



INAOE

Una Nueva Representación de Jeroglíficos a partir de su descomposición en Partes Primitivas

Laura Alejandra Pinilla Buitrago, Jesús Ariel Carrasco Ochoa

Laboratorio de Aprendizaje Computacional y Reconocimiento de Patrones,
Coordinación de Ciencias Computacionales,
Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), México

**Reporte Técnico No. CCC-16-008
Julio de 2016**

©Coordinación de Ciencias Computacionales
INAOE

Luis Enrique Erro 1
Sta. Ma. Tonantzintla,
72840, Puebla, México



Una Nueva Representación de Jeroglíficos a partir de su descomposición en Partes Primitivas

Laura Alejandra Pinilla Buitrago, Jesús Ariel Carrasco Ochoa

Coordinación de Ciencias Computacionales,

Instituto Nacional de Astrofísica, Óptica y Electrónica,

Luis Enrique Erro 1, Sta. Ma. Tonantzintla,

72840, Puebla, México

laurapin@inaoep.mx, ariel@inaoep.mx

Resumen. En los últimos años ha tomado importancia el desarrollo de mecanismos automáticos o semi-automáticos que faciliten el proceso de interpretación de piezas y vestigios arqueológicos. Parte de dichas piezas son inscripciones que contienen jeroglíficos. La interpretación de inscripciones antiguas es una tarea ardua y lenta que comúnmente es llevada a cabo por expertos en el campo, debido a que un mismo jeroglífico puede tener múltiples significados, que pueden variar según el tiempo y la región geográfica en la que fueron pintados. Lo anterior ha influido en que instancias de diferentes categorías sean muy similares, mientras que instancias de una misma categoría tengan alta variabilidad, haciendo del proceso de interpretación una tarea desafiante. En este trabajo de investigación estamos interesados en explorar nuevas formas de representar jeroglíficos que permitan la recuperación de este tipo de imágenes según su contenido visual. Se estudiarán representaciones a partir de diferentes primitivas. En la representación se incluirá información adicional tal como, la distribución espacial de las primitivas en el jeroglífico y las relaciones existentes entre múltiples primitivas, de modo que la representación resultante sea más discriminativa, permitiendo obtener mejores resultados en recuperación de jeroglíficos.

Palabras clave. Bocetos, jeroglíficos, representación visual de imágenes, descriptores, bolsas de palabras visuales.

1. Introducción

Desde tiempos prehistóricos cuando no existía el lenguaje oral y escrito, las representaciones visuales formadas por líneas y trazos (conocidas como *bocetos* (ver *Figura 1.1*)) han permanecido como una forma de comunicación. Los bocetos permiten compartir ideas, transmitir conocimientos, expresar sentimientos o representar aspectos del mundo real. La representación de bocetos es un tema de creciente interés en los últimos años que tiene múltiples aplicaciones, tales como la interpretación, reconocimiento, clasificación o recuperación de planos de arquitectura [1, 2], diagramas de circuitos electrónicos [3, 4], diagramas químicos [5], partituras musicales [6], bocetos faciales [7, 8], objetos en *2D* y *3D* [9], escritura jeroglífica [10, 11], entre otros.

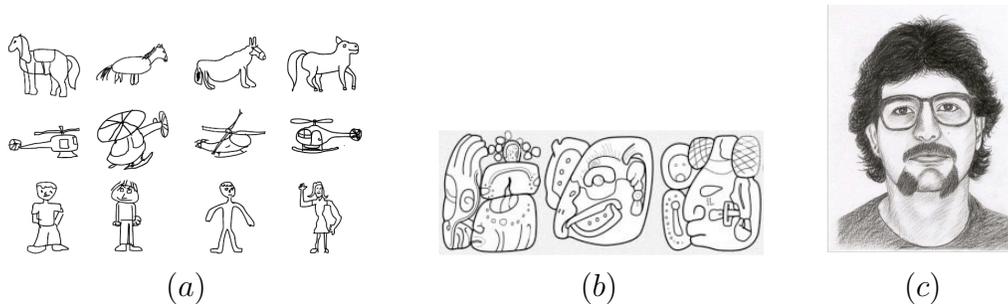


Figura 1.1: (a) bocetos para diferentes categorías [9] (b) Jeroglíficos Mayas [12] (c) boceto de un rostro [13]

Los bocetos pueden estar compuestos por múltiples líneas, las cuales pueden o no intersectarse entre sí. Las líneas tienen diferentes características (*p. ej.* curvatura, grosor, longitud) las cuales dependen del estilo y de la habilidad de la persona que genera el boceto. Un conjunto de bocetos que pertenecen a una misma categoría pueden verse a simple vista muy similares para el ojo humano, sin embargo, si son comparados mediante transformaciones geométricas (traslación, rotación, escalamiento, etc.) y deformaciones no lineales (ruido, articulaciones, oclusión, etc.) pueden ser muy diferentes entre sí [14]. Lo mencionado anteriormente hace de la representación de bocetos un problema desafiante y que despierta el interés de la comunidad científica.

En la literatura, los bocetos son representados atendiendo a alguno de los siguientes tres enfoques: (1) representación mediante descriptores [15, 16], (2) representación mediante

estructuras tipo árbol/grafó [17] y (3) representación basada en el modelo BoV (del inglés *Bag of Visual Words*) [18, 19, 20].

El primer enfoque consiste en la abstracción y codificación de características en un único descriptor (*descriptor global*) o en múltiples descriptores (*descriptor local*). Los descriptores pueden ser clasificados atendiendo a si son aplicados a contornos o regiones, o según su invariancia a diferentes transformaciones geométricas (rotación, escala traslación, reflexión, etc.). Usualmente, un descriptor es desarrollado según las características del conjunto de imágenes en el dominio de aplicación. Entre los descriptores más populares en la literatura aplicados a bocetos se pueden mencionar *SC* (del inglés *Shape Context*) [21] y *HOOSC* (del inglés *Histogram of Orientation Shape Context*) [16]. Otros descriptores populares como *HOG* (del inglés *Histogram of Orientated Gradient*) [22] y *SIFT* (del inglés *Scale Invariant Feature Transform*) también han probado su efectividad en la representación de bocetos, sin embargo, debido a que fueron diseñados para trabajar con imágenes de intensidad (color y textura), no proveen una representación adecuada en imágenes de bocetos [23].

El segundo enfoque para la representación de bocetos consiste en construir estructuras tales como árboles o grafos [24, 25, 26, 27] a partir de la identificación de patrones, trazos básicos o primitivas (*p. eg. líneas, curvas, elipses, etc.*) y sus relaciones (*p. eg. intersección, tangencia, adyacencia, distancia, etc.*); lo anterior da lugar a los nodos y ramas en el árbol/grafó. Este enfoque ha sido utilizado principalmente para representar bocetos de contorno cerrado simple [24] y símbolos [27] (ver *Figura 1.2*).

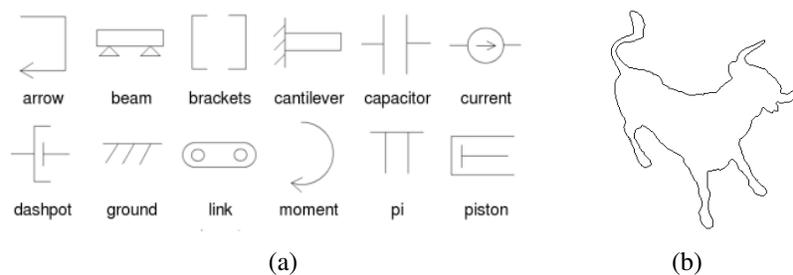


Figura 1.2: (a) Símbolos [27] (b) Imagen de contorno cerrado simple [24]

Una de las principales desventajas que comparten los dos enfoques mencionados anteriormente para el análisis de bocetos radica en que la similitud entre bocetos se realiza a través de una comparación uno a uno, lo que implica que se requieran muchas instancias

de entrenamiento para capturar las diferentes variaciones tanto locales como globales que se pueden presentar al interior de cada categoría. Adicionalmente, un mayor número de instancias en el conjunto de entrenamiento implica un costo computacional mayor. Para solventar este inconveniente, un tercer enfoque denominado bolsa de palabras visuales o bolsa de términos visuales *BoV* (del inglés *Bag of Visual terms*) [28] ha tenido una creciente aceptación en tareas de clasificación y recuperación de imágenes. En esta técnica, los descriptores son tratados como palabras para la generación de un diccionario de términos visuales y la ocurrencia de los descriptores en cada imagen es almacenada en un histograma de frecuencias, el cual constituye la representación de la imagen.

Investigaciones recientes en el análisis de bocetos, aplicadas a dominios específicos, requieren de una colaboración multidisciplinaria. Por ejemplo, el análisis de jeroglíficos en inscripciones antiguas [10, 23, 29], donde áreas como visión por computadora, reconocimiento de patrones y arqueología colaboran conjuntamente. Para el análisis de este tipo de bocetos se debe considerar la variabilidad en la forma, la cual está relacionada con el periodo histórico y la región geográfica en el que los bocetos fueron pintados, dando lugar a que partes de las inscripciones se presenten en posiciones arbitrarias y en distintos tamaños; además, las diferentes variaciones locales y globales provocan que instancias de diferentes categorías sean muy similares, mientras que instancias de una misma categoría tengan alta variabilidad (ver *Figura 1.3a*). Finalmente, es importante mencionar que años de abandono han provocado, en algunos casos, la pérdida parcial de estas inscripciones (ver *Figura 1.3b*). Las características mencionadas anteriormente convierten al análisis de jeroglíficos de inscripciones antiguas en un problema desafiante.

Los enfoques para la representación de bocetos mencionados anteriormente han mostrado ser prácticos y útiles en diferentes dominios de aplicación. Sin embargo, trabajos recientes [10, 18, 30] han mostrado evidencia de que los métodos de clasificación, identificación y recuperación de bocetos que contienen jeroglíficos, no obtienen resultados satisfactorios debido a la complejidad de este tipo de imágenes. Por esta razón, este trabajo de investigación estará enfocado en proponer una representación de jeroglíficos que permita distinguir entre instancias de diferentes categorías que presentan características similares e instancias de la misma categoría que presentan diferentes variaciones. La representación desarrollada debe

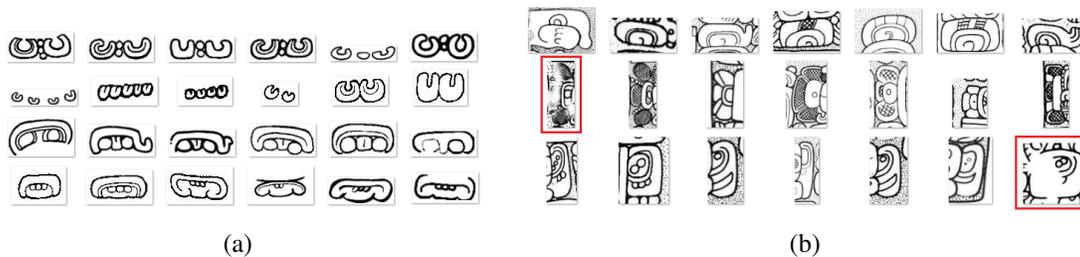


Figura 1.3: Jeroglíficos Mayas: Cada fila contiene instancias que pertenecen a una misma categoría, se puede apreciar las variaciones locales y globales entre ejemplos de una misma categoría. En (a) las instancias en las filas 1 – 2 y 3 – 4 tienen características similares, lo que las hace difíciles de distinguir. En (b) los jeroglíficos marcados en recuadro, corresponden a instancias con pérdida parcial de información visual.

permitir recuperar bocetos de jeroglíficos con una precisión mayor a los trabajos reportados en la literatura.

