

TESIS DEFENDIDA POR
Benjamín Hernández Valencia
Y APROBADA POR EL SIGUIENTE COMITÉ

Dr. Gustavo Olague Caballero
Co-Director del Comité

Dr. Wolfgang Steffen
Co-Director del Comité

Dr. Hugo Homero Hidalgo Silva
Miembro del Comité

Dr. Heriberto Márquez Becerra
Miembro del Comité

Dr. Juan Humberto Sossa Azuela
Miembro del Comité

Dra. Ana Isabel Martínez García
Coordinador del programa de posgrado
en Ciencias de la Computación

Dr. David Hilario Covarrubias Rosales
Director de Estudios de Posgrado

12 de mayo del 2009.

**CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN SUPERIOR DE
ENSENADA**



**PROGRAMA DE POSGRADO EN CIENCIAS
EN CIENCIAS DE LA COMPUTACIÓN**

**RECONOCIMIENTO DE PATRONES EN IMÁGENES NO VISIBLES:
EXPRESIONES FACIALES Y NEBULOSAS PLANETARIAS**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de
DOCTOR EN CIENCIAS

Presenta:

BENJAMÍN HERNÁNDEZ VALENCIA

Ensenada, Baja California, México, Mayo del 2009.

RESUMEN de la tesis de **Benjamín Hernández Valencia**, presentada como requisito parcial, para la obtención del grado de DOCTOR EN CIENCIAS EN CIENCIAS DE LA COMPUTACIÓN. Ensenada, Baja California, México. Mayo 2009.

RECONOCIMIENTO DE PATRONES EN IMÁGENES NO VISIBLES:
EXPRESIONES FACIALES Y NEBULOSAS PLANETARIAS

Resumen aprobado por:

Dr. Gustavo Olague Caballero
Co-Director de Tesis.

Dr. Wolfgang Steffen
Co-Director de Tesis.

En el universo existen fenómenos que para su estudio requieren del reconocimiento y clasificación de imágenes más allá de espectro visible, o bien de imágenes en donde la resolución de las líneas espectrales sea inferior a 1 nm. En ambos casos no son detectadas por el ojo humano. En este trabajo presentamos dos algoritmos de clasificación diseñados para este tipo de imágenes. En el primero abordamos el problema de *Reconocimiento de Expresiones Faciales sobre imágenes infrarrojas*. Nosotros proponemos un nuevo algoritmo autónomo de aprendizaje visual, basado en Computación Evolutiva, que integra en un sólo proceso la selección de la región de interés y la extracción de los rasgos más promisorios. Primero, el algoritmo escoge regiones faciales relevantes en donde la extracción de rasgos es realizada. Segundo, el algoritmo calcula la Matriz de Co-Ocurrencia en tonos de gris usados para describir a dicha región, y selecciona el subconjunto más significativo de ellos en la etapa de clasificación. Finalmente, la clasificación se realiza utilizando un comité de votación simple, basado en la técnica Máquina de Vector de Soporte (SVM). Resultados experimentales sobre tres expresiones faciales, muestran que la eficiencia de nuestro algoritmo es superior en un 16 % a la clasificación realizada por el observador humano. En el segundo trabajo, abordamos el problema de clasificación *Morfo-Cinématica* de objetos celestes llamados *Nebulosas Planetarias* (NP), a partir de su proyección o forma aparente sobre el plano del cielo y una descripción suficiente de su campo bidimensional de velocidades. El algoritmo consiste en extraer los vectores que representen las variaciones principales de los rasgos de los objetos en un espacio ortogonal sin correlación (eigencaracterísticas). Seguido, se intenta encontrar el mapa correcto que relacione las eigencaracterísticas de estos objetos con su correspondiente clase, por medio de SVM. Finalmente, se genera un comité de votación que toma las contribuciones de la clase con la máxima y la segunda máxima probabilidad de pertenencia de la región. Para un universo de 539 NPs sintéticas, 11 clases y un conjunto de 107 imágenes de prueba, el algoritmo responde correctamente al 98 % de la imágenes. Asimismo, se ha probado sobre imágenes reales y los resultados son prometedores.

Palabras Clave: reconocimiento de expresiones faciales, nebulosas planetarias, máquina de vector de soporte, reconocimiento de patrones.

ABSTRACT of the Thesis presented by **Benjamín Hernández Valencia**, as a partial requirement for obtaining the DOCTOR OF SCIENCE degree in COMPUTER SCIENCES. Ensenada, Baja California, México. May 2009.

PATTERN RECOGNITION FOR NON-VISIBLE IMAGERY: FACE EXPRESSION RECOGNITION
IN THERMAL IMAGERY AND MORPHO-KINEMATIC CLASSIFICATION OF PLANETARY
NEBULAE

In nature, the study of some problems require a special process of recognition and classification of images beyond of visible spectrum, or images where the spectral line resolution must be better than 1 nm. In both cases, they are not detected by the human eyes. In this work we presented two fine tuned special classification algorithms for this type of imagery. In the first work, we study the *Face Expression Recognition* problem in *infrared imagery*. We propose a new autonomous visual learning algorithm, based on Evolutionary Computations, that join in a simple step the selection of the regions of interest and the features extraction process. First, our algorithm chooses relevant facial regions where the feature extraction is performed. Then, the algorithm computes the Gray-Level Co-Occurrence matrix used to describe this region and selects the most significant subset descriptors in the classification stage. Finally, the classification is done using a simple support vector machine (SVM) committee. Experimental results of three different facial expression, show that the efficiency of our algorithm is higher by 16 % compared to the one made by the human observers. In the second work, we consider the problem of *Morpho-Kinematic Classification* of celestial objects named *Planetary Nebulae* (NP), starting from its projection or apparent form in the sky plane and a description of its position-velocity fields. The algorithm consists in extracting the vectors that represent the main variation of the object's features in an uncorrelated orthogonal space (eigenfeatures). Next, we try to find the correct map that relates these eigenfeatures with their corresponding class, by using SVM. Finally, a voting committee is created, that takes the contributions of the class with the maximum and the second maximum probability of class deduced from each regions. For an universe of 539 synthetic PNs, 11 classes and a set of 107 test images, the algorithm respond correctly to 98 % of them. It has also been tested on real images and the results are promising.

Key words: face expression recognition, planetary nebulae, pattern classification, SVM

Dedicatoria

A mi esposa, Alma

y a mi hijo Abraham,

porque son la fuerza de conducción de mi vida
y me han regalado,
su amor y su tiempo.

A mi madre, Silvia.

En memoria de mi padre, Baltasar.

A mis hermanos, Ismael, Salomón, Isaias, Gamaliel,
Baltazar, Israel, Elizabeth, Nahum y Silvia.

A mi segunda madre, Adelaida Rosa Maria.

Agradecimientos

A Dios.

Al Dr. Gustavo Olague Caballero por el gran equipo formado, la ayuda brindada a lo largo de la realización de este trabajo, su tiempo y a sus comentarios que van más allá de lo académico.

Al Dr. Wolfgang Steffen.

A los miembros del Comité de Tesis por sus valiosos comentarios.

Al Centro de Investigación Científica y de Educación Superior de Ensenada.

Al Instituto de Astronomía Campus Ensenada, Universidad Nacional Autónoma de México.

Índice general

Resumen	I
Resumen en Inglés	II
Dedicatoria	III
Agradecimientos	IV
Índice general	V
Índice de figuras	VII
Índice de tablas	XI
I. Introducción	1
I.1. Relevancia científica	6
II. Aprendizaje Visual basado en textura FER	8
II.1. Introducción	8
II.2. Trabajos Previos	13
II.3. Características de nuestro algoritmo	17
II.4. Conceptos Básicos	18
II.4.1. Computación Evolutiva	19
II.4.2. Textura a partir de la Matriz de Co-ocurrencia	22
II.4.2.1. Matriz de Co-ocurrencia en Tonos de Gris	26
II.5. GA de aprendizaje visual	31

II.5.1. Selección de las ROIs	32
II.5.2. Extracción de características	34
II.5.3. Clasificación	35
II.5.4. Función de aptitud	35
II.5.5. Parámetros del algoritmo genético	36
II.5.6. Parámetros de entrenamiento	36
II.6. Experimentación y resultados	37
II.6.1. Base de datos OCTBS	37
II.6.2. Resultados	39
II.7. Conclusiones	43
III. Reconocimiento Morfo-Cinemático de NP	45
III.1. Introducción	45
III.2. Trabajos Previos	48
III.3. Nebulosas Planetarias	52
III.3.1. Espectrómetro	56
III.3.2. Espectro de Posición-Velocidad	60
III.3.3. Efectos de Proyección de una NP	63
III.4. Modelado Morfo-Cinemático	66
III.5. Esquema de Reconocimiento	71
III.5.1. Extracción de Eigencaracterísticas	76
III.5.2. Máquina de Vector de Soporte	80
III.5.3. Comité de Clasificación SVM	82
III.6. Experimentación y Discusión	83
III.6.1. Catálogo general de NPs	84
III.6.2. Objetos reales	87
III.6.2.1. Histograma en imágenes astronómicas	88
III.6.2.2. Generación de la imagen de prueba	89
III.6.2.3. Prueba	91
III.7. Conclusiones	96

<i>ÍNDICE GENERAL (continuación)</i>	VII
IV. Conclusiones	98
IV.1. Trabajos futuros	99
Bibliografía del Autor	103
Referencias	105

Índice de figuras

1.	Caracterización del <i>espectro electromagnético</i> . El ojo humano sólo responde al espectro electromagnético entre los 3800 y 7400 Å ($1\text{Å} = 1 \times 10^{-10}m$) que es una fracción muy reducida del espectro.	2
2.	Diagrama de flujo común de un algoritmo FER.	11
3.	Imágenes térmicas provenientes de la base de datos OTCBVS.	17
4.	Diagrama de flujo de nuestro algoritmo de aprendizaje visual.	19
5.	Estructura general de un algoritmo genético.	21
6.	Celda de resolución de los vecinos más cercanos de una posición dada C . Las celdas $\{1, 5\}$, $\{2, 6\}$, $\{3, 7\}$ y $\{4, 8\}$ están a orientadas a 0° , 45° , 90° y 135° respectivamente.	27
7.	Construcción de la matriz de co-ocurrencia.	28
8.	Descripción detallada del cromosoma de aprendizaje visual.	33
9.	Todas estas expresiones faciales parecieran estar en el estado de “felicidad”, pero en la base de datos algunas no se etiquetan como tal.	38
10.	Calcula de la región más prominente de la expresión facial y su escalamiento a una marco de 32×32 pixeles.	39
11.	Ubicación de las 4 regiones de interés del mejor cromosoma de nuestro experimento.	40
12.	Diferentes morfologías de NPs. La imagen de Helix fue tomada por el Anglo-Australian Telescope, las restantes fueron adquiridas por el Hubble Space Telescope (HST).	53
13.	Diagrama esquemático Hertzsprung-Russell de la evolución de una estrella cuya masa inicial está entre 0.5 y $8 M_\odot$	54

