

New world, New Worlds: Visual Analysis of Pre-Columbian Pictorial Collections.

Daniel Gatica-Perez^{1, 2}, Edgar Roman-Rangel^{1, 2}, Jean-Marc Odobez^{1, 2}, and Carlos Pallan³

¹Idiap Research Institute, Switzerland

²École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

³National Anthropology and History Institute of Mexico (INAH), Mexico

{gatica,eroman,odobez}@idiap.ch

<http://www.idiap.ch/project/codices/>

Abstract. We present an overview of the CODICES project, an interdisciplinary approach for analysis of pre-Columbian collections of pictorial materials – more specifically, of Maya hieroglyphics. We discuss some of the main scientific and technical challenges that we have found in our work, and present a summary of our current technical achievements. This overview stresses the importance of thinking globally and acting both locally and globally with respect to developing approaches for cultural heritage preservation, research, and education.

1 Introduction

The work presented in this paper arises from the collaboration between Switzerland’s Idiap Research Institute and Mexico’s National Institute of Anthropology and History (INAH). The initial ideas and contacts were established as early as 2005, and the resulting CODICES project started in the summer of 2008 with the support of the Swiss National Science Foundation (SNSF). Our interdisciplinary work aims to design, implement, and test computational tools that allow for automatic and semi-automatic description, localization, retrieval, and classification of hieroglyphs of a large digital Maya corpus collected in Mexico.

The Maya is one of several pre-Hispanic cultures that flourished in ancient Mesoamerica. It originated and developed in the mid-Preclassic period (c.a., 1,500 - 400 BC), in the territories that currently spread between the Gulf of Mexico and the Isthmus of Tehuantepec, and southern portions of Mesoamerica including Guatemala, Belize, and Honduras (Fig. 1). This culture reached its climax during the late Classic period (c.a., 600 - 900 AD), when some of their activities achieved impressive levels of refinement, including agriculture, astronomy, architecture, arts, and writing. Our work is contributing tools to facilitate the management and analysis of large collections of photographs of artifacts, monuments, and buildings within the Maya Mexican territory, which have been collected by INAH over a number of years.

This overview discusses some of the main scientific and technical challenges that we have found during our work in the project, and presents a summary

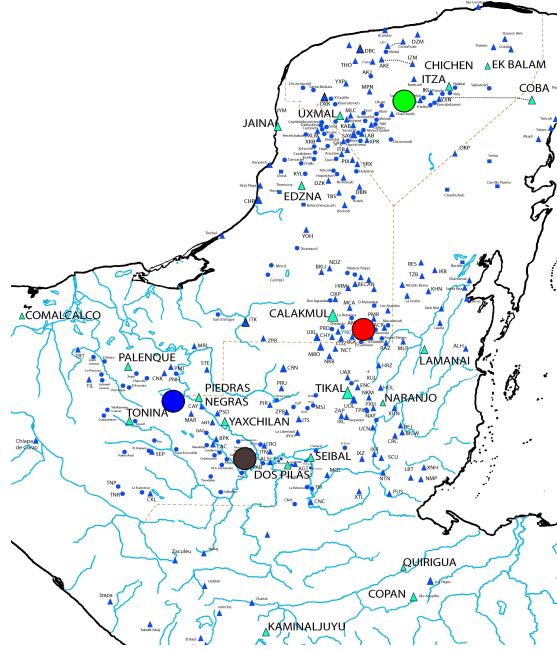


Fig. 1. Maya territories.

of our work so far [7] [8]. Our goal here is to illustrate the kind of computer vision techniques that can be developed to analyze digital versions of ancient materials produced by a culture that, while praised and studied in depth by scholars worldwide – history, archaeology, the arts –, has received less attention from the perspective of multimedia analysis. Our work can be seen as an instance of the local/global research activities that could be envisioned for the future of this domain.

The paper is organized as follows. In Section 2, we briefly introduce the Maya writing system. Section 3 describes the tasks involved in the collection of a Maya hieroglyphic corpus. Section 4 presents a summary of our technical work. Section 5 provides some concluding remarks.

2 The Maya writing system

According to [10], the Maya writing system derives from a large group of languages that developed in southern Mesoamerica. Some of the earliest Maya inscriptions have been dated to be from the late Preclassic period (c.a., 400 BC - 250 AD). During the late Classic period (c.a., 600 - 900 AD), this script system spread all over the entire Maya world, reaching its maximum in the terminal Classic period (c.a., 800 - 950 AD). Since then on it started to diminish, although it continued operational even after the so-called “Maya Collapse” [9] in

few northern sites. Roughly, this writing tradition stayed operational during 17 centuries.

