Report On Possibilities For A Photograph Database

by Adam Engst[1]
Consultant
901 Dryden Road #88
Ithaca, NY 14850

Introduction

As Gould Colman explained it to me, the transfer of the bibliographic information on the archival photograph collection to a computer database is currently open to a large number of possibilities. While there are advantages and disadvantages to programs on a number of machines, there are no external forces currently requiring a certain solution. I was retained to research the possibilities and present my findings, either recommending hardware and software combinations to test or recommending that the Archives wait several years before repeating this process.

I have gone through several steps to come up with this report. First, I tried to determine precisely the needs and desires of the Archives. Second, I researched the software possibilities on three different hardware platforms which are available to the Archives, the Macintosh series, the IBM PC line, and a general category of mainframe. Third and finally, I weighed the advantages and disadvantages of various combinations, adding in my knowledge about the Cornell community, the state of the software industry, and the history of some companies in particular. Most of the information below comes from my notes on telephone conversations held with representatives of the various companies.

Questions

Since the Archives is not locked into using any specific program or computer, I was left to figure out what sort of a system would best fit the needs of the department. As I see it, there are some requirements placed on any system by the size of the data, the nature of the data, the use to which the data is put, and the cost of the hardware, software, and programming time.

Size and speed

Gould told me that the Archives currently has between 50,000 and 100,000 photographs. Assuming one record in the database for each photograph, the system must be able to handle 100,000 records with decent searching speed. This was the first question I asked of the various database companies. Their replies must be taken at face value though, since the only way to really test the each system is to put 100,000 representative records into each and do some searches.

Keywords

Since a small number of photographs are the end result of any search through the database, the system must be able to handle a relatively large number of keywords, or else researchers will have difficulty narrowing down their searches. A number of databases require programming convolutions to be able to deal with a field containing an unknown, but potentially large, number of keywords. Such a limitation does not rule out a database, it simply downgrades it in terms of ease of setup and programming. In addition, selecting the keywords is an extremely important task which must be thought out carefully.

Graphical information

Bibliographic information is useful for providing a brief description of each photograph and locating it within a collection, but for a researcher who is trying to find a certain photograph, possibly out of hundreds of similar ones, bibliographic information will not allow that researcher to select a certain photograph with surety. A graphical method of describing the photographs would decrease the amount of time it would take a researcher to find the right photograph for the use he or she has in mind. The main possibility is displaying images on a videodisc. Few of the database systems can directly control a videodisc player. There are serious cost drawbacks to displaying visual information though, so inability to control a videodisc does not disqualify a database.

Initially, scanning the images and storing them on a CD-ROM would seem to be feasible because each CD-ROM can hold 600 megabytes of information. Unfortunately, CD-ROM is not feasible because a scanned photograph has an average file size of 300K, which, when multiplied by 100,000 photographs, would force you to use close to 50 CD-ROMs holding 600 megabytes each. In comparison, a single videodisc can hold 108,000 images.

Costs

The cost of the software are minimal in comparison to the costs of transferring the images to videodisc, although the purchase of expensive mastering equipment can reduce the overall costs. Another cost which cannot

be ignored is the cost of programming and setup with whatever software is decided on. As a result, ease of programming does play a financial role in the final decision as well.

So the questions that I asked each database company were as follows:

•Can your program handle 100,000 records with a fast search speed for a single record, say under 10 seconds as a worst case scenario?

•Can your program handle unlimited length text fields in an index (to retain searching speed) or is there a simple way around the program's inability to do so?

•Is there any way for your program to access images stored on a standard videodisc player?

•How hard would it be to set up your program with a simple interface for researchers who may be inexperienced with computers?

I also tried to get a feel for each company—how easy they would be to work with, how much help they would be if we needed any technical support, and whether or not they would still be in business in several years. These are intangibles, but potentially useful pieces of information.

A note before I get into the details. I've tried to write this so no technical knowledge is required to understand it. I'm sure that in some places I have failed because there is simply no other way to talk about certain features and actions of computers. In those places, I've included a footnote or tried to explain the term I use within the text. If at any time, you are confused reading this, please call me, and I will attempt to clear up the source of the confusion.

Hardware
The companies with whom I spoke have database programs that run on the Macintosh series of microcomputers, the IBM PC line of microcomputers, and (in the cases of Oracle and NOTIS) almost all minicomputers and mainframes. The Archives currently has several IBM PCs and clones and will be getting several Macintosh SE/30s shortly. In addition, I gather that the department has access to the mainframe resources of the library and the university. So existing hardware does not bias the decision.

With a few exceptions, all of the software packages I researched can handle the large size of the database without a loss in searching speed. Obviously, the minimum (and preferred) hardware configuration in each case does vary slightly, although some packages run fine on less powerful machines, which is a bonus since it will reduce the costs.

In general, and like all generalizations this one is not to be trusted completely, the Macintosh will be the easiest to set up and for both researchers and staff members to use. IBM PC clones have the advantage of being in the majority, although powerful systems are not really much cheaper than Macintosh systems. Microcomputers have the advantage (and disadvantage) of local control—if something goes wrong with the computer you can have it fixed quickly if necessary, whereas you must wait for another department to respond to your problem with a mainframe. On the other hand, if something with a microcomputer fails, you must deal with it, unlike with a mainframe, which will have a staff to deal with problems. Mainframes often suffer from poor interfaces as well, although there are ways of avoiding the poor interfaces.

Using a Mac and HyperCard with a videodisc is simple, while using a PC with a videodisc requires a special device driver, which is a small program that allows the computer to control the videodisc. Such programs are available, often from the videodisc maker, and there are also programmers who could write a custom device driver if necessary.

General hardware conclusions
Based on my experiences with the various types of computers and my knowledge of the Cornell user community, I recommend using a Macintosh. Macs are predominantly easier to work with in the setup phase, and they are far easier for inexperienced users to work with. Since the entire point of this project is to provide easy access to information, I think that the interface is one of the most important parts of the system, and better interfaces can be created on the Macintosh. In any case, my research covers all three platforms, and I hope that my recommendation of a hardware platform is born out by the software possibilities on the Macintosh. In addition, the Macintosh database companies were far more knowledgeable about controlling videodiscs, which is why I often have more information on the Macintosh databases.

Macintosh Software
Company: 1stDesk Systems
Program: 1stTeam
Hardware: Mac
Price: $795
  Their powerful relational database, called 1stTeam, is compatible with HyperCard and can store up to 255 characters in each field. Offhand, 255 characters doesn't sound like it would necessarily be enough for our keywords, but

perhaps their literature will shed more light on the subject. Otherwise, 1stTeam is certainly a possibility because it should be fast enough and can control a videodisc through HyperCard, although it is much less well-known than either 4th Dimension or Omnis 5.

Company: ACIUS
Program: 4th Dimension
Hardware: Mac
Price: $695

4D can handle 100,000 records with no problems, but it would have trouble with indexing. Without indexing, the search speed slows tremendously, but 4D cannot index its unlimited length text fields, which we would use for holding keywords. So setting up the keywords would be a little tricky in 4D. There are a number of different ways around for this problem, but they would require a bit more work programming. In the first French version, there was some kind of external command which could control a videodisc, although they may not still exist in the current version. There is a demo database, called Minifans, which we could look at if 4D turned out to be a likely candidate. Against 4D, I've heard that it is one of the slower databases for the Mac, which is a problem for this project. A test of its speed would definitely be needed before I could recommend it any farther.

ACIUS is one of the major database companies for the Mac, but I have been unable to get through to them at all, which may indicate mediocre customer support. While this is not a complete argument against using 4D, it doesn't bode well for future support needs. I can't really recommend them unless I can get through to talk to them.

