USING STANDARDS AS A BRIDGE BETWEEN TRADITIONAL RESEARCH AND TECHNOLOGIES IN PROTECTING CULTURAL ASSETS

Yu Fujimoto[1], Yasuhiko Horiuchi[2]

[1]Doshisha University in Kyoto, Faculty of Culture and Information Science
Tatara Miyakodani 1-3, Kyotanabe city, Kyoto, Japan. yfujimot@mail.doshisha.ac.jp

[2]Horiuchi Photo Office
Showa-dai cho 2-27-2, Takatsuki city, Osaka, Japan. horiuchi-yasuhiko@cpost.plala.or.jp

Keywords: Standardisation, Work-Oriented Approach, ISO 191XX, UML Activity Diagram, Scripting

Abstract:

The magnitude 9.0 earthquake and tsunami, which hit the northeastern coastal regions of Japan on 11 March 2011, caused immense damage not only to the residents and social infrastructure but also to cultural asset in these areas. In a press release dated 6 May 2011, the Ministry of Education, Culture, Sports, Science and Technology stated that at least 529 registered cultural assets were damaged and we have yet to grasp the conditions of numerous unregistered cultural properties [3]. To survey the damage to unregistered cultural properties, many academic and governmental sectors are preparing for field investigations in the afflicted areas; however, most of these entities have scant time for discussing survey items or guidelines for cooperating with one another. Striking a strong balance between standardisation and urgency is a critical problem. The Consortium for Earthquake-Damaged Cultural Heritage (CEDACH) was launched on 21 March 2011 to survey and protect cultural properties in afflicted areas. CEDACH members are discussing a new method for cultural resource management, which achieves a bridge between traditional research and technologies. Our approach in this paper is a Work-Oriented Approach that focuses on the process of participants' daily works; our approach is a candidate for adoption by the CEDACH task force [4-5]. We have already issued both our application schema, which is designed to conform to the ISO 191XX series and its prototype Survey Data Archivist for Client (SDA for Client) software [15]. However, in this previous work, we have not considered post-processing tasks, such as modifying, editing and publishing. We therefore propose an advanced documentation for recording intermediate procedure. The Unified Modeling Language (UML) activity diagram and command line scripts are used to record these data processing tasks without the need for intermediate files. In this paper, we show an image processing example in which videos are first acquired from a field, still images are extracted from the videos, then integrated together to create a high-resolution still image. Although the series of image processing steps for obtaining high-resolution images could not achieve the expected results because of varying specifications of current personal computers, we conclude that our approach would be more efficient in specifying the workflows.

1. INTRODUCTION

1.1 The impact of the earthquake and tsunami on cultural assets

The magnitude 9.0 earthquake and tsunami, which hit the northeastern coastal regions of Japan on 11 March 2011, caused immense damage to the residents and social infrastructure. In particular, the estimated 35-m-high tsunami deposited much debris in the area [1]. According to the numbers reported by the National Police Agency on 26 April 2011, there have been more than 14,000 deaths and over 11,000 missing victims, numbers that are expected to increase [2]. There are also many missing civil servants throughout the affiliated area, bringing further chaos. Cultural assets in afflicted areas are also seriously damaged. In a press release dated 6 May 2011, the Ministry of Education, Culture, Sports, Science and Technology stated that at least 529 registered cultural assets were damaged [3], and we have yet to grasp the conditions of numerous unregistered cultural properties. In response to this situation, Japanese academics with interests in cultural studies are developing practical and efficient methods for collecting and managing information regarding damaged unregistered cultural assets.

1.2 Consortium for Earthquake-Damaged Cultural Heritage (CEDACH)

When large-scale disasters strike, restoration processes are typically divided into three phases: the rescue phase; the rehabilitation phase and the reconstruction phase. As of the middle of April 2011, the current focus is still on the rescue phase that puts life saving efforts before everything else. Full-scale investigation and protection activities for cultural properties will be included in the reconstruction phase; nonetheless, we need to prepare for full-scale field surveys that will be conducted.

