

# **Design issues in the production of hyper-books and visual-books**

**Nadia Catenazzi\*, Monica Landoni\* and Forbes Gibb\*\***

*\* Environmental Informatics Unit, TP.440 - CEC Joint Research Centre, 21020 Ispra (VA), Italy*

*\*\* Department of Information Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH*

## **Abstract**

*This paper describes an ongoing research project in the area of electronic books. After a brief overview of the state of the art in this field, two new forms of electronic book are presented: hyper-books and visual-books. A flexible environment allows them to be produced in a semi-automatic way starting from different sources: electronic texts (as input for hyper-books) and paper books (as input for visual-books). The translation process is driven by the philosophy of preserving the book metaphor in order to guarantee that electronic information is presented in a familiar way. Another important feature of our research is that hyper-books and visual-books are conceived not as isolated objects but as entities within an electronic library, which inherits most of the features of a paper-based library but introduces a number of new properties resulting from its non-physical nature.*

## **Introduction**

The evolution of publishing tools from movable print to digital typesetting controlled by computer has been principally aimed at improving document quality and lowering unit costs rather than responding to readers' needs. And a by-product of these improvements is the production of enormous quantities of printed paper which are never read and which have limited justification in being defined as information.

Although the computer is largely responsible for this information glut, it can and should be turned to assist the reader by providing a consultation environment alongside the production system, where information can be presented in a more attractive and more functional way than on paper. To ensure this result, it is necessary to provide some conversion tools which are able to translate information acquired from different sources into an electronic environment,

and at the same time to offer a powerful and meaningful paradigm for information representation. The result is an object called the electronic book, a generic term which is becoming widely used, and which encompasses a large family of implementations where the features of a conventional book are mapped into an electronic system to various degrees. The route followed by the authors of this paper is to maintain as closely as possible the familiar metaphor of the book. The result is two complementary proposals, hyper-books and visual-books, both based on the same philosophy but with specific features related to the type of original documents they manage: electronic texts and paper books respectively.

## Electronic books

A working definition of an electronic book has been given by Barker (1991) as 'a collection of pages of electronic information which can be displayed on a suitable viewing screen.' However, this definition is very general and encompasses various interpretations. Our approach to electronic book design is strongly related to the book metaphor. Books are the traditional repositories of information. People know how to read them, how to use a Table of Contents, how to use an index, etc. By maintaining the same model on screen, people's access to electronic information can be facilitated. Instead of formulating complex queries which are used to retrieve books based on, for instance, a document surrogate such as an inverted file, a representation of the book itself is available which can be consulted like a physical book. This approach helps to overcome some of the limitations inherent in electronic support, such as the loss of the physical aspects of the book (for example its thickness) which can be used as an orientation clue. Nevertheless, there is another important group of difficulties which are connected to technological shortcomings, mainly the limited size and resolution of display screens, that makes reading an electronic book less pleasant than a paper book.

## Design issues

The communication of knowledge is a basic requirement for all types of books, and particularly for electronic books where methods of information and knowledge delivery are different from those used in conventional paper books. Essential to electronic book design is an understanding of user profiles and of the purpose of the book, which determine the choice of interaction strategies and types of interface.

It is also important to consider the design processes used for paper books as possible guidelines for electronic-book production, while recognizing that the different nature of electronic books makes it necessary to develop specific procedures. At least two different levels of analysis have to be considered, as suggested in Barker and Manji (1991): a basic one which deals with features related to hardware, software and interfaces; and a conceptual one which considers the user's point of view.

At the first level of analysis, the attention of the designer is focused on practical aspects of the system, such as the study of interface features and media involved in book representation, and the need for various design templates. *In conceptual design, models and paradigms for book representation are considered in order to provide an environment where the user can interact with the information in a natural way. Particular attention is given to the design of empty templates for books and pages which the author will fill in following a relatively simple process. The conceptual design phase also provides a justification for the choices made in the*

interface design. In this context, it is useful to distinguish between various categories of electronic books, because different books have different design requirements.

There are different ways of classifying electronic books according to different criteria. Barker (1991) defines a taxonomy of electronic books based on two criteria: the nature of information, and the types of services they are able to provide. Collis (1991) classifies electronic books as reference books and text books, a distinction based on different usage. Another important criterion for classifying electronic books is based on the kind of presentation offered. Each presentation is dependent on a particular interpretation of the concept of the electronic book. These categories are not necessarily exclusive.