14.	Diagrama de una estrella AGB, progenitora de un NP. El núcleo está compuesto por carbón y oxígeno, C-O. Reacciones de fusión nuclear ocurren en dos capas separadas. En la capa más interior se quema helio para producir carbón y en la siguiente se quema hidrógeno para producir helio. La capa exterior de hidrógeno es la que ocupa el mayor volumen de la estrella.	55
15.	Esquema óptico del espectrómetro MEZCAL.	57
16.	Ubicación de los elementos básicos del espectrómetro “MES-SPM”. La posición de rendija es ortogonal al brazo del instrumento. En el gráfico a) el brazo está orientado norte ↔ sur, entonces la rendija este ↔ oeste. El gráfico b) muestra la orientación inversa.	58
17.	Imágenes adquiridas con el espectrómetro MES-SPM. Las líneas roja y amarilla muestran 2 posiciones distintas de la rendija. Los gráficos b) y c) son imágenes de la posición de la rendija girados con respecto al gráfico a). Los gráficos b') y c') son los espectros bidimensionales del haz de luz que pasa por la rendija. Cortesía de H. Tirado, A. López; Octubre 2004; Observatorio Astronómico Nacional, San Pedro Mártir B.C. México.	59
18.	Efecto Doppler.	60
19.	Espectro P-V.	63
20.	Importancia de los efectos de proyección de la morfología aparente de una nebulosa planetaria. Vea texto.	64
21.	Modelo tridimensional de una NP bipolar con cintura. (a) y (b) distribución de partículas vistas desde el polo y a 45° sobre dos ejes. Las proyecciones en el plano del cielo se observan distintos en su forma debido a la rotación del modelo; (a') puede ser clasificada como una NP esférica, mientras que (b') se observa bipolar. Los espectros (a'') y (b'') corresponde a la posición de la rendija dada en (a') y (b') respectivamente.	65
22.	Simulación del proceso de adquisición de una nebulosa planetaria utilizando SHAPE.	67
23.	Puntos de visión del modelo 3D de una nebulosa planetaria bipolar cerrada con cintura del AMCNP.	68
24.	La base de datos esta compuesta por las proyecciones del modelo 3D de la nebulosa y el cálculo del espectro P-V en 7 posiciones distintas de la rendija.	69
25.	Muestra del AMCNP.	70
26.	Esquema general de reconocimiento y clasificación de las NP.	73
27.	Regiones de interés $\{\Omega_d, \Omega_w, \Omega_n, \Omega_s, \Omega_c\}$. Las regiones sobre el espectro P-V se muestran de diferente tamaño en el eje v para facilitar su ubicación.	74

28.	Espectros similares (a) y (b) pueden ocurrir en NP con morfología distinta dependiendo de la posición de la rendija.	75
29.	Los 10 gráficos superiores muestran las eigenimágenes con mayor varianza para la región Ω_d . Los 10 restantes pertenecen a Ω_w	79
30.	Nebulosas mal clasificadas.	86
31.	Ubicación de la rendija para el conjunto de entrenamiento y prueba para esta proyección.	87
32.	Ubicación de estructuras en el histograma de una nebulosa planetaria. Los gráficos inferiores muestran a la imagen que corresponde a cada estructura así como su composición final. Las estructuras son: (1) fondo del cielo, (2) nebulosa, (3) estrellas. Objeto He1-5, Hortensia Riesgo, <i>et al.</i> SPM 2001.	89
33.	Generación de una imagen de prueba a partir de observaciones reales. La imagen directa I_D sufre una transformación de similitud dada por la posición de la rendija I_P . Una sección del espectro I_S es seleccionado por el usuario y es agregado a la transformación anterior para obtener una imagen en el formato de nuestro catálogo (I_C).	92
34.	Imágenes de prueba obtenidas en el OAN-SPM.	93

Índice de tablas

II.1. Matriz de confusión con los valores de relación-cruzada.	41
II.2. Matriz de confusión de los sujetos de prueba.	41
II.3. Clasificación Humana de imágenes térmicas.	42
II.4. Clasificación Humana de imágenes en el espectro visible.	42
III.1. Matriz de decisión de una nebulosa bipolar, orientada $(20^\circ, 0, 0)$, con un ancho de rendija de $\Delta_s = 19$ centrada en el pixel $x = 82$. La etiqueta real es 1.	83
III.2. Composición de NPDB $\Delta_s = 17$	85
III.3. Matriz de confusión	86
III.4. Matriz de decisión para las 13 imágenes de prueba.	95

Capítulo I

Introducción

El ser humano percibe al mundo que lo rodea esencialmente a través del sentido de la vista. Más del 99 % de la información que percibe es mediante su órgano óptico; el ojo. Sin embargo, éste está limitado en una banda angosta del espectro electromagnético que va aproximadamente desde los 380 a los 740 nm en longitud de onda, ver Figura 1. De la misma forma, no tiene la capacidad de discernir pequeñas diferencias entre longitudes de onda cercanas; en particular, puede percibir diferencias en color del orden de 1 nm en el espectro visible. Asimismo, el ojo humano está limitado en tamaño y sensibilidad, por lo que puede detectar únicamente la intensidad lumínica a partir de cierto umbral, además de que no puede acumular la luz que incide en él por periodos largos, lo que imposibilita intensificar la imagen que está captando. Finalmente, el ojo humano no puede almacenar las imágenes para futuras referencias de forma similar a un placa fotográfica. Dadas estas limitaciones naturales y el deseo intrínseco del ser humano por conocer qué hay después de estas fronteras, se han desarrollado instrumentos, sensores y algoritmos computacionales que le permitan investigar lo que existe *más allá de su capacidad visible*.

En la literatura especializada en el campo de visión por computadora, existe una gran diversidad de algoritmos enfocados a procesar imágenes en el ancho de banda en el cual está sintonizado el ojo humano. Sin embargo, en este trabajo desarrollaremos dos algoritmos

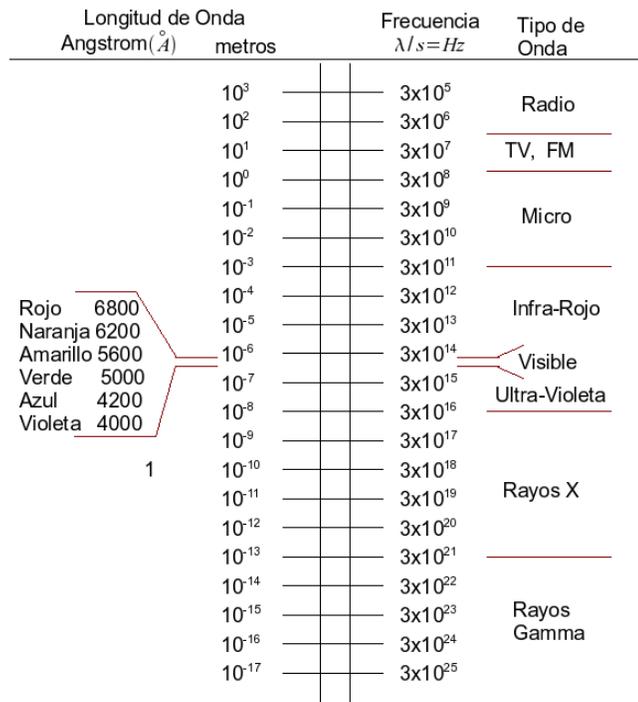


Figura 1: Caracterización del espectro *electromagnético*. El ojo humano sólo responde al espectro electromagnético entre los 3800 y 7400 Å ($1\text{Å} = 1 \times 10^{-10}m$) que es una fracción muy reducida del espectro.

que nos permitirán trabajar con: primero, imágenes más allá del espectro visible y segundo imágenes en donde es necesario que la resolución espectral del fenómeno observado sea inferior a la que puede captar el detector óptico natural del ser humano. Esto es, fenómenos que respondan a longitudes de onda inferiores a 380 nm o imágenes que respondan a longitudes de onda superiores a 740 nm, o bien, a fenómenos que para poder ser observados requieran resolver diferencias inferiores a 1 nm en su espectro electromagnético, lo cual permite detectar cambios en la velocidad relativa entre el observador y el objeto. En ambos casos las imágenes son invisibles al ojo humano.

Reconocimiento de patrones más allá del espectro visible está envuelto en una gama amplia de disciplinas, incluyendo el desarrollo de sensores de rayos gamma y rayos X, ultravioleta, infrarrojos, infrarrojo cercano, ondas milimétricas, ondas micrométricas, apertura sintética de radar y sensores electro ópticos, así como algunos tópicos de procesamiento de señales, visión por computadora e inteligencia de máquina. Ésta es una área fértil para el

análisis y experimentación de aplicaciones educativas, industriales y de investigación.

El empleo de sensores en la banda no visible se ha venido incrementado en esta década. El costo de sensores en el infrarrojo cercano y medio ha disminuido dramáticamente, así como el de los sensores que operan en la banda milimétrica. Es por esto, que el empleo de esta tecnología ya no está confinada para las aplicaciones militares. Áreas como el transporte y sistemas de seguridad basadas en la visión, han comenzado a comercializar con sensores infrarrojos. En consecuencia, esta nueva modalidad conlleva a la necesidad de la *adaptación* de métodos y algoritmos que originalmente fueron desarrollados para sensores en la banda visible y al desarrollo de nuevos métodos y sistemas que exploten las nuevas características de estos sistemas. La tendencia actual de los sistemas de visión emplean más de un simple sensor que capture, siga y reconozca un objeto de interés. Hoy en día estos sistemas, utilizan algoritmos adicionales no redundantes y complementarios que identifiquen de mejor manera los rasgos del objeto.

En este trabajo se presentan dos algoritmos en el área de reconocimiento de patrones en imágenes no visibles. En el primero, titulado *Aprendizaje Visual Basado en Descriptores de Textura para el Reconocimiento de Expresiones Faciales en Imágenes Térmicas*, se utiliza una base pública de imágenes térmicas o infrarrojas (Avidi, OTCBVS) adquiridas con un sensor que responde a espectros de $70,000 - 140,000\text{Å}$, ver Figura 1. En el segundo, titulado *Reconocimiento Morfo-Cinemático de Nebulosas Planetarias*, el algoritmo es sometido a prueba con imágenes astronómicas adquiridas en el Observatorio Astronómico Nacional ubicado en la sierra de San Pedro Mártir B.C, México, utilizando el espectrómetro MES-SPM (Meaburn et al., 2003). Este espectrómetro ha sido probado con sensores que operan en el rango de $3,900 - 9,000\text{Å}$ con una resolución espectral entre $\sim 0,4 - 0,9\text{Å}$.

La motivación general de este trabajo, es adaptar y aplicar métodos de reconocimiento de patrones diseñados para el espectro visible, a imágenes que no presentan la riqueza de textura y color, que se encuentran generalmente en las imágenes adquiridas en la banda del visible. Asimismo, inferir en una primera aproximación, cuál es el grado de dificultad que esta “falta de información” puede ocasionar, en los algoritmos tradicionales de reconocimiento

de objetos y los concernientes a visión por computadora. Por otro lado, las imágenes no contenidas en el espectro visible presentan otras características interesantes; la principal es que en las imágenes térmicas, los efectos de iluminación de la escena son desechados, esto implica que las imágenes son invariantes a la iluminación de la escena. Es evidente que esta tarea es compleja debido a la gran diversidad de algoritmos de reconocimiento de patrones existentes en la literatura. Para acotar esta problemática decidimos estudiar dos problemas distintos. En el primer trabajo desarrollamos un algoritmo novedoso para detectar la clase de emoción que un individuo plasma en una imagen térmica, y se presenta una disertación de la dificultad de tratar de plasmar e interpretar el estado emocional del individuo. A esta temática se le conoce en la literatura como *Reconocimiento de Expresiones Faciales*. En el segundo trabajo, aplicamos estrategias de inteligencias de máquina para objetos astronómicos llamados *Nebulosas Planetarias*, que nos permitiera alcanzar una clasificación morfológica fiable. En él, analizamos los efectos de proyección de estas estructuras, que en forma natural tienen tres dimensiones, y utilizamos distintos modelos sintéticos para lograr adquirir una serie de imágenes y espectros cinemáticos que nos ayudaran a romper dicha ambigüedad. A partir de este catálogo, pudimos clasificar un conjunto de imágenes reales. En ambos casos las imágenes contenían escasa información de textura y los sensores e instrumentos utilizados nos permitían “ver” información que para el ojo humano no era visible.

