Research and development data management system by using geospatial information for Archaeological studies

Yoshinori IWAMOTO*, Ryosuke SHIBASAKI**
Center for Spatial Information Science, The University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
yiwamoto@iis.u-tokyo.ac.jp*
shiba@skl.iis.u-tokyo.ac.jp**

Yu FUJIMOTO
Nara University
yu-fuji24@e-broad.jp

Hiroya TANAKA
Keio University
htanaka@sfc.keio.ac.jp

Abstract: In the field survey of archaeology, huge amount of data are gathered through various methods such as taking photos, handwriting drawings and acquiring digital data like 3-D laser, and these data should be related to each other. However, it is difficult to relate these data together in present condition, because there are some differences in a format, a scale, and a precision. Therefore, a data management tool for archaeological field survey is being keenly demanded.

In this study, we aim to develop the method for data management system, named "Geospatial data manager". This system is a comprehensive data management scheme based on photos to help archaeologist easily and quickly develop data product on the Web. And through this study we propose standard metadata type at investigation in archeology.

Our system enables us to manage various data by attaching absolute location on maps and by relating with each other based on the qualitative proximity.

With the absolute locating component, individual objects are given geometrical positions in a global space. And, the relative data connection component helps provide a relationship between data with hyperlinks.

And also we publish and manage this system on the Web. Moreover, our application help elicit hidden context by raking up information on the relationships among data through the Web. Experiment in it based on the data of investigation of the world heritage of Tyre, Lebanon.

Keywords: WebGIS, Geospatial, Excavation data, Archaeological studies

1. Introduction

Field survey and Excavation in archaeology is conducted to acquire and to collect information on archaeological remains and relics in a systematic way using limited time and human resources. Data to be collected are so diversified ranging from overall structure of archaeological remains and relations of strata, details of individual parts of archaeological remains and information on each relic such as its classification, location and strata of unearthed position to its three dimensional shape and photos. The voluminous and diversified information should be efficiently collected, acquired and organized in such a manner that the relationships among them can be easily retrieved.

In recent years, digital camera, laser scanners, spatial database management system such as GIS (Geographic Information Systems) and 3-D drawing modeling tools such as CAD have made very rapid progress. The advances enable to acquire digital data so easily on archaeological remains and relics. At the same time it also provides a possibility of developing new types of products such as 3-D models. In addition, using Internet, digital data can be easily shared among archaeologists. Through sharing digital archaeological data among larger numbers of researchers, comparative studies and analysis from more diversified viewpoints can be promoted, which will eventually result in greater contribution to the advances in archaeology.

However, in order to improve the efficiency of information gathering, and informational sharing, we should consider how to combine a space representation peculiar to archaeology and the design of data management scheme for huge amount of data.

In this research, the development example of an information system on the WEB for data collection and organization is presented, through a case study the excavation work in Tyre remains and Lebanon.

2. Procedure of excavation survey

Excavation survey is usually divided into 4 stages. 1st stage is pre survey, 2nd stage is main excavation survey, 3rd stage is the arrangement preservation processing and 4th stage is writing the report. Figure 1 shows these procedures and main works and records. (Table.1)

At 1st stage, pre survey has 4 works (liaison and coordination, literature searching, distribution survey, trial excavation). Liaison and coordination means both the archeologist and parties of the administration of the archeology jurisdiction are the prior information exchanges, confer, and adjust the handling of ruins and the relic. Literature searching means searching documents about survey area. Within the range of the searching, not only the history but also a geographic document and the environment in the surrounding reach. And distribution survey is to confirm distribution situation of ruins. Afterwards, in trial excavation, ruins and relics situation are confirmed in detail.

Next is 2nd stage, main excavation survey. Main survey is work for preservation of the record of ruins related to the business enforcement ground. The excavation investigation targets the range to be influenced by the business that turns out by Pre survey. The measurement is geographical features within the range where the location of ruins can be clarified. It makes it to object. Recently, digitalization is advanced though the measurement method. Ground leveling removes the obstacle such as trees and peels off topsoil. At this time, the labor saving and shortening the period might be achieved with a machine. In setting grid, all area is set like grid (several meters) and it aims an accurate excavation. At last excavation survey is occurred. In excavation survey, the excavation record of detailed is acquired

Next is the arrangement preservation processing. This stage is most important and it costs a lot of time. In this stage, re cords of works are acquired at each work of about one line. Of course, this stage is done for the data that doesn't do for al l data acquired in excavation survey but is judged to be important from among that.