There is a large group of electronic books, called portable electronic books, whose main purpose is to reproduce portability (Feldman 1990; Lande 1991). In fact, one of the most appealing features of paper books is the possibility of carrying them everywhere and reading them comfortably. The electronic equivalent has many drawbacks: limited screen size, limited resolution, etc. At the moment, portable electronic books have made their most significant impact in the area of reference publications such as dictionaries and thesauri. A further group of electronic books is more focussed on the logical structure rather than on the physical nature of the book: information is usually displayed in scrolling windows, which may overlap, and the idea of the page as a static block of text is usually not considered. This is the case with Dynatext (EBT, 1990) and SuperBook (Egan *et al*, 1989). Both of these provide full-text indexing, links, navigation and orientation through a dynamic table of contents, and a multi-window text display. An interesting aspect of these two systems is the fact that they provide the capability of automatically importing electronically available text.

Finally, there is a group of electronic books which emphasizes the physical aspects of the book but also considers its logical aspects. This, in our opinion, is the most complete and satisfactory solution because it considers logical structures together with physical clues (for example, pages, book thickness), both of which are indispensable for recognizing an object as a book. This approach becomes particularly important when a paper version exists of the same document and readers are familiar with it.

Benest (1987; 1990) presents interesting examples of this kind of electronic book; a book is presented as an open document in which pages may be turned and flicked through. The thickness of the book is indicated by an active black band along the side and the bottom of the book. The text page may be annotated and highlighted. Book markers also may be inserted.

*Vortex* (Burril, 1989) is another example of an electronic book in this category, but one which proposes different solutions. The most original idea here is concerned with how to solve the problem of providing orientation facilities within the book. In order to indicate the reader's position in the book, the solution adopted in *Vortex* is the idea of closed pages. The text is displayed between two sets of vertical lines which represent the thickness of pages before and after the current page. Section names are indicated between these lines and may be selected in order to go directly to that page.

#### Electronic books in the context of electronic libraries

The increasing interest in electronic books and the availability of practical examples of them indicates that using them could become a common practice. From this perspective, an obvious

and very useful application is in the context of an electronic library which could be seen both as a personal library (for organizing personal collections of electronic books), and as a public library (for delivering the full text of documents online). The library could be accessed remotely from any location through a network using personal or portable computers.

Electronic books do not suffer from the constraints of their physical counterparts, and can be easily replicated as many times as required. The library itself has infinite space: there is no limit to the number of books it can contain. Many of the operations normally performed in physical libraries do not exist in electronic libraries (binding and preservation, shelf maintenance etc.), although there may be new administrative functions. The computer gradually replaces the librarian's role as intermediary between users and the library. Another player, the information manager, emerges in the electronic library environment; (s)he is responsible for the development and the maintenance of the system, and uses expertise drawn both from library and information science.

When organizing electronic libraries, most of the technical operations performed in traditional libraries have to be reconsidered. One of the main issues is the acquisition and processing of material. Material can exist in different formats, and conversion tools are usually essential parts of the acquisition process. Another issue is concerned with the cataloguing of the material in order to store it in the library database. Several issues are concerned with public services: how to provide copies of documents; how to display them; how to inform users; how to assist users; how to manage search and retrieval taking into account gateways to other library systems; how to manage loans, which can be interpreted as temporary provision of copies or free-copy delivery according to the system strategy; and how to protect copyright.

### Two new forms of electronic book.

The authors of this paper are currently working on a research project whose purpose is the design and development of an environment for building electronic books. Referring to the classification above, our electronic books belong to the final category, as both the physical and logical aspects of the book are taken into account. This choice has been based on our view of user-needs (readers need to consult a book on the screen and read portions of it, not just find an entry in a dictionary) and of our selected application area (scientific publications). For that reason, large screens (at least 19 inches) are an important feature in the design of our electronic books. At the moment, low resolution is one of the main obstacles to acceptance of reading electronically displayed text, but technological developments will ensure that this barrier will be removed within a short period of time.

Our electronic books are conceived not as isolated objects but as integral parts of an electronic library, and some of their properties are dependent on this feature. Users typically find a book after a bibliographic search or browsing through library shelves. They then consult it in order to see if it is what they are looking for and, if the book seems to be interesting, they borrow it. Our electronic books are designed to allow this kind of approach: browsing, followed by personalization, and finally printing in hard copy.

Another important aspect of our research is that in addition to the interface-design issues, a flexible publishing environment is provided in order to produce electronic books in a semi-automatic way. More specifically, the design of two different environments is involved: the first consists of a set of tools which allow the conversion of existing electronic texts into

electronic books (*hyper-books*); the second aims to define tools for the conversion of paper books into their electronic counterparts (*visual-books*).

### Hyper-books

More and more information is now available in electronic form, but it is not presented in a natural and user-friendly way. The hyper-book proposes an alternative for representing electronic text which overcomes the problem of limited familiarity with text representation, and defining an environment where electronic books are generated using electronically produced text. An important feature of such electronic books is the possibility of accessing text in a non-linear way, following the hypertext philosophy. It is obviously for this reason they are called hyper-books. The environment where they are produced is known as a *hyper-book builder*.

