

MAAYA: Multimedia Methods to Support Maya Epigraphic Analysis

Daniel Gatica-Perez^{1,2}, Gulcan Can^{1,2}, Rui Hu¹, Stephane Marchand-Maillet³, Jean-Marc Odobez^{1,2}, Carlos Pallan Gayol⁴ and Edgar Roman-Rangel³

¹Idiap Research Institute, Switzerland

²Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

³University of Geneva, Switzerland

⁴University of Bonn, Germany

Abstract

This paper presents an overview of a research project integrating epigraphy and computer science, which has developed a number of visual analysis methods for segmentation, classification, indexing, and retrieval of Maya glyphs in ancient codices. The paper summarizes the technical aspects of the methods developed for these tasks, including preparation of source materials, creation of a data repository, segmentation of glyph strokes, classification of single glyphs, and visualization of glyph collections. The developed methods, based on computer vision and machine learning, are generic and could be applicable to other sources of visual data in the digital humanities.

1. Introduction

The Maya civilization flourished in Mesoamerica since the late pre-classic period extending into the post-classic. Experts on Maya epigraphy have deciphered many of the hieroglyphic writings of the Ancient Maya over the past 200 years (Macri, 2009). Such epigraphic traces are found in codices, monuments, and ceramics across archaeological sites and museums in Mexico and Central America, and in institutions in Europe and the US. As the field of Digital Humanities starts to make its way into this area of scholarly work, it is clear that computational methods and tools will be increasingly needed to facilitate and accelerate the daily work of epigraphers, to support the investigation of new, possibly data-driven, research hypotheses, and to help training students in this discipline.

In this paper, we present an overview of MAAYA, a project that integrated the work of epigraphists and computer scientists with three goals. The first one was the development of computational methods for visual analysis that help segment, classify, index, and retrieve Maya glyphs from the three original surviving codices hosted in European institutions (Dresden, Madrid, Paris). The second goal was the development of online information management tools that support annotation, search, retrieval, and visualization tasks. Finally, the third, long-term goal is the advancement of Maya epigraphy through the use of these tools.

The paper summarizes work on some of the main modules needed to achieve these goals. This includes the preparation of codex source materials and the creation of a data repository (Section 2); methods for automatic visual analysis, including segmentation of glyph strokes (Section 3) and glyph classification

(section 4); and methods for visualization and browsing of glyph data (Section 5). Most of the content of this paper was originally presented in (Hu, 2015; Can, 2016; Hu, 2017; and Roman-Rangel, 2016a). We conclude the paper with some thoughts on future work in Section 6.

2. Data sources and multimedia repository

Although the codices' exact origin and dating are not entirely known, they were produced in all likelihood within the greater Peninsula of Yucatan during the Post-classic period. The primary data sources are existing reproductions, photographs, and digital scans of the three ancient Maya codices preserved in Germany (Saxon State and University Library Dresden, Dresden), Spain (Museo de America, Madrid), and France (Bibliotèque Nationale de France, Paris). The codices originated in the northern Yucatan peninsula. Other datasets used in the project involve monument-related data, but are not discussed here for space reasons.

The production of a Digital Multimedia Repository (DMR) of Maya hieroglyphic texts and iconography involved researchers at the University of Bonn (C. Pallán Gayol, G. Krempel, P. Biró), research partners at New College of Florida (G. Vail, M. Mermell, H. Neville and E. Cassell) and the Comenius University of Bratislava (J. Špótak). This repository focuses on the production of high-quality clean raster and vectorial visualizations for all hieroglyphic and iconographic data within the Maya codices, with each element linked to a corresponding database record. The Digital Multimedia Repository approaches the codices at different scales or levels of detail:

- (1) Entire codex overview;
- (2) Thematic sections into which each codex is subdivided;
- (3) Almanacs;
- (4) Tools (registers);
- (5) Frames;
- (6) Individual elements (main text glyph-blocks; calendric glyph-blocks, captions, icons); and
- (7) Individual signs or individual iconic elements.

The production of both the Digital Multimedia Repository and the derived datasets involves a number of steps. Roughly, the process includes: (a) advanced image enhancement and post-processing of existing documentation of the codex data sources, including upscaling to 400% normal size, edge refinement, high dynamic range (HDR) tone mapping, noise reduction, and unsharp masking; (b) production of templates providing a mapping of all textual and iconographic elements within any given tool of the codices; (c) cropping of enhanced and upsized images segmented at the glyph-block-level; (d) generation of clean-raster images separating the cultural information (brushstrokes) from background noise and preservation accidents; and (e) production of high-quality vectorial images, including color/grayscale forms, simplified binary forms showing only black and white information, and generation of blue

vectorial forms providing informed reconstructions based on available comparative evidence of all instances of the same sign within the codices (see Figure 1).

