Automatic Maya Hieroglyph Retrieval Using Shape and Context Information

Rui Hu
Idiap Research Institute
Switzerland
rhu@idiap.ch

Carlos Pallán Gayol
University of Bonn
Germany
pallan.carlos@gmail.com

Guido Krempel
University of Bonn
Germany
tilalhix@gmail.com

Jean-Marc Odobez
Idiap Research Institute,
EPFL
Switzerland
odobez@idiap.ch

Daniel Gatica-Perez
Idiap Research Institute,
EPFL
Switzerland
gatica@idiap.ch

ABSTRACT

We propose an automatic Maya hieroglyph retrieval method integrating shape and glyph context information. Two recent local shape descriptors, Gradient Field Histogram of Orientation Gradient (GF-HOG) and Histogram of Orientation Shape Context (HOOSC), are evaluated. To encode the context information, we propose to convert each Maya glyph block into a first-order Markov chain and apply the co-occurrence of neighbouring glyphs. The retrieval results obtained based on visual matching are therefore re-ranked. Experimental results show that our method can significantly improve the glyph retrieval accuracy even with a basic co-occurrence model. Furthermore, two unique glyph datasets are contributed which can be used as novel shape benchmarks in future research.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Information Search and Retrieval; I.4 [**Image Processing and Computer Vision**]: Feature Measurement—*Feature representation*

Keywords

Maya hieroglyph, image retrieval, shape descriptors, glyph co-occurrence, Markov model

1. INTRODUCTION

The Maya civilization is one of the major cultural developments in ancient Mesoamerica. It began to flourish during the Pre-Classic period (2000 BC to AD 250), reached the peak of development during the Classic period (AD 250-900), and continued throughout the Post-Classic period.

MM'14, November 3–7, 2014, Orlando, Florida, USA.

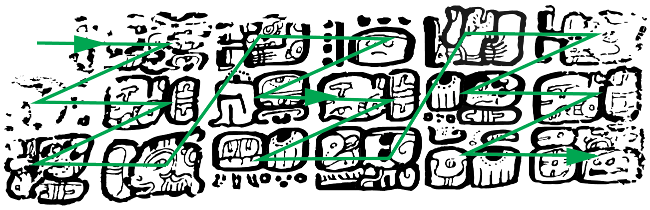


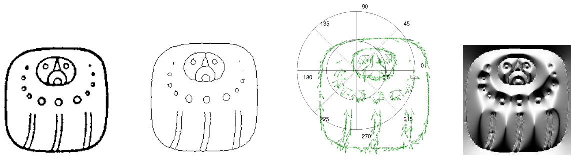
Figure 1: Codex Dresden. Main hieroglyphic text on page 65a. Green arrows indicate reading order of the glyph blocks.

Most Maya texts were written during the Classic period on ceramic vessels and stone monuments to record historical events. There are also Maya texts written on personal items such as wooden lintels, mural paintings, bones, etc. A special type of Maya textual sources are the so called Codices, which are stuccoed and painted bark paper books, folded in laparello format. Such books usually contain both textual and iconographic information, survived from the Post-Classic period. Fig.1 illustrates one exemplified main text from page 65a of the Dresden Codex (high resolution images of this codex are available at the SLUB website [2]).

Maya texts are usually composed of glyph blocks arranged in double-columns. The usual reading order is from left to right and from top to bottom by means of double columns (Fig.1). Due to the flexibility in artistic freedom of Maya writing, diverse forms and different glyph-variants were in used. Thus the graphic conventions within a glyph block can vary greatly. One glyph block could contain one or multiple glyphs; and it could correlate with a morpheme, an individual word, or even an entire sentence. The choice could be due to economic, religious or aesthetic reasons. Individual glyph recognition is crucial to Maya script decoding.

Maya decipherment has about 200 years of scholarly tradition. Over 1000 signs have been discovered. While about 80% of signs can be read phonetically, the meaning of some 40% logograms remain unknown [8]. Epigraphers often spend a tremendous amount of time looking at traditional printed catalogs to identify unknown glyphs. Therefore, the aim of this work is to facilitate these tasks by automatic recognition of Maya glyphs using shape and context information.

