



Universidad Nacional Autónoma de México

Posgrado en Ciencia e Ingeniería de la Computación

Clasificación automática de puntas de proyectil arqueológicas

T E S I S

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS (COMPUTACIÓN)

P R E S E N T A:

Fernando Castillo Flores

T U T O R:

Dr. Francisco Javier García Ugalde
Facultad de Ingeniería–UNAM

C O M I T É T U T O R:

Dr. Ernesto Bribiesca Correa
IIMAS–UNAM

Dr. Fernando Arámbula Cosío
ICAT–UNAM

Ciudad Universitaria, Cd. Mx, a marzo de 2019



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

J U R A D O:

PRESIDENTE:	Dr. Ernesto Bribiesca Correa	ICAT-UNAM
VOCAL:	Dr. Francisco Javier García Ugalde	Fac. Ingeniería-UNAM
SECRETARIO:	Dr. Fernando Arámbula Cosío	ICAT-UNAM
SUPLENTE:	Dr. Alfonso Gasstélum-Strozzy	ICAT-UNAM
SUPLENTE:	Dra. Mariko Nakano Miyatake	ESIME Culhuacán-IPN

Declaración de Autoría

Yo, Fernando Castillo Flores, declaro que la presente tesis titulada: Clasificación automática de puntas de proyectil arqueológicas, y el trabajo presentado en ella, son de mi autoría y, confirmo que:

- Este trabajo fue realizado total, o parcialmente, durante mi estancia doctoral en esta universidad.
- Si alguna parte de esta tesis ha sido previamente utilizada en otro trabajo para la obtención de algún grado o cualquier otro título de ésta Universidad, u otra institución, esto se ha señalado de forma clara y puntual en dicha tesis.
- En las partes donde he consultado otros trabajos públicos, estos siempre han sido reconocidos.
- En las partes donde he citado el trabajo de otros autores, las correspondientes fuentes se proporcionan de forma clara y puntual. Con la excepción de dichas citas, toda esta tesis surge como resultado de mi propio trabajo.
- He reconocido todas las fuentes principales de ayuda.
- En las partes de la tesis que se basan en el trabajo realizado con la colaboración de otras personas, he dado constancia de manera clara la contribución que ha tenido cada una de las partes involucradas.

Firma:



Fecha: Ciudad Universitaria, Cd. Mx. a marzo de 2019

Clasificación automática de puntas de proyectil arqueológicas

Fernando Castillo Flores

Índice general

Declaración de Autoría	III
1. Introducción	1
1.1. Análisis estadístico y clasificación morfológica de puntas de proyectil arqueológicas	1
1.2. Algoritmo de clasificación de puntas de proyectil	4
2. Marco teórico	9
2.1. Clasificación de imágenes y extracción de características	9
2.2. Métodos de clasificación	23
3. Base de datos y desarrollo del algoritmo de clasificación	29
3.1. Base de datos	29
3.2. Interfaz gráfica de usuario	31
3.3. Pre-procesamiento de la imagen	32
3.4. Extracción del mapa de curvatura en el espacio a través de la escala (CSS-map)	33
3.5. Detección de la frontera proximal-distal	35
3.6. Extracción de características de la punta de proyectil; bordes de la hoja	43
3.7. Extracción de características de la punta de proyectil; base	45
3.8. Extracción de características de la punta de proyectil; paredes del pedúnculo	47
3.9. Extracción de características de la punta de proyectil; hombros del pedúnculo	49
4. Experimentos y resultados	51
4.1. Evaluación del desempeño de la clasificación de las componentes de la punta de proyectil	52
4.2. Evaluación del desempeño de la clasificación de la forma global de la punta de proyectil	55
5. Conclusiones	61

DEDICATORIA

A mis padres, hijos, esposa y hermanos

Este trabajo lo hice con mucho cariño pensando en ustedes y también ustedes han sido parte de él.

AGRADECIMIENTOS

Agradezco a Fernando y Margarita, mis padres, por todo lo otorgado a mis hermanos y a mí.

Agradezco a Carolina, mi esposa, por su apoyo incondicional en este trabajo y en cada momento de nuestras vidas.

Agradezco a Camila y Fernando, mis hijos, por todo el apoyo y paciencia que han tenido en éste y todos los demás proyectos que hemos realizado.

Agradezco a Ramón e Ileana, mis hermanos, por todo el apoyo recibido y por compartir tantos momentos conmigo.

Agradezco a la señora Irma Magdalena López, al señor Ventura González y al señor Iván González por todo el apoyo brindado en el ámbito familiar.

Agradezco al doctor Francisco Javier García Ugalde, tutor principal de este proyecto, por todo el apoyo, tanto académico como personal, que ha brindado en esta etapa de mi vida.

Agradezco a los doctores Fernando Arámbula Cosío y Ernesto Bribiesca Correa, miembros de mi comité tutorial, por todo el tiempo, consejos, paciencia y experiencia que amablemente compartieron conmigo durante mi estancia doctoral.

Agradezco al doctor Alfonso Gastelum Strozzi por la valiosa colaboración que otorgó en este proyecto de doctorado.

Agradezco al maestro en arqueología Jesús Filiberto Zarco Navarro por la amable asesoría y colaboración que brindó en el desarrollo de este proyecto.

Agradezco al doctor José Luis Punzo Días, y a todo su equipo de trabajo, por la amable hospitalidad que tuvieron en las visitas que realicé a las instalaciones del INAH-Michoacán y, por todo el tiempo, esfuerzo y colaboración que amablemente brindaron en este proyecto.

Agradezco a la doctora Karina Ruby Pérez Daniel por sus valiosos consejos brindados para el desarrollo de este proyecto.

Agradezco al doctor Rogelio Alcantara Silva por la ayuda y consejos que amablemente ha compartido en estos años de doctorado.

Agradezco al maestro Fernando Lepe Casillas por su valiosa colaboración, la cual, permitió hospedar este proyecto en un dominio *web*.

Agradezco al Proyecto Arqueología y Paisaje del Área Centro Sur de Michoacán y al Centro INAH–Michoacán por el acceso a sus colecciones arqueológicas que amablemente proporcionó a esta investigación.

Agradezco a los propietarios de las bases de datos de imágenes de puntas de proyectil, quienes amablemente autorizaron su uso para esta investigación; esto es: a la Dra. Mercedes Okumura [[Bradley and Okumura., 2016](#)], a Judith Paulos y la asociación *Mount Vernon Ladies Association* [[MVLA, 2018](#)], OverstreetID [[OverstreetID, 2018](#)], RelicShack [[RelicShack, 2018](#)], Ken's Relics [[Kens's Relics, 2018](#)] y J & DEE ARTIFACTS [[J & DEE ARTIFACTS, 2018](#)].

Agradezco al Sistema Nacional De Ciencia y Tecnología (CONACYT) por el apoyo económico que me brindó bajo la modalidad: beca para alumnos de doctorado.

Agradezco a la DGAPA–UNAM por el apoyo económico brindado, en la modalidad “beca para terminación de estudios de doctorado”, durante la realización del proyecto PAPIIT *IN – 106816*.

Lista de Acrónimos

ANN	Artificial Neural Networks
ASIFT	Affine-Scale Invariant Feature Transform
BoVW	Bag of Visual Words
BRIEF	Binary Robust Independent Elementary Features
CLD	Color Layout Descriptor
CSD	Color Structure Descriptor
CSS-map	Curvatura Scale Space Map
DCT	Discrete Cosine Transform
FAST	Features from Accelerated Segment Test
GSIFT	Global-Scale Invariant Feature Transform
HOG	Histograms of Oriented Gradients
HOOSC	Histogram Of Orientation Shape Context
HSV	Hue, Saturation Value
K-NN	K-Nearest Neighbors
LoG	Laplacian of Gaussian
MPEG	Moving Picture Experts Group
PCP	Punto Central Proximal
PFPD	Puntos Frontera Proximal-Distal
PHOG	Pyramid of Histograms of Orientation Gradients
RGB	Red, Green, Blue,
SIFT	Scale Invariant Feature Transform
SPM	Spatial Pyramid Matching
SURF	Speeded-Up Robust Features
SVM	Support Vector Machines

Resumen

Este trabajo presenta la investigación en el área de reconocimiento y clasificación de patrones bidimensionales que he realizado como alumno del Posgrado En Ciencia E Ingeniería De La Computación, de la Universidad Nacional Autónoma de México (UNAM), bajo la tutoría del doctor Francisco Javier García Ugalde y, en colaboración con los grupos arqueológicos dirigidos por el doctor José Luis Punzo Díaz del INAH-Michoacán y por el maestro Jesús Zarco Navarro del posgrado en antropología de la UNAM.

Uno de los procesos más comunes en el ambiente arqueológico, es la clasificación de puntas de proyectil con base a su forma; esto puede vincularse al reconocimiento computacional de patrones. Por otra parte, comúnmente, la clasificación arqueológica de puntas de proyectil se realiza de forma manual, lo que en ocasiones implica que este proceso sea ineficiente, ya que existen cientos de categorías registradas. Para enfrentar este problema hemos desarrollado un algoritmo computacional que realiza la clasificación automatizada de la punta de proyectil con base a su imagen digital. Dicho algoritmo es robusto ante cambios de escala, rotación y calidad de la imagen y, la única condición que presenta es que el fondo de la imagen sea uniforme.

La extracción de características, del algoritmo de clasificación que se ha desarrollado, está basada en el mapa de curvatura a través de la escala y el espacio (CSS-map, *curvatura scale space map*) y de la aplicación del gradiente en regiones específicas del contorno de la punta de proyectil. El entrenamiento y la validación (clasificación) se realizó mediante una base de datos de aproximadamente 1000 imágenes. Además, el extractor de características desarrollado se evaluó con los siguientes métodos de clasificación: árboles de decisión binaria, discriminación lineal, K-vecinos más próximos (KNN, *K-nearest neighbor*) y soporte de maquina vectorial (SVM, *support vector machines*). Los mejores resultados se obtuvieron con el método SVM; así, la precisión y sensibilidad obtenidas al clasificar 420 imágenes, utilizando 630 imágenes para entrenar al sistema (distintas a las 420 utilizadas en la clasificación) fueron de 70 % y 71 %, respectivamente. Por otra parte, el desempeño del algoritmo desarrollado se comparó respecto a otros 3 métodos de clasificación, muy utilizados en la literatura del área, que se basan en la forma de la imagen; esto es: PHOG (*Pyramid Of Histograms Of Orientation Gradients*), HOOSC (*Histogram of Orientation Shape Context*) y momentos invariantes de Hu. Para ello, se clasificaron 9 categorías de las puntas de proyectil, utilizando conjuntos de 45, 90, 180 y 270 imágenes para entrenar y, 315, 270, 180 y 90 imágenes para clasificar. Los resultados obtenidos muestran que el algoritmo desarrollado tiene mejor desempeño que los otros métodos; por ejemplo, al considerar 180 imágenes para entrenar y, 180 para clasificar, la precisión (P) y sensibilidad ($recall$, R) obtenidas para cada uno de los métodos fue: HOOSC $\rightarrow (P = 60\%, R = 58\%)$, Hu $\rightarrow (P = 34\%, R = 32\%)$,

nuestro $\rightarrow (P = 80\%, R = 78\%)$ y PHOG $\rightarrow (P = 73\%, R = 66\%)$.

Finalmente, el algoritmo desarrollado es de fácil acceso para cualquier persona, ya que se implementó una interfaz gráfica de usuario para manipularlo y, ambos (algoritmo e interfaz) pueden ser utilizados a través del portal de internet:

<http://matahari.fi-p.unam.mx/~inge-flechas> [UNAM, 2018]

Capítulo 1

Introducción

1.1. Análisis estadístico y clasificación morfológica de puntas de proyectil arqueológicas

Los artefactos líticos son uno de los elementos más importantes en las investigaciones arqueológicas; ya que, por su propia naturaleza son más resistentes al deterioro provocado por diversos agentes naturales a los que puedan estar expuestos con el paso del tiempo [Sánchez and Carpenter, 2014]. De esta forma, luego de su recuperación y análisis, los arqueólogos pueden entender diversos aspectos de los grupos estudiados como: organización respecto a su elaboración, comportamiento de sus individuos, actividades productivas y grado de desarrollo tecnológico [William Andrefsky, 2005]. Por otra parte, la elaboración de dichos elementos estuvo condicionada, en gran medida, por la materia prima disponible en la región [Sánchez and Carpenter, 2014] y por las personas encargadas de su manufactura (“talladores”) [Arratia and Mirambell, 2005]. A pesar de estas restricciones, son pocos los artefactos líticos que presentan patrones morfológicos regulares; las puntas de proyectil son una de estas pocas excepciones [Valdovinos and Castellanos, 2014], ya que, normalmente eran construidas bajo el mismo proceso tecnológico [Arratia and Mirambell, 2005]. Dichos patrones han sido utilizados por los arqueólogos para clasificar a las puntas de proyectil en categorías conocidas como “tipologías morfológicas” [Hutchings, 2016, Mirambell and Lorenzo, 1974, William Andrefsky, 2005, Rots, 2016].

Existen distintos métodos de análisis asociados a la investigación de puntas de proyectil arqueológicas; algunos de ellos son: métodos estadísticos y de clasificación respecto a la forma, composición química del material, microscopía del desgaste que presenta el artefacto y, recreación en su uso y proceso de manufactura [Rots, 2016, Hutchings, 2016, Rafferty, 2018]. La elección de los métodos a utilizar, además de estar en función de los propios objetivos de la investigación, también lo está de las condiciones de conservación de las puntas de proyectil involucradas, del número de ellas y de los antecedentes de investigación de la región geográfica a la que pertenecen. A pesar de estos factores (variedad de métodos y dependencia en la elección de su uso), los métodos de análisis estadístico y de clasificación con base a la forma, normalmente se utilizan en la mayoría de las investigaciones; ya que, estos permiten sintetizar la información disponible para comprender el fenómeno estudiado. Además, los métodos de clasificación, también permiten realizar una comparación entre distintas colecciones de puntas de proyectil, y obtener diferencias y similitudes entre ellas.

El análisis estadístico de la forma de las puntas de proyectil arqueológicas comúnmente se realiza con base a la técnica conocida como “morfometría geométrica”: *rama de la estadística, que combina herramientas de geometría, computación gráfica y técnicas biométricas para realizar un análisis multivariado de la forma* [Bookstein, 1997]. Así, la morfometría geométrica se realiza respecto a la distribución de un conjunto de puntos conocidos como “puntos de referencia”, o “landmarks”, por su terminología en inglés; los cuales, están situados en distintas posiciones del objeto, normalmente a lo largo de su contorno, pero también en lugares intrínsecamente relevantes al objeto, como el centro de la pupila, si el “objeto” fuera un ojo, o en posiciones de carácter geométrico como el centroide del objeto [Cootes et al., 1995]. Dicho análisis permite conocer parámetros como: la forma promedio de un conjunto de objetos de referencia, el grado de deformación de las partes que conforman el objeto, la conglomeración de los elementos del conjunto estudiado respecto a ciertas variables de interés, etc. A continuación, a manera de ejemplo, se mencionan algunos trabajos arqueológicos que utilizan la morfometría geométrica para el análisis de las puntas de proyectil.

En [Okumura and Araujo, 2014] se utiliza la morfometría geométrica para analizar la evolución de la forma, a través del tiempo, de un conjunto de 94 puntas de proyectil del tipo peduncular pertenecientes al sur de Brasil. Los resultados muestran que los principales cambios en la forma se encuentran en el incremento en las dimensiones del cuello y hombros del pedúnculo, situación que se atribuye a un posible incremento en la población de los grupos étnicos asociados a las puntas de proyectil y/o en errores al reproducir (copiar) estos artefactos. En [Smith and Goebel, 2018] se utiliza la morfometría geométrica para estudiar la correlación que existen entre las puntas de proyectil del tipo “ranurada” (*fluted-point*) vinculadas al norte de Canadá y a la región norte del continente americano que se asocia con la hipótesis arqueológica del “corredor libre de hielo” (“Ice-Free Corridor”), respecto a las puntas de proyectil de la zona sur de Canadá, las “grandes planicies” y el este de Estados Unidos. En [Buchanan et al., 2011] la morfometría geométrica se utiliza para cuestionar la hipótesis arqueológica que sostiene que la forma y el tamaño de las punas de proyectil dependían del tipo de presas para las que eran utilizadas. En dicha investigación se utiliza un conjunto de 74 puntas de proyectil Paleo-indias del tipo “Clovis” y “Folsom”, asociadas a la caza de mamuts y bisontes, las cuales, provienen de las planicies del sur y suroeste de Estados Unidos. Finalmente, en [Serwatka and Riede, 2016] se utiliza la morfometría geométrica para analizar 114 puntas de proyectil pedunculares del final del periodo paleolítico, provenientes de distintas partes del norte de Europa. La investigación pretenden cuestionar las correlaciones culturales y tecnológicas que tradicionalmente se han realizado, con base al estudio morfológico de las puntas de proyectil, entre los grupos ancestrales que habitaron esas regiones; ya que, postula que en los trabajos donde se han realizado dichas correlaciones, los ejemplares utilizados han sido recuperaron a nivel de la superficie y/o no se tiene certeza respecto a los datos reportados en su recuperación.

Por lo general, la metodología en la que se utilizan técnicas de morfometría geométrica, aplicada a las investigaciones de puntas de proyectil, estrictamente no constituye un sistema de clasificación, para el cual, la información analizada debe transferirse y, así, brindar al sistema el conocimiento necesario para implementarlo en otros trabajos de clasificación.

Además de esto, esta técnica presenta algunas desventajas respecto al proceso de clasificación, ya que, para lograr un desempeño satisfactorio, el ambiente de adquisición de la imagen debe ser muy controlado y, normalmente debe implementarse con categorías similares entre sí y rechazando puntas de proyectil que se encuentran dañadas. Así, se ha observado que el método de morfometría geométrica está más enfocados en la descripción estadística de la forma de las puntas de proyectil, que en la clasificación morfológica de éstas. Por otra parte, con base al estudio que se ha realizado del estado del arte y, considerando esta perspectiva computacional de clasificación, se encontró que son pocos los trabajos dedicados al desarrollo de un sistema de clasificación automatizada de puntas de proyectil. Uno de estos trabajos es el de S. Nash y Elton Prewitt [Nash and Prewitt, 2016], en el cual, se utiliza una red neuronal para clasificar a un conjunto de 15 categorías morfológicas. En este trabajo, la red neuronal es alimentada por 8 parámetros descriptivos de la forma de las componentes de la punta de proyectil; sin embargo, una de las principales desventajas es que la extracción de dichas características se realiza de forma manual; esto es, para cada punta de proyectil, cada uno de los 8 parámetros son asignados respecto al criterio de los investigadores; por lo cual, adicionalmente al tiempo que el usuario debe invertir para asignar las características al conjunto a clasificar, esto no descarta la subjetividad que tradicionalmente ha presentado el proceso de clasificación de puntas de proyectil arqueológicas. Otra desventaja que presenta el trabajo, es que el modelo propuesto de clasificación sólo es evaluado en un ámbito muy local, ya que, solamente considera 15 categorías tipológicas, de las cientos que existen y, a pesar de que brinda la posibilidad de incluir más categorías en dicho modelo, el trabajo no señala como se comportaría el desempeño del clasificador al realizar dicha inclusión.

La principal limitación para desarrollar un clasificador genérico de puntas de proyectil arqueológicas, es la gran cantidad de categorías que existen (del orden de cientos), e independientemente de las considerables diferencias que puedan encontrarse entre una base de datos y otra, en algunas de estas bases existen categorías con muy pocos ejemplares. En ocasiones, para resolver estos problemas, presentes en el desarrollo de clasificadores de otros objetos, distintos a las puntas de proyectil, o de otras disciplinas (no arqueológicas), se han implementado técnicas de incremento artificial de datos y, para el caso de redes neuronales profundas, se ha utilizado la técnica de “transferencia de conocimiento” (*fine-tuned*, por su terminología en inglés); sin embargo, al no existir trabajo previo respecto al uso de estas técnicas en el problema de clasificación de puntas de proyectil y, considerando que la acentuación de los problemas antes mencionados es grande, no es posible conjeturar, sin desarrollar trabajo al respecto, la viabilidad del uso de estas técnicas para enfrentar dicha limitación. Por otra parte, algunas de las técnicas de incremento artificial de los datos, utilizadas, por ejemplo en imágenes de escenas naturales, no deben ser utilizadas en la clasificación de puntas de proyectil; ejemplo de ello es el cambio de excentricidad, ya que, desde el punto de vista arqueológico, esta característica es discriminante. Finalmente, la clasificación de puntas de proyectil tradicionalmente se ha enfrentado a otros problemas; algunos de ellos son: bases de datos con algunos de sus ejemplares dañados, o con imprecisión respecto a la categoría a la que pertenecen, similitud acentuada entre algunas categorías y, al no existir un consenso que busque estandarizar su uso, también hay duplicidad entre categorías.

1.2. Algoritmo de clasificación de puntas de proyectil

La principal contribución de este trabajo de tesis fue el algoritmo desarrollado para la clasificación de puntas de proyectil arqueológicas. Además, durante el proceso de investigación se creó una colaboración importante entre el área arqueológica y el área computacional; misma que, al utilizar herramientas de análisis de imagen, aprendizaje de máquina, reconocimiento de patrones y visión computacional, ha permitido enfrentar de manera distinta, de como lo ha hecho la comunidad arqueológica mediante el uso de la computadora, los problemas asociados a la clasificación de puntas de proyectil. Específicamente, el diseño del algoritmo de clasificación se basa en la separación automatizada de la punta de proyectil en sus componentes principales; esto es, base, hoja, hombros y paredes del pedúnculo. Esta metodología ha permitido abarcar un rango muy amplio de categorías morfológicas registradas por los arqueólogos y, así, enfrentar el problema asociado a la gran cantidad que existen de ellas, del orden de cientos, y que en algunas no hay suficientes ejemplares para entrenar al sistema. Este enfoque también permitió enfrentar el problema asociado a la gran similitud que existe entre algunas categorías, pues al elegir sólo las componentes de la punta de proyectil a clasificar, el algoritmo puede despreocuparse de la influencia de la forma global y, sólo considerar aquellos detalles que diferencian estas categorías. Por último, dicho enfoque benefició el desempeño del clasificador al permitir diseñar descriptores específicos para cada una de las componentes en las que se dividió la punta de proyectil (base, hoja, hombros y paredes del pedúnculo). Otra contribución de este trabajo de tesis es la accesibilidad que el algoritmo de clasificación desarrollado presenta para los usuarios, este objetivo se planteó en consideración especial a la comunidad arqueológica; así, dicho algoritmo puede ser ejecutado por un equipo de cómputo con rendimiento promedio, bajo una licencia de uso libre para fines de investigación y es de fácil manipulación mediante una interfaz gráfica de usuario. A continuación se resumen estas características y aportaciones del algoritmo, al presentar los objetivos del trabajo de tesis.

Objetivo general: *Desarrollar un algoritmo computacional para realizar la clasificación automatizada de imágenes de puntas de proyectil arqueológicas con base a su forma. Dicho algoritmo debe enfrentar los problemas actuales que presenta este proceso, además de ser accesible para cualquier usuario, tanto por la sencillez que presente en su manipulación, como por los requerimientos de alimentación del sistema. Finalmente, debe tener un amplio espectro de aplicación y, así, desempeñarse satisfactoriamente en la mayoría de bases de datos arqueológicas.*

Objetivos particulares

- *Realizar una investigación en las áreas de análisis de imagen y visión computacional para determinar aquellos métodos de clasificación, basados en la forma de los objetos en la imagen, que permitan desarrollar de manera óptima, con base a la adaptación e integración de los mismos, el algoritmo de clasificación mencionado en el objetivo general del trabajo.*

- Comparar el desempeño del algoritmo desarrollado con otros clasificadores, basados en la forma de los objetos de la imagen, reportados en la literatura.
- Desarrollar una interfaz gráfica de usuario que permita manipular el algoritmo desarrollado de forma sencilla.
- Crear un vínculo con la comunidad arqueológica para obtener asesoramiento en el desarrollo del proyecto y validación de resultados.

Considerando que la separación de la punta de proyectil en sus componentes principales es la premisa clave que se ha planteado en este trabajo de clasificación, la propuesta para definir dichas componentes se basó en el estudio del estado del arte y en la asesoría de los grupos arqueológicos que colaboraron en esta investigación, dirigidos por el doctor José Luis Punzo Días y por el maestro Jesús Zarco Navarro; dichas componentes son: bordes de la hoja, base, hombros y paredes laterales del pedúnculo; la Figura 1.1 denota cada una de ellas. Por otra parte, la regularidad de los patrones de forma que se encuentran para los distintos tipos de puntas de proyectil, es lo que permite al algoritmo detectar de manera automatizada cada una de estas componentes. Así, las partes del contorno que presentan un alto valor en la curvatura, y cambios de orientación pronunciados en ella, son la principal característica para realizar dicha detección. Dado que el mapeo de curvatura en el espacio a través de la escala (CSS-map, *Curvature Scale Space Map*) [Abbassi et al., 1999] permite detectar estos cambios, se propuso implementar a éste y, al gradiente del contorno de la punta de proyectil, como principales métodos para realizar esta tarea (Secciones 3.4 y 3.5).

La Figura 1.2 muestra, de forma idealizada, las categorías propuestas de cada una de las componentes de la punta de proyectil. La combinación de dichas categoría genera 400 formas distintas de puntas de proyectil, las cuales, generan varios detalles considerados por los arqueólogos como aletas, barbas y zonas de enmangue [William Andrefsky, 2005]. Además, esta base (400 formas de punta de proyectil), puede utilizarse en investigaciones más específicas que incluyen otras características como tamaño del artefacto, excentricidad y grado de simetría. Por otra parte, en dicha figura se puede observar que cada una de las componentes tiene categorías similares, por ejemplo, en el caso de las categorías “incurvados” y “recurvados salientes–entrantes” de la componente bordes de la hoja, o en el caso de las categorías “expandidas” y “divergentes” de la componente paredes laterales del pedúnculo. Dicha similitud es más evidente al considerar imágenes “reales” de puntas de proyectil, respecto a las imágenes “ideales” de la Figura 1.2. En este sentido, en la base de datos que se formó para entrenar y evaluar al clasificador, se decidió incluir este tipo de imágenes (no muy definidas respecto a su categoría) con la finalidad de no “sobre–ajustar” el entrenamiento del sistema y, de esta manera cumplir con el objetivo de asistir al arqueólogo en la situación antes mencionada. En la Sección 3.1 se proporcionan los detalles de la implementación de dicha base de datos.

Dado que la finalidad de la base de datos es tener suficientes elementos para entrenar y evaluar el desempeño del clasificador, y considerando la disponibilidad de los elementos de cada una de las categorías, experimentalmente se llegó a la conclusión de que cada categoría

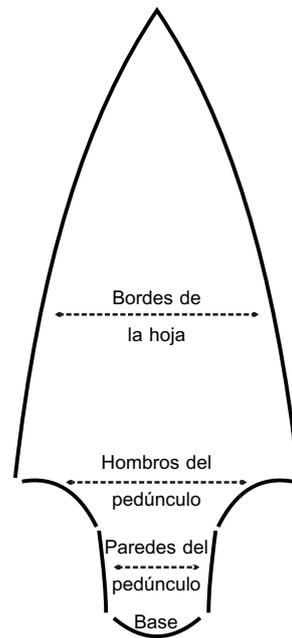


Figura 1.1: Componentes de la punta de proyectil.