La aplicación principal del primer trabajo está enfocada a sistemas expertos que requieran la toma de alguna decisión basado en el estado emocional del individuo, bajo condiciones de ausencia de iluminación. Recientemente el anuncio hecho por la industria automotriz interesada en desarrollar vehículos inteligentes, equipados con sistemas de advertencia para efectuar alguna acción, que detecte la fatiga del conductor, la distracción del mismo, advertencia de abandono del carril, advertencia de colisión delantera, detección del punto oculto y visión nocturna, es un campo de aplicación prometedor. El segundo trabajo tiene su campo de aplicación en el entorno de minería de datos Astronómicos. Con la existencia de telescopios que muestrean en forma sistemática el cielo nocturno, llamados “surveys”, y el lanzamiento y construcción de nuevos telescopios, v.g. el telescopio espacial GAIA y el desarrollo del

“Thirty Meter Telescope”, los procesos de reconocimiento de patrones han empezado a tener relevancia significativa en la comunidad astronómica. Es aquí, en donde este trabajo puede aportar resultados útiles para esta comunidad.

Este trabajo de tesis está organizado de la forma siguiente:

- En la siguiente sección presentamos un resumen de la relevancia científica de los dos trabajos. Es importante enfatizar que dichos trabajos están autocontenidos.
- En el capítulo II, presentamos el trabajo *Aprendizaje Visual Basado en Descriptores de Textura para el Reconocimiento de Expresiones Faciales en Imágenes Térmicas*. Este trabajo presenta un sistema invariante a los efectos de iluminación de la escena, para el reconocimiento de expresiones faciales en imágenes infrarrojas. En él se propone un nuevo algoritmo que integra la selección de la región de interés y el proceso de extracción de rasgos o características, utilizando los paradigmas de cómputo evolutivo. El aprendizaje visual escoge los mejores rasgos a través del estadístico de segundo orden llamado “matriz de co-ocurrencia en tonos de gris”. Utilizando un comité de votación producto de clasificar por medio de una “maquina de vector de soporte” y logra superar en un 16 % a la clasificación realizada por el ser humano.
- En el capítulo III, presentamos el trabajo *Reconocimiento Morfo-Cinemático de Nebulosas Planetarias*. Este trabajo presenta un algoritmo de clasificación que toma en cuenta la proyección sobre el plano del cielo, y su campo de velocidades en dos dimensiones, a partir de modelos tridimensionales de nebulosas planetarias. Se experimenta con 11 clases distintas de estos objetos. El algoritmo consiste en extraer los vectores que representen las variaciones principales de los rasgos de los objetos en un espacio ortogonal decorrelacionado. Para ello se emplea el análisis de componentes principales (eigencaracterísticas). Seguido, se intenta encontrar el mapa correcto que relacione las eigencaracterísticas de estos objetos con su correspondiente clase, por medio de la búsqueda del hiperplano que represente el máximo margen de separación entre clases (máquina de vector de soporte). Finalmente, se genera un comité de votación que to-

ma las contribuciones de la clase con la máxima y la segunda máxima probabilidad de pertenencia de la región. Para un conjunto de 107 imágenes de prueba, el algoritmo responde correctamente al 98 % de la imágenes.

- En el capítulo IV, se presentan las conclusiones generales de los trabajos, a la vez que se proponen los trabajos futuros.

I.1. Relevancia científica

En la sección anterior hemos introducido dos trabajos que emplean imágenes no visibles para el ojo humano. En este apartado, resumiremos los puntos en donde hacemos aportaciones relevantes.

El primer trabajo “*Aprendizaje Visual Basado en Descriptores de Textura para el Reconocimiento de Expresiones Faciales en Imágenes Térmicas*”, las contribuciones específicas son:

- Nosotros ampliamos el estudio del problema reconocimiento de expresiones faciales más allá del espectro visible utilizando imágenes en la banda térmica.
- Presentamos una nueva representación del problema de reconocimiento de expresiones faciales, empleando un lazo cerrado del aprendizaje visual y la computación evolutiva.
- Nuestro trabajo presenta una solución conjunta y sistemática para resolver la selección de la región de interés y la extracción de características, al mismo tiempo.
- Nuestro algoritmo utiliza estadísticas de segundo orden independientes del dominio del problema. Por tanto es posible su adaptación para diferentes problemas del reconocimiento de patrones.

Para el siguiente trabajo titulado “*Reconocimiento Morfo-Cinemático de Nebulosas Planetarias*”, las contribuciones específicas son:

- Nosotros clasificamos a las nebulosas planetarias a partir de la cinemática observada del objeto y su proyección sobre el plano del cielo. En la literatura existe un trabajo de clasificación automática de estos objetos, cuya clasificación está basada en las abundancias químicas de los elementos y no en su forma y campo de velocidades.
- Proponemos un esquema de votación para el clasificador, que toma en su análisis las contribuciones de la primera y segunda “opinión” probabilística para obtener una clasificación fiable.

Capítulo II

Aprendizaje Visual Basado en Descriptores de Textura para el Reconocimiento de Expresiones Faciales en Imágenes Térmicas

II.1. Introducción

Las emociones son fenómenos psicofisiológicos que representan modos de adaptación a ciertos cambios del entorno de la persona, resaltan ciertas conductas en la jerarquía de respuestas del individuo y activan redes asociativas relevantes en la memoria. En términos fisiológicos, las emociones provocan respuestas de distintos sistemas biológicos, como son *las expresiones faciales*, los músculos, la voz, actividad del sistema nervioso autónomo y el sistema endocrino, con el fin de establecer un medio para el comportamiento más efectivo. En particular, existen 8 emociones básicas del individuo, 4 primarias, *felicidad*, *enojo*, *miedo* y *tristeza* y 4 secundarias, *sorpresa*, amor, vergüenza y aversión. Determinar lo que una persona piensa y siente a través de la simple observación de una expresión facial no es una tarea fácil. El ser humano requiere de años de interacción constante con una gran cantidad de personas, para construir modelos mentales discriminantes que lo ayuden a conocer como expresamos emociones. Es así, que nosotros estamos en la posibilidad de comparar los modelos aprendidos de lo que percibimos de otras personas e inferir su estado emocional. Nuestros

modelos son entonces construidos con una gran diversidad de información del contexto relevante. Esto nos permite generar modelos mentales complejos. No obstante, la generalización de la representación que hemos construido es extremadamente difícil, y sufre ajustes finos conforme interactuamos con más personas. Por ejemplo, si interactuamos con personas que pertenecen a diferentes grupos de razas, aldeas o comunidades nuestro marco contextual se amplía y el modelo mental que pudiera ser ajustado a esta nueva información, de alguna forma tiende a estar más completo, ser más general. La dificultad natural de nuestro problema es comparable con la complejidad de la psiquis humana y su estado emocional. En términos de sistemas computacionales, en donde se necesiten determinar el estado emocional de una persona, resulta también ser una tarea extremadamente difícil y es hoy en día un problema no resuelto. Más aún, una solución comprensible de este problema debiera considerar diferentes canales de información que provean de una señal multimodal que dependa del contexto. Éstas pudieran ser la voz, señales de patrones psicológicos, el movimiento del cuerpo, manos y las expresiones faciales. En base a esta lista, los elementos visuales y en particular la instancia de la gesticulación facial, la expresión facial, provee de información relevante con el propósito de inferir el estado emocional de las personas. Esto implica que un sistema de inteligencia artificial (IA) obtenga muchos beneficios, si se apoya en técnicas de visión por computadora que extraigan información visual relevante de las imágenes de las expresiones faciales de las personas. Este problema es conocido en la literatura especializada, visión por computadora, como “*Problema de Reconocimiento de Expresiones Faciales*” (FER¹).

Las investigaciones recientes del problema FER han centrado su atención en el uso de un solo canal de información, específicamente en imágenes o secuencia de ellas tomadas con sensores en el espectro visible. No obstante, algunos otros autores (Wilder et al., 1996; Prokoshi, 2000), han hecho notar que los sistemas de análisis de expresiones faciales basadas en el espectro visible dependen de la iluminación y la pose del sujeto de prueba y de la escena. Trabajos recientes (Hernández et al., 2007; Trujillo et al., 2005; Pavlidis et al., 2001; Sugimoto et al., 2000; Yoshitomi et al., 2000) han mostrado la utilidad y factibilidad del empleo

¹por sus siglas en inglés *Face Expression Recognition*.

de sensores que capturan imágenes más allá del espectro visible. Estos autores sugieren que, dado la dificultad del problema FER, las nuevas investigaciones en esta área busquen nuevas e imaginativas maneras de aplicar los algoritmos FER a fuentes antes no empleadas, tal es el caso de señales infrarrojas o imágenes térmicas.

Un problema interesante para los diseñadores de sistemas de IA, es el tratar de replicar algún proceso humano, v.g. la detección de emociones. Dado que la meta de estos sistemas es tratar de resolver una instancia de un proceso más general, los diseñadores están sesgados a tomar ciertos caminos en la exploración de estos procesos. Ellos intentan simular el proceso como proceso teorizable. Pueden también ser influenciados por sus propias intuiciones derivadas del conocimiento personal y sus experiencias. Lo que puede llevar a los investigadores a considerar el uso de esta información, como información relevante. Más aún, dado que la salida de un sistema de IA es de importancia primordial, debiera ser posible utilizar cualquier tipo de información disponible para él, incluso si se sugiere una fuente que los seres humanos no utilizan de manera natural, v.g. señales térmicas. Es así, que un sistema de IA debiera de considerar señalización multimodal.

En el estado del arte de las técnicas para el problema FER, se identifican tres procesos básico (Fasel y Luetin, 2003) los cuales son, ver Figura 2:

1. Detección de regiones de interés.
2. Extracción de rasgos o características.
3. Clasificación de imágenes.

El primer proceso es la determinación de la región de interés (**ROI**). ROI es la ubicación de un marco espacial dentro de una imagen, de la cual deben de extraerse los rasgos o características. Las técnicas del estado del arte han utilizado métodos holísticos en donde la ROI cubre la cara completa (Lyons et al., 1999), y métodos modulares en donde la información de los rasgos es extraída de regiones específicas (Padgett y Cottrell, 1997; Dubuisson et al., 2001; Donato et al., 1999). El segundo proceso es la extracción de rasgos dada una ROI.

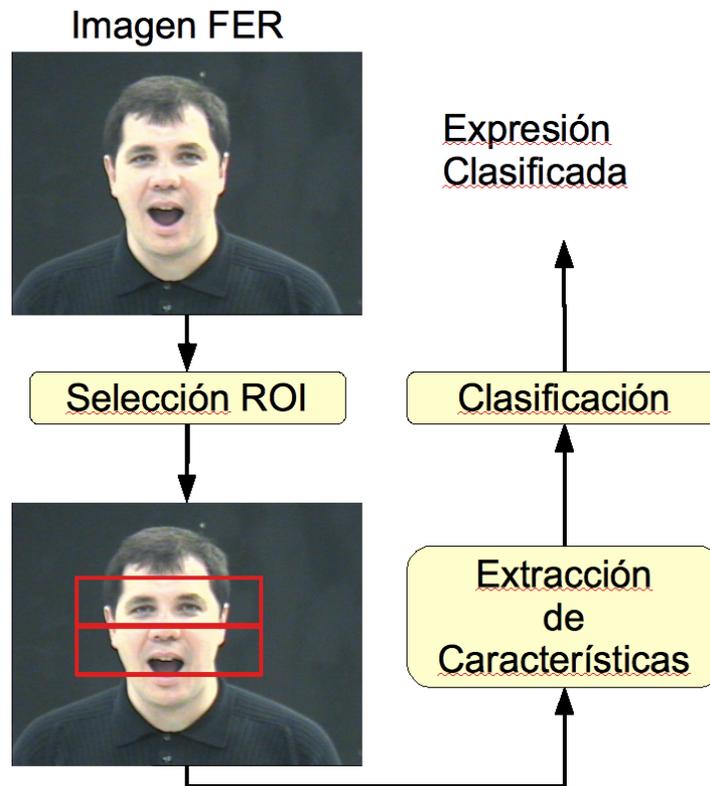


Figura 2: Diagrama de flujo común de un algoritmo FER.