At last, writing the report. The data of ruins and relics is summarized in the report based on records of the photograph, the drawing, and the sketch, etc which is acquired in excavation survey. In the excavation survey, the report, it becomes t he only material that learns a historical meaning of the ruins and relics. It has the publication of the report and the excava tion survey is completed.

In this way, there are a lot of stages in the excavation survey, and the huge amount of data obtained in excavation surv ey.

	1st stage	*2nd stage*	*3rd stage*	*4th stage*
	pre survey	main excavation survey	the arrangement preservation processing	writing the report
Main works	liaison and coordination literature searching distribution survey trial excavation	measurement ground leveling setting grid excavation survey	cleaning classification analysis storage	data select discussion writing publishing
main record	document distribution maps daily report of survey	particular maps grid maps daily report of survey feature catalogue relics label	daily report of survey feature catalogue result of analysis relics label storage label	report

Table.1 Procedure of excavation survey

3. Data type of excavation survey

In this chapter explain data type of excavation survey. First is data format, which is acquired in excavation survey. Second is element of data type. Knowing what by format and what is described is necessary to develop the system.

1) Data format

In recent years, digital camera, laser scanners, and 3-D drawing modeling tools such as CAD have made very rapid progress. The advances enable to acquire digital data so easily on archaeological remains and relics. This content explain main data format in excavation survey (Fig.1).

a) Drawing

Drawings mainly covers comparatively macroscopic information like the whole organization of a site and more detailed 3-D form of individual ruins and relics. Drawings are usually supposed to represent geometry of objects but in reality, drawings are the results of judgments on what is important and what should be recorded. Therefore some parts that do not exist are sometimes drawn complementally by interpretation of an author. Drawings are important as a unique form of expression, which tells situations of ruins and relics objectively and at the same time reflects the interpretation of an author.

b) Hand-written Sketch

Hand-written sketch can express clearly what attracts the attention of an author among ruins and remains. It is used in many cases to express the interpretation of the author, which cannot be expressed with a drawing or a photograph. And it can also express textures and color of ruins and relics.

c) Photograph

Photos can record objectively not only the shape of objects but also the color, textures, and so on. They are used as a complementary recording media of drawings. And they are also used to record excavation works.

d) Document

It describes a classification, an excavation stratum, a presumed age, and so on. And excavation works is also recorded in documents. For example, it records under what situations, what remains were discovered when.on your word processor, please use the font closest in appearance to Times. Type. Sizes for your manuscript are specified in the following:

e) 3D data

In recent years, digital camera, laser scanners, and 3-D drawing modeling tools such as CAD have made very rapid

Hand-written Sketch Photos 3D map

Figure.1 An Example of Hand-written Sketch, Photos and 3D map

progress. The advances enable to acquire digital data so easily on archaeological remains and relics.

2) Element of data

In this content explains what information described by the diary report, the relic label, the report and so on. Information described that is necessary information for the excavation survey, and necessary for the system that develops by this rese arch.

a) Relic label (Fig.2)

This is a survey label put up to one-relic, and the record when excavating it is being written. When taking a picture of the relics, this label takes onto the same photo. Information described in the label is the following.

<project stage>, <area name>, <relics No.>, <strata name>, <create date>

b) Label for storage box (Fig.3)

This is a label put up to the box for storage the relics. This label explain what work was done to the relic on the inside is understood. Information described in the label is the following.

<project stage>, <area name>, <relics No.>, <ruin No.>, <strata name>, <create date>, <process>

c) Dairy report

The diary report is a report that the worker who does the excavation survey writes every day. Information written by the content of the excavation survey besides worker's subjectivity judgments is described here. Mainly information described in the dairy report is the following.

<project stage>, <investigator name>, <comment>, <work descriptions.>, <strata name>, <create date>, <process>, <weather>, <scale>, <material>

d) Project report

The data of ruins and relics is summarized in the report based on records of the photograph, the drawing, and the sketch, etc which is acquired in excavation survey. In the excavation survey, the report, it becomes the only material that learns a historical meaning of the ruins and relics. Project report has huge amount of information. Mainly information described in the dairy report is the following.

