

Towards Multimedia Conference Proceedings

M. CHEYNEY¹, P. GLOOR², D.B. JOHNSON¹, F. MAKEDON¹,
J. MATTHEWS¹ AND P. METAXAS³

¹ *Department of Computer Science, Dartmouth College,
6211 Sudikoff Laboratory, Hanover, NH 03755, USA
E-Mail: {cheyney,djohnson,makedon,jim.matthews}@dartmouth.edu*

² *Union Bank of Switzerland
Bahnhofstrasse 45, 8021 Zurich, Switzerland
E-Mail: gloor@lcs.mit.edu*

³ *Department of Computer Science, Wellesley College,
106 Central Street, Wellesley, MA 02181, USA
E-Mail: PMetaxas@lucy.wellesley.edu*

Abstract:

Academic conferences are a long-standing and effective form of multimedia communication. Conference participants can transmit and receive information through sight and sound, that is, by viewing individuals, text, and graphics, and by hearing the spoken word. This same-time, same-place communication is sufficiently valuable to justify large investments in time and travel funds. Printed conference proceedings are attempts to recapture the value of a live conference, but they are limited by both their delivery medium and by the significant differences from the conference presentation. We addressed this problem in the multimedia proceedings of the DAGS '92 conference. The recently published CD-ROM delivers text, graphic, audio, and video information as an integrated whole, with extensive provisions for random access and hypermedia linking. We believe that this project provides a model for future conference publications and highlights some of the research issues that must be resolved before similar publications can be quickly and inexpensively produced.

1. Introduction

Academic conferences are a long-standing and effective form of multimedia communication. Conference participants can transmit and receive information through sight and sound, that is, by viewing individuals, text, and graphics, and by hearing the spoken word. This same-time, same-place communication is sufficiently valuable to justify large investments in time and travel funds. Traditional printed conference proceedings and session videotapes are attempts to recapture the value of a live conference, but they are limited by both their delivery medium and by the significant differences from the conference presentation. Video and audio tapes require as much time as the recorded events, while printed proceedings cannot convey the dynamic nature of a conference. But more importantly, they are not integrated with each other.

In preparing the DAGS '92 Multimedia Conference Proceedings we aimed to address these shortcomings, and therefore deliver more of the value of an academic conference to our audience. Recently, conference organizers have realized the possibilities that the multimedia productions can have, and some first efforts at multimedia proceedings on CD-ROM have already been published [1, 2]. The results are, however, a collection of papers in two or three different formats that can be used mainly as “digital microfiche,” along with a few digitized movies. Their main problems are a lack of interaction between the components, and a lack of fast searching mechanisms. They were produced to serve as an alternative to printed proceedings, yet printing the articles reveals many missing pictures or graphics. Even though these are efforts in the right direction, we believe that they fall short of the capabilities of multimedia.

The DAGS '92 CD-ROM [7] delivers text, graphic, audio, and video information as an integrated whole, with extensive provisions for random access and hypermedia linking. We believe that this project provides a model for future conference publications and highlights some of the research issues that must be resolved before similar publications can be quickly and inexpensively produced. The experience gained from this effort will be applied not only to multimedia conference proceedings of the future, but also to multimedia textbooks and learning environments. We view this as the strongest point of our experiment and research, and a direction that should be pursued further.

2. Background

In June 1992 Dartmouth College hosted the first annual Dartmouth Institute for Advanced Graduate Studies (DAGS) symposium, on the topic of “Issues and Obstacles in the Practical Implementation of Parallel Algorithms and the Use of Parallel Machines.” The symposium program consisted of eight talks by invited speakers plus thirteen contributed talks, presenting a total of twenty-two papers (one invited talk spanned two papers). The topic is a central problem in the area of parallel computation today. It is therefore appropriate to make the results as widely available and accessible as possible to the parallel computing research and teaching communities. So, in addition to the usual printed proceedings [3], it was decided to publish multimedia proceedings that would capture as much as possible of the conference atmosphere.

3. Description of the System

The multimedia proceedings on CD-ROM contains the twenty-two papers in hypertext format integrated with the eight invited talks in Apple's QuickTime movie format. Hypermedia links and random-access are used extensively in this integrated system. The CD-ROM includes:

- A navigation shell that facilitates hierarchical navigation in hyperspace;
- hypertext versions of the papers presented;
- digitized movies of the invited speakers delivering the conference talks;
- their transparencies including marks that were added during the talks;
- hyperlinks connecting relevant parts of the proceedings; and
- bibliography.

In the future, we plan to include animations of some of the algorithms presented. Furthermore, the system is extendible in the sense that users are able to create their own hyperlinks among objects that they deem relevant, do searches on keywords, and keep notes on the documents.