Related work. Computer vision algorithms have shown the potential of providing new insights in automatic visual



(a) (b) (c) (d)

Figure 2: (a) Original glyph (taken from [12]); (b) thinning; (c) HOOSC spatial quantization of a given pivot; (d) gradient field for GF-HOG computation.

analysis of cultural heritage, which can significantly facilitate the daily work of related scholars. Previous works address Maya glyph retrieval as a shape matching problem.

Shape based retrieval systems include curve fitting [4], point-to-point matching [3], grid based matching [5], etc. These works either do not scale well over large datasets, or only offer limited flexibility over affine transformation. Recently, local shape descriptors [6, 7, 10] have been proposed and used in a Bag-of-Visual-Words (BoVW) framework for shape based retrieval. Such method can be scaled sub-linearly with appropriate search structures, and also offers certain degrees of freedom of affine invariance. Our work follows this approach.

One of the first attempts of automatic Maya glyph retrieval is [9], where shape context is used to represent and match between single glyphs. A shape descriptor (HOOSC) is proposed in [10], and used in the BoVW system for Maya glyph retrieval. The retrieval performance is improved in [11] by using only the intermediate spatial scope, as well as carefully pre-processing glyphs into thinned edges, and evenly pick pivot points. In another direction, GF-HOG [7] was recently developed in the context of sketch based retrieval to overcome the limitation of HOG-like local shape descriptors when applied to shapes rather than images.

These previous Maya glyph retrieval works have framed the glyph retrieval problem without considering any specific structure of Maya writing as a language. In contrast, language modelling has been widely used in machine translation and speech recognition. Glyph context information has recently been applied in [6] for Egyptian hieroglyph analysis with limited performance improvement. To the best of our knowledge, Maya glyph co-occurrence that captures context information in language has never been considered in automatic Maya glyph analysis systems.

Contribution. In this paper we investigate the following questions. First, given the promising results GF-HOG has achieved for sketch-based image retrieval, how would it perform on glyph data? Second, would glyph co-occurrence information help improve the shape based glyph retrieval accuracy? As exemplified by [6], co-occurrence appears not consistent enough for the task of Egyptian hieroglyph retrieval. Another contribution of this paper regards two glyph datasets collected and produced from multiple sources (see section 4.1). Both datasets can be used as shape benchmarks, and one of them can also be used to further study the Maya language model. Expert epigraphers in our team contribute tremendous effort in creating these two datasets.

2. SHAPE BASED GLYPH RETRIEVAL

Each glyph is represented by a set of local shape descriptors. The BoVW approach is then applied.

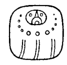
T0501	T0502	T0668	T0757	T0102	T0103
					
/b'a/	*/ma/*	*/cha/*	*/b'a/*	*/ki/*	*/ta/*

Table 1: Thompson numbers, visual examples, and the syllabic values of glyph pairs. Each pair contains two different signs of similar visual features. All images are taken from [12]

2.1 Shape representation

We use the local shape descriptors HOOSC and GF-HOG for our glyph retrieval tasks. Since they achieves the current state-of-the-art performance on shape based Maya hieroglyph retrieval and sketch based image retrieval separately.

HOOSC. Since Maya glyphs often present stroke lines with different degrees of thickness, we first apply a thinning algorithm [11] to get a thinned version of the glyph, see Fig.2 (b) for example. A uniformly sampled subset of points along the contour is then selected as the pivot set. For each pivot point, we calculate the HOOSC descriptor by computing the histogram of orientation within each spatial bin of a local grid centred at the pivot's location, as illustrated in Fig.2 (c). We consider 8 orientations and 2 rings. The outer ring covers the average pairwise distance between each pair of points along the contour; the inner ring covers half of this distance. The final descriptor of a given point is the concatenation of histograms of 16 grids. We use a 8 bin histogram for each grid, which forms a 128-dimensional feature vector.

GF-HOG. Compared with natural images, shapes like glyphs provide only sparse spatial information. Thus, instead of computing HOG on the sparse gradient field along the contour, GF-HOG aims to encode a richer structure information within each local window. This is achieved by interpolating a dense gradient field from the sparse orientation field using a Laplacian constraint. It can be solved using a Possion equation with Dirichlet boundary condition. See [7] for detail. In our work, the gradient field is generated from the thinned glyph (Fig.2 (b)). Fig.2 (d) shows an example of the extracted gradient field. HOG descriptors are then extracted from the dense gradient field around each pivot point using a 4×4 grid. A 8-bin orientation histogram is computed within each grid. This results in a 128-dimensional feature vector. Following [7], we compute HOG descriptor over three window sizes to improve the scale invariance.

