

PETER MÜLLER AND INGMAR DECKER

The Electronic Visualization Library

The Electronic Visualization Library

Peter Müller and Ingmar Decker

March 2, 1998

Abstract

Scientific visualization is a rapidly growing research field with a need for information dissemination. This report describes the Electronic Visualization Library (EVlib), a digital library for scientific visualization. Main design goals of EVlib were an attractive user interface providing various kind of access mechanisms as well as an open interface to other already established information systems.

EVlib stores fulltext versions of documents where they are available. We also provide access to BibTeX entries for every stored document. All available BibTeX entries are combined in BibTeX bibliographies which are registered with the “Collection of Computer Science Bibliographies” at University of Karlsruhe.

Additionally, we have defined a mapping from BibTeX attributes to the Dublin Core attribute set. This mapping is used to provide a gatherer interface for the Harvest information system. This way, existing Harvest installations can immediately use EVlib as an information resource.

1 Introduction

The enormous success of the World-Wide Web and its possibilities to disseminate and access information have led to several initiatives dealing with the development of electronic or digital libraries (DL).

Digital libraries provide a broader access to a broader variety of (digitally represented) objects. They also help users to cope with the so-called “information overload”, mainly by providing enhanced query facilities and graphical user interfaces.

Among other criterias, one can distinguish general DLs, cf. [25], from DLs for specific application domains, cf. [10],[7], each of them having specific advantages.

The *Electronic Visualization Library* (EVlib) is a domain specific DL for scientific visualization and computer graphics. Its development is part of a project called “Information Services for Scientific Visualization”, which is organized by the “Verein zur Förderung eines Deutschen Forschungsnetzes e. V.” (DFN) and funded by the German Ministry of Research and Development (BMBF). The main motivations for establishing an information service and a DL for scientific visualization are:

- Scientific visualization is a rapidly growing research field. Therefore there exists an equally growing need for information in this subject, especially about newly available publications, tools and techniques.

- Information about this research area is spread widely over a multitude of diverse (Internet) resources, such as different libraries, bibliographies, and services, all hosted on different systems and sites. General search engines for the WWW have not proved to be effective enough. Due to their non-specific, general operation, they return huge result sets at low precision. With a domain specific information service it is possible to offer available information in one place with a uniform user interface.
- Currently, no information service dedicated to scientific visualization exists.

The planning and realization of a scientific visualization DL has to address diverse aspects and problems. The main aspects are: user profile and requirements, user interaction, DL content, information acquisition, object representation, storage, archiving, access, and the interconnection to other services. This paper describes and discusses how these problems have been addressed by EVlib.

In Section 2 we briefly explain main requirements that we consider important for EVlib. Section 3 discusses related work, i.e. other domain specific DLs. In Sections 4 and 5 we describe the structure and implementation details of the library. Section 6 deals with problems of interoperability, and Section 7 presents plans for future work. In Section 8 we summarize this paper and draw some conclusions.

2 Concepts and Requirements

DLs are commonly viewed as “systems providing a community of users with coherent access to a large, organized repository of information and knowledge” [21]. The “community of users” is an important point which is sometimes neglected in other definitions, cf. [25, 3].

In this section we explain basic concepts and requirements for DLs in general and for EVlib in particular. First, we describe user and user interaction requirements, because to a certain extent these determine other technical or system requirements. It is important to think about what a DL *should* provide, the next step is to look for what it *can* do.

2.1 User Community and Database Contents

EVlib is designed for a “community” of people who deal with scientific or data visualization.

Scientific and data visualization is a new and rapidly growing research and application area. There are increasing possibilities for this task provided, even on every personal computer and, additionally, there seems to be a trend towards visualization of diverse kinds of information [16],[17]. Today, scientific visualization techniques are used more and more in many application areas, like medicine, mathematics, or economy.

Thus, there is a need for *up-to-date information* about new tools and techniques for both scientists of scientific visualization and scientists and users of those other application areas.

One of the major information source in science is literature: books, articles and technical reports of research institutions. Information about the first is

widely spread over diverse libraries, information systems or Web sites. The same holds for the latter, which – in addition – are often “hidden” somewhere on FTP or Web sites. Thus, there is a need for making this literature available at one place.

