

# Revealing the Mysteries of the Maya Script: Review

Ryan Sharif

December 4, 2015

## 1 Introduction

The Maya language as it was spoken in the 16th Century in the Yucatan Peninsula has been lost for centuries; however, the Maya script endures. How do archaeologists make sense of a system that has been lost for centuries? Ancient writing systems used elaborate and enormous character sets to record phonological, morphological, and sentential information. The task seems impossible. Because of the enormity of the character corpus, researchers working on the translation of these corpora have not been able to completely translate the surviving Mayan literature, despite over two hundred years of effort[4]. Recent advances in machine learning and other statistically based algorithms have allowed researchers to build databases that allow researchers to search and use machine translation make progress in deciphering this writing system.[2] This paper examines the ACM Tech news article [1] and IEEE paper [2] which looks at this process in depth.

## 2 The Mayan script and problems that researchers face

Researchers working on the translation of ancient Mayan writing face several problems: (i) there is a limited amount of digitized data to work with, (ii) photographs and other recorded media are deteriorating, (iii) many of the glyphs that were once in the original pieces have become obscured, and (iv) glyphs artistic style varies depending on the region and time the glyphs were rendered. Although the first problem—available digitized data—may seem the easiest to remedy, researchers face problems

related to storing large RAW images, sharing the data, and cataloging them for easy search and retrieval. Despite these difficulties, the IEEE paper points out that researchers working on the writing system have translated approximately 80% of the writing that survives [3].

Although the first problem—available digitized data—may seem the easiest to remedy, researchers face problems related to storing large RAW images, sharing the data, and cataloging them for easy search and retrieval. The problems related to deterioration, loss of data, and variations in artistic style are difficult problems that may benefit from recent advances in computer vision. Thus, the IEEE paper focuses on the algorithms that have helped to overcome these difficult problems.

## 3 High quality representations of the Mayan script digitized

You need to use different tools for different problems, and the IEEE paper highlights the different approaches the team used. One thing researchers would like to do, especially for problems relating to deterioration, is automatically retrieve the most likely sequence of glyphs given the context of where they appear, i.e., given that we have seen a sequence of glyphs, and given that one may be unreadable which one has the highest probability of occurring here. The team used a dynamic programming algorithm, the Viterbi algorithm, whose results suggest it is well suited for the task. At its core, the algorithm calculates the most likely sequence of events/characters to appear given a training set of data. Because more monumental writings were present in the training set, the algorithm made better predictions on similar monumental data; it performed poorer on data recorded on paper-like media.

Many of the glyphs in the Mayan script look similar, which typically presents a problem for automatic categorization. Moreover, how do we deal with variations in style? As mentioned in the article, the team considered this problem to be separate from contextual prediction. There are times when researchers want a model that best predicts a glyph given no other context, especially when dealing with damaged glyphs. In their experimental

results, the researchers concluded that a model, which used a pivot based method for classifying glyphs. At its core, this method aligns images so that they can be compared using clusters of features, e.g., both glyphs have a particular leg symbol in the same spot. The best model using this method correctly predicted 90% of the experimental data set.

## 4 Discussion

Machine learning and dynamic programming algorithms have yielded solutions to many problems in science. This article shows us that the humanities can benefit from incorporating recent advances to solve similarly difficult problems. Extrapolating the results achieved in the IEEE paper, we can imagine incorporating these techniques in Ancient Egyptian, Mesopotamian, and other ancient writing systems sitting in drawers in museums around the world, invigorating these dead writing systems once more.

## References

- [1] Sarah Bourquenoud. Revealing the mysteries of the Maya script. *News Mediacom*, 2015.
- [2] Rui Hu, Gulcan Can, Carlos Pallan Gayol, Guido Krempel, Jakub Spotak, Gabrielle Vail, Stephane Marchand-Maillet, Jean-Marc Odobez, and Daniel Gatica-Perez. Multimedia Analysis and Access of Ancient Maya Epigraphy. *Signal processing magazine*, 32(4):75–84, 2015.
- [3] Harri J Kettunen and Christophe Helmke. *Introduction to Maya Hieroglyphs: Workshop Handbook*. 2008.
- [4] Martha J. Macri and Gabrielle Vail. *The New Catalog of Maya Hieroglyphs*, volume 2. University of Oklahoma Press, 2009.