Este proceso reduce significativamente la dimensión del problema mediante la aplicación de técnicas tales como: análisis de componentes principales (PCA) (Turk y Pentland, 1991; Donato et al., 1999), filtros Gabor (Donato et al., 1999) y métodos híbridos (Lyons et al., 1999; Padgett y Cottrell, 1997). Finalmente, se utiliza un clasificador entrenado que emplea al vector de características extraídas de cada ROI, para efecto de asignar la expresión facial a una clase ya aprendida. Entre los clasificadores más comunes, podemos mencionar a las Redes Neuronales, modelos de Markov y a las máquinas de vector de soporte. En cada una de estas etapas básicas, los investigadores se han centrado en desarrollar aproximaciones finas de forma aislada. Ignorando con ello, las posibles dependencias entre ellas que pudieran ser explotables o prometedoras para la investigación. Básicamente un diseñador humano, debe tener una respuesta a cada una de las preguntas siguientes.

- *¿Cuál pudiera ser el mejor conjunto de áreas faciales que debieran ser analizadas? Por*

ejemplo, algunas investigaciones utilizan aproximaciones holísticas o bien emplean la región completa de la cara para realizar la extracción de rasgos. Entonces podríamos preguntarnos si estas regiones son suficientes y no redundantes? Más aún, pudiera otra región ayudarnos a proveer información adicional sobre la expresión emocional?.

- *¿Cómo determinamos cuales características son relevantes para el proceso de clasificación de las emociones?*
- *¿Cómo se realiza la clasificación en una imagen después de haber extraído el vector de descriptores de rasgos?*

Para el caso de las imágenes en el espectro visible la respuesta a estas preguntas pueden parecer evidentes y se han propuesto la literatura soluciones eficientes en (Fasel y Luetttin, 2003). Por otro lado, la señales en el espectro infrarrojo podrían parecer artificiales según la percepción de los seres humanos. Esto aumenta la dificultad de intentar hacer la correlación entre este tipo de información y nuestro proceso mental. Es así, que la respuesta a estas preguntas no puede ser directa. Por consecuencia, la manera apropiada de cómo pudiera ser manejada e interpretada la información infrarroja puede no ser evidente. Nosotros creemos que este problema puede ser abordado en forma apropiada y coherente utilizando meta-heurísticas recientes tal como la *computación evolutiva*. EC es una técnica novedosa que nos ayuda a resolver problemas de optimización no lineal y búsquedas complejas (Holland, 1992; Goldberg, 1989). Nosotros proponemos un método que realiza el aprendizaje visual en dos procesos básicos del problema de reconocimiento de expresiones faciales: en el proceso de extraer o determinar las regiones faciales más relevantes y en el proceso de determinar la forma de realizar la extracción de características de la imagen. *Aprendizaje Visual*² es el proceso en el cual un sistema autónomo artificial adquiere conocimiento a través de un conjunto de imágenes de entrenamiento con el objeto de resolver alguna tarea visual dada. En este trabajo, el aprendizaje visual es implementado utilizando el paradigma de Cómputo Evolutivo como el motor de aprendizaje. Nosotros creemos que utilizando el aprendizaje vi-

²conocido en inglés como *Visual Learning*.

sual para desarrollar una solución al problema de FER, emularemos el espíritu del proceso de aprendizaje humano. Utilizando EC como el motor de aprendizaje, nos permite buscar en un espacio amplio de soluciones posibles con el objetivo de obtener o buscar una respuesta óptima a las preguntas antes citadas. Entonces, el sistema podrá probar diferentes caminos para resolver los dos primeros procesos básicos mostrados en la Figura 2. Es así, que nosotros creemos que el aprendizaje visual de nuestro algoritmo, nos ayudará en la adquisición de una comprensión más profunda de cómo esta información pudiera ser utilizada.

II.2. Trabajos Previos

El análisis de rostros de los seres humanos ha sido un tópico de interés de una gran cantidad de proyectos en visión por computadora. Básicamente por las implicaciones que tiene en el desarrollo de sistemas interactivos hombre-máquina; mucho trabajo se ha realizado en el desarrollo de detección de rostros (Sun et al., 2004; Viola y Jones, 2004) y técnicas de reconocimiento de rostros (Zhao et al., 2004). Una buena referencia de las técnicas de análisis faciales puede consultarse en Kong et al. (2005).

El problema de reconocimiento de expresiones faciales también ha recibido una importante atención por parte de los investigadores, como constata en los trabajos de recopilación de Pantic y Rothkrantz (2000b) y más recientemente en Fasel y Luetttin (2003). Los algoritmos FER están enfocados en extraer el contenido emocional a través de patrones visuales en la cara de una persona. Se han desarrollado técnicas para una sola imagen o para una secuencia de ellas. Los sistemas FER que emplean una secuencia de imágenes o vídeo ofrecen la oportunidad de estudiar a profundidad cómo la cara humana genera diferentes gesticulaciones. Trabajos importantes en esta área pueden consultarse en Bourel et al. (2002); Pantic y Rothkrantz (2000a); Li Tian et al. (2001). Sin embargo, no siempre se tiene disponible una secuencia de imágenes como datos de entrada para los sistemas FER. En consecuencia, sistemas FER que emplean imágenes sencillas son una alternativa de solución. Así, una gran variedad de diferentes técnicas basadas en una exposición han sido propuestas, (Lyons et al., 1999;

Dubuisson et al., 2001; Padgett y Cottrell, 1997; Donato et al., 1999; Pantic y Rothkrantz, 2000b).

A pesar del gran interés en las investigaciones del problema FER, pocos han sido los trabajos enfocados al estudio de este problema en imágenes más allá del espectro visible. Sin embargo, después del éxito en el desarrollo de algunas aplicaciones en otros problemas del análisis de rostros (Wilder et al., 1996; Prokoski, 2000; Kong et al., 2005), se ha incrementado el interés por el estudio de nuevos métodos para el problema FER. Sugimoto et al. (2000); Yoshitomi et al. (2000), emplean ROIs predeterminadas correspondientes a áreas alrededor de la nariz, boca, ojos y mejillas usando recocido simulado y mascararas de comparación para lograr una apropiada localización. Las características son calculadas por medio de la generación de imágenes diferenciales obtenidas a partir del promedio del rostro sin emoción y una imagen de prueba dada, utilizando la transformada discreta del coseno. Dichos autores realizan la clasificación utilizando propagación hacia atrás de una red neuronal entrenada. Pavlidis et al. (2001) utilizan una variante del problema clásico FER. Ellos, emplean imágenes térmicas de alta definición para estimar el cambio en las regiones de la cara producidas por el flujo sanguíneo. El sistema tiene por objeto detectar la ansiedad o la mentira, funciones similares a las de un polígrafo, de un individuo sin que este perciba que está siendo analizado. De la misma forma, la ROI es seleccionada por el diseñador del sistema. Trujillo et al. (2005) propone un procedimiento global y local automático para la localización de características, aplicadas al problema FER con imágenes térmicas. Ellos utilizan cúmulos de puntos de interés para localizar las regiones más promisorias donde se encuentran las características y aplican el análisis de componentes principales para reducir las dimensiones del vector de rangos. Estos autores emplean un comité SVM de votación ponderado para la clasificación de la expresión. A diferencia de este trabajo, las técnicas antes mencionadas hacen suposiciones apriori de las regiones más probables, en las cuales se encuentre la información más relevante. Nosotros proponemos modificar este método tradicional, dada la poca información y conocimiento limitado que se tiene del comportamiento de las imágenes térmicas en el problema FER. Además, debido al hecho de que existe un número reducido de trabajos relacionados con

este tema, combinado con los escasos reportes del rendimiento de estos algoritmos y el uso de diferentes bases de datos, provoca que una comparación directa entre estos métodos sea poco útil. Por ejemplo, el trabajo de Pavlidis et al. (2001) no es comparable debido al tipo de información que es extraída después de su análisis. Solo el trabajo que hemos presentado con antelación (Trujillo et al., 2005) emplea la misma base de datos. Es por todas estas razones que hemos decidido comparar la calidad de nuestro algoritmo, con los resultados de nuestro trabajo basado en SVM-PCA, así como los obtenidos de una prueba diseñada y aplicada con una muestra de seres humanos.

Nuestro método consiste en un algoritmo de aprendizaje visual que emplea el paradigma de EC como el motor de aprendizaje y búsqueda. Recientemente, las técnicas de EC aplicadas al problema de algoritmos de aprendizaje y reconocimiento de patrones han producido resultados competitivos y promisorios. Las técnicas EC, cuando son comparadas con los algoritmos no meta-heurísticos, muestran estructuras impredecibles difíciles de entender a priori por la mente humana de los diseñadores. En ese sentido, técnicas visuales utilizando esta meta-heurística podría llevarnos hacia un entendimiento más profundo en el dominio del problema. Estas técnicas son usadas comúnmente como un proceso de extracción de características. La extracción de características engloba la definición, la construcción o la síntesis de un operador que extraiga las características de la imagen en un vector de rasgos con el menor número de grados de libertad posibles, así como la extracción de un subconjunto de ellos. Si centramos la atención solamente en el proceso de selección de rangos, los algoritmos evolutivos inician con un conjunto de posibles características de interés, que definen el espacio de búsqueda del algoritmo de aprendizaje. El algoritmo intenta diferentes combinaciones de estas características y evalúa la calidad de la clasificación en cada paso. Después de muchas generaciones se obtiene un óptimo o subóptimo de las mejores características que definen el problema. Bala et al. (1996) emplea un algoritmo genético (GA) para realizar la selección de característica enfocado a reconocer conceptos visuales. Sun et al. (2004) utilizan PCA en dos diferentes clases de objetos y emplea un GA para seleccionar el mejor conjunto de *eigen*-imágenes para reconocer a los objetos. Es interesante mencionar, que ellos conclu-

yen que no necesariamente las *eigen*-imágenes con mayor peso en sus *eigen*-valores son las necesarias o suficientes para realizar una clasificación útil. Viola y Jones (2004) emplea la técnica de aprendizaje “Adaptive boosting” (AdaBoost) para encontrar el mejor conjunto de imágenes integrales de características, propuestas por ellos mismos, para realizar la detección de rostros.

Un algoritmo de aprendizaje de máquina es muy útil para construir o definir un operador que extraiga los rasgos relevantes para un problema de clasificación. Howard et al. (1999) emplean programación genética (GP) para generar imágenes de características para la detección de “objetivos” en imágenes de radar (SAR³). Bhanu y Lin (2003) utiliza también imágenes SAR para detectar “objetivos” utilizando programación genética en un marco cooperativo co-evolutivo. Ellos inician definiendo una base de conocimiento del dominio del problema, que llaman *características primitivas*, como un conjunto de terminales, entonces la GP puede sintetizar nuevas características que logran alta velocidad en la detección. Zhang et al. (2003) emplean GP para realizar una detección de múltiples clases de objetos pequeños en imágenes grandes. Ellos emplean estadísticas de píxel independientes del dominio como terminales en la GP, y muestran como un simple programa evolucionado resuelve la detección de objetos y su localización.

El aprendizaje visual ha influenciado también el campo de análisis faciales. Teller y Veloso (1996) estudia el reconocimiento de rostros usando una variante de la GP, que ellos nombran arquitectura PADO. Viola y Jones (2004) desarrolla uno de los algoritmos más importantes para el reconocimiento de rostros. Su trabajo muestra que puede detectar rápidamente los rostros humanos en una escena natural o cotidiana, usando descriptores locales rectangulares simples y la técnica de AdaBoost para seleccionar las características más relevantes. Silapachote et al. (2004) emplea también AdaBoost en el aprendizaje para seleccionar las características producidas por los filtros Gabor, en el problema FER. Yu y Bhanu (2006) emplean la respuesta de los filtros Gabor como un conjunto de primitivas para evolucionar su GP aplicado también al problema FER.