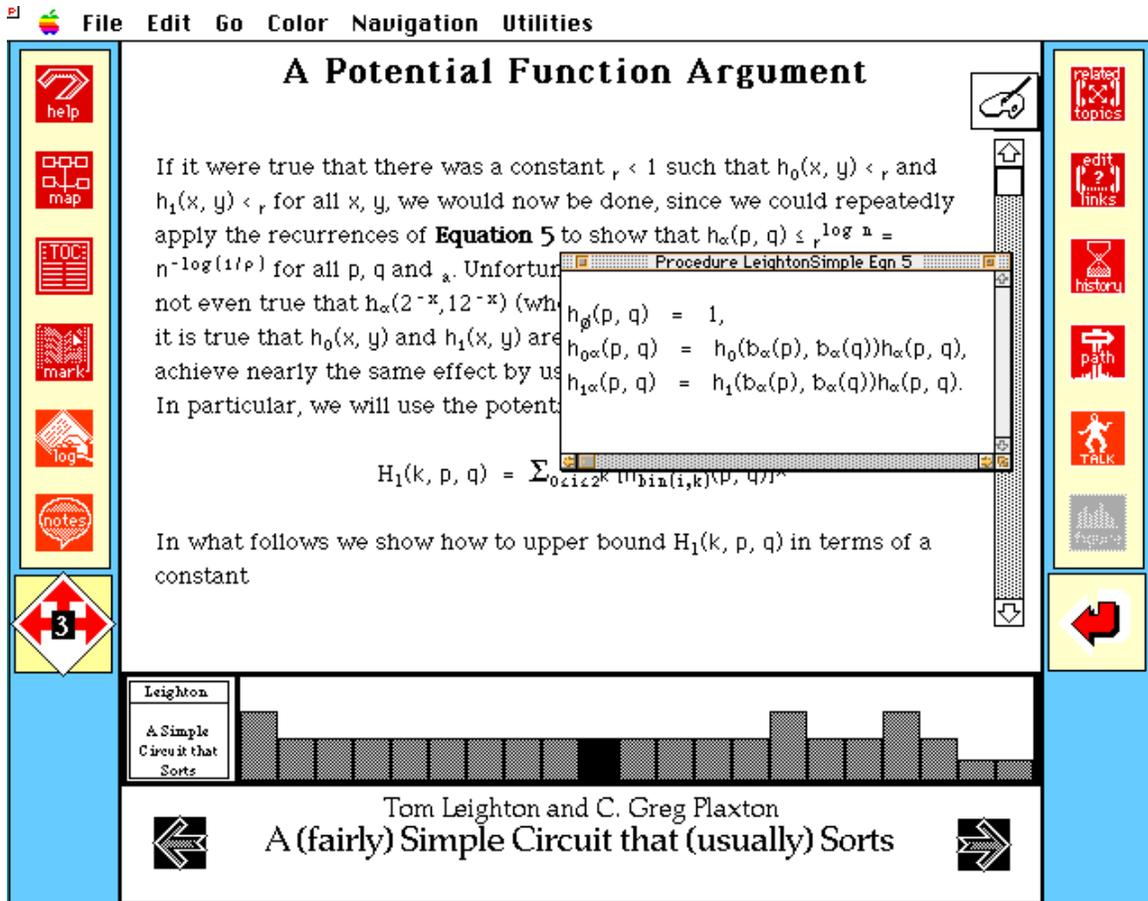


Fig. 1: The Article Screen of the navigation shell.

In a typical session, users can first get a quick overview of the contents of the CD-ROM. Then they can follow a talk on a particular topic that seems interesting by watching the movies of the transparencies and of the speaker (Fig. 2). They can also get an overview of the talk by using the pop-up menu containing the section titles of the talk (by pressing the button "Jump to Highlights" in Fig. 2), or by skimming through the transparencies. If, at some particular point, the speaker mentions a theorem without proof, as is usually the case in conferences, users can jump into the hypertext to read the omitted proof in detail. Then, they can go back to the talk, continue reading the hypertext bringing on the screen several windows containing relevant information, or do a search on some keyword to find out who else mentioned this keyword during the conference. Assuming that the search brought up several candidate sections of papers, users can jump into the new paper and continue reading from there, or even jump into the video

movie of the second speaker and see how the material was presented during that talk. In every place, users can make their own remarks on the subject being read and keep them for later examination or for filling in the background gaps.

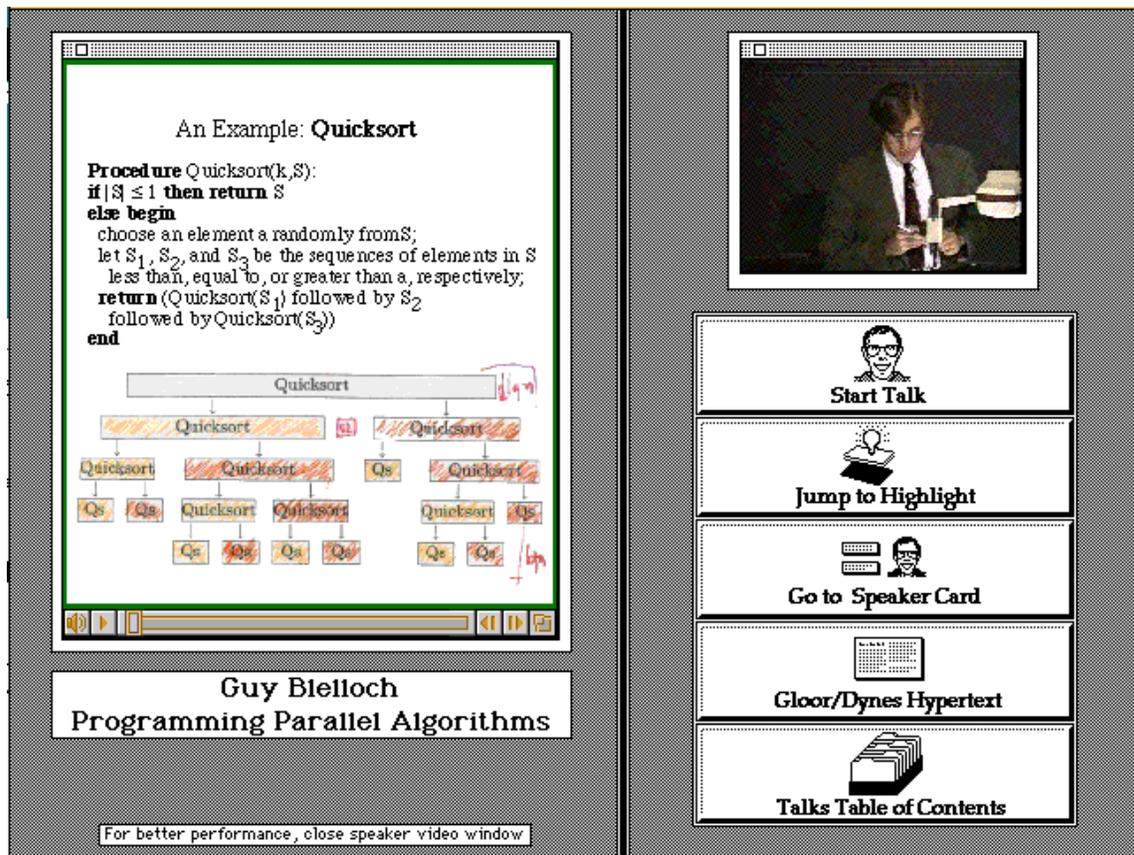


Fig. 2: The Talk Screen. On the left is the transparencies movie, on the top right of the screen is the speaker movie.

There is considerable redundancy and flexibility built in our system; evaluating which parts are crucial and should be kept and enhanced, and which could be dispensed without affecting the overall performance of the multimedia proceedings, is part of future research and evaluation.

4. Guided Tour through "Parallel Computation"

In this section we will briefly describe a possible tour through the hypermedia document to give a first impression of the features of our system.



Fig. 3: The Contents Screen.

The system starts up with a title animation, and then presents a table of contents (Fig. 3) that gives to the users direct access to the eight invited talks. In this example the user decides to study the talk by Charles Leiserson about "The Network Architecture of the Connection Machine CM-5." The user then clicks on the title of this talk on the Contents Screen and the system jumps into the Speaker Screen.