2.2 BoVW based retrieval system

BoVW approach is applied to quantify local features and represent the full shape with a histogram. Specifically, given the local descriptors extracted from each glyph, k-means clustering is applied to generate a set of k clusters. Each cluster is referred to as a 'visual word' and the set of k clusters defines the vocabulary of the system. A frequency histogram which represents the count of the visual words extracted from an image is then used as a global descriptor.

Given a query G and a database glyph D represented by their histograms H^G and H^D, we compute the Cityblock distance to measure the dissimilarity between G and D:

$$d(G, D) = \sum_{1 \leq i \leq k} |H^G(i) - H^D(i)| \quad (1)$$

where each histogram is normalized so that $\sum_{1 \leq i \leq k} H(i) = 1$.

glyph block	glyph string
	
	

Table 2: Generating glyph strings from blocks.

3. GLYPH CO-OCCURRENCE

Using shape alone to represent and distinguish between glyphs is often ambiguous for many reasons. First, different signs often share similar visual features, see for example Table 1. Furthermore, glyphs of the same sign category vary with time, location, as well as with individual styles, calligraphies and renderings of different artists. For example, the first block in table 2 and the third block in Fig.3 are both *cha-ki* (the name of the rain god). Finally, the surviving historical Maya scripts often lose their visual quality over time; as such, it is often the case that noise is introduced and data is partially missing. Therefore, we propose to use glyph co-occurrence to complement the visual information and help improve the retrieval accuracy.

3.1 Glyph co-occurrence in Maya writing

In Maya hieroglyphic writing, glyph blocks were frequently composed of combinations, conflations and mergings of individual signs. Glyph co-occurrence within single glyph blocks could encode valuable information. The reading order within a block usually follows the basic rule of left-to-right and top-to-bottom. We propose to convert each glyph block into a linear string of glyphs, according to this reading order, as shown in Table 2. Given a glyph string, we consider the 1^{st}-order co-occurrence of neighbouring glyphs.

Co-occurrence model. We extract glyph co-occurrence information from the classic Thompson hieroglyphic catalog [12]. Despite its outmoded taxonomy, it remains one of the most comprehensive and widely used list of Maya glyphs. It covers 892 signs extracted from both codices and monuments. Thompson categorises signs into affixes, main signs and portraits (note that this categorisation method is no longer used in recent research of Maya writing). Affixes often co-occur with main signs, portraits as well as affixes to form blocks. In his book, Thompson provides two glyph co-occurrence tables, listing co-occurring signs discovered at that time for each affix, distinguishing whether it is used as prefix or postfix. However, no frequency information is given in these tables, and co-occurrence between main signs and portraits is not listed. We extract a co-occurrence matrix C from these tables:

$$C(S_i, S_j) = \begin{cases} 1 & \text{sign } S_i \text{ appears before sign } S_j, \\ \alpha & \text{otherwise.} \end{cases} \quad (2)$$

where α is a smoothing factor that accounts for missing co-occurring glyph pairs in the two co-occurrence tables. Here we set $\alpha = 0.01$. Note that $C(S_i, S_j) \neq C(S_j, S_i)$.

Co-occurrence for glyph labelling. Denote by $G_{1:n} = [G_1, \ldots, G_i, \ldots, G_n]$ the observed glyph string, and by $S_{1:n}$ the sequence of recognized states, where S_i indicates the Thompson number annotated with glyph G_i. Considering the glyph string $G_{1:n}$ as a 1^{st}-order Markov chain, the probability of labelling it to a sequence of states $S_{1:n}$ is:

$$P(S_{1:n}|G_{1:n}) \propto P(G_1|S_1) \prod_{2 \leq i \leq n} (P(G_i|S_i)P(S_i|S_{i-1})), \quad (3)$$

Figure 3: Example glyph blocks of the MC dataset.

where $P(S_i|S_{i-1})$ denotes the transition probability from the previous state S_{i-1} to the current state S_i. Here we use $C(S_{i-1}, S_i)$ to approximate this probability. $P(G_i|S_i)$ refers to the likelihood of glyph G_i being labelled as sign S_i. To encode this term we use the visual similarity between G_i and the glyph example of S_i in the database, and define $P(G_i|S_i) \propto e^{-d(G_i,S_i)/\lambda}$, where $d(G_i, S_i)$ is computed using Eq.(1), λ is a scale factor set to the average distance of the top 50 ranked results for all queries.