Annotation of the glyph collocations was performed on an online server-like capture tool (relying on Filemaker Pro 12), enabling real-time annotation capabilities. Importantly, to facilitate semi-automatic annotation of the glyph-collocations and glyph-strings within the codices, we developed a glyph-concordance tool able to cross-reference four different glyph catalogues (Thompson, 1962; Evrenov, 1961; Macri, 2009; Zimmermann, 1956), which radically reduces the amount of time required when compared to traditional (manually-based) annotation. It also allows advanced multiple co-occurrence queries by cross-referencing the above mentioned catalogs.

From the various datasets produced or used in the project, the Maya Codex Dataset comprises a selection of annotated raster and vectorial renderings of hieroglyphs which, as the name implies, stems exclusively from the codices, and which includes both syllabic signs (that correspond to syllables) as well as logographic signs (that correspond to concepts). This dataset was developed primarily by a team of epigraphists working at the University of Bonn (Pallán, Krempel), the New College of Florida (Vail) and the University of Bratislava (Špoták). Its production steps are therefore analogous to those outlined above for the Digital Multimedia Repository repository.



Figure. 1. Process to generate vectorial representations of the Dresden Codex: from color images to binary images to reconstructed (blue) shapes. (a) Noisy raster; (b) Clean raster; (c) Grayscale vectors; (d) Binary vectors; (e) Blue-reconstructed vectors; (f) Composite showing blue and grayscale vectors. Image by Carlos Pallán Gayol based on a SLUB Library (Dresden) open-access image¹.

3. Segmentation of glyph strokes

The surviving ancient Maya codices often lose their visual quality over time. In addition to the fading of ink, background noise is often introduced due to deterioration of the material, ink stains, etc. Extracting clean glyph strokes from raw images of the ancient scripts is a challenging yet crucial step for any further automatic document image analysis tasks such as page layout analysis, character recognition, and others. In this section, we present a system for automatically extracting hieroglyph strokes from images of degraded ancient Maya codices, presented at length in (Hu, 2017). Our system adopts a region-based image segmentation framework.

Raw glyph block images are first represented by over-segmented homogeneous image regions using the Simple Linear Iterative Clustering (SLIC) algorithm, which clusters image pixels into local compact, homogeneous, and edge-aware regions (Achanta, 2012). Fig. 2 illustrates the super-pixel representation of examples of glyph block images in two different resolution scales (i.e., different number of regions).

¹ [http://www.slub-dresden.de/sammlungen/digitale-sammlungen/werkansicht/cache.off?tx_dlf\[id\]=2967](http://www.slub-dresden.de/sammlungen/digitale-sammlungen/werkansicht/cache.off?tx_dlf[id]=2967)

The number of super-pixels extracted is a key factor that will affect the glyph binarization result. A larger number leads to binarization results that better preserve details such as thin glyph strokes. However, this also keeps more background noise. In contrast, a smaller number of super-pixels leads to a coarser image presentation resolution, and results a cleaner binarization result. However, it will tend to lose more delicate glyph stroke details. As a result of these observations, we build our system using multiple resolution super-pixel regions. A support vector machine (SVM) classifier is used to label each super-pixel region with a probability to belong to foreground glyph strokes. Pixel-wise probability maps from multiple super-pixel resolution scales are then aggregated to cope with various stroke widths and background noise. Images in the top row of Fig. 3 show the segmentation results using this method. Finally, a fully connected Conditional Random Field (CRF) model (Krahenbühl, 2011) is applied to improve the labeling consistency. Comparing the segmentation results in the bottom row to those of the top row in Fig. 3, we can see that more delicate details are preserved by using the fully connected CRF model.



Figure 2. Super-pixel representation of glyph block images using different pre-defined numbers of regions: (top row) 500 regions; (bottom row) 2000 regions. Each super-pixel is displayed using the average color of pixels within that region. Figure adapted from (Hu, 2017).



Figure 3. Segmentation results using only the appearance component (top row); and with the fully connected CRF model (bottom row). Figure adapted from (Hu, 2017).

