Towards Building a Knowledge Base for Research on Andean Weaving

Denise Y. Arnold, Sven Helmer, and Rodolfo Velásquez Arando

***Abstract*—We are working on a knowledge base to store 3D Andean textile patterns together with rich cultural and historic context information. This will allow ontological studies in museum collections as well as on ethnographic and archaeological fieldwork. We build on an existing ontology, extending it to incorporate more content and make it more accessible. This goes well beyond storing and retrieving textile patterns and enables us to capture the semantics and wider context of these patterns.**

I. Introduction

Andean civilizations have used weaving for conveying information for a very long time. In contrast to alphabetic writing as developed in Europe and the Middle East, information is stored in patterns on materials such as cloth. Research has established that Andean people have employed this medium to document and communicate complex information. While some of these textile messages are understood by today's researchers, others are still waiting to be deciphered. One challenge in continuing this research is that the material for study is only available in a piecemeal fashion and without contextual information.

We plan to prepare information on Andean textiles in a way that puts it into a cultural, geographical, and historical context by providing a knowledge base that researchers can annotate and query. We draw on AI techniques for knowledge representation, reasoning, and sophisticated methods for searching. In addition to storing weaving patterns, our knowledge base will be fed with multimedia data such as digital photos, video, and text, providing context for the textile designs. This ranges from commentaries by living weavers to ethnographic and historical records. Also, previous approaches have been limited to surface features of textiles. We, on the other hand, plan to store 3D-data. This is important because it gives a deeper insight into weaving techniques (enabling us to identify historical and cultural associations) and particularly layered arrangements of the weft and warp also convey information (the surface patterns do not hold the whole information).

This project has a certain urgency to it, as the weaving practices, some of which have been handed down for millenia, are endangered. Younger generations no longer pick up weaving, meaning that this knowledge is being lost. We want to contribute to preserving the rich cultural heritage of Andean weaving. In particular, we have the following aims:

D.Y. Arnold is director of the Instituto de Lengua y Cultura Aymara (ILCA), Bolivia, and honorary professor at Birkbeck, University of London, UK; S. Helmer is with Birkbeck, University of London, UK; R. Velásquez is with the Universidad de Aquino de Bolivia (UDABOL), La Paz, Bolivia, and with ILCA.

- developing an ontological approach that permits a more logical systematization of the data and interpretations of former studies on Andean weaving
- surpassing the (software) limitations of the current 2D analysis techniques of Andean cloth through the development of 3D virtual reality fabric design and simulation tools

The remainder of this paper is organized as follows. The following section contains more background on the role of textiles in Andean culture. Section III briefly describes the approach we want to take in building the knowledge base. Section IV concludes the paper.

II. Background on Andean Weaving

In the Andes, the wider ramifications of textile design go beyond the immediate functional utility of cloth, having implications in cultural practice and learning methods, identity and territory, and importantly of cultural heritage. The international Coroma case, which sought the repatriation of 300 historical textiles taken from a Bolivian rural community without communal permissions, to be displayed in US Museums and put on sale in galleries, shows current limitations in textile documentation, classification methods, collecting methods, and legality, while revealing their value to local communities.

In cloth layout, the relative proportions of figurative and plain cloth reveal historical allegiances to the woven repertories of wider confederations, as well as indicating if textiles were woven in the highlands or valleys, and by older or younger generations. The colors of cloth indicate regional and historical identities, age and family groupings, and regional political identities. Cloth color also indicates the technological repertories of dying processes based on the natural resources available in plants, insects and minerals, and the time of year when these were collected. In the recent past, the detailed designs of flora, fauna, avifauna, and astral bodies, were directly inspired by the detailed observation of local resources, in such a way that textiles are often maps of regional territories, at particular times of the year.