Company: Blyth Software
Program: Omnis 5
Hardware: Mac and PC
Price: $695

Omnis 5 from Blyth Software certainly has the power to deal with 100,000 records, and it can be extended to do even more such as control a videodisc, although the representative didn't think such an external command had been written so far. Alternately, either Blyth could do it for us for free if it was small and fairly easy or an independent programmer might be willing to write such a thing for a fee. Omnis can index variable length text fields, so it would have no problem with a field containing a variable number of keywords.

If the software to control a videodisc was difficult to write or acquire in other ways for Omnis, it can work with HyperCard so that HyperCard uses the

Omnis database while acting as a front-end. However, Omnis can also create simple interfaces easily, so it should not be necessary to link the two together on that account. Omnis's language is supposedly English-like and easier than most database programming languages. An advantage of Omnis over any of the HyperCard extensions is that Omnis is a full-fledged database, and as such, can generate reports and display multiple windows, which would be good for displaying a number of records which met the search criteria.

Omnis needs a minimum of 1 megabyte of memory and is happier with a fast machine and more memory. Overall, I was quite impressed with the possibilities of using Omnis 5, since it seems to meet all the requirements and be fairly easy to work with in addition. The representatives have been extremely knowledgeable and responsive, unlike some of the other companies, such as ACIUS.

Company: Fox Software
Program: FoxBase Plus and FoxBase/Mac
Hardware: PC/Mac
Price: $395/$495

FoxBase can have up to 254 characters in text fields, and can search on unlimited length text fields, but they aren't indexed which slows the search. However, Fox claims that FoxBase can handle up to 1 billion records, and that it is the fastest of all the Mac databases by a great deal (some 30 times faster than 4D), although some of the PC databases come close in speed. Reportedly, a new version of FoxBase/Mac can use HyperCard's external commands (such as the ones to control a videodisc player) directly, which would be a major point in its favor.

FoxBase is generally accepted to be better than DBase III+ on the PC and the Mac version is correspondingly good, if not better since the Mac version can handle unlimited length text fields. FoxBase cannot control a videodisc, although it could work with a CD-ROM. Despite FoxBase's speed and file compatibility between machines, I think it is somewhat too limited in this situation because of its inability to index unlimited length text fields and its inability to control a videodisc. In addition, if it crashes for any reason, it will often corrupt the entire database rather than just losing the last record entered. This is a serious problem because you can never predict crashes.

Company: Odesta Corp.
Program: Double Helix II
Hardware: Mac

Price: $395

Double Helix cannot link to HyperCard and cannot control a videodisc, although there is a version that runs on Vax mainframes (for about $5000) which would help the speed and storage problems. Despite the fact that Double Helix can handle 100,000 records and has a simple method of programming, I doubt that this program is a real answer. Double Helix simply doesn't have enough to recommend it over any of the other major databases except its idiosyncratic programming environment, which may be easier than most.

HyperCard extensions

All of the following products require HyperCard, or one of two HyperCard clones, SuperCard or Plus, which provide the same basic features as HyperCard but with significant extensions. If a HyperCard system is decided on, it would be well worth the time to investigate creating the database in SuperCard or Plus rather than in HyperCard itself. The various extensions listed below may or may not work with SuperCard or Plus, although there is a good chance that they will. The areas in which SuperCard and Plus go beyond the capabilities of HyperCard include reporting, graphics, multiple windows, and color. Whether or not these features are worth moving away from HyperCard is another question entirely, and one that need only be asked if the Archives decides to go with HyperCard rather than one of the full-fledged databases.

Company: Answer Software
Program: HyBase (under HyperCard)
Hardware: Mac
Price: $150

Size and speed are not problems, since HyBase can handle up to 2 billion records and can usually find a single one in about 5 seconds. All fields are unlimited in size, or at least very large. The company claimed that it is not difficult but that some programming experience is helpful. Answer Software could set up the database for us if necessary. However, Gregory Crane at Harvard said that he used HyBase on Project Perseus for a while and found it very difficult to work with.

Dealing with the people at Answer Software was rather difficult and based on Gregory Crane's advice, I don't think that HyBase is a good possibility. It suffers from difficult set up, which is unnecessary for this project.

Company: Discovery Systems
Program: HyperSearch (under HyperCard)
Hardware: Mac
Price: $99

I have not yet received any information from Discovery Systems regarding their HyperSearch package, so I cannot make any specific statements for or against it. However, Library of Congress is using it in their American Memory project, and they seemed pleased with its speed and ease of use.

Company: KnowledgeSet Corp.
Program: HyperKRS (under HyperCard)
Hardware: Mac
Price: $195

HyperKRS works completely within HyperCard so it would be simple to design the database. Nothing else need be done in terms of setup except for generating the index, which is fairly slow, but only needs to be done once. A Mac Plus is all that is required for searching. HyperKRS was designed for CD-ROM, which accounts for its speed. I tested the demo software they sent me and I wasn't remarkably impressed. I had trouble finding anything, mostly because I was unfamiliar with the information for which I was searching. The speed was good but not great, but my Macintosh is not that fast, which is certainly an issue with this program.

On the whole, HyperKRS sounds like it may be the simplest of all the HyperCard extensions to set up initially. After that, I have no real numbers to compare its speed with HyperHIT or Xearch.

Company: NovaSoft Engineering Group
Program: GridFile (under HyperCard)
Hardware: Mac
Price: $195

NovaSoft has a sample application called ClipFile for GridFile which is being used right now to access pictures on clip art CD-ROMs. We would have to do the indexing and setup ourselves, which would be difficult without the aid of a relational database expert. GridFile is extremely fast, though, and is able to search any database for a unique record in 3 disk reads (certainly under 1 second).

The cons of GridFile include the fact that it requires 2 megabytes of memory and a fast hard disk; it does not provide as good data packing as some other databases, which makes the file larger; it slows down on smaller databases in comparison to the others; it doesn't support split files over two or more hard disks; and it would be hard to set up.

The pros of GridFile are that it is blindingly fast (faster even than some mainframe databases) and that it uses HyperCard as a front-end, which can then control a videodisc.

On the whole, I think GridFile is very powerful, but possibly too difficult to work with. There are other programs which provide similar speeds, but are easier to work with and require less hardware.

Company: SoftStream International
Program: HyperHIT (under HyperCard)
Hardware: Mac
Price: $195

Steve Hannaford, the technical support representative for HyperHIT said that HyperHIT has extremely fast searching speed (~1 sec) and can handle unlimited length fields with no problems. The information does not need to be textual—it could be pictures or sounds. The setup is not trivial but not that hard, and the HyperHIT system is entirely contained in external commands that work within HyperCard. Steve didn't think it would be hard for someone without database training to use.

The data file is external to the HyperCard stack which would control the videodisc and thus requires only a small amount of space. Another advantage to the external data file is that there could be a number of different interfaces since the HyperCard stack does not have the data embedded in it. Search time is usually under 1 second with a Mac Plus, and it would drop with a faster machine and hard disk. There is little speed degradation when the file size increases (3 hundredths of a second when going from 1000 records to 10,000 records). In fact, the videodisc might be the bottleneck, depending on how fast it can find each frame.

Steve Hannaford was very helpful and said that he wasn't getting many calls as the technical support person for HyperHIT, which could mean that people aren't having any problems worth calling about. The advantages of HyperHIT are that it is extremely fast despite what sort of machine it runs on, it supposedly isn't difficult to set up (although Gregory Crane will be testing it for Project Perseus soon and will have an opinion on its ease of use), and it will allow simple interfaces and videodisc access through HyperCard. Overall, HyperHIT sounds like a good possibility.