Hyper-books maintain many of the features of paper books and present information in a format which is as close as possible to the paper book format. In addition, other capabilities can be provided such as indexing, links, and history mechanisms. A hyper-book is presented as an open document with the two pages side by side in view. The page contains different components: the footer, the header, and the text itself. The text may contain notes, pictures or other non-textual components. A different font style is used for distinguishing active areas of the text (i.e. links).

The hyper-book supports several book-control functions such as turn page and close book, and a number of reader tools: orientation, personalization, searching, history, navigation and printing facilities, which are described in turn below. In addition, a number of editor tools are provided (see Figure 1). These are used during the building process of hyper-books, but are not visible when the hyper-book is made available to the reader in its final version.

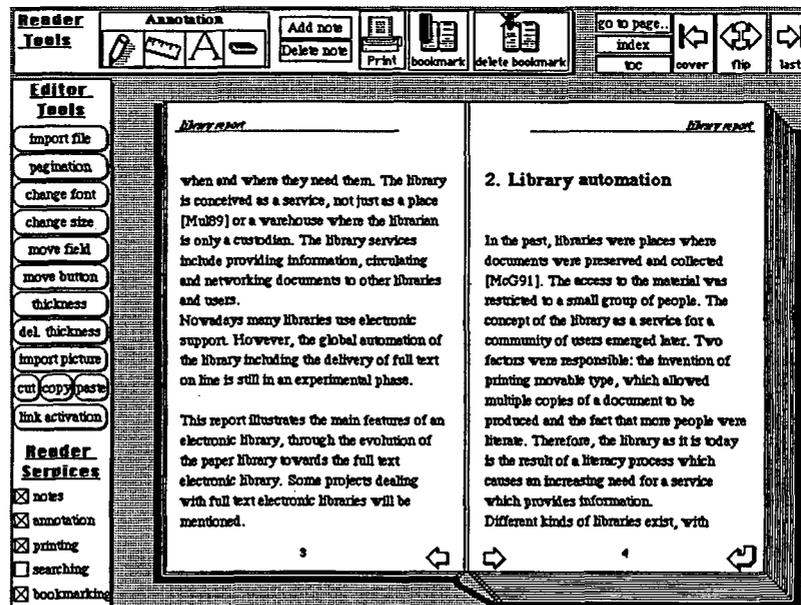


Figure 1: The hyper-book template: designer and reader services

The obvious method of providing orientation in a book is the Table of Contents. In addition, paper books use running headings and the thickness of the paper as orientation clues. Some electronic books show the Table of Contents at all times (Egan *et al*, 1989; Catenazzi, 1991) with a marker on the current position. However, this solution does not respect the book metaphor. Hyper-books provide a Table of Contents only in the preliminaries, with the possibility of jumping there by clicking on a button from any part of the book.

Many approaches have been adopted in electronic books for indicating the user's position within the book, such as the closed pages (Burril, 1989) and black bands (Benest, 1990) mentioned above. Hyper-books provide black bands as well as running heads. A black band, simulating the thickness of the book, is composed of vertical and horizontal lines along the bottom and outer sides of the book, proportional to the number of pages currently opened. It is possible to make large jumps in the book by selecting the black band in a specific place.

Hyper-books provide several kinds of personalization: annotation, highlighting and book-marks. These have the advantage that they may be superimposed on the page without altering the original copy of the book for another user. The types of annotations provided by hyper-books are essentially typing and graphic annotations. In addition, it is possible to write notes in an overlapped window associated with the text. A bookmark is displayed as a small rectangle inserted on the top side of the book, with a number indicating the page where it has been created. Bookmarks may also be removed. Another opportunity offered by hyper-books is the possibility of copying pieces of text and saving them in a personal clipboard.

As far as searching facilities are concerned, the index and Table of Contents represent the two main methods of finding useful material in a physical book. Hyper-books provide both of these, as well as full text indexing. The result of a search using full text indexing is visible in the main body of text through highlighted search terms, as well as in the Table of Contents as the numbers of search-term frequencies for each entry. This approach guarantees that contextual information is always available (Egan *et al*, 1989).

When consulting electronic books, it is useful to remember the path followed during the reading process. This can concern all visited pages or only those which have been marked with bookmarks. Hyper-books provide history mechanisms for both cases.

There are essentially two ways of navigating in a book: hierarchically (from the Table of Contents down to the units of text) and transversally (references across the text). Hyper-books provide both these methods. Transversal links include internal links (reference across the document) and external links (links across the library). The third way of moving in the text is the classical linear way. The hyper-book offers the possibility of going back and forward through the pages of the book, as well as flicking through pages.

