Breathing New Life into Static Materials

Jieh Hsiang  
Department of Computer Science, and  
Digital Center for Digital Humanities  
National Taiwan University  
Taipei, Taiwan  
jhsiang@ntu.edu.tw

Shu-Fen H. Lin  
National Taiwan University Library  
National Taiwan University  
Taipei, Taiwan  
shufen@ntu.edu.tw

Abstract
Many libraries have large quantities of books that are rarely used. A practice that is becoming increasingly popular is to put them in remote repositories, which often only expedites their pace into oblivion. We call such collections static materials because of their extremely low usage rate. In this paper we propose a method which, through digitization, gives them a new life. Our approach starts with identifying a feasible objective that could make a ‘collection’ emerge from certain static materials, then systematically extract and digitize objects of the selected theme from those materials. The resulting collection from such a practice may be unique, significant, and useful.

We demonstrate this idea by describing a digital library of old photographs of Taiwan, extracted almost exclusively from existing pre-1945 Japanese books held in the National Taiwan University Library. This collection, numbered at 39,743, covers a wide spectrum of topics, and fills a visual documentary void in Taiwanese history. This digital library can be accessed at http://photo.lib.ntu.edu.tw/pic/db/oldphoto.jsp

1. Introduction
Many libraries have large volumes of books that are rarely used. Although these materials present a heavy burden to the overcrowded stack space, libraries are reluctant to dispose of them due to their potential cultural significance and rarity. One course of action is to build remote storage to keep the books [1] and retrieve them for use only upon request. This method, however, makes these materials even more difficult to access and expedite their path to oblivion [11].

In this paper we present an approach that creates new and significant collections from a library’s static materials. We use the term static materials to indicate library holdings that have (1) extremely low usage and (2) no more incremental issues. Typical static materials include old and rarely circulated books, and discontinued periodicals. An important observation is that static materials might contain hidden information that, collectively, reveals unexpected significance. We illustrate this point through describing how we constructed a unique Digital Library of Taiwanese Old Photographs exclusively from static materials kept at the National Taiwan University Library. We outline the steps used in creating this digital library and show the significance of the outcome. We also suggest that this strategy can be applied to other similar cases, especially in countries that are formerly colonized by western power and are trying to preserve their cultural heritage.

2. Contents of the digital library of old Taiwanese photographs
Some years ago the National Taiwan University Library was approached by Professor Wu Micha of the Department of History, who posed the problem that there was no significant collection of pre-1945 photographs of Taiwan. He indicated that such a collection of visual records would greatly benefit research on Taiwan history and anthropology.

Taiwan had been ruled by different people throughout its history. The island was a part of the Chinese Qing Dynasty until 1895, when it was ceded to Japan. Japan colonized Taiwan between

1 The first recorded government was the Dutch who, via the Dutch East India Company, ruled a portion of Taiwan between 1624 and 1662 before being driven out by the Chinese.
1895 and 1945, and returned it to the Chinese after losing the 2nd World War. The dramatic change of the island during the Japanese colonial rule had become an increasingly popular research topic in recent years. A comprehensive photographic collection of that era will provide valuable visual evidence to the researchers.

The NTU library did an extensive survey of existing pre-1945 photographic collections, and found that there were indeed no large scale collections of this type. The photo collections that we did find were all about China, but not Taiwan. That is because those collections were mainly kept by news or government agencies, which did not establish a permanent base in Taiwan until the Nationalist government moved its seat to Taipei in 1949. For example, the Central News Agency (CNA), the official news agency of Taiwan, has a reserve of over one million photos, but those about Taiwan were almost all dated after 1949.

It is conceivable that relatively few photos of Taiwan from the 19th century can be found except those made by western travelers such as John Thomson [13]. Local photographing facilities were non-existing in Taiwan at that time. It is, however, somewhat surprising that there was no comprehensive photographic collection(s) from the Japanese era. The Japanese colonial government conducted extensive and systematic surveys of the island since the beginning of its occupation. The field studies were often performed by academic groups (such as the anthropology research team of the Taihoku Imperial University), and were usually accompanied by photographers (e.g. [2]). Although they were not compiled into photo albums, it is conceivable that many of them might have appeared in the government issued books/reports that were produced followed the field studies.

Based on this hunch, we proposed a possible solution to the challenge of establishing a photos library of the time by utilizing the static materials in the NTU Library. The National Taiwan University houses a large collection of old Japanese books inherited from its predecessor, the Taihoku (Taipei) Imperial University. The latter was founded in 1928 and became NTU in 1945. The university, being the only imperial university south of the Japan proper, was also a repository of books and other materials collected or published by the Taiwan’s Colonial Government. We examined the portion of this collection that is related to Taiwan, numbered at about 5,000 titles, and found that a large portion of the books indeed contained photographs scattering among the pages. A further study of the books that contain photographs enabled us to divide them roughly into three categories:

(1) Books that commemorate special events:

The colonial government published volumes, often with pictures, to celebrate important constructions, the visitation of dignitaries, special exhibitions, etc. There were even photo albums published on very special occasions such as a formal visit by the crown prince. Although these books were mainly for propaganda purposes, they often contain high quality photographs that witness events of the day.