In the language of local design patterns, textile features such as borders, figure orientation, style and color use, reveal changing social and cultural correlations. Design layouts also express the key symmetries used in particular cultures, in broader or minor patterns. The language of textile stripes frequently represents the flow of rivers, stone walls or boundary markers. In many weavings, band width (and higher warp count) indicates the quantity of the item represented, whereas

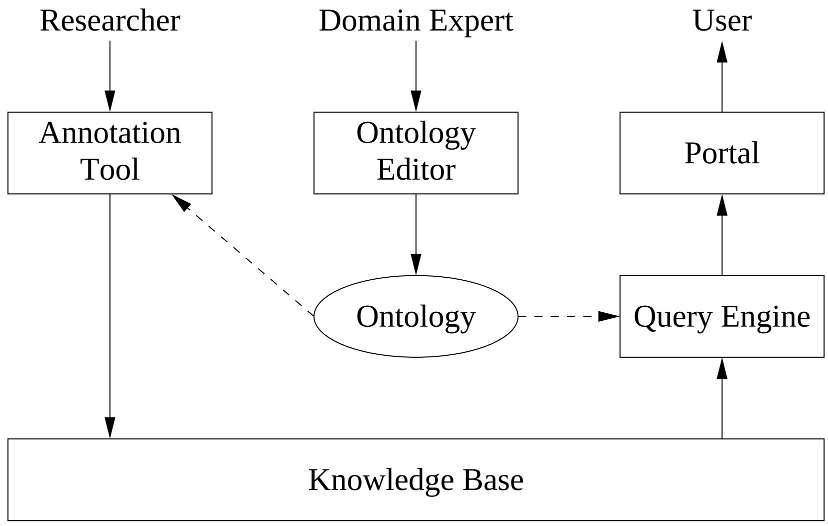

Fig. 1. General architecture of the system

band color indicates its content (water, maize, potato flowers). The relation of wider to narrower stripes also indicates generational hierarchy, commonly expressed as that between "mothers" and "offspring". This language of width and color, found in both woven bags and the knotted *kipu* threads, is used to document the patrimony of raw food crops produced locally, or else preserved as toasted grains. The same language is also used to document the wider exchange patterns between highlands and lowlands, for example of highland salt for valley maize, when the colored stripes are ordered by the ecological zones where each product is found. "Raw" (natural) and "cooked" (dyed) colors have distinct usages.

Many different theoretical tendencies have sought to make sense of, and translate into broader terms, these matters of textile design. These have drawn on linguistics, semiotics, ethnography, archaeology, cultural geography, history of art and science, information studies, and math, among others.

Linguistic approaches tend to perceive Andean weavings as semiotic media, like language and written texts, hence artifacts through which cultural processes are negotiated and entextualized in wider semiotic systems. For instance, Halliday's social semiotics, applied to visual analysis, helped establish a "grammar of visual design" [10] that we intend to apply to textiles. Other linguistic studies argue that textiles embody not just grammars, but more general-purpose meaning systems [12] that work across cultural and linguistic boundaries in larger political formations in common. A current limitation of many linguistics-based studies is that they scrutinize individual acts of weaving as examples of sign-making or *parole*, while ignoring the *langue* of cloth as a more viable social medium for analyzing consistencies in regional repertoires of woven forms and their variations.

These perceived consistencies have also been approached from anthropology and the history of science. Recent studies of the *kipu* knotted threads [12], [14] and cloth [1] view them as regional operative, information and documentary frameworks, that became standardized under Inca state control. For example, Urton's *kipu* studies identify both binary sequences (direction of spin, knotting) of individual artefacts, and segmentary hierarchies of information analysis in wider *kipu* repertories. Frank Salomón, drawing on the visual work of Edward Tufte, perceives *kipus* as information systems organized through cultural schema to express region-wide models that are homologues of decision-making and performance registries. However, both authors, by over-theorizing, ignore the regional ontologies we seek to explore, especially the regional metalanguage of primary fabric structures that appear to organize local cloth repertories [6].

