

# **The Digitization of the Gutenberg Bible and other Historic Books**

Shigeaki Iwai  
Keio University, Tokyo

## 1. Organization, operation and aims of the HUMI Project

The Humanities Media Interface Project at Keio University was established in 1996 as an inter-faculty organization with the aim of conducting basic investigations and experiments on how to realize a digital research library whose primary resources would be the rare book collections of our university. Specialists in various fields of the humanities, including both Western and Oriental cultures, and experts in information technology and library science started collaborative activities seeking the questions and answers concerning the feasibility of the plan. The project formed a consortium with about 15 Japanese private enterprises that are interested in the digital archiving of cultural heritage, and operating costs are covered by their research grants and school budget. The Japanese government also assisted us to start the project.

## 2. Activities of the HUMI Project

Activities of the HUMI Project are manifold, but most of them can roughly be categorized under the following 4 issues: (1) evaluation of digitization technology; (2) digitization of rare books in the Keio University Library; (3) collaborative digitization of the Gutenberg Bible in foreign libraries; (4) application of digital technology to studies in the humanities.

### 2-1. Evaluation of digital technology

For such a specific purpose as providing digital images of rare books, the evaluation of certain fields of information technology is important. High-resolution digital cameras, secure and capacious storage devices,

fast network devices and fine PC monitors are indispensable for our purpose, and in addition, peripheral equipment, e.g. a camera stand and a book cradle, specialized in the digitization of rare books should always be kept in mind. In spite of the rapid development of the digital technology, quick capturing, safe storage, fast transmission and fine display of image data still present many hurdles.

It can be safely said that an ideal image capture method for rare books has not yet been produced, although there are various products on the market. Images produced by using digital scanning camera are widely accepted, but "scanning" is often time-consuming, and that induces its own problems. For example, keeping pages of fragile books open is difficult and less than ideal, and overexposure to light and heat is potentially damaging. A one-shot digital camera on the present market is free from the problem of capturing time, but it has a problem of image quality caused by the mechanism of its image sensor. If a large-format analogue film is scanned with such a high quality scanning device as a drum scanner, the problems of capturing time and image quality are solved, but this method costs too much for a voluminous book. A radical solution to those problems is a successful development of a one-shot, high resolution and highly reproducible digital camera, but there seems to be no sign of its appearance on the market so far.

The first priority when capturing digital images should be put on the preservation and protection of rare books. Each rare book is different in shape and size, which means that each book requires a special supporting cradle. The Gutenberg Bible, for example, is a comparatively large bound book with a heavy upper and lower cover. To avoid strains on the spine and joint during image capture, a custom-made book cradle must be used with a slanting-angled camera stand. Safe light is also one of the most important component parts of peripheral equipment.

A safe and capacious storage system, backup media, and a connection to a high-performance server and client computers through the medium of a fast network, are essential for an efficient manipulation

of digital images. Among these basic technologies, the speed of data transmission and the quality of the display are vital factors for the realization of a digital research and educational environment. The spread of the 100Mbit fast Ethernet system rendered it possible to transfer large image data, at least in a local area. However, the present Internet traffic is still too narrow to enable remote access from outside. If the predicted improvement in the Internet traffic becomes a reality, the usefulness of archived images will be rapidly acknowledged. And this is always in close relation to the performance of a client computer, especially a computer screen. Last year the news of the successful production of extremely high-resolution LCD screens was reported in succession. If such screens are put to practical use and the quality of the screen is endlessly improved in the future, the focus of our attention is pulled back to the issue of image capture again. Needless to say, the acceptable size of a captured image depends on the quality of a target screen. For example, if we make up our minds to use in our digital rare book library the recently announced computer screen whose size is 22 inches with the resolution of about 3,800 x 2,400 pixels, our choice for image capturing methods is inevitably limited; only the images produced by a digital scanning camera or a large-format film scanner satisfy the required quality at present. And this again reminds us of the problems of scanning time, book preservation, cost, storage space, transmission speed, etc. All the issues are closely related to each other and the feasibility of a digital rare book library depends on an appropriate integration of those basic technologies.

## 2-2. The digitization of rare books in the Keio University Library

The second activity of the HUMI Project is the digitization of the rare book collections in our library. Keio University was founded in 1858 by a Japanese educator Yukichi Fukuzawa and its library has been developing in the course of the school's history. The rare book room of the library holds approximately 10 thousand items comprising Western and Oriental materials, and all the pages of those items are expected to be digitized.

From 1997 to 1999, in collaboration with a Japanese governmental

organization called the Information-technology Promotion Agency, a number of those items were digitized and made open to the public. Several digitization methods and web designs were tested in this project, and the practical knowledge acquired in this activity is being utilized and developed in other projects.

2-3. Collaborative digitization of the Gutenberg Bible in foreign libraries

### **2-3-1. Digitization of the Gutenberg Bible in Cambridge, Mainz and London**

The third and one of the most prominent activities of the HUMI Project is the collaborative digitization of the Gutenberg Bibles in foreign libraries. The Gutenberg Bible is known to be the first substantial printed book in the West and a person named Johann Gutenberg is thought to have printed it around 1455 in Mainz, Germany. The Bible is constituted of 2 volumes with 1,284 pages in total. 48 sets are extant in the world and in 1996 Keio University acquired one with its 2nd volume missing. It was this that led to the discussion on the establishment of a digital rare book library in our university, and therefore, the HUMI Project has been particularly involved in this item from its start.