Company: The Voyager Company
Program: VideoStacks
Hardware: Mac
Price: $99.95

VideoStacks is a set of external commands to control a videodisc for a number of different videodisc players. There is a possibility that some of the external commands would be available from Apple free of charge or they might be distributed

with the videodisc itself. Some sort of videodisc drivers will be necessary.

Company: Xiphias
Program: Xearch
Hardware: Mac
Price: $???

Xearch is an external command for searching in HyperCard which Xiphias uses in their CD-ROM-based product, Time Line of History. It sounds like it would be fast enough and they do have a licensing agreement, although I don't yet have the details.

Initially Xearch sounds like it could be quite useful, although I don't have a sense of how easy or fast it is in comparison to HyperHIT or HyperKRS.

General HyperCard software conclusions

I think that all of the various packages mentioned above will probably provide HyperCard with the searching speed necessary to use the database. The main distinction then, lies in the ease with which each is set up. HyperKRS and Xearch are probably the easiest, with HyperHIT, HyBase, and GridFile lining up in increasing order of difficulty. More specific research and testing would need to be done to determine speed and ease of use in order to choose between the various extensions. HyperHIT may be the best compromise between speed and difficulty.

General Macintosh software conclusions

I am of two minds in this category. I think that HyperCard is a wonderful program (not to mention the fact that it is free with all Macs), and it will become integrated into the Macintosh hardware and system software in the next few years, making it even stronger and faster. On the other hand, it really is not a database and does not provide the features that a full-fledged database provides, such as reporting and fast searching. It may be necessary to use HyperCard in some fashion to facilitate access to a videodisc, which lends strength the cases of those database products that can link to Hyper-Card, such as 4D, Omnis 5, and 1stTeam. On the other hand, the structure of the proposed database is very simple and does not really require the full power of a relational database. In the final consideration, I think I would currently recommend Omnis 5 because of its power, flexibility, and ability to link to HyperCard, not to mention the quality of the customer support, with which I was very pleased.

PC Software

Company: Ashton-Tate
Program: DBase IV/Dbase Mac

Hardware: PC/Mac
Price: $795/$495

Neither DBase IV nor Dbase Mac have any internal way of controlling a videodisc, and the representative didn't know of any external ways either, although he thought one might be possible. In addition, neither can index on variable length text fields, which would slow them down a great deal for this purpose. Add these problems to the fact that Ashton-Tate is undergoing major problems as a company and has publicly announced that they will not be upgrading Dbase Mac at all, and you get a company to stay away from.

Company: Borland International
Program: Paradox/Reflex Plus
Hardware: IBM PC/Mac
Price: $725/$279

The Borland representative didn't think that either Paradox, the more powerful PC program, or Reflex Plus, a decent Macintosh database, could control a videodisc. It took several phone calls and some time on hold to get that much information, so I didn't pursue it farther. However, Tim at Turquoise Film/Video Productions (one of the mastering services) said that he was thinking about re-writing his custom database in Paradox because it was fast and fairly easy to work with. He also said that Paradox runs on a number of machines and is probably file compatible with Reflex. As a result, Paradox sounds like the best of the PC databases that I've looked into. Using Paradox would require some additional device driver to control the videodisc, but such a program might be available from a number of sources, including Turquoise Productions. Paradox won most of the speed tests I saw in the course of my research, so I would recommend it over the other PC databases based on what I currently know.

Company: DataEase International
Program: DataEase
Hardware: IBM PC
Price: $700

DataEase cannot control a videodisc, although it supposedly can interface with three scanners for using pictures. However the storage of those scanned images would be ridiculous and dealing with graphics on the PC is more difficult than on the Mac. DataEase does have long text fields which can be searched, although the representative wasn't sure about whether or not they were indexed, which is a major concern. I have a demo disk from them which may answer the indexing question, although I see nothing special about DataEase otherwise.

Company: Image Concepts
Program: C-Quest
Hardware: PC or Unix mainframe
Price: $6000 or $25000

C-Quest is a proprietary system for storing photographic information and controlling a videodisc. It has been around for several years, but doesn't seem to have a devoted following.

Clif Nickerson of Image Concepts was somewhat helpful, although his system is designed more for a stock photograph collection than a historical research collection. The main evidence of this is the way it uses synonyms of keywords, a method which allows the user to search on "Stream" and get "Brook" and "River" and "Run" and "Creek". Unfortunately this is not nearly as useful with proper names of people and places, since they tend to be specific. The only use I can think of it is use modifiers, so you could have Frank Rhodes walking, talking, shaking hands, or making a speech, and search on the action involved. I don't know if that is too much trouble to set up and key in or not.

C-Quest runs under Unix mainframes as well as PC clones. Under the Unix system, C-Quest can display 18 pictures at once; on the PC it can only display one at a time. Its speed is dependent on the number of subjects used in the search, but Clif said something about speeds of under 1 second, which he said was faster than the mainframe database Ingres (and he thought than Oracle). The C-Quest interface is menu-driven and not particularly good. It does not have a simple interface for researchers to use, although one is being proposed. Image Concepts will change, add, or remove fields from the menus for a nominal fee, which is not as good as setting it up oneself. C-Quest is not cheap, by any means, at $6000 for the PC version of the software and $25000 for the Unix version.

Clif said that the easiest way of getting images on disc were to buy a video camera and a writable videodisc, at which point you could do it all in-house. I suspect that quality wouldn't be as good, although there is no way to know without trying. He recommended making a 35mm film image in case the resolution of the monitors increased enough to make it worthwhile to re-master a videodisc.

C-Quest has an impressive list of clients, although I suspect that is from being the only game in town for 4 years, since no one else does this on the PC at all. Evidently, the Library of Congress system is slower than C-Quest, although that doesn't really

mean much without more details. C-Quest can control videodiscs from a number of companies, such as Sony, Pioneer, Philips, and Panasonic.

Overall, I find their system to be somewhat clumsy, expensive, and not really suited to the needs of the Archives. The Archives photographs have specific subjects without synonyms and need a very simple interface for researchers. While $6000 is not truly expensive in relation to the cost of mastering the videodisc, I think it is quite a bit more than you would pay for any other system. It might require less setup initially, although it is still a generic program that would require some customization. I cannot recommend it, especially since I heard from another consultant that the version of C-Quest at the United Nations was actually quite slow.

Company: Microrim
Program: Rbase for DOS
Hardware: IBM PC
Price: $725

Rbase will handle an unlimited number of records but the representative didn't know if it could handle an unlimited length text field. I didn't want to hold any longer to find out if it might be able to control a videodisc since the representative didn't think so. I see no reason to specifically recommend Rbase.

Company: Symantec
Program: Q&A
Hardware: IBM PC
Price: $???

Q&A does not have variable length text fields, but it can have large ones which are indexed so that the search speed doesn't suffer. The speed isn't good, though, at 15-20 seconds average, partly because Q&A is not a high-powered relational database. There is no videodisc access, although the representative thought that an external program might work. Considering the speed problem, I can't recommend looking any further at Q&A.

General PC software conclusions

I think most of the major PC database programs will handle the textual part of the database without trouble. However, it seems as though it will be more difficult to link the textual information in a PC database to the frame numbers of a videodisc. These device drivers do not seem to be readily available or supported by the database companies. For instance, the representative of Ashton-Tate knew nothing about linking to a videodisc, yet supposedly the Library of Medicine is using DBase III+. However, a device driver might come with the videodisc player. I also feel that it will be more

difficult within these PC programs to create a foolproof interface for researchers who are inexperienced with computers. I missed at least two major databases for the PC, Revelation and Nutshell, because I was unable to find phone numbers for them. However, I am not particularly worried that they are the perfect database because none of the other PC database companies had much of an idea what a videodisc even was, much less if their program could control it. The PC database companies were also much harder to reach on the telephone and much less willing to talk. If someone else turns out to be using Revelation or Nutshell, it would be worth checking them out. Otherwise, I think Paradox will be the best on the PC side.