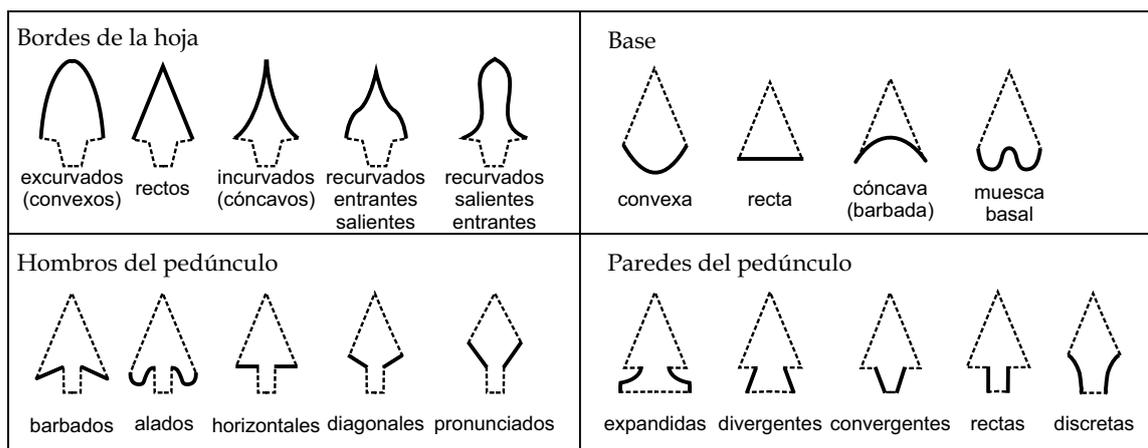


Figura 1.2: Categorías de cada una de las cuatro componentes de la punta de proyectil.

debería estar conformada por al menos 50 imágenes. De esta forma, la parte arqueológica del equipo formó cada una de estas categorías con imágenes tomadas de sus propias colecciones y otra parte importante con imágenes disponibles en internet (previamente autorizadas para su uso en este proyecto); así, se logró formar una base de datos de aproximadamente 1000 imágenes. Haber utilizado las imágenes de internet contribuyó en gran medida a obtener una evaluación objetiva del algoritmo, ya que, debido a la gran variedad de características de estas imágenes, se tiene un escenario similar al que enfrentaría el clasificador al desempeñarse en distintas bases de datos arqueológicas. Además, esto ofrece una referencia importante de las capacidades del algoritmo desarrollado en este trabajo, pues la mayoría de estas imágenes son de menor calidad que las encontradas en bases de datos arqueológicas oficiales.

Desde el punto de vista de la visión computacional, para lograr que el algoritmo desarrollado se aproveche en la gran mayoría de las bases de datos arqueológicas, éste se diseñó de tal manera que fuera invariante ante la escala, y robusto ante la rotación y calidad de la imagen.

Para evaluar el desempeño del algoritmo se propusieron dos criterios, por un lado, se consideraron las componentes de la punta de proyectil de manera individual, lo cual, permitió tener una evaluación intrínseca del algoritmo. Por otro lado, se consideró la clasificación de la forma global de la punta de proyectil, ya que esto ofrece a los arqueólogos un indicativo directo de su desempeño y, además, brinda la posibilidad de comparar al algoritmo con otros descriptores y clasificadores de forma conocidos en la literatura científica, como PHOG (*Pyramid Histogram of Oriented Gradients* [Lazebnik et al., 2006]), HOOSC (*Histogram of Orientation Shape Context*) [Gülcan et al., 2015] y momentos geométricos invariantes de Hu [Hu, 1962]. En el caso de los algoritmos HOOSC y PHOG, la extracción de características se realizó con el método “bolsa de palabras” (BoW, *Bag of Words*) [Sivic and Zisserman, 2003, Krapac et al., 2011, Lazebnik et al., 2006], ya que, éste entrega buenos resultados al combinarlo con este tipo de descriptores (gradientes locales). Para desarrollar el segundo criterio de evaluación, el grupo arqueológico que colaboró en esta investigación propuso 9 categorías de forma global de puntas de proyectil, dichas categorías fueron elegidas con base a su relevancia arqueológica y a la disponibilidad de imágenes en la base de datos. Por último, con la finalidad de tener mayor certeza en los resultados presentados, se propuso diseñar el proceso de evaluación (para ambos criterios) de tal manera que éste fuera iterado en 10 ocasiones y, eligiendo en cada una de ellas, de manera aleatoria, el conjunto de imágenes para entrenar y el resto para clasificar. De esta forma, los resultados presentados en este trabajo surgen de promediar los resultados obtenidos en cada una de dichas iteraciones. Por otra parte, en el segundo criterio de evaluación, el conjunto de entrenamiento en cada iteración debe ser el mismo para los cuatro clasificadores a comparar (HOOSC, Hu, PHOG y el propuesto en esta investigación).

La razón por la cual el método de HOOSC fue elegido para la comparación, es que al ser diseñado para clasificar jeroglíficos mayas, lo convierte en un descriptor generado en un contexto arqueológico y, esto permite tener un buen contraste respecto al descriptor que se ha generado en este trabajo. Por otra parte, el método de PHOG fue incluido porque es uno

de los descriptores locales con base a la forma más representativos y con mejor desempeño. Por último, se incluyó al método de Hu ya que éste es un algoritmo clásico de clasificación de forma. Finalmente, para determinar que método utilizar para entrenar (y clasificar) a estos cuatro descriptores, se ha propuesto realizar pruebas sobre varios clasificadores reportados en la literatura como: soporte de maquina vectorial (SVM, *support vector machines*), K vecinos más próximos (K-NN, *K-Nearest Neighbor*), fronteras de decisión polinómica y árboles de decisión binaria.

Capítulo 2

Marco teórico

2.1. Clasificación de imágenes y extracción de características

Dado que, desde una perspectiva computacional, el tema principal de esta investigación es la clasificación automatizada de imágenes, en este capítulo se explicará en que consiste este proceso y, se presentarán los métodos de extracción de características, y de clasificación, más comunes en la literatura científica del área, así como también aquellos que han sido utilizados en este trabajo.

Para comenzar con la explicación formal, se presenta la siguiente definición de *proceso de clasificación*:

Definición 2.1 [Goodfellow et al., 2016] Sea el conjunto finito $M = \{m_1, m_2, \dots, m_i, \dots\}$, en donde cada uno de sus elementos, m_i , representan un “objeto” que será clasificado en k categorías del conjunto $\{C = c_1, c_2, \dots, c_N\}$; $1 \leq k < N$. Por otra parte, \mathbf{x}_i es un vector n -dimensional que representa las características extraídas del elemento m_i . De esta forma, se define la clasificación del elemento m_i como aquella función $f(\mathbf{x}_i)$ que asocia dicho elemento con k categorías del conjunto C ; esto es: $f : \mathbb{R}^n \rightarrow C^k$.

Respecto a la definición anterior, cuando $k = 1$, la clasificación se denomina uni-categorica, ya que sólo es asociada una categoría a cada elemento m_i . Por otra parte, cuando $k > 1$, la clasificación se denomina multi-categorica, e implica que algunos de los elementos m_i pueden ser asociados con más de una categoría; por ejemplo: para cierta aplicación, algún auto puede ser clasificado en las categorías: auto, deportivo y mercedes.

El diseño de la función f , o el ajuste de sus parámetros, si ésta ha sido previamente definida, se puede realizar con base a alguna de las siguientes dos metodologías: 1) clasificación supervisada y, 2) clasificación no supervisada. La diferencia entre ambas es que en la primera se utiliza un conjunto de elementos pre-clasificados, conocido como “conjunto de entrenamiento”, con el cual, es posible optimizar el desempeño de f al minimizar el error de clasificación de dicho conjunto; este error se calcula de acuerdo a cierta métrica que involucra a la categoría pre-asignada respecto a la categoría que el clasificador asignó. Por otra parte, en la clasificación no supervisada, el diseño de f , o el ajuste de sus parámetros, se realiza con base a la partición del espacio de características, con el método k -medias, u otro similar, de los vectores de características de cierto conjunto de elementos no pre-clasificados. El desempeño de la clasificación no supervisada es más bajo que el desempeño de la clasificación supervisada, sin embargo, la clasificación no supervisada en ocasiones se

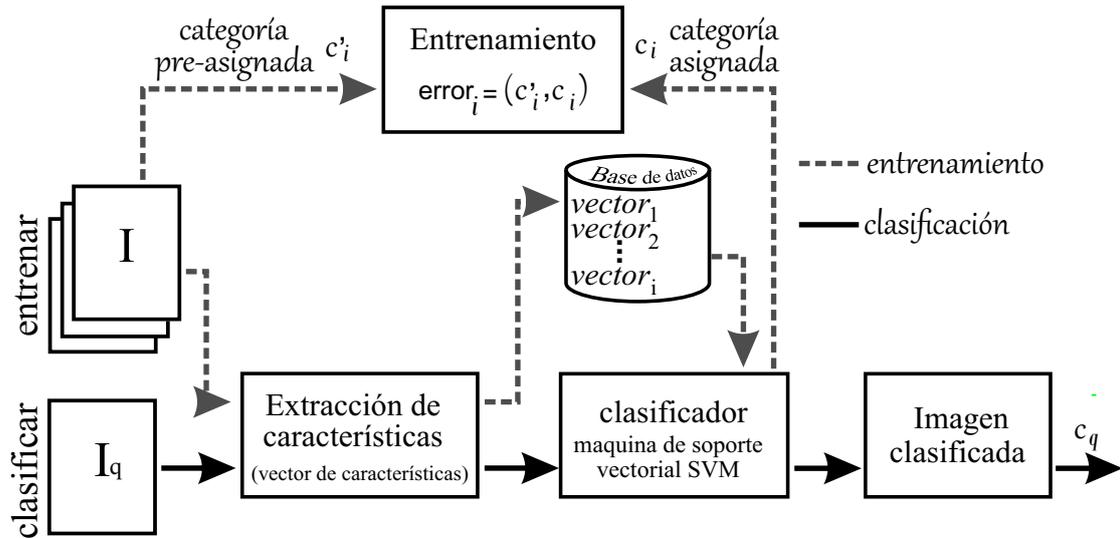


Figura 2.1: Diagrama del proceso de clasificación implementado en este trabajo. $c'_i, c_i, c_q \in C$; en donde, C es el conjunto con las categorías de clasificación.

utiliza porque no siempre se dispone del conjunto de entrenamiento, o porque en algunos sistemas, la clasificación se realiza en “tiempo real” reajustando los parámetros del clasificador de acuerdo a los datos obtenidos secuencialmente.

La Figura 2.1 presenta el diagrama de clasificación implementado en este trabajo (unicategorica y supervisada); como se puede observar en dicho diagrama, las principales etapas del proceso son, por una parte, el entrenamiento del clasificador (en este caso se utilizó un clasificador de maquina de soporte vectorial, SVM), y por otra parte, el proceso de clasificación de alguna imagen de consulta. Además, también se puede observar que la extracción de características es un proceso común para ambas etapas. A continuación se describe cada una de estas etapas y sus métodos más comunes. De acuerdo a la definición del proceso de clasificación (Definición 2.1), en la primera etapa se extraen características del objeto a clasificar. Dichas características deben discriminar de la mejor forma a las categorías involucradas en la clasificación; por ejemplo, si se desea clasificar un conjunto de estudiantes de nivel básico de acuerdo al grado escolar al que pertenecen, la estatura del individuo sería una buena característica en esta tarea. Por otra parte, al considerar imágenes digitales y, dependiendo del campo de aplicación del clasificador, es preferible que las características sean invariantes, o al menos robustas, ante distintas transformaciones como: escala, rotación, traslación, iluminación y ligeros cambios de perspectiva.

Los métodos computacionales que extraen las características de una imagen digital son conocidos como descriptores de imagen¹. Existen distintos tipos de ellos, por lo cual, es común agruparlos respecto al tipo de características que extraen. Los tipos de descriptores más comunes son aquellos que extraen características de color, forma, textura y gradientes locales. A continuación se explicará cada uno de ellos, dando mayor énfasis a los descriptores de forma y gradiente locales, ya que de esta clase de métodos son los que se utilizaron para

¹Dado que en este trabajo sólo se utilizarán imágenes como elementos a clasificar, se hará referencia a los descriptores de imagen simplemente como “descriptores”.

comparar con el método desarrollado en este trabajo.

Descriptores de color. El color es una de las características más importante para el sistema de visión humana, por lo cual, este tipo de descriptores ha sido ampliamente utilizado en tareas de clasificación de imágenes, reconocimiento de patrones, indexación, recuperación de contenido, etc. Normalmente, estos descriptores ofrecen un buen equilibrio entre los resultados obtenidos, y el uso de recursos computacionales, además, gozan de un alto grado de independencia, ya que no dependen, o lo hacen en menor medida, del pre-procesamiento de la imagen. Finalmente, pueden ser combinados con otros descriptores, produciendo un descriptor híbrido, para obtener mayor robustez en la clasificación. En general, se considera que este tipo de descriptores tienen un enfoque de extracción de características globales de la imagen; esto, pese a que en algunos de ellos se incorpora información local espacial de la imagen; sin embargo, dicha información no es selectiva; ejemplo de ello son el descriptor de distribución de color (CLD, *Color Layout Descriptor*) [Yamada and Kasutani, 2001], el descriptor de estructura de color (CSD, *Color Structure Descriptor*) [Manjunath, 2001], el correlograma de color [Huang et al., 1997], etc, o algún método en donde no se considera exclusivamente un clasificador de color, sino una combinación de éste con un detector local, como es el caso del descriptor color-SIFT (*Scale Invariant Feature Transform*) [Abdel-Hakim and Farag, 2006].

Existe un gran número de descriptores de este tipo, por lo cual, el estándar MPEG-7 [Manjunath et al., 2002], en su apartado de descriptores visuales, es una buena opción para iniciar con el estudio de ellos. Por otra parte, debido a que los descriptores de color no fueron incluidos en este trabajo, a continuación, sólo se describen las características de algunos de ellos y, posteriormente, se explica de manera más detallada al histograma y al correlograma de color, ya que estos son descriptores muy representativos de esta categoría.

- **Color dominante** [Manjunath, 2001]. Este descriptor simplifica el contenido de color al cuantizar cada pixel en un máximo de 8 categorías respecto al espacio de color de la imagen; por ejemplo, para una imagen que pertenece al espacio de color RGB, la cuantización se realizará de acuerdo a un conjunto de vectores en tres dimensiones. De esta forma, el vector de características se obtiene como la concatenación de valores $\{q_i, h_i, \sigma_i\}$; en donde, $i = 1, 2, \dots, N$; N representa el número de categorías que cuantizan el espacio de color. q_i es el i -ésimo vector de cuantización, h_i es el valor del histograma normalizado en la posición i y, σ_i es la varianza de h_i .
- **Scalable color** [Manjunath, 2001]. Este descriptor se forma con el histograma de color de la imagen en el espacio de color HSV, el cual, es codificado con la transformada de Haar [Gonzalez and Woods, 2002] para obtener una representación binaria de dicho histograma. Presenta la ventaja de consumir pocos recursos computacionales al comparar los vectores de características de distintas imágenes; además, permite seleccionar el grado de precisión con el que se desea representar la imagen en el vector binario. Sin

embargo, al ser un descriptor basado sólo en el histograma de color, presenta algunas limitaciones que se comentarán al explicar dicho método (histograma de color).

- **Distribución de color (CLD, Color Layout Descriptor)** [Yamada and Kasutani, 2001]. Este descriptor además de considerar al color, adiciona información de la distribución espacial de éste en la imagen; sin embargo, presenta la desventaja, respecto a descriptores locales, que dicha información no es selectiva. Presenta la ventaja de ser un método que requiere pocos recursos computacionales. Su implementación se lleva a cabo al codificar a la imagen en 64 bloques de igual tamaño, los cuales, se forman al dividir el largo y ancho de la imagen en 8 secciones. Para cada uno de estos bloques se extrae el color representativo; el cual, de acuerdo al estándar MPEG-7 [Manjunath, 2001], se define como el promedio del color de los pixeles que conforman el bloque; además, en dicho estándar se propone trabajar con el espacio de luminancia-crominancia YC_bC_r , por lo cual, para cada uno de los planos de color de dicho espacio, se debe extraer el promedio de la intensidad de los pixeles dentro del bloque. La transformada discreta del coseno [Lathi, 1998] (DCT, *Discrete Cosine Transform*) se calcula de acuerdo a los 64 coeficientes de cada plano de color, obteniendo así los tres planos en el dominio de la frecuencia. Finalmente, los coeficientes de frecuencia de cada plano son asignados al vector de características, leídos en forma de "zig-zag" y, comenzando en la esquina superior izquierda (frecuencias bajas). El estándar MPEG-7 recomienda tomar los 6 primeros coeficientes de frecuencia (leídos en "zig-zag"), del plano de luminancia Y y los 3 primeros, también leídos en "zig-zag", de cada uno de los planos de crominancia C_b y C_r ; obteniendo así un vector de características de $6 + 3 + 3 = 12$ dimensiones.

Se puede demostrar que el histograma de color [Swain and Ballard, 1991] y el correlograma de color [Huang et al., 1997] son invariantes ante cambios de escala, rotación y traslación; además, el correlograma también es ligeramente robusto ante oclusiones parciales de los objetos y, ligeros cambios de escena y perspectiva; sin embargo, el costo computacional de éste es ligeramente mayor².

El histograma de color (h) se calcula con base a la frecuencia de la cuantización de los pixeles de la imagen; por simplicidad, considere una imagen representada en un espacio monocromático (tonos de gris) con un rango dinámico de $[0, 255]$, a la cual, se desea extraer el histograma de tonos de gris con 25 recipientes s_1, s_2, \dots, s_{25} ; de esta forma, el histograma evaluado en el recipiente s_1 es igual al número de pixeles de la imagen en un rango de $[0, 9]$ unidades, el histograma evaluado en el recipiente s_2 es igual al número de pixeles en un rango de $[10, 19]$ unidades, etc. Por otra parte, el histograma debe normalizarse en un rango de $[0, 1]$ para que éste pueda ser utilizado como una función de densidad de probabilidad y, además, sea invariante ante cambios de escala. La Figura 2.2 muestra una imagen en tonos de gris (rango dinámico de $[0, 255]$) y su correspondiente histograma (100 recipientes); en ella se puede observar que la tela es de color blanco y el decorado en ella es de color negro;

²La complejidad del histograma de color es $O(n)$, mientras que la del correlograma es $O(dn)$; en donde n es el número de pixeles en la imagen y d es un parámetro del correlograma menor o igual a la magnitud del lado más pequeño de la imagen; d normalmente se define en el orden de las unidades.

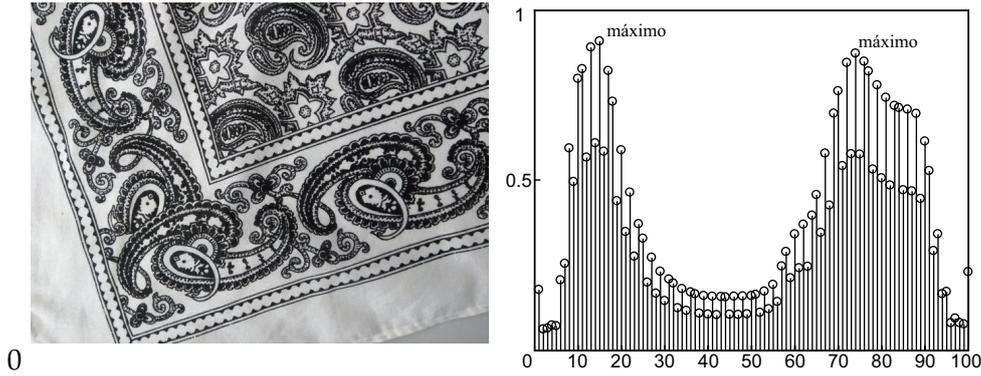


Figura 2.2: La parte izquierda presenta una imagen en tonos de gris con un rango dinámico de $[0, 255]$, mientras que, en la parte derecha se presenta el histograma de dicha imagen al considerar 100 recipientes.

de esta forma, el histograma tiene un máximo ubicado en las primeras unidades del dominio (en el recipiente número 15) y otro en las unidades finales (en el recipiente número 75); los cuales, corresponden al tono del decorado y de la tela, respectivamente.

Dado que el histograma de color no incluye información espacial de la imagen, éste no discrimina adecuadamente imágenes con porcentajes de color similares, pero con una distribución espacial diferente. El correlograma de color [Huang et al., 1997] enfrenta este problema al incluir, en el histograma de color, información espacial con base a la distancia que existe entre píxeles de distintos colores. La Figura 2.3 presenta un ejemplo de esta situación, en ella se pueden observar dos imágenes distintas pero con porcentajes de color similares, por lo cual, sus histogramas de color son muy parecidos entre sí, por otra parte, en el gráfico inferior se observa que el correlograma de ambos difiere considerablemente.

El correlograma de color, γ , de la imagen I , es un conjunto de matrices que se calculan con base a la frecuencia en que los píxeles cuantizados en los valores s_1, s_2, \dots, s_N , se encuentran a una distancia $d = \{d_1, d_2, d_{max}\}$ uno de otro; así, el elemento d_k del conjunto d está asociado con la k -ésima matriz del correlograma; $d_{max} \leq \min(\text{largo}(I), \text{ancho}(I))$.

Por ejemplo, para alguna distancia d_k , $M_{i,j}^{d_k}$ representa el número de píxeles cuantizados al valor s_j y que se encuentran a d_k unidades de distancia (norma L_∞) de algún píxel cuantizado al valor s_i ; de esta forma, la k -ésima matriz se define como:

$$\gamma_{i,j}^{d_k} = \frac{M_{i,j}^{d_k}}{h(s_i) \cdot 8d_k} \quad (2.1)$$

en donde, $h(s_i)$ es el histograma (no normalizado) de la imagen evaluado en el valor s_i y, el denominador de la fracción representa el número de píxeles a una distancia d_k (donde el factor $8d_k$ se debe a la norma L_∞) de cualquier píxel cuantizado al valor s_i .

Descriptores de forma. La forma es otra de las características más importantes para el sistema de visión humana y para la clasificación de imágenes. Estos descriptores se dividen en dos principales grupos; los que extraen las características con base al contorno del objeto y,

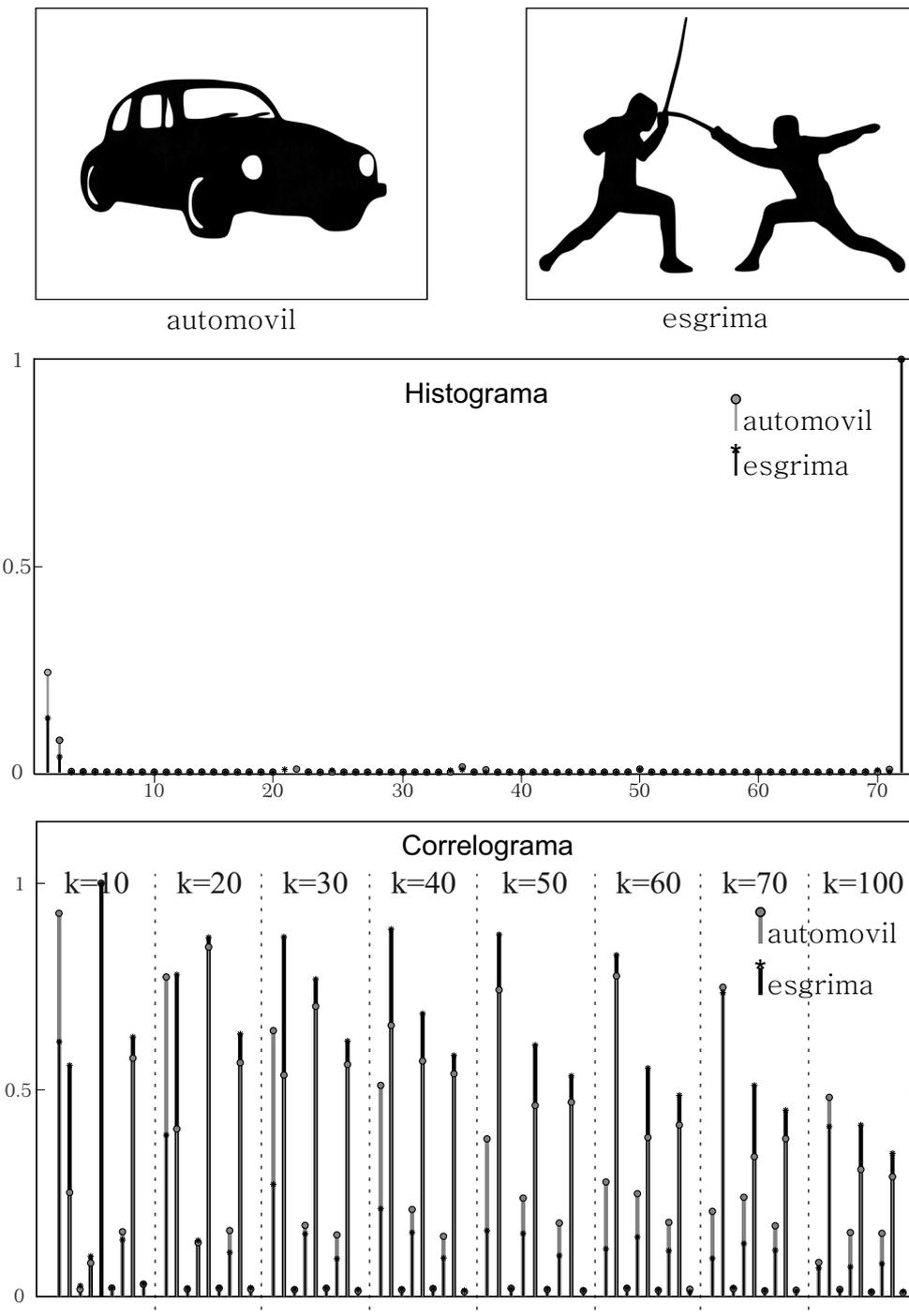


Figura 2.3: Histograma y correlograma de color para dos imágenes distintas entre sí, pero similares en sus porcentajes de color. Como se puede observar, el histograma no es adecuado para discriminar este tipo de imágenes; por otra parte, el correlograma sí es capaz de realizar dicha discriminación.

aquellos que extraen las características con base a la región interna de éste. Algunos de los descriptores más comunes del primer grupo son los códigos de cadena [Freeman and Saghri, 1978, Bribiesca, 1999], el método de aproximación poligonal de la forma [Gu, 1995], los modelos autorregresivos [Kauppinen et al., 1995], los descriptores de Fourier [Persoon and sun Fu, 1997, Zahn and Roskies, 1972] y el mapeo en el espacio, de la curvatura a través de la escala (CSS-map, *Curvature Space Scale map*) [Mokhtarian and Mackworth, 1986, Abbassi et al., 1999]. Por otra parte, varios de los descriptores más comunes del segundo grupo son: algún tipo de momento, ya sea estadístico o de otro tipo, como los momentos geométricos de Hu [Hu, 1962], los momentos de Zernike [Teague, 1980, Mehtre et al., 1997], los momentos de Legendré y/o los momentos de Krawtchouk; además, existen otros descriptores de este grupo como el muestreo de rejillas [Lu and Sajjanhar, 1999], o la transformación armónica polar [Yap et al., 2010].

A continuación se explicará como obtener los momentos geométricos de Hu, ya que estos se utilizaron para comparar su desempeño con el desempeño del clasificador desarrollado en este trabajo. Por otra parte, el mapeo en el espacio, de la curvatura a través de la escala (CSS-map) se explicará detalladamente en el Capítulo 2, ya que, parte del desarrollo del clasificador de puntas de proyectil se basa en este método.

