From Vectors to Objects: Modeling the Nasca Lines at Palpa, Peru[1]

SEE THE CD FOR THE EXTENDED VERSION

MARTIN SAUERBIER
INSTITUTE OF GEODESY AND PHOTOGRAMMETRY,
ETH HOENGGERBERG, CH-8093 ZURICH, SWITZERLAND,
WWW.PHOTOGRAMMETRY.ETHZ.CH
MARTIN.SAUERBIER@GEOD.BAUG.ETHZ.CH

KARSTEN LAMBERS
INSTITUTE OF GEODESY AND PHOTOGRAMMETRY,
ETH HOENGGERBERG, CH-8093 ZURICH, SWITZERLAND,
WWW.PHOTOGRAMMETRY.ETHZ.CH
KARSTEN.LAMBERS@GEOD.BAUG.ETHZ.CH

DEPARTMENT OF PREHISTORY,
UNIVERSITY OF ZURICH, KARL-SCHMID-STR. 4,
CH-8006 ZURICH, SWITZERLAND, WWW.PREHIST.UNIZH.CH

ABSTRACT

The Nasca Lines around Palpa are currently investigated by archaeologists and geodesists from Switzerland, Germany, and Peru. A photogrammetric analysis of high resolution aerial images allowed the accurate 3D recording of more than 1500 geoglyphs, most of which have also been thoroughly described in the field. A digital 3D model of the area of investigation is the first important result of this effort. However, the vectors recorded during the mapping process that mark the preserved outlines of the geoglyphs do not coincide with the geoglyphs themselves, which contain also an interior surface. Furthermore, they show certain traits important for an archaeological analysis, some of which (geometric attributes like location, size, and orientation) can be derived from the processed vector data, while others (contextual attributes like stratigraphic relations and associated objects) have been recorded during fieldwork. An important prerequisite for the intended GIS-based analysis of our data is to define meaningful archaeological objects (i. e., the geoglyphs) and assign them different kinds of attributes like those mentioned above. In this paper we describe the original data, the process of object definition, and our data model and discuss some preliminary results.

Introduction

If archaeological objects, be they single finds, features such as graves or buildings, or whole sites, are to be analyzed using a GIS, the underlying database in which information concerning the objects is stored has to contain both spatial data (defining the location of the object) and attribute data (describing its characteristics) (cp. Bill 1999:11, Wheatley and Gillings 2002:23). The main purpose of the GIS for the Nasca Lines at Palpa, Peru, which is currently being developed and implemented at the Institute of Geodesy and Photogrammetry at ETH Zurich (IGP) is to enable the analysis of the relationship between the geoglyphs and their surroundings in order to better understand the criteria that guided the creation of the ground drawings (Grün, Sauerbier and Lambers 2003, Sauerbier and Lambers 2003). The recording of spatial as well as attribute data on the more than 1500 geoglyphs in the Palpa area required a two fold approach resulting in two separate data sets:

- Spatial data: On a base provided by a series of large scale aerial images suitable for stereoscopic analysis, all geoglyphs were identified and photogrammetrically mapped (Grün and Lambers 2003). The resulting geometrical data was stored as 3D vector data (lines).

- Attribute data: All geoglyphs were then located, numbered and described in the field using a standardized feature sheet (Reindel, Lambers and Grün 2003). The archaeological data was stored in a relational database.

These two data sets cannot be linked easily due to a peculiarity of the spatial data set resulting from the state of preservation of the Palpa geoglyphs. Unlike objects in a cadastral GIS, where geometrical information is derived from measurements of clearly defined objects (e.g. buildings, real estate; see Bill 1999), the Palpa geoglyphs usually cannot be measured in their entirety (as polygons) since many of them are partly damaged by erosion or modern land use. Another difference to cadastral GIS, where measured objects of the same class usually do not overlap, is that many geoglyphs had already been cut or covered by other geoglyphs in prehispanic times. During the image-based photogrammetric recording of the geoglyphs, only the actually preserved or securely deducible parts of geoglyph borders were mapped in order to maintain this important information. This has resulted in a set of digital 3D vectors that do not represent entire geoglyphs, but rather discontinuous sections of their borders. In other words, the spatial data and the attribute data do not then refer to the same category of archaeological entities. In order to enable a meaningful archaeological analysis, the 3D vectors have to be combined and complemented into polygons representing the most likely original shape of each geoglyph based on additional archaeological information. Such reconstructed objects can then be assigned the geoglyph numbers and attribute data recorded in the field. The aim is to store the data in an object-relational database management system (DBMS), which brings about the benefits of an object-oriented approach (methods, inheritance etc.). For analysis, the combined data sets have to be made accessible via an intuitive graphical user interface that should allow the retrieval of all available information about a given geoglyph by simply clicking on its graphical representation.