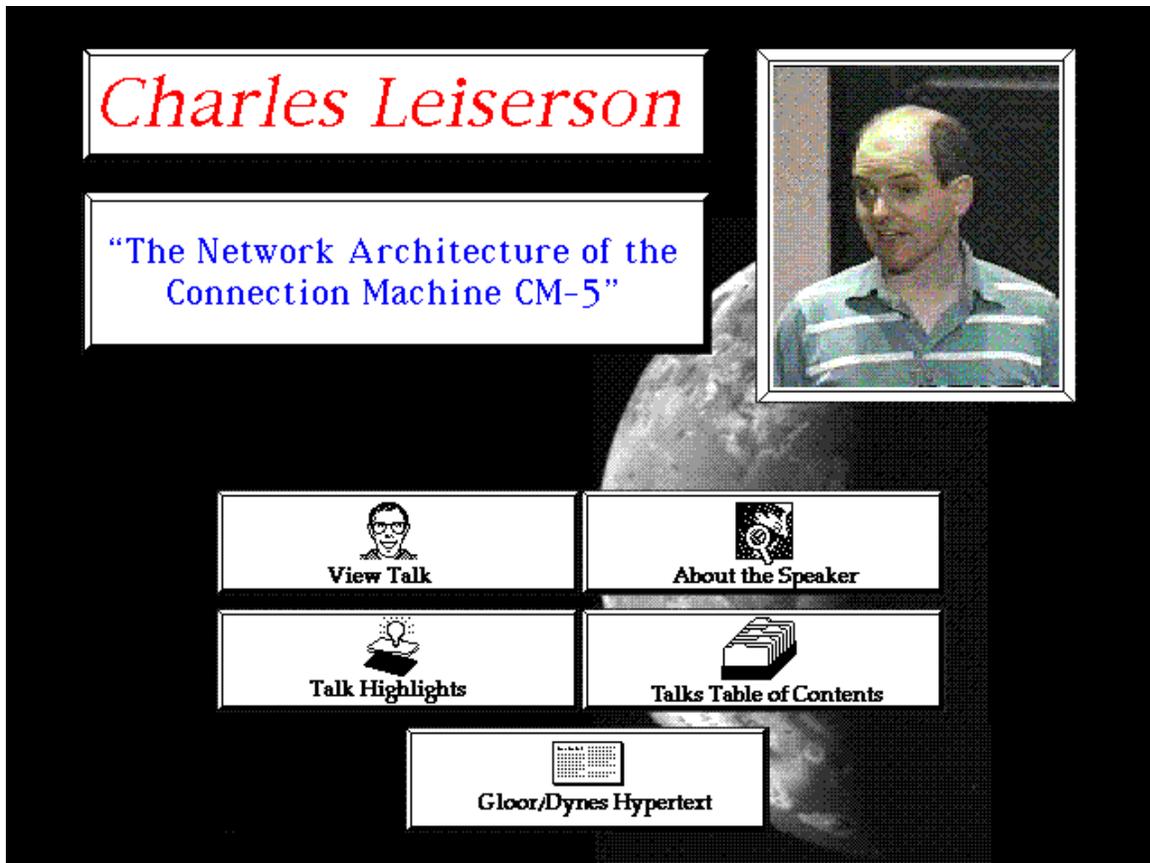


Fig. 4: The Speaker Screen.

On the Speaker Screen (Fig. 4) users can get background information about a particular talk by reading a short speaker biography, obtain a list of the talk's highlights, read the hypertext version of the paper related to the talk, or actually watch the talk. The user here decides to view the talk and clicks on the "View Talk" button.

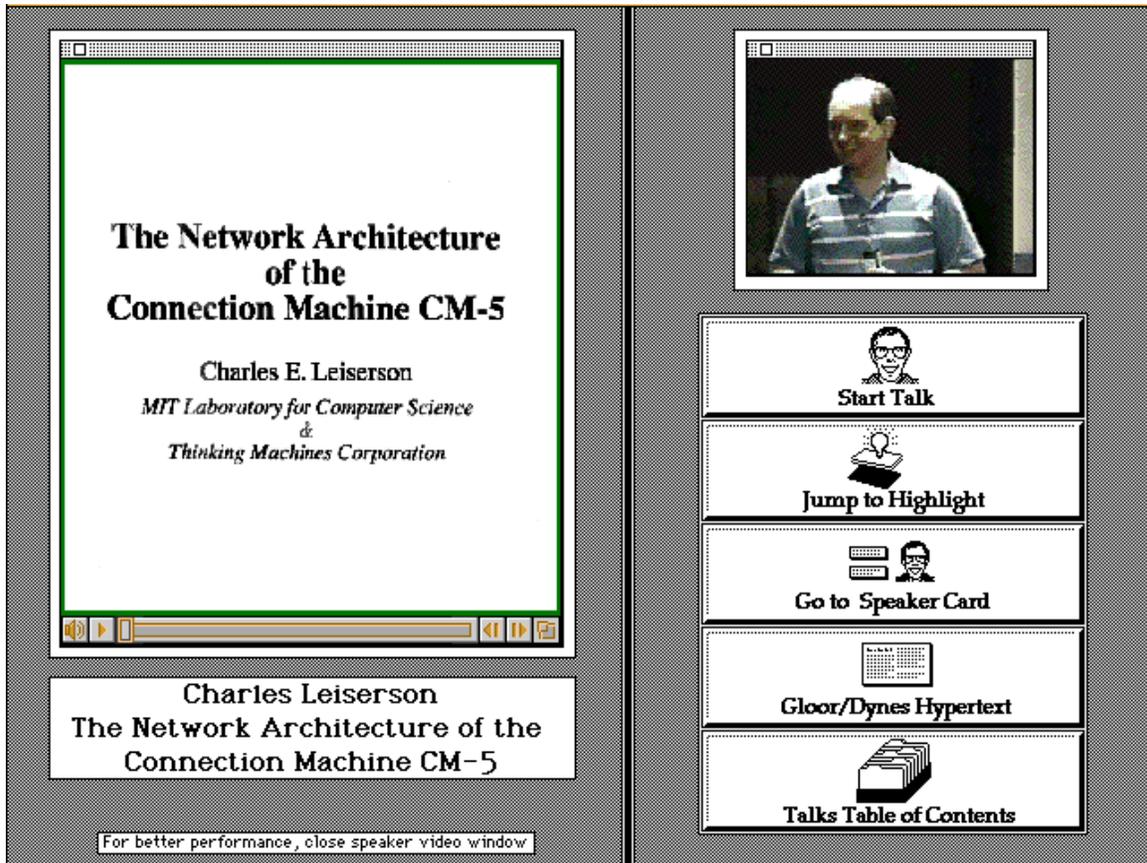


Fig. 5: The Talk Screen.

The system then takes the user in the Talk Screen (Fig. 5) where he/she can watch the actual talk from the beginning by clicking the “Start Talk” button. They can also jump to any position within the talk using the slider controls on the left panel underneath the transparencies.