2.2 User Interaction and Document Access

People are used to traditional or “physical” libraries. The conception of a DL has to consider this fact and has to provide what is liked in those traditional ones. Of course, DLs are expected to enhance and complement these possibilities.

A DL using *WWW technology* can support users in finding and accessing various kinds of relevant information *worldwide*. A user interface with diverse interaction techniques and services can be provided using standard HTML elements like HTML forms. Using well-known and well-tested indexing techniques and structured document representations, it is clear that a DL is expected to provide sophisticated search facilities based on keywords with regard to various types of bibliographic information.

Nevertheless, it is also well-known that searching techniques based on keywords have certain disadvantages. Sometimes, keywords do not represent the contents of a document accurately enough. This is especially true for systems where keywords are automatically extracted. A further problem is misspelled, incomplete or incorrect keywords. For these reasons it seems to be useful to provide other query and access techniques. One example is *browsing*. Users should have the ability to browse through a list of authors or titles, similar to browsing in a catalog or glancing over the books in the shelves of a traditional library.

Furthermore, scientific visualization literature is naturally filled with images and figures which usually exemplify the content of the respective document. A DL can use automatic image extraction mechanisms to provide access to the documents by browsing through an image database (see Section 5.3).

Another important point is the access of documents as *full text* versions, i. e. not only as references. User studies have shown that “Digital library resources that offer full text are perceived far more favourably than the more usual bibliographic data sources” [5]. As far as possible, EVlib offers full text versions of documents (PostScript or PDF).

2.3 Information Acquisition

Acquisition of information for EVlib has to aspire the document collection to meet the following criteria: *topicality*, *completeness*, and *quality*. Since EVlib is conceived to run mainly as an automatic service, a human librarian cannot be employed to guarantee those criteria. On the other hand, full automatic gathering of information with help of robots or spiders would require that the release of newly available information is known (to gain *topicality*), that these tools “know” where to find information (to gain *completeness*), that they are able to select only relevant documents and to generate correct metadata (to gain *quality*).

This is currently impossible. Rather, the acquisition must be based on user participation to meet the ongoing growth of available online information and

the need to provide an up-to-date information service. I. e., users of a domain specific DL contribute available information.

This approach should be facilitated by scientists own interest to make public their research to a broad audience. Regarding the three requirements on the document database, the advantages are:

Topicality. If every user contributes his/her information as soon as it is available, the document database would keep an up-to-date state of available information.

Completeness. If every user contributes his/her information, the database can achieve a high degree of completeness.

Quality. If the contribution of new information is restricted to those users who work in the respective research domain, only relevant information would enter the document database.

Topicality and completeness require a highly motivated user community. As a consequence, a DL has to offer services and features attractive for its users. A problem regarding quality is of course the fact, that the restriction of allowing only users of the respective research domain to contribute information is actually a virtual one: it relies on the users' good will. Another aspect is that the user is required to provide additional bibliographic information each time a document is contributed.

Therefore, a DL should provide utilities which support users in contributing high-quality information, and which guarantee a certain degree of seriousness. For example, every contribution should clearly identify its contributor.

2.4 Document Representation and Database Management

The document database management of a DL must cope with representation, storage, and archiving of large collections of documents. Additionally, there exist more general requirements like scaling and interoperability, cf. [21].

Since digital or online documents may co-exist in different representations (e. g. PostScript and HTML), there is a need to find a document representation which is useful in the library's environment. One possibility is to define a common document format to which every contributed document is converted to, e. g. SGML. Usually, such formats provide structural information including titles, authors, chapters or paragraphs, which can then be used for retrieval purposes. However, using one format introduces conversion problems, because many available document formats (like PostScript) do not provide structural information.

A second possibility is to associate an additional *metadata representation* to each document. This format integrates information common to each variant of a document and is exclusively used for retrieval. These representations are thus independent of a particular document variant, and it is easy to introduce new document formats (e. g. it is easy to add also PDF to the list of available document formats).

New document databases should be open to integrate as well as to be integrated in existing services. This implies the use of existing and widely-used formats, standards and protocols and the provision of interfaces to other systems. This would disseminate provided information widely, leading to the problem of

long term availability of digital information. As a consequence, it is necessary to find solutions for long time storage of large document collections.

3 Related Work

Many different projects and sites providing information about publications exist. These projects range from bibliographic data services, including services for the research area of scientific visualization or computer graphics, to more sophisticated and distributed services which provide more information and more interactivity.