The two pages currently in view, or selected parts of the book, may also be printed in hard copy. All the typographical information of the hyper-book is preserved in the paper copy.

Not all the facilities described above are necessarily provided to the reader. Most depend on the choices made during the building phase, to which we now turn.

### Hyper-book builder

The hyper-book builder is an environment in which hyper-books are produced. Two activities are basically supported: design of the hyper-book interface (designer's role) and interpretation

and presentation of electronic texts according to that interface. Hyper-book production presents many analogies with the production process of a paper book. Essential players in the production of paper books are author, editor, designer and publisher. The roles which are mostly reproduced in this context are that of the designer and that of the editor. The text is assumed to be in its final version, and the author is not considered. The editor and the designer collaborate to produce the definition of typographical aspects.

#### Main issues related to the hyper-book builder

An electronic text may be available in several formats, the simplest being a flat ASCII text. However, there are many programs for writing and manipulating text, each of which stores it in a specific format using a markup language. Markup gives information about the logical structure and/or the physical appearance of a document. Markups can provide useful information when automatic treatment of the document is required, and it should therefore be exploited in the process of translating the electronic text into the electronic book. Each markup has specific features, and different translation rules have to be defined for different markup languages.

In order to realize our first prototype hyper-book, SGML (Standard Generalized Markup Language) was chosen as the markup language. SGML offers a number of advantages over other markup techniques such as ODA, Interleaf or Scribe, or specific markup languages used by word-processors. SGML was accepted as an international standard (ISO 8879) by the International Organization for Standardization in 1986. Macleod (1991) describes it as a system for specifying markup languages rather than as a markup language itself. It allows for the definition of different types of documents, so that a Document Type Definition (DTD) formally defines which elements are allowed in a document and how they are organized. The language is device- and application-independent (Feeney, 1989). Only the logical structure is indicated by the markup, and the same document can be presented in many different ways. This represents an additional advantage of using SGML when, for instance, the same document has to be accessed on different machines.

SGML is thus a standard and has all the advantages of an open system, which solves the problem of exchanging electronic documents, something which is extremely important in the context of implementing an electronic library. Another advantage of using SGML is the availability of ready-developed software for handling it. There are SGML parsers, which verify that a document conforms with its DTD. Translators also exist from other formats into SGML, and the system can be easily expanded. Many initiatives and projects are promoting the use of SGML in the USA and, more recently, in Europe. There are good possibilities that it will be more and more widely used.

As for design principles, considered in terms of types of interface, styles of interaction and facilities and services provided to the users, these may be partially deduced from traditional book design principles and from contemporary experiments with electronic books. Many aspects of paper book design are also valid for electronic books, although some constraints on the one hand, and more capabilities on the other, contribute to create a very different object. The size of the book, for instance, is constrained by the screen dimensions: the font size should be bigger on screen than on paper in order to be legible. The study of book design issues is particularly important, because the hyper-book builder has to provide a number of empty templates. These are pre-formatted empty book structures and pages which have to be filled in by the text, as will now become clear.

### Using the hyper-book builder

The purpose of the hyper-book builder is to provide an environment where the user (at this stage the user is the creator of the book) can build his or her hyper-book through a set of tools described below.

A hyper-book is generated by starting from empty templates and a text in SGML or ASCII form. A template is a model of a hyper-book; it represents it in terms of structure, format and reader services. The contents are not included. The term *structure* always refers to the logical structure. The format is relative to the presentation of the text and to its appearance on the screen. The structure of the hyper-book is determined by the DTD (Document Type Definition) if the input text is in SGML; no structure is defined for ASCII texts.

The format specifies a number of layout features: book size; page size, margins, footer and header; typefaces to be used in different parts of the book; notes; link activation process; design of specific pages (title page, contents page, index, references, etc.).

The reader services include a number of facilities, as described above (orientation, personalization, searching, etc.) which, during the book generation phase, may be defined by the book designer (editor) for the reader. In Figure 1, the editor panel contains a list of services that the designer can select for the reader. Once a service has been selected, corresponding buttons appear in the reader tools panel area.

A number of predefined templates are provided, while a set of control functions enables the user to modify some aspects of the book format and also some reader services, taking into account the reader and the application features. In order to generate hyper-books, these empty templates are filled in by the text. The book can then be further modified.

Different steps are involved in the process of generating hyper-books. The first consists of the automatic translation of SGML texts into the hyper-book format. A template is chosen from the set of templates valid for a specific DTD before starting the process. The translation is a two-step process: the parsing process which verifies that a document conforms with its DTD, and the application of rules which map the different elements of the document into a format specified by the template. If the text is available in ASCII form, it will be simply inserted and paginated in a template.