From the results we can see that our method preserves stroke details and reduces background noise. As an application in (Hu, 2017), we conducted glyph retrieval experiments using manually produced clean raster images (generated by the epigraphers in our team) and our automatically generated glyph

strokes. The experiments show that the machine-extracted glyph strokes achieve comparable retrieval results to those obtained using glyphs manually segmented by epigraphers. This is a very positive result, as it enables the processing of raw data in later tasks.

4. Visual classification of single glyphs

Robust shape representations are critical for visual analysis of glyphs. We have studied and compared two types of shape representations in a bag-of-words based pipeline to recognize Maya glyphs in (Can, 2016). The first representation is a knowledge-driven descriptor, called Histogram of Orientation Shape Context (HOOSC) (Roman-Rangel, 2011). The second one is a data-driven representation obtained by applying an unsupervised Sparse Autoencoder (SA), a classic model in machine learning (Hinton, 1994). Both methods operate on a finite set of points sampled from the binary shape that represents a glyph.

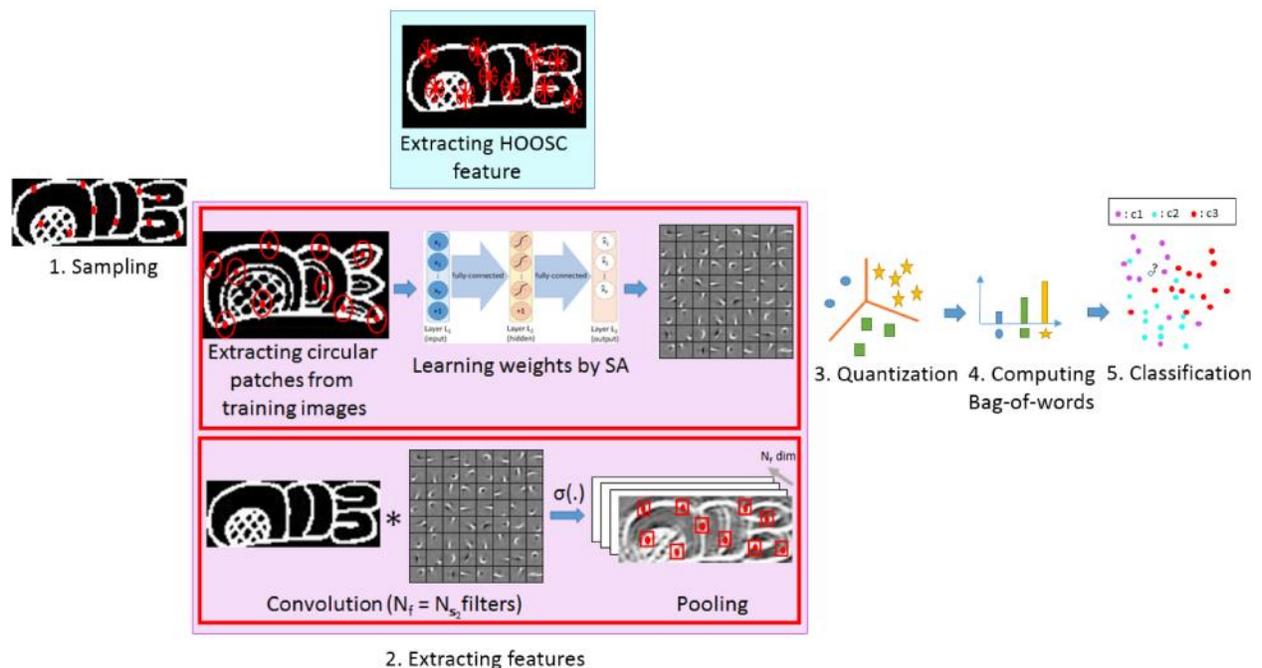


Figure 4. Glyph classification system. For feature computation, the HOOSC method (blue block) or the SA method (purple block) can be used. The SA approach requires a learning phase. Figure taken from (Can, 2016).

Figure 4 illustrates the bag-of-words pipeline used for our comparative study. For the same set of points sampled from each shape, we compute both the HOOSC and SA representations. The HOOSC descriptor is defined on a circular grid centered on each point and computes the histogram of orientations of the contours over the grid. The SA representation requires a learning phase from raw pixel values. The learned filters for the SA representation reveal the type of image content (see gray-level basic shapes that appear in the purple block), and are tuned to the inherent motifs in the shape data.