An example for a bibliographic data service is the “ACM SIGGRAPH online bibliography”, a collection of computer graphics bibliography citations¹. This bibliography contains about 14.000 references in BibTeX format. The references can be searched using keywords contained in certain BibTeX fields. The query formulation facilities are rather rudimentary. Only some references contain abstracts and there is access to neither full-text nor PostScript sources.

A more sophisticated service is the well-known and established “e-Print archive” at Los Alamos National Laboratory². This service provides documents and preprints mainly for the domain of high-energy physics and mathematics. Document acquisition is based on contributions of users. Contributors can submit documents via WWW or e-mail. L^AT_EX and HTML formats are used because of openness and stability. Besides usual query facilities based on keywording and lists ordered by author, title or year, one can get an overview of the newest contributions. PostScript, PDF or DVI versions of the documents are usually available. Unfortunately, there is no such archive for scientific visualization or computer graphics.

The “Networked Computer Science Technical Reports Library” (NCSTRL) “is an international collection of computer science technical reports from CS [Computer Science] departments”³ and thus a distributed bibliographic information service. Document acquisition is provided by each participating institution. A simple query can be run for the whole distributed repository, but querying for a specific aspect a user has to know the collections in which he/she has to search. Documents are provided in different versions, depending on the participating institution.

NCSTRL is based on Dienst [20], a service specification for digital libraries. Dienst provides features which lead us to use it as a basis for EVlib. The following section describe our use of Dienst.

4 Using Dienst

Dienst consists of three components: the *document database* which stores the documents, the *Dienst server* which provides database management facilities, and the *Dienst CGI stub* which provides an interface to the Dienst server via WWW. We have chosen Dienst as a software platform for the following reasons:

¹<http://www.siggraph.org/publications/bibliography/bibliography.html>

²<http://xxx.lanl.gov>

³<http://www.ncstrl.org>

- The Dienst specification is based on a document model suitable for digital documents. It already provides support for various formats. For example, a document might exist in a PostScript as well as in a HTML version.
- Documents can be viewed in various ways depending on available formats. It is possible to get an overview of thumbnail pages or to directly access available online formats.
- The Dienst distribution already provides a WWW interface framework to search and browse the document collection in various ways. For example, it is possible to search the whole collection by keywords or to use a search form to search in specific bibliographic fields (for example, only in the author field). Browsing is possible by author name and publication year.

Dienst uses *document representations* as a metadata scheme for documents. These representations hold bibliographic and formal attributes as provided by the respective contributor of the document. They are used for retrieval as well as administrative purposes. The original Dienst specification does not exactly fulfill the requirements of a fully automatic publication service. Therefore we have made several extensions and modifications which are discussed in the following section.

5 Dienst Extensions and Modifications

Originally, Dienst uses RFC 1357 [9] as its document representation format. We have changed this to BibTeX [15], because BibTeX offers the following advantages:

- BibTeX already provides a well-defined set of bibliographic attributes which allows to give complete references. It is also possible to define new attributes which are silently ignored by BibTeX parsers but which are recognized by our service.
- BibTeX is widely in use in the scientific community due to its tight relationship to L^AT_EX. Thus, if a user requests the BibTeX description of a document he/she can directly use it in his/her own publications for further references.
- Individual BibTeX representations can be collected into one file, forming a *BibTeX database*. Information services based on such databases already exist (eg. [1, 11]). This simplifies integration of such services as described in Section 6.

Another important advantage in using BibTeX is that attribute semantics are well-known in the scientific community. Many scientists are also familiar with the syntax of specific attributes (e. g. for specifying multiple authors).

We added a few attributes to the original set of BibTeX attributes which are necessary for retrieval and administration purposes. Table 1 lists the most important additions.

The BibTeX fields are filled once a contributor submits a new document. The original Dienst distribution does not support fully automatic document submission. Instead, a Dienst representative is notified each time a new document is

Attribute	Description
abstract	A short abstract of the content of the document. This attribute is used for information retrieval purposes.
keywords	A list of keywords. In journals this is often designated as indexing terms.
language	The language in which the document is written.
postscript-url	URL to a PostScript version of the document.
pdf-url	URL to a PDF version of the document.
url	URL to an electronic version of the document in neither PostScript nor PDF format.