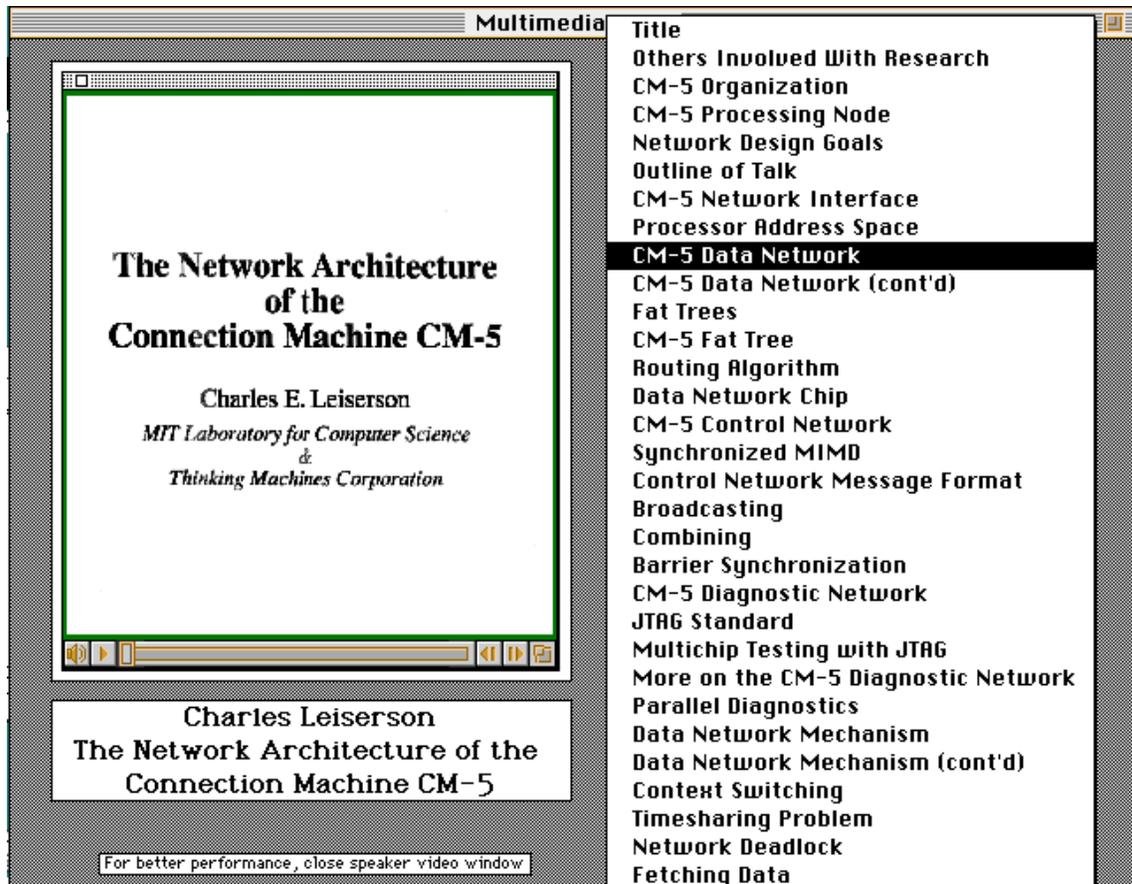


Fig. 6: Highlights of the talk.

The user also has the option to jump directly to the beginning of a subsection within the talk using the “Jump to Highlight” button (Fig. 6). Upon pressing that button, the user is presented with an ordered pop-up-list of the highlights of the talk. In this example the user chooses to watch Charles Leiserson explain the “CM-5 Data Network.”

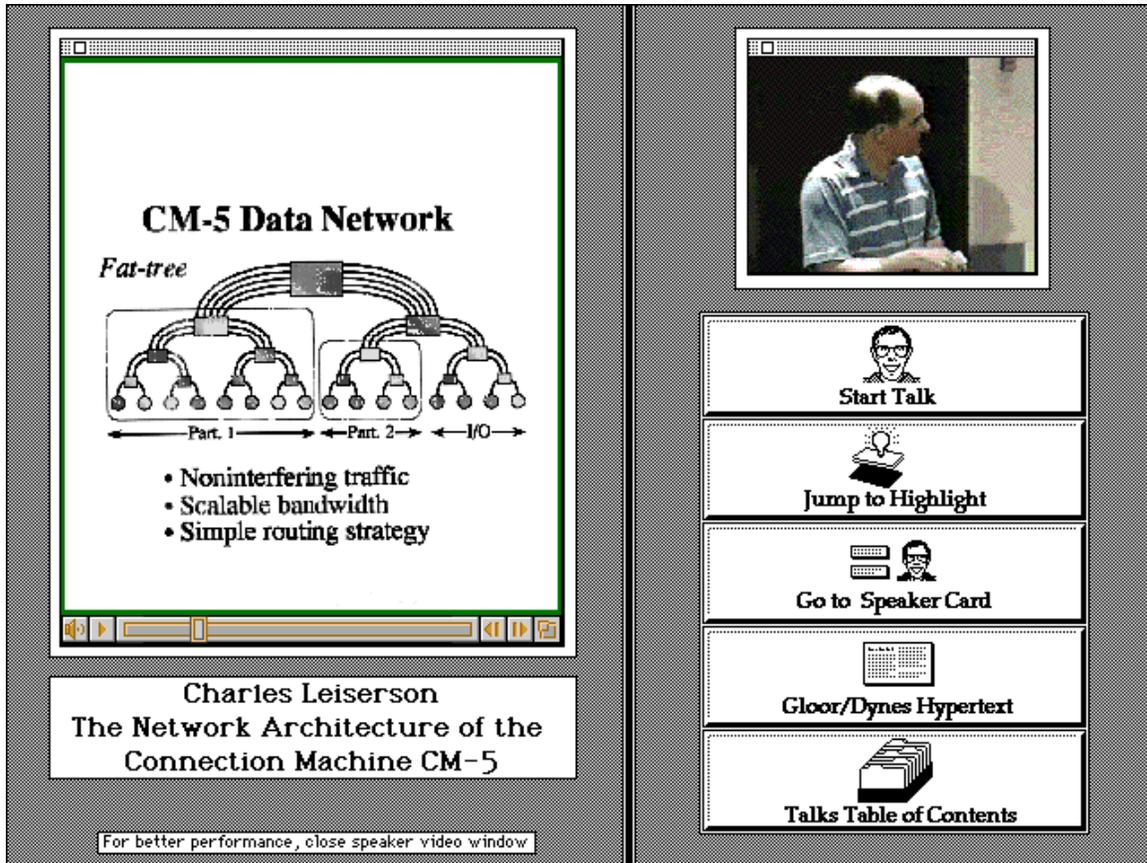


Fig. 7: Dynamic access of a random section of the talk is provided.