The next step consists of manipulating the resulting book in order to improve its appearance. The changes deal with the templates and the text itself. The resultant text is characterized by containing active areas (links). The user is provided with the ability to modify the link structure: new links may be added, and existing ones may be deleted or changed. In order to complete the generation process, full text indexes may be provided. There are several reasons which justify the choice of full text indexing: firstly, people often remember exactly a word in a text, for instance a date or a name, and they need a mechanism for finding that word; secondly, the cost is relatively low compared to keyword indexing, and the indexing process can be automatically implemented.

When the generation process is complete, the book is ready for the end-user, all the tools used in the generating phase having been removed.

Our prototype hyper-book is under development on a Macintosh IIx; the screen size is 19 inches and the resolution is 75 dpi (dots per inch); the software platform is HyperCard 2.0.

The SGML software currently used is XGML Translator, a tool for specifying conversion produced by Exoterica Corporation.

Figure 2 shows the conversion process. DTD x represents a given DTD. T1, T2, T3 are templates defined for that DTD.

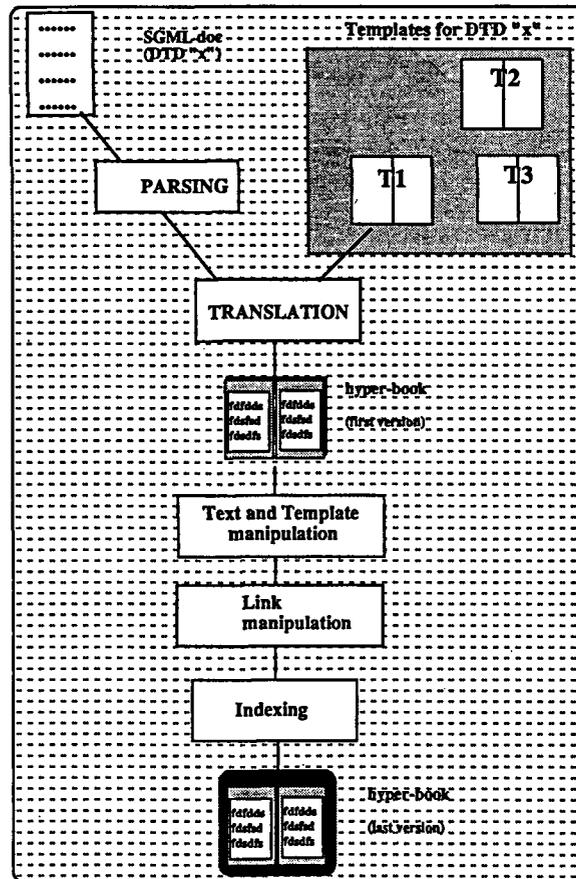


Figure 2: The process of converting electronic texts into hyper-books

### Visual-books

The visual-book represents a particular interpretation of the electronic book, based mainly on the visual aspects of a conventional paper (real) book: its physical features, such as dimensions, thickness, page form and general design style. The visual-book is the result of a process of conversion of existing paper books into electronic form.

The visual-book project is part of a more general project called the Virtual Electronic Library, under development at the Environmental Informatics Unit of JRC at Ispra, Italy; the

visual-book is a practical approach to real-book translation. The physical aspects of the real book are the keys to the construction and representation of the visual-book. The focus is on the book's appearance, interpreted as a way of conveying the cognitive background the reader already has. Book functionalities are studied in relation to the book's use, so information is presented in a natural and familiar way, and at the same time the reader's activities are supported in a comfortable environment.

An important point in visual-book development is the choice of which original paper book to translate. Not every kind of publication is suitable for electronic translation. The computer has some limitations when compared to paper support, such as portability and ease of reading, but at the same time offers advantages, related to its active role. An electronic environment allows changes and updating of original information, provides different views/reading of the same document, integrates multimedia sources of information, permits interchange of data, and offers software support online. All these facilities, while being useful for some types of publication, are not appropriate for every kind of book; different kinds of reading requirements make electronic translation more or less useful for the reader. In particular, books which are commonly read in their entirety and in a linear way, such as novels, are not as suited to electronic translation which can generate confusion and disturb the reading. Indeed, the main requirement of such books is that of legibility, and different views can disturb the reading flow. Furthermore, if the active environment offered by the computer is not exploited, the screen is reduced to being no more than a surrogate for a role that paper can play better. The relation between book and reader in this context is very close, and the presence of the computer is more a hindrance than a help.