3.2 Retrieval system

We solve glyph recognition through retrieval, i.e. by ranking the database glyphs D (signs in Thompson catalog [12]) based on their similarity with the query. When only shape information is considered, the score of a query glyph G_i being labelled by sign D, is computed by their shape likelihood:

$$Score^{shape}(S_i = D) \propto P(G_i|S_i = D). \quad (4)$$

In our model, we propose to rank the glyphs according to

$$Score^{shape+context}(S_i = D) = \max_{S_{1:i-1;i+1:n}} P(S_{1:n}|G_{1:n}) \quad (5)$$

which means, given $S_i = D$ finding the sequence of Thompson labels $S_{1:n}$ that provides the maximum probability to label $G_{1:n}$, under the model in Eq.(3). This can be efficiently computed using the Viterbi algorithm. In other words, the score of the glyph G_i being recognized as $S_i = D$ now takes into account all observed glyphs in the string, with the effect that a glyph D that normally co-occurs with glyphs that are visually likely at neighbouring positions will receive a higher weights.

4. EXPERIMENTS

Two sets of experiments are conducted. In the first case, we compare the two local shape descriptors GF-HOG and HOOSC for shape based single-glyph retrieval. In the second case, we incorporate context information within single blocks to improve the glyph retrieval accuracy.

4.1 Datasets

Two glyph datasets are contributed, both requiring dedicated expertise of expert epigraphers to collect and compile.

Maya Syllabic dataset (MS). MS is an adapted version of the Syllabic Maya dataset used in [11], which contains manually segmented glyphs from various resources by experts. With the support of experts in our team, we refined the dataset by correcting a few glyph categories and removing uncertain glyphs. Note that the grouping of glyphs is based on the Thompson catalog of Mayan glyph signs, not on shape. The resulting dataset contains 1044 glyphs of 25 sign categories. We randomly selected 20% glyphs from each category to form the query set, resulting in a total of 207 glyphs. All remaining glyphs formed the test set.

Maya Codex dataset (MC)[1]. MC is a new high-quality Maya hieroglyph dataset developed by experts in our team. Currently, it contains 174 glyphs from 72 blocks segmented from two Maya codices (Dresden and Madrid), along with

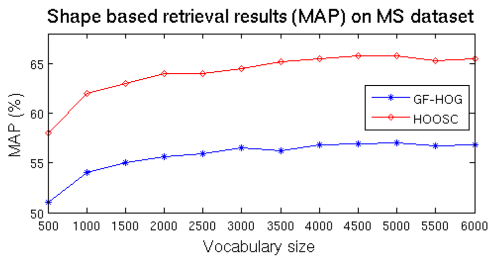


Figure 4: Shape based retrieval results (MAP) with varying vocabulary size on the MS dataset.

their segmented individual glyphs and corresponding annotations. See table 2 and Fig.3 for examples. Our work has defined consistent conventions to generate high-quality vectorial representations of the three existant Maya codices (Dresden, Madrid, Paris) using Adobe Illustrator. Note that these three codices are named after the places where they are currently exhibited. Given that individual signs are structured within blocks, this dataset can also be used to study the Maya language model. Currently, we only have a limited amount of examples available, since epigraphers require about 30 minutes to render a high quality glyph-block depending on complexity and preservation factors. We are constantly enlarging and improving this dataset. To automatically label glyphs from the MC dataset, we use them as queries to match with the 892 signs of the Thompson catalog [12], which we have scanned and manually segmented.

4.2 Results

Fig.4 shows the Mean Average Precision (MAP) of using GF-HOG and HOOSC on the MS dataset. From the curves we can see that HOOSC outperforms GF-HOG for around 8% on average with varying vocabulary sizes.