Furthermore, several academic and governmental sectors have already entered some of the afflicted areas to perform rescue surveys, because many anticipate that important cultural properties may be discarded by mistake or auctioned to third parties. These first rescue attempts aim to prevent such problems. Groups are also concentrating on collecting information about the conditions of damaged cultural assets. For example, Historical Resources Networks (HRN), which was established in response to the Hanshin–Awaji earthquake that hit Kobe city in 1995, immediately tried to be in contact with museums located in the most damaged areas and reported information about the well-being of buildings, employees and cultural properties.

The Consortium for Earthquake-Damaged Cultural Heritage (CEDACH) was launched by archaeologists, historians and cultural and information scientists who regularly discuss this incredible disaster on social networking sites such as Facebook and organised a task force for surveying and protecting cultural properties damaged by the earthquake and tsunami [4-5]. The CEDACH task force is comprised of a Data Management Team (DMT) and an On-site Technical support Team (OTT). The CEDACH-DMT provides local Cultural Resource Management (CRM) officers with guidelines for measuring and conserving damaged cultural properties, and is currently working on developing a database schema and data collecting system for field surveys, which require a good balance between urgency and standardisation. In addition, all acquired datasets should have reproducibility and further validation by third parties.

To help realise these goals, we have developed a Work-Oriented Approach (WOA). In this paper, we discuss WOA and propose a new guideline for emergency data collection, which provides a bridge between traditional research and new technologies.

2. APPLICABILITY OF THE ISO 191XX SERIES TO CULTURAL ASSETS

2.1 Problems with digital information

Cultural assets are continually in danger of destruction because of disasters, urbanisation, civil wars and so on. Survey records are essential in such cases to reconstruct cultural assets. As noted above, Japanese academics are now confronting such a problem after the earthquake and tsunami that hit the northeastern coastal regions of Japan. In reconstructing damaged (or lost) cultural assets from digital copies, survey records must conform to specific standards and must be verified by independent third parties. CEDACH aims to develop such guidelines for CRM.

For most of this decade, information and communications technology (ICT) has been regarded as an essential tool for cultural studies. In archaeology, for example, measurements are generally conducted using ranging and sensing technologies; however, as the application of digital technologies to cultural field surveys has increased, information management methods and standardisation have become serious issues for researchers engaged in cultural field surveys.

Contrary to traditional hardcopy information management, digital information management on the computer is more complicated; although digital information seems to have a longer lifespan than paper-based information, the case strongly depends on data management guidelines. Much of the raw data collected tends to be stored and managed on individually owned computers with individual file management rules. Suppose the lifespan of one personal computer is five years, important data should be copied to other secure storage devices during this period or important survey data might be completely lost when the computer is scrapped or crashes. Furthermore, individual data management rules prevent smooth data exchange among different storage devices, as the data is often isolated. Therefore, we face two significant questions regarding digital information: (1) where should we store files? and (2) what standard(s) should we use for these files?

2.2 ISO 191XX in Japanese archaeology

Table 1: Important standards in ISO 191XX series

Name of standard	Description
ISO 19107 (Temporal Schema)	Defines primitive and complex spatial objects.
ISO 19108 (Spatial Schema)	Defines temporal objects such as "instant" and "period".
ISO 19109 (Rules for application schema)	Specifies how application schema should be designed.
ISO 19115 (Rules for metadata)	Defines objects for describing exchange metadata.
ISO 19118 (Encoding)	Specifies rules for encoding instances in XML documents.
ISO 19136 (GML: Geography Markup Language)	

There are currently three international information standards used in Japan for cultural properties management: (1) the Dublin Core metadata standard; (2) the CIDOC Conceptual Reference Model (CRM) for existing museum databases and (3) the ISO 191XX series, a series of information standards for geospatial phenomena, for defining geospatial phenomena [6-8]. The Dublin Core metadata standard has been adopted by the National Diet Library (NDL) and is used for bibliographical information management. CIDOC-CRM is used for museum management. The ISO 191XX series is considered a new information standard for Japanese archaeology. The Archaeological GIS Modelling Group (AGMG) and Nara National Research Institute for Cultural Properties (NNRICP) are working together to develop application schemas in accordance with this standard.