Esta propuesta de investigación se encuentra organizada de la siguiente forma. En la *Sección 2* se revisa el trabajo relacionado con este trabajo de investigación. En la *Sección 3* se describe la motivación y justificación de este trabajo de investigación. En la *Sección 4* se presenta el problema de investigación. La propuesta de investigación conformada por la hipótesis, preguntas de investigación, objetivos, metodología y cronograma de actividades, es presentada en la *Sección 5*. Por último, en la *Sección 6* presentamos el trabajo realizado y los resultados alcanzados hasta el momento.

2. Trabajo Relacionado

Como ya se mencionó, en la literatura los bocetos son representados siguiendo alguno de los siguientes enfoques: (1) representación a través de descriptores, (2) representación mediante estructuras tipo árbol/grafó y (3) representación basada en el modelo BoV. En la *Figura 2.1* aparecen algunos trabajos que pertenecen a cada uno de los enfoques.

Dado que nuestro interés es proponer una nueva representación para un tipo específico de boceto: los jeroglíficos; en la siguiente sección se presentan los trabajos directamente relacionados con este problema.

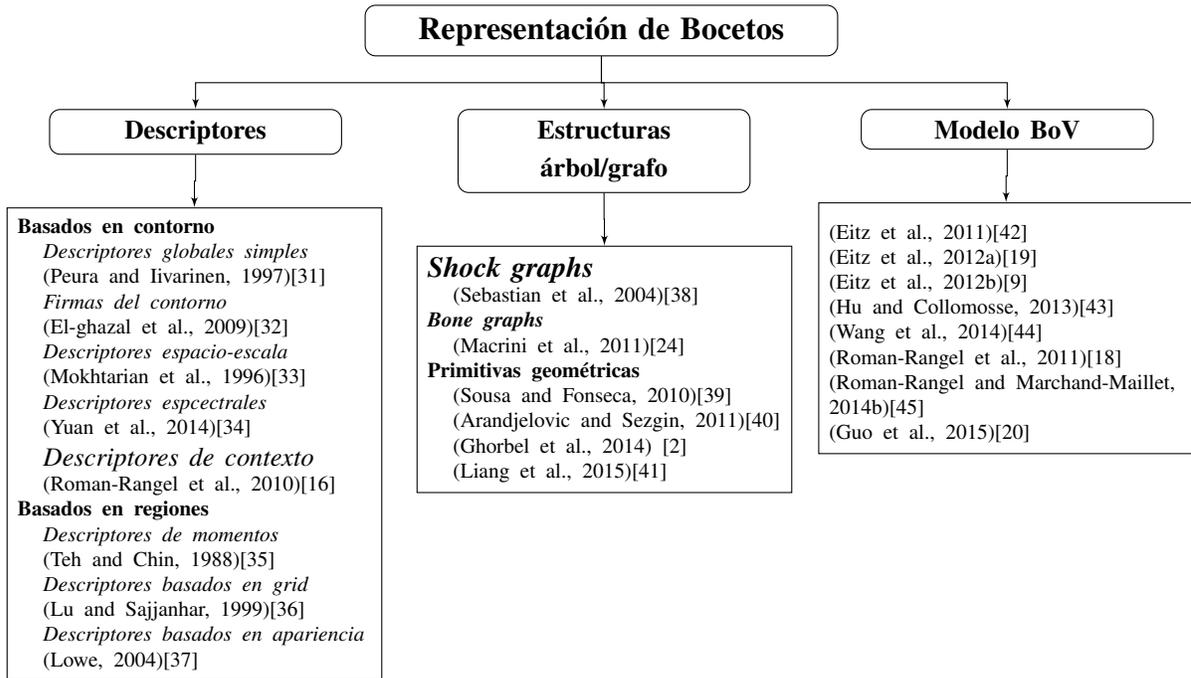


Figura 2.1: Representación de bocetos

2.1. Representación de Jeroglíficos en inscripciones antiguas

En [46] se presenta un método para el reconocimiento de símbolos (o *signos*) mesoamericanos, específicamente un conjunto de símbolos denotado como ‘Uno’ (ver *Figura 2.1*); los cuales corresponden a los símbolos más básicos de los cuatro tipos identificados en varios monumentos desarrollados en diferentes culturas Mesoamericanas (Olmeca, Maya, Teotihuacan, Azteca, etc.) [47]. El método en [46] representa a los símbolos a través de un subconjunto de píxeles que pertenecen al contorno de la imagen, los cuales son usados junto con métricas de distancia que permiten evaluar si cumplen con un conjunto de reglas que permiten determinar si el símbolo pertenece a la categoría denominada ‘Uno’; entre otras, las reglas evalúan la presencia de un eje de simetría en la forma y la morfología de los segmentos de línea/curvas a ambos lados del eje de simetría. El sistema únicamente logra reconocer símbolos completos (no permite oclusión) (ver *Figura 2.1*). El método es apropiado para discriminar entre símbolos caracterizados o no como ‘Uno’.

En [12] el descriptor SC (del inglés *Shape Context*) [15] es usado para representar jeroglíficos Mayas. Dado un jeroglífico Maya y sea M el número de píxeles que conforman el

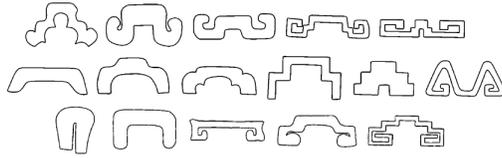


Figura 2.2: Símbolos ‘Uno’ provenientes de distintas regiones culturales y períodos de Mesoamérica. (Tomado de [46])

contorno de la imagen, el descriptor SC es calculado desde un número s de píxeles obtenidos aleatoriamente a partir de M ($s < M$), representando la imagen a través de s descriptores. Para cada píxel p_i ($i = 1, 2, \dots, s$) se obtiene un descriptor local SC_i , en función de la orientación y la proximidad de los $s - 1$ píxeles restantes; para ello, se construye un espacio de coordenadas en escala logarítmica-polar con origen en p_i , dividido en 12 regiones angulares y 5 intervalos de distancia (la distancia es normalizada), dando lugar a 60 regiones (r_j , $j = 1, 2, \dots, 60$). Así, para cada p_i se obtiene un descriptor $SC_i \in \mathbf{R}^{60}$, el cual corresponde a un histograma que considera el número de puntos en cada región (ver Figura 2.3).

La distancia o similitud entre dos imágenes \mathcal{P} y \mathcal{Q} es calculada en función de las correspondencias (punto-a-punto) entre los descriptores obtenidos a partir de los puntos en ambas imágenes.

Aunque el descriptor SC ha sido usado exitosamente en múltiples aplicaciones [48, 49], la complejidad y la alta variedad de contenido de los jeroglíficos impide que la representación resultante sea adecuada para representar a este tipo de imágenes, lo que da lugar a que muchas de las instancias no puedan ser recuperadas exitosamente.

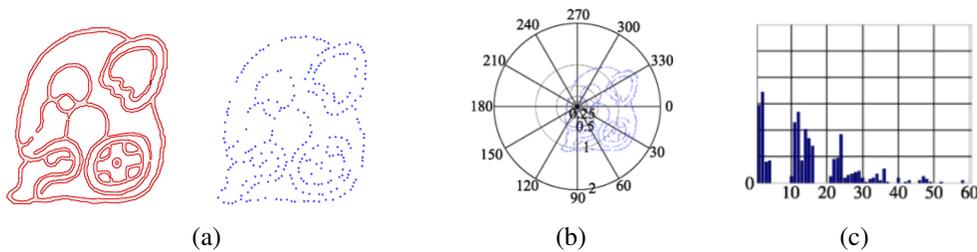


Figura 2.3: a) Jeroglífico Maya (contorno marcado en rojo) y puntos pivote obtenidos a partir del contorno. b) Espacio de coordenadas polares con origen en un punto pivote p_i . c) histograma SC_i para el punto p_i .

En [16], los descriptores SC [21], GSC (del inglés *Generalized Shape Context*) [48] y el

HOOSC (del inglés *Histogram of Orientation Shape Context*) son evaluados en recuperación de jeroglíficos. A diferencia del descriptor SC, el descriptor GSC no cuenta el número de píxeles en cada región, sino que cada región es descrita por un vector 2-D obtenido a partir de la suma de los vectores tangentes de los píxeles en cada región, dando lugar a un descriptor GSC_i de dimensión 120 para cada p_i . Sin embargo, si en cada región obtenemos vectores tangentes cuya dirección es opuesta, estos vectores se anulan dando lugar a estimaciones poco fiables. En [16] se propone el descriptor denominado HOOSC, el cual es una combinación de los descriptores SC y HOG (del inglés *Histogram of Oriented Gradients*) [22]. HOOSC al igual que SC calcula un descriptor para cada p_i considerando la distribución de los $s - 1$ píxeles en el diagrama logarítmico-polar; adicionalmente cada región r_j es descrita por un histograma de orientaciones locales de los píxeles en cada región, donde el histograma es representado en g intervalos ($g = 8$), dando lugar a un descriptor $HOOSC_i$ de dimensión 480 para cada p_i (ver *Figura 2.4*). Debido a que los jeroglíficos tienen diferentes grados de complejidad y por lo tanto diferente número de píxeles en el contorno, calcular la similitud en función de las correspondencias punto-a-punto puede llegar a ser computacionalmente costoso y poco eficiente a medida que la cantidad de imágenes es mayor; por lo tanto, en [16] los descriptores fueron combinados con el modelo BoV; este enfoque ha mostrado ser eficiente (en términos de representación y costo computacional) en recuperación de imágenes [48, 50]. Los mejores resultados en recuperación de jeroglíficos en términos de precisión y recuerdo, fueron obtenidos haciendo uso del descriptor HOOSC.

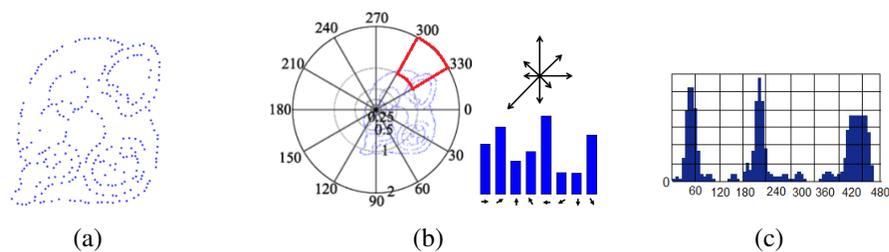


Figura 2.4: (a) Puntos pivote obtenidos, (b) Histograma de orientaciones en la región r_j (marcada en rojo), (c) Histograma $HOOSC_i$ en el punto p_i

En [12, 16] los descriptores son calculados a partir de s píxeles pivote obtenidos desde el contorno de los jeroglíficos (ver *Figura 2.3a*) y la descripción se realiza en función de los $s - 1$ píxeles restantes. En [18, 51] varias mejoras son incorporadas al descriptor HOOSC

(denotado como HOOSC4): (1) Adelgazamiento del contorno de la forma, (2) selección aleatoria de pixeles con distribución uniforme, (3) descripción de cada p_i en función de los M pixeles en la imagen adelgazada, (4) cambio en la distribución logarítmico-polar, sólo son considerados los intervalos de distancia 2, 3 y 4, y (5) incorporación de la posición (x_i, y_i) (en un intervalo $(0, 1]$) del pixel p_i en el descriptor, dando lugar a un descriptor de dimension $288+2$ para cada p_i ; lo anterior evita que descriptores obtenidos desde contornos den lugar a descripciones ruidosas que aumentan la variabilidad al interior de la categoría. Por otro lado, la dimensión del descriptor es disminuida al enfocarse en las regiones que contienen mayor información. Esto último surge a partir de la observación de que muchas de las regiones en el primer y último intervalo de distancia presentan poco contenido de información, inclusive algunas de las regiones (principalmente en el último intervalo) no contienen pixeles (ver *Figura 2.4*). Por otro lado, la información espacial de cada pixel es capturada al agregar su posición relativa en la imagen. Es necesario resaltar que los jeroglíficos pueden tener diferente significado bajo transformaciones geométricas tales como reflexión o rotación.

