LARGE-SCALE MULTI-MEDIA DATABASE IN THE HUMANITIES:
Case Presentation on PROJECT EMPEROR-I

Ching-chih Chen
Professor & Associate Dean
Graduate School of Library & Information Science
Simmons College, Boston, MA 02115

Principal Investigator & Project Director
PROJECT EMPEROR-I

ABSTRACT

The dynamic development of optical videodisc technology in the last few years has suggested great potential for the construction and interactive use of large-scale multi-media databases in subject areas such as the humanities and social sciences, in a way not possible before. PROJECT EMPEROR-I applies the latest in videodisc technology to present and interpret a major historical/archaeological period of China's Qin Dynasty together with its magnificent terracotta figures of warriors and horses. The experience gained from this exciting project together with the techniques used to create numerous interactive courses to construct and retrieve large-scale image databases, and to develop further a digitized image collection can be used to illustrate the potential of these new information technologies in humanities and social sciences.

INTRODUCTION

The overview article titled "Micro-based Videodisc Applications" (Chen, 1985) discussed how dynamic and explosive developments in microcomputer and related high tech areas in recent years have created great potential uses for videodisc technology, which were unattainable before. As we know, videodisc technology has been with us longer than microcomputers, however until recently, it has been limited in applications and has been mostly applied in recreation and entertainment areas. However, when combined interactively with computers, particularly with low-cost high-performance microcomputers, the potential for information management and retrieval is boundless. Furthermore, with the advent of computer hardware and software technologies numerous exciting applications related to large-scale multi-media databases in the humanities become possible. PROJECT EMPEROR-I can serve as a perfect illustration.

PROJECT EMPEROR-I

"PROJECT EMPEROR-I: China's Treasure Revealed via Videodisc Technology" is a major research and development project which has been supported by the Humanities Project in Libraries, US National Endowment for the Humanities (NEH) during the time period from October, 1984 to 1988. It applies the latest in videodisc technology in presenting and interpreting a major, historical/archaeological period in China.

The site and artifacts recorded and presented are from the period of the First Emperor of China. During the Emperor's brief fifteen-year reign, his accomplishments included the implementation of a unified written script, the connecting of strategic portions of the Great Wall, the unification of the various warring states, the standardization of weights and measures, the development of an extensive transportation system, and the building of his magnificent tomb near Xian. The discovery of the 7,000 plus life-size terracotta figures of warriors and horses near his tomb (Figure 1) is considered one of the most magnificent and significant archaeological finds in this century, and has captured the fascination of thousands and thousands of people all over the world (Chen, 1986).

The presentation and interpretation of this fascinating subject is done through the use of hybrid technologies -- videodisc and microcomputer.

WHY VIDEODISC FOR PROJECT EMPEROR-I

Optical videodisc technology holds great promise for information storage and retrieval because of the following characteristics:

● High density storage capacity - 108,000 frames of visual images on a two-sided analog disc and one-hour dual sound tracks;

● Quick random access - about 1.5 to 3 seconds to retrieve any one of the 54,000 frames of images on each side of the disc;

● Multi-media integration - capable of including information derived from multi-media sources, such as slides, videotapes, films, audio information and music, textual information, etc.;

● High resolution colour and graphics - multiple colour and resolution of the highest broadcasting quality;

● Perfect archival medium;

● Interactive capability with microcomputers - enabling the creation of innovative and interactive programs for achieving a new mode of education and research.

1

2

Figure 1.   Qin Terracotta Figures of Warriors and
             Horses (Courtesy of PROJECT EMPEROR-I)

OBJECTIVES OF PROJECT EMPEROR-I

"PROJECT EMPEROR-I is unique in that it is a project that brings together the East and the West, the past and the present, humanities and high tech, and scholarship and applications. Not only does it demonstrate how new technologies can help to enhance better understanding and appreciation of humanities by delivering enormous multi-media, multi-formatted and multi-dimensional information in a way not possible before, but also it introduces a new interactive educational and learning model. When this visual, audio and textual information.on videodisc, is used

interactively with a micro-based system with developed coursewares, the system user can consume, enjoy and digest online information at his/her own pace and choice" (Chen, 1987a). The creation of an effective image retrieval database which has either analog images or high-resolution digital image data points to an exciting future for the public consumption of humanities information, and humanities research, education and training.