Glyph retrieval on the MC dataset is a more challenging task. Indeed, in spite of the large visual variability of single glyph categories, most signs of the catalog [12] are represented by only one instance. Furthermore, catalog signs of different categories often share similar visual feature. Finally, query glyphs are often noisy due to degradations, and partially missing due to occlusion by other glyphs from the same block. We conduct two retrieval experiments (without or with block context, see Eq.(4) and Eq.(5)) on this dataset. For each experimental condition, Fig.5 displays the results as the average over all queries of the ranking of the correct match over 892 different sign categories. The results confirms that HOOSC achieves slightly better results than GF-HOG. It also demonstrate that incorporating context through glyph co-occurrence improves the ranking significantly (around 40 on average) for both descriptors. This shows that our approach based on glyph co-occurrence is highly beneficial even if the applied co-occurrence model is rather approximate.

5. DISCUSSION AND FUTURE WORK

This paper presented a new Maya hieroglyph retrieval method using shape and glyph context information, by means of which two glyph datasets were also contributed.

We evaluated two local shape descriptors using the BoVW approach. Experimental results showed that HOOSC achieves higher retrieval accuracy than GF-HOG on both datasets. Future work evaluating these two descriptors will take larger shape datasets of more diverse resources into consideration.

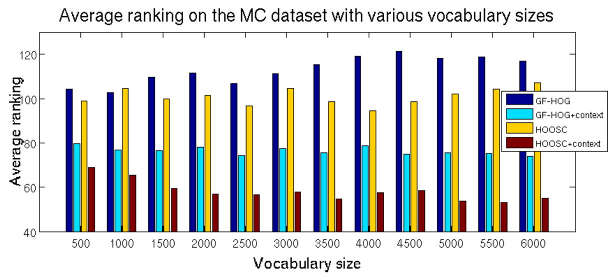


Figure 5: Average ranking using two shape descriptors and context information on the MC dataset.

Although the co-occurrence information we applied here is at an incipient stage, the results suggest that the context within Maya glyph blocks indeed encodes valuable information. In the future, more complete glyph context information will be extracted, which can be used in various Maya hieroglyph analysis tasks, such as assisting the decipherment of signs whose semantic meaning remain problematic, by analysing their context from various resources.

Our ongoing work also includes automatic Maya text detection, block segmentation, as well as glyph detection and segmentation within blocks. This would help both epigraphers and public users for research and education purposes.

6. ACKNOWLEDGMENTS

We thank the Swiss National Science Foundation (SNSF) and the German Research Foundation (DFG) for their support. All images in Fig.1, Table 2 and Fig.3 were produced by C. Pallan and G. Krempel as part of the MAAYA project.

7. REFERENCES

[1] MC dataset. `www.idiap.ch/dataset/maya-codex`.
[2] SLUB library. `http://digital.slub-dresden.de/werkansicht/dlf/2967/1/`.
[3] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *PAMI*, pages 509–522, 2002.
[4] A. Del Bimbo and P. Pala. Visual image retrieval by elastic matching of user sketches. *PAMI*, pages 121–132, 1997.
[5] M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa. A descriptor for large scale image retrieval based on sketched feature lines. In *SBM*, pages 29–36, 2009.
[6] M. Franken and J. van Gemert. Automatic egyptian hieroglyph recognition by retrieving images as texts. In *ACM Multimedia*, pages 765–768, 2013.
[7] R. Hu and J. P. Collomosse. A performance evaluation of gradient field hog descriptor for sketch based image retrieval. *CVIU*, pages 790–806, 2013.
[8] H. Kettunen and C. Helmke. *Introduction to Maya Hieroglyphs - XVII European Maya Conference*. University of Helsinki, 2012.
[9] E. Roman-Rangel, C. Pallan-Gayol, J.-M. Odobez, and D. Gatica-Perez. Retrieving ancient maya glyphs with shape context. In *ICCV workshop on eHeritage and Digital Art Preservation*, pages 988–995, 2009.
[10] E. Roman-Rangel, C. Pallan-Gayol, J.-M. Odobez, and D. Gatica-Perez. Analyzing ancient maya glyph collections with contextual shape descriptors. *IJCV*, pages 101–117, 2011.
[11] E. Roman-Rangel, C. Pallan-Gayol, J.-M. Odobez, and D. Gatica-Perez. Searching the past: an improved shape descriptor to retrieve maya hieroglyphs. In *ACM Multimedia*, pages 163–172, 2011.
[12] J. E. S. Thompson. *A catalog of Maya HIeroglyphs*. University of Oklahoma Press, 1962.