These representations have been evaluated both on Maya glyph data (a 10-class set of single glyphs, which is a small subset of all glyph categories that are present in the codices), and a larger-scale hand-

drawn sketch dataset, spanning 250 classes of general objects (Eitz, 2012) to test the generalization ability of the descriptors. We evaluated the performance of both representations with different parameter settings.

From the experiments, the HOOSC descriptor performed better than the SA representation on the smaller dataset (glyph), for which the HOOSC descriptor was originally designed for. In contrast, when other data is considered (10-class experiments on sketches) and more classes are present (250 sketch classes), the SA representation was able to surpass the HOOSC descriptor (29.3% vs. 20.7% on the 250-class experiments). This shows the importance (and the necessity) of significant data volumes for data-driven approaches. We also observed that using a larger number of hidden units in the SA method improved the performance.

With dense sampling, i.e., when more points are used as the basis to compute the descriptors, the HOOSC descriptor improved substantially when considering large enough spatial regions, reaching 35% accuracy on the 250-class case, in par with that of (Eitz, 2012). On the other hand, the SA representation did not benefit from this factor, due to the quadratic increase of parameters during the learning phase.

Overall, this study, using both Maya hieroglyphics and generic hand-drawn shapes, informs about the performance trends of the studied methods, and could also be informative for their possible use on other cultural heritage sources of visual shapes. This includes other Maya glyph datasets and other sources, for instance Egyptian hieroglyphs.

5. Browsing collections of glyph blocks

A major objective of research in archaeology consists in interpreting traces left by past activities. Once discovered, every such trace is documented for its intrinsic properties such as geographic location or estimated date of origin. This allows archaeologists to discover high-level patterns that can then be interpreted for their high-level meaning, e.g., semantics. Whereas the diversity and volume of collected traces may make this task very complex and tedious, machine intelligence could support this action by its capabilities to reason over large and diverse sets of items at once. We therefore advocate for the use of machine learning capabilities in the context of data exploration and, in particular, for the creation of systematic similarity-based organization of this data.

In the case of Maya hieroglyph decipherment, epigraphers document their findings based on the place the writing was found or the type of artifact it appeared on. From this factual information, they hypothesize further interpretations of the content, based on any contextual information they may exploit. Here, information technologies may be of help by providing large browsable catalogs in which to situate the piece in question in terms of estimated date of origin, similar locations, or more perceptual information such as visual similarity. In the further particular case of Maya codices, this strategy has allowed to break down the codices into pieces at various interpretation level, such as pages, t'ols, glyph-blocks, iconic blocks, or calendrics.

We have thus built several browse and search structures over these collections by defining similarity measurements on inferred intrinsic or perceptual properties (Sun, 2015; Marchand-Maillet, 2016). We could thus provide epigraphers with interfaces allowing tasks such as browsing codices per item (i.e., page, t'ol, glyph-block, etc.), browsing per string of glyph codes a.k.a., T-string, browsing glyphs and glyph-blocks per visual similarity, browsing per annotation similarity, searching T-Strings, counting co-occurrences, etc. This is illustrated in Fig. 5.

These interfaces are built upon mathematical descriptors, which for the case of browsing by visual similarity correspond to a collection of numeric values that captures statistics about the visual information of a hieroglyph of interest, like the HOOSC descriptor described in Section 4 (Roman-Rangel, 2011; Hu, 2015). Having such descriptors makes easier to perform efficient comparison of visual elements from a mathematical point of view. In this regard, we have described Maya hieroglyphs at the level of detail of glyph-blocks in our project. For this, the descriptors must be robust enough to account for visual variations of the individual glyph-signs (e.g., rotation or scaling), and they must also capture visual similarity across different signs (Roman-Rangel, 2011). Furthermore, ideally they must be as short as possible, i.e., of low-dimensionality, such that it is computationally efficient to perform browsing and retrieval operations on them.