III. Our Approach

In this section we present the general architecture of our system (see Figure 1 for an overview). Central to the design is the ontology, which provides the data model and terminology for the rest of the system. The content of the domain model will have to come from domain experts in the area of Andean weaving. Researchers will be able to load (annotated) data into the knowledge base, while users will be able to extract data via a portal, which hides the complexity of the underlying inference-based query engine.

A. Ontology

Attempts to standardize weaving terminology are currently mainly driven by (museum) collections, which tend to overlook the significances of weaving practices and cloth techniques and structures in their ethnographic contexts. Many collections are still cataloged in databases (such as Merlin) that consist of photographic records of samples, backed by registers of their probable precedence, date, size, material, collector and date of acquisition, with modest descriptions of any techniques used. In our ontology we want to describe the complex relationships between textiles and their ethnographic, geographical, historical, and archaeological contexts, enabling us to combine material from museum collections and from new fieldwork. While respecting infodiversity issues, our mediating systems must reinterpret the semantics of existing data structures in diverse collections, incorporating data regarding period, precedence, region and community, with weaving technique.

While there are projects which have developed ontologies from scratch, doing so is a very labor-intensive process. We will draw on existing developments in this area, using the reference model CIDOC-CRM as a basis [5]. Advantages of this model are that an RDF representation of the ontology already exists and that it was designed with the intention of being extensible. Our aim is to articulate an Andean textile heritage knowledge model with the recommendations in CIDOC-CRM, while modeling certain aspects in more detail. Textile patterns are much more than just images, We need to be able to describe weaving techniques and to express the productive basis (pastoral economy) as well as the multiple productive processes (rearing animals, shearing, classifying/cleaning wool, spinning, plying, washing, dying, looming up, weaving, and finishing cloth). In addition to this, we want to provide a platform for weavers to document and protect their repertory of designs. Rights management is only covered very rudimentarily in CIDOC-CRM, as are political, intercultural and regional contexts.

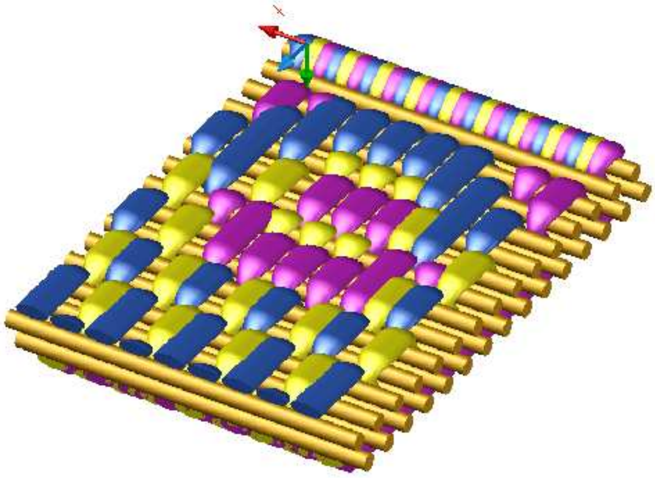

Fig. 2. Example of 3D interface

B. Interfaces

The system will provide three different interfaces: a portal for users to access the data, an annotation tool for researchers to add and update data, and last but not least an ontology editor for domain experts to create and maintain the ontology.

The portal consists of a graphical user interface (GUI) and more advanced (textual) options for experienced users. An important aspect of the GUI is not only ease-of-use, but the option to describe and search for 3D-textile patterns. This was motivated by Frame, who developed a structural method that relates early cloth's technical features (deep structures) to certain recurring surface designs [7]. The *Instituto de Lengua y Cultura Aymara*, ILCA, in Bolivia, recently developed a new tool, called *Sawu 3-D*, based on Computer Aided Geometric Design (CAGD) techniques, that can analyze textile structures in warp-faced cloth, and express them in 3D simulations with warp and weft counts, and relevant color coding (cf. [13]). For an example of the 3D-interface see Figure 2. We want to navigate seamlessly from these textile structures to specific samples, possible iconographic details, video illustrations of how these were woven, video comments on their meaning, bibliographic references, the historical context of archaeological objects, and the time mapping of related pieces.