When the user released the mouse button, the system jumped directly to the “CM-5 Data Network” section within Charles Leiserson’s talk (Fig. 7). After following the talk for a while, the user decides to read the paper associated with the talk. He/she can do that by clicking at the “Gloor/Dynes Hypertext” button.

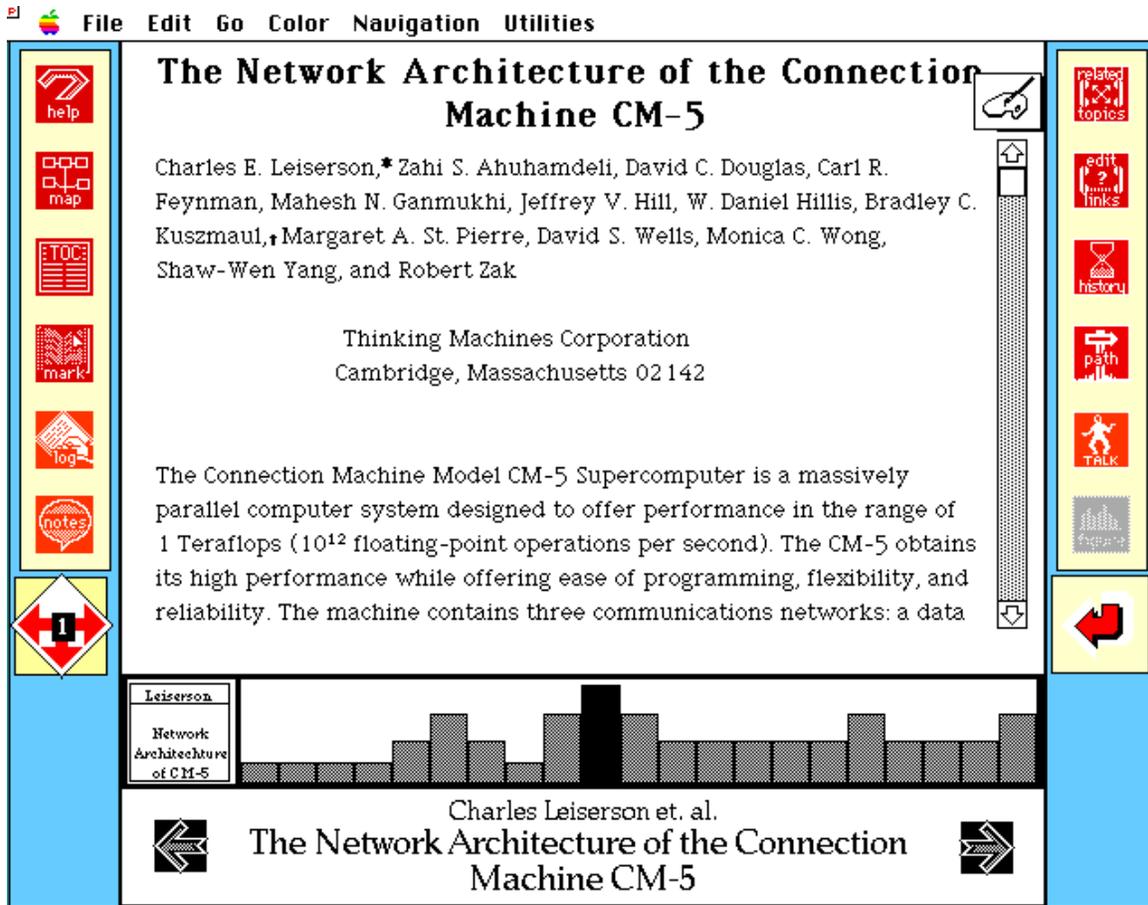


Fig. 8: The hypertext version of the paper.

The system now jumps directly to the beginning of the paper associated with the talk. On the left of the Article Screen (Fig. 8) is a palette of tools that provide

(from top to bottom) help , a map  of the hypertext contents, a table of contents , bookmark facilities , a notebook , and marginal notes . Below is the “navigation diamond”  that marks the hypertext node level and its connections to other node levels.

On the right side is another palette with buttons that provide (from top to bottom) a list of topics  related to the current text node, the ability to add hyperlinks 

between text nodes, a history  of the nodes visited by the user in this session, the ability to create a path through the nodes (or to use a predefined one that takes them through the proceedings in a linear fashion) , a link to the Talk Screen , and a button to the figures  related to the current node (de-highlighted when there is none). Under all that is a button that returns to the node last visited.