In the following sections we describe a preliminary workflow for object definition that represents an intermediate step towards a system that stores both attribute and spatial data in a single database and allows read/write-access from a graphical user interface.

Technical Requirements and Software Solutions

The archaeological research problems to be addressed, which have been described in Grün, Sauerbier and Lambers 2003 and Sauerbier and Lambers 2003, bring about specific requirements concerning the technical design of the GIS. An important specification is that the basic data should be accessible in a standard data format in order to allow the different users (Grün, Sauerbier and Lambers 2003:Fig.5) to access the database using GIS software of their own choice. Furthermore, a standard data format is important for the publication of the data (on CD or web based, e.g. in XML format). Therefore, it was decided that Oracle 9i DBMS would be used as a central database of the GIS. Most GIS software packages have an interface to Oracle DBMS, and the data is also accessible via the SQLPLUS interface, through the use of embedded SQL, or from programs written in C++ or Java.

Three different software packages were evaluated for the GIS being developed for use as a tool for data manipulation, query and analysis, and as graphical user interface that provides access to the data. These were TopoBase 2.03 by C-Plan, GeoMedia 4.0 by Intergraph and ArcGIS 8.1 by ESRI. The most important criteria for evaluation were user friendliness, the efficiency of tools for data manipulation, the querying and editing capability for raster and vector data, and cost. Furthermore, the GIS software should be able to be customized, i.e. it should allow enhancement of its functionality through use of programming language. For fieldwork in Palpa, as well as for publication and presentation of the project, a further important requirement of the GIS is that it provides an efficient tool for the production of 2D paper maps at different scales.

As a result of the evaluation it was decided that ArcGIS 8.1 (and later on release 8.3, which was not yet available for the work described here and was therefore replaced by ArcView 3.2 for some steps) would be adopted as a basis for the GIS, though this did not preclude the use of other applications for specific tasks.

Development of a Workflow for Object Definition

As mentioned earlier, the first requirement for a GIS-based analysis of the Palpa geoglyphs is to combine and integrate the two distinct data sets (spatial and attribute data) in order to obtain meaningful archaeological objects. The workflow for object definition is then as shown in Figure 1.

[1] Spatial data

The preserved borders of the geoglyphs which are visible in the stereopairs of the aerial images were mapped on an analytical plotter Wild S9 using XMAP 4.00 by Aviosoft as illustrated in Figure 2 (Grün and Lambers 2003).

The resulting vectors were stored in proprietary XMAP files (XMP). In order to process the data, these files are converted to DXF format using the XMPDXF converter by Aviosoft. Then, in ArcGIS 8.1 the DXF-file is converted to a coverage using the DXFARC command. It should be noted that although it is also possible to obtain objects of area type directly in XMAP, the conversion of polygonal topologies using XMPDXF causes a loss of information.

The next step is the completion of the line vectors to obtain closed polygons. For this purpose, the high resolution orthoimage (25 cm footprint) is shown as a background layer in the ArcInfo user interface, with the vector layer in the foreground as indicated in Figure 3.

Based on what is visible in the orthoimage and what is known about the geoglyph stratigraphy from fieldwork, the gaps in the geoglyph borders are then closed in order to represent the most likely original form of the geoglyphs. This requires the manual edition and insertion of arcs and nodes in order to create closed polygons. Due to the somewhat limited functionality of ArcGIS 8.1 concerning object topology, vectors representing borders of two adjacent geoglyphs have to be marked twice, resulting in redundant data. This problem is expected to be solved in release ArcGIS 8.3. The editing process is accomplished in a two-dimensional coordinate system. The third dimension has to be added later by intersection of the objects with the DTM.