Developing application schemas is especially essential in the ISO 191XX series; by applying ISO 19109, suitable standards for different application domains can be defined. Using this set of standards, a common and correct understanding of the content and structure of cultural asset may be attained. When developing application schemas in the ISO 191XX series, it is recommended to use Unified Modeling Language (UML) class diagrams; all objects should be defined with other standardised schemas, including ISO 19107 and ISO 19108. Although instances should be encoded in conformance to either ISO 19118 or ISO 19136 as they are exchanged among other application systems, all instances would be managed by application-specific methods, such as those incorporated into a relational database management system (RDBMS) (Table 1) [9-13].

Conforming to the ISO 191XX series, AGMG issued the first version of an application schema for archaeological information on 30 March 2011 [14]. This standard specialises in making survey reports as deliverables. Because archaeological site measurements are generally consigned to third party survey companies, such information standards are essential.

Despite the fact that this application schema was developed pursuant to international rules, its scope is limited to Japanese archaeology. More specifically, classes are defined according to traditional Japanese-style survey reports and applying current digital technology to them is not easy. The reason for this is the underlying traditional yet dominant anti-standardisation sentiments among literary intellectuals. They tend to demand that private sector companies make traditional survey reports; thus, the application schema as issued by AGMG has no alternative but to focus on and mimic traditional survey reports. Consequently, this application schema best serves survey companies rather than archaeologists.

3. WORK-ORIENTED APPROACH (WOA)

3.1 Basic Concept of WOA

To develop a new application schema better suited to archaeologists, we propose a Work-Oriented Approach (WOA) that attaches importance to what one surveyor acquires in his or her daily work. This approach consists of three layers of standards that correspond to three types of cultural survey workflows. The bottom layer, which consists of the rules for acquisition, defines a standard protocol that specifies how digital information is created and edited. The second layer defines an application schema for day-to-day data management, such as rules for storing the data. The top layer defines an application schema as a substitute for publishing traditional survey

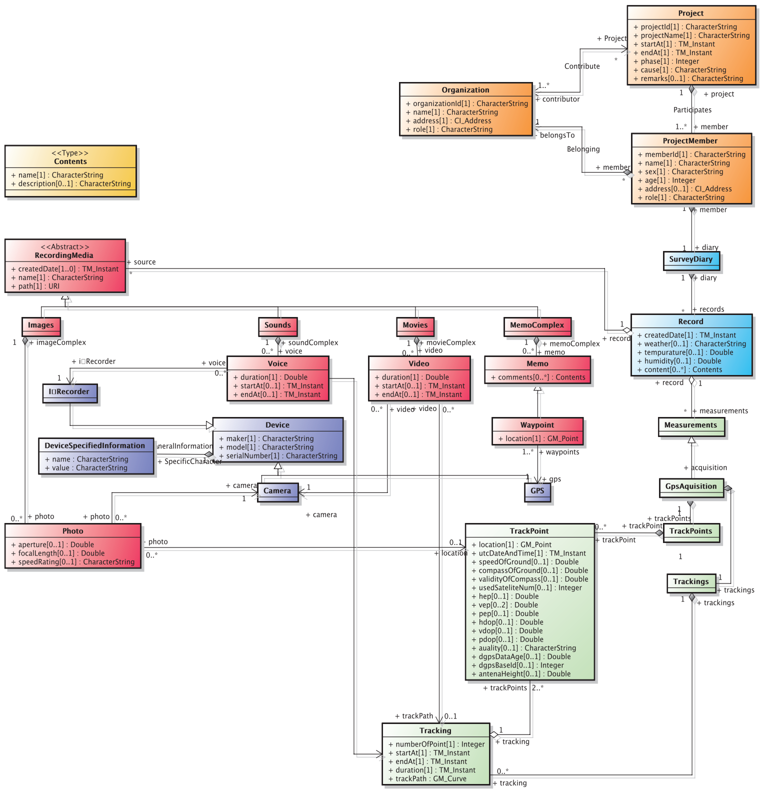

Figure 1: The application schema for second layer of WOA, which was adopted for SDA for Client.