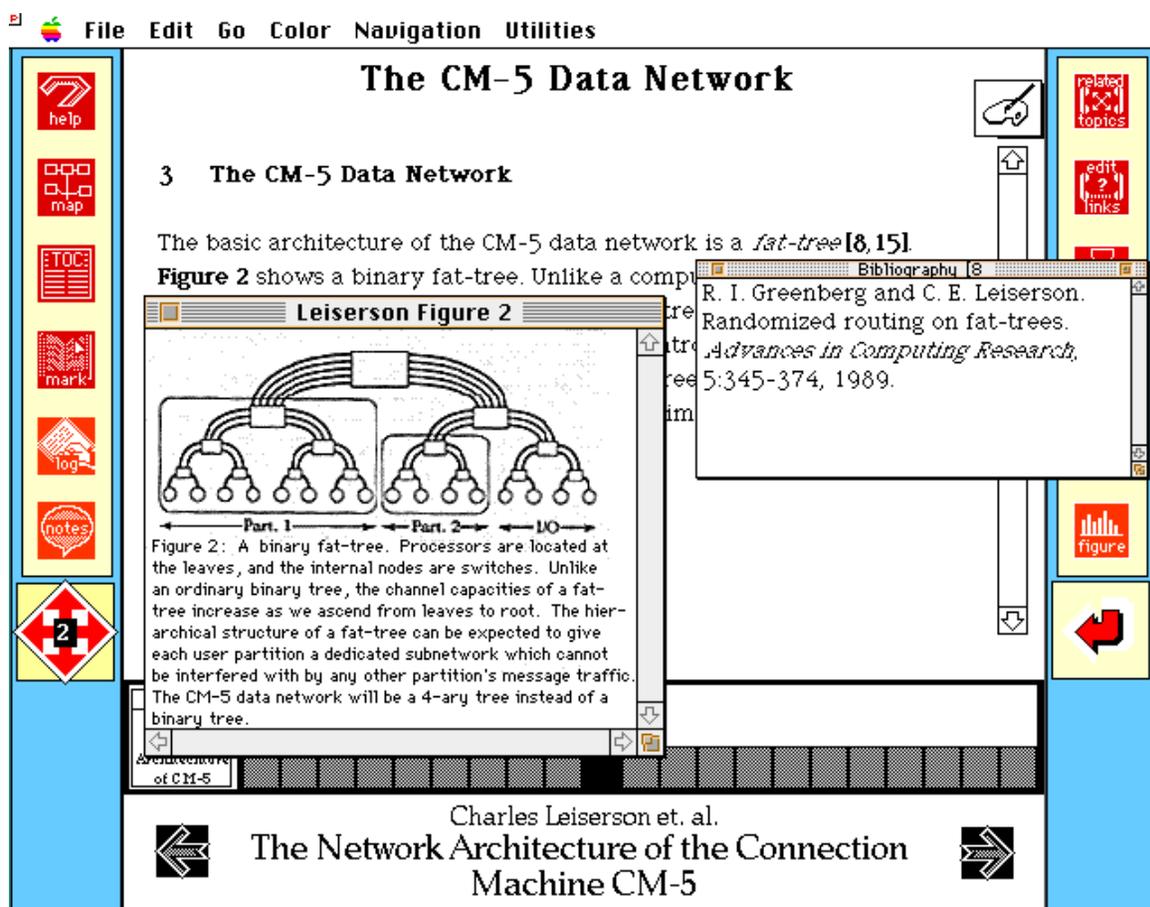


Fig. 9: Figures, bibliographies and footnotes appear in separate windows.

Clicking a few times on the navigation diamond and following the level 2 nodes, the user can get an idea of the contents of the paper. A bibliographical reference or a figure mentioned in the paper can be opened by clicking on the word (Fig.

9). To get a feeling of the structure of this paper, the user decides to bring up the overview map by clicking the “map” button.

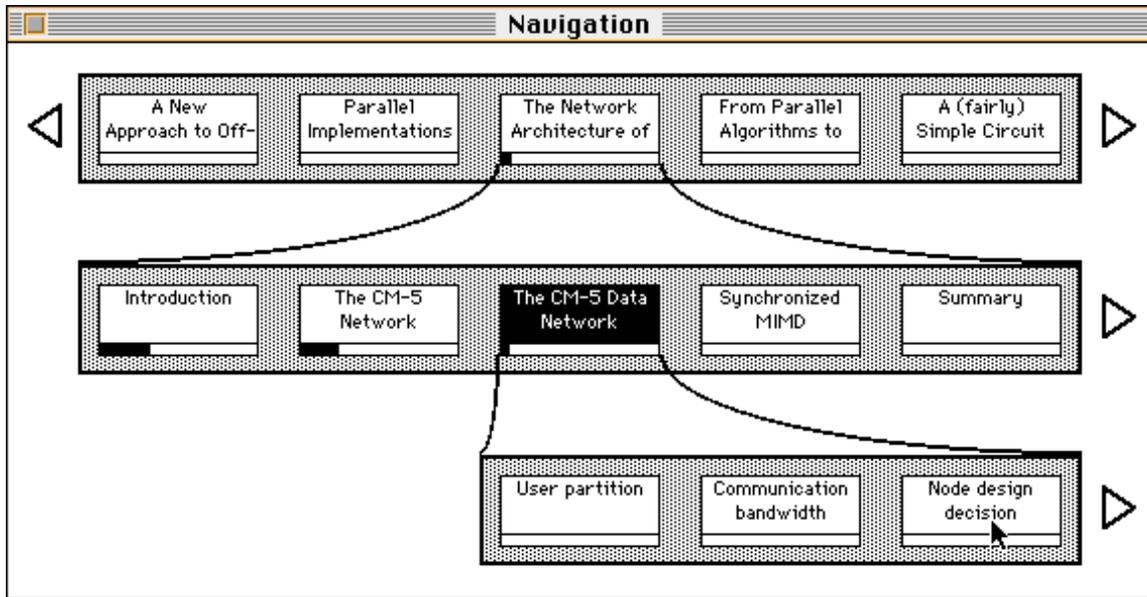


Fig. 10: The map and the hypertext version of the paper.

The system brings up the overview map that shows a graphical structure of the whole hypertext document in relation to the actual viewpoint of the user (Fig. 10). The user learns that he/she is on the section entitled “The CM-5 Data Network” of the “The Network Architecture of...” paper. Information about previous and next sections also appears, along with information about the subsections of the current section. The user decides to jump directly into the “Node design decision” subsection by “option-clicking” this box on the map.

There are numerous other features in our system that can not be shown in this brief guided tour. The reader is referred to the user’s manual of the DAGS ‘92 CD-ROM [7] for a more detailed description of the user interface.

5. Development Stages

The first step in producing the multimedia proceedings was to collect the raw material (video, sound, text), and edit it to produce the videos and the hypertext. This section describes how this was accomplished.

5.1 Producing Digital Talks

The speakers' presentations were videotaped, their overhead transparencies were copied, and their papers were collected. To make the synchronization easier, we used two video cameras for the videotaping; one focusing on the speaker and one on the projected transparencies (Fig. 11). It turned out that the second was by far the more useful. After the collection, all these materials were converted to digital form for computer-based processing using a commercial video card. Due to the space constraints of CD-ROM, our intended delivery medium, we decided to deliver only the eight invited talks in a full audio/visual form.