Mainframe Software

Company: Oracle
Program: Oracle (runs under a HyperCard front end on the Mac)
Hardware: Mac/PC/minicomputers/mainframes
Price: variable depending on version—from $299 to $1299

Oracle for the Macintosh is a port of the most popular database program in the world. It retains complete compatibility with all other Oracle databases on all other machines, which is a plus if this data will be shared with other people. In addition, a Macintosh running the HyperCard front-end to Oracle can use any Oracle database on any machine. Because HyperCard is the front end to the actual database, Oracle can control a videodisc through HyperCard. Should the Archives wish, they could probably find a mainframe or minicomputer on which they could use Oracle.

Oracle's advantages are speed, portability, and ease of use with HyperCard, although it might be a bit more expensive than the Archives would want initially. There is a Developer's Version for the Mac for $299, which would allow us to test its capabilities (with the only limitation being that this version cannot link to other Oracle databases on other machines). The main disadvantages to Oracle are that it is potentially more expensive (although the Macintosh version is quite cheap) than other databases, and that it may simply be too complicated for the relatively simple database information we have. I gather that setting up a database in Oracle is not all that easy. In addition, I've heard that Oracle for the Macintosh is not that fast and occasionally does strange things to data files.

Oracle as a company is excellent, with toll free support and guaranteed stability. They are the

largest database company in the world and the third largest software company in the world.

Company: Northwestern University
Program: NOTIS
Hardware: IBM mainframe
Price: free

NOTIS has a number of advantages, although it also sports major several disadvantages. NOTIS is currently installed and running in the library, so there are no added software or hardware costs to the system other than a Macintosh and videodisc from which people can search in the Archives. It is relatively fast and can certainly handle another 100,000 records in its database. NOTIS has the advantage of being accessible from anywhere on campus, but researchers may not use it unless they can also see the videodisc images because it is difficult to search for photographs based solely on bibliographic information. It has been in use at Cornell for some time now, so many people are familiar with its interface, although its interface is also one of its main disadvantages.

Searching and moving between the various results of a search in NOTIS is difficult and completely not intuitive. Its other main disadvantages include the fact that it would be very difficult to link it to a videodisc, if it is possible at all, and the problem of portability of data since NOTIS does not have the ability to export its information to another program, something which all of the microcomputer databases can do and which is very important for future expansions or modifications. It might be possible to sidestep NOTIS's poor interface with a HyperCard interface currently being worked on at Mann Library. In addition, there is a commercial product that will be available soon from Texas A&M and Apple, called MacNOTIS, which also provides a better interface to NOTIS. Because both of these products use HyperCard, it is theoretically possible to have the HyperCard interface control a videodisc while using the information from the NOTIS database. Howard Curtis of Mann Library thought that this was possible, although extremely clumsy and prone to break whenever either NOTIS or HyperCard changed much. Other people thought that it would be an unworkable situation even if it was theoretically possible. Howard also said that it might be possible, though difficult, to program HyperCard to download records from NOTIS to the Mac, which would allow the records to be used by microcomputer databases.

NOTIS is an easy solution because it requires no new hardware or software, but putting the records into NOTIS removes them from a certain level of accessibility. It would be difficult and clumsy to attach a videodisc to a Macintosh running one of the HyperCard interfaces, if it is indeed possible at all. More research would need to be done to determine the reality of such a setup. Even worse, it would be hard to transfer those files to any microcomputer system. However, there are some ways of moving from microcomputer databases to a format which NOTIS can read, which points towards putting the records into a microcomputer database first, and then, if there is interest, transferring a copy to NOTIS. As much as NOTIS seems like the simplest solution, I don't feel comfortable recommending it given the possibility for videodisc access and the inaccessibility of the data once it is in NOTIS. I realize that NOTIS data can be shared by other mainframe cataloguing databases, but they don't (on the whole) provide the kind of features that microcomputer databases do. Being able to move data between systems is important, and customized mainframe databases are a blockade to such a move.

Videodisc Mastering Services

Company: Image Premastering Services
Program: Videodisc services
Hardware: NA
Price: variable

Image Premastering Services claims they are known for having the highest image quality for still frame transfers. Some of their main clients have been the United Nations, the Library of Congress, the Mayo Clinic, and the American College of Radiology. They can handle absolutely any original—for the Library of Congress they laid down 30,000 glass plate negatives without cracking any. Of course, slides are the cheapest method, and run anywhere from 55¢ to $1.35 per slide. Other original media are correspondingly more expensive, although presumably the cost goes down with quantity. In addition, there are some basic initial costs which cannot be avoided. These costs total $3150, although that is minor compared to the cost of transferring 100,000 photos to the disc.

$2000 for the master disc
$500 for a check disc
$150 for the videotape
$150 for a duplicate/backup, kept at their site
$350 as a basic setup fee

They claim that Stokes is mainly a slide copy service and makes a 35 mm film negative, which is a second-generation picture of the original. If the

image is from a print, then the videodisc image is third-generation picture and suffers correspondingly in quality. Stokes does color-correction, so the colors may be bright, but they are likely to be inaccurate. Stokes is also generally cheaper because everything is automated in their process.

Image, on the other hand, is specifically dedicated to mastering videodiscs and they have patented technology for the process. They use a 2 foot lens over a 12 foot optical bench, which gives them two advantages.

1) The light comes in at a perfect 90° angle, which gives much better edge definition to the image.

2) They use an aerial image transfer, which somehow projects the image so that there is no film grain in the resulting videodisc image. It also allows them to easily perform custom sizing.

The Image representative recommended the Mac and said that a military project used HyperCard with Oracle. He had heard something about 4D, but didn't know of anyone who was using it. He didn't recommend using the PC at all because the hardware is more expensive and is harder to set up the software to interface easily with the videodisc.

Company: Stokes Mastering Services
Program: Videodisc services
Hardware: NA
Price: variable

I spoke with John Stokes and Jim Couch of Stokes Mastering Service. In regard to costs, Stokes estimated that a basic image transfer of positive images would be somewhere between $2 and $3 per image, although his estimate for complete costs (ie. in-house handling and database work) was closer to $4.50 per image. The work is done on-site and includes a person to come and do it with Stokes's somewhat specialized equipment. There is not much difference between Stokes's doing the work and it being done in-house except for the fact that he claimed they had higher quality control, which is fairly likely. I suspect this is somewhat cheaper than Image Premastering's prices, although not by as much as I had originally thought.

They have a number of projects going on, the most notable of which is for the Library of Medicine, whose cost was about $2.40 per image. That price is slightly inaccurate because it was a test run in some ways and the library got two sets

of negatives and slides for each of 70,000 images. The person to talk to at the Library of Medicine is Lucy Kiester, phone number 301-496-5962. The Library of Medicine is using a PC with DBase III+ for their database. Stokes claimed that videodisc access was incredibly simple with any database and that you didn't need a custom driver, but he did admit that you had to write some software. The Library of Congress is also working with Stokes, which is curious since the representative at Image Premastering said that the Library of Congress was working with them. Perhaps there are two different departments? In any case, Stokes claimed that the Library of Congress is using some in-house computer system rather than an off-the-shelf software package. Stokes said something about how that was their policy. This is not necessarily true since the American Memory project is using a Macintosh and HyperCard.

One advantage of Stokes's method is that you can get negatives of each image as well, which allows you to reduce handling of the original images by making additional negatives. The Library of Medicine uses these negatives for public access to avoid giving out their originals.

Interestingly enough, Stokes said that more people are doing the imaging first, then the database work, partly because Stokes can transfer the images to videodisc faster than the database can be set up. I was unsure about the real reasons for this, but they could be determine by talking to some of the people Stokes referred me to.