En [23] la detección de jeroglíficos en inscripciones es evaluada haciendo uso de los descriptores SIFT (del inglés *Scale-Invariant Feature Transform*)[37] y HOOSC junto con los métodos de detección de puntos de interés DoG (del inglés *Difference of Gaussians*)[37] y Harris-Laplace [52]; estas combinaciones son aplicadas tanto en contornos gruesos como en contornos adelgazados de la forma. La mejor representación es la obtenida de la combinación *Harris-Laplace+HOOSC*. DoG y Harris-Laplace fueron diseñados para trabajar sobre imágenes en escala de grises; su aplicación en imágenes binarias como los jeroglíficos da lugar a un escaso conjunto de puntos con alta inconsistencia visual, por lo que descriptores calculados a partir del conjunto de puntos de interés no son suficientes para generar una representación adecuada de jeroglíficos.

En [53, 54] siguiendo ideas usadas en la representación de documentos de texto se propone un método para la detección de *stopwords* en documentos visuales (imágenes) a partir del cálculo de la entropía para cada una de las palabras visuales en el vocabulario, eliminando las palabras visuales con mayor entropía. Este método permite reducir la dimension de la representación, manteniendo el mismo o mayor (en algunos casos) poder discriminativo en la recuperación de jeroglíficos Mayas. En [54] el número de términos visuales a eliminar es

calculado de forma automática.

En [55] se explora una configuración diferente en el diagrama logarítmico-polar a partir del descriptor HOOSC4, unificando los tres primeros intervalos de distancia, y conservando el cuarto. En [55] se utiliza un método de codificación *sparse* denominado k -SVD [56] el cual encuentra una combinación lineal de palabras visuales (funciones de base) para codificar los descriptores HOOSC de las imágenes, la representación es obtenida haciendo uso de diferentes estrategias de *pooling*: AVP (*Average Pooling*), Max-NWP (*Max-N Weight Pooling*), Max-NBP (*Max-N Binary Pooling*) y Max-NIP (*Max-N Integer Pooling*). Este método es comparado con el método de cuantización tradicional VQ (del inglés *Vector Quantization*)[57]. La representación resultante usando codificación *sparse* obtienen resultados competitivos en recuperación de jeroglíficos Maya respecto a los obtenidos con el método de cuantización VQ.

En [58] se propone un método para la representación de jeroglíficos Mayas mediante bolsas de frases visuales BoP (del inglés *Bag of Visual Phrases*). Sea D el vocabulario obtenido a partir de un conjunto de imágenes y w_i cada uno de los términos visuales en el vocabulario, la representación de un jeroglífico es obtenida al considerar la probabilidad de observar pares de términos visuales (w_j, w_k) en una vecindad, incorporando de esta forma la distribución espacial-local de los términos visuales en la imagen (ver *Figura 2.5*). El método fue probado en recuperación de jeroglíficos Mayas usando diferentes formas de representación mediante términos visuales: BoV [50], SPM [59], n-Gramas [60] y la representación BoP propuesta. Los resultados dan evidencia de que la representación BoP permite obtener mejores resultados en recuperación de jeroglíficos.

En [10] se presenta un sistema de recuperación de jeroglíficos en inscripciones Egipcias (ver *Figura 2.6*). Los jeroglíficos son extraídos desde una inscripción mediante un algoritmo de detección de textos basado en partes salientes [61]. La representación de jeroglíficos es obtenida siguiendo tres estrategias: 1) calcular un único descriptor para todo el jeroglífico, (2) calcular múltiples descriptores a partir de puntos de interés, y 3) descriptores a partir de puntos de interés junto al modelo BoV. La similitud entre las imágenes utilizando la representación (1) y (3) se obtiene al calcular la distancia entre los vectores que representan las características de las jeroglíficos, mientras que la representación (2) calcula la similitud a

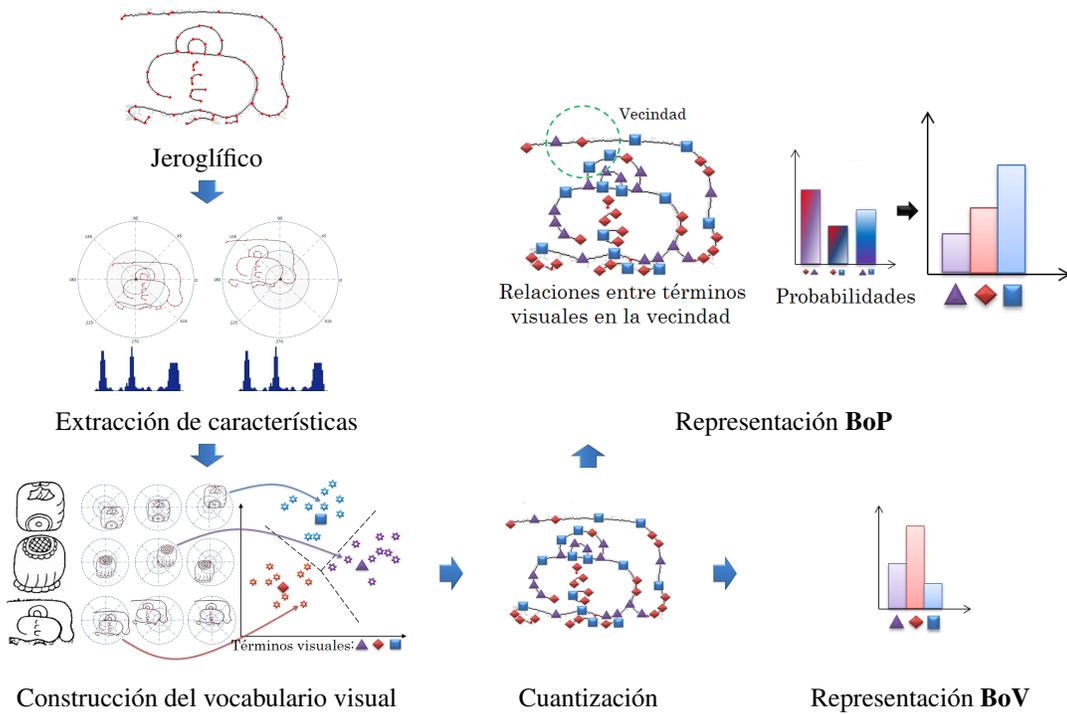


Figura 2.5: Esquema para representación de jeroglíficos Mayas a través de bolsas de términos visuales *BoV* o bolsas de frases visuales *BoP*

partir de las correspondencias punto-a-punto entre los descriptores. Los descriptores utilizados para la representación son: *SC*, *SS* (del inglés *Self-Similarities*) [62], *HOG*, *HOOSC* y *HOOSS*; este último es propuesto por los autores y es una combinación de los descriptores *HOG* y *SS*. Los resultados muestran que los descriptores *HOG* y *HOOSC* calculando la similitud entre jeroglíficos a partir de las correspondencias punto-a-punto obtienen los mejores resultados para la recuperación de jeroglíficos Egipcios.

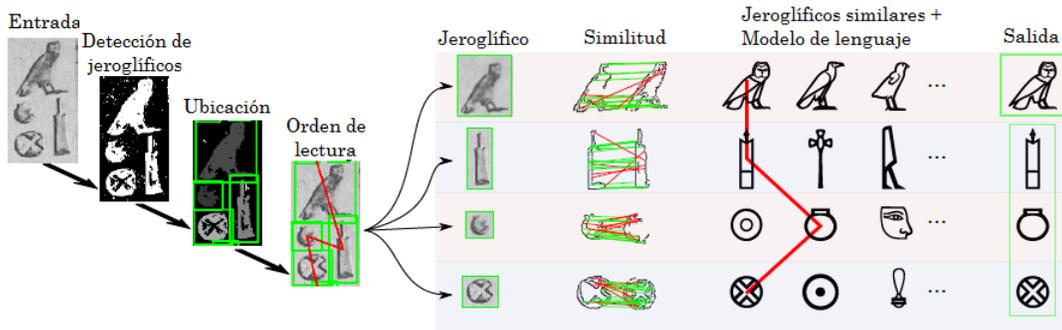


Figura 2.6: Reconocimiento de jeroglíficos Egipcios (Tomado de [10])

2.2. Discusión

Los trabajos descritos en la sección anterior presentan diferentes métodos de representación de jeroglíficos en inscripciones antiguas. Las principales características de estos trabajos están resumidas en la *Tabla 2.1*.

Analizando los trabajos reportados en la *Tabla 2.1* y considerando sus características, podemos resaltar los siguiente:

- El método en [46] no está diseñado para trabajar con símbolos incompletos ni ruido, éstas son características comunes en imágenes tomadas desde inscripciones antiguas. Además, las reglas definidas solo permiten identificar un conjunto único de símbolos denotados como ‘*Uno*’, cuyas características difieren de otros tipos de jeroglíficos más complejos como los Mayas y los Egipcios.
- Descriptores de contexto como SC y HOOSC han sido muy utilizados para representar formas complejas como los jeroglíficos en inscripciones antiguas [16, 23, 55]. Sin embargo, este tipo de descriptores son calculados desde un conjunto denso de puntos, los cuales son obtenidos aleatoriamente. La selección aleatoria de los puntos no garantiza que los descriptores sean capturados desde puntos que representen información importante de la imagen.
- Descriptores como SC y HOOSC han sido diseñados para codificar las características de la imagen desde diferentes perspectivas (puntos), lo que implica que la información que varía y la que permanece invariante alrededor de cada punto sea codificada en un mismo vector. Lo anterior introduce ruido a la descripción cuando se presentan diferentes variaciones alrededor de puntos correspondientes entre imágenes que pertenecen a la misma categoría [45].
- Aunque los métodos presentados en [10, 12] consideran en cierta forma la información espacial de los puntos a partir de los cuales se calculan los descriptores, ya que la similitud es calculada en función de la mejor correspondencia punto-a-punto, es necesario recordar que este enfoque es poco práctico cuando el número de puntos que describen a la imagen y el número de imágenes se incrementa. Por otra parte, a pesar de que

Tabla 2.1: Características de los métodos de representación de jeroglíficos

Método	Primitiva	Jeroglíficos	Representación	Información espacial de las primitivas	Información relación entre primitivas
Frauel et al. (2006b)[46]	Segmentos de línea	Mesoamericanos	descriptores	Si	No
Roman-Rangel et al. (2009)[12]	Puntos	Mayas	descriptores	Si	No
Roman-Rangel et al. (2011)[18]	Puntos	Mayas	BoV	Si	No
Franken and van Gemert (2013)[10]	Puntos	Egipcios	descriptores BoV	Si	No
Roman-Rangel and Marchand-Maillet (2013a)[58]	Puntos	Mayas	BoP	No	Si
Roman-Rangel and Marchand-Maillet (2014a)[54]	Puntos	Mayas	BoV	No	No
Roman-Rangel and Maillot (2015)[30]	Puntos	Mayas	BoV	No	No
Método Propuesto	Puntos, regiones y/o segmentos de línea	Mayas, Egipcios	BoV	Si	Si

el método en [18] agrega la posición relativa de cada punto (x_i, y_i) en la descripción, esta estrategia presenta inconvenientes, puesto que pueden existir puntos ubicados en posiciones distintas, cuya información en la vecindad es similar. Lo anterior implica que dos descriptores que deben considerarse como similares puedan verse aún más diferentes entre sí.

- La mayoría de los métodos que permiten representar formas complejas como los jeroglíficos Mayas y jeroglíficos Egipcios, son desarrollados a partir de los puntos como primitivas.