Los momentos de Hu [Hu, 1962] se componen de 7 medidas estadísticas que son invariantes ante escala, rotación y traslación. Estos se obtienen con base a los momentos geométricos centrados normalizados; por lo cual, primero serán definidos dichos momentos.

El momento geométrico de orden (p, q) de la imagen $I(x, y)$ se define como:

$$m_{p,q} = \sum_{x=1}^M \sum_{y=1}^N I(x, y) \cdot x^p \cdot y^q \quad (2.2)$$

en donde, $p, q = 0, 1, \dots$. Mientras que, M y N son, el ancho y largo de la imagen, respectivamente. Por otra parte, si en dichos momentos se considera al centroide del objeto (o de la escena en la imagen) como el origen de coordenadas, estos serán invariantes ante traslaciones; así, se definen los momentos centrales como:

$$\mu_{p,q} = \sum_{x=1}^M \sum_{y=1}^N I(x, y) \cdot (x - \bar{x})^p \cdot (y - \bar{y})^q \quad (2.3)$$

en donde, (\bar{x}, \bar{y}) representa el centroide del objeto; esto es:

$$\bar{x} = m_{1,0} / m_{0,0}$$

$$\bar{y} = m_{0,1} / m_{0,0}$$

Los momentos centrados y normalizados respecto a la escala se definen como:

$$\bar{\mu}_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^k} \quad (2.4)$$

en donde, $k = (p + q)/2$ y $p + q \geq 2$.

Finalmente, los momentos de Hu se definen con base a $\bar{\mu}_{p,q}$, como:

$$\begin{aligned}
 H_1 &= \bar{\mu}_{2,0} + \bar{\mu}_{0,2} \\
 H_2 &= (\bar{\mu}_{2,0} - \bar{\mu}_{0,2})^2 + 4\bar{\mu}_{1,1} \\
 H_3 &= (\bar{\mu}_{3,0} - 3\bar{\mu}_{1,2})^2 + (3\bar{\mu}_{2,1} - \bar{\mu}_{0,3})^2 \\
 H_4 &= (\bar{\mu}_{3,0} + \bar{\mu}_{1,2})^2 + (\bar{\mu}_{2,1} + \bar{\mu}_{0,3})^2 \\
 H_5 &= (\bar{\mu}_{3,0} - 3\bar{\mu}_{1,2})(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})[(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})^2 - 3(\bar{\mu}_{2,1} + 3\bar{\mu}_{0,3})^2] + \\
 &\quad (3\bar{\mu}_{2,1} - \bar{\mu}_{0,3})(\bar{\mu}_{2,1} + \bar{\mu}_{0,3})[3(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})^2 - (\bar{\mu}_{2,1} + 3\bar{\mu}_{0,3})^2] \\
 H_6 &= (\bar{\mu}_{2,0} - \bar{\mu}_{0,2})[(\bar{\mu}_{3,0} + \bar{\mu}_{1,2}) - (\bar{\mu}_{2,1} + \bar{\mu}_{0,3})^2 + 4\bar{\mu}_{1,1}(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})(\bar{\mu}_{2,1} + \bar{\mu}_{0,3})] \\
 H_7 &= (3\bar{\mu}_{2,1} - \bar{\mu}_{0,3})(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})[(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})^2 - 3(\bar{\mu}_{2,1} + 3\bar{\mu}_{0,3})^2] + \\
 &\quad (3\bar{\mu}_{1,2} - \bar{\mu}_{3,0})(\bar{\mu}_{2,1} + \bar{\mu}_{0,3})[3(\bar{\mu}_{3,0} + \bar{\mu}_{1,2})^2 - (\bar{\mu}_{2,1} + 3\bar{\mu}_{0,3})^2]
 \end{aligned} \tag{2.5}$$

Detectores de puntos característicos y extractores de características con base a gradientes locales. Actualmente este tipo de descriptores son ampliamente utilizados en problemas de clasificación, segmentación, detección y reconocimiento de objetos en imágenes y vídeo. Se conforman de dos etapas: detección de puntos característicos y extracción de características; además, se consideran descriptores locales, ya que la extracción de características se realiza sólo en la vecindad de los puntos característicos, lo cual, proporciona robustez ante la oclusión del objeto a clasificar y/o cambios drásticos de escenario (fondo) en la imagen. Finalmente, también son robustos ante rotación, traslación, cambios de iluminación y, algunos de ellos, ante el cambio de escala.

Específicamente, los puntos característicos son “regiones relevantes” de la imagen; dependiendo del descriptor (detector), este término, regiones relevantes, adquiere sentido; por ejemplo, para detectores como el de Harris [Harris and Stephens, 1988], Harris–Laplace [Mikolajczyk and Schmid, 2004], Hessian–Laplace [Mikolajczyk and Schmid, 2004] y FAST [Rosten and Drummond, 2006] dicho término es asociado con las esquinas de los objetos en la imagen, por otra parte, para detectores del tipo estructuras de líneas y curvas (*curvilinear structure detector*) [Steger, 1998], estas regiones se enfocan en sus bordes; mientras que, para detectores como SIFT [Lowe, 1999], SURF [Bay et al., 2008], KAZE [Alcantarilla et al., 2012] y BRIEF [Calonder et al., 2010] éstas representan zonas que contrastan con su entorno (“manchones”).

Considerando su relevancia histórica, el detector de esquinas de Harris [Harris and Stephens, 1988] fue el primero en ser publicado. Se basa en el cálculo de los valores propios de la matriz de segundos momentos del gradiente de la imagen. La mayor limitación que enfrentó este descriptor, es su no invarianza ante el cambio de escala; sin embargo, la propuesta de superar este problema fue uno de los principales factores que motivó el desarrollo de este tipo de descriptores. Tony Lindeberg contribuyó en dicho objetivo al desarrollar una metodología para la descomposición multi–escala de la imagen, la cual, utilizó en el diseño de un algoritmo para la detección de puntos característicos a través de la escala [Lindeberg, 1998]. Este algoritmo se basa en la detección de puntos críticos (máximos y mínimos) de la función resultante de la convolución de la imagen con la laplaciana de una gaussiana (LoG). Por otra parte, Mickolajczyk y Schmid, con base a los resultados de esa investigación, adaptaron el

detector de Harris para brindarle invarianza ante el cambio de escala y, así, desarrollaron dos descriptores conocidos como Harris–Laplace y Hessian–Laplace [Mikolajczyk and Schmid, 2001].

La detección multi–escala de puntos característicos, además de permitir que el descriptor sea invariante ante ella (escala), ofrece una forma de organizar y discrimina la información en la imagen. La Figura 2.4 proporciona un ejemplo de ello; en ella se puede observar una imagen a distintas escalas etiquetadas del 0 al 4³. Para cada una de las imágenes, los puntos característicos se han delimitado con círculos verdes. En estos círculos, además de realizar dicha delimitación, se asocia su tamaño (radio) con la escala del elemento que delimitan; esto de acuerdo al detector SIFT, el cual, ha sido utilizado para generar esta figura. Por otra parte, para la escala 1 (primera octava), se observa que los puntos característicos corresponden a detalles pequeños, en tamaño, como las fichas que debe comer “Pacman” (*Pac-Dots*). Conforme la escala aumenta, estos detalles pierden relevancia, volviéndose progresivamente difusos hasta desaparecer y, surgen detalles de mayor tamaño; por ejemplo, observe que en la imagen que corresponde a la cuarta escala ya no es posible distinguir dichas fichas, pero los “fantasmas” (personajes que en la imagen original se muestran en color rojo, anaranjado, rosa y cian) se vuelven relevantes al aparecer como un “manchón” delimitado por el punto característico.

Además de los descriptores mencionados anteriormente, varios más de este conjunto, han sido desarrollados (vea por ejemplo: [Wu et al., 2017, Wu et al., 2013, Hietanen et al., 2016]); sin embargo, explicar a algunos de ellos está fuera del objetivo de este trabajo, por lo cual, a continuación sólo se describirán los aspectos fundamentales del descriptor SIFT, y algunas de sus adaptaciones, ya que éste es el más representativo y, los descriptores PHOG y HOOSC están basados en él. Este par de descriptores (PHOG y HOOSC), en conjunción con la metodología para el desarrollo de vectores de características conocida como “bolsa de palabras visuales” (BoVW, “*Bag of Visual Words*”), serán explicados también, ya que fueron utilizados para comparar la eficiencia obtenida con dichos métodos, respecto a la obtenida con el clasificador generado en esta investigación.

Como se dijo anteriormente, el descriptor SIFT (*Scale-invariant feature transform*) [Lowe, 1999, Lowe, 2004] es el más emblemático de esta categoría; desde su publicación en 1999, ha sido ampliamente utilizado en diversas actividades como registro de imagen, visión estereográfica, indexación, clasificación, etc; además, se han realizado varias adaptaciones de él y, ha servido como modelo para muchos otros descriptores de este tipo. Como referencia a su importancia, se menciona que de acuerdo a una consulta realizada en Google Scholar, el artículo [Lowe, 2004] aparece en 47000 citas de publicaciones científicas a la fecha en que se escribió este texto (agosto del 2018). Este descriptor presenta la ventaja de ser invariante ante cambios de escala, rotación y traslación; además de ser robusto ante cambios de iluminación y ligeros cambios de perspectiva. Por otra parte, está formado por el par de etapas que se presentan en estos métodos; esto es: detección de puntos característicos y extracción

³Estrictamente cada una de estas etiquetas (y cada una de las imágenes de la figura) no representan una sola escala, sino un conjunto de ellas conocidas como octava [Lowe, 1999]; de esta forma, cada octava tiene asociadas 4 escalas y, cada una de las 4 imágenes, corresponde a la primera escala de cada octava.



Figura 2.4: Ejemplo de la descomposición multi-escala de una imagen. El cuadro superior izquierdo corresponde a la imagen a escala original. El cuadro a la derecha de éste, corresponde a la imagen escalada por el factor más pequeño de los cuatro considerados. El cuadro inferior izquierdo representa la imagen escalada a un factor mayor que el anterior y así sucesivamente. Cada escala corresponde al inicio de una octava de acuerdo al método de SIFT [Lowe, 2004]. Finalmente, para cada una de las cuatro imágenes escaladas se han delimitado en color verde los puntos característicos detectados con dicho método.

de características. La detección de puntos característicos se realiza con base a la descomposición multi-escala de la imagen. Para ello, se convoluciona a la imagen con una función gaussiana $g(x, y, \sigma)$ a la que se incrementa progresivamente su escala (σ), generando así un proceso iterativo. De acuerdo al trabajo de Lowe [Lowe, 1999, Lowe, 2004], la resta de un par de estas imágenes convolucionadas, una después de la otra respecto al dominio de la escala, representa una aproximación a la convolución de la imagen con la función laplaciana de la gaussiana (*LoG*) en dicha escala, la cual, como se comentó anteriormente, es la base para la detección de los puntos característicos en la imagen. Dichos puntos tienen asociada una posición en 3 coordenadas (2 espaciales y una de escala) y una orientación. La orientación proporciona invarianza ante la rotación y, se calcula de acuerdo a la imagen convolucionada con aquella gaussiana con la escala más cercana a la escala del punto característico. Así, se extrae la magnitud y dirección del gradiente de los píxeles de esta imagen y, con ellos, se forma un histograma de 36 recipientes que representan la cuantización de la dirección de dichos gradientes; la contribución de cada píxel al histograma es ponderada por la magnitud de su gradiente y por una función gaussiana de varianza igual a 1.5 veces la escala del punto característico y, de media igual a 0, evaluada en la distancia que hay del punto característico al píxel considerado. Finalmente, la orientación del punto característico es asignada respecto al pico más alto del histograma. En caso de existir picos de altura mayor al 80 % que este pico, estos también son considerados como orientaciones, por lo cual, un punto característico determinado en la primera etapa del proceso de detección, puede convertirse en dos, o más, puntos característicos al final de dicho proceso; los cuales, tendrán la misma ubicación (espacio y escala) pero distinta orientación.

El vector de características de cada punto característico es calculado de acuerdo al gradiente de los píxeles alrededor de él, los cuales, pertenecen a la imagen convolucionada con la gaussiana de escala más cercana a la escala de dicho punto. Para ello se considera una vecindad de 16×16 píxeles, dividida en bloques de 4×4 (píxeles) y, para cada uno de estos bloques, se calcula el histograma de la orientación de los gradientes considerando 8 recipientes de cuantización; además, la contribución de cada píxel al histograma es ponderada por la magnitud de su gradiente y por una función gaussiana de varianza igual a 8 (un medio del ancho de la vecindad del descriptor) evaluada en la distancia que hay del punto característico al píxel considerado. Así, el vector de características está formado como la concatenación de estos 16 histogramas, lo que da lugar a un vector de $16 \times 8 = 128$ dimensiones. Por otra parte, para que el descriptor sea invariante ante la rotación, las coordenadas de la vecindad y la dirección de los gradientes son normalizados de acuerdo a la orientación del punto característico. Finalmente, al normalizar al vector de características se asigna invarianza ante cambios de iluminación lineales en la imagen y, para dar robustez ante cambios de iluminación no lineales, el histograma normalizado es sometido a un proceso de umbralización para sustituir la altura de los picos mayores a 0.2 unidades, por dicho valor (0.2); después de esto, el histograma es normalizado una vez más.

Han surgido varias modificaciones, o adaptaciones, al algoritmo de SIFT (vea por ejemplo: [Wu et al., 2013]), las cuales, pretenden complementarlo en otros aspectos como color,

textura, cambios considerables de perspectiva en la escena y rapidez de cómputo. A continuación se describen algunas de estas modificaciones.

- **ASIFT (Affine-SIFT)** [Morel and Yu, 2009]. Como su nombre lo indica, este algoritmo proporciona invarianza ante transformaciones affine de la escena que forma la imagen. Para implementarlo se realizan una serie de estas transformaciones y se aplica el algoritmo de SIFT a cada una de las imágenes transformadas. Dado que la mayoría de los dispositivos de captura de imagen, como cámaras fotográficas, no operan bajo un modelo de transformación affine, sino geométrico, se podría considerar que el contexto de aplicación del método es reducido; sin embargo, aún con este tipo de imágenes se obtiene buena robustez ante cambios drásticos de perspectiva. Por otra parte, la razón del costo-beneficio obtenida con el método es aceptable al considerar que el tiempo de cómputo es entre tres y cuatro veces mayor que en el caso de SIFT.
- **GSIFT (Global-SIFT)** [Mortensen et al., 2005]. Este método permite enfrentar el problema de ambigüedad que surge al comparar los puntos característicos de dos, o más, imágenes, tales que, en cada una de ellas existen varias zonas similares; por ejemplo, las que se encuentran en objetos texturados, o con un patrón espacial claramente definido, como en el caso del tablero de ajedrez. El método añade a cada vector de características, computado con SIFT, información global con base a un histograma polar de distribución radial logarítmica [Belongie et al., 2002], de los valores de curvatura de los píxeles en la imagen; en donde, el radio del conjunto de histogramas más alejado del punto característico es igual a la magnitud de la diagonal de la imagen.

De acuerdo a este tipo de descriptores (gradientes locales), el reconocimiento de un objeto, o escena, en un par de imágenes, se realiza mediante la comparación de sus puntos característicos, lo cual, permite formar parejas con aquellos que su distancia está por debajo de cierto umbral; posteriormente, dicha comparación es validada con base a una transformación geométrica de las parejas formadas [Lowe, 2004]. Sin embargo, para el caso de la clasificación de imágenes, este método no es adecuado, ya que normalmente en dicho proceso no se considera un “objeto intrínseco” (en distintos escenarios o bajo distintas perspectivas) para formar una categoría, sino que, debido a la propia definición de categoría, ésta se compone de “objetos”, tales que, pueden ser intrínsecamente diferentes, pero comparten la esencia que los identifica respecto a otros “objetos”; por ejemplo, esta diferencia se puede apreciar al considerar el “objeto intrínseco” auto azul con placas XYH1Z y, en la categoría “auto”, como aquella formada por varias imágenes de automóviles tomadas en las calles de cierta ciudad. De esta forma, en el proceso de clasificación, a diferencia del proceso de reconocimiento, al tener distintos objetos (intrínseco) que forman una categoría, debe considerarse un método que, a partir de los vectores locales de características, obtenga los patrones globales de la imagen para discriminar las distintas categorías. El método conocido como “bolsa de palabras visuales” (*BoVW*, *Bag of Visual Words*) [Sivic and Zisserman, 2003, Krapac et al., 2011, Lazebnik et al., 2006] ha sido ampliamente utilizado para esta tarea. Este método está formado por tres etapas: 1) extracción de características (descriptores locales), 2) codificación de palabras y, 3) fusión de palabras (*pooling*). Implícitamente, la primera etapa (extracción

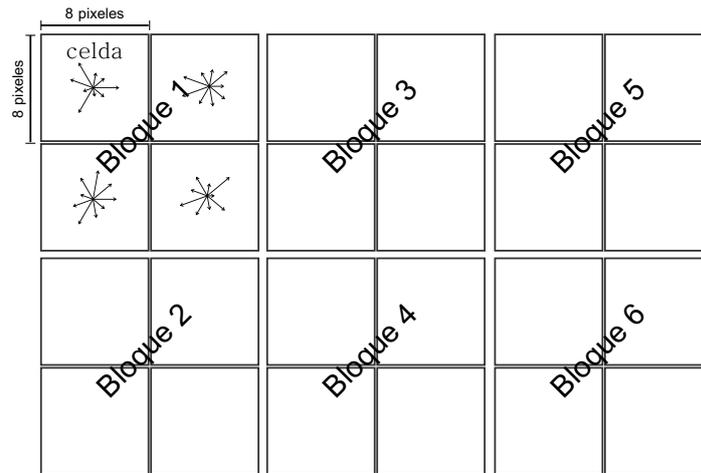


Figura 2.5: Diagrama del proceso de extracción de características con el algoritmo HOG.

de características) ya se ha comentado en esta sección; así, no se repetirá dicha explicación y se aprovechará el espacio generado para comentar los dos descriptores (HOG y HOOSC) utilizados en esta investigación para realizar la comparación con el clasificador desarrollado. Posterior, se explicarán de forma secuencial las dos etapas restantes (codificación y fusión de palabras).

El descriptor HOG (*Histograms of Oriented Gradients*) [Dalal and Triggs, 2005] es quien extrae las características en el método de clasificación PHOG (*Pyramid Histogram of Oriented Gradients*) [Lazebnik et al., 2006]). Este descriptor (HOG) no realiza la detección de puntos característicos, por lo cual, estos se establecen con base a una rejilla que divide la imagen en un conjunto de celdas cuadradas. En cada celda se calcula el histograma de la dirección del gradiente de los píxeles que la conforman, así, se obtiene un vector de características al concatenar los histogramas de un grupo de $K \times K$ celdas llamado bloque; la Figura 2.5 ejemplifica este proceso. Opcionalmente, los bloques se pueden traslapar para restar relevancia al cambio drástico en la dirección del gradiente entre bloques contiguos. De acuerdo a lo explicado anteriormente, este descriptor tiene los siguientes parámetros: longitud del lado de la celda, número de celdas laterales que conforman un bloque y número de recipientes del histograma; al implementar este método en el presente trabajo, estos parámetros se fijaron como 8 píxeles, 2 celdas y 9 recipientes, respectivamente.

El descriptor HOOSC [Gülcan et al., 2015] también se utilizó para comparar al clasificador desarrollado en este trabajo. En dicho descriptor cada vector de características se extrae con base a la dirección de los gradientes de los píxeles dentro de un conjunto de celdas agrupadas de forma polar en anillos concéntricos respecto a las posiciones de cierta rejilla; además, los radios de un conjunto de anillos surgen al seccionar equidistantemente el radio del mayor anillo entre el número total de anillos. De forma similar a HOG, se calculan los histogramas de direcciones y estos son concatenados; la Figura 2.6 ejemplifica este proceso. Los parámetros de este descriptor son el número de recipientes del histograma, el número de celdas en las que se divide un anillo, el número de anillos, el tamaño de la rejilla y el radio del anillo de mayor tamaño (del conjunto de anillos concéntricos); al implementar

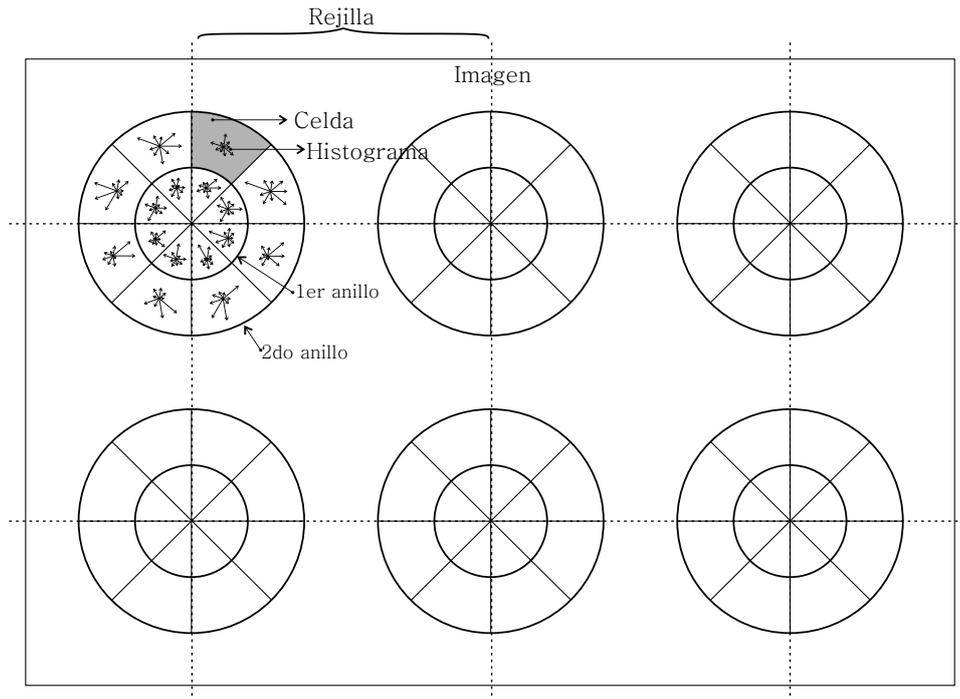


Figura 2.6: Diagrama del proceso de extracción de características con el algoritmo HOOSC.

este método en el presente trabajo, estos parámetros se fijaron, respectivamente, como: 8 recipientes, 8 celdas por anillo, 2 anillos, 20 píxeles para el largo de la rejilla y un parámetro variable para el tamaño del radio mayor (Sección 4.2).

Como se dijo anteriormente, en la segunda etapa del método BoVW cada vector de características es representado por una “palabra”, la cual, pertenece a algún diccionario. Para formar el diccionario se utiliza la conglomeración de los vectores de características de un conjunto de imágenes de entrenamiento; de esta forma, el diccionario contiene, de forma abstracta, los patrones de dichos vectores; la Figura 2.7 muestra el diagrama de este proceso. Dicha conglomeración se puede realizar con distintos métodos como: cuantización rígida, vectores de Fisher [Perronnin et al., 2010], super-codificación de vectores (*super vector encoding*) [Zhou et al., 2010] y codificación tipo núcleo-libro (*kernel codebook encoding*) [van Gemert et al., 2008]; en este trabajo se utilizó la cuantización rígida.

Finalmente, en el proceso de fusión o agrupamiento, mejor conocido como “*pooling*” por su terminología en inglés, el conjunto de palabras visuales de una imágenes son “fusionadas” en un solo vector de características que contiene la información global y, el cual, se utilizará en los procesos de entrenamiento y clasificación. Existen distintos métodos de fusión de palabras visuales, el histograma es uno de los más comunes por su sencillez y aceptable desempeño. A pesar de las ventajas de este método, en ocasiones es necesario mejorar los resultados obtenidos. La mayoría de alternativas que se han propuesto para esta tarea, consisten en incluir información espacial de las palabras. Uno de los métodos más comunes, y utilizado en este trabajo, es la comparación espacial en forma de pirámide (SPM, *Spatial Pyramid Matching*) [Lazebnik et al., 2006], este método concatena los histogramas

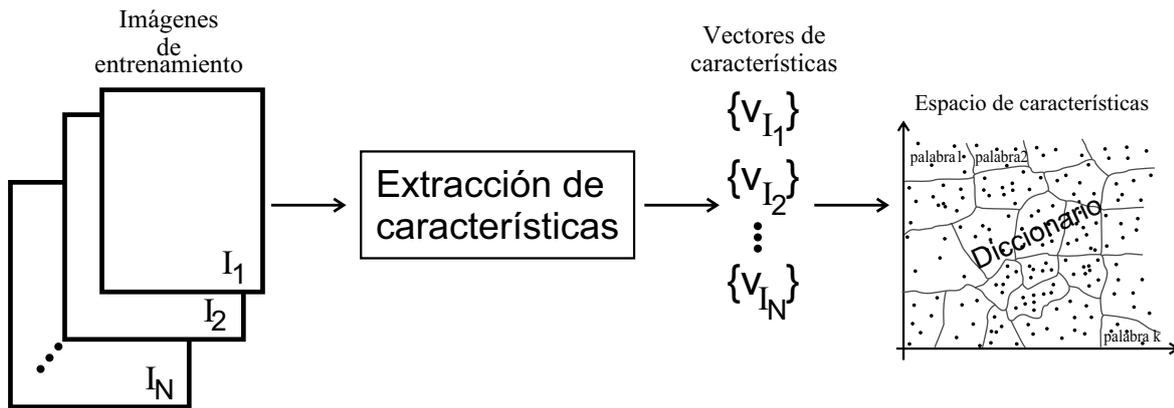


Figura 2.7: Diagrama que ejemplifica la construcción de un diccionario de k palabras con base a los vectores de características de N imágenes de entrenamiento.

de palabras extraídos sobre las regiones que se generan al particionar iteradamente la imagen formando $2^n \times 2^n$ rectángulos, donde n es el nivel de la pirámide; esto es: para $n = 0$, el histograma se calcula sobre toda la imagen (no hay partición), para $n = 1$, se divide la imagen a la mitad del largo y del ancho, generando así 4 rectángulos, para $n = 2$, el largo y el ancho se dividen en 4 partes (generando 16 rectángulos) y, así sucesivamente.

2.2. Métodos de clasificación

A continuación se describen algunos de los métodos de clasificación más comunes.

K vecinos más próximos (K-NN, K-nearest neighbors). Este es uno de los métodos de clasificación más sencillos. Consiste en asignar, al vector de características de la imagen a clasificar, los K vectores de características más cercanos a él (vecinos), respecto al conjunto de vectores de características de las imágenes de entrenamiento; de esta forma, la imagen se clasificará en la categoría de mayor frecuencia dentro del conjunto de categorías asociadas a dichos K vectores. Para medir la cercanía entre vectores se utiliza alguna métrica como la distancia euclidiana, tablero de ajedrez (*chessboard*), Manhattan (*city-block*), Mahalano-bis, etc. Por ejemplo, para clasificar al vector \mathbf{x}_i ; el cual, consta de n componentes; esto es: $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^n]$, considerando la norma euclidiana y el conjunto de l vectores de entrenamiento $\{\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_l\}$; en donde, $\mathbf{t}_j = [t_j^1, t_j^2, \dots, t_j^n]$ y $1 \leq j \leq l$, deben calcularse las distancias:

$$d(\mathbf{x}_i, \mathbf{t}_j) = \sqrt{\sum_{s=1}^n (x_i^s - t_j^s)^2} \quad (2.6)$$

Así, la clasificación se realizará con base a las categorías de los vectores \mathbf{t}_j que involucran a las K menores distancias del conjunto $d(\mathbf{x}_i, \mathbf{t}_j)$.