As far as hardware goes, Stokes sounded like he doesn't really know very much about the Mac. He claimed that the Mac had no advantage over the PC in ease of use if the software was designed well, although I disagree with that rather strongly. Based on a number of years of working with novices on both systems, the PC is less intuitive and clumsier than the Mac when it comes to user interfaces. In any case, Stokes has developed a database package under Informix (which is Oracle compatible, or so he said) which runs on a number of different machines. If we decide to use Oracle, we could buy the Oracle package and then Stokes would provide us with his custom database for only the cost of support. This is curious because he could very easily create a stand-alone Oracle database and then just sell that (or give it away if he wanted) without the customer having to buy their own copy.

Stokes wouldn't really comment on Image Premastering except to give me the name and

number of Bill Perry (202-857-7537) at National Geographic, which did independent tests of both Stokes and Image Premastering (and one other, actually whose name Stokes did not mention). In addition, Stokes claimed that their quality has improved since then. Evidently, the American College of Radiology had 11"x14" X-rays which Stokes claimed were optimized for the Image Premastering system and those came out better.

As far as quality goes, any transfer from a positive image will lose quality in the transfer process, much as copying a tape or videotape loses quality. The contrast of a videodisc is usually around 20 to 25 with a maximum of 45, whereas a transparency is about 1000 and a slide about 250. Thus a great deal of contrast is lost when going to videodisc in any case. The transfer to a negative reduces this contrast lost by spreading out the contrast rather than clipping it, although it also puts it through several generations of imaging. Stokes is aiming at a contrast factor of two to three times better than high definition video, which is as good as a monitor will get in the near future. He said that it is very difficult to match the original exactly, and that matching the original better is their main task right now.

Company: Turquoise Film/Video Productions
Program: Videodisc services
Hardware: NA
Price: variable
    Turquoise said that they can provide anything up to a turn-key system. Their background is in motion picture processing, and they moved from that to providing software and hardware as well. Their database is an in-house one currently but it can import and export to a number of other formats. They are thinking about re-writing in Paradox, which can also run on a number of different machines. They charge an average of $2 per image, including hardware, and they can shoot either in St. Louis or on-site. They use a special motion picture film to get better quality images, but I have no sense how the quality of their images compares to the quality of either Stokes's or Image Premastering's images. It doesn't seem that there is anything remarkable about Turquoise in relation to the other two mastering services, but should the Archives decide to have a videodisc mastered externally to Cornell, it would be a good idea to talk more specifically to all three companies.

Other Cornell Projects
    I spoke with Anne Carnell about the project in the University Photography Department, and she said that they are having someone in Publications develop an in-house program. This program will catalog and store all of the information on their photographs, but it will also provide billing, usage, and reporting capabilities. They didn't think one of the commercial programs could provide all that, something which I doubt, given the power of some of these databases. They are looking at videodisc in the near future, for much the same reason as the Archives, perhaps in the next year or so. She didn't seem to have a wonderful grasp on what is entailed with the entire technology, since she didn't know about the problem with file sizes for scanned images, and she didn't know how many images could be stored on a videodisc.

    I also spoke with Dave Watkins, who is the head of Media Services and in doing so found my way back to the original Stokes project mentioned in the memos to and from Chris Pelkie. They are currently selecting slides, negatives, and prints for a free sample videodisc to be supplied by Stokes. Partly to test the quality of Stokes's service, they are trying to assemble a number of different types of images for inclusion. If the Archives wishes participate in this project to see how it works out, they should contact Dave Watkins. He is looking for up to three hundred of the most difficult type of images. The deadline for submission to this project is January 1st, 1990, which is fast approaching.

    Other departments that may be interested in this project include the Department of Entomology, which has 30,000 slides that they wish to use for diagnostic work, obviating the need to go to the slide collection itself. These slides must be of the highest quality because of their use and the fact that the disc will be sold to other universities. Plant Pathology and Veterinary Medicine may wish to do similar things. Evidently the Hotel School has a videodisc of wine labels and the Vet School and the Law School are still looking into the possibilities of some sort of image database on a videodisc.

    If the Archives wishes to look more closely at the specifics of a videodisc system in future, I strongly recommend that the department provide a number of photographs to Dave Watkins for this test project. The project will provide a videodisc which can be shown to potential donors and with which we can test the pros and cons of various software packages. Such an opportunity should not be passed up lightly!

    I met with Margaret Webster, who runs the Architecture School's Slide Library. She has for some time been planning a videodisc project to keep track of the 350,000 slides in the library. They hired an outside consultant, Nancy Humphries of ETECH, to research the possibilities and provide a system. Margaret has been putting records into the database and will be setting up the pilot project after the Slide Library moves in January

of 1990. They will be using a PC database called btrieve/xtrieve (with which I'm not familiar because it was never compared in the literature with the more well-known databases) along with a specialized graphics board in the PC that will allow them to manipulate the images electronically. Manipulation of images is something which I did not explore particularly because of the expense involved, but is certainly a possibility for the Archives. The question that must be answered to justify the cost of such a board is the use to which these photographs are being put. If the photographs are ending up in publications designed and executed on a personal computer, or the images must be frequently manipulated, then a graphics board makes sense. However, if the publications in which these photographs appear use traditional methods of production, then additional 35mm negatives would be more useful. Margaret Webster knows more than most people on campus about videodisc systems because their system has been in the research phase for close to two years now.

The final person from Cornell with whom I spoke was Mike Oltz from the Interactive Multimedia Group. He gave me some bits of information that may be useful. He thought that the Architecture School and the History of Art department were looking into something similar and were probably working along different lines. As it turns out, History of Art has dropped their project entirely, whereas the Architecture project is the closest to reality of any of the ones I've heard of. The exception to this is the Medical School, which has a system for pathology training using a Macintosh pseudo-database called Guide. They started out with no funding at all and now have close to five million dollars of computer equipment, which does lend hope to the Archives getting funding for a videodisc.

The Interactive Multimedia Group has a sample videodisc from Image Premastering Services, which is currently lost, but Mike will try to find it and let me know via email. He mentioned that Revlon, the makeup people, are also doing something like this. If we wanted to learn more we should talk to the advertising photography department.

There is no reason to assume that any of the other systems are better than a system the Archives could come up with, although the ability to perform information transfer might be useful in the future to avoid re-keying records. Similarly, standard information in each record would help in translating the records from another system when other departments wished to archive various photographs. Database information can be shared between PCs and Macs without too much trouble, so the specific machines used by different departments should not really matter, although the ability to transfer between the various databases matters.

Cornell Coordination

One problem that has come up time and time again in my research is that there are a number of Cornell departments working completely independently on similar videodisc projects. While such a lack of communication is not unusual at Cornell, it is regrettable, particularly in a field such as this where the information really is fairly finite. It would be extremely useful if there could be a single person who would, if nothing else, have copies of all the various pieces of information collected by the different departments. That way, whenever anyone was thinking about starting such a project, the information would be more or less at hand and would include names of the people at Cornell who are good resources. This person would merely disseminate information and would refrain from making any recommendations as far as hardware or software go in order to avoid the politics.

The Interactive Multimedia Group would seem to be a logical group to coordinate or various videodisc information, but because they exist completely on soft money for specific projects, they are not set up to handle any sort of coordination. Geri Gay said that Media Services was one place coordination could come from, and some part of CIT Services would be another. People to talk to in CIT include Larry Fresinski, Donna Tatro, and if all else fails, Stuart Lynn.

Another area in which the various departments could pool resources would be in setting up facilities at Cornell for transferring images to videodisc. I gather that there are close to a million images at Cornell that could be put on videodisc if the process was cheaper and easier. Dave Watkins in Media Services is the person to talk to about such a project. Margaret Webster in the Architecture Slide Library would also be very interested.