<project name>, <project ID>, <project period>, <project stage>, <principal investigator>, <associate participant>, <reference>, <relics name>, <investigator name>,<result of analysis>, <inference age>, <figure No.> , <table No.>, <grid No.> , <area No.> , <grid name>, <area name>, <local name>, <lat&lng>, <comment>, <work descriptions.>, <strata name>, <create date>, <process>, <scale>, <material> and so on.

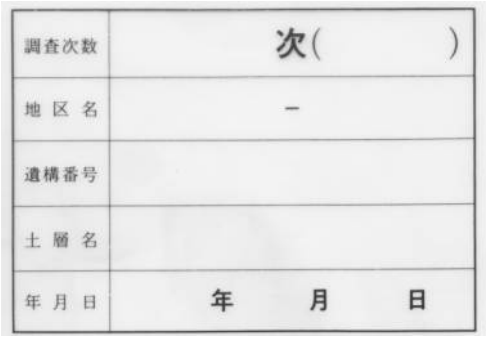

Figure.2 Relic label

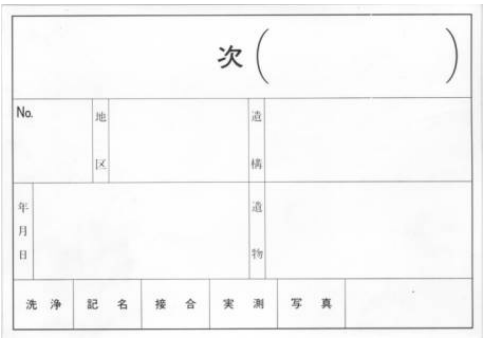

Figure.3 Label for storage box

3) Classification of data

In archaeology, acquired data can be characterized by four attributes (Fig.4). In a report, data are associated and organized based on these four attributes. The following four aspects are included for one data in archeology. It is Location, strata, time, and semantics. These four classifications are mutually connected. The archeologist arranges data to four types, and discovers a new finding there.

a) Location (Fig.5)

Data can be classified based on locations in ruins and relics. And ruins and relics are sometimes considered as a cube with six faces in archaeology. Therefore, one location may have six pieces of data.

b) Strata (Fig.6)

Data can be classified based on strata from which they are found. Excavated stratum, their location and the details are recorded by Stratigraphic relationships.

c) Time

Data is characterized by not only the age of ruins and relics but also excavation date.

d) Semantics

Data can be associated with each other along with archaeological inferences or stories inspired by the comparative analysis of ruins and relics. Classification of semantics has many cases.

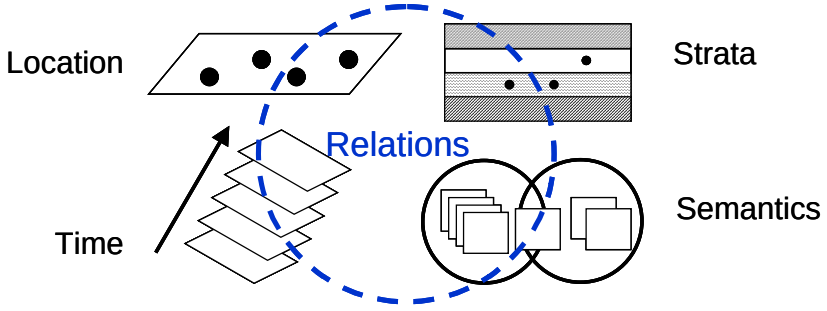

Figure.4 Classification of data

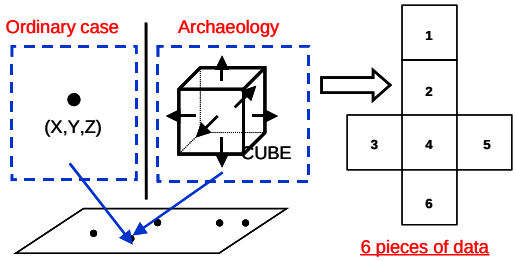

Figure.5 Location

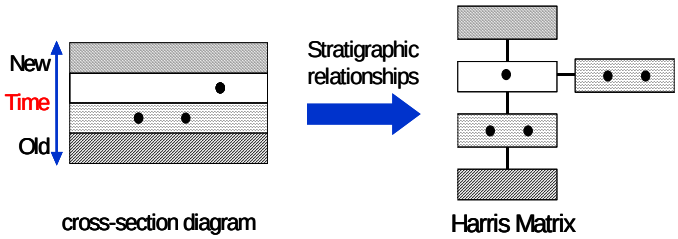

Figure.6 Stratigraphic relations

4. Data modeling

The modeling of the data treated with this system is done based on the result of the investigation of Chapter 2 and Chapter 3. As reflecting the result, this system has high usability for archaeologists. And purpose of data modeling is systematically managing data for digitalization.