Los diagramas de Voronoi se pueden considerar como una consecuencia de este método cuando $K = 1$; estos son muy utilizados en aplicaciones de clasificación, optimización de procesos y toma de decisiones. Además, son especialmente útiles cuando el espacio de

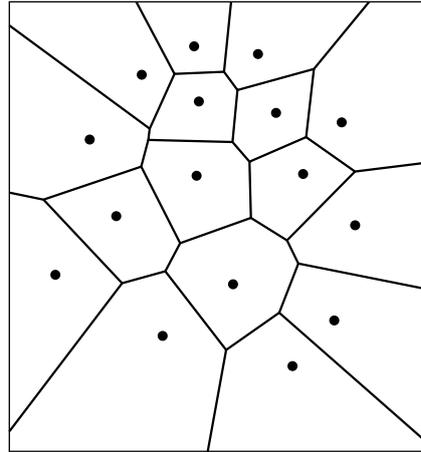


Figura 2.8: Ejemplo de diagrama de Voronoi.

características es de dimensión 2, el cual, normalmente es vinculado a las coordenadas espaciales de un objeto, o lugar. Así, el espacio coordinado se divide en regiones, o celdas, que contienen a un solo vector de características, en donde, la distancia de cualquier posición dentro de la celda a la coordenada del vector de características contenido en ella, es menor que al considerar la coordenada de cualquier otro vector de entrenamiento. La Figura 2.8 presenta un ejemplo de diagrama de Voronoi.

Clasificador Bayesiano. Como su nombre lo indica, este clasificador, al estar basado en el teorema de Bayes, tiene un enfoque estadístico. De esta forma, la clasificación se basa en la probabilidad condicionada de cada una de las clases c_j , respecto al vector de características \mathbf{x}_i , del i -ésimo elemento a clasificar, ésta es conocida como probabilidad *a-posteriori*; esto es:

$$P(c_j|\mathbf{x}_i) = \frac{p(\mathbf{x}_i|c_j)P(c_j)}{p(\mathbf{x}_i)} \quad (2.7)$$

El factor $P(c_j)$, llamado probabilidad *a-priori*, indica la probabilidad intrínseca de cada una de las clases (qué tan probable es la presencia de cada clase en el sistema). Por otra parte, $p(\mathbf{x}_i|c_j)$ es la densidad de probabilidad condicional del vector \mathbf{x}_i , de componentes aleatorias, tal que su distribución está condicionada por el estado natural del sistema (por las clases). Tanto $P(c_j)$ como $p(\mathbf{x}_i|c_j)$ pueden estimarse con base al conjunto de entrenamiento, aunque éste no siempre resulta ser un proceso sencillo o preciso. Por otra parte, $p(\mathbf{x}_i)$ se define como:

$$p(\mathbf{x}_i) = \sum_{j=1}^l p(\mathbf{x}_i|c_j)P(c_j) \quad (2.8)$$

donde l es el número de clases. $p(\mathbf{x}_i)$ se convierte en un término constante al definir \mathbf{x}_i , por lo cual, no interviene en la clasificación y, sólo representa un factor de normalización a la unidad.

Por lo general, este tipo de clasificadores ofrecen buenos resultados y costos computacionales muy bajos, ya que, una vez que se ha determinado $P(c_j)$ y $p(\mathbf{x}_i|c_j)$, sólo debe realizarse

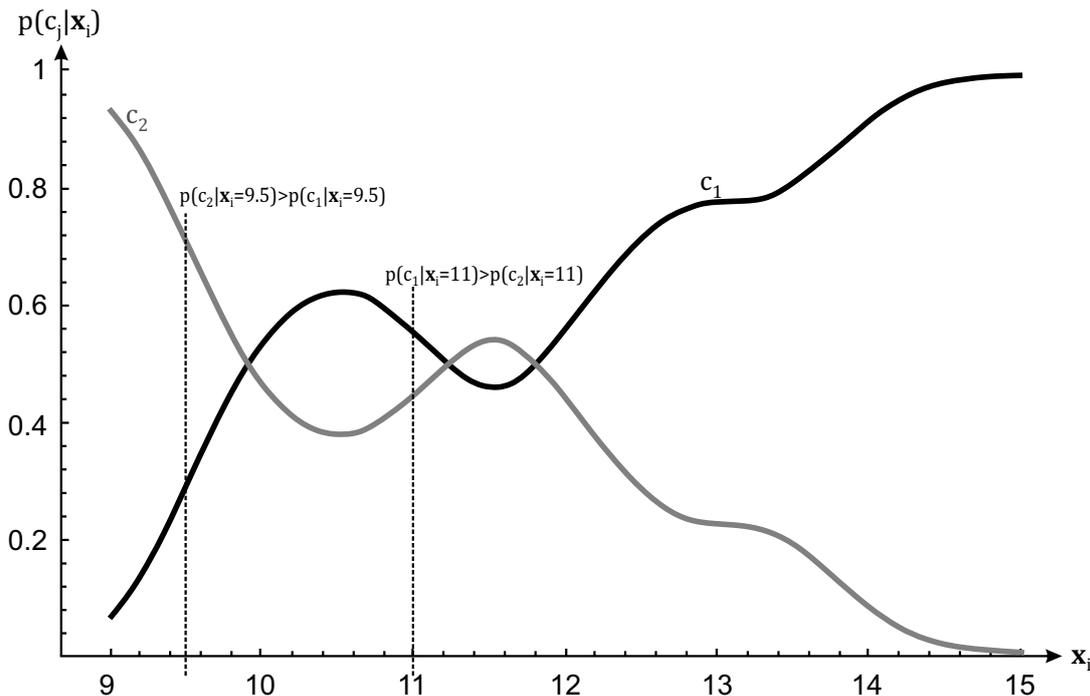


Figura 2.9: Ejemplo de un clasificador de Bayes de un sistema con dos categorías. En este caso, $P(c_1) = 2/3$ y $P(c_2) = 1/3$. Imagen y ejemplo obtenidos de [Duda et al., 2000].

el producto entre ellos, obteniendo así los valores $P(c_j|\mathbf{x}_i)$, tales que, el mayor de ellos es vinculado con la categoría asignada. Sin embargo, su uso no es tan común, porque, como se comentó anteriormente, no siempre es sencillo, o posible, determinar las probabilidades $P(c_j)$ y $p(\mathbf{x}_i|c_j)$. En general es un método recomendado cuando el vector de características está conformado por pocos elementos, del orden de unidades y, preferentemente cuando éstas representan características directas de los objetos a clasificar, como dimensiones, luminosidad, etc.

La Figura 2.9 [Duda et al., 2000] muestra un ejemplo de un clasificador bayesiano representado con base a las gráficas de $P(c_j|\mathbf{x}_i)$ (probabilidad *a-posteriori*). En dicho sistema, el vector de características (\mathbf{x}_i) es de una sola componente, además, sólo intervienen dos categorías ($j = 1, 2$), $P(c_1) = 2/3$ y $P(c_2) = 1/3$ y, el par de densidades de probabilidad condicionada $p(\mathbf{x}_i|c_j)$ son representadas en la gráfica de la figura. En dicho sistema, si se considera, por ejemplo, $x_i = 9.5$ (como se indica en las gráficas de $P(c_j|\mathbf{x}_i)$ con la línea punteada vertical), el elemento es clasificado en la segunda categoría (c_2), ya que, $P(c_2|\mathbf{x}_i = 9.5) > P(c_1|\mathbf{x}_i = 9.5)$. Por otra parte, si se considerara, por ejemplo, $\mathbf{x}_i = 11$, en ese caso el elemento es clasificado en la primera categoría. Finalmente, en las gráficas de probabilidades $P(c_j|\mathbf{x}_i)$ se observa que la suma de ambas, para todo valor del dominio es igual a la unidad; esto de acuerdo a los axiomas de definición de probabilidad.

Árboles de decisión binaria. En este tipo de estructuras la clasificación se realiza al recorrer al árbol en forma descendente (del nodo raíz hasta alguna hoja); dicho recorrido se basa

en el resultado lógico de la indagación que cada nodo realiza a alguna de las componentes del vector de características del elemento a clasificar. La Figura 2.9 muestra un árbol binario utilizado para clasificar un sistema de 4 categorías y , vectores de características de 5 componentes; en ella se puede observar que la componente del vector de características considerada al iniciar la clasificación (en el nodo raíz) es x_3 , en donde, la indagación en el nodo implica que si x_3 es menor que 0.3514, el flujo de la clasificación se dirige hacia el nodo inferior del lado izquierdo, el cual, es una hoja del árbol que representa a la categoría c_3 , así, el proceso de clasificación termina en ese momento asignando dicha categoría al elemento a clasificar. Por otra parte, si x_3 no es menor que 0.3514, el flujo debe dirigirse hacia el nodo inferior del lado derecho y, el proceso continuará hasta llegar a una hoja del árbol.

Este tipo de clasificadores ofrece varias ventajas, por ejemplo, el tiempo de procesamiento en la clasificación es bajo (dicho proceso sólo depende de preguntas lógicas), permite la interpretación directa del proceso de clasificación y de las propias categorías clasificadas y, el experto en el tema puede involucrarse directamente en el diseño del clasificador. Adicionalmente, las características utilizadas, además de cuantitativas, también pueden ser cualitativas; aunque, normalmente, el diseño de aquellos árboles que utilizan características cuantitativas es más preciso, pues se disponen de métodos para optimizar la demanda de recursos computacionales y la precisión de la clasificación. Adicionalmente, al tratarse de características cualitativas, normalmente deben iterarse varios modelos de árboles para elegir aquel que mejor resultados presente. Por otra parte, los árboles de decisión no necesariamente deben ser binarios, sin embargo, normalmente se prefiere a estos porque, son más sencillos de interpretar y, siempre puede obtenerse una equivalencia en forma de árbol binario de cualquier otro que no lo sea [Duda et al., 2000].

Los árboles de decisión pertenecen a la categoría de “clasificadores supervisados”, de esta forma, su diseño se realiza con base a un conjunto de entrenamiento y, a un parámetro conocido como “impureza”. Este parámetro mide que tan homogénea es la distribución de categorías dentro de un nodo; idealmente, conforme se desciende en los niveles del árbol, todas las categorías dentro de un nodo deben ser iguales, con lo cual, se obtendría una impureza igual a cero. Sin embargo, en la mayoría de problemas reales de clasificación esta característica (impureza igual a cero) no se presenta; por lo cual, la labor consiste en asociar, a cada nodo del árbol, la componente del espacio de características y la indagación lógica de tal forma que la impureza total del sistema de clasificación sea mínima; existen varios métodos para realizar este proceso, siendo el descenso del gradiente uno de los más comunes. Además, existen también varias métrica para medir la impureza, sin embargo, en la mayoría de los casos se obtienen valores muy similares con cualquiera de ellas, por lo cual, en este trabajo sólo se presentará la definición basada en la entropía, ya que, ésta es una de las más comunes:

$$i(N) = - \sum_j P(c_j) \log_2 P(c_j) \quad (2.9)$$

en donde, N se refiere al N -ésimo nodo, c_j es el conjunto de categorías, y $P(c_j)$ es la razón del número de elementos en el nodo N asignados (pre-clasificados) a la categoría c_j .

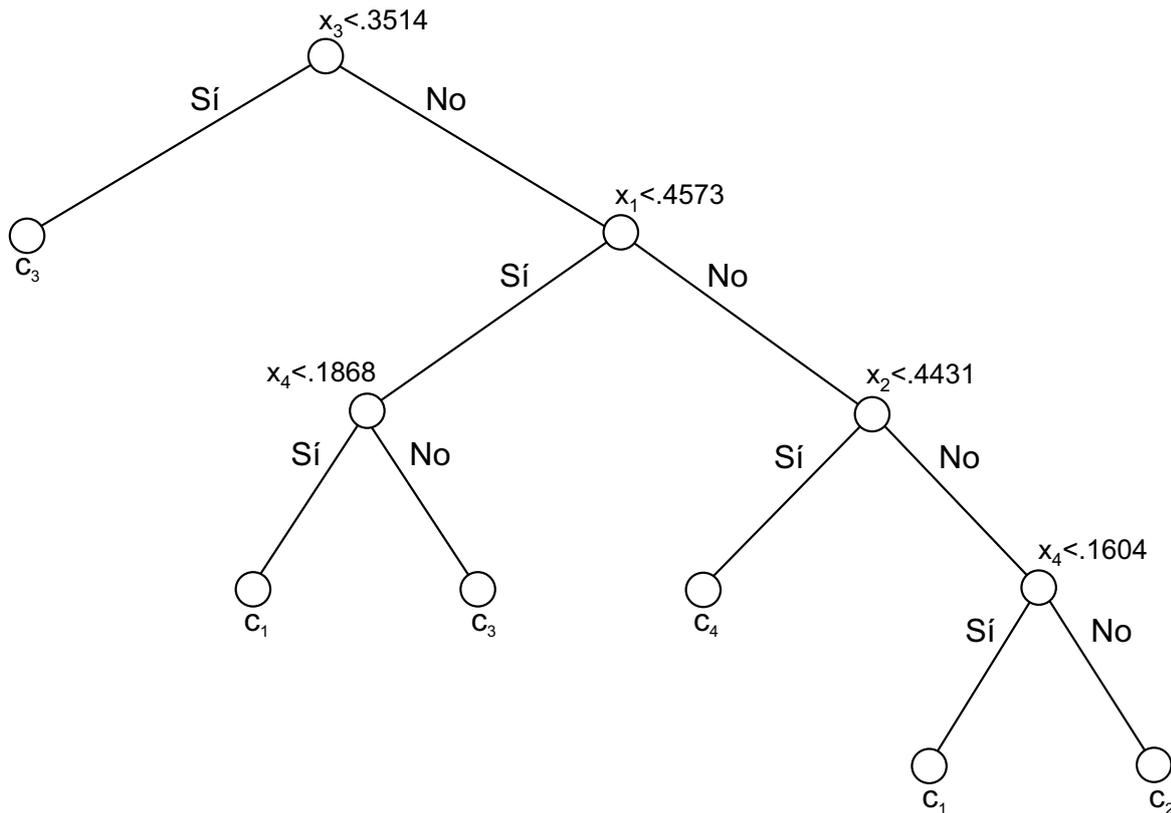


Figura 2.10: Ejemplo de clasificador en forma de árbol de decisión binaria. El sistema de clasificación está conformado por 5 categorías y un espacio de características de 4 componentes.

Redes neuronales artificiales (ANN, *artificial neural networks*). Actualmente, las redes neuronales artificiales (ANN), se utilizan en varias tareas computacionales como: clasificación de caracteres, diagnósticos médicos, reconocimiento de rostro, reconocimiento de voz, etc. Las redes neuronales artificiales son métodos de clasificación supervisados, con los cuales, normalmente se obtienen buenos resultados; sin embargo, necesitan ser entrenados con bases de datos mucho mayores que las utilizadas en otros métodos de clasificación y, normalmente los equipos utilizados para su implementación deben tener una buena capacidad de cómputo, ya que consumen una gran cantidad de recursos de procesamiento. Además de los buenos resultados que se obtienen con este tipo de clasificadores, otra de sus ventajas es que no necesariamente se debe desarrollar un extractor de características intrínseco al sistema a clasificar, sino que es posible, y común, extraer la información con alguna operación matemática; por ejemplo, para el caso de imágenes, esta operación puede ser la convolución, la transformada de Fourier, indicadores estadísticos, etc. Esto implica que sea suministrada una gran cantidad de información a la red neuronal, de la cual, parte de ella (de la información suministrada) no necesariamente tiene una interpretación directa para los humanos. De esta forma, es la propia red neuronal durante la etapa de entrenamiento, quien discrimina dicha información, de acuerdo a su relevancia en el problema de clasificación a resolver.

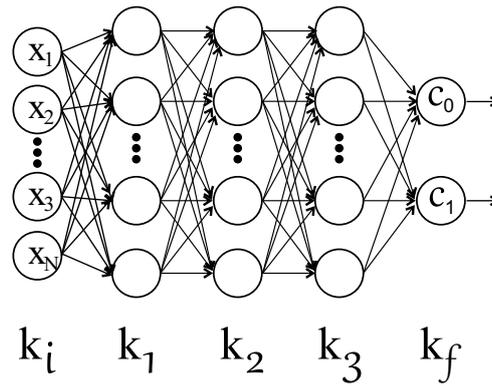


Figura 2.11: Esquema de una red neuronal artificial (*artificial neural networks ANN*) con tres capas intermedias $\{k_1, k_2, k_3\}$.

A pesar de estas ventajas, y considerando que el algoritmo desarrollado para realizar la clasificación que se propone en esta tesis, no necesariamente debe ser ejecutado en “tiempo real”, o en un dispositivo de bajo rendimiento computacional, no fue posible implementar una red neuronal, porque, como se comentó en la Sección 1.2 y 1.1, los elementos a clasificar (puntas de proyectil) cuentan con cientos de categorías y, algunas de ellas no disponen de una cantidad suficiente de ejemplares para realizar un entrenamiento adecuado de dicha red neuronal. Por otra parte, a continuación se describe el funcionamiento básico de una red neuronal.

En una red neuronal, cada una de las componentes del vector de características del elemento a clasificar es modificada progresivamente al multiplicarse por elementos interconectados llamados “neuronas”. Las neuronas se estructuran en niveles conocidos como “capas”. Toda red neuronal consta de una primera capa, o capa de entrada, que es el propio vector de características del elemento a clasificar y, de una capa de salida, que es el resultado, o decisión, que emite la red. Además de este par de capas (entrada–salida), todas aquellas capas intermedias, son conocidas como “capas internas de la red”; las cuales, pueden ir desde una, hasta decenas o cientos, dependiendo de la complejidad, y propósitos, del problema, de las características de la base de datos de entrenamiento y, de los recursos computacionales disponibles en la clasificación. La Figura 2.11 muestra el esquema general de una red neuronal conformada por 3 capas intermedias $\{k_1, k_2, k_3\}$.

Capítulo 3

Base de datos y desarrollo del algoritmo de clasificación

El presente capítulo se divide en dos partes, en la primera de ellas se explica cómo fue conformada la base de datos de imágenes de puntas de proyectil utilizada en el trabajo. En la segunda, se describe cada una de las etapas del algoritmo de clasificación; esto es: 1) pre-procesamiento de la imagen, 2) extracción del mapa de curvatura en el espacio a través de la escala (CSS-map), 3) detección de los puntos que definen la frontera proximal-distal en la punta de proyectil y, 4) segmentación y extracción de características de las componentes que conforman a la punta de proyectil. Además de esto, en esta segunda etapa también se describe la interfaz gráfica de usuario del programa.

3.1. Base de datos

La base de datos utilizada para el desarrollo de la primera etapa de diseño del algoritmo (separación automática de la punta de proyectil en sus componentes principales y extracción de características de dichas componentes), inicialmente se conformó por una gran variedad de categorías, de las cuales, no todas ellas contaban con varios ejemplares, por lo cual, para la segunda etapa del algoritmo (entrenamiento y evaluación del clasificador) dicha situación debió modificarse al incluir imágenes de varios sitios de internet. Específicamente, en la primera etapa de desarrollo se utilizó la base de datos del grupo arqueológico que colaboró en esta investigación, la cual está conformada por 32 puntas de proyectil del Instituto Nacional de Antropología e Historia con sede en Michoacán [INAH, 2018], 58 imágenes del Laboratorio de estudios de evolución humana del Instituto de Biociencias de la universidad de São Paulo, las cuales, amablemente fueron facilitadas para propósitos académicos, por la doctora Mercedes Okumura [Bradley and Okumura., 2016], y 34 imágenes pertenecientes al programa *George Washington's Mount Vernon archaeology program*, a su vez, amablemente facilitadas por Judith Paulos, miembro de la asociación *Mount Vernon Ladies Association* [MVLA, 2018]. Dado que estas dos últimas colecciones provenían de instituciones arqueológicas, la calidad de las imágenes fue buena y, en general, se adaptaban perfectamente a las demandas del trabajo; por otra parte, dado que en ambas colecciones las imágenes contenían la regla métrica arqueológica, esta tuvo que ser removida manualmente, recortando la imagen, para evitar confusión en el algoritmo de segmentación. En el caso

Tabla 3.1: Instituciones, compañías y sitios web que compartieron parte de sus bases de datos de puntas de proyectil para utilizarse en este trabajo. La columna “muestras” bosqueja cada una de estas colecciones. Por otra parte, en la última columna se indica el número total de imágenes utilizadas de cada colección.

Dueño	Muestra	Total
OverstreetID		555
RelicShack		241
São Paulo		58
Ken’s Relics		46
MVLA		34
INAH–Michoacán		32
J & DEE ARTIFACTS		11

de la base de datos del INAH–Michoacán, las imágenes fueron adquiridas por el autor de esta tesis, lo cual, permitió controlar las condiciones de captura de acuerdo a las demandas requeridas por el trabajo; para esto se estableció el fondo de la imagen de color verde y una resolución de 4608×3456 píxeles. Para el caso de la colección de la Universidad de São Paulo, todas las imágenes tienen fondo blanco y un tamaño de 3872×2592 píxeles. Finalmente, en la colección George Washington’s Mount Vernon, las imágenes tienen fondo negro y un tamaño aproximado promedio de 1300×1000 píxeles. En la tabla 3.1 se pueden apreciar algunas imágenes de estas colecciones (y de otras descritas posteriormente) y el total de cada una de ellas utilizadas en este trabajo de tesis.

Como se mencionó anteriormente, para realizar la segunda etapa del trabajo (entrenamiento y evaluación del clasificador), fue necesario incrementar la base de datos de tal manera que, cada categoría de punta de proyectil tuviera suficientes ejemplares para realizar estos dos procesos satisfactoriamente. Esto se logró al considerar imágenes de varios sitios de internet, previa autorización para uso académico, de coleccionistas, foros, proyectos de

investigación y difusión cultural relacionados con puntas de proyectil. La tabla 3.1 muestra el nombre de estos sitios, algunas muestras de las imágenes y el total de ellas utilizadas en este trabajo. La colección OverstreetID [OverstreetID, 2018] contribuyó de manera significativa a la base de datos, ya que ésta contaba con varios ejemplares de las categorías con formas de puntas de proyectil poco comunes. Además, ésta es una de las colecciones más completas de puntas de proyectil, ya que, de acuerdo a su sitio de internet, cuenta con aproximadamente 6000 imágenes. Estas imágenes tienen fondo blanco o negro, y todas ellas tienen una resolución promedio de 300×475 píxeles. El mayor problema que se enfrentó en esta base de datos es que en algunas de sus imágenes no existe un muy buen contraste entre el fondo y la punta de proyectil, por lo cual, algunas de éstas debieron segmentarse manualmente. La colección del sitio web RelicShack [RelicShack, 2018] también contribuyó de forma importante a nuestra base de datos de imágenes, esta se compone de imágenes con fondo blanco o negro y, dos resoluciones principales: una de 1067×1600 píxeles y otra de un promedio aproximado de 300×400 píxeles; además, todas estas imágenes presentaban la regla métrica arqueológica, la cual fue removida manualmente al recortar esa parte de la imagen. Por otra parte, también se utilizó la colección del sitio web Ken's Relics [Kens's Relics, 2018], la mayoría de estas imágenes tienen un fondo blanco uniforme y, en algunas de ellas, éste es rojo; además, éstas imágenes tienen distintas resoluciones de aproximadamente 800×500 píxeles, las cuales, están en función de la excentricidad de la punta de proyectil. Finalmente, se utilizó la colección del sitio web J & DEE ARTIFACTS [J & DEE ARTIFACTS, 2018], todas ellas tienen un fondo blanco y una resolución aproximada de 700×250 píxeles, la cual, también depende de la excentricidad de la punta de proyectil.

3.2. Interfaz gráfica de usuario

Con la finalidad que el algoritmo desarrollado para realizar la clasificación de puntas de proyectil pudiera ser manipulado de manera sencilla por cualquier persona, especialmente considerando la comunidad arqueológica, se desarrolló una interfaz gráfica de usuario. Además, ésta permite estructurar el proceso de clasificación al ofrecer la opción de seleccionar, y combinar, las categorías a clasificar y, automáticamente almacenar en un directorio los resultados obtenidos. La Figura 3.1 muestra esta interfaz gráfica, en ella se han delimitado con recuadros punteados sus componentes principales; de esta forma, la primera componente permite manipular las imágenes a clasificar, las cuales se presentan en la lista superior izquierda, por su parte. En la segunda componente son desplegadas estas imágenes. La tercera componente permite seleccionar y combinar las categorías y, finalmente, en la cuarta componente son desplegados los resultados del proceso; esto es, la categoría asignada por el algoritmo, las componentes principales de la punta de proyectil que han sido separadas y el mapa de curvatura CSS.

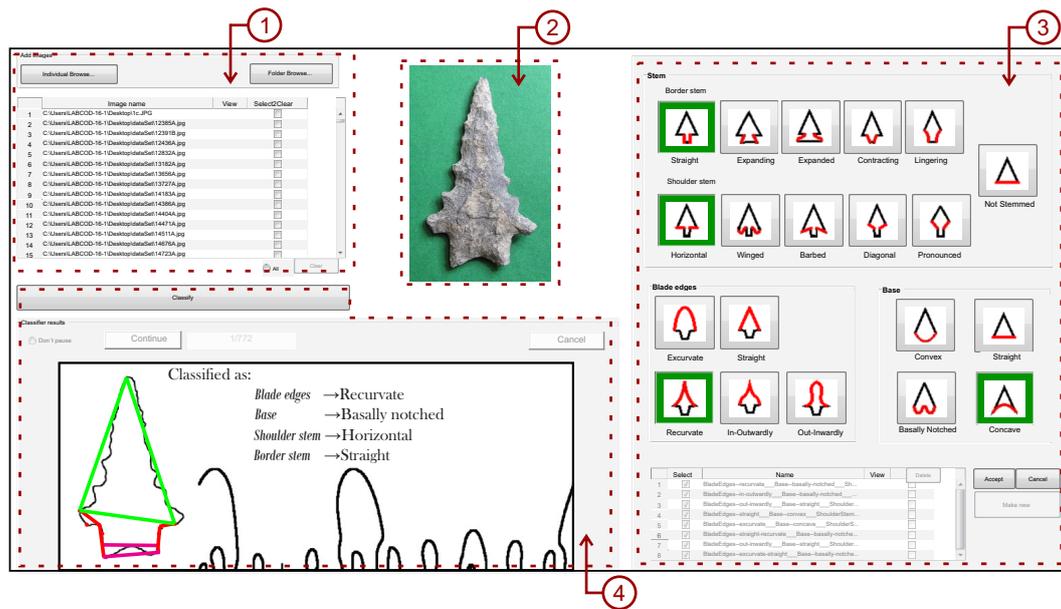


Figura 3.1: Interfaz gráfica de usuario del algoritmo de clasificación de puntas de proyectil. Ésta permite organizar las imágenes, seleccionar y combinar las categorías de la punta de proyectil a clasificar, visualizar los resultados y automáticamente almacenar las imágenes clasificadas en directorios correspondientes a las categorías asignadas.