Other Non-Cornell Projects

I've found names of people at other institutions who have done something along these lines or are thinking about it. Talking to them might help the final decision because you can get an opinion from someone in a similar position. I did not get more detailed information from these people since it is often easier in this situation to use academic channels for sharing information, and the Archives already has contacts in some of these institutions, whereas I would be going in cold. So it doesn't make sense for me to talk to everyone immediately unless it seems that they have something important to offer to the decision-making process right now. That step can come if and when the Archives decides on a specific system or type of system. If I know of a way of contacting the people below, I've mentioned it. Most of this information comes via electronic mail, so I can ask for additional contact information if desired. My apologies for the lack of organization, but no method proved

itself better than formatting for maximum readability.

Elizabeth Wood mentions joint project between the Emergency Medicine and Radiology Department of Los Angeles County and the University of Southern California Medical Center that had their mastering done by Image Premastering.

Elizabeth H. Wood
Computer Services Librarian
Norris Medical Library
University of Southern California
ewood%phad.hsc.usc.edu@usc.edu

The AV Department of Hornbake Library at the University of Maryland is developing a videodisc in conjunction with the National Agricultural Library. I wonder if this is the Forestry Service collection mentioned above.

David Austin mentions two other projects. First, Andrew Eskind at the Eastman House is working on something to do with a videodisc. Second, Jim Sheldon at the MIT Media Lab is working on a videodisc of Edweard (sic) Muybridge motion pictures in conjunction with the Addison Gallery of American Art, Phillips Academy, Andover, MA, 01810. Finally, David says "Also, make sure you check the SN/G: Report on data processing projects in art (1988). It is not yet on-line but available in hard copy, maybe even at Cornell. It is a list of projects registered with the Scuola Normale Superiore, Pisa, Italy and the Getty Art History Information Program, Los Angeles."

David Austin
U29716@UICVM

Jim Sheldon
jls@media-lab.media.mit.edu

The AVIADOR (Avery Videodisc Index of Architectural Drawings on RLIN) project at the Avery Architectural and Fine Arts Library at Columbia sounds very similar to what the Archives might want to do. In addition, RLG is working on a way of linking a videodisc to an RLIN terminal, which would be very interesting. Janet Parks sent me a copy of their literature on the videodisc system.

Janet Parks
Curator of Drawings
Avery Architectural and Fine Arts Library
Columbia University
New York, NY 10027
212-854-6738

Jane Kleiner mentions several videodisc projects,

one of which is the Emperor I collection done by Ching Chi Chen at Simmons. It is quite sophisticated and includes sound as well. She thinks MIT has an architectural collection on videodisc and adds that the National Agricultural Library has a collection of historical photographs from the Forestry Service on videodisc.

Jane Kleiner
notjpk@lsuvm

Lennie Stovel mentions that there would be more information in the Library of Congress' literature on their Prints and Photographs Division's videodiscs, although he does not give a specific contact.

Lennie Stovel
Library Systems Analyst
Research Libraries Group
bl.mds@rlg.bitnet

David Finkelstein at Stanford University Academic Information Resources says that they are currently digitizing a large slide collection, which will eventually reside on videodisc. They are currently using HyperCard because of its ease of use but are looking into more powerful database programs such as Ingress, which is in use but is not necessarily well-liked at MIT's Athena Project.

David Finkelstein
Academic Information Resources
Stanford University
davef@jessica.stanford.edu

Steve Cisler from Apple Computer has available a technical report on basic videodisc production. It is called "Multimedia Production: A Set of Three Reports". and includes: "Casual Multimedia Production", "Videodisc Basics", and "Videodisc Production of the Visual Almanac". It was done by Apple's Multi-media Lab for the production of the forthcoming Visual Almanac. It is written for the non-technical person and includes mastering costs, sources of replicators, techniques. Steve mentioned that some people at the Visual Resources Association felt that the image quality was not high enough for scholars. He also said that there is a new method of distributing videodisc images with a certain type of network called Broadtalk. Steve can be contacted for more information on Broadtalk, and he will send a copy of the videodisc report to anyone who sends him a request on university letterhead along with a self-addressed mailing label. I have the report and recommend it highly for anyone who is actually starting on the specifics of producing videodisc.

Steve Cisler
Apple Library

10381 Bandley Drive MS: 8C
Cupertino, CA 95104
sac@apple.com

In response to another question, Steve Cisler gave
the address of several replicators for videodiscs which I
have yet to contact. These are as follows.

Crawford Communications
506 Plasters Ave
Atlanta, GA 30324
404-876-8722

Pioneer Communications
1058 E 230th St.
Carson, CA 90745

3M Optical Recording
223-5S 3M Center
St. Paul, MN 55144
612-733-2142

Cynthia Read-Miller at the Ford Museum is using
a videodisc and microcomputer catalog set up by a
company called Argus.

Bernard Littau at UC Davis is putting together a
radiology learning system for the veterinary school. He
warns about several problems with videodisc production.
First, it is extremely expensive to master the first disk
from videotape. It requires a great deal of staff and
equipment time just to make the videotape, and then the
frame numbers on the resulting videodisc must be
matched with the photograph database. Second, he feels
that videodiscs are more suited to video sequences since
that was what they were originally designed for, and,
videodiscs have limited resolution for displaying still
images in comparison to a digital image stored on CD-
ROM and displayed on a computer monitor. Unfortu-
nately, as subsequent conversations with Bernard proved,
100,000 photographs is simply too many to put in CD-
ROM format because it would require 30 or more CD-
ROM discs. Third, he said that when they did the
videodisc, they were forced to use three frames for each
image by the mechanics of the process of transferring the
images to videotape. As such, they ended up using three
frames of the videodisc for each image, reducing the
storage capacity by three. Using three frames per image
had the advantage of safety if one or two of the images
were bad for some reason or other. I wonder if this
problem appears if a commercial mastering service does
the work since Bernard's project was done in-house, I
believe.

Bernard Littau
VM Radiological Sciences
School of Veterinary Medicine, University of

California
Davis, CA 95616
916-752-0184
Internet: vmrad@ucdavis.edu
BITNET: vmrad@ucdavis

There is an integrated image database package
with its own Programmers Application Language
developed by PCM, Inc. The package is called PC
ALBUM and runs on an IBM-PC.

PCM, Inc.
8330 Boone Blvd. Suite 430
Vienna, VA 22180
703-356-4600 or 800-654-5845

Ernst Robl recommends an expensive system
called INMAGIC. It is sold by a company of the same
name in Cambridge, MA. The Los Angeles Public
Library uses it to catalog their extensive photograph
collection and speaks very highly of it. The version
which runs on IBM-PC type microcomputers is $1000,
and there is a version which runs on VAX mainframes as
well. Ernst says that INMAGIC allows a considerable
amount of individual configurations and handles variable
length data well. It can accept data from other sources,
which is good for compatibility reasons. In fact, the LA
Public Library has staff members do the cataloguing on
laptop computers in the stacks rather than bring the
collection out to a terminal.

Ernst has served a couple of terms as chair of the
Picture Division of the Special Libraries Association and
has authored an introductory book on picture
librarianship, *Organizing Your Photographs* [Am-
photo,1986]. In connection with the above, he has
visited a large variety of institutional and commercial
picture collections. (The Picture Division no longer
exists as an individual entity with SLA, but its interests
have been taken over by several other divisions.) His
book points out some general issues to consider in the
cataloging of photos, although the sections on computers
are fairly basic because of its audience.