Table 1: Additions to BibTeX attributes.

contributed. Since EVlib is supposed to be a fully automatic publication service, we have modified Dienst such that no human intervention in the contribution process is necessary. This implies the following additions:

- *User administration.* In order to gain a certain level of responsibility, contributors have to register with the publication service (cf. Section 5.1). Contributors are linked to a document, hence errors can be immediately reported to the responsible person.
- *Submission process.* We completely changed the submission process of the original Dienst distribution to a fully automated process. This requires to move some services to external tools (cf. Section 5.2).

5.1 User Administration

One important aspect of allowing everyone to contribute documents is to ensure some level of responsibility. It is necessary that contributors are able to manage and administer their contributions. On the other hand, users must have the ability to identify a responsible person for a given contribution, for example to point out writing errors. For that reason we have added a user administration service to EVlib by extending some of Dienst's functionality.

EVlib is just one of many information services useful for scientific visualization one can think of. We are also working on several other services including an event manager and a cooperative information notification service. Each of these services also require to manage some user specific data.

As a consequence, we have separated core user administration functionality from service specific administration functionality. Core user administration functionality is now achieved by a separate service which is transparently integrated in the other services. This allows users to register for all services with one single service.

5.2 Document Submission Process

The submission process is divided into two parts: *information contribution* and *postprocessing*.

In order to submit a new document to EVlib, it is necessary that the contributor provides some information about the contribution. The first step is to choose a document *type* and *category*. Document types are inherited from BibTeX. We assume that scientists know these types (again, due to the extensive use of BibTeX in scientific research).

The document category is a rough classification of the document. We currently provide nine categories: *surface rendering*, *volume rendering*, *vector field visualization*, *information visualization*, *visualization systems*, *computer graphics*, *computational geometry*, *image processing* and *miscellaneous*. We use these categories to provide some structure. Additionally, the contributor is able to designate a \LaTeX file which is used to extract information about the document.

In a second step, the contributor has to fill-out an attribute-based form. Attributes depend on the chosen document type and follow the extended BibTeX specification. The latter implies that we also offer the possibility to fill out only those attributes which are required to obtain a valid BibTeX record, although this would restrict the possible retrieval quality. We therefore recommend to provide additional bibliographic information like abstract or keywords.

If the user has designated a \LaTeX file in the first step, the extracted information is automatically placed in respective fields of the document submission fill-out form. This simplifies the contribution process, because the user now only needs to modify these fields without the need to extract all the information by hand.

In a third step this information is preprocessed to

1. *prevent duplicates*. A simple algorithm is used to check for possible duplicates of the new document within the already existing pool of documents. We use the title of the document to create a checksum which is then compared with those generated for other documents. If two checksums are equal, the contributor is notified that a possible duplicate exists.

We allow the contributor to decide whether to proceed with the contribution process or not. It sometimes makes sense to allow duplicates, for example if a technical report eventually gets published. Whereas the technical report might still be electronically available, this usually is not the case with the published version.

2. *check spelling*. Abstract and title are passed through a spell checking program. Misspelled or unknown words are then marked. This should improve the quality of the submitted information.

Finally, the contributor confirms that the given information should be entered. At this point the respective BibTeX record is generated and spooled in an incoming directory. This allows to provide faster feedback to the contributor, because every further processing is done offline.

We also generate a *document identifier* (DID). It is used to uniquely identify a document within EVlib. A DID consists of three parts library abbreviation, year and running number connected with dashes: EVL-*yyyy-nnn*. Every BibTeX record uses the DID as key value.

Periodically, a “publishing process” is executed which looks for new contributions and which eventually stores them in their final location within EVlib’s document database.

This publishing process is responsible to download specified fulltext versions of the respective document (if available) and to generate additional information for advanced access and browse methods. For example, if a PostScript version is available, images are extracted and entered into an image database (cf. Section 5.3).

Finally, the indexes of EVlib are updated to include the new contribution. Typically, new contributions are fully integrated into EVlib within a few minutes.

5.3 Query by Images

As an additional retrieval and browse facility we started to provide a query-by-image service. We are going to explore the use of *content-based image retrieval* (cf. [13, 19]) techniques as a document access technique. Since one can often remember pictures of a document better than authors or titles, we think that a search for pictures of a desired document might be of great benefit.