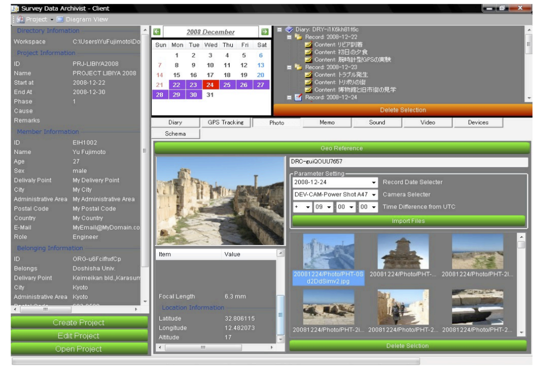

Figure 2: A screen shot of SDA for Client

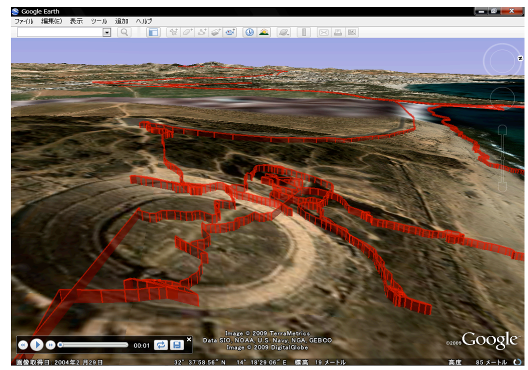

Figure 3: GPS tracking data exported to Google Earth by using a function of SDA for Client.

reports. The application schema issued by AGMG corresponds to this top layer. Instantiated datasets at the top layer are converted (or remapped) from instances at the second layer by model conversion. By separating the working levels from the purposes, the workflows and the rules that are applicable to each, the integration of traditional customs and cutting-edge technologies becomes much easier than attempting to develop and propagate a single global standard.

Applying the concept of the second layer of WOA, we have already issued both an application schema that is designed to conform to the ISO 191XX series and its prototype software, Survey Data Archivist for Client (SDA for Client) [15]. In the application schema shown in Figure 1, orange coloured classes define information related to survey projects; blue ones define survey diary as records of comments about daily work; and purple ones define digital devices used for data capturing. Red and green coloured classes are survey objects, which are acquired in day-to-day work; red ones represent multimedia objects and green ones are measurements objects. In accordance with this schema, SDA for Client properly stores digital copies of these survey objects at specified directories and manages additional information. For example, photos (e.g. JPEG images) are stored at specified locations at first, then each photos' metadata, which is written in Exchangeable Image File format (EXIF) is extracted. Finally, the metadata and GPS positions (acquired by GPS logger) are combined and rewritten to an XML document in conformance to ISO 19118. Additionally, the software has the feature that allows the exporting of any geographic information dataset to Google Earth (Figure 2).

Although our previous work was originally designed for archaeological field surveys, it is also applicable to rescue surveys. As such, CEDACH is adopting the WOA concept as a methodology that achieves a good balance between standardisation and urgency. We are currently improving our SDA for Client for this purpose.

3.2 Method for specifying working processes

To date, efforts for standardisation have primarily focused on the ISO 191XX series; in general, the bottom layer has not yet been discussed. As noted above, the bottom layer's focus is on recording and storing intermediate