On the other hand, publications which are consulted rather than read, such as encyclopedias, manuals, scientific documents, and in general any reference books, and which suggest a different kind of reading geared to problem solving, are more suited to electronic support. The reader of these types of publication is more involved in using their contents than in simply reading them. For this reason, every attempt to enrich and exploit the existing information is useful to the reader, and electronic support plays an important role by offering more facilities than the paper version.

Another specific and interesting application area, especially for visual-book translation, is that of ancient manuscripts, where page image is often more relevant than content, and the original document is too precious and fragile to be consulted and read as a normal document.

When looking for a meaningful application for the visual-book, two areas have therefore been considered: scientific documentation and ancient manuscripts. Technological constraints have made parameters such as the book's dimensions, printing and paper quality particularly relevant when choosing an application. The area we are examining at present is that of scientific publication, this being the most promising common application for visual-book and hyper-book.

#### Basic concepts of the visual-book

The main idea of a visual-book is to offer the reader an object as similar as possible to the paper book. An important principle in visual-book design is therefore to determine which paper book features are most useful and familiar, and to reproduce them in an electronic environment, particular attention being paid to the functionalities a visual-book could acquire from electronic support, and to the choice of those which would be most useful and consistent

with the specification of the book concept. One of the main advantages of electronic support is its dynamic nature which allows the modification and updating of data. On the other hand, a set of functionalities already available in paper books but not currently in their electronic translation are widely used by conventional readers, functionalities such as bookmarks, notes in the margins or elsewhere in the text, highlighting interesting parts, easy access to pages frequently consulted following different clues, information about the ratio between pages already read and those remaining, control of the reading progress, browsing, and searching for interesting sections.

The translation of the paper book into a visual-book has to take into account all these functionalities with respect to the available technology.

#### Visual-book features

The visual-book is composed of a set of visual-pages, i.e. images of original pages connected by the same relations as in the original book. Page images are shown on the screen following the book metaphor: a pair of images, the left and the right pages, are displayed in a graphical replica of a book, maintaining as much as possible the original size and shape.

This kind of representation solves, in a natural way, problems such as:

**Orientation:** maintaining the same aspect as in reality, so the reader has a familiar clue while navigating the visual pages. The page number, the page location, the position of pictures and graphics and other visual clues are consistent with the real-book version, thus overcoming orientation problems which may be caused by the browsing facilities offered by a dynamic medium.

**Personalization:** this is more a technical than a conceptual problem. In fact highlighting, underlining, and adding notes are all possible and naturally interpreted (maintaining, once more, the same appearance as on the paper) in the visual-book but, at this development stage, the difficulty lies in managing pictorial objects overlapping written text.

**Searching:** this is based on the Table of Contents and indexes already present in the original book. There is no text indexing because of the pictorial nature of the visual-book. Another important strategy in visual-book searching is serendipity, the art of finding relevant or interesting information which you were not specifically looking for. By maintaining the real-book appearance, it is possible to reproduce in the visual-book the same effect as in real-book browsing.

**History:** this can be maintained as a visual map of the visited pages, using suitable icons to represent them.

#### Similar work

An interesting interpretation of the visual-book concept is the one developed by the Library of Congress research service (Arms, 1990). The idea is to merge the existing index file with the document image file, and at the same time provide the document image display (soft copy) and the document hard copy (printed version). The objective of this experiment is to evaluate the impact of sophisticated electronic technologies for translating documents to optical digital storage and retrieving them in order to preserve and manage library printed material. At the same time, it is interesting to examine the effects of the technology on copyright, royalties and the publishing community. An interesting feature, when developing the visual-book, is the

acquisition module and the connected compressor and decompressor module. The storage strategy is based mainly on optical-memory technology.

Another possible interpretation of the visual-book, as a reconstruction of a real document in an electronic medium, is the work developed by INIST - the Institut de l'Information Scientifique et Technique (Gourd, 1991). This organization manages a large number of written scientific publications, after scanning them in electronic format. The system used for document management is GEAC (the automated library cataloguing system used in some UK universities). Digitized documents are then presented in hard-copy form when a request is made, thus allowing the system to by-pass presentation problems. The importance of this approach lies in the solution to storage problems and, in general, the addressing of the problem of managing large document collections.

Another similar system is ELNET (King, 1991), a database in Japanese composed of 34 newspapers and more than 250 journals, all digitized on optical support, and available as hard copy. It has also faced the problem of data storage, and the solution (optical storage) is close to the one chosen by INIST. This system is part of a family of similar work all aimed at information delivery through fax or equivalent tools, starting from a paper original, and avoiding display problems by only providing hard copies of the original document.

#### Visual-book builder

The chosen interpretation of the visual-book is based on the collection of page images organized in a way which simulates the real-book structure. The elementary model is composed of pages, interpreted as pictorial objects and the relationship between the previous and following pages, so as to allow them to be turned. It is possible to isolate at least three main components in visual-book construction: the photo laboratory, the pictorial database, and the authoring environment.