1) UML class diagram

This research defines data class for excavation (Fig.7). The class is classified into 4. Super class of this data modeling is “project” class. “project” class consists of 2 sub class, “Excavation object” class and “excavation data” class. “location data” class relates to all classes. Each class relates itself.
Features of this data modeling are such. First, “location data” class relates to all classes. Because all class have location data. And to separate location information means it data can be easily handled. The feature of the second point is to divided “excavation object” class and “excavation data” class. Generally, when the excavation survey ends, objects are returned to former place or condition. So, excavation object and excavation data are the same classes, data is lost, too. Of course excavation object and excavation data are related as object ID and data card ID in these class.

a) Project class
This class is super class. This class describes the outline concerning excavation survey. For example project name, project period, principal investigator and so on. The data that might not change by the excavation work and the excavation objects is described. The documents that researched with pre survey are described here. That is, information related to the entire excavation survey is described. Recursive relation
b) Excavation object
This class describes excavation objects. Generally, the excavation object is buried and returned in excavation survey when survey is finished. So this class means conceptual. This class describes object ID, data card ID and location data. Data card ID puts up excavation data and correspondence.
c) Excavation data
Excavation class is conceptual, many information about excavation objects is describes in this class. Of course, even if it is empty excluding expect data card ID and object ID, this class doesn't have the problem. This class can overwrite and add data in these attributes. Location data” class relates to all classes. Follow character of Excavation survey, this class describes grid No, grid name, area No, area name, local name, object name, latitude, longitude. A feature thing is to be described object name in this class.
d) Location data
Location data" class relates to all classes. Follow character of Excavation survey, this class describes grid No, grid name, area No, area name, local name, object name, latitude, longitude. A feature thing is to be described object name in this class. At the excavation survey, ruin that found is likely to be recognized as a common word that expressed location.

2) XML for the excavation survey

```
<?xml version="1.0" encoding="utf-8"?>
<excavationXML>
  <project>
    <project name> ・・・</project name>
  </project>                                   Project
  <excavationobject>
  </excavationobject>                          Excavation object
  <excavationdata>
  </excavationdata>                            Excavation data
  <locationdata>
  </locationdata>                              Location data
</excavationXML>
```

Figure.8 XML for the excavation survey

Based on UML class diagram, this research proposes XML for excavation survey (Fig.8). To define "XML for the excavation survey" is easily possible misappropriating acquired data to other systems.