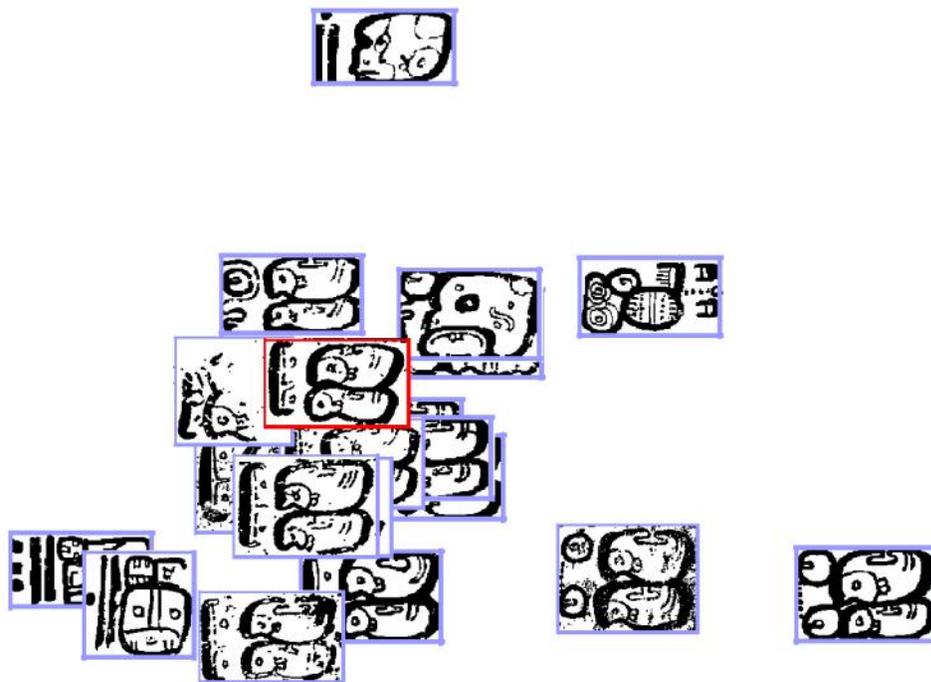


Figure 5. Example of a browser of glyph blocks that uses visual similarity to place similar blocks in adjacent positions.

In our work, we have investigated the impact of using histograms of local descriptors like HOOSC and neural representations for such representations (Roman-Rangel, 2016a). Local descriptors are statistics

describing different locations of an image. Since they correspond to handcrafted methods tailored for specific purposes, e.g., description of hieroglyphs in our case, they have high potential for effective browsing. Contrarily, neural representations are methods that rely on learning strategies that can learn a wide variety of complex mathematical functions. Therefore, they are able to estimate more robust representations of glyph-blocks in comparison to local descriptors, and they can perform well in practice in different learning tasks. Namely, our work shows that they obtain high retrieval performance when dealing with glyph-blocks.

One limitation of the neural representations is that they require large amounts of data to ensure accuracy in their training. Since the amount of available Maya glyphs is limited to the existing codices, we risk to end up with models representing only a subset of the potentially larger existing variations of hieroglyphs, i.e., glyph-blocks might be under-represented by the resulting neural models, with the potential consequence of not recognizing new instances of, otherwise assumed, known signs. To address this issue, we have investigated the impact of combining local descriptors with neural representations (Roman-Rangel, 2016b). In this case, we use neural models to compute representations of specific locations within an image, for a latter combination of such representations that provide a full glyph-block description. Our results show that this approach is promising, as it obtains state-of-the-art performance.

As previously mentioned, one characteristic of the image representation is that they are of low-dimensionality. This not only accounts for efficient computational processing, but also could allow us to show glyph-blocks in 2-D browsing interfaces. For this, we are currently investigating dimensionality reduction techniques to produce mathematical representations of glyph-blocks, as short as two dimensions, such that signs that are visually similar to one another are placed closer to those that are visually different. We hope to integrate such approach in future versions of our current browsing interfaces.

6. Conclusion

We presented an overview of the MAAYA project, summarizing work on data preparation and annotation, glyph stroke segmentation, glyph classification, and glyph visualization and browsing. Based on what the project achieved, there are several opportunities for future research in this domain. To close the paper, we briefly comment on two of them.

First, the generation of a digital dataset of glyphs present in the three codices is a milestone of our work, as it has enabled the development and objective evaluation of a variety of methods of visual analysis and visualization. The data and methods could thus be relevant for other researchers working in ancient Maya epigraphy. More generally, the computer vision and machine learning methods summarized here could be applicable to other sources of data in the digital humanities, as they are generic and thus potentially amenable to other problems.

Second, there is a need to deepen the understanding of how scholars and other publics (e.g. novice students in the discipline and the general public) could best benefit from the methods we developed.