³por sus siglas en inglés *Synthetic Aperture Radar*.

II.3. Características de nuestro algoritmo

Este trabajo introduce un nuevo algoritmo de aprendizaje visual para el problema de reconocimiento de expresiones faciales. Nuestro algoritmo emplea imágenes térmicas como la señal de entrada para la clasificación, ver Figura 3. Asimismo, este trabajo emplea el mismo marco conceptual del estado del arte, utilizado en los sistemas FER. Nosotros aplicamos el paradigma de computación evolutiva para aprender en forma automática y autónoma las primeras dos tareas de este problema: selección de la ROI y la extracción de características. La extracción de características incluye tanto al proceso de construcción del vector de rasgos como la selección de los más característicos. El proceso de clasificación es llevado a cabo con el método de máquinas de vector de soporte (SVM). Para reforzar la clasificación empleamos un comité de votación que involucra a todas las SVM que se generen. Las características son generadas a través de las estadísticas de segundo orden calculadas por medio de la matriz de co-ocurrencia en tonos de gris (GLCM). Este algoritmo está basado en características independientes del dominio del problema. Por tal motivo puede ser fácilmente exportable a otro problema de reconocimiento de patrones.



Figura 3: Imágenes térmicas provenientes de la base de datos OTCBVS.

El algoritmo clasifica 3 expresiones faciales a partir de un conjunto de imágenes térmicas extraídas de la base de datos OTCBVS (Avidi, OTCBVS), como las mostradas en la Figura 3. En particular, la base proporciona 2 emociones primarias, *Felicidad* y *Enojo* y una secundaria *Sorpresa*. La estructura general de nuestro algoritmo es el siguiente, vea figura 4:

1. Utiliza un GA para optimizar la selección de la ROI, así como la extracción de ca-

racterísticas. El GA selecciona un conjunto Ω de n regiones de interés en la cara del individuo.

2. Simultáneamente, los parámetros π_{ω_i} de la matriz de la GLCM son optimizados $\forall \omega_i \in \Omega$, para $i = 1, \dots, n$. En este momento, el GA realiza la construcción del vector de rasgos por medio del ajuste de los parámetros de la GLCM.
3. La selección de características crea un vector $\vec{\gamma}_{\omega_i} = (\beta_1, \dots, \beta_m)$ de los m descriptores distintos obtenidos de la GLCM Ψ . En donde $\{\beta_1, \dots, \beta_m\} \subseteq \Psi$. Cada $\beta_j \in \vec{\gamma}_{\omega_i}$ representa el valor promedio del j -ésimo descriptor en la región ω_i . El conjunto $\Gamma = \{\gamma_{\omega_i}\}$ es empleado para el proceso de clasificación. Es en este sentido, que la selección de la ROI y la extracción de características están codificados en el mismo GA; así, solucionamos los dos primeros procesos del problema FER en un sólo paso.
4. Una máquina de vector de soporte ϕ_i es entrenada para cada ω_i . Seguido se calcula la calidad de la clasificación utilizando el estadístico de validación-cruzada empleando la variante “K-particiones” (K-fold). El conjunto de todas las SVM $\Phi = \{\phi_i\}$ componen el comité de votación que decide cual es la expresión facial que describe el rostro dado. La calidad total de todas las SVMs, son incluidas en nuestra función de aptitud de nuestro GA.

II.4. Conceptos Básicos

Nuestra exposición bibliográfica de la teoría básica que empleamos en nuestro algoritmo de aprendizaje, no intenta ser exhaustiva. Nuestra intención es poner en perspectiva a nuestro algoritmo y hacerlo más entendible para un espectro más amplio de diseñadores. Es así, que mostramos una revisión básica de la computación evolutiva y la construcción de descriptores a través del cálculo de la matriz de co-ocurrencia.

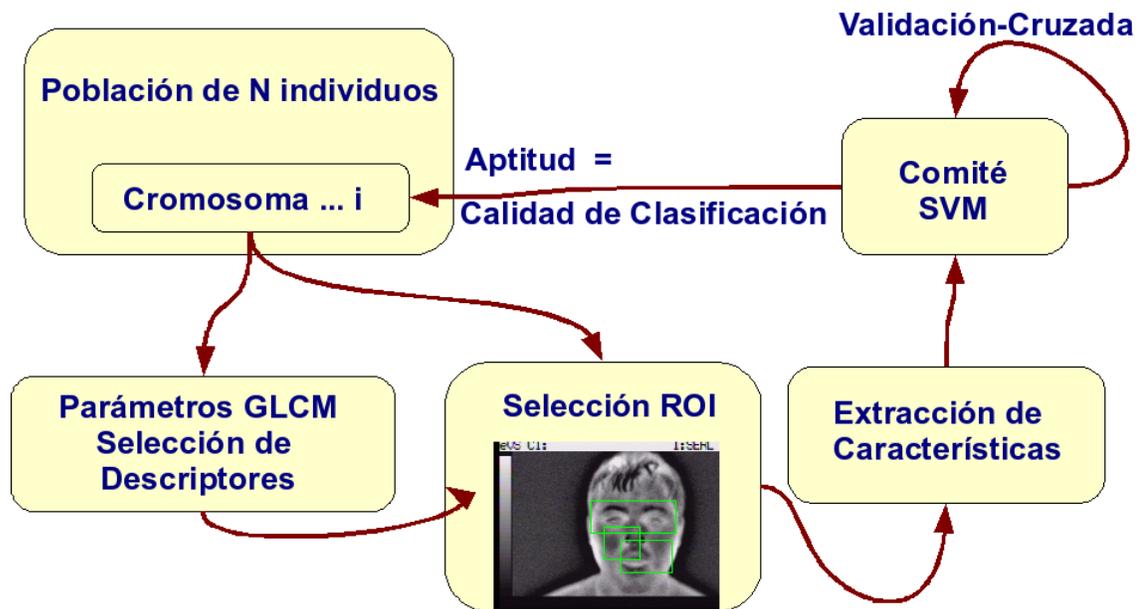


Figura 4: Diagrama de flujo de nuestro algoritmo de aprendizaje visual.

II.4.1. Computación Evolutiva

Desde la década de los 60s algunas investigaciones concluyeron que las técnicas tradicionales de IA son inadecuadas para comprender la complejidad de sistemas adaptativos. Los sistemas de IA basados en grandes bases de conocimiento y lógica de predicción, son extremadamente difíciles de manejar y contienen estructuras rígidas que lo imposibilitan para adaptarse cuando las condiciones del ambiente cambian, es decir, son poco adaptables. Esta es la razón primordial por lo que los investigadores optaron por buscar otras técnicas, para encontrar soluciones a problemas complejos no lineales, difíciles de analizar, entender y en algunos casos definir. Bajo este ambiente surge el paradigma de computación evolutiva. En años recientes las técnicas basadas en meta-heurísticas y algoritmos bioinspirados han tomado una importancia relevante. Esto debido a que, sorpresivamente, han probado su “habilidad” para buscar, optimizar, adaptarse al medio y emular el aprendizaje. Más aún, no solo la solución del óptimo es importante, si no también el análisis de la evolución del algoritmo pudiera dar cierta información para entender el problema, como el comprender porque el algoritmo toma un camino y no otro.

La computación evolutiva se fundamenta en dos grandes estudios: 1) la *Teoría de Evolución de las Especies*, postulada por Charles Darwin en 1859, la cual se basa en la supervivencia de las especies que mejor se adapten a su medio ambiente, y 2) en la mecánica de evolución postulada por diversos principios biológicos extraídos de la Genética y la Biología (Holland, 1992; Koza, 1992; Michalewicz, 1998; Olague, 1994). La simulación de estas teorías ha traído como resultado el desarrollo de algoritmos y técnicas de optimización numérica nuevas y poderosas (Gen y Cheng, 1999). Al conjunto de estas técnicas se les conoce, hoy en día, como *Computación Evolutiva* (EC⁴). En la literatura especializada, son 4 las divisiones más representativas en EC que se conocen: 1) *Algoritmos Genéticos* (GA), propuesto por Holland (1992), creador del *Teorema Fundamental* donde esboza matemáticamente la factibilidad de estas técnicas; 2) *Estrategias Evolutivas* (ES⁵), desarrolladas por Rechenberg (1973) y Schwefel (1995), que centran su atención en la forma de explotar el proceso evolutivo (selección, apareamiento, mutación y auto-adaptación); 3) *Programación Evolutiva* (EP⁶), desarrollado por Fogel et al. (1996), que aplica estos conceptos a una máquina de estado finito o autómatas, y por último ; 4) *Programación Genética* (GP), propuesto por Koza (1992), que propone la explotación de una estructura de árbol aplicando estos elementos.

El proceso de selección natural postula “que sólo el individuo dentro de una población que se adapte mejor a los cambios de su medio ambiente, es el que sobrevive”. En este sentido, Rechenberg propone la hipótesis “que el método de evolución orgánica representa la estrategia óptima para la adaptación y supervivencia de cosas dentro de su ambiente”, y concluye, “por tanto, esto debiera tener validez mientras se tomen los principios biológicos de la evolución para la optimización de sistemas”.

El algoritmo genético es una técnica estocástica, basada en el mecanismo de selección orgánica e incluye en su terminología diversos principios biológicos. En general, son 5 los componentes básicos de este algoritmo y Olague (1994) los sintetiza en la forma siguiente:

⁴por sus siglas en inglés *evolutionary computation*.

⁵por sus siglas en inglés *evolutionary strategies*.

⁶por sus siglas en inglés *evolutionary programming*.

1. Una *representación genética* de la solución del problema.
2. Una forma o técnica de crear una población inicial de las soluciones.
3. Una *función de evaluación* o clasificación de las soluciones en términos de su *aptitud*.
4. *Operadores genéticos* que alteren la composición genética de los hijos durante la reproducción.
5. Valores de los parámetros del algoritmo genético.

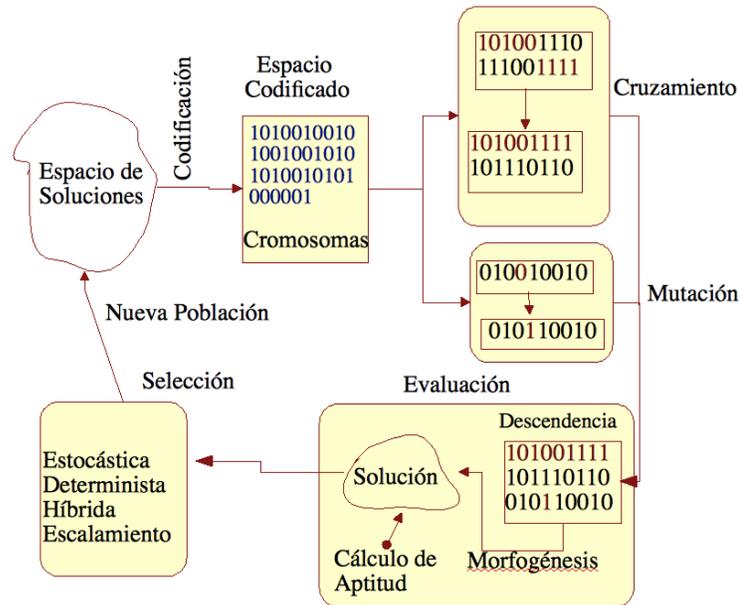


Figura 5: Estructura general de un algoritmo genético.