3.3. Pre-procesamiento de la imagen

En esta etapa del algoritmo primero es segmentada la punta de proyectil, posteriormente se extrae su contorno y, finalmente, éste se alinea y escala a un tamaño estandarizado. La segmentación de la imagen se implementó en el espacio de color RGB, ya que las imágenes de nuestra base de datos se encuentran en dicho espacio, sin embargo, esto no es determinante en el algoritmo y puede ser fácilmente adaptado a otro espacio de color, o a imágenes en tono de gris. El algoritmo de segmentación está dividido en dos partes, la primera de ellas considera aquellas imágenes con fondo de color, mientras que la segunda considera aquellas con fondo blanco o negro. Las imágenes con fondo de color normalmente tienen un buen contraste entre el fondo y la punta de proyectil, de esta forma, la segmentación se puede realizar de manera sencilla aplicando un umbral a la intensidad de los píxeles, específicamente, dicho umbral se aplica a la imagen que resulta de la resta de los dos planos de color más contrastantes en la parte del fondo. Por otra parte, las imágenes con fondo blanco o negro, normalmente no presentan un buen contraste entre la punta de proyectil y el fondo, en este caso se aplica un algoritmo de crecimiento de regiones con una semilla ubicada en cualquier parte del fondo y un criterio de crecimiento basado en la intensidad de los píxeles que se encuentran dentro de una vecindad circular de cuatro píxeles de radio. Para determinar si el fondo de una imagen es de color o blanco/negro, se usa la desviación estándar del promedio de la intensidad de los píxeles de la orilla de la imagen; dicha orilla es considerada como el marco de la imagen de un ancho de 4 píxeles (para cada plano RGB). De esta forma, para las imágenes con fondos de color, dicha desviación estándar es más alta que para imágenes con fondo blanco o negro; este umbral fue determinado experimentalmente

en 12 unidades. Adicionalmente, con la finalidad de que el algoritmo también pudiera ser utilizado con imágenes delineadas en sus orillas (enmarcadas), situación medianamente frecuente en imágenes arqueológicas, se utilizó la transformación de Hough [Duda and Hart, 1972] para detectar y eliminar dichos marcos. A pesar de este proceso de segmentación, y debido a la calidad de algunas imágenes, aproximadamente el 13% de las imágenes de la base de datos fueron segmentadas manualmente, ya que el algoritmo no las segmentó de forma correcta.

Posteriormente a esta primer etapa de segmentación, la punta de proyectil es alineada al colocar su eje principal de forma perpendicular respecto a la horizontal. Dicho eje está conformado por la línea que se forma con el centroide de la punta de proyectil y la punta de la hoja. Para determinar la posición de la punta de la hoja se extraen las distancias de las líneas rectas que van del centroide a cada uno de los píxeles del contorno que están por encima de éste, posteriormente, se calcula el promedio de las 15 distancias más largas y se toman del conjunto antes mencionado aquellos píxeles que están por encima del centroide, que sean mayores al 98% de dicho promedio. Finalmente, la coordenada de la punta de la hoja se define como el promedio ponderado de las coordenadas de estos píxeles (los que fueron > 98% del promedio); los pesos de la ponderación son las respectivas distancias de cada uno de estos píxeles al centroide. Este proceso fue diseñado de esta forma para no considerar como punta aquellas partes alrededor de ella que por irregularidades son más distantes al centroide. Por otra parte, dicho proceso, también evita colocar la punta en secciones pequeñas del contorno (en la parte superior de la punta de proyectil) localizadas en posiciones muy distintas a la punta, aunque esta situación no sucede con frecuencia. Como se puede observar en el proceso descrito anteriormente, la imagen debe estar “pre-orientada” de tal forma que la parte de la punta de la hoja coincida con la parte superior de la imagen. Finalmente, el contorno (discreto) de la punta de proyectil debe ser escalado para que contenga aproximadamente 2000 píxeles.

3.4. Extracción del mapa de curvatura en el espacio a través de la escala (CSS-map)

La parte esencial del algoritmo de clasificación propuesta en este trabajo, depende de la detección de las componentes de la punta de proyectil (base, hoja, hombros y paredes del pedúnculo). Partiendo del postulado de que en las zonas del contorno que pertenecen a las fronteras de dichas componentes se presentan algunos de los mayores valores de curvatura, así como también los cambios más pronunciados en su orientación, se ha decidido utilizar al mapa CSS (*Curvature Scale Space Map*) [Abbassi et al., 1999] y al gradiente para la detección de estas zonas. A continuación se desarrollan los elementos teóricos de ambos métodos y su vinculación con el algoritmo.

El mapa CSS es un algoritmo que detecta los puntos de inflexión del contorno a diferentes escalas. Dicho mapa se extrae de acuerdo al trabajo de Abbassi [Abbassi et al., 1999]: en el cual, primero se calcula la curvatura del contorno del objeto y se identifican los cambios

de orientación (cambios de signo) en dicha curvatura. Posteriormente, se incrementa la escala al convolucionar al contorno con un filtro gaussiano de desviación estándar σ (escala); también en esta nueva escala se detectan los cambios de curvatura. Este proceso se itera al aumentar a un paso constante σ ; con lo cual, en el mapa CSS se presenta, en la dirección horizontal, los cambios de orientación de la curvatura y, en la dirección vertical, la escala. A continuación se presenta el entorno matemático necesario para obtener dicho mapa.

Sea $\Gamma(u) = (x(u), y(u))$ la representación vectorial de una curva paramétrica de parámetro u . La curvatura, $K(u)$, de $\Gamma(u)$ se define como:

$$K(u) = \frac{\dot{x}(u)\ddot{y}(u) - \ddot{x}(u)\dot{y}(u)}{(\dot{x}^2(u) + \dot{y}^2(u))^{3/2}} \quad (3.1)$$

donde $\dot{x}(u)$ y $\ddot{x}(u)$ representan, respectivamente, la primera y segunda derivada de $x(u)$. Similarmente para $\dot{y}(u)$ y $\ddot{y}(u)$. Sea $(X(u, \sigma), Y(u, \sigma))$ las componentes de la convolución de $\Gamma(u)$ con un filtro gaussiano unidimensional $(g(u, \sigma))$; esto es:

$$\begin{aligned} X(u, \sigma) &= x(u) * g(u, \sigma) \\ Y(u, \sigma) &= y(u) * g(u, \sigma) \end{aligned} \quad (3.2)$$

De esta forma, $X_u(u, \sigma)$ y $X_{uu}(u, \sigma)$ son, respectivamente, la primera y segunda derivadas parciales de $X(u, \sigma)$, con respecto a u . De acuerdo con las propiedades de la convolución, estas derivadas pueden ser calculadas como:

$$\begin{aligned} X_u(u, \sigma) &= x(u) * g_u(u, \sigma) \\ X_{uu}(u, \sigma) &= x(u) * g_{uu}(u, \sigma) \end{aligned} \quad (3.3)$$

donde g_u y g_{uu} son, respectivamente, la primera y segunda derivadas parciales, con respecto a u , de $g(u, \sigma)$. Similarmente para $Y_u(u, \sigma)$ y $Y_{uu}(u, \sigma)$. Dado que $g_u(u, \sigma)$ y $g_{uu}(u, \sigma)$ son conocidas de forma analítica, la curvatura con respecto a la escala está dada por:

$$K(u, \sigma) = \frac{X_u(u, \sigma)Y_{uu}(u, \sigma) - X_{uu}(u, \sigma)Y_u(u, \sigma)}{(X_u(u, \sigma)^2 + Y_u(u, \sigma)^2)^{3/2}} \quad (3.4)$$

En un contorno cerrado, como es el caso de las puntas de proyectil, toda sección cóncava estará delimitada por dos secciones convexas. Por otra parte, al extraer el mapa CSS, y conforme la escala es incrementada, el contorno se vuelve más "suave" y "compacto", de esta forma, gradualmente este par de puntos se aproximan hasta traslaparse y convertir a la sección cóncava en convexa. Este proceso crea "lóbulos" en el mapa CSS, conforme más pronunciadas sean las inflexiones del contorno, estos lóbulos serán más grandes. Dicho proceso se ilustra en la Figura 3.2, en ella se presenta una punta de proyectil a diferentes escalas y su correspondiente mapa CSS; los cambios de curvatura se han resaltado al marcar sobre el contorno las partes cóncavas en gris y las convexas en negro. Por otra parte, en el mapa CSS de esta figura, se puede apreciar que el par de lóbulos más grandes corresponden a las dos partes con más inflexión de la punta de proyectil (los bordes de la hoja) y, por último, la

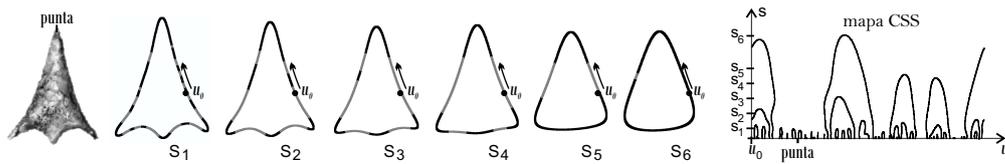


Figura 3.2: Evolución de la escala de la punta de proyectil al extraer el mapa CSS. Los cambios en la orientación de la curvatura se han resaltado al marcar sobre el contorno las partes cóncavas en gris y las convexas en negro.

simetría idealizada que tiene ésta, implica que dicho par de lóbulos sean también simétricos, tanto en forma, como en distancia respecto a la punta de la hoja. Estas características son las que permitirán al algoritmo detectar las distintas componentes de la punta de proyectil.

En el algoritmo desarrollado en este trabajo, el mapa CSS se computó respecto a un sub–muestreo de la versión que fue normalizada del contorno de la punta de proyectil a un tamaño de 2000 píxeles (Sección 3.3); así, con dicho sub–muestreo se obtiene un contorno conformado por 600 píxeles. Este proceso permite estandarizar varios parámetros del algoritmo y reduce el tiempo de procesamiento. La norma euclidiana es utilizada para establecer la distancia de muestreo, ya que ésta asiste para que el mapa CSS obtenido sea de buena calidad. Finalmente, el tiempo de procesamiento de la convolución (Ecuación 3.3) aumenta considerablemente al incrementar la escala del mapa CSS, de esta forma, el algoritmo se diseñó de manera que la convolución sólo se realizara alrededor del contorno de los lóbulos, lo que reduce dicho tiempo en un 90 %. Por otra parte, la primera escala del mapa CSS se fijó en 24 unidades ($\sigma_0 = 24$) y, su incremento, en pasos de una unidad. Se utilizó el valor de 24 unidades para la primera escala porque en nuestra experimentación, se pudo comprobar que las escalas más pequeñas que ésta, no representan la forma global de la punta de proyectil y generan mucho ruido en el mapa CSS.

Finalmente, para utilizar al mapa CSS en las etapas posteriores, éste debe ser pre–procesado. En dicho pre–proceso se vinculan las secciones de los lóbulos que fueron cortados en los extremos laterales del mapa de curvatura; la Figura 3.2 presenta un par de estos lóbulos en su correspondiente mapa de curvatura. Además, debido a que se utilizó un contorno sub–muestreado para la extracción de dicho mapa, algunos de los lóbulos quedan abiertos en la parte superior; los cuales, se “cerraron” con la función de Matlab *filledgeaps* desarrollada por Peter Kovesei [Kovesei, 2015].

3.5. Detección de la frontera proximal–distal

Para identificar a las componentes de la punta de proyectil es necesario encontrar la frontera entre la parte proximal y distal de ésta. En realidad, la parte distal sólo está conformada por la hoja de la punta de proyectil, sin embargo, se usó este término porque puede ser directamente asociado con la idea de “frontera proximal–distal”. La hoja está formada por los bordes y por la punta, la Figura 3.3 muestra cada una de estas partes; sin embargo, para la detección de dicha frontera, sólo es necesario considerar la parte de los bordes de la hoja. Por

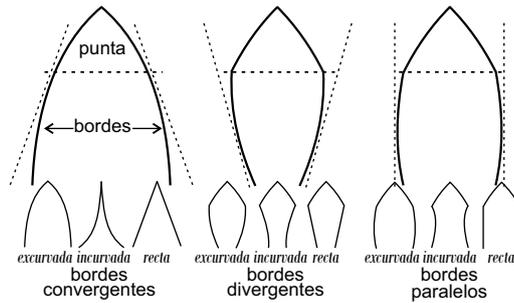


Figura 3.3: Orientaciones del borde de la hoja.

otro lado, en las categorías de la componente bordes de la hoja, ambas partes (bordes y punta) se consideran como una única componente (Figura 1.2), ya que en este caso, el número de categorías a combinar se reduce, mientras que no se afecta a la precisión de la clasificación, porque en la mayoría de los casos, la punta sigue la forma natural de los bordes y, en aquellos casos en donde esto no sucede, como en las formas “recurvadas entrantes–salientes” y “salientes–entrantes”, estos se consideran en el propio conjunto de categorías de esta componente (Figura 1.2). Para la identificación de la frontera proximal–distal, en el caso de que la punta de proyectil sea peduncular (el significado de dicho término se explicará más adelante), deben tomarse en cuenta las siguientes formas de los bordes de la hoja: estos pueden ser convergentes o divergentes hacia la punta, o paralelos entre sí, además, pueden ser convexos, cóncavos o rectos, la Figura 3.3 muestra las combinaciones que pueden obtenerse. En el caso de que la punta de proyectil sea no peduncular, estas consideraciones no intervienen en la detección de dicha frontera; el proceso que corresponde a este caso se describirá en la Sección 3.

La zona proximal está formada por la base y por el pedúnculo (en caso de que la punta de proyectil sea peduncular). La función principal del pedúnculo es sujetar a la punta de proyectil a un mástil, de esta forma, también es conocido como “área de enmangue”. Además, el pedúnculo puede tener otras funciones como retención del proyectil en el objetivo, incrustación y/o desangrado de la presa. A diferencia de la zona distal, en esta zona (proximal) se han considerado sus partes por separado (base y pedúnculo), ya que el pedúnculo tiene sus propios componentes como aletas, barbas y muescas. Si se traza una línea horizontal, en el plano principal de la punta de proyectil (plano frontal), con la finalidad de separar las zonas proximal y distal, los puntos en donde esta línea interseca al contorno de la punta de proyectil son conocidos como puntos de la frontera proximal–distal (PFPD). La Figura 3.4 muestra las zonas proximal y distal de una punta de proyectil peduncular, la línea de frontera entre estas zonas (línea punteada), y los puntos asociados a dicha frontera (β^{izq} y β^{der} en este caso). Los PFPD serán utilizados para identificar estas zonas.

El proceso de detección de los PFPD es distinto para las puntas de proyectil pedunculares y no pedunculares. De esta forma, el mapa CSS es utilizado para diferenciar cada una de ellas. Para los tipos pedunculares, este mapa presenta un par de lóbulos cuasi–similares en forma y distancia respecto al punto central de la zona proximal (PCP); este punto corresponde a la intersección entre la parte proximal del contorno y la parte baja de la sección

distancia vertical, en el mapa CSS, del punto más alto del contorno del lóbulo (ver los mapas CSS de las Figuras 3.4 y 3.5) y, la posición es el valor de u que corresponde a este punto; esto teniendo en cuenta que u es la variable paramétrica que recorre el borde de la punta de proyectil en sentido anti-horario, cuyo origen (u_0) está situado en la parte derecha de dicha punta a la altura del centroide (Figura 3.4).

Por otra parte, la Ecuación 3.6 ofrece un umbral adaptativo para evaluar la simetría que guardan la posición del par de lóbulos, específicamente, este umbral es más tolerante con los lóbulos de mayor tamaño, ya que la cantidad de estos es menor (respecto a la cantidad de lóbulos más pequeños), dado que representan zonas de inflexión muy pronunciadas. Esto implica que es más probable que la diferencia en distancias de un par de lóbulos pequeños, elegidos al azar, sea menor, sin que estos realmente representen la forma peduncular de la punta de proyectil. La Figura 3.5 ilustra esta situación, en ella se presentan tres puntas de proyectil de distintas características pedunculares. La última de ellas (derecha) es no peduncular, por lo cual, su mapa CSS no debe presentar este tipo de lóbulos; sin embargo, debido a las propias irregularidades en la zona proximal de su contorno, ésta genera un mapa con un par de lóbulos con valor asociado $\Delta_d = 20$, por lo cual, dependiendo del umbral asignado en el criterio de simetría relacionado con la diferencia en la posición relativa de los lóbulos, estos podrían ser erróneamente considerados como pedunculares. Por otra parte, la punta de proyectil de enmedio es peduncular, por lo cual, su par de lóbulos candidatos así deben ser considerados; sin embargo, debido a que dicho pedúnculo no es tan simétrico, su valor $\Delta_d = 23$ y, su asignación nuevamente dependerá de este umbral. Por último, la punta de proyectil de la izquierda tiene una forma peduncular con buena simetría, por lo cual, la asignación correcta de ésta casi no dependerá del umbral en cuestión. De esta forma, podemos observar que el umbral adaptativo de la Ecuación 3.6, el cual se muestra en las gráficas de esta misma figura, permite clasificar correctamente como “peduncular” a la punta de proyectil de enmedio, ya que sus lóbulos, cumplen con la tendencia de tener mayor altura que los lóbulos no pedunculares de la punta de proyectil derecha, para la cual, dicho umbral, debido a la baja altura de sus lóbulos, permite que esta sea clasificada correctamente como “no peduncular”.

Si en el proceso anterior la punta de proyectil se establece como peduncular, cada uno de los dos PFPD (uno por lado) son detectados con base al gradiente en dirección horizontal ($\nabla_x[u]$) de aquellas secciones del contorno, que inician en el cuello del pedúnculo, el cual, corresponde a la posición del lóbulo peduncular (círculos blancos de la Figura 3.4) y terminan 300 píxeles después, avanzando hacia los bordes de la hoja (sentido horario para el lado izquierdo y anti-horario para el derecho). En la punta de proyectil de la Figura 3.4 estas secciones del contorno se muestran remarcadas en negro y, sus respectivos gradientes se presentan en la parte derecha de abajo de la figura. La longitud de 300 píxeles se determinó experimentalmente para garantizar la inclusión de los PFPD de todas las categorías. Considerando que el origen del sistema de coordenadas se fijó en el centroide de la punta de proyectil, y que la coordenada horizontal se incrementa hacia la derecha, el gradiente del lado izquierdo presenta el siguiente comportamiento: primero es predominantemente negativo, después, probablemente cero y, finalmente positivo. Por otra parte, la secuencia

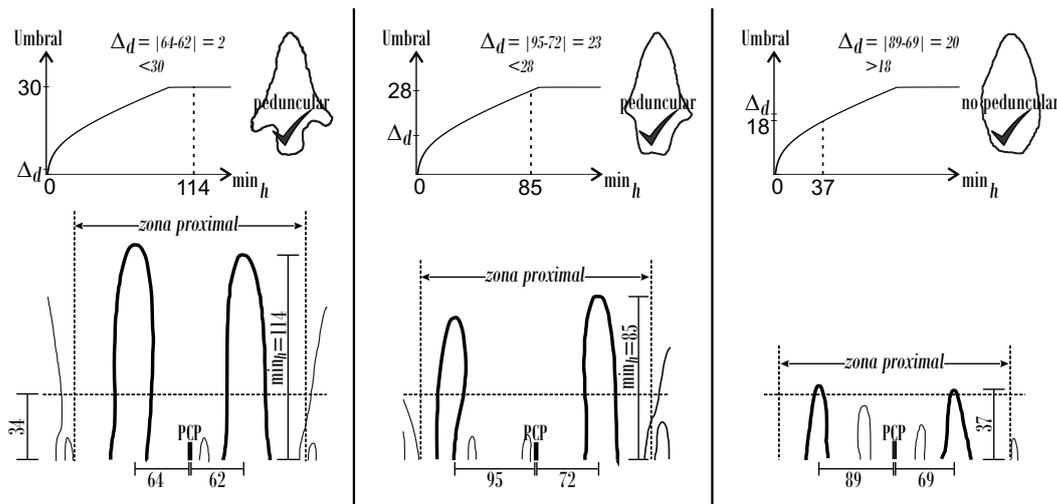


Figura 3.5: Ejemplo de las ventajas que se obtienen al utilizar el umbral adaptativo de la Ecuación 3.6 para discriminar puntas de proyectil pedunculares y no pedunculares.

del gradiente del lado derecho es la siguiente: positivo, después, quizá cero y, por último, negativo (ver Figura 3.4).

La primera parte de estos gradientes representan los hombros del pedúnculo, los cuales, comienzan en el cuello del pedúnculo (vinculado con la correspondiente posición en el contorno de los lóbulos pedunculares) y terminan al inicio de los bordes de la hoja. La parte intermedia del gradiente (la que es igual a cero), se forma al considerar los bordes de la hoja del tipo paralelos y algunos elementos de la categoría alada de la componente hombros del pedúnculo. Finalmente, la última parte del gradiente, representa la parte convergente de la punta de proyectil, por lo cual, en el caso de los bordes de la hoja del tipo convergentes, ésta parte comienza al inicio de la hoja y, en el caso de los bordes de la hoja paralelos, comienza al inicio de la punta. De esta manera, es posible identificar los PFPD al delimitar cada una de estas partes del gradiente. Además, el rango de estos gradientes es $\{-1,0,1\}$ ya que los píxeles del contorno están situados de forma contigua. En este sentido, dicha delimitación puede realizarse al ajustar cada región del gradiente a una función escalón unitario discreta desplazada en el dominio u . Los parámetros de desplazamiento son asociados con las posiciones antes mencionadas de la punta de proyectil. Las funciones de ajuste del gradiente para cada uno de los lados de la punta de proyectil se definen como:

$$\begin{array}{c|c}
 \text{lado izquierdo} & \text{lado derecho} \\
 \hline
 f^{(\text{izq})}[u] = -H[-u + \alpha^{(\text{izq})}] + H[u - \beta^{(\text{izq})}] & f^{(\text{der})}[u] = H[-u + \alpha^{(\text{der})}] - H[u - \beta^{(\text{der})}] \\
 0 \leq u \leq 300 & 0 \leq u \leq 300 \\
 (3.7) & (3.8)
 \end{array}$$

en donde H es la función escalón unitario discreta definida como: $H[u] = \begin{cases} 0; & u < 0 \\ 1; & u \geq 0 \end{cases}$. Por

otra parte, los super-índices “izq” y “der” representan, respectivamente, el lado izquierdo y derecho de la punta de proyectil. Finalmente, α y β son los parámetros que desplazan el escalón. Las gráficas del gradiente de la Figura 3.4 también muestran cada par de escalones que conforman las funciones de ajuste $f^{(izq)}[u]$ y $f^{(der)}[u]$.

El conjunto de valores α y β que serán considerados para encontrar el mejor ajuste de $f^{(izq)}[u]$ son: $\alpha^{(izq)} \in \{u | \nabla_x^{(izq)} [u] = -1\}$ y $\beta^{(izq)} \in \{u | \nabla_x^{(izq)} [u] = 1\}$. De forma similar para $f^{(der)}[u]$; esto es: $\alpha^{(der)} \in \{u | \nabla_x^{(der)} [u] = 1\}$ y $\beta^{(der)} \in \{u | \nabla_x^{(der)} [u] = -1\}$; donde ∇ es el operador gradiente. Por otra parte, el mejor ajuste se logra con el par de valores α y β , de cada lado, que minimizan las siguientes funciones de error:

$$\begin{aligned} \text{error}^{(izq)} &= p_{-H,-1}^{(izq)} \text{error}\{(f^{(izq)}[u], \nabla_x^{(izq)} [u]) | u \in R_{-H,-1}^{(izq)}\} + \\ & p_{-H,0}^{(izq)} \text{error}\{(f^{(izq)}[u], \nabla_x^{(izq)} [u]) | u \in R_{-H,0}^{(izq)}\} + \\ & p_{H,0}^{(izq)} \text{error}\{(f^{(izq)}[u], \nabla_x^{(izq)} [u]) | u \in R_{H,0}^{(izq)}\} + \\ & p_{H,1}^{(izq)} \text{error}\{(f^{(izq)}[u], \nabla_x^{(izq)} [u]) | u \in R_{H,1}^{(izq)}\} \end{aligned} \quad (3.9)$$

$$\begin{aligned} \text{error}^{(der)} &= p_{H,1}^{(der)} \text{error}\{(f^{(der)}[u], \nabla_x^{(der)} [u]) | u \in R_{H,1}^{(der)}\} + \\ & p_{H,0}^{(der)} \text{error}\{(f^{(der)}[u], \nabla_x^{(der)} [u]) | u \in R_{H,0}^{(der)}\} + \\ & p_{-H,0}^{(der)} \text{error}\{(f^{(der)}[u], \nabla_x^{(der)} [u]) | u \in R_{-H,0}^{(der)}\} + \\ & p_{-H,-1}^{(der)} \text{error}\{(f^{(der)}[u], \nabla_x^{(der)} [u]) | u \in R_{-H,-1}^{(der)}\} \end{aligned} \quad (3.10)$$

Los sub-índices “-H” y “H” representan cada uno de los dos escalones que intervienen en la función de ajuste; esto es: $-H[-u + \alpha^{(izq)}]$ y $-H[u - \beta^{(der)}]$ para $-H$ y $H[u - \beta^{(izq)}]$ y $H[-u + \alpha^{(der)}]$ para H , además, los sub-índices -1 , 0 y 1 se refieren a las correspondientes partes de estos escalones que adquieren estos valores. El conjunto p 's son constantes utilizadas para ponderar a las funciones de error (éstas serán explicadas después). Finalmente, el conjunto R 's son regiones del dominio definidas como:

lado izquierdo

$$\begin{aligned} R_{-H,-1}^{(izq)} &= 0 \leq u \leq \alpha^{(izq)} \cap \{u | \nabla_x^{(izq)} [u] = [-1, 0]\} \\ R_{-H,0}^{(izq)} &= \alpha^{(izq)} < u < \beta^{(izq)} \cap \{u | \nabla_x^{(izq)} [u] = [-1, 0]\} \\ R_{H,0}^{(izq)} &= \alpha^{(izq)} < u < \beta^{(izq)} \cap \{u | \nabla_x^{(izq)} [u] = [0, 1]\} \\ R_{H,1}^{(izq)} &= \beta^{(izq)} \leq u \leq 300 \cap \{u | \nabla_x^{(izq)} [u] = [0, 1]\} \end{aligned} \quad (3.11)$$

lado derecho

$$\begin{aligned}
R_{H,1}^{(\text{der})} &= 0 \leq u \leq \alpha^{(\text{der})} \cap \{u \mid \nabla_x^{(\text{der})} [u] = [0, 1]\} \\
R_{H,0}^{(\text{der})} &= \alpha^{(\text{der})} < u < \beta^{(\text{der})} \cap \{u \mid \nabla_x^{(\text{der})} [u] = [0, 1]\} \\
R_{-H,0}^{(\text{der})} &= \alpha^{(\text{der})} < u < \beta^{(\text{der})} \cap \{u \mid \nabla_x^{(\text{der})} [u] = [-1, 0]\} \\
R_{-H,-1}^{(\text{der})} &= \beta^{(\text{der})} \leq u \leq 300 \cap \{u \mid \nabla_x^d [u] = [-1, 0]\}
\end{aligned} \tag{3.12}$$

El símbolo \cap denota el operador de intersección de conjuntos. La expresión de la parte izquierda de este operador establece el dominio que corresponde a cada parte del escalón. Así mismo, la expresión del lado derecho (del símbolo \cap) se encarga de filtrar estos dominios para computar el error solamente con elementos del gradiente dentro del rango de valores que corresponde a cada uno de los escalones; posteriormente se explicará las ventajas que ofrecen estas funciones y que son parte importante del algoritmo propuesto en este trabajo. $\text{error}\{(f[u], \nabla[u]) \mid u \in R\}$ es el error entre $f[u]$ y $\nabla[u]$, definido como el error medio cuadrático, evaluado en la región R ; por ejemplo, $\text{error}\{(f^{(\text{izq})}[u], \nabla_x^{(\text{izq})}[u]) \mid u \in R_{-H,-1}^{(\text{izq})}\}$ se obtienen como:

$$\text{error}\{(f^{(\text{izq})}[u], \nabla_x^{(\text{izq})}[u]) \mid u \in R_{-H,-1}^{(\text{izq})}\} = \frac{\sum_{u \in R_{-H,-1}^{(\text{izq})}} (f_{-H}^{(\text{izq})}[u] - \nabla_x^{(\text{izq})}[u])^2}{|R_{-H,-1}^{(\text{izq})}|} \tag{3.13}$$

donde $|R_{-H,-1}^{(\text{izq})}|$ es la cardinalidad del conjunto $R_{-H,-1}^{(\text{izq})}$.