Considerando lo ya mencionado y lo presentado en la *Tabla 2.1*, en esta propuesta de investigación doctoral estamos interesados en generar una representación de jeroglíficos a partir de diferentes tipos de primitivas (*p. ej.* puntos, segmentos de línea o regiones), incorporando en la representación, información extraída a partir de relaciones (*p. ej.* adyacencia, distancia o inclusión) entre múltiples partes primitivas y su distribución espacial en la imagen, de modo que sea posible reflejar en la representación aquellas características que permiten discriminar jeroglíficos que pertenecen a diferentes categorías, pese a su similitud parcial.

3. Motivación y Justificación

Los jeroglíficos en inscripciones antiguas constituyen una de las estructuras fundamentales en el sistema de escritura antiguo. Los jeroglíficos son de gran interés para profesionales en artes y humanidades (*p. ej.* historiadores, arqueólogos, antropólogos, curadores, fotógrafos, entre otros), quienes están interesados en leerlos, interpretarlos y descifrarlos; de tal manera que puedan ser clasificados y etiquetados según su significado, ubicación geográfica y/o periodo histórico, obteniendo a partir de ellos la mayor cantidad de información posible [63]. Sin embargo, muchas de las inscripciones encontradas aún no han sido descifradas, obligando a los profesionales a estudiar y buscar de forma manual, en los catálogos existentes, información que les facilite su interpretación [18].

Comúnmente, la interpretación de inscripciones antiguas requiere que componentes básicos individuales (jeroglíficos) sean identificados, sin embargo, los jeroglíficos son a menudo

formas complejas que presentan diferentes grados de variabilidad, influenciados por la evolución en el estilo de escritura, el cual está relacionado con el periodo histórico y la región geográfica en la que fueron pintados. Es común encontrar jeroglíficos con diferentes variaciones que tiene el mismo significado, mientras que existen jeroglíficos con diferente significado que son muy similares entre sí [63]. Esto hace del proceso de identificación e interpretación, una tarea desafiante.

Por lo tanto, es necesario proveer herramientas automáticas o semi-automáticas que faciliten el proceso de recuperación manual. Entre los métodos de representación de jeroglíficos, descriptores de contexto como SC [15] y HOOSC [16] permiten hacer una buena representación de jeroglíficos [16, 23]. Sin embargo, la manera en que las características de la imagen son codificadas da lugar a que la representación resultante llegue a ser muy diferente cuando se presentan variaciones locales y globales (similitud parcial), entre imágenes que pertenecen a la misma categoría.

Los trabajos existentes para la representación de jeroglíficos tienen un alcance limitado considerando la complejidad de las imágenes. Por lo tanto, es necesario proponer una nueva representación de jeroglíficos, la cual logre capturar y describir características mejor que los métodos reportados en la literatura.

4. Problema de Investigación

Como ya se mencionó, los jeroglíficos son formas complejas que pueden presentar diferentes variaciones entre instancias que pertenecen a una misma categoría y similitud parcial entre instancias que pertenecen a diferentes categorías. Por lo tanto, el problema de investigación consiste en proponer una nueva representación para jeroglíficos de inscripciones antiguas a partir de su descomposición en partes primitivas, de modo que la información y/o la relación de una o múltiples partes primitivas den lugar a características discriminativas para la representación de este tipo de imágenes, obteniendo de esta manera mejores resultados en recuperación de jeroglíficos, respecto a los reportados en el estado del arte.

5. Propuesta de Investigación

5.1. Hipótesis

Los jeroglíficos son comúnmente representados a partir de puntos como primitivas, sin embargo la calidad de la representación depende en gran medida del tipo de información que provee el conjunto de puntos utilizados para describir la imagen, de la manera en que se realiza la descripción y de cómo se organiza dicha información para la representación del jeroglífico. Por otro lado, los puntos no son el único tipo de primitiva que puede ser obtenido a partir de los jeroglíficos, un jeroglífico puede ser descompuesto en otro tipo de primitivas tales como segmentos de línea o regiones. Con todo esto, nuestra hipótesis de investigación es la siguiente:

Es posible obtener una representación para jeroglíficos en inscripciones antiguas a partir de características obtenidas desde y entre partes primitivas (puntos, segmentos de línea o regiones), de modo que dicha representación permita distinguir entre jeroglíficos similares de distinta categoría, así como variaciones entre jeroglíficos de la misma categoría; de manera que se puedan obtener mejores resultados en recuperación de jeroglíficos respecto a los métodos reportados en la literatura.

5.2. Preguntas de investigación

Considerando el problema de investigación en la *Sección 4* así como nuestra hipótesis de investigación, surge la siguiente pregunta de investigación:

Q_1 ¿Cómo construir una representación de jeroglíficos, la cual capture información discriminativa, de modo que se puedan obtener mejores resultados de recuperación de jeroglíficos respecto a los reportados en la literatura?

A partir de esta pregunta de investigación y considerando las debilidades de los trabajos relacionados con la representación de jeroglíficos presentados en la *Sección 2*, planteamos las siguientes preguntas de investigación:

Q_2 ¿Cómo descomponer un jeroglífico en partes primitivas de modo que éstos representen

información importante de la imagen, la cual sea útil para la representación de jeroglíficos?

Q₃ ¿Cómo abstraer información a partir de una o múltiples partes primitivas de modo que se capturen características importantes de los jeroglíficos?

Q₄ ¿Cómo se puede organizar la información obtenida desde partes primitivas, generando una representación que permita obtener mejores resultados en recuperación de jeroglíficos, respecto a trabajos reportados en la literatura?

Considerando estas preguntas, los objetivos de investigación doctoral son los siguientes.

5.3. Objetivo general

Desarrollar una representación de jeroglíficos de inscripciones antiguas a partir de su descomposición en partes primitivas, de modo que la representación permita obtener resultados estadísticamente mejores, en recuperación de jeroglíficos, respecto a los resultados reportados en el estado del arte.

5.4. Objetivos específicos

- Proponer un método para descomponer jeroglíficos en partes primitivas, las cuales obtengan información de la imagen útil para la representación de jeroglíficos.
- Proponer un método para obtener información desde una o múltiples partes primitivas, obtenidas a partir de la descomposición del jeroglífico.
- Proponer un método para organizar la información desde y entre partes primitivas, dando lugar a una representación de jeroglíficos en inscripciones antiguas.
- Utilizar un esquema de recuperación de jeroglíficos el cual haga uso de la nueva representación propuesta. De modo que los resultados obtenidos sean estadísticamente mejores que los reportados en el estado del arte.

5.5. Contribuciones esperadas

- Un método para descomponer jeroglíficos en partes primitivas que contengan información importante de la imagen, la cual sea útil para la representación.
- Un método para extraer información desde una o múltiples partes primitivas, que incorpore información espacial de las primitivas así como las relaciones entre ellas.
- Un método para generar una representación de los jeroglíficos a partir de la información extraída desde partes primitivas en la imagen.
- Un esquema de recuperación de jeroglíficos el cual haga uso de la nueva representación propuesta.

5.6. Metodología

Para alcanzar los objetivos y dar respuesta a las preguntas de investigación planteadas en esta propuesta doctoral se seguirá la siguiente metodología:

1. Estudio crítico de trabajos presentados en la literatura sobre representación de jeroglíficos
 - a) Identificar las características (ventajas y desventajas) de los trabajos reportados en la literatura relacionados con la representación de bocetos, con el objetivo de:
 - Identificar qué tipo de información es extraída y cómo ésta es utilizada y/o organizada para la representación de jeroglíficos.
 - Analizar si la forma en que la información está siendo extraída y organizada para la representación de jeroglíficos es apropiada según las características de este tipo de imágenes.
 - Identificar qué tipo de información adicional puede ser extraída e incluida para generar una mejor representación de jeroglíficos
2. Para proponer un método para descomponer jeroglíficos en partes primitivas, dado que el contorno de la imagen muchas veces es procesado de modo que se simplifique la

información de la imagen. Primero exploraremos diferentes variantes para la simplificación del contorno de una imagen:

- Adelgazar la forma del jeroglífico a partir de métodos que llevan a cabo operaciones morfológicas [64].
- Podar formas adelgazadas de jeroglíficos, con el objetivo de eliminar ruido. Esto puede ser llevado a cabo midiendo la importancia de cada rama en la representación. Para esto consideraremos diferentes formas para medir la importancia de las ramas, por ejemplo:
 - Longitud de la rama.
 - Información reconstruida a partir de la rama.

3. Para descomponer jeroglíficos en partes primitivas, exploraremos tres variantes:

a) Obtener puntos del jeroglífico como partes primitivas:

- 1) A diferencia de los trabajos reportados en la literatura que calculan los puntos de manera aleatoria, calcularemos de forma determinista un conjunto de puntos de interés. Para esto se explorará el uso de:
 - Puntos dominantes a partir de curvas complejas.
 - Puntos invariantes a diferentes escalas.
 - Puntos críticos.

b) Obtener segmentos de línea a partir del jeroglífico, como partes primitivas, para lo cual probaremos:

- 1) Obtener puntos que delimiten los segmentos de línea.
 - Puntos dominantes a partir de curvas complejas.
 - Puntos de unión y puntos finales en jeroglíficos adelgazados.
- 2) Encontrar una estrategia para obtener segmentos de línea que capturen información importante del jeroglífico.
 - Segmentos de línea que corresponden a ramas finales, ramas simples y ramas de unión en la forma adelgazada del jeroglífico.

- b) Para proponer un método que permita extraer información desde y entre segmentos de línea y/o regiones como partes primitivas, exploraremos:
- 1) Calcular descriptores a partir de segmentos de línea y/o regiones excluyendo la información que se encuentra alrededor de cada primitiva, utilizando descriptores tales como SC, HOOSC, HOG.
 - 2) Determinar una forma para obtener una posición que represente a la primitiva en la imagen, por ejemplo:
 - Centroide.
 - Punto medio del segmento de línea.
 - 3) Encontrar relaciones entre múltiples partes primitivas, por ejemplo:
 - Identificar relaciones tales como adyacencia o inclusión.
 - Identificar relaciones entre las primitivas con mayor o menor frecuencia de ocurrencia en la imagen.
5. Proponer una manera para combinar la información desde una o múltiples partes primitivas en una representación de jeroglíficos.
6. Evaluación experimental en recuperación de jeroglíficos, haciendo uso de cada una de las representaciones obtenidas.
- a) Se utilizará un *framework* para recuperación de jeroglíficos.
 - b) Los resultados obtenidos en recuperación de jeroglíficos se compararán con métodos del estado del arte en términos de precisión y recuerdo.

5.7. Cronograma de Actividades

Para dar cumplimiento a los objetivos planteados siguiendo la metodología propuesta seguiremos el cronograma de actividades presentado en la *Tabla 5.1*. En esta tabla, cada celda coloreada representa el tiempo que tomará cada actividad. Las celdas de color oscuro representan las actividades que ya han sido llevadas a cabo, mientras que las celdas de color claro

representan a las actividades pendientes. Las actividades están especificadas en la primera columna. Las demás columnas especifican los años de duración de este trabajo de investigación, divididos en bimestres.