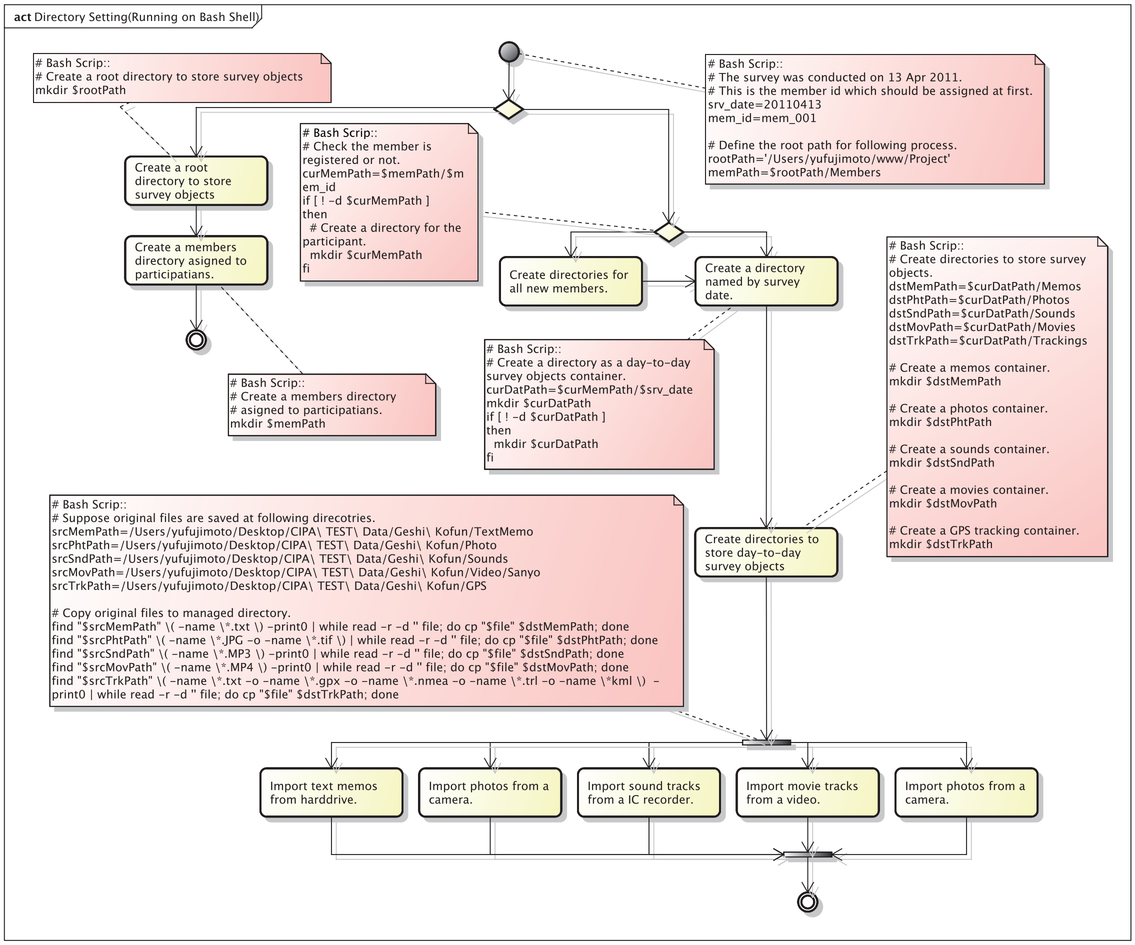

Figure 4: An example workflow described using UML activities and bash scripts

data created during a certain project. The nature of the bottom layer is different from that of the top and second layers. The ISO 191XX series, which is adapted for these layers, defines the rules for modelling a static structure and its encoding. In other words, the standard provides rules for designing database schemas. Therefore, the series is not applicable to the bottom layer, because the aspects of dynamic practical workflow are beyond its scope.

To overcome this limitation, we consider the efficiency of UML activity diagrams and scripting. By using UML activity diagrams and describing the scripts for the data processing of comments, the data handling process becomes clear. Figure 4 shows an example UML activity diagram for creating directories and copying files acquired in a field survey. As a result of this process, directories are structured in the same manner as the SDA for Client.