WHAT HAS PROJECT EMPEROR-I ACCOMPLISHED?

PROJECT EMPEROR-I has accomplished, if not surpassed, all of the original goals and objectives set for the project. In meeting those objectives, we negotiated for official permission from the Ministry of Culture of the People's Republic of China to gather relevant information from museums, by interviewing experts, by filming at the archaeological diggings at Xian and videotaping at most major museums with Qin artifacts, and through archival film footage, as well as by digesting many Chinese and western publications, all of which formed source materials -- over 10,000 still slides, over five hours of archival film footage, and over 60 hours of videotaping -- from which selected materials were used to create the set of two double-sided videodiscs. These two discs contain 216,000 frames of visual images and two hours of audio information in both English and Chinese (further description is given in the following section).

Furthermore, we solicited and received hardware support from a variety of firms, including Digital Equipment Corporation and SONY American, Inc., and have utilized the equipment to begin the dissemination effort and to further develop databases and interactive courses. The software for the image database construction and interactive course development has also been donated -- C-Quest from Image Concept Inc. and DIRECTOR from Videologic, Inc. Subsequently, work is continuing on constructing an extensive image database on the selected images from the discs, specifically those from the "Still-picture Library". It is also expected that seven interactive courses will be completed soon. These courses include one for the general public and six for more advance interactive use by researchers and students of the subject field. They will enable users to digest the material from the discs and learn in a self-directed/self-paced mode.

PROJECT EMPEROR-I'S PRODUCTS

"The First Emperor of China" Videodisc

As already stated, PROJECT EMPEROR-I's initial product is a set of two 12" NTSC CAV videodiscs, entitled, "The First Emperor of China: Qin Shi Huang Di, 秦始皇帝 ." Each two-sided disc contains 108,000 frames of visual images and one hour of audio track with narration and/or interviews in both English and Chinese, together with authentic musical interludes. The first

disc includes over 200 segments of motion video from films and videotapes, and over 4,000 still frame slides. This visual information, together with matching narrations in both English and Chinese, is generally arranged in "chapters" like an electronic book, each chapter dealing with an identifiable topic: for example, "Introduction to the First Emperor of China"; "The Great Wall"; and "Collection of the Qin Terra-Cotta Museum of Warriors and Horses." The visual and audio information is programmed to be retrieved quickly by either chapter or frame search. The second disc essentially contains both oral and visual history information, including interviews with ten of the world's foremost experts on the subject. Both visual and audio information are arranged again in chapters by historically researched questions asked of each expert. With the high density storage capacity of videodiscs, it is safe to claim that our discs are probably the most comprehensive electronic database on the subject existing at this time.

While the discs are not intended to be stand-alone products, they do contain three short stand-alone pieces which can be either accessed linearly like most video productions or interactively selected from the displayed menus in the microcomputer system through the use of the developed coursewares. These three introductory stand-alone pieces are:

● The First Emperor of China (about 8 minutes);

● The Great Wall (about 5 minutes);

● The Excavation of the Terra-cotta Figures (about 5 minutes).

These three segments are likely to have wide appeal for the general public, thus will be widely used for interactive course development.

It is also fair to note that even without the interactive courseware, the set of two videodiscs alone are an invaluable, large-scale, multi-media database on the subject, since it is the first time, enormous amounts of inaccessible information sources in visual, audio, and textual formats, and much newly created multi-media information are being brought together in a convenient way for quick random access.

Electronic Database

In addition to the set of two completed videodiscs, a comprehensive image database is being constructed for selected visual images on the videodiscs, which are deemed to be most significant for program development, education and research use.