(2) Books as reports of surveys:

The Japanese Colonial Government conducted extensive surveys of a great diversity of subjects such as architecture, historical monuments, landscape, natural resources, industries, agriculture and fishery, forestry, local customs and beliefs, etc. The surveys were conducted periodically and systematically throughout the Japanese colonial era. Publications of these surveys were often supplemented with photographs as an aid to give more visual description of the findings. These photos cover a wide range of subjects and span over the entire fifty years of Japanese rule.

(3) Books about indigenous Taiwanese:

The Chinese residents of Taiwan during the Qing Dynasty were mainly Han people from Fujian and Guangdong provinces of China, while the indigenous Taiwanese were Austroesians. The plain indigenous people (Pinpu) either gradually assimilated into the Han society or moved in-land. The mountain tribes, on the other hand, were feared and considered savages. The Qing government’s approach was to confine them in their mountain dwellings and keep them away from the more fertile land where Han people lived. But little effort was made to understand or to assimilate them. In order to rule the island more effectively, the Japanese made extensive anthropological studies of the Austroesians. After rebellions that were brutally suppressed (see, e.g. [5]), the mountain people were more or less pacified, although assimilation efforts largely failed. Many photographs about the indigenous Taiwanese were included in the anthropological surveys sponsored by the Colonial Government. They were among the earliest visual records of the indigenous Taiwanese and were invaluable
documentaries of their earlier culture before the invasion of modern civilization.

3. Digitization process and considerations

The staff in the Library spent a few months poring over the 5,000 volumes of books and labeled all the photos in the books. A label assigned with a unique identifier\(^2\) was attached to each photo for the reference of scanning and the metadata construction. This was to ensure that the link between the image and its correspondent metadata record can be established once the digitization process begins. We eventually came up with 39,743 photographs in total. The next step was to digitize these photos and to establish their metadata. The photos were scanned at either 300 dpi or 600 dpi with 24 bits/pixels. They were kept in tiff format and later compressed into lower resolution jpeg files for web browsing. The image quality of the photos varies due to the divergent printing quality of the original source.

For metadata we followed the Historical Photos Metadata Format (HPMF) developed for photographs by the Council for Cultural Affairs\(^3\) so that they are compatible with the other photographic collections in Taiwan. HPMF was developed for the National Repository of Cultural Heritage Initiative (NRCH) [7], which eventually established a digital library containing over 1 million digital artifacts of Taiwanese culture (http://nrch.cca.gov.tw). In order to manage the great variety of subjects in this digital library, CCA developed 15 different metadata formats, one for each of the 15 categories of materials contained in the NRCH [3]. HPMF is based on the Dublin Core Element Set [14], with each attribute further expanded to accommodate to the needs that are pertinent to photographic collections.\(^4\) As with other DC-based formats, HPMF records are easy to produce since they do not contain many attributes. But it is also less powerful than other standards such as IPTC Photo Metadata [12] or the MARC-based metadata used in LC’s American Memory [10].

The metadata records of the photos were created by the library staff, who carefully read the descriptions that accompanied the photos and then produced the records. Because of the rich information and variety of the photos, we also created a subject classification scheme [9] which contains 15 major subject categories and further subdivided to 71 more precise classes. Each photo is classified into at least one of the subjects.

A retrieval system was built for this photographic collection. In addition to providing keyword and metadata search, our system also emphasizes on the ease of use and faceted navigation of query results. The facets include such attributes as title, year, keyword, related person, source, etc. The system can be accessed at http://photo.lib.ntu.edu.tw/pic/db/oldphoto.jsp

In the following we show two photographs taken from this collection. Picture 1\(^5\), published in 1932, shows the Dawu (達悟) people of the Lanyu Island. Dawu are Austroesian Taiwanese who are closely related to the Batane people of the Philippines. It is interesting to note that the visor in the picture is made from melted Mexican silver coins.

![Picture 1. Dawu people of Lanyu](http://photo.lib.ntu.edu.tw/pic/db/detail.jsp?dtd_id=13&tid=32&id=30277&pk=seq&showlevel=2)

The next photo\(^6\), taken in 1925, shows a sampan sailing on a major irrigation canal, the Chia-Nan Grand Aqueduct, of southern Taiwan.

---

\(^2\) In addition to the institution code and ‘materialType’ code (HP -- abbreviation for ‘historical photographs’), the Library utilizes the accession number of each book as the primary part of the file name, then adds the number of page that the photo appears, followed with a number indicating the order of the photo that appears in that page.