After the completion of the vectors, a polygon topology is created automatically using the "Clean" command in ArcGIS 8.1. This topology is stored in an arc.adf file that contains in a binary format the number of arcs and the arcs comprising a polygon. The resulting polygons shown in Figure 4 should not be confused with the geoglyphs. On the one hand, all geometrically possible polygons are created, including such areas enclosed by geoglyphs that are not themselves a geoglyph. On the other hand, due to the many overlaps between the geoglyphs, each polygon usually represents only a part of a given geoglyph, while it may belong at the same time to several overlapping geoglyphs.

Thus, the next step in the workflow is the revision of the polygon layer. For this step, ArcView 3.2 is presently used because of its comfortable editing functionality. It is planned in the future to accomplish this task in ArcGIS 8.3. Initially, all dispensable polygons that do not belong to any geoglyph are deleted. This is done manually by graphically selecting them, resulting in a shapefile with single polygons, as shown in Figure 5.

In a second stage, all polygons belonging to a given geoglyph are assigned an object ID that corresponds to their geoglyph number in the archaeological database. The resulting data is stored in a table. Where geoglyphs overlap, the corresponding polygons belong to two different archaeological objects. In such a case, the respective polygons have to be copied and stored as several objects in order to allow the assignation of several object IDs to a polygon. The result of this step is an assemblage of polygons, each of which carries an object ID that clearly indicates to which geoglyph it belongs.

In a final step, all geoglyphs with the same object ID are merged using ArcView's geoprocessing extension ("Dissolve" command) into a single geometric object using an aggregation by the attribute "ID". As a result, ESRI shapefiles and coverages, and their corresponding tables containing the objects are obtained. These now represent the real world geoglyphs, with their IDs as attributes, as indicated in Figure 6.

[2] Attribute data

The attribute data registered in the field was stored in a MS Access 2000 database. This data, exemplified in Figure 7, consists of partly standardized descriptions concerning, among other things, the form, construction technique, orientation, stratigraphy, and context of the geoglyphs.

The attributes are of different data types (text, numerical, Boolean) and some are constrained to a limited set of predefined values (e.g. the construction technique, form). The unique primary key of the table which identifies each record is the geoglyph number. The goal is to migrate the relational MS Access data to an object relational Oracle database. A prerequisite here is the storage of attributes according to the object-oriented data model (Lambers and Sauerbier, forthcoming), so a fully automated migration process is not possible. Another constraint is the non-compatibility between some of the MS Access data types and corresponding Oracle data types. For example, the Access data type MEMO (character field of unlimited size), used for textual descriptions, has no direct counterpart in Oracle. Long character fields in Oracle can be either up to 4 GB (CLOB, BLOB), with these being not compatible with ArcGIS 8.1, or up to 4000 Bytes (VARCHAR2), which restricts the length of textual descriptions. In this case it was decided to use VARCHAR2 since all geoglyph descriptions are shorter than 4000 Bytes.

The workflow for data migration starts with a conversion of the Access database (MDB) to an XML file which contains the table structure. For this step, the Oracle Migration Workbench, a free extension to Oracle 9i, is used (Fig.8).

In a second step, the table structure is converted from the XML file to an SQL-DDL script which contains the SQL statements for creating the table structure, constraints, primary and foreign keys, and references in an Oracle database. This step requires the setting of the corresponding data types, in parts manually, in a Migration Workbench menu. Afterwards, it is possible to create SQL statements for the data itself which can then be added to the Oracle database via the Oracle Migration Workbench or the SQLPLUS-Worksheet. The result is a database with a table structure based on the system independent conceptual data model, which has to be adapted and implemented iteratively.