Large-scale image processing and management represents an exciting challenge for PROJECT EMPEROR-I. While there are many picture of Qin artifacts, the magnificent terra cotta figures and pits, and many other images dealing with the same topics, they

present very different visual pictures. Using the terra cotta figures as an illustration, one can imagine the great variety of pictures, such as:

● one single complete warrior or horse;

● numerous full-size warriors or horses;

● combinations of types of warriors and horses;

● close-up views of a certain part(s) or view(s) of figures;

● various views of a figure, such as front, back, profile, upper, lower, etc.

This list can go on. It is important to keep in mind that an image generally encompasses numerous concepts which cannot be adequately described by "keywords." Thus, the traditional keyword approaches to the organization of these images, without adequate and detailed descriptions and modifiers, are usually not sufficient for relevant and precise retrieval of specific images (Chen, 1987).

For PROJECT EMPEROR-I, an image database is being constructed using the C-Quest Image Organization and Retrieval Software of Image Concept, Inc., on an IBM XT or IBM AT. C-Quest is a flexible software with hierarchical and synonym capability. Each record includes approximately a dozen fields, such as disc side number, frame number, type of object, date of object, size of object, material of object, date of discovery, site where discovered, current location of object, information source, publication source, and comments. Most of these fields and a combination of a great number of subjects (as many as 25), with or without modifiers, can be searched easily. The image database will later be joined together with the developed interactive courses. Thus, a system user will be able to interactively and simultaneously retrieve the desired textual, visual and audio information by using micro-based interactive videodisc systems, such as the Digital Equipment Corporation's Interactive Video Information System (IVIS) or SONY's VIEW system (Chen, 1987b).

Interactive Courseware

Several interactive courseware packages on a number of important topics are also being developed for both general program use and the more interactive research and educational use. Each course will lead the course/system user through a series of lessons step by step, with choices being made from both the menu and sub-menus presented to them on the microcomputer monitor screen. Concurrently, the desired visual, audio and textual information in whatever combinations selected by each system user is provided online as ordered via direct computer access or access to the videodisc. Browsing and image manipulation capabilities are also available. In addition, information such as bibliographies for further readings, full-

text retrieval of certain selected readings, a term dictionary, and a chronology of historical events are also being built in as part of the course selection (Chen, 1986). Each interactive videodisc course utilizes the system features, such as branching, windowing, multiple-choice, and quick access capabilities to the existing massive amount of multi-media information on both the videodiscs and the associated microcomputer.

These self-paced interactive videodisc course modules, are developed to allow individuals to pursue the subject to the extent of their interests. Individual system users may peruse the subject at their own pace and within their own time-frame Clearly, each system user can select from the displayed menu and/or sub-menu on a microcomputer monitor the specific information which she or he would be interested in pursuing. For example, the main menu permits one to select "topics," "explorations," "reference," or "messages".

● Topics

Under "topics," a system user can choose any one of the following multiple topics also shown on the screen in Figure 2:

- Historical background and introduction
- Significance and symbols
- The art of war
- Social and economic organization
- Technology, art and architecture
- Etc.

Further selection can be made under each of the above listed sub-areas. For example, under "Historical background and introduction," one can choose any of the following sub-topics:

- Introduction to the First Emperor of China
- Introduction to the Great Wall
- Chronology of Qin history
- Major achievements of the First Emperor of China
- Reasons for Qin collapse
- Etc.

● Explorations

For this selection, one is able to explore all the multi-dimensional (rotational) images and videos stored on the videodisc. One is also able to browse the images on the disc by providing the capability to move to the next image, to move back to the previous image, to jump multiple frames of images, to zoom in for close-up image or to zoom out as shown in Figures 3 and 4.