Linguists classify the Maya writing system as a member of the so-called logo-syllabic writing systems. The systems in this class encompass two distinct types of signs: syllabographs and logographs. The former are visual elements without specific meaning, only used to represent phonetic values (i.e., sounds or phonemes), and they usually correspond to a consonant-vowel (CV) or single aspirated vowel structure (Vh). The latter are visual symbols encoding high-level meaning: they approximate our notion of “word-signs”, and the majority of them have a consonant-vowel-consonant (CVC) structure. Fig. 2 shows 5 visual examples of syllabographs and logographs. Roughly speaking, logographs account for 80% of all the Maya hieroglyphs currently known.

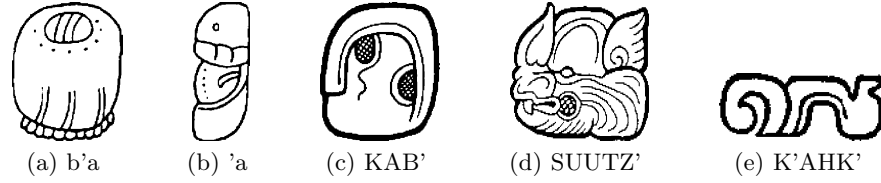


Fig. 2. Examples of Maya hieroglyphs. Syllabographs b'a and 'a; and logographs KAB' (earth), SUUTZ' (bat) and K'AHK' (fire). Images taken from [4] and [11].

Syllabographs might be thought to be simpler than logographs since they only encode sounds instead of ideas. However, this is not true in terms of visual complexity: both syllabographs and logographs may be rich in visual details. Besides the intrinsic complexity of each Maya hieroglyph, the challenge can be increased by additional resources that enrich the inscriptions. Just to mention few examples: *conflation* occurs when two glyphs are visually fused, while retaining their same relative size; *infixation* occurs when one glyph is reduced in size and inserted within another; *superimposition* appears when one glyph partially covers another whose main characteristics remain visible as background; and *pars pro toto* occurs when one glyph is represented by only a fraction of its characteristic or diagnostic features. Fig. 3 shows examples of these phenomena.

Usually Maya hieroglyphs do not appear as single instances but they are arranged inside glyph-blocks where usually logographs are phonetically complemented by syllabographs, either on initial position (prefix or superfix) or in final position (postfix or suffix). In turn, glyph-blocks are found inside complex inscriptions whose organization resembles to a set of pairs of columns. Reading an inscription can be thought as following a scanning pattern in a grid indexed by a system of coordinates, where letters refer to columns and number refer to rows. For instance, the reading of the inscription showed in Fig. 4, with 4 columns and 2 rows would be: A1, B1, A2, B2, C1, D1, C2, D2.

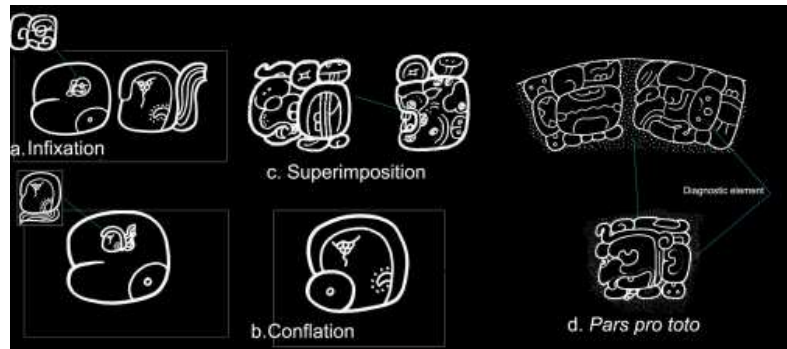


Fig. 3. Examples of complexity in the Maya writing system.