Tabla 5.1: Cronograma de Actividades

Actividad	2015						2016						2017						2018					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Estudio critico del estado del arte																								
Elaboración de la propuesta doctoral																								
Defensa de la propuesta Doctoral																								
Diseñar un algoritmo para generar una representación de jeroglíficos haciendo uso de puntos como primitivas																								
Evaluar y comparar el algoritmo propuesto																								
Diseñar un algoritmo para generar una representación de jeroglíficos haciendo uso de segmentos de línea como primitivas																								
Evaluar y comparar el algoritmo propuesto																								
Diseñar un algoritmo para generar una representación de jeroglíficos haciendo uso de regiones como primitivas																								
Evaluar y comparar el algoritmo propuesto																								
Diseñar un algoritmo para generar una representación de jeroglíficos que incluya información de y entre múltiples partes primitivas																								
Evaluar y comparar el algoritmo propuesto																								
Redacción de artículos para revistas y conferencias																								
Escritura de la tesis Doctoral																								
Revisión de la tesis por el comité Doctoral																								
Realizar las correcciones de la tesis																								
Defensa final																								

6. Resultados Preliminares

En esta sección se presentan los resultados preliminares obtenidos siguiendo la metodología descrita en la *Sección 5.6*, dando cumplimiento parcial de este modo al primer y segundo objetivo de investigación, *i.e.* proponer un método para descomponer jeroglíficos en partes primitivas (puntos en los resultados preliminares que presentaremos) y la codificación de características e información importante (información espacial) desde estas partes primitivas, para generar una representación de jeroglíficos en inscripciones antiguas.

En la *Sección 6.1* presentamos un método de representación de jeroglíficos a partir de descriptores calculados desde puntos de la imagen como partes primitivas, en combinación con el modelo BoV. Los puntos son obtenidos mediante métodos de selección de puntos dominantes a partir del contorno adelgazado de la imagen, evitando de este modo la selección aleatoria de puntos. Los puntos dominantes son puntos de alta curvatura, los cuales conservan las principales características de la forma, preservando de este modo información importante para la percepción visual de la imagen [65].

6.1. Representación de jeroglíficos Mayas a partir de puntos como primitivas

De los enfoques utilizados para la representación de bocetos, únicamente dos de ellos han sido utilizados para la representación de jeroglíficos: 1) representación a través de descriptores y (2) representación basada en el modelo BoV. Ambos enfoques usan algunos puntos de la imagen como primitivas, donde la calidad de la representación depende en gran medida del conjunto de puntos obtenidos a partir de la imagen.

Trabajos previos han dado evidencia de que la representación de jeroglíficos haciendo uso de descriptores calculados desde un conjunto de puntos seleccionados aleatoriamente con distribución uniforme [18] produce buenos resultados en la recuperación de jeroglíficos. Sin embargo, la selección aleatoria de puntos con distribución uniforme da lugar a un conjunto de puntos, un conjunto de descriptores y por lo tanto una representación diferente en cada ejecución. Además, usar puntos obtenidos aleatoriamente no garantiza que la representación resultante sea obtenida a partir de puntos con información importante de la imagen.

Para evitar el proceso aleatorio de selección de puntos, proponemos un método de detección de puntos de interés calculados mediante métodos de detección de puntos dominantes. El método propuesto inicialmente incluye una etapa de poda de contornos adelgazados. Posteriormente, se obtiene un conjunto de puntos de interés particionando la imagen adelgazada en segmentos de línea, obteniendo para cada segmento un subconjunto de puntos de interés mediante un método de detección de puntos dominantes. Los puntos dominantes obtenidos para todos los segmentos de línea dan lugar a un conjunto de puntos de interés, a partir de los cuales se calculan descriptores de contexto de forma (HOOSC). Finalmente, la representación de jeroglíficos es obtenida haciendo uso del modelo BoV, agrupando descriptores similares usando k -means y seleccionando el centroide de cada grupo para construir un vocabulario de k términos visuales. Cada descriptor en la imagen es entonces etiquetado como el término visual más parecido en el vocabulario, representando cada jeroglífico mediante un histograma de frecuencia de términos visuales en la imagen.

De esta forma el método propuesto está dividido en seis etapas: *a)* adelgazamiento del jeroglífico Maya, *b)* podado de ramas no deseadas, *c)* partición del jeroglífico adelgazado y

podado en segmentos de línea, *d*) obtención del conjunto de puntos de interés, *e*) cálculo de descriptores HOOSC a partir del conjunto de puntos de interés y *f*) construcción del vocabulario de términos visuales y representación de jeroglíficos como histogramas de palabras visuales (ver *Figura 6.1*).

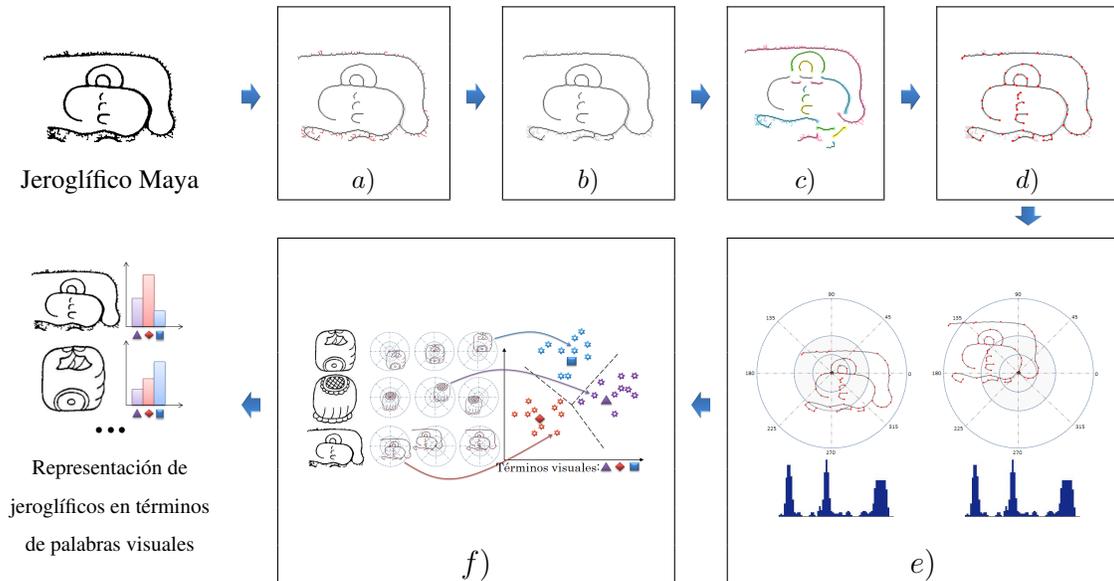


Figura 6.1: *a*) Adelgazamiento del jeroglífico. *b*) Podado de ramas adicionales. *c*) Partición del jeroglífico delgado en segmentos de línea, *d*) Obtención del conjunto de puntos de interés *e*) cálculo de descriptores HOOSC a partir del conjunto de puntos de interés y *f*) construcción del vocabulario de términos visuales y representación de jeroglíficos en términos de palabras visuales.

6.1.1. Adelgazamiento y poda de jeroglíficos Mayas

Trabajos previos [18] dan evidencia de que para describir a los jeroglíficos se obtienen mejores características para describir a los jeroglíficos si se trabaja con jeroglíficos adelgazados en comparación con las obtenidas a partir de contornos gruesos (ver *Figura 6.2b* y *Figura 6.2c*). Sin embargo, durante el proceso de adelgazamiento, pequeñas perturbaciones y ruido presente en el contorno de la imagen pueden producir ramas no deseadas (ver *Figura 6.2c*), las cuales no representan partes importantes de la forma original. Por lo tanto, proponemos incluir una etapa de poda, para eliminar ramas no deseadas en la representación adelgazada de los jeroglíficos (ver *Figura 6.2d*).

Una vez que un jeroglífico ha sido adelgazado, siguiendo el método presentado en [64],

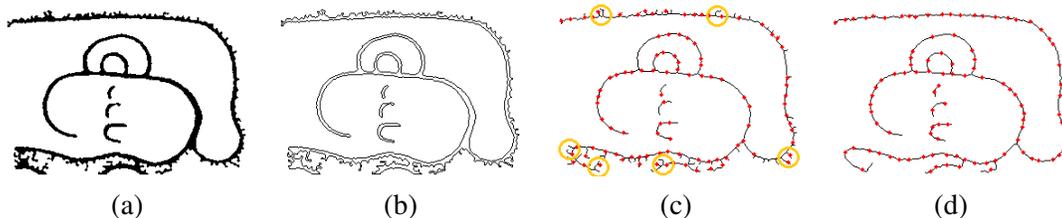


Figura 6.2: (a) Jeroglífico Maya. (b) Extracción del contorno, (c) contorno adelgazado, algunas ramas adicionales están marcadas en círculo. (d) Puntos obtenidos mediante selección aleatoria uniforme a partir del contorno adelgazado y podado.

diferentes componentes (píxeles y ramas) pueden ser identificados. Sea p un píxel en el jeroglífico adelgazado y considerando la vecindad 3×3 centrada en p , un píxel p es un *pixel de unión* b_p si hay tres o más píxeles del jeroglífico adelgazado en esta vecindad; un píxel p es un *pixel final* e_p si hay un único píxel en esta vecindad; y un píxel p es un *pixel de curva* c_p si tiene exactamente dos píxeles en esta vecindad.

Una rama B_j es definida como un conjunto de píxeles consecutivos y adyacentes p_i , $i = \{1, 2, \dots, N\}$ (N es el número de píxeles en B_j), una rama es una curva de un píxel de ancho que puede ser *final* o *simple*. Una rama es una *rama final* si no contiene *píxeles finales* ni *píxeles de unión* excepto p_1 y p_N (ver Figura 6.3c). Una rama es una *rama simple* si ambos p_1 y p_N son píxeles finales. Una *rama final* es producida durante la etapa de adelgazamiento por ruido presente en el contorno de la forma o por la presencia de partes salientes en la imagen; mientras que una rama simple representa una región en la imagen, a partir de la cual se obtiene una *curva abierta simple*. Una curva es abierta y simple si los píxeles donde inicia y termina la curva son diferentes $p_1 \neq p_N$ y no se cruza consigo mismo en ningún punto.

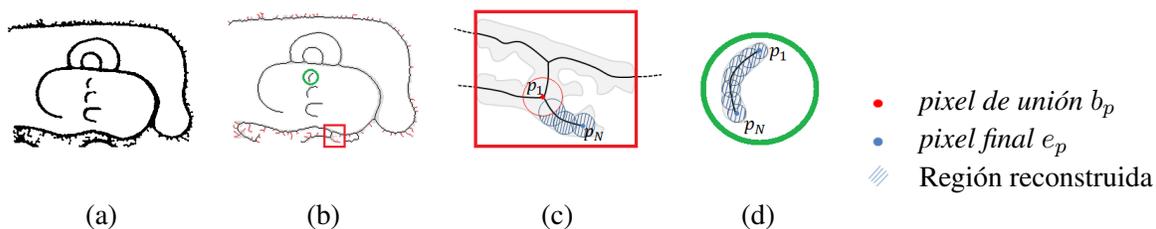


Figura 6.3: (a) Ramas consideradas en la etapa de poda. (b) *Rama final* y una *rama simple* marcadas con un cuadrado y un círculo, respectivamente. (c) y (d) Región reconstruida a partir de una *rama final* y una *rama simple*, respectivamente

Después de adelgazar el contorno se producen algunas ramas no deseadas, las cuales deben ser podadas, una rama debe ser podada si no es importante para la representación de

la imagen. Diferentes medidas pueden ser utilizadas para determinar la importancia de una rama [66]. Nosotros proponemos remover *ramas finales* y *ramas simples* que no provean información útil (es decir, que no sean importantes), dando lugar a una representación más “limpia” de los jeroglíficos Mayas. Para esto procedemos de la siguiente manera:

Sea $A(I)$ la región del *foreground* de la imagen original, consideramos que B_j no representa información útil si la región $R(B_j)$ reconstruida a partir de B_j es más pequeña que cierto porcentaje α de $A(I)$; siendo α un parámetro definido por el usuario. De esta manera, las ramas que cumplen esta condición son podadas.

$$numPixeles(R(B_j)) < \alpha * numPixeles(A(I))$$

La región que representa una rama puede ser reconstruida si, para cada pixel p_i en la rama B_j conocemos la transformada de distancia en la imagen original. La transformada de distancia en p_i , denotada como $dt(p_i)$ es la distancia entre p_i y el pixel más cercano en el *background* de la imagen, así la región reconstruida a partir del pixel p_i denotada como $R(p_i)$ contiene los pixeles al interior del disco con origen en p_i y radio $dt(p_i)$

Como se puede apreciar en la *Figura 6.3d* la región reconstruida a partir de una *rama simple* B_j es calculada como $R(B_j) = \{R(p_1) \cup R(p_2), \dots, R(p_N)\}$ (ver *Figura 6.3d*). Mientras, que la región reconstruida a partir de una *rama final* no considera los pixeles de $R(p_1)$ y es calculada como $R(B_j) = \{R(p_2) \cup R(p_3), \dots, R(p_N) - R(p_1)\}$ (ver *Figura 6.3c*).