● Reference

A system user can select "reference" to look for more information on a topic as well. Under "reference," one is able to go to the "glossary" for definition of terms, or to obtain
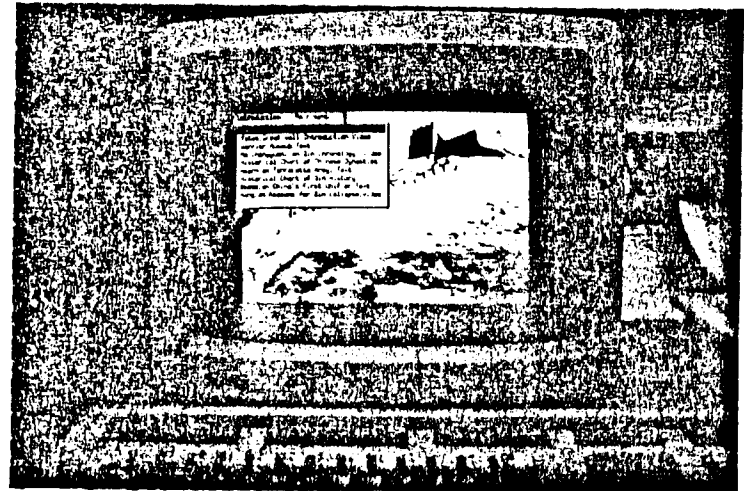

Figure 2. Topical Screen on a Microcomputer
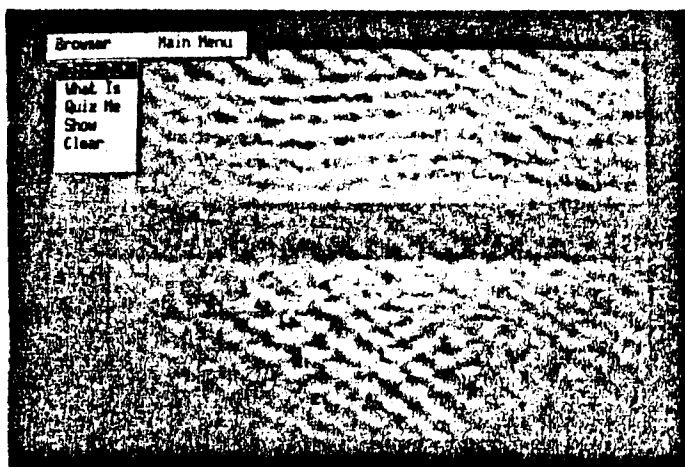

Figure 3. Slide Browsing under "Exploration"

Figure 4. "Zoom in" Effect of the Same Image
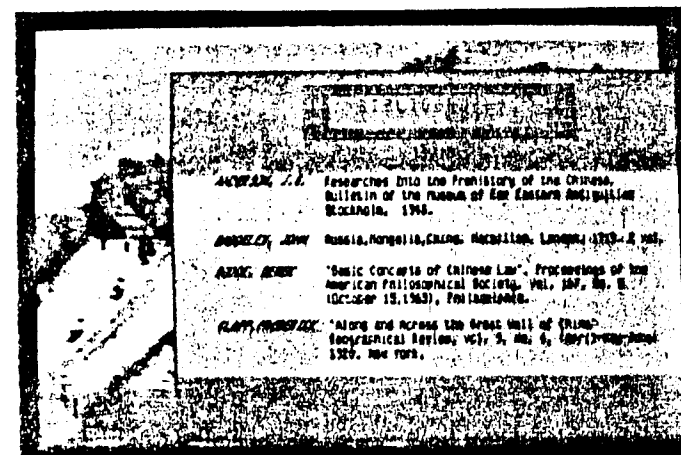Shown in Figure 3.

more bibliographic information on a topic (see Figures 5). Furthermore, many of the listed bibliographic items can also be immediately retrieved online on a full text basis, therefore one can either browse the first few pages of a requested article on the screen, or can "read" the whole piece.

● Messages

This final category offers a system user the capability to either send messages to other system user(s) or to receive messages and/or announcements (see Figures 6).

IMAGE DIGITIZATION

PROJECT EMPEROR-I also explores numerous opportunities in image digitization and its potential impact on information retrieval, conservation, preservation and restoration activities, and information utility.