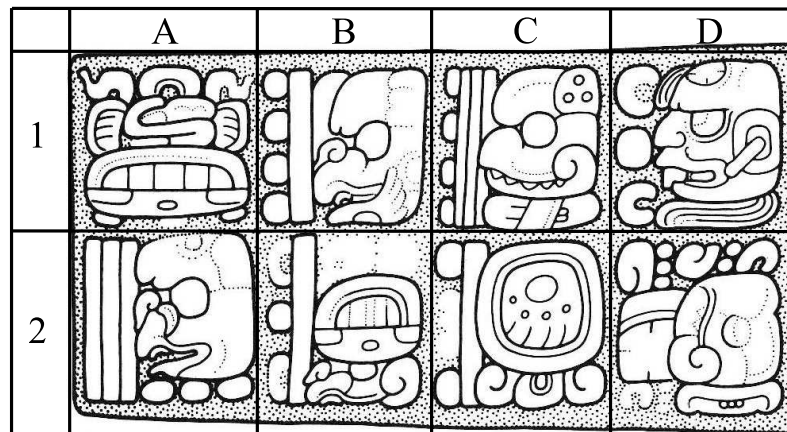


Fig. 4. Reading order of Maya inscriptions in a paired columnar format.

Currently an approximate of 1000 distinct signs have been cataloged, from which only 80% have been deciphered and are readable. Maya archeologists continue exploring and discovering new hieroglyphs at sites and monuments, generating high-quality digital versions of them, which provides plenty of data to researchers in the field of the Maya culture. However, this rapid generation of digital data also posits the need for automatic tools than can help them classify all the new signs discovered.

3 Constructing a digital Maya hieroglyphic corpus

Through the project “Hieroglyphic and Iconographic Maya Heritage” (*Acervo Jeroglífico e Iconográfico Maya*, AJIMAYA), INAH has collected a large collection of photographs of monuments and other buildings within the Maya Mexican territory. Deciphering the inscriptions in these images is an eight-step laborious process:

1. Digital photographs, taken at night under raking-light illumination to bring out the level of detail that facilitates the study of eroded monuments.
2. Line drawings, traced on top of photographs taken under different light conditions, to capture the inner features that are diagnostic.
3. Manual identification of glyphic signs with the aid of glyphic catalogs.
4. Manual transcription, i.e., rendering the phonetic value of each Maya sign into alphabetical conventions.
5. Transliteration, i.e., representing ancient Maya speech into alphabetic form.
6. Morphological segmentation, which breaks down recorded Maya words into their minimal grammatical constituents (morphemes and lexemes).
7. Grammatical analysis to indicate the function of each segmented element.
8. Translation, which involves rendering ancient Maya text on a modern target language, e.g., English.

In Fig. 5 we show an example of the first, second and third steps. Some of the data used in this work has been generated through this process.

4 Our contribution

As part of the CODICES project, we have focused on shape representation approaches for retrieval of hieroglyphs. In this section we explain our two main contributions to the area of cultural heritage. First, we provide details about the syllabic dataset that has been collected. Then, we briefly explain a shape descriptor (HOOSC) that has been designed to deal with the Maya dataset. We follow the explanation by commenting our current results. Finally, we discuss about the need of automatic tools to support archaeological research.



Fig. 5. First, second and third steps in the deciphering process of Maya inscriptions.

4.1 Compilation of a syllabic dataset

The first contribution of our project is the compilation of a dataset of segmented instances of Maya syllabographs, which is meant to be used as testbed for computer vision techniques. Due to the difficulty of manually locating, segmenting, and annotating these instances, we focus only on syllabographs, reserving logographs for future work. With the goal of gathering enough data, this dataset contains instances of the 24 most popular syllabic classes within the AJIMAYA corpus. To the best of our knowledge, this is the biggest dataset of Maya syllabographs that has been analyzed with automatic techniques.

More precisely, the Maya syllabic dataset comprises 1270 segmented syllabographs distributed over 24 visual classes, where each class is referred to by its Thompson catalog number [11]. The dataset also contains 2128 extra segmented glyphs in a “negative class”. The sources for the selected instances are: the AJIMAYA project, the Macri and Looper syllabic catalog [4], the Thompson catalog [11], and the website of the Foundation for the Advancement of Mesoamerican Studies (FAMSI) [6]. Table 1 shows one visual example for each positive class.

For experiments, we have divided the dataset into two subsets, selecting at random 80% of the instances from each positive class and labeling them as “candidates” (G_C), and labeling the remaining 20% as “queries” (G_Q). The purpose of this segmentation is to evaluate the generalization, from candidates to queries, of any computer vision method that is tested on this dataset.