El conjunto p 's pondera la función de error para contemplar las propias características de las dos partes que conforman al gradiente. Dado que $p_{-H,-1}^{(\text{izq})}$ y $p_{-H,0}^{(\text{izq})}$ ponderan el error del primer escalón de $f^{(\text{izq})}[u]$ y, $p_{H,0}^{(\text{izq})}$ y $p_{H,1}^{(\text{izq})}$ ponderan el error del segundo escalón de esta misma función. Además, por normalización, la suma de los cuatro pesos debe ser igual a 1 y, se desea dar la misma relevancia al error de ajuste en ambos escalones, esto es:

$$p_{-H,-1}^{(\text{izq})} + p_{-H,0}^{(\text{izq})} = 1/2 \text{ y } p_{H,0}^{(\text{izq})} + p_{H,1}^{(\text{izq})} = 1/2$$

Específicamente, los valores de p para ambos lados son:

lado izquierdo	lado derecho
$1^{\text{er}} \text{ escalón} \Rightarrow p_{-H,-1}^{(\text{izq})} = 0.25; p_{-H,0}^{(\text{izq})} = 0.25$	$1^{\text{er}} \text{ escalón} \Rightarrow p_{H,1}^{(\text{der})} = 0.25; p_{H,0}^{(\text{der})} = 0.25$
$2^{\text{do}} \text{ escalón} \Rightarrow p_{H,0}^{(\text{izq})} = 0.45; p_{H,1}^{(\text{izq})} = 0.05$	$2^{\text{do}} \text{ escalón} \Rightarrow p_{-H,0}^{(\text{der})} = 0.45; p_{-H,-1}^{(\text{der})} = 0.05$

(3.14)

Con base a la experimentación, para la mayoría de los casos, sin las condiciones $p_{H,0}^{(\text{izq})} > p_{H,1}^{(\text{izq})}$ y $p_{-H,0}^{(\text{der})} > p_{-H,-1}^{(\text{der})}$, a β se le asignaría un valor incorrecto, casi hasta el final del gradiente. Este error ocurre porque en la parte del gradiente que corresponde a la parte convergente de los bordes de la hoja, existen varias transiciones hacia cero, en ocasiones, inclusive existen

más 0's que 1's, para el lado izquierdo y, que -1 's para el derecho. Estas transiciones son causadas por la propia forma afilada de los bordes de la hoja y por la representación discreta del contorno; esto es: el ángulo que se forma entre el eje principal de la punta de proyectil y la línea recta que representa al borde de la hoja, en la mayoría de los casos tiende a ser muy agudo, lo que causa transiciones verticales, y no horizontales, en esta parte del gradiente. Por otra parte, $p_{-H,-1}^{(izq)} = p_{-H,0}^{(izq)}$ y $p_{H,1}^{(der)} = p_{H,0}^{(der)}$ porque en la parte del hombro del pedúnculo, al existir considerablemente una menor cantidad de transiciones hacia cero, éstas no afectan el proceso de ajuste.

A pesar de esto, en algunas ocasiones β es colocada al final del gradiente, en cuyo caso debe utilizarse α . Por otra parte, α también debe utilizarse si los bordes de la hoja son paralelos. Para evaluar cual de los dos parámetros debe ser utilizado, experimentalmente se estableció un umbral que implica el uso de β si éste es menor que el 90% de la correspondiente señal gradiente en la que participa, de lo contrario, se utiliza α . Sin embargo, este umbral implica utilizar a β , y no a α , en algunos bordes de la hoja del tipo paralelos de tamaño pequeño y, desafortunadamente, si el umbral es reducido para corregir este error, se descartaría, incorrectamente, el uso de β en varias imágenes de las categorías "alada" y "barbada" del hombro del pedúnculo, ya que el contorno de éstas es más largo que el de las otras categorías. Para resolver este problema, en la condición para discriminar a ambos parámetros, también se ha incluido la diferencia entre α y β ; esto es: β también será rechazada si esta diferencia es mayor que el 20% del largo del gradiente.

Por último, las expresiones a la derecha del símbolo \cap filtran a la región R para considerar solamente el error respecto a aquellos elementos del gradiente que son iguales a la función escalón que se considera en dicha región. Esto permite atenuar, en el proceso de ajuste, la influencia del ruido del contorno y peculiaridades en la forma de algunas puntas de proyectil como el "aserrado" de los bordes de la hoja, ya que ambos casos originan pequeños cambios, no contemplados, en la dirección del gradiente. Además, de no considerar dicho proceso de filtrado, los pesos p 's podrían amplificar estos errores aumentando así la posibilidad de colocar los parámetros del escalón en posiciones incorrectas.

Por otra parte, si la punta de proyectil es clasificada como no peduncular, se estableció que el proceso utilizado para identificar los PFPD es diferente para las puntas de proyectil de base incurvada (cóncava y muesca basal) o no incurvada (convexa y recta). El proceso para identificar dichas sub-categorías de la base (incurvada/no incurvada) se describe en la Sección 3.7. En el caso de que la base sea incurvada, los PFPD se establecen como las posiciones de los bordes de la hoja situados a la misma altura que el PCP. Por otra parte, si la base es no incurvada, estos puntos se establecen como el máximo de la curvatura en las zonas de cada lado, próximas al PCP. Dado que en la categoría "recta", de la base de la punta de proyectil, la frontera proximal-distal está muy bien definida, este par de máximos son grandes y su posición está bien delimitada. Sin embargo, en la categoría "convexa", de la base de la punta de proyectil, conforme está es más curvada, la frontera proximal-distal no es tan clara, por lo cual, la detección de dichos máximos se complica, hasta llegar a la situación, en la cual, no es posible realizar dicha identificación y, los PFPD se establecen de manera pre-definida como el par de puntos del contorno (uno por lado) a una altura similar

al 95 % de la longitud del eje principal de la punta de proyectil.

3.6. Extracción de características de la punta de proyectil; bordes de la hoja

El vector de características de la componente bordes de la hoja está formado por:

1. El área de los lóbulos asociados a los bordes de la hoja
2. La tortuosidad (T)
3. La distancia de las posiciones del contorno que corresponden a los lóbulos asociados a los bordes de la hoja, respecto a la punta de ésta

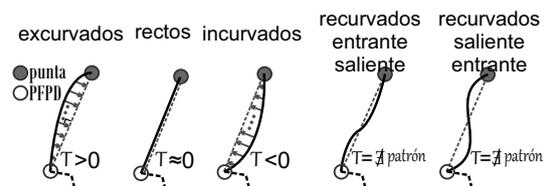


Figura 3.6: Tortuosidad (T) de cada categoría de los bordes de la hoja.

La inflexión de los bordes de la hoja de la categoría “incurvados” y de las dos categorías “recurvados” (ver Figura 3.6) genera, en el mapa CSS, un par de lóbulos grandes y simétricos entre sí, tanto en forma como en distancia respecto a la punta de la hoja; el mapa CSS de la Figura 3.2 presenta un ejemplo de estos lóbulos, ya que los bordes de la hoja de su correspondiente punta de proyectil son del tipo incurvados. El conjunto de candidatos a ser considerados lóbulos del borde de la hoja, está formado por aquellos que son mayores que 78 unidades y están dentro de la correspondiente zona distal de la punta de proyectil; esta zona está formada por el complemento de la zona proximal (Sección 3.5, Figura 3.4). Para determinar si en este conjunto existe un par de lóbulos asociados al borde de la hoja se aplican las mismas condiciones de simetría utilizadas en la detección de los lóbulos pedúnculares (Ecuaciones 3.5 y 3.6). Como se puede deducir de este procedimiento y, con base a la asesoría del grupo arqueológico que colaboró en el proyecto, se decidió no considerar como inflexos aquellos bordes de la hoja que no fueran simétricos (esto es: aquellos que sólo presentaran un lado inflexo), ya que en la mayoría de estos casos no se puede garantizar que la inflexión unilateral represente una intención objetiva de obtener esa forma en ambos lados de la hoja; esto es, dicha inflexión pudo ser causada por daños en la punta de proyectil o por una deformación en la piedra. En este caso, el área de los lóbulos de los bordes de la hoja se establece como cero; de forma similar, para los bordes rectos y convexos dicho valor del área también se establece como cero.

Dado que el área de los lóbulos del borde de la hoja sólo discriminan entre bordes inflexos y no inflexos, es necesario incluir otros parámetros que permitan discriminar a todas las categorías. En este caso, la tortuosidad (T) discrimina a las categorías “excurvados”, “rectos” e “incurvados”. Una manera intuitiva de definir la tortuosidad es como aquella medida de la deflexión de una curva respecto a la línea recta que une los extremos de ésta [Pavlidis, 1980]; en este caso, se considera la deflexión de los bordes de la hoja, los cuales, comienzan en el punto de la frontera proximal–distal (PFPPD) y terminan en la punta de la hoja (ver figura 3.6). Formalmente, T se calcula como el promedio de la distancia mínima (Δ) que se

forma desde cada pixel del borde de la hoja hasta la línea recta que une los extremos de ésta; esto es:

$$T = \frac{1}{N} \sum_{i=1}^N \Delta_i \quad (3.15)$$

en donde N es el número de pixeles que integran el borde de la hoja. En la ecuación 3.15, la tortuosidad es invariante ante rotación y traslación; por otra parte, para el descriptor desarrollado se ha logrado obtener invarianza ante la escala, al normalizar a T con respecto a la distancia de muestreo que se utilizó para decimar el borde de la punta de proyectil y obtener un borde conformado por 600 pixeles que se utilizó para generar el mapa CSS (Sección 3.3). Además de la ecuación 3.15, existen otras formas de definir la tortuosidad (ver por ejemplo: [Bribiesca, 2013, Bullitt et al., 2003, J. Capowski and Freedman, 1995]). En estas definiciones, la tortuosidad, además de ser invariante ante escala, rotación y traslación, también es invariante ante la reflexión especular; lo cual, normalmente es conveniente para un descriptor que será utilizado en un problema genérico de clasificación; sin embargo, considerando la clasificación de puntas de proyectil, las categorías “convexas” y “cóncavas”, serían equivalentes bajo esta transformación (reflexión especular); por lo cual, no se han considerado definiciones de T similares a las que aparecen en esos trabajos. Finalmente, en trabajos como [Bullitt et al., 2003, J. Capowski and Freedman, 1995] el valor calculado de T se complementa con la curvatura y/o la frecuencia local del cambio de orientación de la curva; sin embargo, se decidió no considerar estos parámetros, ya que el enfoque local de T resulta ser contraproducente en nuestro descriptor; esto es debido a la propia naturaleza del borde de la punta de proyectil, el cual, comúnmente presenta cambios locales en la trayectoria ocasionados por el deterioro que el artefacto adquiere con el paso del tiempo.

Debido al sentido que se le da a la línea recta y a la orientación del sistema coordenado, la tortuosidad, para el borde de la hoja del lado izquierdo en la imagen, es mayor que 0 para la categoría “excurvados”, cercana a 0 para la categoría “rectos”, menor que 0 para la categoría “incurvados”, e indefinida para el par de categorías recurvados, ya que en las imágenes de éstas no se aprecia una tendencia clara de este parámetro. En la Figura 3.6 se bosquejan cada una de estas situaciones. Por otra parte, el valor de T cambia de signo para el borde de la hoja del lado derecho, por lo cual, para lograr un discernimiento entre ambos lados, se cambió el signo T que se obtuvo para el lado derecho de la hoja. Cabe hacer notar, que no es necesario realizar este procedimiento de cambio de signo para las otras dos componentes de la punta de proyectil en las que se utilizó el parámetro T (base y hombros del pedúnculo).

Por último, para discriminar a las categorías “incurvados”, “recurvados entrantes–salientes” y “recurvados salientes–entrantes”, se consideró la distancia de los lóbulos del borde de la hoja a la punta de ésta. Dicho parámetro se incluye en el vector de características como el promedio del valor obtenido en ambos lados. Por otra parte, dicho valor se estableció como 0 en el caso de las imágenes que no presentan lóbulos asociados a los bordes de la hoja.

3.7. Extracción de características de la punta de proyectil; base

El vector de características de esta componente de la punta de proyectil está formado por:

1. El área del lóbulo asociado a la base
2. La tortuosidad (T)
3. α y β ; parámetros de ajuste del gradiente
4. $\Delta_{\alpha\beta} = \beta - \alpha$

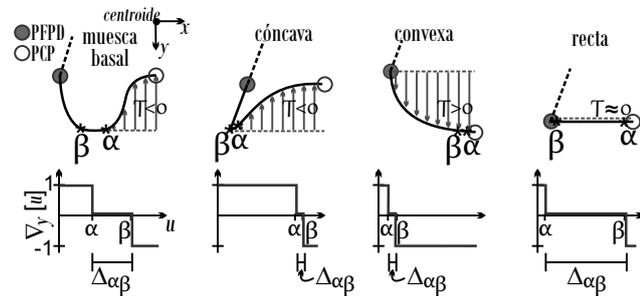


Figura 3.7: Proceso de ajuste del gradiente y tortuosidad (T) para cada categoría de la base.

De manera similar al análisis realizado para los bordes de la hoja, en el caso de la base, se forma un lóbulo en el mapa de curvatura CSS cuando ésta posee inflexión. Las categorías “muesca basal” y “cóncava” al ser “inflexas” presentan dicho lóbulo en su mapa CSS (Figura 3.7) y, son conocidas como “bases del tipo incurvadas”. Por otra parte, las categorías “convexa” y “recta” al no presentar dicha inflexión, se conocen como “bases del tipo no incurvadas”. Para considerar a un lóbulo del mapa CSS como “lóbulo de la base” éste debe tener una altura mayor a 56 unidades y, la distancia de su posición respecto al PCP, no debe ser mayor a 25 unidades.

En las categorías “convexa” y “recta”, el área del lóbulo de la base es igual a cero, por lo cual, debe usarse un criterio distinto para discriminarlas; la tortuosidad (T) es utilizada en este caso. En la Figura 3.7 se puede apreciar que la tortuosidad (T) es mayor que 0 para la categoría “convexa” y, cercana a 0 en el caso de la categoría “recta”. El par de puntos que forman la línea recta que es utilizada para calcular la tortuosidad se determinan en función del tipo de punta de proyectil; así, para los casos en que la punta de proyectil sea de base no incurvada y no peduncular, dichos puntos se toman como el par de PFPD (izquierdo y derecho), por otra parte, cuando la base es no incurvada pero la punta de proyectil es peduncular, los puntos se establecen con el mismo procedimiento que se utilizó para determinar los PFPD para bases no incurvadas y puntas de proyectil no pedunculares (Sección 3.5). Finalmente, si la punta de proyectil es incurvada (independientemente de si es peduncular, o no) el par de puntos se establecen con el parámetro β utilizado en las funciones de ajuste de la base; más adelante se explicará como se obtiene dicho parámetro.

El par de características $\{\alpha, \beta\}$ y $\Delta_{\alpha\beta}$ se utilizan principalmente para discriminar a las bases incurvadas (categorías “muesca basal” y “cóncava”). En la categoría “cóncava” la zona incurvada se prolonga hasta los bordes de la hoja (vea la segunda categoría de la Figura 3.7), o hasta las paredes laterales del pedúnculo si la punta de proyectil es peduncular. En cambio, en la categoría “muesca basal”, dicha zona sólo abarca una parte de la base; esto es, se forma una “aleta basal”, la cual, conecta a la zona que posee inflexión con la hoja (o con la pared del pedúnculo). Esta característica se puede cuantificar si se postula que en la

categoría “cóncava” la aleta ocurre al final de la base y tiene ancho cero. El inicio y fin de la aleta serán vinculados, respectivamente, con los parámetros α y β y, su ancho con $\Delta_{\alpha\beta}$.

Para determinar las posiciones del contorno asociadas a α y β (para ambos lados de la punta de proyectil), se realizó un proceso de ajuste similar al implementado para establecer los PFPD en el caso de puntas de proyectil pedunculares (Sección 3.5). Debido a la simetría de la base de la punta de proyectil respecto al eje coordenado vertical (y), y que en este proceso de ajuste se utilizará el gradiente en dicha dirección, se debe considerar únicamente la función de ajuste del lado derecho de la punta de proyectil implementada en dicho proceso (proceso para establecer los PFPD en el caso de puntas de flecha pedunculares); esto es: sólo es necesario considerar las Ecuaciones 3.8 y 3.10; así, el super índice “der”, que hace referencia a la función de ajuste del lado derecho de dichas ecuaciones, puede ser omitido. Como se dijo anteriormente, en este ajuste se utiliza el gradiente de la dirección vertical ($\nabla_y[u]$) de la sección de contorno que inicia en el PCP (circulo blanco del contorno de las categorías “muesca basal” y “cóncava” que se presentan en la Figura 3.7), y que se extiende hasta las posiciones en los bordes de la hoja a la misma altura que este punto (circulo gris del contorno de estas categorías, en la Figura 3.7); de esta forma, en la Ecuación 3.10 la expresión $\nabla_x^{(der)}[u]$ es remplazada por $\nabla_y[u]$. El conjunto de regiones R utilizadas para evaluar el error, son las correspondientes regiones derechas definidas en la Ecuación 3.12, sólo que en la última de ellas ($R_{-H,-1}$), el límite superior del dominio (300) es remplazado por el número de elementos en $\nabla_y[u]$. Finalmente, con base a la experimentación realizada, las cuatro componentes de peso (p 's) se establecen igual a 0.25.

Con la finalidad de que el proceso sea robusto ante la excentricidad de la punta de proyectil, α y β se toman como el porcentaje de la longitud de la señal gradiente utilizada para calcularlos. Por último, en el vector de características se incluye sólo al parámetro α más pequeño, de entre el par que surge al evaluar ambos lados de la punta de proyectil; lo mismo se hace para el parámetro β . Se realiza este criterio porque existen algunas puntas de proyectil de base incurvada, pero asimétricas en dicha región; esto es, por un lado, ya sea izquierdo o derecho, tienden a ser del tipo “muesca basal” y por el otro “cóncavo”. En esta simplificación, se consideró el punto de vista del grupo arqueológico que colaboró en esta investigación, sugiriendo que se debe clasificar a estos ejemplares en la categoría “muesca basal”, ya que normalmente la presencia de la muesca indicaba una intención legítima de realizarla; a diferencia del extremo cóncavo, el cual, en esta situación pudo ser originado por desgaste del artefacto o al fracturarse esa parte de la punta de proyectil durante, o después, de su vida útil. La parte baja de la Figura 3.7 modela el proceso de ajuste del gradiente de cada una de las categorías de la base.

Las constantes α y β de las categorías no incurvadas (“convexa” y “recta”, es decir, las que no presentaron lóbulo central en el mapa CSS; Figura 3.7) se extraen con este mismo proceso, sólo que el par de funciones gradiente se extraen en la región que se extiende desde el PCP, hasta los puntos correspondientes que se utilizaron para extraer la tortuosidad de la base. Si se consideran imágenes ideales, como las de la Figura 3.7, la forma “convexa” tiene valores pequeños para α y β , mientras que en la forma “recta” α es pequeña y β es grande. Sin embargo, hay casos donde el ruido del contorno ocasiona que algunas de estas

condiciones no se cumplan; afortunadamente, esto no afecta significativamente al clasificador porque, como se dijo anteriormente, la tortuosidad es un parámetro más discriminante para este par de categorías.

3.8. Extracción de características de la punta de proyectil; paredes del pedúnculo

El vector de características de esta componente de la punta de proyectil está formado por:

1. El área del par de lóbulos pedunculares
2. El ángulo de las paredes del pedúnculo
3. $\Delta_{\alpha,\beta} = \beta - \alpha$
4. La distancia horizontal (d) de las paredes del pedúnculo a los PFPD

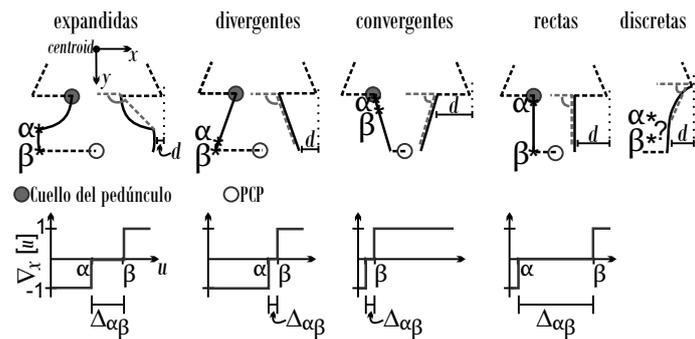


Figura 3.8: Características de las categorías de las paredes del pedúnculo.

Como se explicó en la detección de la frontera proximal–distal (Sección 3.5), las puntas de proyectil pedunculares tienen asociados un par de lóbulos simétricos en la región del mapa CSS que corresponde al pedúnculo. Esta característica permite discriminar entre las puntas de proyectil pedunculares y no pedunculares; para estas últimas, el área de dichos lóbulos se establece como 0.

El ángulo de las paredes del pedúnculo es una de las características más importantes para discriminar esta componente. La Figura 3.8 muestra el ángulo que corresponde al lado derecho de la punta de proyectil de cada categoría. En esta figura se puede observar que el ángulo se mide respecto a la horizontal y la línea recta que se forma al unir los extremos de la pared del pedúnculo. El inicio de la pared del pedúnculo se establece en el cuello del pedúnculo (posición del contorno de la punta de proyectil asociada a los lóbulos pedunculares del mapa CSS) y, el final como alguna de las correspondientes posiciones asociadas a los parámetros de ajuste α y β . De forma similar a los casos descritos anteriormente, ambos parámetros, α y β , surgen del proceso de ajuste del gradiente a funciones escalón unitario. En este caso, se considera el gradiente en dirección horizontal de la sección del contorno que inicia en el cuello del pedúnculo (círculo gris en las imágenes de la Figura 3.8) y termina en el PCP (círculo blanco en dichas imágenes). Las funciones de ajuste y error utilizadas en este proceso son las mismas que se utilizaron en la Sección 3.5 para establecer los PFPD de las puntas de proyectil del tipo peduncular (Ecuaciones 3.7, 3.8, 3.9 y 3.10). Además, también el conjunto de regiones R es el que se utiliza en este proceso (Ecuaciones 3.11 y 3.12), sólo que en este caso, en el conjunto $R_{H,1}^{(izq)}$ y $R_{-H,-1}^{(der)}$, la frontera superior del dominio (300) es remplazada, respectivamente, por el número de elementos en los gradientes $\nabla_x^{(izq)}[u]$ y $\nabla_x^{(der)}[u]$.

Finalmente, los pesos para ponderar a la función de error se establecieron empíricamente como:

$$\begin{aligned} 1^{\text{er}} \text{escalón} &\Rightarrow p_{-H,-1}^{(\text{izq})} = p_{H,1}^{(\text{der})} = 0.05; \quad p_{-H,0}^{(\text{izq})} = p_{H,0}^{(\text{der})} = 0.45 \\ 2^{\text{do}} \text{escalón} &\Rightarrow p_{H,0}^{(\text{izq})} = p_{-H,0}^{(\text{der})} = 0.2; \quad p_{H,1}^{(\text{izq})} = p_{-H,-1}^{(\text{der})} = 0.3 \end{aligned} \quad (3.16)$$

Los valores del subconjunto p 's del primer escalón se establecieron de esa forma por las mismas razones que en el proceso de ajuste utilizado para determinar los PFPD de las puntas de proyectil pedunculares (Sección 3.5). Por otra parte, los valores p 's del segundo escalón se establecieron de esa manera para colocar a β en la posición correcta cuando se considera a la categoría "expandida". Esta categoría presenta una sección que se ha denotado como $\overline{\alpha\beta}$, la cual corresponde al contorno comprendido entre α y β (vea la imagen de la categoría expandida de la Figura 3.8); idealmente, la orientación de $\overline{\alpha\beta}$ debe ser vertical y, consecuentemente, la parte del gradiente que le corresponde debe ser igual a 0, sin embargo, la forma "natural" de la punta de proyectil ocasiona que en varios casos ésta sea un poco convergente hacia la base, generando (en dicha parte del gradiente) algunas transiciones hacia 1, en el lado izquierdo, y -1 en el derecho, por lo cual, de no ser por el conjunto de valores p que se eligió, β sería asociada a valores más pequeños de los que le corresponden.

Como se dijo anteriormente, α o β son utilizados para establecer el final de las paredes del pedúnculo. En la categoría "divergente", al no existir una diferencia considerable entre ambos parámetros, cualquiera de ellos puede ser utilizado para este fin. Por otra parte, en la categoría "recta" debe utilizarse β , ya que α está muy cerca del cuello del pedúnculo. En la categoría "expandida" es preferible utilizar α , pues la sección $\overline{\alpha\beta}$ normalmente es larga y ocasiona que el ángulo del pedúnculo sea menor y se aproxime ligeramente a los valores que tendría en la categoría "recta". Por último, en la categoría "convergente", de manera ideal, tanto α como β son pequeños, sin embargo, en la mayoría de los casos, β toma valores que permiten una correcta discriminación al encontrarse a la mitad del pedúnculo, o al final de éste, ya que el ángulo del pedúnculo al no ser tan agudo (aproximadamente 80°) ocasiona que en la región "convergente", el gradiente haga varias transiciones hacia cero, aumentando así el valor de β . De esta forma, con base a un umbral obtenido experimentalmente se determinó cual de ambos parámetros debe ser utilizado; específicamente, se utiliza α si ésta es mayor que el 7% del número de elementos del dominio del gradiente, en caso contrario β es utilizada.

Los dos últimos elementos del vector de características para esta componente de la punta de proyectil, $\Delta_{\alpha,\beta}$ y d , se han incluido para discriminar entre sí a las categorías "expandida" y "divergente". La categoría "expandida" genera las "muescas laterales", como ejemplo de ellas puede referirse a las dos últimas puntas de proyectil de la Figura 4.4. Estas muescas se realizan al lasquear de forma transversal la parte baja de los bordes de la hoja, lo cual, genera la región $\overline{\alpha\beta}$ de la pared del pedúnculo; así, la distancia horizontal (d) entre dicha región y los bordes de la hoja es pequeña; por el contrario, el tamaño de dicha región ($\Delta_{\alpha,\beta}$) es grande. Adicionalmente, la categoría "divergente" genera las "muescas angulares", como

ejemplo de ellas puede referirse a las dos primeras puntas de proyectil de la Figura 4.4. Estas muescas se producen al remover toda la parte angular inferior de la lasca que da origen a la punta de proyectil, lo cual, ocasiona que la distancia horizontal (d) entre la pared del pedúnculo y los bordes de la hoja sea grande y, además, que $\Delta_{\alpha,\beta}$ sea pequeño. Para extraer el parámetro d es necesario definir la posición horizontal que representa a la pared del pedúnculo, dicha posición se estableció considerando ambos parámetros de ajuste α y β , por lo cual, la distancia d se calcula como el promedio de las distancias de las líneas horizontales que van de las posiciones en el contorno de estos parámetros a los PFPD.