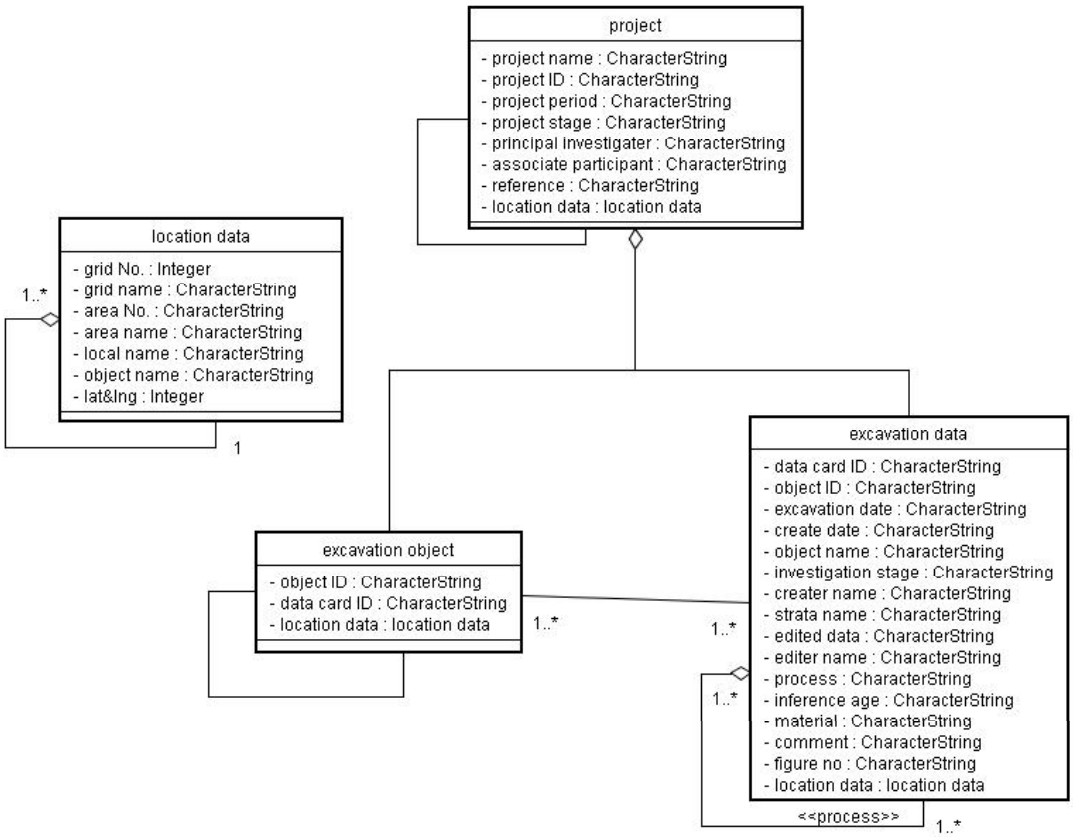

Figure.7 UML class diagram

5. Excavation Data management system "Geospatial Data Manager"

In order to manage excavated data, it is necessary to add some useful searching keys and to express relations between the data. Relics are clearly connected to places where these are found. The primary key is location.
Recorded data of ruins and relics such as pictures, sketches, and documents are given searching keys of location, strata, time, and semantics. And they are related with location by Hyperlink to manage these relations.
In general, the problem as digitalization is a complexity of input data term. This system manages data by the image and the text format. For example, the handwritten memo is scanned and it processes it as an image.
And, these excavated data should be shared by many researchers. In this point, this system works on the WEB. As a result, users are able to show and renew these data, or to add relational information or to search necessary information by the searching keys, or to rearrange these data on real-time basis.

1) Outline of this system

Through the excavation work in Tyre remains and Lebanon, our case study, we collected three dimensional laser data, pictures and so on.
In this system, the map is expressed with Scalable Vector Graphics (SVG), which is the language based on Extensible Markup Language (XML) that World Wide Web Conso

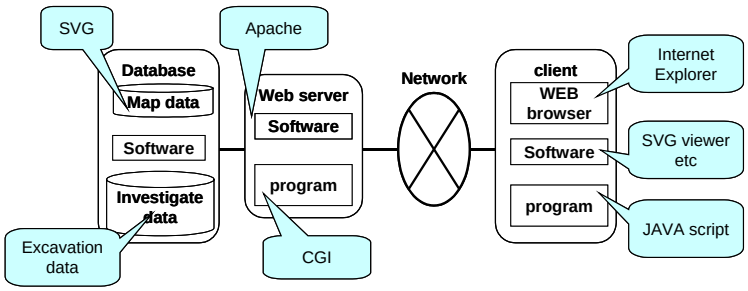

Figure.9 Outline of this system

rtium (W3C) suggests for web graphics. The peculiarities of SVG are twofold; (1) high capability of map expression that vector graphics can be handled, map scale can be easily enlarged or reduced, and that the structure of layers can be applie d, and (2) high capability of data retrieval that words in a map can also be searched. In addition, by using the WEB an arc haeologist can write in information at any time and the archaeological data can be made open to many people.

2) System flow

This figure shows the system flow (Fig.10). At first, researcher uploads image data to server. In this case, researcher uploads digital photos as the excavation data. Secondly server program (CGI) analyze EXIF information iand write these data in the data file.

At same time, another CGI program analyzes image data. Usually, relics photos includes relics label. This program cut out relics label from digital photo and make new label image file. And this program analyzes handwritten information in relics label by using OCR.

Finished these processes, server make input form and display input screen on the browser. At this time Information on EXIF and the label analyzed in the previous process has already been described in the input form. So researcher can easily to input excavation data. After input data, researcher sends this form to server.