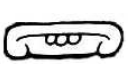




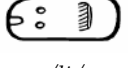










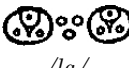


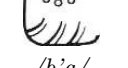


T1	T17	T23	T24	T25	T59
					
/u/	/yi/	/na/	/li/	/ka/	/ti/
T61	T82	T92	T102	T103	T106
					
/yu/	/li/	/tu/	/ki/	/ta/	/nu/
T110	T116	T117	T126	T136	T173
					
/ko/	/ni/	/wi/	/ya/	/ji/	/mi/
T178	T181	T229	T501	T534	T671
					
/la/	/ja/	/a/	/b'a/	/la/	/chi/

Table 1. Thompson numeration, visual examples, and syllabic values (sounds) for the 24 classes of the Maya syllabographs in our dataset.

4.2 The HOOSC descriptor

The Histogram of Orientation Shape-Context (HOOSC) descriptor has been proposed in [8] to describe and retrieve Maya syllabographs in small datasets. This descriptor is robust as it takes advantage of two traditional approaches for image description: a log-polar regional formulation from the Shape Context (SC) [1], and a distribution of orientations from the HOG descriptor [3].

In a nutshell, the HOOSC represents the same shape (glyph) several times from different points; these points are uniformly and randomly selected along the contours of the shape. This can be thought of as looking at the same shape from different perspectives, thus resulting in an aggregated robust description.

More specifically, using a log-polar grid divided in 12 angular and 5 spatial intervals as in [1] (Fig. 6), and whose external boundary spans twice the average pairwise distances of all the points to be described, a feature vector representing each of the selected points is computed as follows:

1. Impose the current point in the center of the log-polar grid, such that all the remaining points are placed around, and the grid includes only those points that are up to twice the average pairwise distances away from the center.
2. Compute the distribution of local orientations of all the points contained in each of the log-polar regions. To take into account uncertainty in orientation estimation and to avoid hard binning effects, this distribution is calculated through a kernel-based approach for orientation density estimation [8].

3. Normalize the resulting vector for each of the 5 spatial intervals, independently for one another, such that the resulting vector sums up to 5.

Since there are 60 log-polar regions and each of them is characterized by a 8-bins histogram of local orientations, the resulting vector has 480 dimensions.

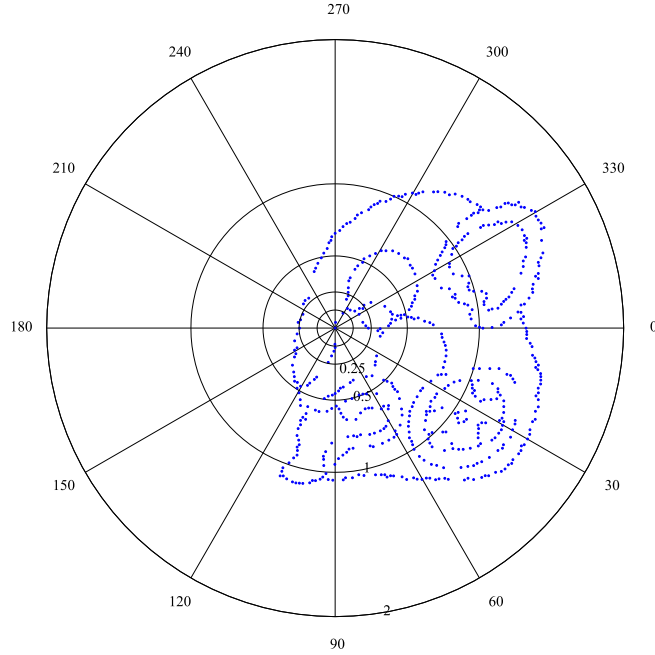


Fig. 6. Log-polar grid with 60 regions used for shape description.

Shape retrieval with the HOOSC. Different shapes might have different degrees of complexity, and therefore different number of points when sampled from their contours. Therefore, a direct comparison of two shapes is rather difficult, and solving a point-to-point correspondence matching could be computationally expensive in some cases. To avoid this, we have used a bag-of-visual-words representation (*bov*) which efficiently compare glyphs.

More precisely, we use the k -means algorithm to quantize the descriptors and to build a bag of visual words. Then we compare pairs of shapes based on the distance between their respective *bovs*. To perform retrieval experiments, we rank the *bov* of candidate-shapes according to their L1 similarity with respect to the *bov* of a given query-shape [5] [8]. We found empirically that 2500 visual words perform well for the HOOSC descriptor.