Finalmente, todos los parámetros de esta componente se toman como el promedio de los valores obtenidos en cada lado de la punta de proyectil. Además, para obtener robustez ante la excentricidad, α y β se consideran como el porcentaje de la longitud de la señal gradiente.

3.9. Extracción de características de la punta de proyectil; hombros del pedúnculo

El vector de características de esta componente de la punta de proyectil está formado por:

1. El área del par de lóbulos pedunculares del mapa CSS
2. El ángulo del hombro del pedúnculo
3. La tortuosidad (T)
4. α ; parámetro de ajuste del gradiente
5. $\Delta_{\alpha,\beta} = \beta - \alpha$

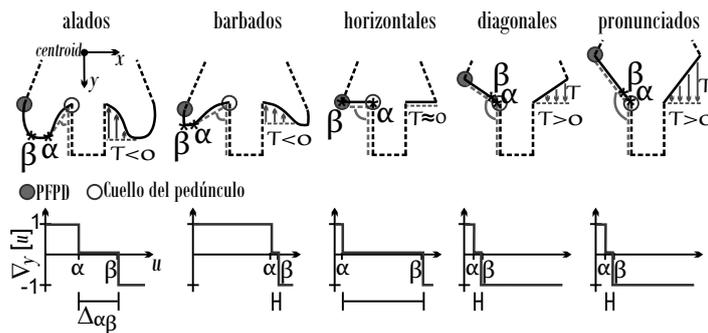


Figura 3.9: Características de las categorías de los hombros del pedúnculo.

El área de los lóbulos pedunculares en el mapa CSS se utiliza con el mismo enfoque que en el caso de la componente “paredes laterales” del pedúnculo. Por otra parte, las otras cuatro características se extraen mediante los parámetros α y β , los cuales, se extraen en el proceso de ajuste del gradiente; el cual, se realiza de forma similar que en la componente de la base y de las paredes del pedúnculo. Para realizar el proceso de ajuste de esta componente se utiliza el gradiente, en dirección vertical, a partir de la región del contorno que inicia en el cuello del pedúnculo (círculos blancos en las imágenes de la Figura 3.9) y termina en los PFPD (círculos grises en dichas imágenes). En la parte inferior de la Figura 3.9 se bosquejan los gradientes que corresponden a cada una de estas cinco categorías. Las funciones de ajuste, error y regiones delimitadoras del error son las mismas que se utilizaron en el proceso de ajuste de la base (Ecuaciones 3.7, 3.10 y 3.12). En la Ecuación 3.10 la expresión $\nabla_x^{(\text{der})}[u]$ debe ser remplazada por $\nabla_y[u]$ y, en la última región de la Ecuación 3.12 ($R_{-H,-1}$), el límite superior del dominio (300) debe ser remplazado por el número de elementos en el gradiente $\nabla_y[u]$. Finalmente, los pesos p 's en este proceso de ajuste, experimentalmente se

establecieron como:

$$\begin{aligned} 1^{\text{er}} \text{ escalón} &\Rightarrow p_{H,1} = 0.05; p_{H,0} = 0.45 \\ 2^{\text{do}} \text{ escalón} &\Rightarrow p_{-H,0} = 0.45; p_{-H,-1} = 0.05 \end{aligned} \quad (3.17)$$

Como se puede observar en la Figura 3.9, el ángulo de cada uno de los hombros del pedúnculo, idealmente, permite discriminar entre las categorías “horizontales”, “diagonales” y “diagonales pronunciados”. En el caso de las categorías “alados” y “barbados”, éstas, deben ser discriminadas con el resto de parámetros. El ángulo de cada uno de los hombros del pedúnculo se mide con respecto a la línea recta que une al cuello del pedúnculo y alguna de las posiciones correspondientes a los parámetros α , β , o los PFPD. Dicha posición se elige en función de cuál de los tres parámetros represente mejor la “inclinación del hombro”. Además, ésta posición también es utilizada para medir la tortuosidad (T) de esta componente. Específicamente, T se calcula respecto a la línea horizontal que va de este punto (α , β o PFPD) a la pared lateral del pedúnculo (componente horizontal en la imagen del cuello del pedúnculo; ver Figura 3.9).

En el caso de la categoría “alados”, la correspondiente posición de α es la que mejor representa la inclinación del hombro. Por otra parte, en la categoría “rectos”, β es una mejor opción para dicho propósito, ya que la posición de α es cercana al cuello del pedúnculo. En la categoría “barbados”, no existe diferencia considerable entre ambos parámetros α y β ; así, cualquiera de ellos puede ser utilizado. Finalmente, en las categorías “diagonales” y “diagonales pronunciados”, las posiciones de α y β son cercanas al cuello del pedúnculo; por lo cual, en este caso deben usarse los PFPD. Con base a la experimentación de este trabajo, se determinó que α debe ser utilizada si ésta es mayor que el 30 % de la longitud del dominio de la señal gradiente. Si esta condición no se cumple y β es mayor que el 50 % de dicha señal, se utiliza este segundo parámetro y, si ambas condiciones no se satisfacen se utilizan los PFPD.

En las categorías “alados” y “barbados”, el enfoque asignado a α y β es similar al utilizado en el caso de las categorías “cóncava” y “muesca basal” de la componente base. Esto es, α y β representan, respectivamente, el inicio y el final de la “aleta” del hombro del pedúnculo. Por otra parte, similarmente al caso de la base, ambos parámetros asisten en la clasificación de las otras categorías de esta componente. Para obtener robustez ante la excentricidad, en las dos últimas componentes del vector de características (α y $\Delta_{\alpha,\beta}$) α y β se normalizan de acuerdo a la razón entre la distancia horizontal medida desde sus respectivas posiciones en el contorno hasta el cuello del pedúnculo, y la distancia horizontal medida desde los PFPD hasta el cuello del pedúnculo. Finalmente, de manera similar a las otras tres componentes, las características de los hombros del pedúnculo (excepto el área) se calculan como el promedio de los valores obtenidos en ambos lados de la punta de proyectil.

Capítulo 4

Experimentos y resultados

El desempeño del algoritmo se evaluó de dos formas distintas. Por un lado, se evaluó, individualmente, el desempeño de la clasificación de cada componente de la punta de proyectil; por otro lado, se evaluó el desempeño de la clasificación de su forma global; este segundo enfoque se propuso para comparar el algoritmo de clasificación que se generó en este trabajo, respecto a los descriptores de forma PHOG, HOOSC y momentos geométricos de Hu. Los resultados de dicha evaluación son presentados con las siguientes tres métricas: precisión (P), sensibilidad (*recall*, por su terminología en inglés) (R) y la medida- F ($F1$ -*measured*, por su terminología en inglés). A continuación se define cada una de ellas.

La precisión obtenida para una categoría que forma parte del proceso de clasificación, se define como el índice de elementos recuperados que pertenecen a esa categoría (VerdaderosPositivos), respecto al total de elementos clasificados bajo dicha categoría (VerdaderosPositivos + FalsosPositivos); esto es:

$$P = \frac{\text{VerdaderosPositivos}}{\text{VerdaderosPositivos} + \text{FalsosPositivos}} \quad (4.1)$$

La sensibilidad (*recall*, R) obtenida para una categoría que forma parte del proceso de clasificación, se define como el índice de elementos recuperados que pertenecen a esa categoría (VerdaderosPositivos), respecto al total de elementos de dicha categoría presentes en la base de datos que se clasificó (VerdaderosPositivos + FalsosNegativos); esto es:

$$R = \frac{\text{VerdaderosPositivos}}{\text{VerdaderosPositivos} + \text{FalsosNegativos}} \quad (4.2)$$

Como se puede observar en la Ecuación 4.1, la precisión evalúa la asertividad del sistema al clasificar cierta categoría. Sin embargo, este parámetro, si no es presentado con la sensibilidad de dicha categoría, no es concluyente para generar una decisión objetiva del desempeño del sistema; ya que, un alto valor en la precisión (considerando cierta categoría), pudo haberse obtenido al clasificar pocos elementos de dicha categoría, respecto al total que existen de ella en la base de datos. Así, estos dos parámetros (precisión y sensibilidad) deben presentarse juntos al evaluar el desempeño de un clasificador. De esta forma, la medida- F ($F1$ -*measured*) es un índice que presenta un balance entre ambos parámetros. Dicho índice se define como:

$$\text{medida-}F = 2 \frac{\text{Precisión} \times \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}} \quad (4.3)$$

Para brindar mayor confiabilidad estadística en los resultados presentados por medio de estas métricas, se decidió presentar dichos resultados como el promedio de 10-iteraciones de cada uno de los procesos relacionados a ellos.

4.1. Evaluación del desempeño de la clasificación de las componentes de la punta de proyectil

Para evaluar el desempeño del algoritmo respecto a la clasificación de cada una de las componentes de la punta de proyectil, el equipo de arqueólogos que participó en esta investigación, formó cada categoría con 50 imágenes. Esto también permitió determinar con que método de entrenamiento se obtienen los mejores resultados (KNN, discriminación lineal, árboles de decisión y SVM). Finalmente, esto permitió evaluar la dependencia del algoritmo respecto al número de imágenes de entrenamiento. La Figura 4.1 presenta los resultados de estos enfoques. Ambos gráficos contienen el promedio de la medida- F de las categorías de cada componente. Además, los resultados de la gráfica izquierda se obtuvieron al utilizar 30 imágenes por categoría durante el entrenamiento y el resto de ellas (20 imágenes por categoría) durante la clasificación. Por su parte, en los resultados de la gráfica derecha se utilizó el método SVM.

Como resultado de la experimentación, en la gráfica izquierda de la Figura 4.1 se puede observar que el método de clasificación SVM representa la mejor opción para ser utilizado con nuestro algoritmo, ya que en tres de las cuatro componentes de la punta de proyectil, obtiene los mejores resultados y, en el caso de los bordes de la hoja obtiene el segundo mejor resultado. Por otra parte, en la gráfica derecha se puede observar que a partir de 20 imágenes de entrenamiento por categoría, a excepción de los "bordes de la hoja", el algoritmo no depende en gran medida del número de imágenes consideradas en el entrenamiento; esta característica quizá no sea tan relevante en este enfoque de clasificación, pues, como se mencionó en la introducción de esta tesis, una imagen puede ser utilizada para entrenar distintas componentes y, además, el total de categorías es sólo la suma de las categorías de cada componente. Sin embargo, en el enfoque de clasificación global, esto puede ser muy importante, ya que en este caso, existen cientos de categorías (formadas por la combinación de las categorías de las componentes) y no siempre existen las suficientes imágenes en la base de datos para entrenar a algunas de ellas.

Para conocer de manera explícita el desempeño del algoritmo al clasificar cada una de las categorías que conforman las distintas componentes de la punta de proyectil, en la Figura 4.2 se presentan los resultados obtenidos respecto a las tres métricas utilizadas en la evaluación del desempeño (precisión, sensibilidad y medida- F). Nuevamente se utilizaron 30 imágenes por categoría durante el entrenamiento y las 20 restantes (por categoría) para la clasificación. Además de esto, para enriquecer el material estadístico de análisis, en la Figura 4.3 se presentan las matrices de confusión de este proceso. En las gráficas de la Figura 4.2 se puede observar que el desempeño general del algoritmo, considerando de manera conjunta a las cuatro componentes de la punta de proyectil, es aproximadamente igual a 70 %, tomando en cuenta que dicho valor se obtiene al promediar los resultados de la medida- F de

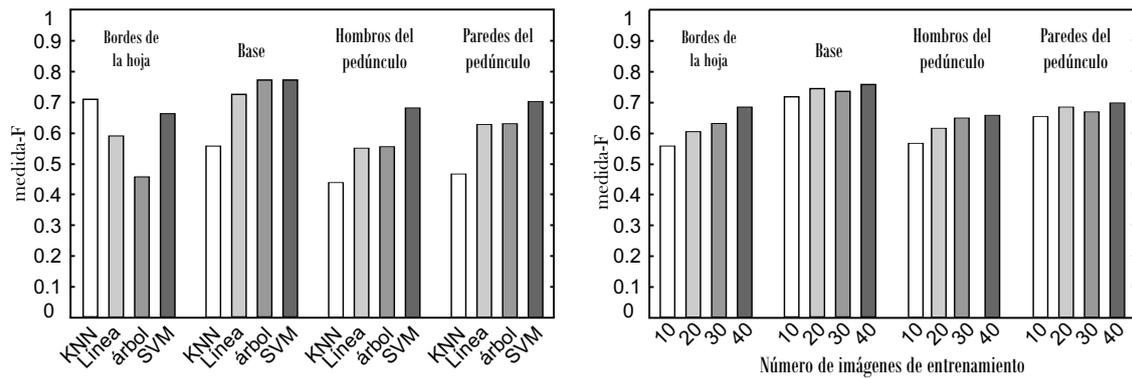


Figura 4.1: Gráfica izquierda: desempeño de los cuatro métodos elegidos como candidatos para entrenar el descriptor generado; estos resultados se obtuvieron al utilizar 30 imágenes por categoría durante el entrenamiento y el resto durante la clasificación (20 imágenes por categoría). Gráfica derecha: dependencia del algoritmo respecto al número de imágenes de entrenamiento por categoría (10, 20, 30, 40); en este caso se utilizó el método SVM con un núcleo cuadrático. En ambas gráficas cada una de las cuatro componentes de la punta de proyectil son presentadas como el promedio de la medida-F de sus respectivas categorías.

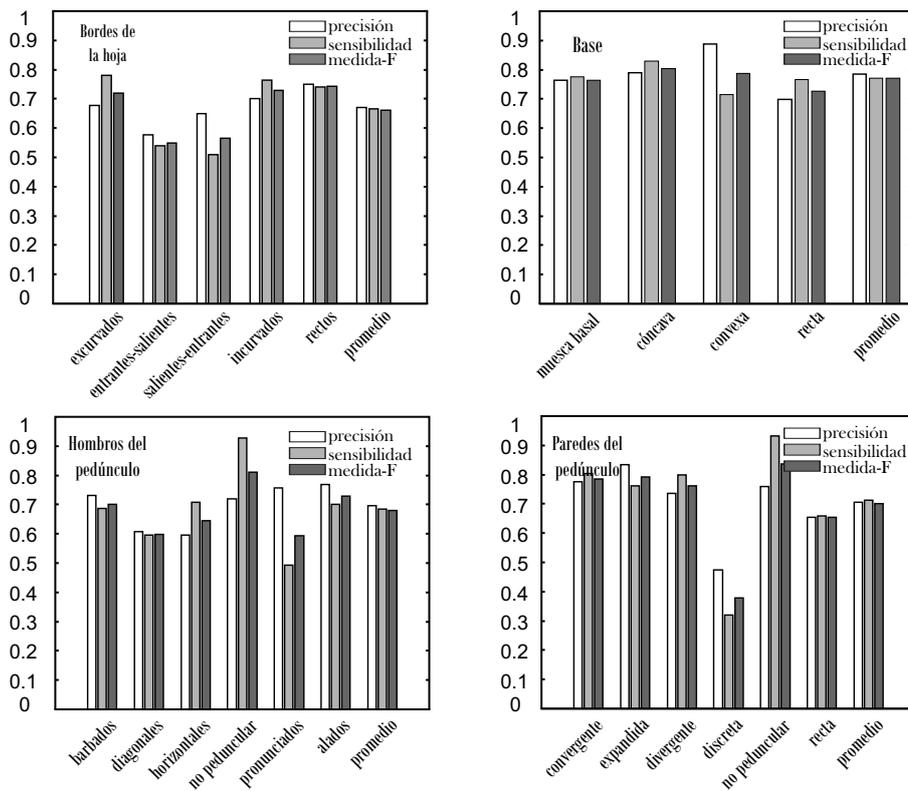


Figura 4.2: Resultados del desempeño del algoritmo desarrollado. Cada gráfica presenta la precisión (P), la sensibilidad (R) y la medida-F para las categorías de cada componente de la punta de proyectil. Estos resultados se obtuvieron al utilizar 30 imágenes por categoría durante el entrenamiento y el resto durante la clasificación (20 imágenes por categoría).

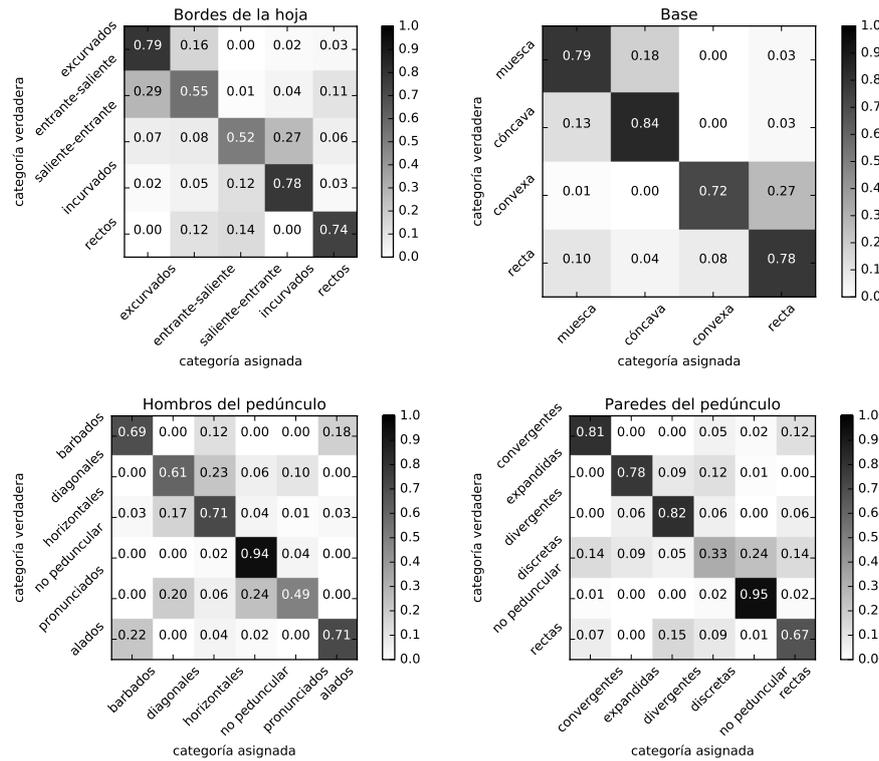


Figura 4.3: Matriz de confusión de cada una de las componentes obtenidas en el mismo proceso descrito para obtener los resultados de la Figura 4.2.

todas las categorías. Adicionalmente, la medida F del promedio de las categorías de la base (última barra de esta categoría) es ligeramente superior a las demás (78 %) y, éste decrece un poco al considerar los “bordes de la hoja” (67 %).

En la gráfica de la categoría “bordes de la hoja” se puede observar que la precisión y la sensibilidad que se obtuvieron en las categorías “entrantes-salientes” y “salientes-entrantes” tienen valores más bajos que en el resto de categorías de esta componente. El comportamiento de las partes de la matriz de confusión que corresponden a este par de categorías se asocia a la premisa mencionada en la Sección 1.2: “para construir la base de datos utilizada en este trabajo, se decidió incluir algunas imágenes para las cuales, la categoría asociada a ellas no es tan evidente; esto con la finalidad de evitar sobre-ajustar el entrenamiento del clasificador y, así, cumplir con el objetivo de asistir al arqueólogo en la clasificación de aquellas imágenes que parecen estar en la frontera de dos, o más categorías.” De esta forma, es más probable que dichas imágenes sean clasificadas erróneamente por el sistema en una categoría similar. Específicamente, algunas imágenes de la categoría “entrantes-salientes” fueron asignadas a las categorías “excurvados” y “rectos”, lo cual, es responsable de los bajos valores de sensibilidad obtenidos en dicha categoría. Por otra parte, la baja precisión que presenta esta categoría (“entrantes-salientes”), se debe a que algunas de las imágenes de la categoría “excurvados” y, otras más de la categoría “rectos”, fueron asignadas a ella. Se observa un comportamiento similar en la categoría “salientes-entrantes”, para la cual, algunos de sus elementos fueron asignados a la categoría “incurvados” y, viceversa.

Como se dijo anteriormente, la componente “base” presenta resultados satisfactorios. Esto adquiere aún más importancia al considerar el parecido de las componentes “muesca basal” y “cóncava” de esta categoría. Por otra parte, respecto a la componente “hombros del pedúnculo”, se puede observar que la categoría “diagonales pronunciados” tiene un bajo valor de sensibilidad (aproximadamente 50%), una vez más, en la matriz de confusión se puede observar que los errores de clasificación suceden principalmente al asignar algunos de sus elementos a las categorías “diagonales” y “no peduncular”. La explicación a dichos errores, al considerar las imágenes que se han asignado como elementos de la categoría “diagonal”, está relacionada con la similitud que existe entre ambas categorías (“diagonal” y “diagonal pronunciada”). Sin embargo, en el caso de la asignación a la categoría “no peduncular”, esta explicación no es tan obvia y, está relacionada con el hecho de que las imágenes erróneamente clasificadas fueron consideradas por el algoritmo como “no pedunculares” en la etapa de detección de los PFPD, pues al no tener una inflexión tan pronunciada en la parte del pedúnculo, los lóbulos pedunculares del mapa CSS no son tan sobresalientes, lo que impide superar el umbral de altura mínima y/o las condiciones de simetría de las Ecuaciones 3.5 y 3.6. Por otra parte, el algoritmo muestra un desempeño aceptable en el resto de las categorías de esta componente, incluyendo las del tipo “barbadas” y “aladas”, las cuales, son muy similares entre sí. En el caso de la componente “paredes del pedúnculo”, se puede apreciar que la categoría “pedúnculo discreto” no tiene un desempeño satisfactorio, ya que algunas de sus imágenes son asignadas a la categoría “no peduncular”. La explicación a este comportamiento es la misma que corresponde al error que se comete al asignar algunos de los elementos de la categoría “diagonales pronunciados” (de la componente “hombros del pedúnculo”) como “no pedunculares”. Además, otras de estas imágenes (de la categoría “pedúnculo discreto”) son asignadas a las categorías “rectas” y “convergentes”, ya que existe cierta similitud entre ellas. Finalmente, las categorías “expandidas” y “convergentes” son clasificadas exitosamente a pesar de la gran similitud que existe entre ellas.

4.2. Evaluación del desempeño de la clasificación de la forma global de la punta de proyectil

La evaluación del desempeño del algoritmo respecto a la clasificación de la forma global de la punta de proyectil, se realizó para comparar a éste con otros algoritmos conocidos en la literatura de clasificación de forma, como: PHOG, HOOSC y momentos invariantes geométricos de Hu. Adicionalmente, este enfoque permite evaluar al algoritmo en un ambiente más directo respecto al trabajo cotidiano de los arqueólogos. El equipo arqueológico que colaboró en la investigación conformó las categorías de la forma global de la punta de proyectil, en esta ocasión cada categoría se formó con 40 imágenes, ya que, como se dijo en la introducción de esta tesis, las categorías que consideran la forma global son más complicadas de integrar porque algunas de ellas cuentan con pocos ejemplares en las bases de datos. Considerando esta situación, y la relevancia arqueológica de las distintas formas de punta de proyectil, en este trabajo se propusieron 9 categorías representativas para realizar esta evaluación. La Figura 4.4 muestra dichas categorías.



Figura 4.4: Se presentan las 9 categorías que se utilizaron para evaluar el desempeño del algoritmo respecto a la forma global de la punta de proyectil.

Dado que los momentos invariantes de Hu se extraen de forma directa (sin ajustar parámetros), estos constituyen el propio vector de características de dicho descriptor. Además, el clasificador SVM fue utilizado con este descriptor, pues presentó el mejor desempeño respecto a otros métodos de clasificación como *KNN*, discriminación lineal, cuadrática y cúbica y, árboles de decisión binaria.

El método “bolsa de palabras visuales” (BoVW, *bag of visual words*) fue utilizado con los descriptores PHOG y HOOSC. Por otra parte, nuevamente el clasificador SVM fue utilizado, ya que éste obtuvo los mejores resultados respecto a los métodos de clasificación mencionados anteriormente con el descriptor de Hu. Este par de descriptores (PHOG y HOOSC) tienen asociados parámetros de ajuste que permiten mejorar su desempeño, por lo cual, se experimentó con un conjunto de valores para estos parámetros y, así, utilizar aquellos con los que se obtengan los mejores resultados. Respecto a la bolsa de palabras visuales, se evaluó el número de palabras utilizadas para conformar el diccionario; para ellos se consideraron desde 100 hasta 500 palabras, con pasos de 100 palabras. Con base a los resultados obtenidos, se decidió utilizar 300 palabras (para ambos descriptores), ya que a partir de este valor, el desempeño es similar (y mejor que al considerar 100 y 200 palabras), con la ventaja de que el tiempo de procesamiento es menor respecto a los casos de 400 y 500 palabras. Los parámetros evaluados para el método de PHOG fueron los siguientes: tamaño de la celda (longitud de los lados del cuadrado) considerando el conjunto de valores: $\{4, 8, 16, 24, 32, 40\}$ y, los niveles de la pirámide; considerando desde 0 hasta 3 niveles. Los mejores resultados se obtuvieron con un tamaño de celda igual a 32 unidades y un nivel de la pirámide igual a 2. Adicionalmente, el tamaño del bloque y el número de orientaciones para cuantizar al ángulo del gradiente, se fijaron en 2×2 y 9 orientaciones, respectivamente. En el caso del descriptor HOOSC, el radio del círculo utilizado para extraer los gradientes se evaluó con los siguientes valores: $\{1, 1/2, 1/4\}$ de la longitud del eje principal de la punta de proyectil; el valor correspondiente a $1/4$ fue con el que se obtuvieron los mejores resultados. Adicionalmente, y siguiendo el criterio utilizado en el trabajo de HOOSC [Gülcan et al., 2015], el resto de parámetros de este método se fijó en: 2 para el número de anillos de extracción de vectores de características, 8 para el número de particiones de un anillo y 8 para las orientaciones en las que se cuantiza al ángulo del gradiente.

Para conocer la dependencia de cada uno de estos cuatro métodos (HOOSC, momentos de Hu, nuestro y PHOG) respecto al número de imágenes utilizadas en el entrenamiento, la clasificación se evaluó al considerar los siguientes números de imágenes de entrenamiento por categoría: $\{5, 10, 20, 30\}$. El proceso fue diseñado considerando que dichos métodos fueran entrenados con el mismo conjunto de imágenes. El valor de 5 se ha incluido en esta

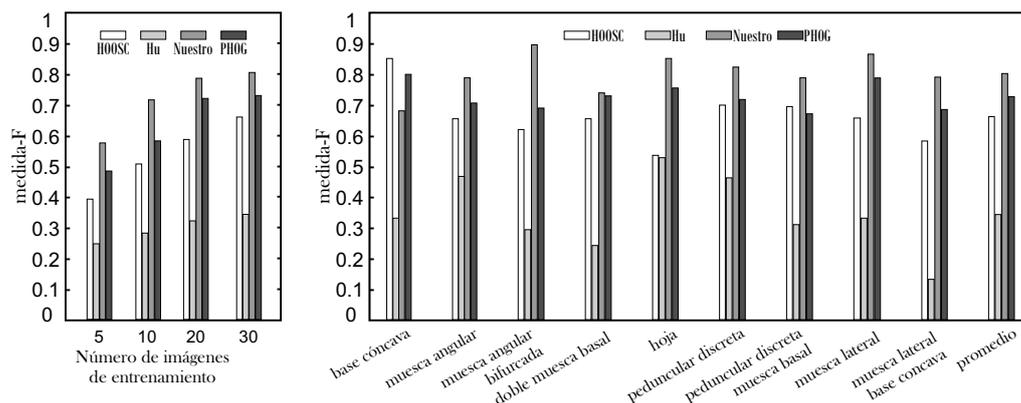


Figura 4.5: Gráfica izquierda, desempeño del algoritmo en función del número de imágenes por categoría utilizadas en el entrenamiento; estos resultados son el promedio de la medida- F de las 9 categorías consideradas en la forma global de la punta de proyectil. Gráfica derecha, resultados de la clasificación (medida- F) de cada una de las nueve categorías y el promedio de ellas; en este caso, se utilizaron 30 imágenes por categoría durante el entrenamiento y el resto (10 imágenes por categoría) durante la clasificación.

evaluación porque, como se dijo al inicio de este capítulo, en este enfoque de clasificación global, varias bases de datos no cuentan con suficientes imágenes para entrenar algunas de sus categorías; de las cuales, algunas de ellas sólo cuentan con un número de imágenes similar a éste valor (5).