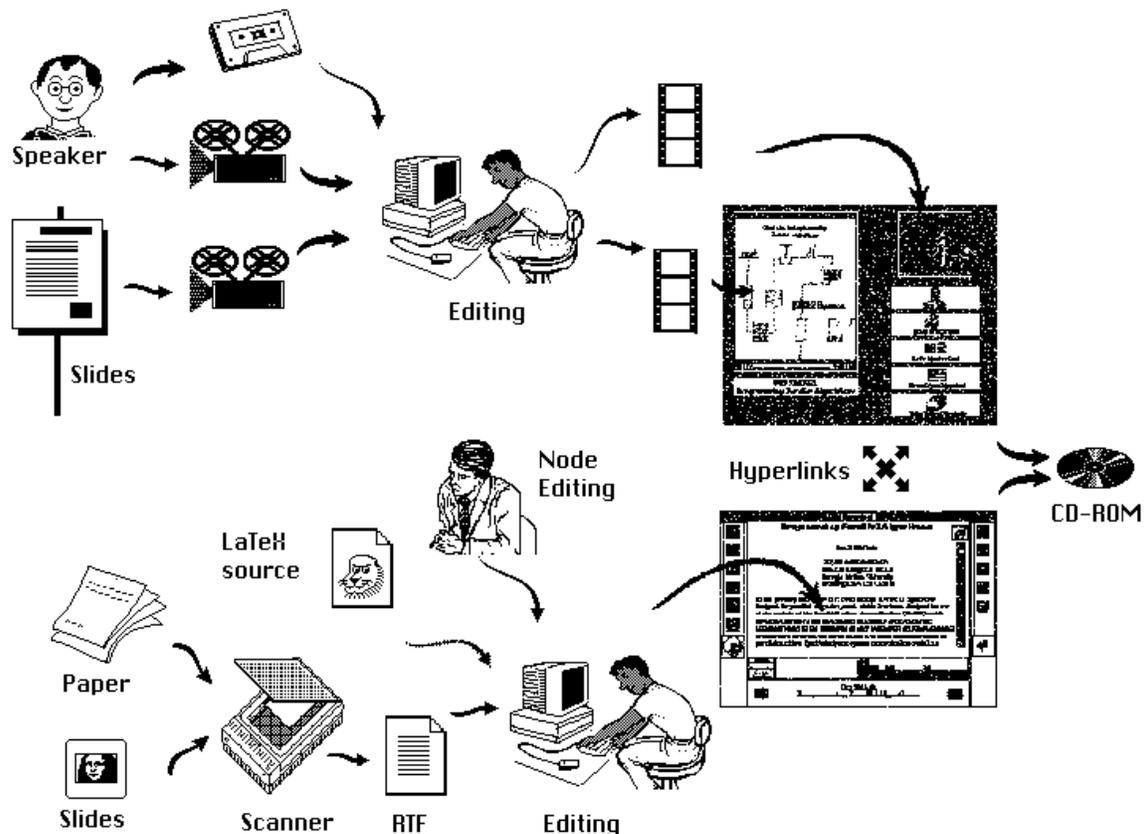


Fig. 11: Description of the developmental stages.

Despite the huge storage provided by the CD technology we knew that we could not fit even the eight one-and-a-half hour video tracks on a CD-ROM. However, most of speakers in conferences are usually standing next to the projector so that

they can change transparencies and point to them. The audience gets most of the information by looking at the transparencies. Therefore we decided to display a short video loop of each speaker by selecting a chunk of the video with similar beginning and ending frames so that the loop transition would not be distracting to the viewers. This was certainly the right decision because little valuable information was lost this way; it turns out that one or two minutes of video is sufficient to convey a sense of a speaker's appearance and mannerisms, while at the same time is less distracting to the person studying the talk. Furthermore, breaking the synchronization between the audio and video tracks allowed us to edit the audio without introducing video skipping. Editing both the audio and the video data of the speakers in a way that preserves synchronization and smooth transition between edited frames seems a formidable task for which we had neither tools nor time to undertake.

The audio track of each talk was digitized and then edited to remove pauses and noise words such as “umm”s and “ahh”s. The edited talks were roughly half as long as the originals, and much more listenable. To improve the quality, we amplified most of them using a commercial sound processing application.

The overhead transparencies were scanned and edited for clarity and contrast; they were also made smaller to fit in the appropriate window of the interface program. One of the problems we encountered in this process was the fact that, after diminution, the transparencies that were poorly handwritten had low readability. So, some of the transparencies were retyped.

Then, using the original videotapes as guides, we synchronized the transparencies to the edited audio tracks using a commercial video editor. The resulting “movie” reproduces the most important features of a talk, the speaker's words and transparencies, and preserves their temporal connection. These “movies” were indexed to allow random access to a list of primary topics, and to allow more sensitive linking between the papers and the talks.

The video loops were digitized with a low-end video capture board and compressed to provide efficient playback from a CD-ROM. The loops were kept small to enhance playback performance and to keep the lack of synchronization between the speaker's lips and words from becoming a distraction. Due to the low resolution of the 120 by 180 pixel QuickTime video window the user has to

look very closely to notice the missing synchronization between lip movement and the voice track.

Given the variety of the playback speeds of the commercial CD-ROM drives available, some computer configurations will have a hard time displaying both the transparencies video and the speakers video loops on the screen at a comfortable speed. For that reason, we have given to the user the ability to stop the speaker video loop and replace it with a static color picture of the speaker.

5.2 Producing the Hypertext

By far the most time-consuming stage was to convert the twenty-two papers into hypertext format. Wanting to produce a truly hypermedia product and not to sacrifice its usability, we decided to present the papers in hypertext form, using an advanced hypertext engine. We used the Gloor/Dynes hypertext system [4] that was developed at MIT for the CD-ROM version of the “Introduction to Algorithms” textbook [5] and is based on Apple's Hypercard software.

We first broke the twenty-two papers into hypertext “nodes,” and assigned each node a “node level” that reflected its degree of generality. For example, all the abstracts of the papers are on level 1. This way, a user can quickly become familiar with the themes of the presented papers by visiting all the first level nodes of the Proceedings (Fig. 12). All the beginnings of sections within the paper are on level 2, so a user can get an overview of the contents by following these nodes.

This “chunking” process was handled manually by computer scientists with expertise in the subject area. It appears that one cannot automate this process unless the authors have written their papers following some carefully predefined specifications. We also tried to have authors provide chunking information, but the results were not sufficiently consistent from paper to paper, so in the end all chunking was performed at Dartmouth.