A partir de estos elementos el algoritmo genético se encarga de mantener a la población de los individuos, en cada una de las generaciones. Primero, se genera un conjunto aleatorio de soluciones llamada *población*. Cada individuo en la población es llamado un *cromosoma*, que representa por él mismo una solución del problema. El cromosoma evoluciona a través de iteraciones sucesivas llamadas *generaciones*. Durante cada generación, el cromosoma es *evaluado*, empleando su valor de *aptitud*. Para crear nuevos cromosomas llamados *descendencia* y producir nuevas generaciones, existen dos operadores genéticos; *mutación* que es el

proceso de generar un individuo nuevo a partir de la modificación de la estructura de uno ya existente y *apareamiento* que genera dos individuos nuevos a partir de dos padres mediante el intercambio o cruzamiento de parte de la estructura genética de cada padre, *intercambio genético*. Finalmente se seleccionan aquellos que se encuentren mejor adaptados al medio, en base a su valor de aptitud o en base a diferentes procesos estocásticos. A estos procesos se les conoce como *operaciones evolutivas*. Después de varias generaciones el proceso converge a la mejor solución, que representa el *óptimo* o *subóptimo* del proceso. En la Figura 5 se presenta en forma esquemática este proceso.

II.4.2. Textura a partir de la Matriz de Co-ocurrencia

El análisis de textura en imágenes digitales ha sido una área de investigación prominente desde los años 70s en el campo de visión por computadora. Se han desarrollado una variedad de técnicas y operadores que describen la textura de una imagen. En particular, nos centraremos en el método de la *matriz de Co-ocurrencia*.

El concepto de textura es intuitivo, no tiene una definición precisa: la superficie de un objeto en el espectro visible e infrarrojo está texturizado a una cierta escala. En un sentido amplio, la palabra textura se refiere a una superficie que caracteriza a un objeto dado un tamaño, una forma, una densidad, un arreglo o una proporción de sus partes elementales. Usualmente una textura denota un elemento liso o rugoso, blando o duro, fino o grueso, mate o brillante, etc. La textura se divide en dos categorías; *táctil* y *visual*. La textura táctil es la producida por la sensación al tocar una superficie. La textura visual, se refiere a la impresión que el ser humano tiene al observar una imagen. Generalmente producida por la variación espacial de la intensidad, color y orientación del objeto con respecto al sistema sensor utilizado para producir la imagen. Es así, que en este trabajo la palabra textura se refiere exclusivamente a la *textura visual*.

En la literatura especializada existe una gran variedad de definiciones de textura, que dependen de la aplicación particular o el enfoque que cada autor desarrolla en su aplicación, por

ejemplo:

- Tamura et al. (1978) dice: que podemos mirar a la textura como lo que constituye una región macroscópica. Estas estructuras son simplemente atribuibles a la repetición de patrones cuyos elementos o primitivas son organizadas acorde a una regla de colocación.
- Sklansky (1978). Una región en una imagen tiene un textura constante si un conjunto de estadísticas locales ú otras propiedades locales de una función de cuadro son constantes, varían lentamente o son aproximadamente periódicas.
- Haralick (1979). La imagen de una textura, ellos la consideran como no-figurativa y celular. Una imagen de textura es descrita a través de números y tipos de sus primitivas (tonales) y la organización espacial o disposición de sus primitivas. Una característica fundamental de la textura es: que no puede ser analizada sin un marco de referencia de sus primitivas tonales que son indicadas o implicadas. Para una superficie suave en tono de gris, existe una escala tal que cuando es examinada, ésta no tiene textura. Entonces cuando la resolución aumenta, adquiere una textura fina y luego se convierte en una textura gruesa.
- Richards y Polit (1974). La textura se define para sus propósitos como un atributo de un campo que no tiene ningún componente que parezca numerable. Las relaciones de fase entre los componentes no son evidentes. Ni debe el campo tener un gradiente obvio. El intento de esta definición es enfocar la atención del observador hacia las propiedades globales del desplegado – v.g. sus rugosidades, protuberancias o suavidad total. Físicamente, patrones no numerables (aperiódicos) son generados mediante procesos estocásticos a diferencia de los determinísticos.
- Zucker y Kant (1981). Aquí, la textura es aparentemente una noción paradójica. Por un lado, es usada frecuentemente en el procesamiento visual de la información, especialmente para propósitos prácticos de clasificación. Por otro lado, nadie ha tenido éxito en

producir una definición comúnmente aceptada de la textura. La solución de esta paradoja, proponen, dependerá de la riqueza del modelo, desarrollado para el procesamiento de la información visual temprana, en donde un aspecto central será generar sistemas con muchos niveles de abstracción distintos. Estos niveles pudieran probablemente incluir la intensidad actual al principio e ir progresando a través de los descriptores de los bordes y orientación de la superficie, y quizá con descriptores volumétricos. Dadas estas estructuras multi-nivel, parece claro que éstas debieran ser incluidas en la definición y en el cálculo de descriptores de textura.

A partir de estas definiciones, es claro que el problema de textura y su definición, dependen del grado de abstracción con el cual se está trabajando, cual es objetivo final del algoritmo, cual es el sistema de percepción que se esta empleando y que debieran de generarse diferentes descriptores, cuyo peso en el proceso de discriminación, dependa del grado de abstracción.

En el campo de visión por computadora existen una gran variedad de procesos y métodos para el tratamiento de la textura. Si la textura se define como una función de la variación espacial de la intensidad del pixel, entonces la identificación de sus propiedades, v.g. su homogeneidad, entropía, contraste, etc., pueden ser aplicables para procesos de clasificación de textura. Esto es, producir mapas de clasificación en una imagen en donde se identifiquen regiones uniformes, descriptores de textura más o menos similares, asociadas a una clase dada. Un segundo tipo de problema de la textura es la segmentación. La segmentación consiste en obtener el límite exacto o región dentro de la imagen en donde una clase es definida. Para los mapas temáticos de la tierra del campo de la geofísica, específicamente en detección remota, es de suma importancia el problema de segmentación. Finalmente, la síntesis de la textura, que intenta reproducir un modelo de textura previamente aprendido, es de suma importancia en el campo de graficación, en donde la meta es redibujar las superficies de los objetos lo más foto-realista posible.

De manera similar que la diversidad de definiciones y las bastas áreas de aplicación de la textura, en visión por computadora, existen también una gran variedad de métodos para su

proceso. Los más comunes son:

- Los métodos basados en **modelos**. Los métodos de modelado *estocástico* intentan modelar el comportamiento de la textura, a partir de la búsqueda de una distribución conjunta de probabilidad sobre sus *campos variables*, utilizando una distribución conjunta representativa de la textura en cuestión. Estos métodos se basan en la hipótesis de que una imagen digital es una realización de un proceso estocástico sobre campos variables. Asimismo, asume que la intensidad de cada pixel depende únicamente de la intensidad de su vecindario. Los modelos Markov-Gibbs Random Field (MGRF) son usados ampliamente para modelar texturas (Besag, 1974; Hassner y Sklansky, 1980; Zhu et al., 1998).

Los modelos basados en *fractales*, son usados cuando una imágenes presentan auto-similaridad, repetición de un patrón geométrico complejo en diferentes escales. Estos modelos son independientes de la escala, y son usados para representar texturas de la naturaleza (Pentland, 1984; Keller et al., 1987).

- Los modelos *espectrales* fundamentan su análisis en el hecho de que el sistema de visión humano analiza la textura a través de la descomposición de una imagen en su frecuencia y orientación de sus componentes (Campbell y Robson, 1968). Coggins y Jain (1985) utiliza la transformada de *Fourier* para generar filtros entonados en diferentes frecuencias y orientaciones. Cada filtro contiene una frecuencia y una orientación en la cual responde la textura, *filtros selectivos*. Así, son capaces de segmentar y clasificar imágenes naturales.

Los *bancos de filtros* selectivos son otro método espectral para el análisis, clasificación y síntesis de la textura. Un banco de filtros consiste en una gama de filtros paralelos, en donde cada filtro es entonado en una frecuencia espacial, orientación y escala particular. La técnica de generar estos filtros es muy variada (Leung y Malik, 2001; Malik et al., 1999; Cula y Dana, 2001; Portilla y Simoncell, 2000; Varma y Zisserman, 2002). Existen también los basados, en filtros Gaussianos, pasa baja Gaussiano, Laplacianos,

el Laplaciano de una Gaussiana, los basados en coeficientes “wavelet” y los filtros Gabor. Recientemente algunos autores, Varma y Zisserman (2003) y Gimel’farb et al. (2003), han cuestionado la efectividad de los bancos de filtros, debido a que: 1) pueden generarse una gran cantidad de bancos que respondan tanto a los rasgos característicos de una textura como a los que no lo son y 2) al hecho de que los filtros usualmente no son ortogonales, por lo que la información de los filtros está correlacionada y dificulta la separación precisa de la descripción de la textura. Otros métodos espectrales pueden consultarse en Reed y Wechsler (1990).

- Los métodos *estadísticos* definen a la textura en relación de un conjunto estadístico de primer o segundo orden o de orden superior. Las estadísticas de primer orden, como la media, desviación estándar y el cálculo de diferentes momentos del histograma, capturan las propiedades de los píxeles individuales. Las estadísticas de segundo orden, calculan la relación espacial de interdependencia de dos píxeles en una orientación y magnitud particular. Ejemplo de ello son los métodos basados en la matriz de *Co-ocurrencia* (Haralick, 1979; Hernández et al., 2007), función de auto-correlación y patrones locales de operadores binarios (Ojala et al., 1996), son las técnicas más populares. En la sección siguiente explicaremos en detalle la técnica basada en la matriz de co-ocurrencia.

II.4.2.1. Matriz de Co-ocurrencia en Tonos de Gris

Haralick et al. (1973), introduce el concepto de *Matriz de Co-ocurrencia en Tonos de Gris* (GLCM⁷) para cuantificar la distribución espacial y la dependencia de dos tonos de gris, o números digitales, en una región local dada. A partir de esta información, la GLCM puede generar un conjunto de estadísticos que han demostrado un funcionamiento confiable en tareas tales como clasificación (Kjell, 1994) y en el problema de recuperación de información de imágenes (Howarth y Rüger, 2004; Ohanian y Dubes, 1992).

⁷por sus siglas en inglés *Gray-Level Co-Occurrence Matix*.

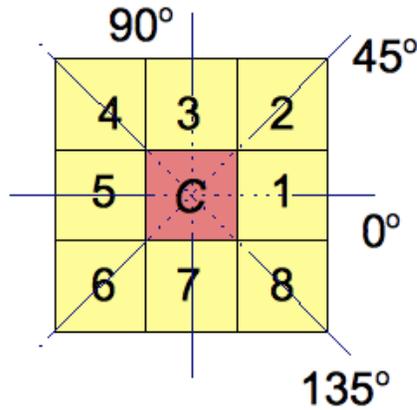


Figura 6: Celda de resolución de los vecinos más cercanos de una posición dada C . Las celdas $\{1, 5\}$, $\{2, 6\}$, $\{3, 7\}$ y $\{4, 8\}$ están a orientadas a 0° , 45° , 90° y 135° respectivamente.