Improving the HOOSC descriptor The HOOSC descriptor was tested in a small dataset in [8]. However, when its performance is evaluated on the larger syllabic Maya dataset, a drop of almost 10% in the retrieval precision was detected. Recently, we have worked in an improved version of the HOOSC, which not only allows to maintain the retrieval performance but also to improve it in almost 20% compared with the original HOOSC descriptor.

The recent improvements are: a preprocessing filter to thin the contours of the glyphs which results in more stable inputs; an efficient approach to select only key points for the description while remaining accurate; an effective detection and selection of the most informative spatial scope of the descriptor, which allows for shorter feature vectors; and the inclusion of the explicit spatial position of each described point.

4.3 Results

Initially, we compared the performance of different shape descriptors in retrieval tasks [1] [5] [8]. This comparison is made in terms of mean average precision (*mAP*), computed after querying and retrieving candidate hieroglyphs in a small dataset in the order of hundreds of glyphs [8]. The initial results obtained with these descriptors indicate that our method is more suitable to describe complex shapes than the other two approaches. Namely, the SC and the Generalized Shape Context (GSC) result in a *mAP* value of 0.322 and 0.279, respectively, whereas the HOOSC descriptor reached a *mAP* of 0.39.

More recently, the experiments performed on the compiled syllabic Maya dataset, using an improved version of the HOOSC descriptor, resulted in a *mAP* of 0.54. Fig. 7 shows one query example randomly chosen from each syllabic class and the 15 candidates best ranked by the improved HOOSC. Note that in most of the cases our method retrieves relevant glyphs within the first positions.

4.4 Towards a tool for learning about Maya hieroglyphics

Multimedia and Computer Vision techniques can help facilitate the daily work of researchers in the field of Maya archaeology. In our research, we are targeting two specific tools: an automatic analyzer of visual variability and a visual retrieval system.

Automatic analyzer of visual variability. This tool will allow Maya researchers to analyze the visual evolution of the inscriptions through time and across different regions of the Maya territory. In [8] we conducted a preliminary analysis of the intra-class visual variability for 8 syllabic classes. The instances in these classes are labeled as belonging to one of three epochs and one of four distinct regions of the Maya territory. Table 2 shows the average intra-class variability computed for three different periods of time.

Table 3 shows the average intra-class variability for the four Maya regions.