So far we have implemented a prototype interface for browsing through the pictures of the documents in the EVlib document database. The pictures are extracted automatically from the PostScript files.

Traditional image retrieval techniques are based on text or keywords assigned to the pictures in a database. In many cases, this keywording is less useful or not practicable. Due to the automatic process of contributing documents to EVlib, there is no textual annotation to the extracted pictures at all.

Thus one has to try to retrieve pictures based on their visual contents. So far it is not generally possible to extract semantic meaning from pictures, so the query techniques have to rely on the analysis and the comparison of certain visual features of the pictures. Features which are often used are color [14], color areas and composition [26, 4, 18], and texture [23].

So far, the prototype interface and retrieval component only provides queries based on similar color distribution (similar to [14]) and composition (similar to [18]), and simple browsing. A user can click on images which are similar to the desired one (see Figure 1).

These are quite weak search mechanisms, and the facilities for retrieving pictures have certainly to be improved. Nevertheless, the possibility of browsing through pictures has already proved to be effective for getting an overview of the database contents. Sometimes it lead users to interesting documents, because they recognized pictures of interest.

We think that this component will enhance the query facilities of EVlib, especially due to the visually oriented context.

6 Interoperability and Openness

An important aspect of todays information systems is their ability to interact with other systems. Especially new systems should be open for interoperability with already existing and widely accepted ones.

With BibTeX we are using a document representation format which is immediately usable by scientists in their own publications. Additionally, each time a new document enters EVlib's document database its BibTeX record is automatically appended to a category specific BibTeX file. As a result, nine BibTeX files

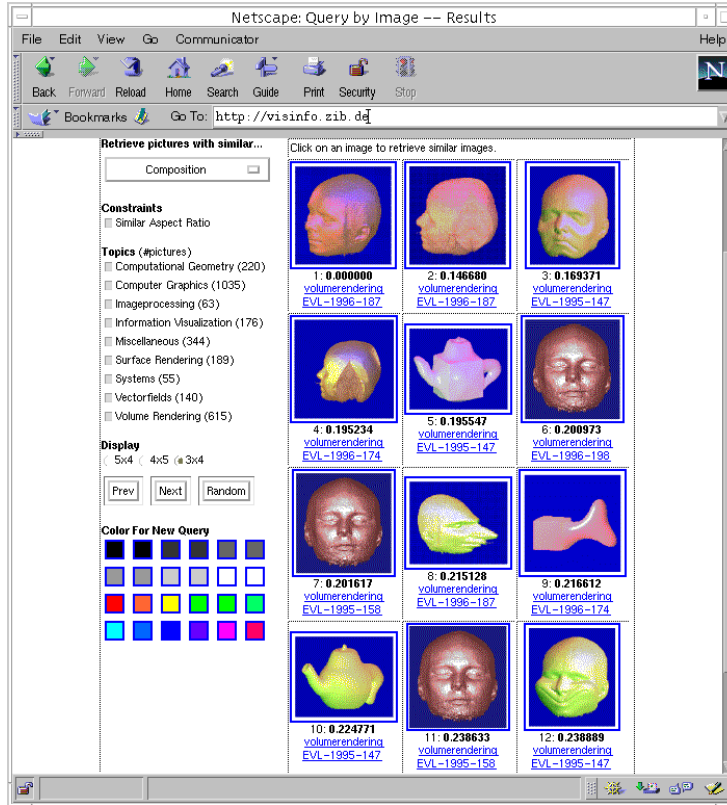


Figure 1: *Image browser interface prototype.* One can click on an image to retrieve similar ones according to the chosen similarity measure (upper left corner). Additionally, below each image is a link to the respective document where the image was extracted from.

exist, each representing the complete set of available references in the respective category.

We have registered these files with “The Collection of Computer Science Bibliographies” [1] located at University of Karlsruhe, Germany. Within this collection the files are combined to form the “Electronic Visualization Library Catalog.” As a side-effect the catalog is searchable and integrated in an attractive service.

Another attractive information service is Harvest [8] which is used in many domain specific information projects [10, 6, 24]. Harvest uses the SOIF format to represent information objects. This format is also used by glimpse [22] to allow attribute based queries. Since glimpse is the indexer back-end of EVlib it is easy to set up a Harvest gatherer process which simply provides access to the already existing SOIF database. This allows every provider of Harvest brokers to register with our EVlib gatherer to integrate our information resource.