Formalmente la matriz de co-ocurrencia $M_{i,j}(\pi)$ define una función de densidad conjunta $f(i, j | \vec{V}, \pi)$, en donde i y j son los niveles en tonos de gris de dos píxeles separados por un vector \vec{V} y $\pi = \{\vec{V}, R\}$ son los parámetros de $M_{i,j}(\pi)$. La GLCM calcula cual es la frecuencia que aparece un pixel en la orientación del vector $\vec{V}(d, \theta)$ y que tanto difiere en intensidad de otro dado por $\Delta = i - j$, para una región R de una imagen I . R es una región espacial que define una ventana cuadrangular de los vecinos en la cercanía de un pixel. A cada elemento se le denomina *celda de resolución*, ver figura 6. El vector \vec{V} está definido por la distancia d entre cada pixel, generalmente euclidiana, en la orientación del ángulo θ . Así, la ecuación II.1 define la construcción de la matriz de co-ocurrencia.

$$M(i, j) = \sum_{p,q \in R} \begin{cases} 1 & \text{Si } I(p, q) = i \text{ y } I(p \pm d' \cos(\theta), q \pm d' \sin(\theta)) = j \\ 0 & \text{otro valor} \end{cases} \quad (\text{II.1})$$

En forma específica, la información contextual de la textura está representada adecuadamente por la matriz de frecuencias relativas $M(i, j)$, en las cuales dos celdas de resolución separadas por una distancia d ocurren en una imagen. La primera con un tono de gris i y la segunda con un tono de gris j . Esta matriz de frecuencias de las dependencias espaciales de los tonos de gris están en función de la relación angular entre las celdas de resolución

en el vecindario, así como de la distancia entre ellos. Consideremos que se desea calcular la GLMC de una sección de la imagen $I_{5 \times 5}$ en la vecindad de un pixel cualesquiera. $I(p, q)$ tiene 3 niveles distintos de tonos de gris consecutivos. Entonces para encontrar la frecuencia del tono i cuando ocurre el pixel j a una distancia $d = 1$ y una orientación $\theta(135^\circ)$ se tiene, ver Figura 7:

- $M(i, j)$ es una matriz en donde el i -ésimo renglón corresponde a un tono de gris del espacio vertical de la imagen y la j -ésima columna corresponde a un tono de gris del espacio horizontal de la imagen. Para una imagen en tonos de gris la matriz M tiene 256×256 elementos como máximo.
- Seguido se identifica cada celda de resolución que se encuentra orientada por $\vec{V} = (1, 135^\circ)$. Es importante enfatizar que θ representa únicamente la orientación de la celda de resolución y no el sentido del vector. Es por ello que para un ángulo de $\theta = 135^\circ$ existen dos celdas, ver figura 6. El símbolo \pm en la ecuación II.1 representa ésta consideración. Una matriz de co-ocurrencia construida de esta manera siempre será una matriz simétrica.

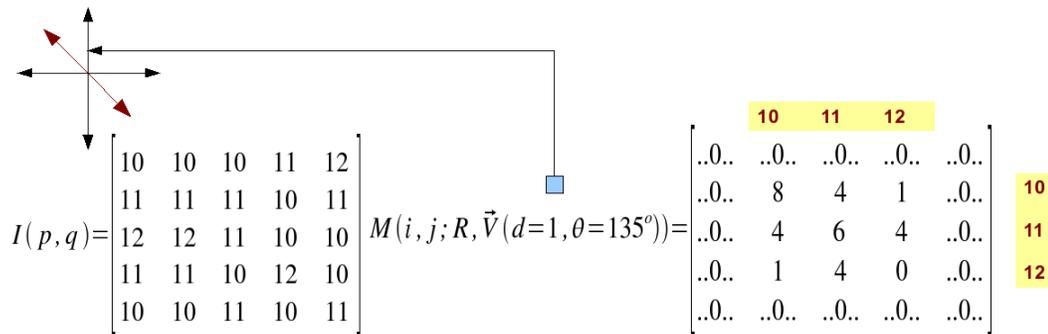


Figura 7: Construcción de la matriz de co-ocurrencia.

Si se requiere aumentar la resolución espacial entre i y j , entonces se debe construir una nueva matriz de celdas de resolución que duplica el número de celdas así como el número de ángulos posibles para analizar. Por ejemplo, si $d = 2$ el número de celdas será de 16 y se pueden analizar 8 ángulos diferentes. Por otro lado, es importante considerar que en la ecuación II.1, d' denota la distancia euclidiana, por lo que se deberán

hacer ajustes dependiendo de la ubicación de la celda de resolución. Por ejemplo para los ángulos de 0° y 90° la distancia $d' = d$ y para los ángulos de 45° y 135° el valor será de $d' = d\sqrt{2}$. Finalmente, debe tomarse en cuenta que los índices de la región de la subimagen en estudio pueden o no pertenecer a R .

- Entonces se cuenta el número de veces que ocurre el i -ésimo tono de gris en relación con el j -ésimo número digital, del conjunto de celdas de resolución dadas por \vec{V} . En el ejemplo de la Figura 7 la diagonal principal de M muestra la frecuencia con que ocurre cada valor de intensidad orientada a 135° y a una distancia 1, v.g., el valor de intensidad 10, 11 y 12 se presentan 8, 6 y cero veces en la subimagen. De la misma forma el valor $i = 10$ en presencia de $j = 11, 12$ orientado por \vec{V} se presenta 4 y 1 veces en la imagen y así sucesivamente.

Si normalizamos la frecuencia dividiéndola entre el número de ocurrencias, GLMC define la función de densidad de probabilidad conjunta, $f(i, j|\vec{V}, \pi)$, de que tan probable es que la intensidad i aparezca cuando j está en una orientación dada. De la misma forma, GLMC también denota la función de probabilidad marginal f_x si sumamos todos los renglones y f_y si sumamos las columnas.

A partir de GLMC, Haralick et al. (1973) propone una serie de descriptores que lo ayudan a identificar regiones de interés en el problema de clasificación de imágenes. En este trabajo nosotros utilizamos los siguientes descriptores de textura calculados a partir de la GLMC para clasificar las expresiones faciales:

- *Entropía*. La entropía es una medida del desorden de un sistema. Imágenes con escenas homogéneas tienen una entropía baja, mientras que imágenes con alto desorden tienen una entropía alta. La ecuación siguiente muestra el cálculo de la entropía:

$$H = 1 - \frac{1}{N_c \cdot \ln(N_c)} \sum_i \sum_j M(i, j) \cdot \ln(M(i, j)) \cdot 1_{M(i, j)} \quad (\text{II.2})$$

en donde N_c es el número de ocurrencias y

$$1_{M(i,j)} = \begin{cases} 0 & \text{Si } M(i, j) = 0 \\ 1 & \text{otro valor} \end{cases} \quad (\text{II.3})$$

- *Contraste.* Esta es una medida de la diferencia entre los valores de intensidad de los pixeles vecinos. Las frecuencias situadas fuera de la diagonal principal de $M(i, j)$ contribuyen exponencialmente mientras más alejadas estén de ella. La expresión es:

$$C = \frac{1}{Nc(L-1)^2} \sum_{k=1}^{L-1} k^2 \sum_{|i-j|=k} M(i, j) \quad (\text{II.4})$$

en donde L es el número de niveles de gris.

- *Homogeneidad.* Ésta es una medida de que tan uniformemente estructurada está una región con respecto a las variaciones de intensidad.

$$Ho = \frac{1}{Nc^2} \sum_i \sum_j M(i, j)^2 \quad (\text{II.5})$$

- *Homogeneidad Local.* Esta medida proporciona la homogeneidad de una imagen usando un factor de peso, tal que para valores no homogéneos fuera de la diagonal principal, $i \neq j$, este descriptor nos entrega valores pequeños.

$$G = \frac{1}{Nc} \sum_i \sum_j \frac{M(i, j)}{1 + (i - j)^2} \quad (\text{II.6})$$

- *Directividad.* Proporciona valores altos cuando dos regiones con el mismo valor de intensidad están separados por medio de una translación.

$$D = \frac{1}{Nc} \sum_i \sum_j M(i, j) \quad (\text{II.7})$$

- *Uniformidad.* Esta es una medida de la uniformidad de los niveles de los tonos de gris.

$$U = \frac{1}{Nc^2} \sum_i M(i, i)^2 \quad (\text{II.8})$$

- *Momentos*. Expresa la información estadística de la co-ocurrencia de la intensidad de los píxeles, como la varianza. La magnitud de este descriptor se incrementa cuando la mayoría de los valores están fuera de la diagonal principal de $M(i, j)$.

$$Mom_k = \sum_i \sum_j (i - j)^k M(i, j) \quad (\text{II.9})$$

- *Momentos Inversos*. Se comporta de manera opuesta al descriptor anterior.

$$Mom_k^{-1} = \sum_i \sum_j \frac{M(i, j)}{(i - j)^k}, i \neq j \quad (\text{II.10})$$

- *Máxima Probabilidad*. Este operador extrae las más probables diferencias entre los valores de intensidad de la región en estudio.

$$\max(M(i, j)) \quad (\text{II.11})$$

- *Correlación*. Esta es una medida de la dependencia lineal de la intensidad de los píxeles considerando una posición específica y en relación a cada uno de los otros valores.

$$S = \frac{1}{Nc\sigma_x\sigma_y} \left| \sum_i \sum_j (i - m_x)(j - m_y)M(i, j) \right| \quad (\text{II.12})$$

$$m_x = \frac{1}{Nc} \sum_i \sum_j iM(i, j)$$

$$m_y = \frac{1}{Nc} \sum_i \sum_j jM(i, j)$$

$$\sigma_x^2 = \frac{1}{Nc} \sum_i \sum_j (i - m_x)^2 M(i, j)$$

$$\sigma_y^2 = \frac{1}{Nc} \sum_i \sum_j (j - m_y)^2 M(i, j)$$

II.5. Algoritmo genético de aprendizaje visual para el reconocimiento de expresiones faciales

Nuestro algoritmo es un método novedoso para atacar el problema FER. Emplea al paradigma de EC con el propósito de aprender en forma autónoma y automática, cuales son

las regiones más promisorias de la cara de un individuo que describen mejor una expresión facial. Asimismo, encuentra cuáles son los descriptores de textura óptimos para realizar esta tarea, en dos sentidos: construyendo un vector de características y seleccionando un conjunto de ellas. También en forma automática y autónoma. Dado que nuestro problema es de clasificación empleamos la técnica de máquina de vector de soporte tanto para la evaluación final del resultado, así como un parámetro en nuestra función de aptitud del algoritmo genético.

En forma específica, el GA busca el mejor conjunto Ω de las n ROIs de la cara del individuo en cada una de las imágenes. Optimiza el proceso de extracción a partir de los parámetros de la GLMC dado por el conjunto $\pi_i \forall \omega_i \in \Omega$, en donde $i = 1, 2, \dots, n$. Selecciona el mejor subconjunto $\{\beta_1, \dots, \beta_m\}$ que contiene los valores promedios de cada descriptor de todos los posibles descriptores Ψ que forman el vector de características $\vec{\gamma}_i = (\beta_1, \dots, \beta_m)$ para cada $\omega_i \in \Omega$. Con esta representación, nosotros realizamos en un solo paso la selección de ROI y la extracción de características. En este sentido, el GA está aprendiendo la mejor estructura, de todas las posibles, del sistema FER en un solo lazo cerrado del esquema de aprendizaje. Nuestro algoritmo elimina la necesidad de un diseñador humano, ya que éste normalmente combina o realiza la selección de las ROIs y la extracción de características. Ésta decisión es llevada a cabo por nuestro mecanismo de aprendizaje. Cada posible solución es codificada en un solo cromosoma, vea la Figura 8. El cromosoma completo contiene 90 variables binarias de diferente tamaño y ésta compuesto de dos secciones lógicas: la primera codifica las variables para la búsqueda de la ROI y la segunda define a los descriptores locales apropiados de la GLCM para cada una de las regiones de interés. El proceso se explica a continuación.

