illustrated in Figure 14a.

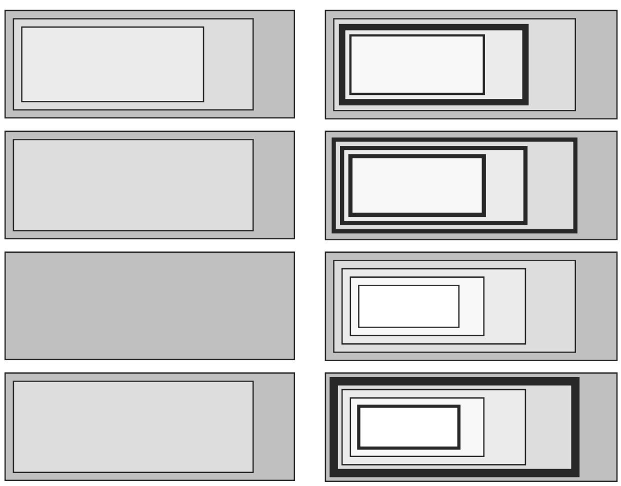



a. Nested rectangles for single message per level.




b. Nested rectangles when there are multiple messages on some levels.

**Figure 13. Visualisation of discussion threads.**



a. collaboration connected more with coordinating project tasks and submissions

b. intensive collaboration for creating a joint understanding of the problem

**Figure 14. Patterns of collaboration.**

Such visualisation techniques, combined with multimedia analysis of (i) video sequences of communications, (ii) elements of the 3D models in the scenery, and (iii) 2D representation of the project media, are expected to be part of the next generation CSCW environments.

## 4. EPILOGUE

So is multimedia data mining just a new combination of buzz-words or is it an exciting new area, capable to produce new methodologies and influence related interdisciplinary fields. The answer is left to the reader.

## 5. ACKNOWLEDGMENTS

Author would like to thank the Australian Research Council for the support of the related projects.

## 6. REFERENCES

[1] Berthold, M. R., F. Sudweeks, S. Newton and R. Coyne "Clustering on the Net: Applying an Autoassociative Neural Network to Computer-Mediated Discussions," *Journal of Computer Mediated Communication*, 2 (4), http://www.ascusc.org/jcmc/vol2/issue4/bert-hold.html (1997).

[2] Fluckiger, F. Understanding Networked Multimedia: Applications and Technology, Prentice Hall, London (1995).

[3] Ip, H. H. S. and A. W. M. Smeulders, eds., *Multimedia Information Analysis and Retrieval*, Springer, Heidelberg, (1998).

[4] Jones, S. ed., *Doing Internet Research*, Sage Publications, Thousand Oaks, CA, 29-55 (1999).

[5] Knuth, D E., *The art of computer programming, Vol 1: Fundamental algorithms*, Addison-Wesley, Reading, MA, 311-312 (1973).

[6] Maher, M. L., S. J. Simoff and A. Cicognani, "Potentials and limitations of Virtual Design Studio," *Interactive Construction On-line*, January, a1 (1997).

[7] Maher, M. L. and S. J. Simoff, "Knowledge discovery from multimedia case libraries", in Smith, I., ed., *Artificial Intelligence in Structural Engineering*, Springer, Berlin, 197-213 (1998).

[8] Maher, M. L., S. J. Simoff and A. Cicognani, *Understanding Virtual Design Studios*, Springer, Heidelberg (2000).

[9] Sudweeks, F., M. McLaughlin and S. Rafaeli, eds, *Network and Netplay: Virtual Groups on the Internet*, AAAI/MIT Press, Menlo Park, CA, 191-220 (1998).

[10] Simoff, S. J. and M. L. Maher, "Ontology-based multimedia data mining for design information retrieval", *Proceedings of the ACSE Computing Congress*, Cambridge, MA, 310 - 320 (1998).

[11] Simoff, S. J. and M. L. Maher, "Deriving ontology from design cases", *International Journal of Design Computing*, 1, http://www.arch.usyd.edu.au/kcdc/journal/vol1 (1998).

[12] Simoff, S. J. and M. L. Maher, "Knowledge Discovery in Hypermedia Case Libraries - A Methodological Framework," *Proceedings of the Knowledge Acquisition Workshop, 12th Australian Joint Conference on Artificial Intelligence, AI'99*, 213-225 (1999).

[13] Simoff, S. J. and Maher, M. L. "Analysing Participation in Collaborative Design Environments," *Design Studies*, 21, 119-144 (2000).

[14] Zaïane, O. R., J. Han, Z.-N. Li, S. H. Chee, S. H. and J. Y. Chiang. "MultiMediaMiner: A system Prototype for MultiMedia Data Mining, *Proceedings of ACM SIGMOD International Conference on Management of Data*, 581 - 583 (1998).

[15] Zaïane, O. R., J. Han, Z.-N. Li, J. Hou, "Mining Multimedia Data" *Proceedings CASCON'98: Meeting of Minds*, Toronto, Canada, 83-96 (1998).