[3] Data integration and user interface

With the spatial data stored in the ESRI file system and the attribute data stored in the Oracle database, a link between the two data sets has to be established that enables a graphical selection of objects and a display of the appropriate descriptive data. For this purpose, ArcView 3.2 offers the possibility of connecting to different DBMS, e.g. Microsoft SQL Server, IBM Informix/DB2 and Oracle via ArcSDE or ODBC (Open Data Base Connectivity). In this case, the ODBC interface has been employed to access the Oracle database tables for read only access in ArcView 3.2. The constraint for a correct natural join of two relations is a common attribute in both tables (Heuer and Saake 2000). Then, the tuples with equal common attribute values are attached, as per Figure 9. This common attribute is the object ID that has been assigned to each geometrical object and which corresponds to the geoglyph number, the primary key in the central table of the archaeological database. Having activated the database connection by using the "Add database table" menu in ArcView, SQL-queries on database tables can be performed. However, manipulation of database contents is not possible.

Conclusions and Future Tasks

At the end of the workflow described, we obtain as an intermediate step a read only, graphically accessible set of archaeological objects, in our case the geoglyphs, with all available data shown in a table under ArcView 3.2.

To complete the desired functionality, some further steps will have to be undertaken. Firstly, we want to achieve a read/write access from the GIS software to the database. For this reason, the most recent release of ArcGIS 8.3 will serve as an user interface in the future. It does not only offer the possibility of read/write-access to the database, but also allows the definition of topologies with certain constraints and rules for efficient error checking. Another aim is to store the geometrical data not in a file system but also in the database. For this purpose, Oracle provides the spatial cartridge (Sharma 2002) which allows the storage of geometrical objects independent of any GIS package's data structure. A further issue is the step from the system-independent, conceptual data model to the logical data model, along with its physical implementation in the database as well as the subsequent extension for the integration of additional data resulting from the analysis.

Acknowledgements

We are grateful for the generous financial support given by both the Swiss-Liechtenstein Foundation for Archaeological Research Abroad (SLSA, Zurich), for fieldwork, and ETH Zurich for analysis. We also would like to thank Prof. Clive Fraser for his contribution to the linguistic improvement of this paper.

1 The figures mentioned in the text are available on the accompanying CD only.

REFERENCES

BILL, R., 1999. Grundlagen der Geo-Informationssysteme, Vol. 1. Heidelberg.

GRÜN, A. and LAMBERS, K., 2003. The geoglyphs of Nasca: 3-D recording and analysis with modern digital technologies. In Acts of the XIVth UISPP conference, University of Liège, Belgium, 2-8 September 2001, section 1: theory and methods - general sessions and posters, Oxford (BAR Int. Ser. 1145):95-103.

GRÜN, A., SAUERBIER, M. and LAMBERS, K., 2003. Visualisation and GIS-based analysis of the Nasca geoglyphs. In Doerr, M. and Sarris, A. (eds.), The digital heritage of archaeology - proceedings of the 30th CAA conference, Heraklion, Crete, April 2002, Athens:161-167.

HEUER, A. and SAAKE, G., 2000. Datenbanken: Konzepte und Sprachen. 2nd revised ed., Bonn.

LAMBERS, K. and SAUERBIER, M., 2003, in press. A data model for a GIS-based analysis of the Nasca Lines at Palpa (Peru). International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XIXth CIPA International Symposium, Antalya, Turkey.

REINDEL, M., LAMBERS, K. and GRÜN, A., 2003, in press. Photogrammetrische Dokumentation und archäologische Analyse der vorspanischen Geoglyphen von Palpa, Südperu / Documentación fotogramétrica y análisis arqueológico de los geoglifos prehispánicos de Palpa, costa sur del Perú. Beiträge zur Allgemeinen und Vergleichenden Archäologie 23.

SAUERBIER, M. and LAMBERS, K., 2003. A 3D model of the Nasca lines at Palpa, Peru. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, International Workshop on Visualization and Animation of Reality-based 3D Models, Tarasp, Switzerland, Vol. XXXIV-5/W 10, CD-ROM.

SHARMA, J., 2002. Oracle Spatial. Technical White Paper. Oracle Corporation, USA (http://technet.oracle.com/products/spatial/pdf/9iR2_spatial_twp.pdf).

WHEATLEY, D. and GILLINGS, M., 2002. Spatial technology and archaeology: the archaeological applications of GIS. London.