For image digitization with resolution of broadcasting quality, we have explored the use of AT&T Taga Board and WORM technology with IBM AT compatible (Figure 7). Once the images are digitized, they can easily be manipulated for various purposes. For example, any part of a given image can be either

Figure 5. Bibliographical Display after Selection of "Bibliography
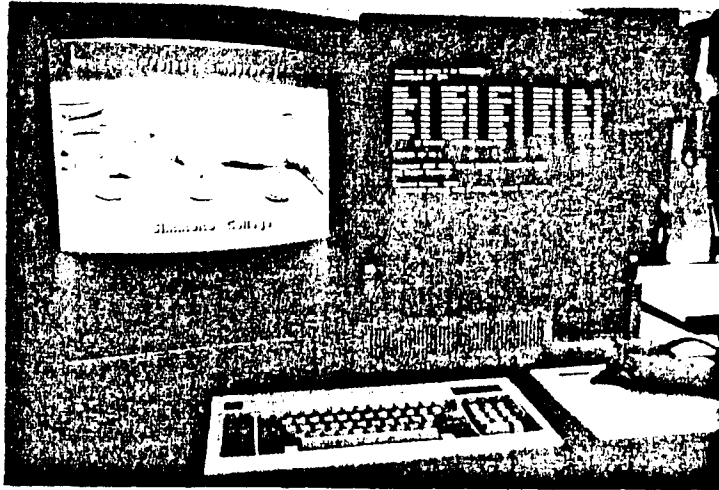

Figure 6. Selection Category of "Messages"

Figure 7. The AT Configuration with Taga Board & WORM Disk

enlarged or reduced in size, altered in position and colour, etc.

For the high-resolution image digitization (with 1K by 1K or 2K by 2K by 24 bits of color), we are currently working with Image Understanding Systems (IUS) of San Jose, CA on an exciting collaborative project in using the IUS' powerful OASIS (The Object Analysis and Interpretation System) that is based on a dedicated Sun 3 family high performance workstation using its native image display and monitor. OASIS includes the entry, management, display, analysis and output facilities (see Figure 8) that can provide PROJECT EMPEROR-I with additional exciting R&D application. For this experimentation, we are also using IUS' image data management and query tools, TIME (the Total Information Management and Exploitation software).

The high-resolution digitized image, produced by using the Ekonix digitizor, are retrievable by various searchable fields, such as title, time period, location of objects, subject matters, etc. Descriptive textual information of a given image and the image itself are again viewable at the user's selection. Furthermore, any portion of the image can also be enlarged, or
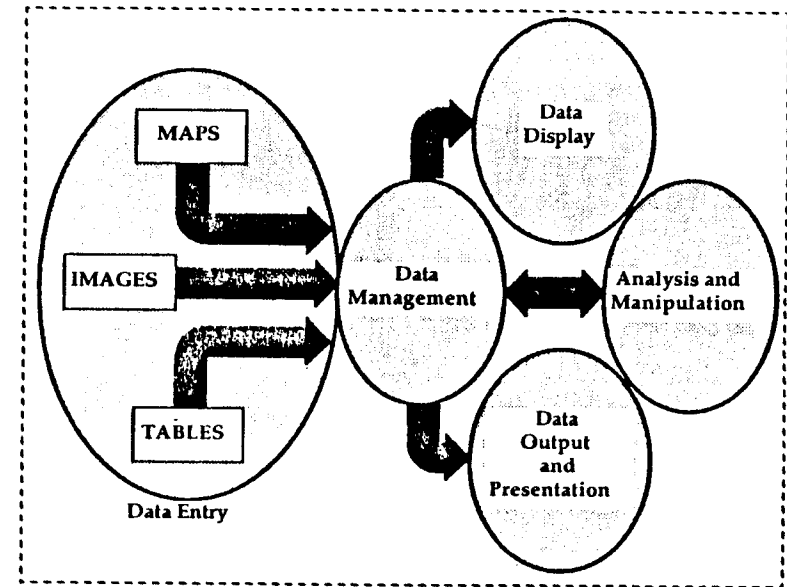
11

Figure 8. System Components of OASIS
(Courtesy of Image Understanding Systems, Inc.)

shrunk, or altered to meet the information needs of a user. The TIME system offers each user a spatial, textual and temporal database management and retrieval system for images, maps, tables and textual information with many attractive features, such as the tracking of raster, vector and text/tables, etc.