This can be framed as a human-computer interaction (HCI) problem. The collaboration across computing and epigraphy produced a couple of designs of interfaces intended for epigraphists, as well as initial results on how the presented information is perceived and analyzed by experts. However, the space to improve these initial prototypes and to adapt them to different potential users is significant. This is an area that future work needs to address.

Acknowledgments. This work was funded by the Swiss National Science Foundation (SNSF) and the German Research Foundation (DFG) through the MAAYA project. We thank all collaborators for their work.

References

(Achanta, 2012) R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 11 (2012), 2274–2282.

(Can, 2016) G. Can, J.-M. Odobez, and D. Gatica-Perez. Evaluating Shape Representations for Maya Glyph Classification, *ACM Journal on Computing and Cultural Heritage*, Vol. 9, Issue 3, Sep. 2016

(Eitz, 2012) M. Eitz, J. Hays, and M. Alexa. How Do Humans Sketch Objects? *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, 31, 4, 44:1–44:10, 2012.

(Evrenov, 1961) E.B. Evrenov, Y. Kosarev, and B. A. Ustinov.. *The Application of Electronic Computers in Research of the Ancient Maya Writing*. USSR, Novosibirsk, 1961.

(Hinton, 1994) G. E. Hinton and R. S. Zemel. Autoencoders, minimum description length, and Helmholtz free energy. In *Proc. Conf. on Advances in Neural Information Processing Systems (NIPS)*, 1994.

(Hu, 2015) R. Hu, G. Can, C. Pallan Gayol, G. Krempel, J. Spotak, G. Vail, S. Marchand-Maillet, J.-M. Odobez, and D. Gatica-Perez. Multimedia Analysis and Access of Ancient Maya Epigraphy. *IEEE Signal Processing Magazine*, Vol. 32, No. 4, pp. 75-84, Jul. 2015.

(Hu, 2017) R. Hu, J.-M. Odobez, and D. Gatica-Perez, Extracting Maya Glyphs from Degraded Ancient Documents via Image Segmentation. *ACM Journal on Computing and Cultural Heritage*, Vol. 10, Issue 2, Apr. 2017.

(Krahenbühl, 2011) P. Krahenbühl and V. Koltun. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials. In *Proc. Conf. on Advances in Neural Information Processing Systems 24 (NIPS) 2011*.

(Macri, 2009) M. Macri and G. Vail. *The New Catalog of Maya Hieroglyphs, Volume Two: The Codical Texts*. University of Oklahoma Press, 2009.

(Marchand-Maillet, 2016) S. Marchand-Maillet, E. Roman-Rangel, H. Mohamed, and F. Nielsen. Quantifying the Invariance and Robustness of Permutation-based Indexing Schemes. In Proc. Int. Conf. on Similarity Search and Applications (SISAP), Tokyo, 2016.

(Roman-Rangel, 2011) E. Roman-Rangel, C. Pallan, J.-M. Odobez, and D. Gatica-Perez. Analyzing ancient Maya glyph collections with contextual shape descriptors. *International Journal of Computer Vision*, 94, 1, 101–117, 2011.

(Roman-Rangel, 2016a) E. Roman-Rangel, G. Can, S. Marchand-Maillet, R. Hu, C. Pallan Gayol, G. Krempel, J. Spotak, J.-M. Odobez, and D. Gatica-Perez. Transferring Neural Representations for Low-dimensional Indexing of Maya Hieroglyphic Art. In Proc. ECCV Workshop on Computer Vision for Art Analysis (VISART), Amsterdam, 2016.

(Roman-Rangel, 2016b) E. Roman-Rangel and S. Marchand-Maillet. Indexing Mayan Hieroglyphs with Neural Codes. In Proc. Int. Conf. on Pattern Recognition (ICPR), Cancun, 2016.

(Sun, 2015) K. Sun, J. Wang, A. Kalousis, S. Marchand-Maillet. Space-Time Local Embeddings. In Proc. Conf. on Neural Information Processing Systems (NIPS), Montreal, 2015.

(Thompson, 1962) J. E. S. Thompson, *A Catalog of Maya Hieroglyphs*. University of Oklahoma Press, 1962.

(Zimmermann, 1956) G. Zimmermann, *Die Hieroglyphen der Maya Handschriften. Abhandlungen aus dem Gebiet der Auslandskunde*, Band 62- Reihe B, Universität Hamburg. Cram, De Gruyter & Co., 1956.