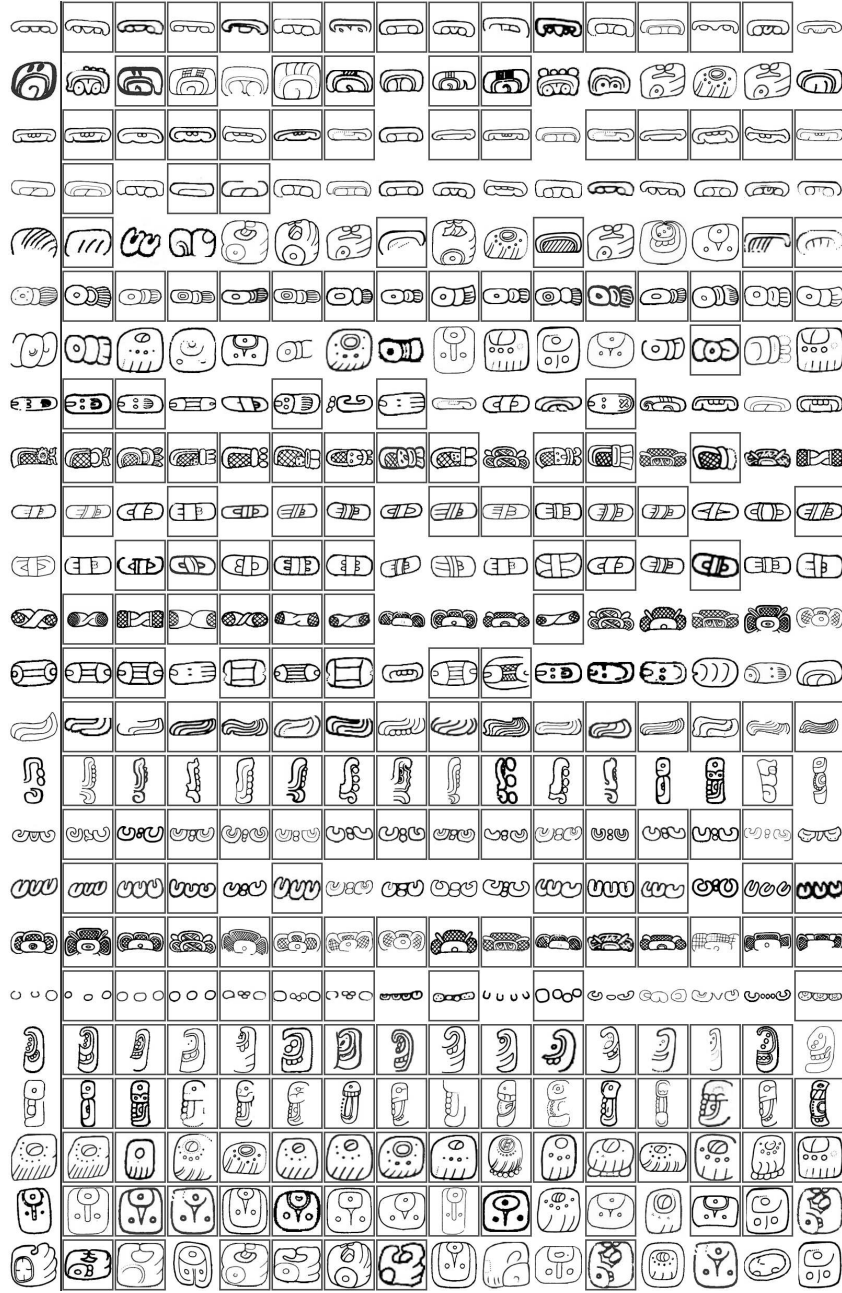


Fig. 7. Retrieval results. The First column shows one random query for each class, followed by its Top 15 retrieved candidate-glyphs shown in terms of decreasing similarity. Relevant glyphs are enclosed in a square.

Period	Time (A.D.)	average
Early Classic	250 - 600	0.277
Late Classic	600 - 900	0.238
Terminal Classic	900 - 1500	0.228

Table 2. Average intra-class visual variability for syllabographs over three time periods of the ancient Maya world.

In map	Region	average
●	Petén	0.251
●	Motagua	0.258
●	Usumacinta	0.349
●	Yucatán	0.214

Table 3. Average intra-class visual variability for syllabographs across four regions of the ancient Maya world. The colors in the bullets correspond to the colors in the map of Fig. 1.

Overall, this preliminary analysis suggests that visual representations converged as time passed, but perhaps diverged as glyphs got spread towards the borders of the Maya world.

A second possible analysis with this tool is that of inter-class similarity. We define a distance measure between classes A and B, as the average of all the distances from each instance of class A to each instance of class B. We use the inverse of this distance as similarity measure and as link strength to construct the graph shown in Fig. 8.

Visual retrieval machine. An accurate visual retrieval system is the main motivation for our current research. This tool will allow archaeologists to quickly search in large collections for instances of a given visual query. Fig. 9 shows one snapshot taken from the first version of this system.

When such a tool is further improved, it will ameliorate the time usually invested by archaeologists in manual search, and it also could help as a training tool for novice scholars learning about the Maya writing system. A video demo of this preliminary tool is available at the website of the project [2].

5 Conclusions

In this overview, we have presented the main developments of the CODICES project. Our work has spanned data collection, shape-based analysis, and the initial steps towards a visual retrieval tool that can be used by archaeologists. In particular, the HOOSC descriptor is an approach that has shown competitive performance w.r.t. state-of-the-art approaches, and represents a suitable starting point to address some of the complexities of Maya hieroglyphics. Further technical details, and future research opportunities were discussed at the workshop.

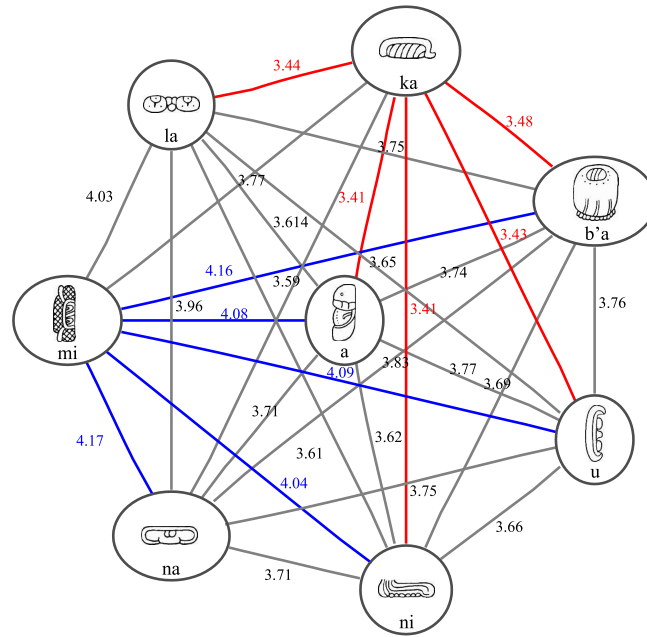


Fig. 8. Inter-class similarity. Each node shows an example of one syllabic class, edges are weighted with the similarity between the two classes it connects to.