The module that allows page-image acquisition is the photo laboratory. Currently we realize the acquisition by testing different grey levels during the scanning process. Image processing can be performed using software such as PhotoShop on a Macintosh. The idea is to define the command sequence leading to a good-quality final image. This is the basic step toward producing a complete automated image-processing module. Theoretical problems arise in defining the concept of legibility, but this stage is mainly based on testing available technology for text image-processing, which includes scanning, compressing and filtering the text image.

At the present status of development, we acquire pages through a DataCopy scanner, using 64 grey levels with a resolution of 100 dpi. The set of books used for our experiment is composed of elements made of less than 100 pages, and of pocket size. Paper quality is average, in order to reproduce a realistic example, and style and typographical components have been chosen to assure a good level of legibility, i.e. the written text has to be clearly printed and its dimension must allow for comfortable reading.

The choice of the scanner has been determined by the quality and properties of the objects to be scanned (black and white text) as well as by the limitation in screen resolution (75 dpi against paper with more than 2,000 dpi). Black and white text does not require a sophisticated colour scanner, which is usually more suitable for photo acquisition, but several grey levels are useful in order to make displayed text more legible. Different scanners, according to these

requirements, have been tested, and other attempts will take place as and when new products appear on the market. Another evolutionary area is that of monitors, where various experiments are also foreseen in parallel with the development of display technology. For this reason, the acquisition phase has to be considered as a changeable module, guaranteeing standard output as the basis for the rest of the project.

After this step, images are saved as (Macintosh) PICT files in order to be opened by PhotoShop as well as by HyperCard. PICT is an almost standard format, and guarantees that resulting data is accessible by a large number of applications. At this stage, storage problems are not considered, so compression techniques are not applied to these picture files. At the same time, observations have resulted in a set of experimental values for acquisition parameters such as the grey levels and the dpi in order to obtain a good, readable image. The result displayed on the screen, a 21 inch colour monitor with a 75 dpi resolution, is legible and clear but not so much as to make the user prefer to read the electronic version instead of the original paper one. The purpose of the visual-book project is first and foremost an investigative one, and results are used as a starting point for exploring other related concepts.

As for the pictorial database, a picture repository has been created in order to collect the results of the photo laboratory. A hierarchical organization is given to pictorial objects which reconstructs the book's original structure. In addition, a file containing details of the logical structure of the book is constructed during image-book creation. The purpose of this descriptive phase is to realize a prototype as independent as possible of particular kinds of book presentation and structure. Currently, the emphasis is on developing an image storage strategy. A library composed of visual-books requires a large amount of memory in order to keep and manage an appropriate quantity of visual-books. The most promising approach is the one followed in the set of projects mentioned above, i.e. optical storage. At the same time, software and hardware compression techniques can be applied to the picture files, and in particular those solutions suitable for black and white images of text.

The authoring environment is the module that allows an image book designer to reconstruct the visual-book on the screen. The basic requirement for this environment is to maintain the original pagination of the document, and perform it in a (semi-) automatic way, following the description defined in the pictorial database construction. This offers a natural and well-known access key to the document which already exists in paper form, giving a consistent way to browse and read. Some other reader services have been added around the simple image book: personalizing, searching and history, developed under HyperCard. Page pictures are inserted in the correct position within the book at run time by opening the corresponding picture file and showing it on the top of the HyperCard stack whose background mimics an appropriately open book.

The book is presented using two templates: the first simulates the cover page, the second imitates two open pages. A lateral band suggests the current position while reading or consulting the book. Arrows are shown on the screen for turning the visual-pages backwards and forwards. The Table of Contents is interpreted as text (using OCR or inserting it manually) that is translated into active references for readers looking for particular subjects. The index is a more complex object. Many conventions have to be exploited in order to translate in a meaningful way the classical paper index: references such as *see also* or stylistically emphasized items need particular care.

## Hyper-books and visual-books in an electronic library

The natural evolution of our prototype is the organization of hyper-books and visual-books into a library, and their distribution on line. The resulting electronic library is called the Virtual Electronic Library (VEL). The acquisition problem has already been solved: hyper-books and visual-books are the result of this process. The next step consists of organizing and making them accessible. Online catalogues are the main access method in the VEL. The cataloguing process could be automated for hyper-books by exploiting information provided in their original markup.