In the figure, rounded rectangles represent activities, whereas dog-eared rectangles indicate notes that serve as a space for writing down remarks. Although the Object Constraint Language (OCL) and natural language are used for notes, command line scripts may also be described, serving as actual implementations corresponding to activities. Therefore, an activity could be performed by using the equivalent command line script. Scripts denoted in the figure are written for the bash shell (i.e. runnable on a UNIX shell); the first line in each note indicates the scripting language used. Although scripting languages should be unified, we do not want to constrain the scripting language, because various software packages are used for survey projects. The most important point here is that the system for running scripts and the scripts themselves should be clearly specified in an acceptable and uniform manner.

3.3 A case study at Geshi–Kofungun

Various results can be properly duplicated and evaluated by using UML activity diagrams and scripting

Table 2: Specifications for the case study

CPU	2.66 GHz Intel Core 2 Duo
Operating System	Mac OS X (Version 10.6.6)
Memory	4GB 1067 MHz DDR3
Assigned memory for ImageJ	3067 MB
Threads for processing	2
Number of images per stack	4 images

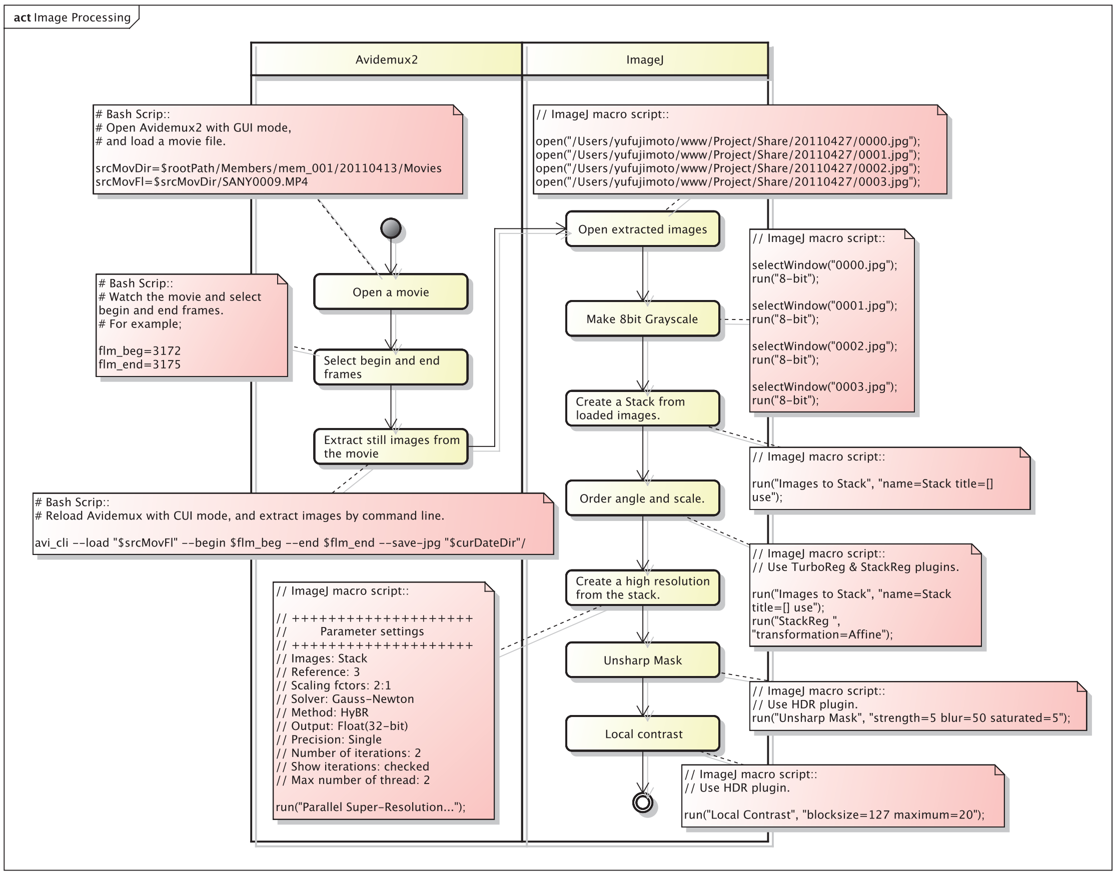

Figure 5: A workflow for creating a high-resolution image from a movie file

command lines. In this section, we evaluate our approach in the case of image data of Geshi–Kofungun, a small group of mounded tombs built in the first half of the 7th century located on the Doshisha University campus. At this archaeological site, we recorded videos and attempted to create a high-resolution single image for a survey report as a part of developing new methodologies for rescuing cultural assets. At present, many field surveyors are paying much attention to the efficiency of using videos. One Japanese company is developing a three-dimensional measuring system for ordinal movie cameras [16]. Since this tendency to use video cameras will likely grow in the near future, developing new methodologies for field survey work using video is important. In the Geshi–Kofungun case study, we tested our approach on a MacBook Pro with specifications summarized in Table 2. We used two open-source software systems: (1) Avidemux2 for editing movies and (2) ImageJ for creating high-resolution images [17–18]. Because both of these software systems are command-line driven or support macro scripting, they are suitable for our approach. In addition, open-source software are free available; thus, negating any financial problems.

Figure 5 (in particular the left side of the figure) shows the workflow for creating a high-resolution image from a movie file; as an initial condition, the acquired movie file is copied to a directory according to the process shown in Figure 4. As depicted in the Figure 4, the first step is to open a movie file (named ‘SANY0009.MP4’) using Avidemux2 in GUI mode, then manually select an important scene and specify the start and end frames for

Figure 6: A scene extracted from a movie file (left) and a final product created from five low-resolution images (right)