Ernest H. Robl
Systems Specialist (Tandem System Manager),
Library Systems
027 Perkins Library, Duke University
Durham, NC 27706
(919) 684-6269 w; (919) 286-3845
ehr@ecsvax

Russell Grau mentions a project he worked on
with a company called Laser Recording Systems. The
project consisted of taking images, scanning them onto a
WORM drive, and then accessing the images via biblio-
graphic information stored in a database. The whole

thing ran on IBM-PC type microcomputers. The person Russell worked with was named Tom Corsten, but he may not be there any more.

Laser Recording Systems, Inc.
270 Sparta Ave.
Sparta, New Jersey 07871
201-729-3055

Russell Grau
916-920-9092

Gordon Fair mentions that Oracle for Macintosh can work with SuperCard as well as HyperCard. Unfortunately, SuperCard is much slower than HyperCard and a project like this does not require SuperCard's color and animation abilities.

Gordon Fair
gf07+@andrew.cmu.EDU

There is a package called Videodisc ShowMaker that will allow you create a database of entries with keywords and then search over the fields in the database. It is intended for a substitution for a slide projector in classes that require many images (it was originally designed for graphic arts education). It is a collection of stacks for the novice HyperCard/Macintosh user; it interacts with the videodisc players using videodisc drivers from Apple; and it can handle large databases of images (around 2000). The current version of Show-Maker uses the HyperCard Find function and is not very fast, but it does what it is supposed to. It is going to be released in December through a company called Ztek, which supplies interactive-video software. If interested, get in touch with them or with the professor in charge, Mark Sanders. My only problem with VideoDisc ShowMaker is that it is definitely not fast enough for 100,000 images. Some sort of HyperCard extension software would be required to increase the search speed.

Mark Sanders
msanders@vtvm1.cc.vt.edu
msanders@vtvm1
(703) 231-6480

Bob Samson at the University of Texas at Arlington might be setting up a system using the Series 2000 Laser-Optic Filing System from TAB Products. I don't know much about this project, but I gather that the system is a digital system, so I don't know how they are getting enough storage space for 350,000 photographs even though it comes with either a 5.25 or 12 inch optical disk. The system also includes a computer (no indication of what kind), a scanner, a high-resolution monitor for viewing the images, and possibly a laser printer for creating hard copy. He would be using this to store 350,000 images from the photographic archives of a local newspaper. Since many of the photographs are quite old, he wishes to avoid physical contact when possible.

Bob Samson
University of Texas at Arlington
B366RCS@UTARLVM1
817-273-3000

Lucy Kiester (phone: 301-496-5962) at the National Library of Medicine is finishing up a videodisc project and used Stokes Mastering Service to transfer her images to disc.

Bill Perry (phone: 202-857-7537) at the National Geographic Society has also done some work with Stokes Mastering Service.

Mike Segel recommends using the Informix database (which runs on many different microcomputer and mainframe systems) because it allows you to store BLOBs (Binary Large Objects) in the data base. I don't know if storing images as BLOBs would take up less space, but if it didn't the space requirements would be prohibitive.

Mike Segel
segel@quanta.eng.ohio-state.EDU

Ed Heath is an intern at the Library of Congress and is working on the American Memory project. He sent me quite a bit of information on American Memory. The project uses a Pioneer Laservision Player and the machine is controlled by a Macintosh IIx using Hyper-Card and Discovery Systems' HyperSearch. The photographs were mastered onto the videodisc before the project started for another reason so Ed didn't know too much about the specifics. American Memory deals with keywords by using free text searches with a "Visual Materials" thesaurus developed by the Library of Congress. Otherwise it is The American Memory setup also includes a CD-ROM.

Ed Heath
Special Projects
University Computing
George Mason University
Fairfax, VA 22030
(703)323-2941 • EHEATH@GMUVAX

Lloyd Davidson tells of an article in *BYTE* magazine (January 1988—("A Better Way to Compress Images", BYTE 13/1, 215-218, 220-223)) in which a method using fractal geometry achieves graphic compression at ratios of over 10,000 to 1. Such compression ratios would easily allow a CD-ROM to store a great many images and would make them far more feasible for

extremely large image collections. He also mentions a second article about the same researchers in the November 4, 1989 issue of the *New Scientist*, p.40.

Lloyd A. Davidson
Seeley G. Mudd Library for Science and
Engineering
Northwestern University
Evanston, IL 60208
L_Davidson@nuacc.acns.nwu.edu

Overall Conclusions and Comments

Taking everything I currently know into consideration, I would recommend using a Macintosh SE/30 with 2 megabytes of RAM and at least an 80 megabyte hard disk. That will satisfy any of the programs and leave plenty of room for expansion. As far as the programs go, I currently recommend Omnis 5 with HyperCard to provide videodisc access if no external routine for this are easily available. Of course, testing would be necessary before a final decision. No matter what software is used, there will be a fair amount of programming time necessary to set it up and get it running. In addition, the time it will take to enter 100,000 records into the database will be considerable. I cannot make any recommendations as to the mastering services because I do not have enough hard evidence to work with. Ideally, Cornell would set up its own facility for transferring images to videodisc.

Since transferring images to a videodisc is so expensive, I can only recommend that the Archives search for donors. In the meantime, deciding on a database and starting to enter the data would be useful whether or not a videodisc is ever monetarily feasible. Once the data is entered into a microcomputer database, it would not be too difficult to move it to another system, should a standard appear or merely a better method of working with a videodisc. I see no reason to wait on starting the database for this reason, and the cost of transferring images to videodisc will not drop much in the future, if at all. If Cornell set up a facility to transfer images to videodisc, it might be cheaper, although one never knows.

An important procedure for the moment is to think about the format of the database. The fields of bibliographic information are set, but some thought must be given to the keywords. The problem is with keywords because there are simply too many different possibilities, since everyone thinks different keywords are important. The Architecture School has a thesaurus, which helps, but they will still need to add some keywords and ignore others. Three to four levels of hierarchical keywords (ie. Post 1905 - People - Professors - Professor Kaplan) are probably as detailed as you want to go at first, since it is too easy to come up with keywords which only make

sense to the cataloguer after four levels in the hierarchy. A good way to figure out a system is to find a picture and then work backwards so you see what steps you went through to find it. Hopefully this will be solved easily by simply using the categories of information that are already set up for the current system. Looking at the Architecture School's system might be helpful in determining the number and type of keywords.

Some more questions to keep in mind while designing such a system include who will be using the system, how will they be using it, and what is the end result going to be? Answering these questions before the database is designed will help in the design stage to make sure that the database is really set up correctly for the purposes at hand.

As a caveat, let me merely mention the problem of copyright. I assume that Cornell owns the copyright on the photographs in the Archives, but if not, it is technically a breach of copyright to transfer these images to a videodisc.

Further questions

I have left a number of questions unasked in my inquiries because of the preliminary nature of the investigation. Most of these deal with the specifics of the videodisc access, since that is the great unknown in the whole project. The questions into which I have yet to delve are as follows (along with my current opinions).

•Is the videodisc access feasible soon or farther in the future?

My feeling is that the videodisc access will be very nice once it is set up, but it will be expensive to master the disc. Assuming the prices quoted by the mastering services, a videodisc of 100,000 images could easily run over $300,000, if not more. Unless a munificent source of funds appears, I suspect that the videodisc will simply be too expensive for the moment. I don't think prices will drop much in the future, because the imaging and material handling work involved will remain more or less the same. In the event that several hundred thousand dollars should become available, the database should be able to support a videodisc. Otherwise, the files would have to be exported and imported into another program, a procedure which can be difficult and time consuming.

•What would be the best videodisc players for the Archives's purposes?

I didn't check into this at all, although I know there are a number of models that would work with either a Macintosh or a PC. All the prices that I've seen are in the $2500 range. The writable videodisc which you can

record to directly is quite a bit more expensive, at $13000 to $15000.

> •Would the Archives wish to use an outside mastering service or set up an in-house mastering service?