Next process, CGI program for address matching analyzes and adds coordinates in map. This figure shows this process (Fig.11). The feature of this program is as follows. The way adds coordinates data are 2ways. First is by using latitude and longitude. Latitude and longitude is got by digital camera with GPS. Another way is by using address data table. Address data table has location name and these coordinates. If uploaded data doesn't have latitude and longitude, researcher selects location name in this data table.

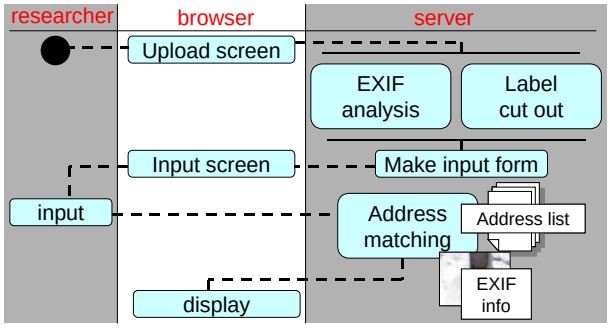

Figure.10 The system flow

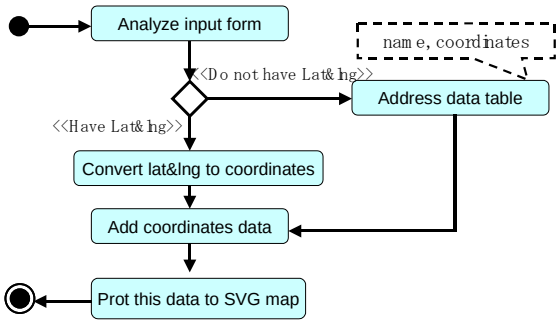

Figure.11 Address matching process

3) Screenshot of this system

This figure shows interface at the system (F ig.12). This inter face consists of mainframe, data frame and thumbnail index. At mainfra me express 4views of classification of data (location, strata, time, semantics). This fram e can change view by using select box. User can access excavation data from all frames. For example from point on map, point on str ata, picture in time screen, semantics screen, and thumbnail index.

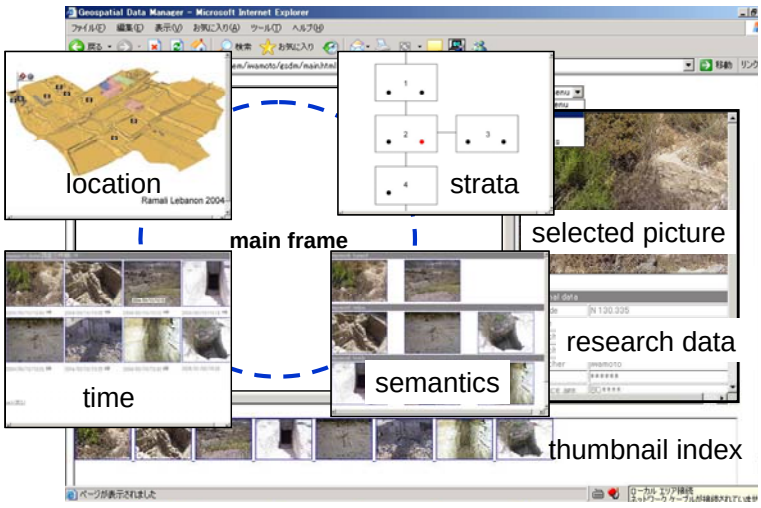

Figure.12 Interface of Geospatial data manager

6. Conclusions and future works

In this research, based on archaeological c lassification, we developed a system, which constructs various enormous information fro m excavation survey systematically.

As for the topics for future, we'll ask archaeologists use practically to receive comments on the system for the improve ment. Also, we believe that archaeological data associated with hyper links to location, strata, time and semantics will be come a basis in discussing a new electronically structure of archaeological documents in the future.

References

[1] Edward C Harris, Ozawa Kazumasa, 1995. "Principles of Archaeological Stratigraphy", Yuzankaku-shuppan

[2] FUJIMOTO Yu, IZUMI Takura, 2004, "Examination and Application of 3D Archaeological Site Measuring System -Investigatio n at Ramari site of Tyre in Republic in Lebanon", Geoinformation Forum JAPAN 2004 Student Forum 2004