Una vez que los jeroglíficos Mayas han sido adelgazados y podados, el objetivo es proveer un conjunto de puntos de interés que dé lugar a un conjunto de descriptores útiles.

6.1.2. Descriptores desde jeroglíficos Mayas

Los métodos de detección de puntos dominantes han sido ampliamente utilizados en diferentes aplicaciones debido a que permiten representar las principales características de una curva usando un conjunto reducido de puntos, los cuales son identificados como puntos de máxima curvatura local.

Un método de detección de puntos dominantes llamado DCE (del inglés *Discrete Curve Evolution*) fue introducido en [67]. DCE remueve en cada iteración el pixel con menor

curvatura hasta que el conjunto de puntos dominantes contiene el número deseado de píxeles.

Los métodos de detección de puntos dominantes están diseñados para trabajar sobre *curvas cerradas simples* o *curvas abiertas simples*. Sea C un conjunto de Z píxeles adyacentes y consecutivos $C = \{p_1, p_2, \dots, p_Z\}$ de un píxel de ancho. Una curva es una *curva abierta* si los píxeles donde inicia y termina la curva son diferentes ($p_1 \neq p_Z$) (ver *Figura 6.4a*). Una curva es *curva cerrada* si el píxel donde inicia y termina la curva es el mismo ($p_1 = p_Z$) (ver *Figura 6.4b*). Una curva es una *curva simple* si no se cruza consigo misma, pudiendo ser abierta o cerrada.

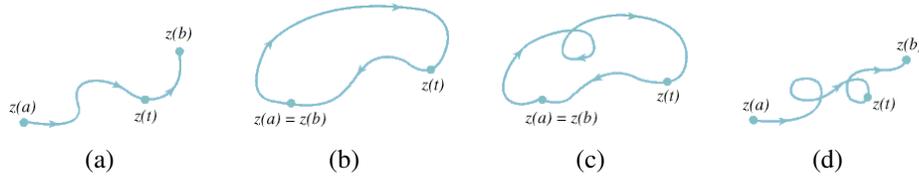


Figura 6.4: (a) Curva simple, (b) Curva cerrada simple, (c) Curva cerrada compleja y (d) curva abierta compleja.

Sin embargo, los jeroglíficos Mayas son bocetos complejos, los cuales una vez adelgazados están compuestos por múltiples curvas, muchas de ellas al estar intersecadas entre sí dan lugar a que se obtengan curvas que no sean ni *curvas cerradas simples*, ni *curvas abiertas simples* (curvas complejas *Figura 6.4c-d*). Por lo tanto, los jeroglíficos necesitan ser descompuestos en *curvas simples*, de modo que un conjunto de puntos dominantes pueda ser obtenido usando el método DCE.

En este método, una *curva simple* o *segmento simple de línea* Sc_t es considerado como un conjunto de píxeles conectados $p_i, i = \{1, 2, \dots, z\}$ (z es el número de píxeles en el segmento), que no contengan *píxeles de unión* ni *píxeles finales*, excepto los píxeles ubicados en los extremos del segmento.

Para cada segmento $Sc_t \in Sk, t = \{1, 2, \dots, T\}$, donde Sk es el jeroglífico adelgazado y T es el número de segmentos obtenidos, obtenemos un conjunto Dp_t de puntos dominantes haciendo uso del método DCE. El parámetro de entrada requerido por DCE (número de puntos a obtener), se calcula como el porcentaje ω (indicado por el usuario) del número total de píxeles en el segmento.

Así, el conjunto de puntos de interés en un jeroglífico es obtenido al considerar los puntos

dominantes de todos los segmentos.

$$Dp(Sk) = Dp_1 \cup Dp_2 \cup \dots \cup Dp_T$$

6.1.3. Construcción del vocabulario visual y representación de jeroglíficos en términos de palabras visuales

Una vez que el conjunto de puntos de interés $DP(Sk)$ ha sido obtenido para cada imagen en la base de datos, un descriptor HOOSC es calculado para cada punto de interés. Sin embargo, como ya se ha señalado, imágenes que pertenecen a la misma categoría pueden tener descriptores similares pero no idénticos, producto de las variaciones que puedan presentarse entre las imágenes. Teniendo por objetivo obtener un vocabulario visual que permita describir a todo el conjunto de imágenes, los descriptores son agrupados asumiendo que aquellos que pertenecen a un mismo grupo son similares entre sí y por lo tanto pueden ser representados mediante el mismo término visual. El centroide de cada grupo puede ser usado como el descriptor que representa a los descriptores en el grupo, para ello el algoritmo *k-means* es utilizado, donde el número de grupos se corresponde con el tamaño k del vocabulario visual, definido por el usuario.

Una vez que el vocabulario visual ha sido obtenido, las imágenes son representadas en términos de este vocabulario visual. Para ello, cada descriptor en la imagen es asignado al término visual más cercano en el vocabulario, este método de cuantización (o codificación) es conocido como VQ (del inglés *Vector Quantization*) [50]. Así, las ocurrencias de cada término visual en la imagen son consideradas para la construcción de un histograma, el cual es usado como la representación final de la imagen.

Ya que el número de descriptores puede ser diferente entre imágenes, el histograma de frecuencias es normalizado, permitiendo que los histogramas que representan a los jeroglíficos puedan ser directamente comparados.

6.1.4. Resultados experimentales

En nuestros experimentos utilizamos dos bases de datos de jeroglíficos. La primera base de datos obtenida de [23] está compuesta de 240 jeroglíficos distribuidos en 24 categorías (10

jeroglíficos por categoría). En la *Figura 6.5* se presenta un ejemplo de cada categoría.

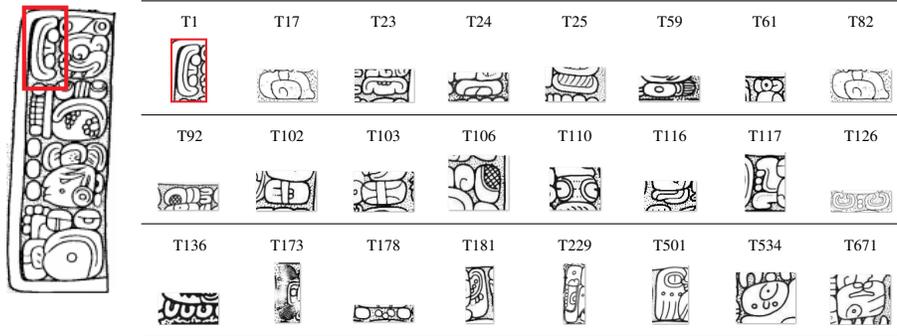


Figura 6.5: Ejemplos de instancias por cada categoría en la base de datos compuesta por 240 jeroglíficos

La segunda base de datos es un subconjunto de la base de datos usada en [18] compuesta por 1270 jeroglíficos y distribuidos en 24 categorías, la cual después de ser procesada por arqueólogos dio lugar una base de datos con 1043 jeroglíficos distribuidos en 25 categorías. La *Figura 6.6* muestra un histograma que refleja el número de jeroglíficos que contiene cada categoría y un ejemplo por cada una.

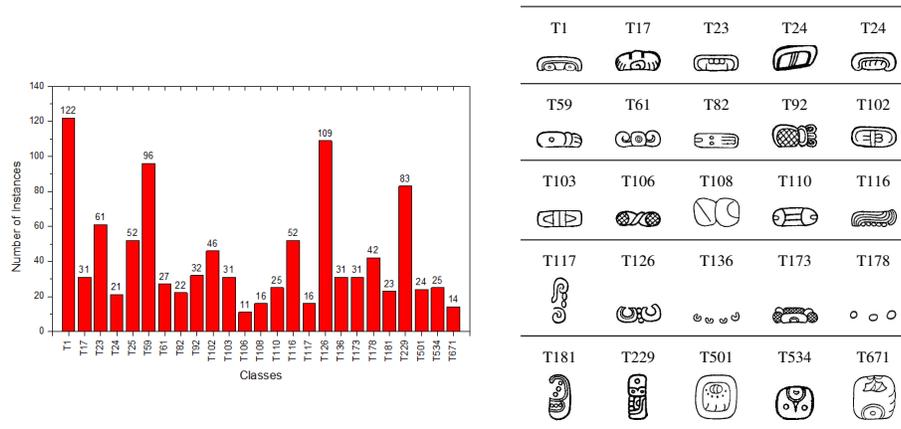


Figura 6.6: Base de datos que contiene 1043 jeroglíficos Mayas. El histograma de la izquierda muestra el número de instancias por categoría. A la derecha se muestra una instancia por cada categoría

En esta sección presentamos los resultados obtenidos en recuperación de jeroglíficos usando la representación propuesta basada en la descripción a partir de puntos de interés obtenidos por medio de puntos dominantes y contornos adelgazados y podados, en comparación con los resultados obtenidos mediante la representación presentada en [18]. Para ambos

métodos usamos dos formulaciones del descriptor HOOSC: HOOSC4 y HOOSC128 introducidos en [18] y [45], respectivamente.

Para la poda del contorno adelgazado utilizamos $\alpha = 0.005$, este valor fue obtenido probando valores entre 0.002 y 0.01, con incrementos de 0.001. El valor seleccionado para α provee los mejores resultados visuales en ambas bases de datos. En el método de representación propuesto, los descriptores HOOSC4 y HOOSC128 fueron calculados a partir del conjunto de puntos de interés, obtenidos como puntos dominantes. En [18] los descriptores son calculados a partir de puntos seleccionados de forma aleatoria y uniforme. En ambos métodos usamos $\omega = 5\%$, 10% y 20% de los puntos del jeroglífico, considerando el número de puntos en el jeroglífico adelgazado en el método [18] y el jeroglífico adelgazado y podado en nuestro método. Para la construcción del diccionario de términos visuales, seleccionamos aleatoriamente 200 descriptores para cada una de las 24 categorías, los cuales fueron utilizados para obtener 1000 términos visuales usando *k-means*.

Los resultados obtenidos a partir de la recuperación de jeroglíficos son mostrados en la *Tabla 6.1* y *6.2* en términos de la precisión promedio (mAP) para la primera y segunda base de datos, respectivamente. Los resultados presentados son el promedio obtenido de 10 repeticiones. La primera fila muestra los resultados obtenidos mediante el método en [18] denotado como *THIN+UDS*, mientras que la segunda fila presenta los resultados obtenidos haciendo uso de la representación propuesta denotado como *THIN + PRUN + DP*. Los resultados son presentados para ambos descriptores (HOOSC4 y HOOSC128), considerando los diferentes porcentajes de puntos utilizados. El tiempo de ejecución requerido por cada método es presentado en la parte inferior de cada resultado.