The outside services are likely to be more expensive, although they would also probably give higher quality results. If there is enough interest, Media Services might set up a videodisc mastering system at some point, which would be ideal. Their system might be somewhat cheaper and would certainly be closer. In any event, price and image quality seem to be directly related, so the nicer your pictures look, the more it will cost. If in-house mastering is deemed unfeasible, then some testing between the three mastering services would be in order.

In addition to these specific questions, I'm afraid that this report brings forth more questions yet that can only be answered by careful thought on the part of the Archives. I've made recommendations and hopefully discovered sources of information that will help answer these questions, but much more work will need to be done before a project like this becomes reality. For instance, it took Margaret Webster almost two years to start her pilot project. If I can be of assistance at any later point in the project, please feel free to call me.

Appendix
Addresses and phone numbers

> 1stDesk   Systems
> 7 Industrial Park Rd.
> Medway, MA 02053
> 508-533-2203
> 800-522-2286
> Makers of 1stFile and related programs.

ACIUS
20300 Stevens Creek Blvd
Cupertino, CA 95014
408-252-4444
Makers of 4th Dimension. They are very hard to reach.

Answer   Software
20045 Stevens Creek Blvd.
Suite 1E
Cupertino, CA 95014
408-253-7515
Makers of HyBase, a HyperCard database extension.

Ashton-Tate
20101 Hamilton Ave.

Terrance, CA 95052
213-329-9989
213-329-8000
Makers of DBase IV and DBase Mac

Blyth   Software
3655 Campus Dr.
San Mateo, CA 94403
415-571-0222
Makers of Omnis 5. I spoke with Jennifer Blome.

Borland   International
1800 Green Hills Road
Scotts Valley, CA 95066-0001
408-438-8400
Makers of Paradox and Reflex Plus

Anne Carnell
1159 Comstock Hall
Cornell University
Ithaca, NY 14853
255-7675
Anne works in the University Photography Department.

Howard Curtis
Mann Library
Information Technology Section
Cornell University
Ithaca, NY 14853
255-9570

DataEase   International
7 Cambridge Ave.
Trumbull, CT 06611-9983
203-374-8000
Makers of DataEase

Discovery Systems
7001 Discovery Blvd.
Dublin, OH 43017
614-761-4197
Makers of HyperSearch, a HyperCard database extension.

DucSoft
238 Columbus Ave
Sandusky, OH 44870
419-626-6797
Makers of Applications and Routines for 4th Dimension, but nothing for videodiscs.

Fox   Software
27493 Holiday Lane
Perrysburg, OH 43551
419-874-0162
Makers of FoxBase+ and FoxBase/Mac

Image Concepts
P.O. Box 211
West Boylston, MA 01583
508-481-6882
Clif Nickerson, Marketing Manager. Note that the
phone number is different from the old literature.
Image Concepts make C-Quest, a videodisc
program for the PC and Unix boxes.

Image Premastering Services
1781 Prior Avenue North
St. Paul, MN 55113
612-644-7802
A videodisc mastering service.

Interactive Media Center
Geri Gay or Mike Oltz
Cornell University
Ithaca, NY 14853
255-5530

KnowledgeSet Corp.
888 Villa St., Suite 500
Mountain View, CA 94041
415-968-9888
Makers of HyperKRS and HyperIndexer,
HyperCard extensions.

Microrim
3925 159th Ave NE
Redmond, WA 98073-9722
206-885-2000
Makers of Rbase.

NovaSoft Engineering Group
2343 Ridgewood Ave.
Edgewater, FL 32032
904-423-5189
Makers of GridFile, a HyperCard database
extension.

Odesta Corp.
4084 Commercial Ave.
Northbrook, IL 60062
312-498-8852
312-498-5615
Makers of Double Helix II

Oracle Corp.
20 Davis Drive
Belmont, CA 94002
800-345-3267
Makers of Oracle database software. Spoke with a
Robert Silverberg, ext. 2019

SoftStream International
19 White Chapel Drive

Mount Laurel, NJ 08504
800-262-6610
609-866-1187
Marketing company for HyperHIT, a HyperCard
database extension

SoftStream—Steve Hannaford
19 White Chapel Drive.
Mount Laurel, NJ 08054
215-543-5194
Technical support representative for HyperHIT.

Stokes Mastering Service
Austin, TX
512-458-2201
A videodisc mastering service.

Symantec Corp.
10201 Torre Ave.
Cupertino, CA 95014
408-253-9600
Makers of Q&A

Turquoise Film/Video Productions
St. Louis, Missouri 63088
314-843-1998
A videodisc mastering service.

Voyager Company
239 Manning Ave.
Los Angeles, CA 90025
800-446-2001
Makers of VideoStacks, a set of videodisc drivers
and other software.

Dave Watkins
Media Services
B-27 MVR
Cornell University
Ithaca, NY 14853
255-5431
Dave is the head of Media Services and is working
with the videodisc project.

Margaret Webster
Architecture Slide Librarian
B-30 Sibley Dome
Cornell University
Ithaca, NY 14853
255-3300

Xiphias
12464 Washington Blvd.
Marina Del Rey, CA 90292
213-841-2790
Makers of Xearch, a HyperCard searching
extension

A videodisc is an optical disc which is read by a laser. The images are *analog,* which means essentially that they are stored as a snapshot consisting of shades of gray or color, rather than being divided into individual dots which can be either on or off, which is how an image would be stored in *digital* format.

CD-ROM stands for Compact Disk - Read Only Memory. It is a digital format, which means that any pictures are made up of individual dots which can be either on or off, black or white. As a result, pictures take up a great deal of space on a CD-ROM, so much space that a project this size would not be feasible. Mainframe storage systems would be required to store so many photographs.

There are two types of databases, relational databases and flat-file databases. Flat-file databases work just like a file cabinet in that each record is stored separately. Relational databases can share information between files, so you would not need to duplicate information if you had a database of addresses and a database of phone numbers because the two files could share the person's name. In addition, relational databases tend to be faster and more powerful.

HyperCard is a program described as a "software erector set" by its author. It is free with every Macintosh and allows non-programmers to create sophisticated programs, called stacks. HyperCard works on the metaphor of a stack of note cards, although it has a great deal of easily-accessed power which seems unrelated to a stack of cards. HyperCard is not a database, but it is an information manager and manipulator.

An external command is a small program that can be inserted into another program to give the second program additional functionality. They are extremely common with HyperCard and provide numerous ways of enhancing HyperCard.

A front-end is what you see and work with, whereas the back-end is the part which actually does the work. For instance, the front-end of a washing machine is the control panel where you set the type of wash and the amount of time. The back-end is the drum and vibration mechanism which actually washes the clothes. You have to be able to use the front-end, but you don't have to know how the back-end works to get your clothes clean.

I don't quite understand their technology and am merely trying to repeat it verbatim in hope that someone more well versed in the photographic arts will understand.

I don't know what the units in question are since Stokes didn't mention them.

A board is a piece of hardware which plugs into a slot in the computer and provides some sort of added functionality. Add-on boards usually fulfill a specific need which most people do not care about, which is why such functionality is not built in to the computer itself. The IBM PC clones and the Macintosh II line can easily accept such boards.

WORM stands for Write Once, Read Many. Essentially, a WORM drive is just like a CD-ROM drive except for the fact that the user can write to the drive once. It is very useful for archiving information because it cannot be erased afterwards.

[1] Prepared for the Department of Manuscripts and University Archives at Cornell University Olin Library, Ithaca, NY 14850. Distributed with the permission of the Department of Manuscripts and University Archives December 14th, 1989. Paper presented at the IASSIST 90 Conference held in Poughkeepsie, N.Y. May 30 - June 2, 1990.