extraction. In this example, the sequence from the 3172nd frame to the 3175th frame is selected. Next, close Avidemux2 GUI mode and perform the extraction command via the shell. As a result, five JPEG files are extracted from the movie file and stored in the 'current date' directory. Contrary to day-to-day field work, data processing work, such as image editing, is often conducted in teams; therefore, such intermediate image files should typically be managed on a shared globally accessible directory.

After the above process is completed, ImageJ is launched and the extracted five images are loaded. To create a high-resolution image, the StackReg (and TurboReg) and Parallel Super-Resolution plugins should first be installed; StackReg is a plugin for recursive alignment of a stack of images and Parallel Super-Resolution is a plugin that creates a high-resolution image from a stack of images [19–20]. In addition, in this study, the HDR plugin is used for tone mapping [21]. The processes for creating a high-resolution image and performing tone mapping are illustrated on the right-hand side of Figure 5.

Equivalent commands for these processes are written in the ImageJ macro language, which is similar to JavaScript. ImageJ has a function to record macros, similar to the concept in Microsoft Excel; by using this functionality, scripts are easily defined. Many of the ImageJ macros can include parameters for image processing, while some plugins do not include such parameters. In the latter case, required parameters are denoted as commented-out lines; for example, in Figure 3, since the *run("Parallel Super-Resolution…")* operation is unable to include parameters, parameters are described as 10 lines of comments before the operation.

Figure 6 shows the final product created by using our process. Because the Parallel Super-Resolution plugin supports small low-resolution images, the results were limited; we determined that a more powerful computer would be needed for this process. Furthermore, a higher scale factor and many more images would be required to achieve better results.

4. CONCLUSION

We found that our attempt to describe workflows using UML activity diagrams and command line scripts was beneficial. The final product (i.e. images) could be easily reproduced from the original raw data, the directory structure, the command line scripts and the workflows.

We conclude that our approach is also applicable for creating planning documents. In large-scale collaborative projects, it is difficult to merge the results of various methodologies adopted by each project team, especially in urgent rescue surveys. In contrast to merging different methodologies, defining uniform rules for the documentation of working processes is easier. Furthermore, this method serves as a bridge between traditional research and technologies. As a result, our approach is effective for managing various datasets acquired and created by third party projects teams.

Generalisation of our approach, however, remains a future task. There are cases in which some members of a project team have strong technical skills. In other cases, designing UML activity diagrams and scripting command line operations may be difficult for participants who are not so strong in using or relying on technology. Furthermore, we also need to evaluate our approach for measurements that are acquired by ranging and sensing equipment.