Tabla 6.1: mAP (*mean average precision*) para la base de datos que contiene 240 jeroglíficos Mayas

Método	Descriptor					
	HOOSC4			HOOSC128		
	5%	10%	20%	5%	10%	20%
<i>THIN + UDS</i> [18]	0.164	0.244	0.308	0.129	0.190	0.234
Runtime (min.)	18.45	18.65	20.64	7.57	7.07	8.66
<i>THIN + PRUN + DP</i>	0.284	0.312	0.340	0.146	0.180	0.241
Runtime (min.)	21.01	21.08	22.23	7.73	8.58	8.32

En las *Tablas 6.1* y *6.2* los mejores resultados son marcados en **negrita**. Como se puede

Tabla 6.2: mAP (*Mean average precision*) para la base de datos que contiene 1043 jeroglíficos Mayas

Método	Descriptor					
	HOOSC4			HOOSC128		
	5 %	10 %	20 %	5 %	10 %	20 %
<i>THIN + UDS</i> [18]	0.174	0.261	0.353	0.149	0.217	0.299
Runtime (hrs.)	20.34	21.39	23.88	7.98	8.69	10.10
<i>THIN + PRUN + DP</i>	0.202	0.294	0.379	0.173	0.248	0.338
Runtime (hrs.)	21.76	23.14	24.61	10.78	11.30	12.12

observar, el método propuesto obtiene los mejores resultados en mAP para casi todas las combinaciones, excepto en la base de datos de 240 jeroglíficos Mayas, usando el 10 % de puntos de interés a partir del jeroglífico adelgazado y calculando descriptores HOOSC128. Sin embargo, es necesario resaltar que los mejores resultados se obtienen usando descriptores HOOSC4, aunque el tiempo requerido para representar y recuperar jeroglíficos haciendo uso este descriptor es más del doble que el requerido al usar descriptor HOOSC128. Lo anterior se debe a la dimensión de los descriptores.

En las *Figuras 6.7 y 6.8* el *average precision* vs el *standard recall* son presentados para ambas bases de datos. Como podemos observar, nuestro método siempre obtiene los mejores resultados considerando el primer elemento recuperado (*standard recall* entre 0.0 y 0.2), inclusive se puede apreciar que la precisión promedio de esas primeras instancias recuperadas es, en la mayoría de los casos, mucho mayor que la obtenida con el método en [18]

7. Conclusiones

Este documento presenta la propuesta de investigación doctoral a desarrollar, la cual se centra en proponer una representación de jeroglíficos de modo que sea posible llevar a cabo tareas de recuperación de jeroglíficos mejor que utilizando los métodos existentes en la literatura.

Como resultado preliminar, en esta propuesta de investigación se ha desarrollado una representación de jeroglíficos Mayas siguiendo el enfoque más exitoso reportado en la literatura. A diferencia de los trabajos reportados en la literatura, el método propuesto combina adelgazamiento y poda de jeroglíficos, con segmentación de la imagen en curvas simples para

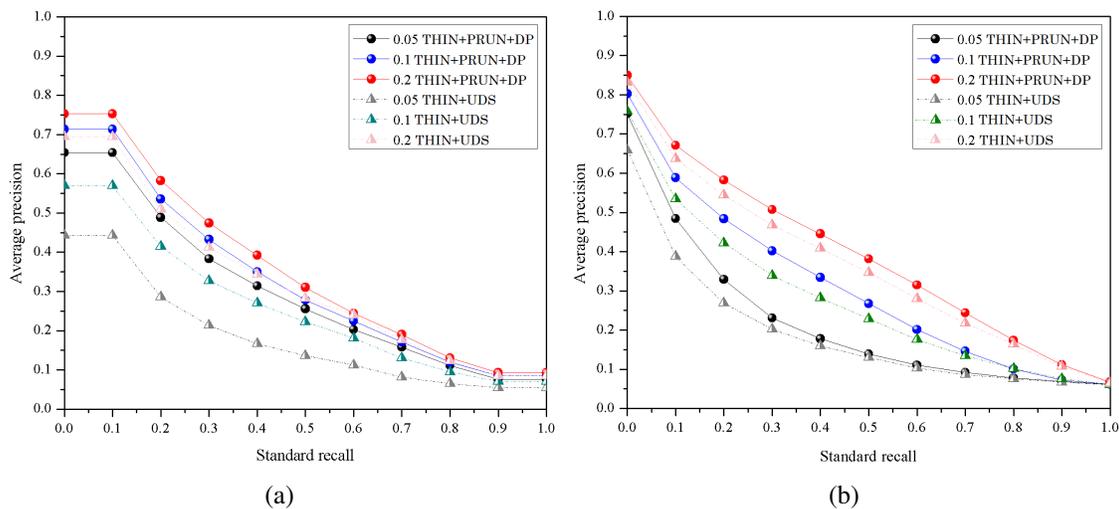


Figura 6.7: *Average precision vs standard recall* (a) base de datos que contiene 240 jeroglíficos Mayas y (b) base de datos que contiene 1043 jeroglíficos Mayas usando el descriptor HOOSC4

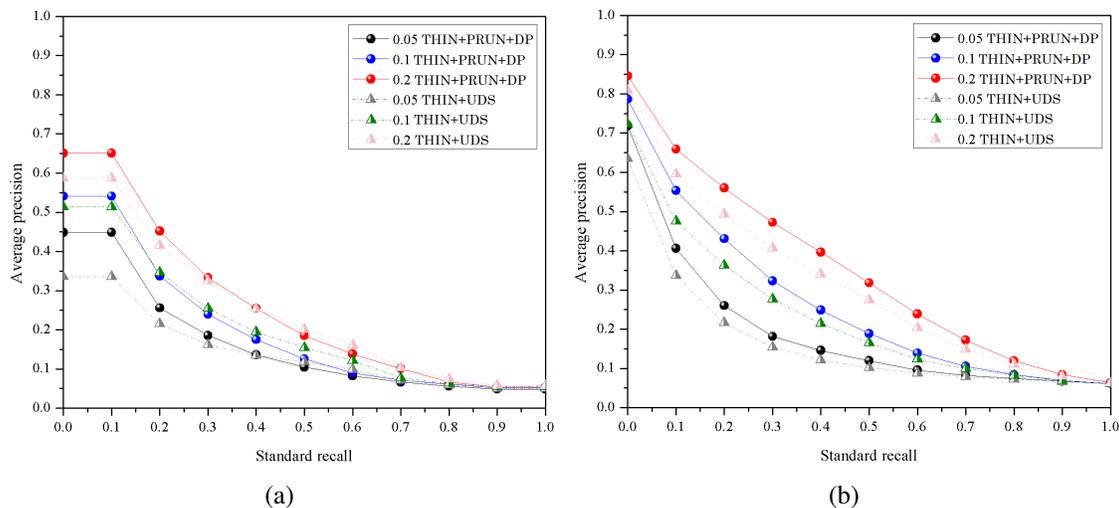


Figura 6.8: *Average precision vs standard recall* (a) base de datos que contiene 240 jeroglíficos Mayas y (b) base de datos que contiene 1043 jeroglíficos Mayas usando el descriptor HOOSC128

la obtención de puntos dominantes, los cuales son utilizados como los puntos de interés de la imagen. Finalmente, la representación de los jeroglíficos es obtenida a partir de descriptores calculados desde el conjunto de puntos de interés. El rendimiento del método propuesto es evaluado, en términos de recuperación de jeroglíficos, haciendo uso de dos bases de datos. Basándose en los experimentos, los resultados dan evidencia de que el método propuesto permite extraer información importante del jeroglífico, dando lugar a una mejor representación y

por ende a una mejor recuperación de este tipo de imágenes respecto a métodos reportados en la literatura. Nuestra propuesta tiene como principal ventaja que evita el proceso de selección aleatoria de los puntos de interés, que se usa comúnmente en la literatura, permitiendo que una mejor representación de jeroglíficos sea obtenida haciendo uso de un menor número de puntos de interés de la imagen.

Por otro lado, según nuestro conocimiento, a partir de la revisión bibliografía realizada, la representación de jeroglíficos solo ha sido explorada haciendo uso de *puntos* como primitivas. En esta propuesta de investigación, estamos interesados en explorar diferentes formas de representar jeroglíficos donde además de los puntos, otro tipo de primitivas (*p. ej.* segmentos de línea o regiones) sean empleados para generar la representación. En la representación se incluirá información adicional tal como, la distribución espacial de las primitivas en el jeroglífico y las relaciones existentes entre múltiples primitivas, de modo que la representación resultante sea más discriminativa y permita obtener mejores resultados en recuperación de jeroglíficos.

Los resultados preliminares presentados son alentadores y por lo tanto consideramos que los objetivos planteados en esta propuesta doctoral pueden alcanzarse en el tiempo estipulado siguiendo la metodología propuesta.

Notación

<i>Notación</i>	<i>Significado</i>
b_p	Pixel de unión
e_p	Pixel final
c_p	Pixel de curva
B	Rama en la forma adelgazada
$Dp(.)$	Puntos dominantes
$dt(.)$	Transformada de distancia
D	Vocabulario de términos visuales
w	Término visual en el vocabulario
K	K – Vecinos más cercanos
k	Tamaño del vocabulario de términos visuales (k – <i>means</i>)

Acrónimos

<i>Acrónimo</i>	<i>Significado</i>
SC	<i>Shape Context</i>
GSC	<i>Generalized Shape Context</i>
HOG	<i>Histogram of Oriented Gradient</i>
HOOSC	<i>Histogram of Orientation Shape Context</i>
DOG	<i>Difference of Gaussian</i>
SS	<i>Self-Similarities</i>
SIFT	<i>Scale-Invariant Feature Transform</i>
DCE	<i>Discrete Curve Evolution</i>
k -nn	<i>k-Nearest Neighbor</i>
BoV	Bolsa de Palabras Visuales
BoP	Bolsa de Frases Visuales
VQ	<i>Vector Quantization</i>
SPM	<i>Spatial Pyramid Matching</i>
AVP	<i>Average Pooling</i>
Max-NWP	<i>Max-N Weight Pooling</i>
Max-NBP	<i>Max-N Binary Pooling</i>
Max-NIP	<i>Max-N Integer Pooling</i>
LLC	<i>Locality-constrained Linear Coding</i>

Referencias

1. Xuejin Chen, Sing Bing Kang, Ying-Qing Xu, Julie Dorsey, and Heung-Yeung Shum. Sketching reality: Realistic interpretation of architectural designs. *ACM Trans. Graph.*, 27(2):11:1–11:15, May 2008.
2. Achraf Ghorbel, Eric Anquetil, Jean Camillerapp, and Aurélie Lemaitre. Imisketch: An interactive method for sketch recognition. *Pattern Recognition Letters*, 35(0):78 – 90, 2014.
3. J.-P. Valois, M. Cote, and M. Cheriet. Online recognition of sketched electrical diagrams. In *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, pages 460–464, 2001.
4. T.M. Sezgin and R. Davis. Sketch recognition in interspersed drawings using time-based graphical models. *Computers and Graphics*, 32(5):500 – 510, 2008.
5. Tom Y. Ouyang and Randall Davis. Chemink: A natural real-time recognition system for chemical drawings. In *International Conference on Intelligent User Interfaces (IUI '11)*, 2011.
6. A. Rebelo, G. Capela, and JaimeS. Cardoso. Optical recognition of music symbols. *International Journal on Document Analysis and Recognition (IJ DAR)*, 13(1):19–31, 2010.
7. Ann Theja Alex, Vijayan K. Asari, and Alex Mathew. Local difference of gaussian binary pattern: Robust features for face sketch recognition. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 1211–1216, Oct 2013.
8. Mingjin Zhang, Jie Li, Nannan Wang, and Xinbo Gao. Recognition of facial sketch styles. *Neurocomputing*, 149, Part C(0):1188 – 1197, 2015.
9. Mathias Eitz, Ronald Richter, Tamy Boubekeur, Kristian Hildebrand, and Marc Alexa. Sketch-based shape retrieval. *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, 31(4):31:1–31:10, 2012.