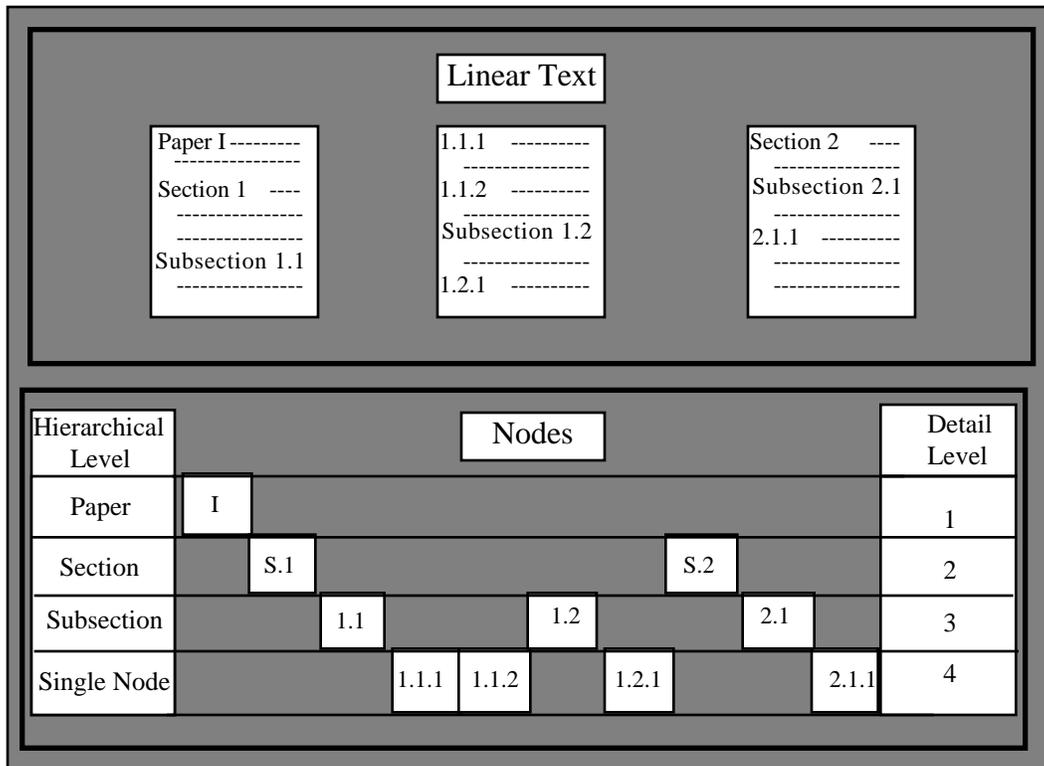


Fig 11: Converting linear text into hierarchical hypertext

Then, we converted all the papers to Hypercard format. Every author provided Postscript versions of the papers and half of them provided electronic versions in TeX or LaTeX format; the latter were converted to Hypercard form with homemade utilities and manual cleanup. Unfortunately, we could not use Postscript files for the hypertext engine, so the remaining papers were scanned, processed by optical character recognition software, and then manually corrected. A great number of errors were introduced by the scanning and recognition process, and some pages were simply retyped. A number of text features required special treatment. Hypercard does not support subscripts or superscripts, so, special fonts were used in their place. Uncommon symbols were similarly provided by custom fonts. This still leaves out very complex equations, which were scanned and displayed as graphics using custom software.

Figures were scanned and edited for clarity. The text was manually marked to provide links to citations, tables, and figures. Each hyperlink in the text appears as a bold-faced word. The system supports multiple windows containing

scanned figures, tables or bibliographic information. Finally, hyperlinks leading to referenced sections were introduced.

The final step was to integrate all these elements into a single user interface, and to test the resulting system. The interface was designed to be simple, usable and attractive. Extensive color graphics and on-line help facilities were built into the navigation shell. The system was tested on a number of machines with different capabilities and configurations. Special care was taken to optimize the transfer data rate so it performs well on a variety of commercially available CD-ROMs.

6. Research Issues

The production of the DAGS '92 Multimedia Proceedings has presented us with two subjects for future research:

- Improving the creation process of multimedia proceedings
- Evaluating the effectiveness of multimedia proceedings.

6.1 Multimedia Authoring Systems

How can such proceedings be assembled in a short period of time, with less manual labor? Even though, in the beginning, some production delays were caused by evaluation of the best alternative manual solution in problems we faced, the experience gained has resolved most of these problems and we now need to focus on the automation of the manual steps.

Some of the time-consuming steps, such as digitizing papers and transparencies, could be eliminated if the source material was available in electronic form. Removing pauses and noise words from audio sources automatically should also be possible, with sufficiently sophisticated techniques. We currently experiment with more advanced software that can help in that direction [6]. Our goal is the design and implementation of a generally usable multimedia authoring system specifically targeted towards the creation of multimedia proceedings.

Converting linear text to a hierarchy of hypertext nodes may be the most intractable problem; it is difficult to see how to replace the human expert, in the near term. It is clear that projects such as this will not be undertaken if they require the multiple man-years of labor we invested; so progress must be made

in automating this process. We are currently working on a carefully specified list of rules and instructions that could help the authors divide an article in hypernodes. Ultimately, we hope that they would use this list in preparing the final version of the paper. This approach, if successful, will eliminate the most difficult problem of future productions.

6.2 Usability Testing

The second subject for research concerns the effectiveness of the final result. We believe that the multimedia proceedings provide most of the content of live talks and linear papers, with the advantage of random talk access and hypertext linking. Of particular interest is the use of this technology in the production of hypermedia books. It is often the case that a good textbook author is also a good speaker; hypermedia textbooks, based on the results of our experiment, are expected to have positive impact on the writing of future textbooks. We intend to produce hypermedia textbooks that are based on a combination of multimedia versions of lectures and hypertext textbooks using the same technology as in the DAGS '92 Multimedia Proceedings. But this proposition needs careful evaluation, and the cost/benefit of certain features (such as video loops and hypertext) needs to be further scrutinized in order to justify the substantial investment into production resources.

7. Conclusions

We believe that the DAGS '92 Multimedia Proceedings is a step towards academic publications that more fully reproduce the experience of a live conference or classroom. Even though it was published on CD-ROM, it can be viewed as a first step towards publishing conference proceedings and journals on the Internet. Currently such publications are one of a kind, expensive to produce, and with clear but unmeasured advantages over their traditional counterparts. Our experience highlights these shortcomings, but also suggests that, with further research, improved systems can be built with less effort, and greater rewards.

References

- [1] Rada, R. (1993) (prod. chair) *Proceedings CD-ROM of the First ACM International Conference on Multimedia, 1-6 August, Anaheim, California.*
- [2] MacSciTech (1992). *Proceedings CD-ROM of the 1992 MacSciTech Conference on Scientific and Engineering Applications for the Macintosh, 15-17 January, San Francisco, California.*
- [3] Johnson, D.B., Makedon, F. & Metaxas P. (1992). (eds.) *Proceedings of the 1992 DAGS Conference,, 23-27 June, Hanover, New Hampshire.*
- [4] Gloor, P., Dynes, S. & Lee, I. (1993). *Animated Algorithms*, MIT Press, Cambridge, Massachusetts.
- [5] Cormen, T., Leiserson, C. & Rivest, R. (1990). *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts.
- [6] Matthews, J., Gloor, P. & Makedon, F. (1993). *VideoScheme: A Programmable Video Editing System for Automation and Media Recognition. Proceedings of the ACM Multimedia '93, Anaheim, California.*
- [7] Gloor, P., Makedon, F. & Matthews, J., (eds.) Johnson, D.B., & Metaxas, P.T., (contr. eds.) (1993) *Parallel Computation: Practical Implementation of Algorithms and Machines*, TELOS/Springer-Verlag, Santa Clara, CA.