5. ACKNOWLEDGEMENT

We heartily appreciate the members of CEDACH for their comments and suggestions that were of inestimable value to our study. In addition, we believe that our approach will contribute to the restoration of damaged cultural assets caused by the earthquake and tsunami, which hit the northeastern coastal regions of Japan.

6. REFERENCES

[1] Enomae, T. and Higashijima K.: *Cultural heritage damaged by the 2011 Tohoku earthquake and the need for recovery aid*, Proceedings of 39th Annual Meeting of the American Institute for Conservation, Philadelphia, PA, USA, 31 May–3 June 2011.

[2] National Police Agency: *Information about East Japan great earthquake*, http://www.npa.go.jp/archive/keibi/biki/index.htm.

[3] Ministry of Education, Culture, Sports, Science and Technology : *Information related to East Japan great earthquake*, http://www.mext.go.jp/a_menu/saigaijohou/index.htm.

[4] Kondo, Y. Kaneda, A. Fujimoto, Y. Seino, Y. Yamaguchi, H. Uozu, T.: *Cultural resource management in the disaster zone: A report on a voluntary initiative for the information-gathering of the earthquake-damaged cultural heritage in Japan*, The 39th Annual Conference of Computer Applications and Quantitative Methods in Archaeology, Beijing, China, April 2011.

[5] CEDACH: The Consortium for Earthquake-Damaged Cultural Heritage, http://cedach.org/, CEDACH.

[6] ISO: *Information and documentation -- The Dublin Core metadata element set*, ISO 15836:2009, International Organization for Standardization, 2009.

[7] ISO: *Information and documentation -- A reference ontology for the interchange of cultural heritage information*, ISO 21127:2006, International Organization for Standardization, 2006.

[8] ISO/TC211: *Geographic Information / Geomatics*, http://www.isotc211.org/.

[9] ISO: *Geographic information -- Spatial Schema*, ISO 19107:2003, International Organization for Standardization, 2003.

[10] ISO: *Geographic information -- Temporal Schema*, ISO 19108:2002, International Organization for Standardization, 2002.

[11] ISO: *Geographic information -- Rules for application schema*, ISO 19109:2005, International Organization for Standardization, 2005.

[12] ISO: *Geographic information -- Encoding*, ISO 19118:2004, International Organization for Standardization, 2004.

[13] ISO: *Geographic information -- Geography Markup Language (GML)*, ISO 19136:2007, International Organization for Standardization, 2007.

[14] Nara National Research Institute for Cultural Properties: *Product specification for making Geo-spatial dataset based on Archaeological features modelling*, CAO News Center for Archaeological Operation, vol.144, March 2011. (独立行政法人国立文化財機構奈良文化財研究所文化財センター，2011，「遺跡情報モデルに基づく地理空間データ作成のための製品仕様書」，『埋蔵文化財ニュース』)

[15] Fujimoto, Y.: *Information standards for cultural heritage with the ISO 191XX series*, Proceedings for the 22nd CIPA Symposium held at Kyoto TERRSA, October 2009.

[16] ITTC: *Real Time 3D Measurement by Single Camera*, http://www.ittc.co.jp/English/sproduct/photocalc/SingleView3D/SingleView3D-2.htm, International Technology Transfer Corporation.

[17] ImageJ: *Image processing and analysis in JAVA*, http://rsbweb.nih.gov/ij/, ImageJ.

[18] Avidemux: *Avidemux–main page*, http://fixounet.free.fr/avidemux/, Avidemux.

[19] Thévenaz, P. Ruttimann, U.E. Unser, M.: *A Pyramid Approach to Subpixel Registration Based on Intensity*, IEEE Transactions on Image Processing, vol. 7, no. 1, pp. 27-41, January 1998.

[20] Wendykier, P.: *Parallel Super-Resolution Project*, http://sites.google.com/site/piotrwendykier/software/par allelsuperresolution, 2009.

[21] Downey, M.: *mikedowney.co.uk*, http://www.mikedowney.co.uk/blog/2001/10/02, 2010.