The Experience of Images Digitization at the National Palace Museum:
The Digital Archives Project of Chinese Antiquities

C. Alan Cheung and Mei-li Yang
National Palace Museum
Taipei, Taiwan, R.O.C.

Abstract

The mission of the Digital Archives Project of Chinese Antiquities (DAPCA) is assigned to complete the digitization of entire collections at the National Palace Museum (NPM). Images digitization is one of three main complicated tasks within this project. Since the technique of digital photography has been widely applied in images digitization, DAPCA built the first imaging system in September 2003 and the second one was delivered by October 2004.

This case intends to introduce the strategies of images digitization applying on antiquities and the experience of imaging systems' evaluation. First, the introduction of DAPCA is delivered. The development of imaging systems is also included. Second, this paper describes the workflow of imaging systems. In addition, the limitations of the systems are also posted as reference. Finally, several issues of system operations and future plans of this project are mentioned. These issues are related to image quality control, uses of digital images and the master file format.

Introduction

The Digital Archives Project of Chinese Antiquities (DAPCA) is one of eight sub-projects in "the Research and Development of the National Palace Museum Cultural Artifacts Digital Archiving System". It was started on January 1st 2002 and is scheduled as a five years project. Its main objective is to digitize the entire collection of antiquities at NPM in both textual data and image data and to make them available for public use through the application of informatics technologies.

The Collection

NPM is proud of its rich and historic artifacts, which are classed to three categories: "antiquities", "paintings and calligraphies" and "books and documents". The objects in the Museum's collection, mostly collected by many dynasties' emperors, are both great in number and variety. As of November 2004, the grand total of the collection is 655,156.

The number of antiquities is nearly 70,000 and the period of time is spread from 5,000 B.C. to now. Because most of them do not have qualified images for applications of research and publishing, 60,000 items of antiquities require the capture of digital images. To reach this, the speed and quality of the images digitization are considered as the critical success factors.

The Project

DAPCA includes three main tasks, which are text digitization, images digitization and system development. Figure 1 shows the simple structure of DAPCA. The other assignments involve collection management system maintenance and thesaurus development.

To digitize the entire collection of NPM, especially antiquities' images digitization, we developed two digital imaging systems near the treasure storages to improve digitizing efficiency. The quality strategy of images digitization in DAPCA follows the International Color Consortium (ICC) color management workflow.

Due to the position of storages, the arrangement of images digitization begins from bronzes to jades followed by ceramics and curios and ending with others as the final stage. This order is helpful for setting the detail schedule of practical imaging when designing the digitizing workflow. At the end of 2004, 9,675 items of bronzes and coins were digitized and 35,418 files of image were captured. The total capacity of files is more than 2,200 GB. It is predictable that the grand total items of digitized antiquities will be 25,000 at the end of 2006.

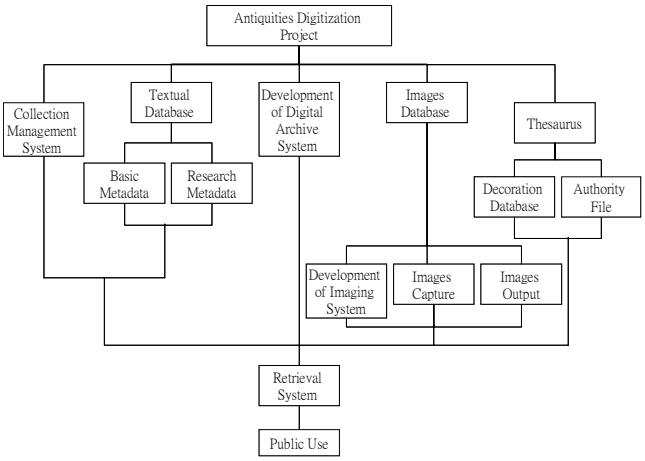

Figure 1. DAPCA structure

Development of Imaging Systems

The strategy of images digitization in DAPCA is concerned about both speed and quality, which also affect the evaluation of the digitization system. Several considerations are pointed out in the prior preparation of decision-making.

To begin with, we decided to produce digital images in-house instead of by outsourcing. Because DAPCA is scheduled as a five-year project, in-house is more cost-effective than outsourcing. It not only saves budget but also gathers images digitization experience and knowledge.

Secondly, we thought about the methods of digitization, which include direct digitization and indirect digitization. To take a photo, develop a film and then get a digital image by drum scanning used to be the mainstream way to digitize images in the last few decades. This indirect digitization way to produce images is still unable to replace an 8 X 10 transparency film with high quality drum scanning. However, as the technologies change with each passing day, the digital camera is becoming one kind of qualified equipment for images digitization. Digital photographic technologies win our trust by reducing time and providing equal, if not better image quality through direct imaging procedures.