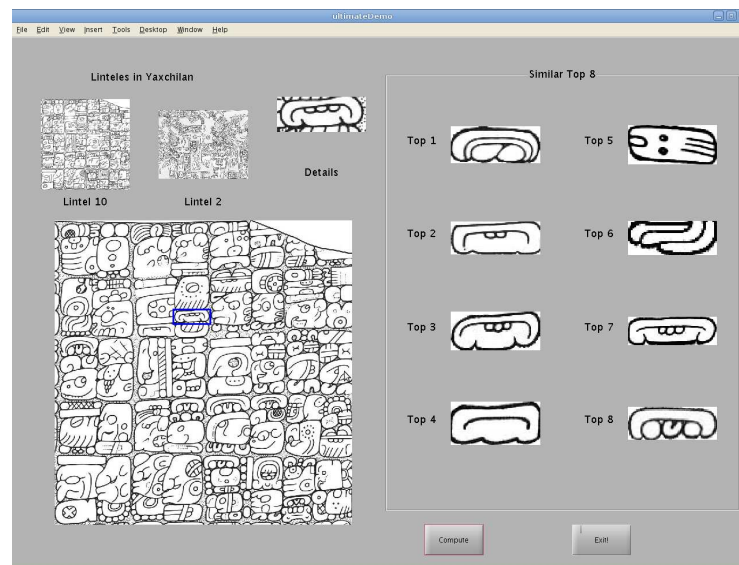


Fig. 9. The system retrieves, from a database, the segmented instances most similar to the selected glyph. Searching within an inscription given a segmented query is also possible.

6 Acknowledgments

We thank the support of the Swiss National Science Foundation through the CODICES project (grant 200021-116702), and to INAH through the AJIMAYA project.

References

1. S. Belongie, J. Malik, and J. Puzicha.: Shape Matching and Object Recognition Using Shape Contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 24(4):509–522, April (2002)
2. CODICES website, <http://www.idiap.ch/project/codices>
3. N. Dalal and B. Triggs.: Histograms of Oriented Gradients for Human Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 886–893. IEEE Computer Society, (2005)
4. M. Macri and M.Looper.: *The New Catalog of Maya Hieroglyphs. Volume 1, The Classic Period Inscriptions*. University of Oklahoma Press : Norman (2003)
5. G. Mori, S. Belongie, and J. Malik.: Efficient Shape Matching Using Shape Contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 27(11):1832–1837, November (2005)
6. M. Pitts and L. Matson.: *Writing in Maya Glyphs*. Foundation for the Advancement of Mesoamerican Studies, Inc., (2008)
7. E. Roman-Rangel, C. Pallan, J.-M. Odobez, D. Gatica-Perez.: Retrieving Ancient Maya Glyphs with Shape Context. In *Proceedings of the IEEE Workshop on eHeritage and Digital Art Preservation. 12th International Conference on Computer Vision*. Kyoto, (2009)
8. E. Roman-Rangel, C. Pallan, J.-M. Odobez, D. Gatica-Perez.: Analyzing Ancient Maya Glyph Collections with Contextual Shape Descriptors. *Special Issue in Cultural Heritage and Art Preservation. International Journal of Computer Vision*. 94(1):101–117. (2011).
9. R. Sharer.: *Daily Life in Maya Civilization*. The Greenwood Press "Daily life through history" series. Westport, CT: Greenwood. (1996)
10. S. D. Stuart, B. MacLeod, S. Martin, Y. Polyukhovich.: Glyphs on Pots: Decoding Classic Maya Ceramics. In: *Sourcebook for the 29th Maya Hieroglyphic Forum*. The University of Texas at Austin, Department of Art and Art History. (2005)
11. J. E. S. Thompson.: *A Catalog of Maya Hieroglyphs*. Norman: University of Oklahoma Press., (1962)