However, to actually offer a usable Harvest service it is necessary to agree upon SOIF attribute names and semantics. It does not make sense to reinvent another EVlib specific set of attributes. Instead it would be better to use a standardized set of attributes. Dublin Core [29], currently developed, provides

such a standardized set of attributes. The main idea of Dublin Core is to provide a very simple and, thus, easy-to-use attribute set for electronic information resources. Although Dublin Core is not yet officially standardized it is already used in many projects⁴.

We have defined a mapping from BibTeX to Dublin Core attributes and provide the latter in our SOIF records. We believe that in the near future more and more Harvest installations will be based on Dublin Core. These installations can then integrate our service immediately.

7 Future Work

The current EVlib implementation does not allow a contributor to withdraw a contribution. In addition, currently it is not possible to modify a contribution once it is entered into the document database. One of our next steps is to include this functionality into EVlib.

We also want to add another contribution interface based on electronic mail. Users should have the possibility to submit complete BibTeX bibliographies without the need to manually fill out HTML forms. We assume that the service would benefit from such an interface, because many institutions already have a publication process which could be easily extended to send out an electronic mail with a respective BibTeX record.

The prototype query-by-image interface needs to be extended. We would like to add better image retrieval techniques, for example techniques based on image structure. Additionally, allowing users to draw an initial query image or to let them provide a URL of a query image is a future prospect.

An important question regarding digital information resources is their long-term availability. This is a question of archiving possibilities. At Konrad-Zuse-Zentrum we have access to a storage system which would allow us to provide long-term availability. However, access to this storage system is rather slow. Consequently, we have to think about a storage and archiving concept. An idea is to automatically migrate contributions to slower storage devices dependent on their access rate.

Another idea we would like to implement is the provision of EVlib volumes. Volumes would combine all contributions of a given year and could be saved on CD-ROM.

8 Conclusions

With EVlib we have established the first digital library dedicated to scientific visualization.

The choice to use Dienst as software platform allows to use certain features like searching and browsing as well as generation of thumbnail overview pages immediately. The modification of Dienst to use BibTeX as document representation has offered the possibility to use bibliography services. Due to its meaning in conjunction with L^AT_EX, authors can use these BibTeX representations in their own work. With the definition of a mapping from BibTeX attributes to the Dublin Core attribute set it was possible to open EVlib for other services.

⁴See http://purl.org/metadata/dublin_core/projects.html for a project list.

For example, we have set up a Harvest gatherer process which enables any Harvest installation to include our information offer as a new information resource. We also use several other publicly available tools to improve EVlib's service, for example to add spell and BibTeX format checking.

EVlib has been developed using a combination of existing tools and utilities based on a document retrieval software (Dienst). Major changes were only necessary to automate the submission process which introduces the need for user administration.

EVlib is online since September 1997. Although only a few contributors have registered, about 800 documents and references of scientific visualization literature have been submitted at a growing number. We believe that this service will be of great benefit for the visualization community.

9 Acknowledgements

The authors wish to thank H.-C. Hege and M. Zöckler for fruitful discussions regarding useful features of a digital library for scientific visualization.

References

- [1] Alf-Christian Achilles. The Collection of Computer Science Bibliographies⁵. Information service. Last visited Nov. 25, 1997.
- [2] ACM. *Computer Graphics Proceedings, Proceedings of SIGGRAPH 95*, 1995.
- [3] Nabil R. Adam and Yelena Yesha. Introduction. *International Journal on Digital Libraries*, 1(1):1–2, April 1997.
- [4] Jonathan Ashley, Ron Barber, Myron Flickner, et al. Automatic and semi-automatic methods for image annotation and retrieval in qbic. In [27], pages 24–35, 1995.
- [5] Christine Barry. The Digital Library – The Needs of Our Users⁶. Paper presented at the International Summer School on the Digital Library, 5th August 1996, Tilburg University, Netherlands. WWW document, last visited 18.11.1997.
- [6] M. Bischoff. Vernetzte Broker für fachliche Dokumente (distributed brokers for research documents)⁷. Project description. Last visited Jan. 8, 1998.
- [7] Dietrich Boles, Markus Dreger, and Kai Großjohann. Medoc information broker – harnessing the information in literature and full text databases. In *Proceedings NIR Workshop, SIGIR96*, October 1996.
- [8] C. Mic Bowman, Peter B. Danzig, Darren R. Hardy, Udi Manber, and Michael F. Schwartz. Harvest: A scalable, customizable discovery and access system. Technical Report CU-CS-732-94, Department of Computer Science, University of Colorado – Boulder, August 1994.