The last consideration was how to build the digital imaging system. The challenges were not only the shortage of imaging equipment providers, but also the difficulty of scientific experimentation of equipment evaluation. The way we chose was to invite an image scientific consultant from outside of NPM first, and define the image needs of DAPCA. Then, we asked system suppliers to perform demonstrations, and the performance of the imaging systems were evaluated by consultants and users. After completing the evaluation of all digital imaging systems, the best specification was described and formatted for the purchasing department.

1st Digital Imaging System

The development of the 1st digital imaging system was started at the beginning of 2003 and completed on September 1st 2003. The qualification of digital images for DAPCA had several requirements and the most important one was the publishing purpose, which satisfies the page size of the antiquities catalogs in NPM. The five components of the digital imaging system are camera body, digital back, lighting system, workstation, and color management system, which were evaluated after the system presentation. The customer service and knowledge of images digitization were also appraised as reference during system testing. The specification was consisted from different system providers instead of a single system because we wanted each component to be their best. The modules of the 1st imaging system are listed:

Camera body: Contax 645AF 120 DSLR
Lens: Carl Zeiss 80mm/F2 AF, Carl Zeiss 120mm/F4 1:1 Makro, Carl Zeiss 35mm/F3.5 AF
Minolta Color Meter IIIF
Minolta Flash Meter V
Digital back: Phase One H5 (6 Mega pixels)
Lighting system: BALCAR · NEXUS Asymmetric 3200 (Electronic flash light)
Workstation1: Dell Precision Workstation 650
Monitor1: Barco Reference Calibrator V CID421
Workstation2: Apple G4 1.25GHz Dual CPU/1GB RAM
Monitor2: Apple Cinema Display 20" LCD
Color management system:

- GretagMacbeth ColorChecker DC ColorChecker
- GretagMacbeth ProfileMaker Pro 4.1
- GretagMacbeth Eye-One Pro Bundle
- GretagMacbeth iCColor 210 series
- CGS O.R.I.S. Color Tuner RIP
- EPSON Stylus Pro 10000CF
- GretagMacbeth Judge Ⅱ
- KODAK Gray card, KODAK Q13, KODAK Q14

2nd Digital Imaging System

Because the performance of the 1st digital imaging system satisfied us with its efficiency and quality, in the end of 2003 we decided to speed up the images digitization of DAPCA through adding another imaging system. We also considered the concept of original size duplication, using a large CCD to capture larger size antiquities and a small size CCD for smaller antiquities, and using a cooling light to avoid damaging the organic antiquities. Because there is only one studio for digital photography, the operational environment of the second digital imaging system should be comparable with the first one and the equipment of color management would be redundant. After an evaluation process similar to the 1st imaging system, the 2nd imaging system's main difference was in its camera body, digital back and lighting system as follows:

Camera body: Hasselblad ELD 555
Lens: Hasselblad CF/CFE Planar 80mm f2.8, CF/CFi Makro-Planar 120mm f4, CF/CFE Distagon 40mm f4
Digital back: Phase One H25 (22 Mega pixels)
Lighting system: BALCAR continuing cool light

3D Images Digitization Project

In the beginning of 2004, NPM launched an initial project to create 3D images, which surpass the present limitations of multimedia display and provides a virtual exhibition of fine arts to visitors. Many different fields' experts were involving in building the NPM 3D Virtual Collection Exhibition System and those institutes included the National Taiwan University, Academia Sinica and the Industrial Technology Research Institute. The five antiquities chose for this project are: "Jadeite cabbage with insects", "Mao-Kung Ting", "Revolving Vase", "Ivory Sphere" and "Boat Carved from an Olive Pit". It is a great successful application of 3-dimensional technology, and the utilization of panorama and object movie techniques provides the availability for the digital archive of antiquities images.

Current Issues of Practical Imaging

The objective of images digitization is to produce high quality digital images for persistent utility; therefore, the quality issues should be considered a high priority. Directly capturing images from the digital imaging system seems to be adopted more and more; however, there is still ambiguity in imaging performance measures. Subjective assessments are not easy to implement and objective methodologies are not satisfied either. In the case of DAPCA, there is not enough time to probe the scientific experimentation of both except well implementing current commercial equipment to develop the best practices. Workflow, quality measurement and storage are three critical parts of implementing the digital imaging system. These are the core issues of images digitization and deserve improvement.