It is not enough to just load the raw data into the database. An annotation tool will provide functionality to preprocess data and establish connections to the existing data set with the help of the ontology. The 3D-interface will also help with inserting new patterns (from fieldwork) or establishing connections between existing patterns (from previously overlooked collection pieces), by allowing the comparison of textile samples with samples from other times and places. This helps in uncovering historical and regional connections that were not known before.

Creating and maintaining an ontology is a challenging task (even if an existing approach is used as a basis). We opt for Protégé as our ontology editor for the following reasons: it is a widely-used, open-source editor, it is extendible (a wide range of plugins are available), and it supports many different standards such as CLIPS, CORBA, and JDBC.

C. Query Engine

In addition to providing traditional database querying capabilities, we need to tackle three different tasks: first, the query engine has to be able to handle multimedia data such as audio and video files, second, we also need support for 3D-pattern-matching, and third, we need to be able to search on a semantic level, meaning we have to integrate an inference engined into the system.

Many vendors of (object-)relational database systems have integrated multimedia querying capabilities into their systems in the form of so-called Extenders or Cartridges (for an overview see [9]). For our purposes the functionality and performance of commercial systems in handling multimedia data is sufficient.

Integrating the 3D-textile pattern matching into the system is probably the most challenging task, as it is an inherently hard problem and there are no extensions for database systems that cover exactly these requirements. An approach based on voxels (three-dimensional pixels) seems very promising at the moment, as there exist querying techniques that are invariant with respect to certain geometric transformations such as translation and rotation [2]. An alternative to this solution could be a neural-network-based pattern matching approach. Neural networks have been used quite successfully for textile pattern-matching before, albeit for fault detection in fabrics [11]. Also most developments in weaving software and database systems to date have been done with reference to conventional looms (Jacquard), with a simple technical picking motion, and not to Andean looms, with complex 3D interlayering of cloth.

For the inference engine we need a solution that integrates it into the system efficiently (i.e. in terms of performance and scalability). Currently the most promising approaches are either using an inference engine that is already integrated into a relational system, such as Oracle 11g RDF/OWL [3], or employing an engine that can be implemented efficiently on top of a relational system, such as Minerva [15].

D. Storage Manager

For the storage manager of our system we face similar challenges as for the query engine. We need to be able to store, retrieve, and index multimedia data, 3D-textile patterns, and the ontological information in an efficient way. The central component of our storage manager will be an (object-)relational database system, as it offers scalability and transactional processing.

Querying in multimedia-capable databases is supported by indexing frameworks that offer advanced index types for multimedia types [9]. The Multimedia Data Cartridge of Oracle offers services for managing index structures based on Generalized Search Trees (GiST) [8], which also makes it possible to extend the indexing capabilities of the database system as needed.

Data objects in the voxel-approach described above are represented by so-called *shape descriptors*. Certain (numerical) features are extracted from an object and stored in a *feature vector*. Basically, this maps each objects to a point in a metric space. Assuming that this mapping is done correctly, the distance between two points in the metric space then measures the similarity of the two original data objects. In [2] such a technique based on partitioning the data objects (and extracting features from each partition) is described. The job of the database system, then, is to store and retrieve the feature vectors efficiently.

Most modern database systems offer bulk-loading functionality for efficiently inserting the initial data set into the database. Benchmarks run on standard Linux desktop PCs have shown that a Lehigh University Benchmark (LUBM) data set consisting of 25 million triples can be loaded and indexed in roughly 40 minutes [4].

IV. Conclusion and Outlook

We are about to start a project to build a complete system as outlined in the previous section. Parts of the system, such as the Sawu 3-D tool for analyzing textile structures in three dimensions, have already been developed, while other parts are currently in the planning stage. The full-fledged project will also allow collaboration on a larger scale with museums in Peru, Chile, and Bolivia.