After completing the digitization of the Gutenberg Bible at Keio University in 1997, the HUMI Project has actively pursued cooperative ventures with research libraries and museums abroad. One of our aims is to accumulate images of different versions of the Gutenberg Bible for bibliographical comparisons, and another is to promote international exchanges of rare book images. In November 1998 we visited the Cambridge University Library for the digitization of a complete set of the Bible, and this was the first time the entire pages of the Gutenberg Bible were digitized. In November 1999 we were invited to the Gutenberg Museum in Mainz, Germany for the digitization of one and half sets of the Bible — the first volume of a set is missing. In March 2000 we visited England again for the digitization of 2 complete sets of the Bible in the British Library. A long discussion with the British Library on the digitization system, particularly on the book cradle for the sake of preservation led to a

great improvement of our system.

## **2-3-2. Technical details of the digitization of the Gutenberg Bible**

### **(a) Digital camera**

The digital input method brought into those libraries and museum is a digital camera which is characteristic in that it is a "1-shot 3-sensor type". The size of a produced image is 2,048 x 2,048 pixels and the color depth is 12 bit. This digital camera was developed by a Japanese telecommunication company (NTT: Nippon Telegraph & Telephone Corporation) in collaboration with an optical company (Olympus Optical Co.). Since this camera is an experimental prototype it is still not on the market. These companies are members of our consortium and they proposed to us to test the performance of the camera through the actual digitization of rare books. The advantages of this camera are speed and quality. An image is captured in less than 1 second and stored in the disk within 15 seconds, and since it has 3 sensors for each RGB color the image quality is far better than the one produced with a commercial digital camera with 1 sensor. Another advantage is that an NTSC real time video finder is integrated into the system. This enables the operator to have quick and accurate adjustments of composition and focus. 1 set of the Gutenberg Bible with 1,284 pages can safely be captured within 4 days in combination with a specialized camera stand, book cradle and other peripheral equipment.

### **(b) Peripheral equipment**

In addition to the digital camera, the digitization system used for the Gutenberg Bible consists of four components of peripheral equipment. The first component is the strobe flashlight, which has an advantage in shortening the exposure time. Needless to say, this lighting system can only be used with a "shot type" input method. At the British Library heat, brightness and the strength of ultraviolet rays caused by the flashlight during the image capture were kept monitored.

The second component is a book cradle specialized for a safe support of the Gutenberg Bible. To avoid a severe strain to the spine and joint during image capture, and to achieve the maximum flatness of the captured page without using anything like an air suction system, a

stable and precise book cradle is required. The HUMI Project designed many prototypes and the latest one was used at the British Library.

The third component is a camera stand which is also specialized for the Gutenberg Bible. To avoid a possible danger of the camera falling on the book, a slanting-angled camera stand was produced. A flexible movement of the camera position is also an important function of it.

The fourth component is a storage system connected with a 100Mbit network. The image data captured with the digital camera is immediately stored in a RAID disk and duplicated in another file server for the distribution of the data to other backup systems. A reliable and capacious backup media is DVD-RAM.

### **2-3-3. Products**

The captured digital images are appropriately processed and formatted into a set of master data and stored in a secure site of our university. The duplications of the original data and the formatted master data are handed to the institution that permitted us to digitize their Gutenberg Bibles. By using those master data we produce digital materials for research and education purposes. The online digital facsimile is the primary product for all institutions. According to the agreements between Keio and other institutions, not all the copies of the Gutenberg Bible are online at present, but the Keio copy and the British Library copies are already open to the public on each web site. The Cambridge copy is on the web but access to it is restricted to the users in both universities. The Mainz copies are not yet online. On the initiative of the HUMI Project, the 4 parties are now discussing a new agreement that will permit other institutions to put images online within each internal network.

From last September to the end of December the HUMI Project and the British Library jointly organized an educational display entitled "The Digital Gutenberg Bible" in the British Library. The display of about 10 minutes illustrates the uniqueness of the Gutenberg Bible in terms of its bibliographical interest. The LCD screen used for the

display was provided by a Japanese company (NTT Corporation), a futuristic device as previously mentioned. Educational programs of this sort are planned to be published in the form of such portable media as CD-ROM or DVD-ROM.

#### 2-4. Studies in the humanities enhanced by digital technology

As previously mentioned, one of our aims of digitizing different copies of the Gutenberg Bible is to advance bibliographical studies. It has been very difficult to conduct minute comparisons of the Bibles because they are scattered in different countries and treasured by their owners. When we successfully digitized a complete set of the Bible in Cambridge, we started visiting other libraries with the images in a laptop for the purpose of on-the-spot comparisons. The images were useful but it was still time-consuming to make comparisons between an original page and a digital image. In order to accelerate the speed of examination, and to increase the precision of comparisons we developed a program, which enables a researcher to detect in a very short time such a slight difference as a stop-press-correction, which is a correction of letters during printing process. The same pages of different copies are overlapped on the computer screen and flipped at high speed. If there is a difference between two copies, the spot is visible as an irregular movement while the rest of text is quite static. The next step we will make is to develop an automatic collation program, and our technical staff intend to produce it by applying image processing technique.

#### 3. Plans for the future

Keio University envisages the establishment of new research centers for the application of digital technology to the humanities. At Bury St. Edmunds in England, Keio will establish such a facility for the digitization of Western materials this year, and as previously mentioned a digital rare book library is being planned for our campus. In order to make our plans more elaborate and sophisticated, we feel it is necessary to keep close contact with institutions and individuals that share common interests with us.

[End of the paper]