Workflow of the Digital Imaging System

The workflow of the digital imaging system of DAPCA includes preparatory procedures, considerations, shooting and checking, follow-up procedures, output and proofing, inspecting image quality and copying, storing, and embedded watermarks. (Figure 2) Firstly, each object had to be cleaned and inspected during preparation procedures. The next stage was the creation of a detailed list of the objects and numbered label for each object. The list records the photographer, handler, date, number of photos, name of the image file, and other necessary information. The label is used to provide scale as well as identification. The list was essential because it was used to enter information after the shoot, while the label was used to match each object.

One of the limitations of the shoot was that curators had to be on hand in order for any handling or touching of the objects. This included the move from storage to studio, taking the objects apart, placing the object on the platform for capturing, and any adjustments required by the photographer.

Curators and researchers checked the accession numbers of the objects to make sure they were correct, and these were entered into the name of image file. The photographer took care of the lighting, composition, and shooting. Immediately after a photograph was taken, an assistant would check the image for foreign matter or dust that might impact the quality of the picture, and the photographer would check the image for problems with the photography, including exposure, the gray scale, composition, and lighting. These considerations would impact the decision of whether or not another shot needed to be taken.

On average, each object was shot from three different angles, although sometimes carved objects required any where from four to six photos. These photos included the most obvious: frontal, side, and back views, as well as inscriptions and special decorations, which were a special consideration for each object. Each photo was taken with the numbered label in order to prevent errors in identification and mislabeling of the digital image. On average, one digital imaging system during each working day would capture about 15 objects, but it also depended on the object type.

Following the shoot, the color and curves of the image file were adjusted. Using a platform of a digital camera back, software for adjusting the image, a high-resolution screen, and the workstation, the photographers adjusted the color according to their experience and their better judgment. Also, the sharpening, curving, and color levels were adjusted for the particular printing purpose. These adjustments do not have a generally accepted standard, so these adjustments were made based on experience and the character of the object. The three image formats were processed after adjusting the original image RAW file. These are 48bits, 24bits-RGB-tiff and JPEG.

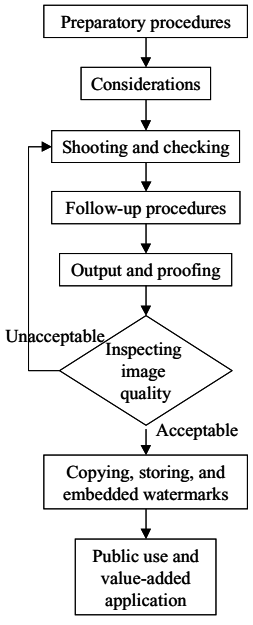

Figure 2. Images digitization workflow

After the adjustments were made, proofing was done on an eight-color inkjet printer on two types of paper. The image quality was inspected after this, using a standard

viewing booth, which follows the international standard, ISO 3664. These files were then sent to the Information Management Center. To protect the intellectual property rights of the Museum, each tiff file was embedded with an invisible watermark and each JPEG file was embedded with a visible one.

Limitations of Digital Imaging Systems

Color reproduction is one of the most important quality characteristics. To follow international standards, we implemented ICC color management across the workflow of images digitization to ensure consistent color between cross-medias. The ICC profile types included camera profile, monitor profile and printer profile. Each image was embedded with AdobeRGB1998.icc in tiff file format because of its wide gamut and acceptance by many commercial devices.[2]

While the ICC profile quality is a big issue in the real world,[1] it is not an easy way to evaluate profiles no matter if it was built in-house or commercially provided, especially in camera profile and printer profile. The gamut of the camera profile built in-house is not as wide as commercial one and the printer profile has the same situation. Moreover, as the antiquity is a 3-dimensional object, the photographic light setting changes in every different shoot. Therefore, the camera profile should be recreated in each single scene by theoretical concept. In this case, it is impossible to follow this guideline because the large number of antiquities has to be digitized in a short time. To overcome this, we decided to take an accommodation and accept commercial ICC profiles for all devices. After all, the device manufacturer does have a standardization lab to make better ICC profiles.

Lighting standards and the photographic lighting selection are two limitations of digital imaging systems. In the first, the lighting environment in which curators watch antiquities is not the same as the photographic lighting. The color temperature should be concentrated and consistent. In general office lighting, there is about 9000 degrees Kelvin but the most recommendations are in daylight between 5000 to 6500 degrees Kelvin. The different lighting sourcing will cause different reflections from the same object and confuse the curator's adjustment. The photographic lighting selection will also affect the color quality of digital images. Comparing the electronic flashlight with a high frequency cool continued lighting, the flashlight has much better spectrum than the cool light. We were told to use the cool light to protect fragile material and organic objects. However, there is no available way to evaluate the UV and energy of light and tell us how to assess and select a better lighting system. The problem between image quality and antiquity conservation is difficult to solve and needed scientific analysis so that the safest way could be chosen to compensate psychology.