It is also our hope that our experience in this application will help IUS to develop a micro-based version in the near future for those who can only afford to have a microcomputer.

MANAGERIAL ISSUES AND PROBLEMS

A major interactive videodisc project such as PROJECT EMPEROR-I involves many different professionals - media specialists such as photographers, video specialists, computer video specialists and others; art historians, archaeologists,

12

museum specialists, computer specialists, computer graphic artists, computer programmers, information specialists, translators, calligraphers, musicologists and many others – during different phases of the projects. Each has his or her own strong views on problems and issues related to the project. The ability to work harmoniously and efficiently among project staff is the key to the success of the project. Thus, it is important to stress that technology related problems are usually the easiest to solve, while problems related to human interaction among a diversified project staff and personnel management are the most challenging problems of all. It may seem surprising to many, but it is generally true that, although the project is very much high-tech oriented, it is fair to say that human factors are most likely to determine the eventual success of a project. The overall responsibility of the project director is enormous indeed!

CONCLUSION

PROJECT EMPEROR-I is a perfect R&D project for fully demonstrating the great potential of interactive videodisc technology for quick online access to multi-media and multi-formatted information, for better information delivery outside the traditional boundaries of time and geography, and for offering opportunities for a new mode of interactive learning and education. Because the subject matter of this project is in the humanities -- art history and archaeology -- it helps to offer tangible products in the fields of humanities, which demonstrate how this type of new information technologies can create profound changes in the way information is accessed, retrieved, consumed and researched. Increased and easy access to large-scale visual, audio and textual information sources can also basically change the roles of many traditional institutions such as libraries and museums.

REFERENCE

Chen, Ching-chih, "Micro-Based Optical Videodisc Applications," Microcomputers for Information Management 2 (4): 217-240, December 1985.

Chen, Ching-chih, "Sample Interactive Courseware Development on the Use of PROJECT EMPEROR-I Videodiscs," in Proceedings of 10th International Online Meeting, London, December 4, 1986. London: Learned Information, 1986. pp. 223-228.

Chen, Ching-chih, "Interactive videodisc technology and microcomputers: A case presentation on PROJECT EMPEROR-I," in The Application of Micro-Computers in Information, Documentation and Libraries (Proceedings of the International Conference on Microcomputer applications in Libraries, Information and documentation, Baden-Baden, March 1986). edited by K.-D. Lehmann and H. Strohl-Goebel. Amsterdam: Elsevier Science Publishers, 1987. pp. 657-661. (1987a)

Chen, Ching-chih and Robert D. Stueart, "Interactive Videodisc Technology and Microcomputers: A Case Presentation of PROJECT EMPEROR-I," in Proceedings of the First Pacific Conference on New Information Technology, Bangkok, June 16-18, 1987 edited by Ching-chih Chen & David I. Raitt. Newton, MA: MicroUse Information, 1987. pp. 33-40. (1987b)

Chen, Ching-chih, "Large-scale image processing," Bulletin of the American Society for Information Science 13 (6): 15-16 (August/September 1987). (1987c)

PROJECT EMPEROR-I has been extensively covered in conferences, public media, computer and video journals, and subject and library literature. For further information, write to:

Dr. Ching-chih Chen
Principal Investigator & Project director, PROJECT EMPEROR-I
Professor & Associate Dean
Graduate School of Library & Information Science
Simmons college
Boston, Massachusetts 02115, USA