\(^3\) http://km.cca.gov.tw/download/rule/metadata_hphoto.pdf, retrieved 3 May 2010

\(^4\) For example, for this Digital Library of Taiwanese Old Photographs, we further expanded the DC attribute of “type” into “originalSurrogate” and “worksType”, and further expanded the DC attribute of “format” into “medium” and “extent” (for describing quantity, dimension, and color).
4. Significance of the outcome

The significance of the resulting collection far exceeded our expectation. Not only is the collection unprecedented in terms of quality, quantity, and variety, the photos are also easy to search and use because of the rich descriptive information extracted from the books.

We summarize the important features of this collection of photographs as follow:

(1) **Volume of photos of an important era**

The total number of photographs is 39,743, covering Taiwan from 1895 to 1945. After eliminating duplications\(^7\), there are more than 37,000 photographs in the final Database of Taiwanese Old Photos. It is the largest (and only) collection of its kind as far as we know.

(2) **Wide spectrum of subjects coverage**

The collection covers a surprisingly wide spectrum of subjects. We did a careful subject classification \(^9\). The major subjects alone are divided into 15 categories. They are Local Beliefs, Indigenous Peoples, Geology and Geography, Landscape, Historical monuments, Architecture, Industries, Agriculture and fishery, Forestry, Education, Daily Life of People, Flora, Fauna, Security and medicine, and Transportation, which are further refined into 71 sub-categories.

(3) **Informative metadata descriptions**

Photographs, especially old ones, are hard to use because of the lack of descriptive and background information. In our case, however, the photos were extracted from books, which usually accompanied a photograph with some description. This has made metadata creation much easier and much more extensive than many other photographic collections. It also made search and retrieval more productive.

(4) **Clear copyright**

The use of digital collections is often plagued by the issue of copyright. This is not a problem in our case. It is because the photographs were extracted from publications of the Japanese Colonial Government of Taiwan, which ceased to exist in 1945. Since the official publications of a defunct agency fall into public domain after 50 years from its demise under the laws of both Taiwan and Japan, our digitizing and using of them does not violate copyright laws. This is another important reason why static materials are good candidates for producing produce special-purpose digital collections than more contemporary materials.

5. Concluding remarks

In this paper we demonstrated how an extraordinary collection of pre-1945 Taiwanese photographs was created from static materials of the National Taiwan University Library. Not only the largest and the most comprehensive of its kind, this collection has also become an important source of visual material for researchers. Its contents are available on the web at two locations: as a part of the National Repository of Cultural Heritage (http://nrch.cca.gov.tw) and as a stand-alone collection at the NTU Library. While the former is for popular use, the latter is mainly for scholars (although open to the public). The NTU site has received about 12,000 hits last year (from 8 countries), according to Google Analytics. Although we do not have the usage data of the CCA site, their NRCH website has been one of the most heavily used source of cultural materials in Taiwan. Photos from our collection are now a common sight in related scholarly talks and papers (e.g. \(^8\)). We have also received requests from publishers and museums to include our photos in books and exhibitions. We have already authorized 291 photos for commercial use as of June 2010. Evidently our work has filled a void in the visual documentary of Taiwanese history.

We should emphasize that the digital library described in this paper is merely an example of
how a remarkable collection can be created exclusively out of static materials in a library. One can think of many situations where such an approach may apply. For instance, countries that were formerly colonized and established after the 2nd World War may have similar kinds of static materials and similar needs. The former colonial governments usually have kept comprehensive records, but they might be scattered in libraries and archives and are difficult to access. While one could consider digitizing the records in their entirety, but that often comes with a prohibitively high cost and may not be a priority on the national agenda or the holders of the materials. However, by selecting only relevant portion of the collection to digitize, the cost may be reduced significantly and the result may be beyond expectation.

The approach that we advocate can be summarized as follows:

1. First identify a need.
2. Identify a collection that may contain material that satisfies the need. We recommend starting from considering static materials in a library because (a) they are usually not protected by copyright, (b) they are rarely used (and the library that houses them often welcome such a “reincarnation” project because it brings visibility and usage), and (c) they contain important material that are not commonly available.
3. Determine the digitization and metadata standards suitable for the purpose. While scanning and metadata are essential, we also recommend constructing comprehensive descriptive information whenever possible.
4. Digitize and build a user-friendly system. Additional work such as subject classification will enable the construction of a subject-browsing search interface and make the search result easier to browse and navigate. Therefore, we also suggest that a subject classification should be done whenever possible.

6. Acknowledgements

The authors wish to thank the National Taiwan University Library, the Council for Cultural Affairs, and the National Science Council of Taiwan for their generous support of this project. They also thank the anonymous referees for their constructive comments.

7. References