We are facing several challenges in implementing a knowledge-based system for research on Andean weaving. First we have to reorganize and consolidate existing textile collection data, which is quite diverse and spread out over many different locations. For a start, we will focus on three main historical axes: Inca (Cusco, La Paz, Killakas), Tiwanaku (Chilean and Peruvian coast), and Yampara (San Pedro de Atacama to the eastern valleys of Santa Cruz). Knowledge gained from this study will be used as a basis for the second step, developing an ontology that will serve as a common language to integrate data from many heterogenous sources. Finally, we plan to provide a visual interface to describe and search for complex three-dimensional textile patterns. This interdisciplinary project relies on the knowledge of ethnographer-linguists, archaeologists, museum curators, weavers, and computer scientists.

References

[1] D.Y. Arnold, J. de Dios Yapita, and E. Espejo. *Hilos sueltos: Los Andes desde el textil*. Plural and ILCA, La Paz, 2007.

[2] J. Aßfalg, H.P. Kriegel, Peer Kröger, and Marco Pötke. Accurate and efficient similarity search on 3d objects using point sampling, redundancy, and proportionality. In *Proc. of 9th Int. Symposium on Advances in Spatial and Temporal Databases (SSTD)*, pages 200–217, Angra dos Reis, Brazil, 2005.

[3] E.I. Chong, S. Das, G. Eadon, and J. Srinivasan. An efficient SQL-based RDF querying scheme. In *Proc. of the 31st Int. Conf. on Very Large Data Bases (VLDB)*, pages 1216–1227, Trondheim, Norway, 2005.

[4] S. Das, E.I. Chong, Z. Wu, M. Annamalai, and J. Srinivasan. A scalable scheme for bulk loading large RDF graphs into Oracle. In *Proc. of the 24th International Conference on Data Engineering (ICDE)*, pages 1297–1306, Cancún, Mexico, 2008.

[5] M. Doerr. The CIDOC conceptual reference module: An ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3):75–92, 2003.

[6] I. Emery. *The Primary Structures of Fabric: An Illustrated Classification*. Whitney Library of Design, New York, 1986.

[7] M. Frame. The visual images of fabric structures in ancient Peruvian art. In *The Junius B. Bird Conference on Andean Textiles*, pages 47–80, The Textile Museum, Washington, D.C., 1986.

[8] J.M. Hellerstein, J.F. Naughton, and A. Pfeffer. Generalized search trees for database systems. In *Proc. of the 21st Int. Conf. on Very Large Data Bases (VLDB)*, pages 562–573, Zürich, Switzerland, 1995.

[9] H. Kosch and M. Döller. Multimedia database systems: Where are we now? In *Int. Assoc. of Science and Technology for Development - Databases and Applications (IASTED-DBA)*, Innsbruck, Austria, 2005.

[10] G. Kress and T. van Leeuwen. *Reading Images: The Grammar of Visual Design*. Routledge, London, 1996.

[11] H.Y.T. Ngan and G.K.H. Pang. Novel method for patterned fabric inspection using Bollinger bands. *Optical Engineering*, 45(8), 2006.

[12] F. Salomón. *The Cord Keepers: Khipus and Cultural Life in a Peruvian Village*. Duke University Press, Durham, 2004.

[13] V. Sathendra and S. Adanur. Fabric design and analysis system in 3d virtual reality. In *National Textile Center Annual Forum*, Charlotte, North Carolina, 2002.

[14] G. Urton. *Signs of the Inka kipu: Binary Coding in the Andean Knotted-String Records*. University of Texas Press, Austin, 1997.

[15] J. Zhou, L. Ma, Q. Liu, L. Zhang, Y. Yu, and Y. Pan. Minerva: A scalable OWL ontology storage and inference system. In *Proc. of First Asian Semantic Web Conference (ASWC)*, pages 429–443, Beijing, China, 2006.