In addition to the features described above (infinite copies, infinite space), the VEL presents a number of innovative features. Hyper-books are not isolated objects, and links exist between documents. Visual-books provide a limited link capability but they also offer a powerful tool for the complete conversion of paper books into electronic support, thereby solving problems of storage and maintenance of large collections of paper books. Intelligent tools are being developed by participants in our project in order to simulate the role of the librarian as consultant: help in formulating bibliographical queries, help in understanding the result of a search, general information about the system. The concept of loan in an electronic environment is of course completely different from what one usually encounters: loan in this context is comparable to having access to the electronic library material. As documents can exist in as many copies as requested, books are always available. Once the book has been accessed, the user can put it in his or her personal library and it can then be manipulated: annotation and cut and paste operations can be allowed, just as if the book were the user's property. A charging system may need to be defined in order to manage the purchase of books. Once the book has been bought, the user becomes its owner.

Users have different interests and objectives. By observing user-behaviour, a VEL will be able to construct individual user profiles. Knowledge about user-needs may be used in order to optimize retrieval strategies and acquisition policy. Such reader profiles can be used to notify new publications of interest to the user, and to rank documents, identified during search, in order of relevance. Reader profiles can also be used to filter library material and reduce the confusion caused by an excess of information, or to create contacts among users sharing the same interests.

## Conclusion

During the first stage of development, a number of difficulties have been experienced. Some of them are related to the difficulty of finding SGML documents and their paper counterparts to be used as test applications. Other problems are more dependent on the development environment: HyperCard. These are essentially the limited text-processing capability (for example, text cannot be justified), limited graphics management (images are currently included using external functions), inefficiency in performing heavy computational operations (repagination, thickness redrawing) and in managing grey-level images. Future versions of HyperCard should improve text-formatting capabilities, and speed up executions thanks to a compiled programming language.

Display problems connected with limited monitor resolution will be solved by the evolution in this technological field, while storage aspects will be faced in the next development phase. An evaluation strategy will be exploited in order to formalize the real impact of our prototypes.

Finally, the actual publishing of hyper-books and visual-books is extremely important in the context of building an electronic library, and this represents the future objective of our work.

## References

- Arms, C. (1990). 'The context for the future', in Arms, C. (ed), *Campus Strategies for Libraries and Electronic Information*, Bedford MA, Digital Press.
- Barker, P. (1991), 'Interactive electronic books', *Interactive Multimedia*, 2, 1.
- Barker, P. and Manji, K. (1991), 'Designing electronic books', *Educational and Training Technology International*, 28, 4.
- Benest, I.D., Morgan, G. and Smithurst, M.D. (1987), 'A humanised interface to an electronic library', in Bullinger, H.-J. and Shakel, B., *Human-Computer Interaction – INTERACT '87*, North Holland, Elsevier.
- Benest, I.D. and Duric, D. (1990), 'Some design issues in the automated office metaphor', *Conference Proceedings of the European X-Window System User Group (Guildford)*, Cambridge, UK European X User Group.
- Burril, V.A. and Ogden, J.A. (1989), 'VORTEXT: The hard-backed screen', *Electronic Publishing*, 2, 3.
- Catenazzi, N. and Argentesi F. (1991), 'Super-Library: an environment for the conversion of linear text into hypertext', *RIAO 91 (Recherche d'Information Assistée par Ordinateur) Conference Proceedings*, Barcelona.
- Collis, B.A. (1991), 'The evaluation of electronic books', *Educational and Training Technology International*, 28, 4.
- EBT (Electronic Book Technologies Inc., Providence RI, USA) (1990), *Dynatext: Electronic Book Indexer/Brower* (descriptive leaflet).
- Egan, D.E., Remde, J.R., Gomez, L.M., Landauer, T.K., Eberhardt, J. and Lochbaum, C.C., (1989, 'Formative design-evaluation of SuperBook', *ACM Transactions on Information Systems*, 7, 1.
- Feeney, M. (1989) 'SGML', *Library and Information Briefings*, June, pp. 99-115.
- Feldman, T. (1990), 'The emergence of the electronic book', *British National Bibliography Research Fund Report No. 46*, London, BNBRF and the Publishers Association.
- Gourd, F. (1991), 'Il sistema interattivo di fornitura di documenti numerizzati all'INIST', Convegno SMAU/CNR', La Biblioteca Elettronica, meeting of 3 October, Milan (unpublished summary of presentation).
- King, S.V. (1991), 'Il sistema TELNET', Convegno SMAU/CNR, La Biblioteca Elettronica, meeting of 3 October, Milan (unpublished summary of presentation).
- Lande, N., (1991), 'Toward the electronic book', *Publishers Weekly* 238, 42.
- Macleod, I.A., Barnard, D.T., Hamilton D. and Levison M. (1991), 'SGML documents and non-linear text retrieval', in *RIAO 91 (Recherche d'Information Assistée par Ordinateur) Conference Proceedings*, Barcelona.