File Format and Storage Issues

Since the 22mega pixel one-shot digital back came out, the storage space of digital images is a new problem. It is not only for long-term preservation but also for a variety of applications. At the end of 2004, 2.2 tera-bytes of image data was produced and four data backups were made in different media. The cost of digital image management was increased sharply and was expended in both storage system and crew. The storage will never be enough even when accompanied with the development of digital disks and magnetic tape technologies.

The master file of the one-shot digital camera back is camera RAW, which is true data recorded from the CCD. The other image files could be distributed from the camera RAW. Whether to only keep the camera RAW and distribute by request or to process at once and use forever are different views. We chose a safe way to produce all of different kinds of image files for easy usage and reduced time to process in the future. Table 1 shows the file formats and purposes of image systems of DAPCA.

Table 1. DAPCA Image File Formats.

	1st Imaging system (6M)	2nd Imaging system (22M)	Purposes
RAW file	13MB	49MB	As digital film and master file for archiving or reproducing
48Bits RGBtiff	36MB	130MB	Main Archive File
24Bits RGBtiff	N/A	64MB	For large size printing or up to A3 size publishing
24Bits RGBtiff20MB	18MB	18MB	For A4 size publish
JPEG 200K	50~350KB	50~350KB	For the Internet display
Each image produces	67MB	261MB	

Output Quality Inconsistent

Digital image is convenient for information sharing on the Internet or physical hardcopy. DAPCA defines the main purpose of digital images is to print out at least A4 size in catalog. Though we have some printing cases to verify the publishing effect of digital images, the output quality is not easy to be controlled. Because there is lack of CMYK standards in Taiwan, the output quality is difficultly evaluated between publishing providers.

The Future

There are many researches focusing on new technologies applied on image reproduction such as the multi-spectral imaging system. But the questions are when this system is available and how precise it can become. Can it apply in the image reproduction of 3D objects or only for 2D painting? We hope a sound solution will soon be available.

Conclusion

The characteristics of digital digitization of DAPCA are a large number of a variety of antiquities and the ability to digitize in a short period. We developed two digital imaging systems to enhance the efficiency of digitization and implemented ICC color management workflow for color consistence. Even though the color management in practical imaging is not perfect; it is still the only way toward the solution of cross-media color reproduction at present.

In the future, we will try to implement the application of 3-dimensional imaging technologies in a digital archive of antiquities. We believe 3-dimenasional images can represent the better display of antiquity legacy. In addition, the quality improvement of image reproduction is still the main point during the digital imaging process. In this case, three core issues of images digitization are workflow, quality measurement and storage that deserve improvement for producing better digital images. Finally, we hope to extend the next phase of DAPCA with another five years from 2007 to 2011 after the 1st phase is completed. Its valuable experience of images digitization will also be outstretched and be the foundation for further study on imaging science.

Acknowledgements

The authors would like to express their gratitude to the staff of DAPCA for their dedication to this project. We would also like to acknowledge the work of many members of the Information Management Center who provided the information of the NPM 3D Virtual Collection Exhibition System, particularly Chiu Ping Wang. Without their help it would not have been possible to gather this experience.

References

1. Abhay Sharma, Understanding Color Management, Thomson Learning, NY, 2004, pg. 326.
2. Beuce Fraser, Chris Murphy, and Fred Bunting, Real World Color Management, Peachpit Press, CA, 2003, pg. 266.

Biographies

C. Alan Cheung received a M.I.S. from Central Queensland University in 2001. He joined the Digital Archives Project of Chinese Antiquities at the National Palace Museum in 2002. He has had a variety of roles in museum information management and systems development. Currently he is a research assistant and works with curators to aid collections management system operating and artifacts imaging. His research interests concern applications of digital images and imaging system development. E-mail: alanpop@npm.gov.tw.

Mei-li Yang, the curator of the Department of Antiquities at the National Palace Museum, specialize in ancient Chinese jade-objects and archaeology. The chief of the Digital Archives Project of Chinese Antiquities at the National Palace Museum. 221, Chih-shan Road Section 2, Shih-lin, Taipei, Taiwan, 111 R.O.C., Tel.: +886-2-2882-1230 ext. 514; Fax.: +886-2-2882-1440; E-mail: YML@npm.gov.tw