En la Figura 4.5 se presentan los resultados del desempeño de los cuatro métodos. La gráfica izquierda de esta figura, muestra la dependencia de cada uno de ellos respecto al número de imágenes por categoría utilizadas en el entrenamiento. Dichos resultados se presentan como el promedio de la medida- F de la clasificación de las nueve categorías utilizadas. Por otra parte, la gráfica derecha muestra el desempeño de estos algoritmos respecto a cada una de las categorías. Los resultados de esta gráfica se obtuvieron con un proceso en el que se utilizaron 30 imágenes por categoría para entrenar y el resto (10 imágenes por categoría) para clasificar. Adicionalmente, la tabla 4.1 presenta los resultados de la gráfica izquierda de la Figura 4.5 y, además, incluye las otras dos métricas consideradas en el trabajo (precisión y sensibilidad); esto es, presenta los resultados de nuestro algoritmo, y de los otros tres con los que éste fue comparado, respecto al promedio de las nueve categorías a clasificar en función del número de imágenes por categoría utilizadas en el entrenamiento ($\{5, 10, 20, 30\}$). Finalmente, esta tabla también presenta, para cada uno de los cuatro algoritmos, el tiempo promedio de cómputo necesario para clasificar una imagen; esta información será comentada al final de la sección.

Como se puede observar en las gráficas de la Figura 4.5 y en la tabla 4.1, nuestro algoritmo y el método de PHOG obtuvieron los mejores resultados. Específicamente, la gráfica izquierda de esta figura, muestra que nuestro algoritmo presenta el mejor desempeño respecto a todos los valores considerados en la evaluación del número de imágenes de entrenamiento; por otra parte, PHOG ocupa la segunda posición en cada uno de estos valores. Además, en dicha gráfica se puede observar que, considerando nuestro método, al utilizar

Tabla 4.1: Desempeño del método de clasificación desarrollado en este trabajo, y de los métodos utilizados para compararlo, respecto a la clasificación de las nueve categorías propuestas de la forma global de la punta de proyectil (Figura 4.4). La primera columna, “# imgs”, indica el número de imágenes por categoría utilizadas para entrenar; mientras que, la última columna representa el tiempo promedio de cómputo necesario para clasificar una imagen.

método	# imgs	precisión	sensibilidad	medida-F	tiempo [s]
HOOSC	5	0.44	0.39	0.38	0.55
	10	0.53	0.50	0.50	
	20	0.60	0.58	0.58	
	30	0.68	0.66	0.65	
Hu	5	0.27	0.25	0.24	0.54
	10	0.29	0.29	0.28	
	20	0.34	0.32	0.32	
	30	0.35	0.35	0.34	
Nuestro	5	0.63	0.60	0.57	1.00
	10	0.73	0.71	0.71	
	20	0.80	0.78	0.78	
	30	0.82	0.80	0.80	
PHOG	5	0.50	0.48	0.48	0.93
	10	0.60	0.58	0.58	
	20	0.73	0.72	0.71	
	30	0.75	0.72	0.72	

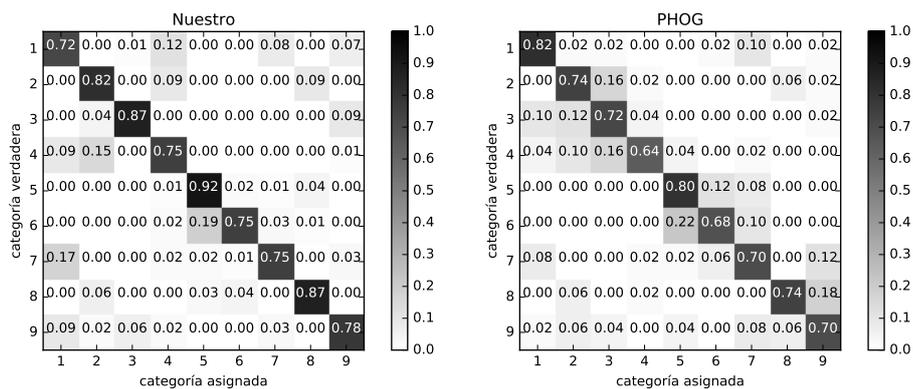


Figura 4.6: Matriz de confusión de nuestro algoritmo y del método de PHOG. Las etiquetas de la matrices corresponden a: 1→base cóncava, 2→muesca angular, 3→muesca angular bifurcada, 4→doble muesca basal, 5→hoja, 6→peduncular discreta, 7→peduncular discreta con muesca basal, 8→muesca lateral, 9→muesca lateral base cóncava.

20 y 30 imágenes de entrenamiento por categoría, los resultados son muy similares, con un valor de 80 % para la medida- F . Esto da una referencia del número de imágenes necesarias para entrenar una categoría asociada con la forma global de la punta de proyectil; esto es importante al tener en cuenta la posibilidad de adaptar el algoritmo para que el usuario pueda entrenar sus propias categorías. Finalmente, el desempeño que se obtuvo con nuestro algoritmo al considerar sólo 5 imágenes de entrenamiento por categoría, es igual al 57 % (medida- F), el cual, puede ser un resultado alentador para quien desee utilizar el algoritmo en bases de datos que contengan categorías con muy pocas imágenes (del orden de unidades). En esta misma gráfica se puede observar que en esta evaluación el método HOOSC presenta un desempeño regular y, los momentos invariantes de Hu no presentan resultados satisfactorios.

En la gráfica derecha de la figura 4.5 se observa que el algoritmo desarrollado en esta investigación ("Nuestro"), presenta, ligeramente, mejores resultados que el método de PHOG. Esto se puede apreciar, por una parte, en un contexto general, el cuál se representa como el promedio de las categorías (última barra del gráfico) y, por otra parte, respecto a cada categoría, excepto en el caso de la categoría "base cóncava". En lo subsecuente el análisis se enfocará en ambos métodos. En la Figura 4.6 se presentan sus matrices de confusión. Dichas matrices permiten confirmar el buen desempeño de ambos métodos, ya que en ellas se aprecia que los pocos errores cometidos en la clasificación, normalmente ocurren entre categorías muy similares, como es el caso de las categorías "base cóncava" y "peduncular discreta con muesca basal", "muesca angular" y "doble muesca basal" y, "peduncular discreta" y "hoja". Por otra parte, la parte de las matrices de confusión que corresponde a la categoría "muesca angular bifurcada" (categoría número 3), muestra que de ambos métodos, sólo PHOG clasifica algunos de estos elementos en la categoría "muesca angular" (categoría número 2). Como se puede observar en la Figura 4.4, estas dos categorías sólo difieren por la muesca de la base. El hecho de que el algoritmo desarrollado en esta investigación realizara de forma correcta esta clasificación, es favorable para mostrar que éste cumple con el objetivo de distinguir a las categorías aún si éstas sólo tienen pequeñas diferencias entre ellas.

Finalmente, como parte de la evaluación del desempeño de los cuatro métodos presentados en esta experimentación, en la Figura 4.7, y en la última columna de la tabla 4.1, se presentan los tiempos de procesamiento de cada uno de ellos. Para la Figura 4.7, el tiempo total se encuentra dividido en las tres etapas principales del proceso: 1) extracción de características, 2) entrenamiento y, 3) clasificación. Para tener una referencia de los resultados reportados, se debe considerar que los experimentos se realizaron en una computadora personal con sistema operativo win7© procesador Intel® Core(TM) i7-6700 (3.4Ghz), 16GB de memoria RAM y las tres principales etapas del procesamiento se implementaron en MATLAB© 9.10 (R2016b). Además, dichos resultados se obtuvieron respecto a la base de datos utilizada para clasificar las 9 categorías representativas de la forma global de la punta de proyectil y, estos se presentan como el promedio de todas las imágenes que intervienen en dicha base de datos. Como puede apreciarse en esta figura, el método desarrollado en esta investigación es el más lento en la etapa de extracción de características (aproximadamente 0.98 seg.), sin embargo, en las otras dos etapas es el más rápido, ya que su vector de

características es sólo de 16 componentes. A pesar de que el método de Hu también tiene asociado un vector de dimensión pequeña (7 componentes), éste no obtiene tan bajos valores en dichos procesos porque la aglomeración de los vectores de características que genera es muy baja. Por otra parte, el método de PHOG es el más rápido en la etapa de extracción de características, sin embargo, en las otras dos etapas es el más lento, pues sus vectores de características son de dimensión 6300 (se utilizó un diccionario de 300 palabras y el segundo nivel de la pirámide). Es por ello que el tiempo de procesamiento total del algoritmo desarrollado en esta investigación y del método de PHOG es similar (el algoritmo desarrollado en esta investigación es aproximadamente 0.07 seg. más lento); dicha información es relevante si se consideran que ambos métodos fueron los que presentaron mejores resultados de desempeño. Finalmente, se puede observar que HOOSC y los momentos invariantes de Hu tienen similares, y los mejores, tiempos de procesamiento, sin embargo, esto no es muy alentador, ya que ellos no presentan los mejores resultados en el proceso de clasificación.

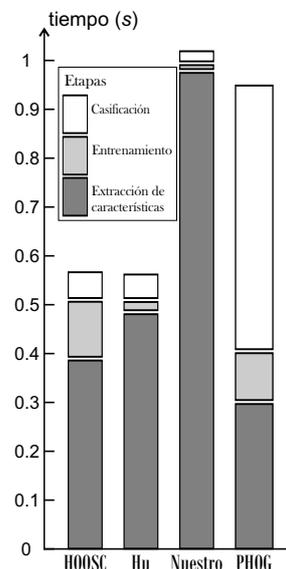


Figura 4.7: Tiempo de procesamiento de cada uno de los cuatro métodos evaluados. Las barras verticales muestran el tiempo total dividido en las tres etapas principales: 1) extracción de características, 2) entrenamiento y, 3) clasificación.

Capítulo 5

Conclusiones

En este trabajo se desarrolló un algoritmo para la clasificación automática de puntas de proyectil arqueológicas, el cual, es capaz de desempeñarse satisfactoriamente en la mayoría de bases de datos disponibles. En contraste con el paradigma arqueológico de clasificación de estos artefactos, el diseño secuencial de este algoritmo, el cual, considera la división automatizada de la punta de proyectil en cada una de sus componentes básicas, permite comparar bases de datos muy diversas, enfrentando así el problema relacionado con la gestión y entrenamiento de los cientos de categorías existentes. Esta metodología favorece la realización de estudios más globales en este tema de investigación arqueológica (clasificación de puntas de proyectil). Finalmente, los resultados obtenidos muestran que el algoritmo desarrollado presenta un desempeño aceptable, el cual, ha sido cuantificado por los indicadores estadísticos “medida- F ”, “precisión” y “sensibilidad”. Así, el valor promedio de la medida- F , considerando todas las categorías que intervienen en el proceso de clasificación, es aproximadamente igual al 70 %. Sin embargo, se pretende mejorar dicho desempeño al modificar ligeramente algunos de los aspectos del algoritmo, como aquel que concierne a la clasificación de la componente “bordes de la hoja”.

Una de las limitaciones del algoritmo desarrollado se deriva del hecho que la clasificación morfológica de puntas de proyectil no permite contextualizar directamente algunos aspectos de estos artefactos como: grupo ancestral asociado, funcionalidad del artefacto, temporalidad, evolución de su vida útil, propiedades aerodinámicas, etc. Por lo cual, se pretende ampliar este trabajo de tesis, integrando, al algoritmo de clasificación, métodos de análisis y gestión de la información relacionados con las bases de datos de las puntas de proyectil. Esto favorecerá el acceso a la información en este tema de estudio, lo cual, permitirá a los investigadores del área realizar estudios más globales de estos artefactos líticos. Finalmente, se desea mejorar algunos aspectos del programa desarrollado, como ofrecer al usuario la posibilidad de entrenar las familias de sus bases de datos, al sólo suministrar las imágenes que conforman a ésta. Esto permitirá que el proceso sea más sencillo e intuitivo de manipular para los arqueólogos.

El software desarrollado es libre para usos académicos y de investigación; éste puede ser solicitado desde nuestro sitio web: <http://matahari.fi-p.unam.mx/~inge-flechas>. La disponibilidad del software desarrollado puede ayudar a generar un ambiente colaborativo entre distintos grupos arqueológicos interesados en esta área de investigación. Además, consideramos que es una oportunidad favorable para enriquecer la base de datos utilizada

en la investigación, la cual, también deseamos liberar para usos académicos y de investigación una vez que ésta se encuentre completa con imágenes que no restrinjan su distribución por cuestiones de derechos de autor.

Lista de publicaciones orales y escritas del trabajo

83rd Annual Meeting of the Society for American Archaeology. Washington D.C., 12 abril del 2018. **Automatic classification of digital images of archaeological arrowheads**; modalidad: "Poster".

Quinto Congreso Internacional El Patrimonio Cultural y las Nuevas Tecnologías. Ciudad de México, 7 de noviembre del 2018. **Algoritmo computacional para la clasificación automatizada de puntas de proyectil arqueológicas**; modalidad: "Presentación oral".

ACM Journal on Computing and Cultural Heritage. **Computer algorithm for archaeological projectile points automatic classification**. JOCCH-18-0099.R1. Aceptado para su publicación el 10 de diciembre del 2018.

Bibliografía

- [Abbassi et al., 1999] Abbassi, S., Mokhtarian, F., and Kittler, J. (1999). Curvature scale space image in shape similarity retrieval. *Multimedia Systems*, 7:467–476.
- [Abdel-Hakim and Farag, 2006] Abdel-Hakim, A. E. and Farag, A. A. (2006). Csift: A sift descriptor with color invariant characteristics. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1978–1983, New York, NY, USA. IEEE.
- [Alcantarilla et al., 2012] Alcantarilla, P. F., Bartoli, A., and Davison, A. J. (2012). Kaze features. In *European Conference on Computer Vision*, pages 214–227.
- [Arratia and Mirambell, 2005] Arratia, L. G. and Mirambell, L. (2005). *Reflexiones sobre la industria lítica*. CNCA-INAH, México, D.F.
- [Bay et al., 2008] Bay, H., Ess, A., Tuytelaars, T., and Gool, L. (2008). Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110:346–359.
- [Belongie et al., 2002] Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on pattern analysis and machine intelligence*, 24:509–522.
- [Bookstein, 1997] Bookstein, F. L. (1997). Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Medical Image Analysis*, 1:225–243.
- [Bradley and Okumura., 2016] Bradley, B. and Okumura., M. (2016). Getting from the point; combining geometric morphometrics and technology to investigate projectile point assemblages from southeastern brazil. In *8th World Archaeological Congress, Quioto*.
- [Bribiesca, 1999] Bribiesca, E. (1999). A new chain code. *Pattern Recognition*, 32:235–251.
- [Bribiesca, 2013] Bribiesca, E. (2013). A measure of tortuosity based on chain coding. *Pattern Recognition*, 46:716–724.
- [Buchanan et al., 2011] Buchanan, B., Collard, M., Hamilton, M. J., and O’Brien, M. J. (2011). Points and prey: a quantitative test of the hypothesis that prey size influences early paleoindian projectile point form. *Journal of Archaeological Science*, 38:852–864.
- [Bullitt et al., 2003] Bullitt, E., Gerig, G., Pizer, S. M., Lin, W., and Aylward, S. R. (2003). Measuring tortuosity of the intracerebral vasculature from mra images. *IEEE TRANSACTIONS ON MEDICAL IMAGING*, 22:1163–1171.

- [Calonder et al., 2010] Calonder, M., Lepetit, V., Strecha, C., and Fua, P. (2010). Brief: Binary robust independent elementary features. In *European Conference on Computer Vision*, pages 778–792.
- [Cootes et al., 1995] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models-their training and application. *Computer vision and image understanding*, 6:38–59.
- [Dalal and Triggs, 2005] Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition*, pages 886–893. Computer Society Conference, IEEE.
- [Duda and Hart, 1972] Duda, R. O. and Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *ACM Communications*, 15:11–15.
- [Duda et al., 2000] Duda, R. O., Hart, P. E., and Stork, D. G. (2000). *Pattern Classification*. Wiley-Interscience, New York, NY, USA, segunda edition.
- [Freeman and Saghri, 1978] Freeman, H. and Saghri, A. (1978). Generalized chain codes for planar curves. In *4th International Joint Conference on Pattern Recognition*, pages 701–703, Kyoto, Japan.
- [Gonzalez and Woods, 2002] Gonzalez, R. C. and Woods, R. E. (2002). *Digital Image Processing*. Prentice Hall.
- [Goodfellow et al., 2016] Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press. <http://www.deeplearningbook.org>.
- [Gu, 1995] Gu, C. (1995). *Ph.D thesis. Multivalued Morphology and Segmentation-based Coding*. Lab. (LTS) of Swiss Federal Institute of Technology at Lausanne (EPFL).
- [Gülcan et al., 2015] Gülcan, C., Odobez, J. M., and Daniel, G. P. (2015). Evaluating shape representations for maya glyph classification. *ACM Journal on Computing and Cultural Heritage*, 9.
- [Harris and Stephens, 1988] Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *Fourth Alvey Vision Conference*, pages 147–151.
- [Hietanen et al., 2016] Hietanen, A., Lankinen, J., Kämäräinen, J.-K., Buch, A. G., and Krüger, N. (2016). A comparison of feature detectors and descriptors for object class matching. *Neurocomputing*, 184:3–12.
- [Hu, 1962] Hu, M.-K. (1962). Visual pattern recognition by moments invariants. *IRE Transactions on Information Theory*, 8:179–187.
- [Huang et al., 1997] Huang, J., Kumar, S. R., and Mitra, M. (1997). Image indexing using color correlograms. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 762–768, San Juan, Puerto Rico, USA. IEEE.

- [Hutchings, 2016] Hutchings, W. K. (2016). *When Is a Point a Projectile? Morphology, Impact Fractures, Scientific Rigor, and the Limits of Inference*, pages 3–12. Springer, Dordrecht.
- [INAH, 2018] INAH (2018). Instituto nacional de antropología e historia. <http://www.inah.gob.mx/es/>.
- [J & DEE ARTIFACTS, 2018] J & DEE ARTIFACTS (2018). Ancient indian arrowheads–native indian artifacts–relics–tools. <http://www.j-and-dee-artifacts.com>.
- [J. Capowski and Freedman, 1995] J. Capowski, J. K. and Freedman, S. F. (1995). A numeric index based on spatial frequency for the tortuosity of retinal vessels and its application to plus disease in retinopathy of prematurity. *Retina*, 15:490–500.
- [Kauppinen et al., 1995] Kauppinen, H., Seppanen, T., and Pietikainen, M. (1995). An experimental comparison of autoregressive and fourier-based descriptors in 2d shape classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:201–207.
- [Kens’s Relics, 2018] Kens’s Relics (2018). Ken’s relics–native american artifacts. <http://www.kensrelics.com>.
- [Kovesi, 2015] Kovesi, P. (2015). Matlab and octave functions for computer vision and image processing. <http://www.peterkovesi.com/matlabfns/>.
- [Krapac et al., 2011] Krapac, J., Verbeek, J., and Jurie, F. (2011). Modeling spatial layout with fisher vectors for image categorization. In *International Conference on Computer Vision*, pages 1487–1494. IEEE.
- [Lathi, 1998] Lathi, B. P. (1998). *Modern digital and analog communication systems*. Oxford University Press, New York, NY, USA, tercera edition.
- [Lazebnik et al., 2006] Lazebnik, S., Schmid, C., and Ponce, J. (2006). Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In IEEE, editor, *Computer Vision and Pattern Recognition*, pages 2169–2178. Computer Society Conference.
- [Lindeberg, 1998] Lindeberg, T. (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30:79–116.
- [Lowe, 1999] Lowe, D. G. (1999). Object recognition from local scale-invariant features. In IEEE, editor, *International Conference on Computer Vision*, pages 1150–1157.
- [Lowe, 2004] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110.
- [Lu and Sajjanhar, 1999] Lu, G. and Sajjanhar, A. (1999). Region-based shape representation and similarity measure suitable for content-based image retrieval. *Multimedia System*, 7:165–174.
- [Manjunath et al., 2002] Manjunath, B., Salembier, P., and Sikora, T. (2002). *Introduction to MPEG-7: Multimedia Content Description Interface*. Wiley, New York, USA.

- [Manjunath, 2001] Manjunath, B. S. (2001). Color and texture descriptors. *IEEE Transactions on circuits and systems for video technology*, 11:703–715.
- [Mehetre et al., 1997] Mehetre, B. M., Kankanhalli, M. S., and Lee, W. F. (1997). Shape measures for content based image retrieval. *A Comparison. Information Processing & Management*, 33:319–337.
- [Mikolajczyk and Schmid, 2001] Mikolajczyk, K. and Schmid, C. (2001). Indexing based on scale invariant interest points. In *Eighth IEEE International Conference on Computer Vision. ICCV*, pages 525–531, Vancouver, BC, Canada.
- [Mikolajczyk and Schmid, 2004] Mikolajczyk, K. and Schmid, C. (2004). Scale & affine invariant interest point detectors. *International Journal of Computer Vision*, 60:63–86.
- [Mirambell and Lorenzo, 1974] Mirambell, L. and Lorenzo, J. L. (1974). *Materiales líticos arqueológicos: generalidades: Consideraciones sobre la industria lítica*. INAH, Mexico.
- [Mokhtarian and Mackworth, 1986] Mokhtarian, F. and Mackworth, A. (1986). Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:34–43.
- [Morel and Yu, 2009] Morel, J.-M. and Yu, G. (2009). Asift: An algorithm for fully affine invariant comparison. *SIANM Journal on Imaging Sciences*, 2:438–469.
- [Mortensen et al., 2005] Mortensen, E., Deng, H., and Shapiro, L. G. (2005). A sift descriptor with global context. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 174–190, San Diego, CA, USA. IEEE.
- [MVLA, 2018] MVLA (2018). Mount vernon ladies association. <https://www.mountvernon.org/preservation/mount-vernon-ladies-association>.
- [Nash and Prewitt, 2016] Nash, B. S. and Prewitt, E. R. (2016). The use of artificial neural networks in projectile point typology. *Lithic Technology*, 41:194–211.
- [Okumura and Araujo, 2014] Okumura, M. and Araujo, A. G. (2014). Long-term cultural stability in huntergatherers: a case study using traditional and geometric morphometric analysis of lithic stemmed bifacial points from southern brazil. *Journal of Archaeological Science*, 45:59–71.
- [OverstreetID, 2018] OverstreetID (2018). Official overstreet indian arrowhead identification online database. <http://typology.arrowheads.com/index.php>.
- [Pavlidis, 1980] Pavlidis, T. (1980). Algorithms for shape analysis of contours and wave forms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:301–312.
- [Perronnin et al., 2010] Perronnin, F., Sánchez, J., and Mensink, T. (2010). Improving the fisher kernel for large-scale image classification. In *Lecture Notes in Computer Science*, pages 143–156. European Conference, Springer.

- [Persoon and sun Fu, 1997] Persoon, E. and sun Fu, K. (1997). Shape discrimination using fourier descriptors. *IEEE Transactions On Systems, Man and Cybernetics*, 7:170–179.
- [Rafferty, 2018] Rafferty, J. (2018). Reverse engineering stone atlatl dart points. *Lithic Technology*, 43:151–65.
- [RelicShack, 2018] RelicShack (2018). Arrowheads and indian artifacts. <http://www.relicshack.com/Arrowheads>.
- [Rosten and Drummond, 2006] Rosten, E. and Drummond, T. (2006). Machine learning for high-speed corner detection. In *European Conference on Computer Vision*, pages 430–443.
- [Rots, 2016] Rots, V. (2016). *Projectiles and hafting technology*. Springer, Dordrecht.
- [Sánchez and Carpenter, 2014] Sánchez, G. and Carpenter, J. P. (2014). *Paisajes, Territorios e Industrias Líticas: adaptaciones humanas de larga duración en el desierto de Sonora*, pages 75–114. INAH, México.
- [Serwatka and Riede, 2016] Serwatka, K. and Riede, F. (2016). 2d geometric morphometric analysis casts doubt on the validity of large tanged points as cultural markers in the european final palaeolithic. *Journal of Archaeological Science*, 9:150–159.
- [Sivic and Zisserman, 2003] Sivic, J. and Zisserman, A. (2003). Video google: A text retrieval approach to object matching in videos. *International Conference on Computer Vision. IEEE*, 2:1470–1477.
- [Smith and Goebel, 2018] Smith, H. L. and Goebel, T. (2018). Origins and spread of fluted-point technology in the canadian ice-free corridor and eastern beringia. *Proceedings of the National Academy of Sciences of the United States of America*, 115:4116–4121.
- [Steger, 1998] Steger, C. (1998). An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:113–125.
- [Swain and Ballard, 1991] Swain, M. J. and Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, 7:11–32.
- [Teague, 1980] Teague, M. R. (1980). Image analysis via the general theory of moments. *Journal of Optical Society of America*, 70:920–930.
- [UNAM, 2018] UNAM (2018). Projectile points classifier (ppc). <http://matahari.fi-p.unam.mx/~inge-flechas>.
- [Valdovinos and Castellanos, 2014] Valdovinos, V. and Castellanos, R. (2014). *Un Tallador de Puntas en el Noreste de México. Arqueoetnografía en el Norte de Tamaulipas*, pages 147–188. INAH, México.
- [van Gemert et al., 2008] van Gemert, J. C., Geusebroek, J., Veenman, C. J., and Smeulders, A. W. M. (2008). Kernel codebooks for scene categorization. In *Lecture Notes in Computer Science*, pages 696–709. European Conference, Springer.

- [William Andrefsky, 2005] William Andrefsky, J. (2005). *Lithics. Macroscopic Approaches to Analysis*. Brendan2016, New York, 2nd edition.
- [Wu et al., 2013] Wu, J., Cui, Z., Sheng, V. S., Zhao, P., Su, D., and Gong, S. (2013). A comparative study of sift and its variants. *Measurement science review*, 13:122–131.
- [Wu et al., 2017] Wu, S., Oerlemans, A., Bakker, E. M., and Lew, M. S. (2017). A comprehensive evaluation of local detectors and descriptors. *Signal Processing: Image Communication*, 59:150–167.
- [Yamada and Kasutani, 2001] Yamada, A. and Kasutani, E. (2001). The mpeg-7 color layout descriptor: A compact image feature description for high-speed image/video segment retrieval. *ICIP*, 1:7–10.
- [Yap et al., 2010] Yap, P.-T., Jiang, X., and Kot, A. C. (2010). Two-dimensional polar harmonic transforms for invariant image representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32:1259–1270.
- [Zahn and Roskies, 1972] Zahn, C. T. and Roskies, R. Z. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions On Computer*, 21:269–281.
- [Zhou et al., 2010] Zhou, X., Yu, K., Zhang, T., and Huang, T. S. (2010). Image classification using super-vector coding of local image descriptors. In *Lecture Notes in Computer Science*, pages 141–154. European Conference, Springer.