⁵<http://liinwww.ira.uka.de/bibliography/>

⁶<http://www.kcl.ac.uk/kis/schools/education/cbticer.html>

⁷<http://www.eprint.de/>

- [9] Danny Cohen (ed.). A Format for E-mailing Bibliographic Records. Request for Comments (RFC) 1357, Jul. 1992.
- [10] Wolfgang Dalitz, Martin Grötschel, and Joachim Lügger. Information Services for Mathematics in the Internet (Math-Net). Preprint SC-97-25, Konrad-Zuse-Zentrum für Informationstechnik Berlin (ZIB), 1997. To be published in Proceedings of the 15th IMACS World Congress 1997.
- [11] Stefano Foresti, Nelsen H. F. Beebe, and Eric Grosse. BibNet – A Bibliography Network Project⁸. Information service. Last visited Nov. 25, 1997.
- [12] William Grosky, Ramesh Jain, and X. Mehrotra, editors. *Handbook of Multimedia Information Management*. Prentice Hall, 1997.
- [13] Amarnath Gupta and Ramesh Jain. Visual information retrieval. *Communications of the ACM*, 40(5):71–79, May 1997.
- [14] James Hafner, Harpreet S. Sawhney, Will Equitz, Myron Flickner, and Wayne Niblack. Efficient color histogram indexing for quadratic form distance functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7):729–736, July 1995.
- [15] Nicholas J. Higham. BibTeX: A versatile tool for L^AT_EX users. *SIAM News*, 27(1):10, 11, 19, January 1994.
- [16] IEEE. *Proceedings Information Visualization'96*, 1996.
- [17] IEEE. *Proceedings Information Visualization'97*, 1997.
- [18] Charles E. Jacobs, Adam Finkelstein, and David H. Salesin. Fast multiresolution image querying. In [2], pages 277–286, 1995.
- [19] H. V. Jagadish. Content-based indexing and retrieval. In [12], pages 69–93. 1997.
- [20] Carl Lagoze, Erin Shaw, James R. Davis, and Dean B. Krafft. Dienst: Implementation Reference Manual. Technical Report TR95-1514, Computer Science Department, Cornell University, May 1995.
- [21] Clifford Lynch and Hector Garcia-Molina. Interoperability, scaling, and the digital libraries research agenda, August 1995. A Report on the May 18-19, 1995, IITA Digital Librarys Workshop.
- [22] Udi Manber and Sun Wu. GLIMPSE: A Tool to Search Through Entire File Systems. Technical Report TR 93-34, Department of Computer Science, University of Arizona, October 1993.
- [23] A. Pentland, R. W. Picard, and S. Scarloff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996.
- [24] PhysDep: Electronic information management and metadata in physics⁹. WWW document. Last visited: Jan. 9, 1998.

⁸<http://www.math.utah.edu/ftp/pub/bibnet/faq.html>

⁹<http://www.physik.uni-oldenburg.de/EPS/EurophysNet/PhysDep/dep-links.html>

- [25] Bruce Schatz and Hsinchun Chen. Building Large-Scale Digital Libraries. *IEEE Computer*, 29(5), May 1996.
- [26] J. R. Smith and S.-F. Chang. Tools and techniques for color image retrieval. In [28], pages 426–437, 1996. <ftp://ftp.ctr.columbia.edu/CTR-Research/advent/public/papers/96/smith96b.ps.gz>.
- [27] Society of Photo-Optical Instrumentation Engineers (SPIE). *Storage and Retrieval for Image and Video Databases III. Proceedings SPIE 2420*, February 1995.
- [28] Society of Photo-Optical Instrumentation Engineers (SPIE). *Storage and Retrieval for Still Image and Video Databases IV. Proceedings SPIE 2670*, February 1996.
- [29] Stuart Weibel, Renato Iannella, and Warwick Cathro. The 4th Dublin Core Metadata Workshop Report. *D-Lib Magazine*, June 1997.