10. Morris Franken and Jan van Gemert. Automatic egyptian hieroglyph recognition by retrieving images as texts. In *ACM Multimedia Conference, MM '13, Barcelona, Spain, October 21-25, 2013*, pages 765–768, 2013.
11. Rui Hu, Carlos Pallan Gayol, Guido Krempel, Jean-Marc Odobez, and Daniel Gatica-Perez. Automatic maya hieroglyph retrieval using shape and context information. In *Proceedings of the ACM International Conference on Multimedia, MM '14*, pages 1037–1040. ACM, 2014.
12. E. Roman-Rangel, C. Pallan, J. Odobez, and D. Gatica Perez. Retrieving ancient maya glyphs with shape context. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 988–995, Sept 2009.
13. Z. Khan, Yiqun Hu, and A. Mian. Facial self similarity for sketch to photo matching. In *Digital Image Computing Techniques and Applications (DICTA), 2012 International Conference on*, pages 1–7, Dec 2012.
14. Junwei Wang, Xiang Bai, Xinge You, Wenyu Liu, and Longin Jan Latecki. Shape matching and classification using height functions. *Pattern Recognition Letters*, 33(2):134 – 143, 2012.
15. Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *In NIPS*, pages 831–837, 2000.
16. Edgar Roman-Rangel, Carlos Pallan, Jean Marc Odobez, and Daniel Gatica Perez. Analyzing ancient maya glyph collections with contextual shape descriptors. *International Journal of Computer Vision*, 94(1):101–117, 2010.
17. Pedro Sousa and Manuel J. Fonseca. Geometric matching for clip-art drawing retrieval. *Journal of Visual Communication and Image Representation*, 20(2):71 – 83, 2009. Special issue on Emerging Techniques for Multimedia Content Sharing, Search and Understanding.
18. Edgar Roman-Rangel, Carlos Pallan Gayol, Jean Marc Odobez, and Daniel Gatica Perez. Searching the past: An improved shape descriptor to retrieve maya hieroglyphs. In

Proceedings of the 19th ACM International Conference on Multimedia, MM '11, pages 163–172, New York, NY, USA, 2011. ACM.

19. Mathias Eitz, James Hays, and Marc Alexa. How do humans sketch objects? *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):44:1–44:10, 2012.
20. Jun Guo, Changhu Wang, and Hongyang Chao. Building effective representations for sketch recognition. 2015.
21. Serge Belongie, Jitendra Malik, and Jan Puzicha. Shape matching and object recognition using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):509–522, Apr 2002.
22. Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893 vol. 1, June 2005.
23. Edgar Roman-Rangel, Jean Marc Odobez, and Daniel Gatica Perez. Evaluating shape descriptors for detection of maya hieroglyphs. In JesúsAriel Carrasco-Ochoa, JoséFrancisco Martínez-Trinidad, JoaquínSalas Rodríguez, and GabriellaSanniti di Baja, editors, *Pattern Recognition*, volume 7914 of *Lecture Notes in Computer Science*, pages 145–154. Springer Berlin Heidelberg, 2013.
24. Diego Macrini, Sven Dickinson, David Fleet, and Kaleem Siddiqi. Bone graphs: Medial shape parsing and abstraction. *Computer Vision and Image Understanding*, 115(7):1044 – 1061, 2011.
25. Yin Lin, Liu Wenyin, and Changjun Jiang. A structural approach to recognizing incomplete graphic objects. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 1, pages 371–375 Vol.1, Aug 2004.
26. J. Lladós, E. Martí, and J.J. Villanueva. Symbol recognition by error-tolerant subgraph matching between region adjacency graphs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(10):1137–1143, Oct 2001.

27. WeeSan Lee, Levent Burak Kara, and Thomas F. Stahovich. An efficient graph-based symbol recognizer. In *Proceedings of the Third Eurographics Conference on Sketch-Based Interfaces and Modeling, SBM'06*, pages 11–18, Aire-la-Ville, Switzerland, Switzerland, 2006. Eurographics Association.
28. Gabriella Csurka, Christopher R. Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray. Visual categorization with bags of keypoints. In *In Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, 2004.
29. Yann Frauel, Octavio Quesada, and Ernesto Bribiesca. Detection of a polymorphic mesoamerican symbol using a rule-based approach. *Pattern Recognition*, 39(7):1380 – 1390, 2006.
30. Edgar Roman-Rangel and Stephane Marchand Maillet. Shape-based detection of maya hieroglyphs using weighted bag representations. *Pattern Recognition*, 48(4):1161 – 1173, 2015.
31. M. Peura and J. Iivarinen. Efficiency of simple shape descriptors. In *In Aspects of Visual Form*, pages 443–451. World Scientific, 1997.
32. Akrem El-ghazal, Otman Basir, and Saeid Belkasim. Farthest point distance: A new shape signature for fourier descriptors. *Signal Processing: Image Communication*, 24(7):572 – 586, 2009.
33. Farzin Mokhtarian, Sadegh Abbasi, and Josef Kittler. Efficient and robust retrieval by shape content through curvature scale space. pages 35–42, 1996.
34. Zhanwei Yuan, Fuguo Li, Peng Zhang, and Bo Chen. Description of shape characteristics through fourier and wavelet analysis. *Chinese Journal of Aeronautics*, 27(1):160 – 168, 2014.
35. C.-H. Teh and R.T. Chin. On image analysis by the methods of moments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 10(4):496–513, Jul 1988.
36. Guojun Lu and Atul Sajjanhar. Region-based shape representation and similarity measure suitable for content-based image retrieval. *Multimedia Systems*, 7(2):165–174, 1999.

37. DavidG. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. ISSN 0920-5691.
38. T.B. Sebastian, P.N. Klein, and B.B. Kimia. Recognition of shapes by editing their shock graphs. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(5):550–571, May 2004.
39. Pedro Sousa and Manuel J. Fonseca. Sketch-based retrieval of drawings using spatial proximity. *Journal of Visual Languages and Computing*, 21:69–80, 2010.
40. Relja Arandjelović and Tevfik Metin Sezgin. Sketch recognition by fusion of temporal and image-based features. *Pattern Recognition*, 44(6):1225 – 1234, 2011.
41. Shuang Liang, Jun Luo, Wenyin Liu, and Yichen Wei. Sketch matching on topology product. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(8):1723–1729, 2015.
42. M. Eitz, K. Hildebrand, T. Boubekeur, and M. Alexa. Sketch-based image retrieval: Benchmark and bag-of-features descriptors. *Visualization and Computer Graphics, IEEE Transactions on*, 17(11):1624–1636, Nov 2011.
43. Rui Hu and John Collomosse. A performance evaluation of gradient field hog descriptor for sketch based image retrieval. *Comput. Vis. Image Underst.*, 117(7):790–806, July 2013.
44. Xinggang Wang, Bin Feng, Xiang Bai, Wenyu Liu, and Longin Jan Latecki. Bag of contour fragments for robust shape classification. *Pattern Recognition*, 47(6):2116 – 2125, 2014.
45. Edgar Roman-Rangel and Stephane Marchand-Maillet. Hoosc128: A more robust local shape descriptor. In JoséFrancisco Martínez-Trinidad, JesúsAriel Carrasco-Ochoa, JoséArturo Olvera-Lopez, Joaquín Salas-Rodríguez, and ChingY. Suen, editors, *Pattern Recognition*, volume 8495 of *Lecture Notes in Computer Science*, pages 172–181. Springer International Publishing, 2014.

46. Yann Frauel, Octavio Quesada, and Ernesto Bribiesca. Detection of a polymorphic mesoamerican symbol using a rule-based approach. *Pattern Recognition*, 39(7):1380 – 1390, 2006.
47. Octavio Quesada-García. *Tres signos. Escritura Antigua de Mesoamerica*. Universidad Nacional Autonoma de Mexico, 2006.
48. G. Mori, S. Belongie, and J. Malik. Efficient shape matching using shape contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(11):1832–1837, Nov 2005.
49. Marçal Rusiñol and Josep Lladós. Efficient logo retrieval through hashing shape context descriptors. In *Proceedings of the 9th IAPR International Workshop on Document Analysis Systems, DAS '10*, pages 215–222, New York, NY, USA, 2010. ACM.
50. J. Sivic and A. Zisserman. Video google: a text retrieval approach to object matching in videos. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1470–1477 vol.2, Oct 2003.
51. Daniel Gatica-Perez, Edgar Roman-Rangel, Jean-Marc Odobez, and Carlos Pallan. New world, new worlds: Visual analysis of pre-columbian pictorial collections. In Costantino Grana and Rita Cucchiara, editors, *Multimedia for Cultural Heritage*, volume 247 of *Communications in Computer and Information Science*, pages 90–102. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-27977-5.
52. Krystian Mikolajczyk and Cordelia Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1):63–86, 2004.
53. Edgar Roman-Rangel and Stephane Marchand-Maillet. Stopwords detection in bag-of-visual-words: The case of retrieving maya hieroglyphs. In Alfredo Petrosino, Lucia Maddalena, and Pietro Pala, editors, *New Trends in Image Analysis and Processing – ICIAP 2013*, volume 8158 of *Lecture Notes in Computer Science*, pages 159–168. Springer Berlin Heidelberg, 2013. ISBN 978-3-642-41189-2.

54. Edgar Roman-Rangel and Stephane Marchand-Maillet. Automatic removal of visual stop-words. In *Proceedings of the ACM International Conference on Multimedia, MM '14*, pages 1145–1148, New York, NY, USA, 2014. ACM.
55. Edgar Roman-Rangel, Jean Marc Odobez, and Daniel Gatica Perez. Assessing sparse coding methods for contextual shape indexing of maya hieroglyphs. *Journal of Multimedia*, 7(2), 2012.
56. M. Aharon, M. Elad, and A. Bruckstein. k-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11): 4311–4322, Nov 2006.
57. S. Lloyd. Least squares quantization in pcm. *IEEE Trans. Inf. Theor.*, 28(2):129–137, September 2006. ISSN 0018-9448.
58. E. Roman-Rangel and S. Marchand-Maillet. Bag-of-visual-phrases via local contexts. In *Pattern Recognition (ACPR), 2013 2nd IAPR Asian Conference on*, pages 867–871, Nov 2013.
59. S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 2, pages 2169–2178, 2006.
60. E. Zhang and M. Mayo. Improving bag-of-words model with spatial information. In *Image and Vision Computing New Zealand (IVCNZ), 2010 25th International Conference of*, pages 1–8, Nov 2010.
61. Sezer Karaoglu, JanC. van Gemert, and Theo Gevers. Object reading: Text recognition for object recognition. In *Computer Vision – ECCV 2012. Workshops and Demonstrations*, volume 7585 of *Lecture Notes in Computer Science*, pages 456–465. Springer Berlin Heidelberg, 2012.
62. E. Shechtman and M. Irani. Matching local self-similarities across images and videos. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1–8, June 2007.

63. Harri Kettunen and Christophe Helmke. *Introduction to Maya Hieroglyphs*. 2011.
64. Louisa Lam, S.-W. Lee, and C.Y. Suen. Thinning methodologies-a comprehensive survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(9):869–885, Sep 1992.
65. C. Di Ruberto and A. Morgera. Aco contour matching: A dominant point approach. In *Image and Signal Processing (CISP), 2011 4th International Congress on*, volume 3, pages 1391–1395, Oct 2011.
66. HongZhi Liu, Zhong-Hai Wu, Xing Zhang, and D. Frank Hsu. A skeleton pruning algorithm based on information fusion. *Pattern Recognition Letters*, 34(10):1138 – 1145, 2013.
67. Longin Jan Latecki and Rolf Lakämper. Polygon evolution by vertex deletion. In *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision, SCALE-SPACE '99*, pages 398–409, London, UK, UK, 1999. Springer-Verlag.