II.5.1. Selección de las ROIs

La primera sección del cromosoma codifica completamente a la ROI. El GA tiene una estructura jerárquica que incluye tanto a los parámetros de control como a la descripción de la región en estudio. La sección de los genes de control c_i determina el estado de su correspondiente ROI, caracterizada por el bloque de aleles ω_i . Los estados que puede tomar son: activo o prendido $c_i = 1$ y desactivado o apagado $c_i = 0$. Cada ω_i define la posición espacial

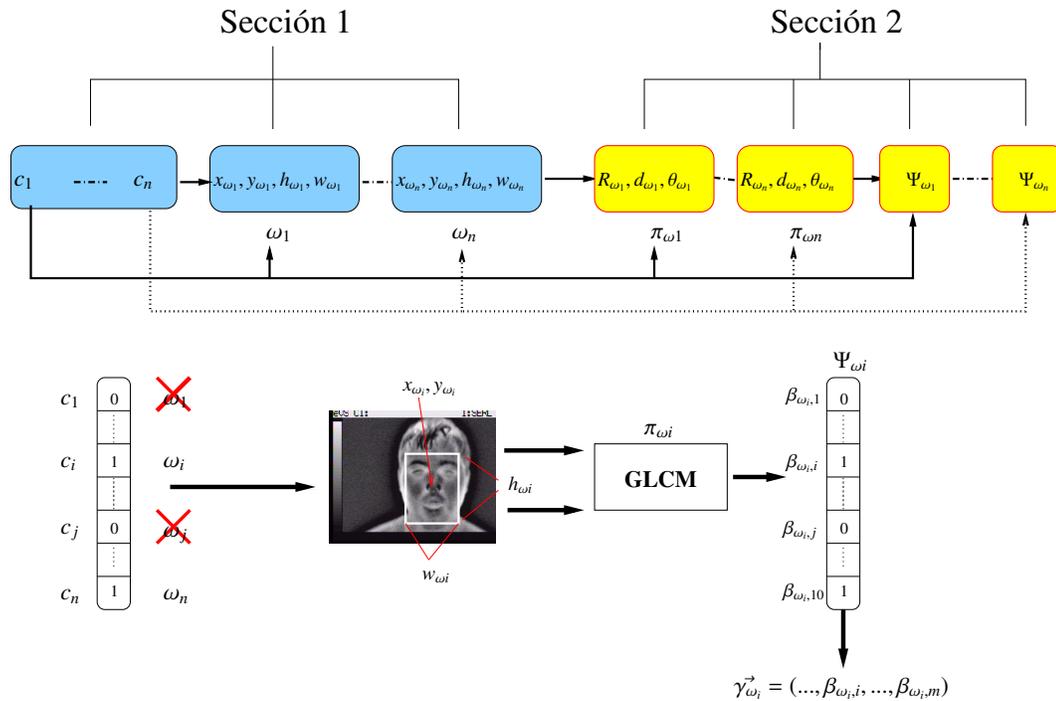


Figura 8: Descripción detallada del cromosoma de aprendizaje visual.

y el tamaño del marco de su correspondiente ROI. Cada ROI es definida con cuatro grados de libertad, los cuales son: *altura* y *ancho* de la región rectangular y dos variables para definir el estado *dentro* de ella. Cada uno de estos marcos no tienen rotación en nuestro algoritmo. Sin embargo es posible usar otro tipo de regiones tales como elípticas. En particular, la estructura es definida por una única estructura compuesta por:

- Cada una de las primeras 5 variables $\{c_1, \dots, c_5\}$ son de un bit de longitud. Éstas controlan la activación de la ROI que debe ser usada en el proceso de extracción de características.
- Cada ROI es definida por $\{w_1, \dots, w_5\}$. Existe siempre una relación unívoca entre c_i y w_i . Cada alelo w_i está definido por cuatro parámetros $w_i = \{x_{\omega_i}, y_{\omega_i}, h_{\omega_i}, w_{\omega_i}\}$, en donde cada uno de ellos consta de cuatro bits. Las coordenadas centrales del marco de referencia en la imagen están dados por $(x_{\omega_i}, y_{\omega_i})$. La altura de la venta es h_{ω_i} y el ancho por w_{ω_i} .

II.5.2. Extracción de características

La segunda sección del cromosoma, codifica las variables involucradas en la selección de los mejores rasgos que utiliza nuestro algoritmo de aprendizaje visual. El primer grupo contiene el conjunto de parámetros π_i que definen la construcción de GLCM calculados para cada marco de referencia ROI $\omega_i \in \Omega$. El segundo grupo funciona como una cadena de decisión que activa o desactiva el uso de un descriptor en particular $\beta_j \in \Psi$ para cada ROI. Esta segunda cadena de control contiene 10 bits, uno para cada descriptor. Debido a que cada una de estas variables ó parámetros están asociadas a una ROI, entonces dependen del estado de la variable estructural c_i . Es así, que solamente son consideradas cuando su ROI correspondiente está en el estado de activo, $c_i = 1$. Es por esto que, nuestro algoritmo está basado en la extracción de características de dominio independiente de la información de textura extraída a partir de la GLMC. Lo anterior provoca que nuestro algoritmo realice aproximaciones imparciales, lo cual es útil para diferentes problemas del aprendizaje visual (Zhang et al., 2003). En forma real, nuestro algoritmo puede ser fácilmente adaptado para manejar diferentes tipos de datos para el problema de reconocimiento de patrones, en donde se necesite encontrar regiones prominentes y rasgos relevantes. En concreto, el cromosoma es definido como sigue:

- El conjunto paramétrico ϕ_{ω_i} es codificado $\forall \omega_i \in \Omega$. Cada ϕ_{ω_i} tiene tres parámetros, $\phi_{\omega_i} = \{R_{\omega_i}, d_{\omega_i}, \theta_{\omega_i}\}$. Estos tres parámetros definen completamente a la matriz de co-ocurrencia, ver ecuación II.1. En donde R es una región cuadrada alrededor de un pixel cualesquiera, d es la distancia a la cual los vecinos cercanos se encuentran y θ la orientación de dichos pixeles. En resumen, ϕ_{ω_i} caracteriza a las celdas de resolución del GLCM.
- El cromosoma cuenta con una cadena de decisión o control que utiliza un sólo bit. Esta cadena activa o desactiva un descriptor $\beta_{\omega_i, j} \in \Psi$ para una ROI dada. Entonces Ψ_{ω_i} determina el número de elementos del vector de características $\vec{\gamma}_{\omega_i}$, con el propósito de buscar la mejor combinación de los descriptores de la GLCM. En esta representación, $\beta_{\omega_i, j}$ es el valor promedio del j -ésimo descriptor calculado en la ROI ω_i .

II.5.3. Clasificación

En nuestro algoritmo nosotros empleamos SVM para realizar la tarea de clasificación. Dado que tenemos ω_i regiones distintas para evaluar la calidad de la clasificación, empleamos un comité simple de votación, ver Trujillo et al. (2005). El comité Φ , está formado por todo el conjunto entrenado ϕ_i de la máquina de vector de soporte, uno para cada ω_i . ϕ_i mide la calidad o precisión de la clasificación. Además, el conjunto compuesto por los vectores de características de los descriptores de textura $\Gamma = \{\vec{\gamma}_{\omega_i}\}$ alimenta al comité Φ , en el cual cada $\vec{\gamma}_{\omega_i}$ corresponde exactamente a ϕ_i . Es así, que el comité SVM, utiliza un esquema de votación para determinar cual es la clase que le corresponde a cada expresión facial.

II.5.4. Función de aptitud

La meta de nuestro algoritmo es encontrar las mejores regiones de interés posibles y extraer las características de los descriptores de textura que realizan un buen reconocimiento de las expresiones faciales. Debido a lo anterior, la mayoría de las funciones de aptitud tienden a tener una polaridad inversa al valor de la calidad de la clasificación de cada ROI seleccionada. Adicionalmente, si dos soluciones diferentes tiene el mismo valor de clasificación, es preferible que se seleccione el cromosoma que utilice el menor número de descriptores de textura. Con esta política, se promueven resultados compactos en nuestra representación y se mejora el rendimiento computacional de nuestra aplicación. Tomando lo anterior en consideración la función de aptitud f es definida como:

$$f = 10^2 * A + 0,25 * Z \quad (\text{II.13})$$

en donde A es el promedio de los valores de todas la SVM en Φ para cada uno de los individuos de la población. En forma específica, $A = \frac{1}{|\Phi|} \sum_x A c_x$ sumado $\forall \phi_i \in \Phi$, en donde $A c_x$ es la precisión de la clasificación dada por ϕ_j . Z es el número total de descriptores inactivos en el cromosoma dados por $Z = \sum_i \sum_j \beta_{\omega_i, j} \omega_i$ para $i = 1, \dots, 5$ y $j = 1, \dots, 10$. La función de

aptitud está basada en el trabajo de Sun et al. (2004).

II.5.5. Parámetros del algoritmo genético

Nuestro algoritmo genético fue diseñado con los parámetros siguientes:

- *Tamaño de la población e inicialización.* Se emplea una inicialización aleatoria para generar los cromosomas de la población. El algoritmo utiliza un conjunto de 100 individuos.
- *Método de supervivencia.* Para que un individuo sobreviva se emplea la estrategia elitista. Los N padres de la generación $t - 1$ así como sus respectivos δ descendientes son combinados en un solo conjunto $N + \delta$. Entonces únicamente los N mejores individuos pasan a formar la población en la generación t .
- *Operadores genéticos.* Debido a que el cromosoma está codificado en forma binaria, nosotros empleamos solo un punto para el apareamiento y un solo bit para el proceso de mutación. La probabilidad de apareamiento es de 0,66 y la de mutación es de 0,05. Cada padre es seleccionado utilizando el método de torneo. El torneo está compuesto por 7 individuos.

II.5.6. Parámetros de entrenamiento

Para implementar el método de la máquina de vector de soporte se utilizaron las librerías propuestas por Chang y Lin (2001). Para cada $\phi_i \in \Phi$ los parámetros son los mismos para toda la población. Estos parámetros son:

- *Núcleo.* Se utiliza un núcleo Gaussiano, conocido en inglés como *radial basis function* (RBF), dado por:

$$k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (\text{II.14})$$

Este núcleo muestra un muy buen rendimiento para clasificar problemas no-lineales respecto de los otros grupos de núcleos.

- *Conjunto de entrenamiento.* Se utilizan 92 imágenes diferentes para las tres expresiones en estudio. Cada una de estas imágenes posee diferentes ángulos de exposición con respecto a la cámara infrarroja.
- *Validación.* Para calcular la exactitud o calidad de cada SVM, nosotros utilizamos la técnica estadística conocida como *validación-cruzada de k-dobles*⁸ con $k = 6$. Dado que nuestra base de datos contiene un número reducido de imágenes, la exactitud es calculada utilizando éste estadístico y omitimos cualquier otra prueba (Goutte, 1997). En la estimación validación-cruzada de k -dobles, el conjunto de entrenamiento es dividido en k subconjuntos más o menos del mismo tamaño. Entonces se extrae una de estas submuestras para la validación o prueba y se emplean los $k - 1$ subconjuntos restantes para el entrenamiento. Esto se repite k veces usando exactamente un subconjunto de prueba distinto. El promedio resultante de todas las pruebas-entrenamiento es usado como el estimador de exactitud o calidad de la clasificación ϕ_i de la máquina de vector de soporte.

II.6. Experimentación y resultados

II.6.1. Base de datos OCTBS

En el experimento se empleó la base de datos proveniente del laboratorio Imaging, Robotics and Intelligent Systems Laboratory UT/IRIS⁹, en el apartado de imágenes térmicas llamada OCTBVS, (Avidi, OCTBVS). Esta base contiene las imágenes de 30 personas diferentes. Cada persona es tomada con tres diferentes expresiones faciales que denotan el estado emocional: **Sorpresa**, **Felicidad** y **Enojo**. Estas imágenes son utilizadas por los diseñadores

⁸conocida en inglés como *k-fold cross-validation*.

⁹<http://imaging.utk.edu>

como la base de comparación o comprobación, y se consideran ciertas o son la “base de verdad”, para desarrollar algoritmos que ataquen el problema FER. Las imágenes de la base de verdad, expresan la emoción que el sujeto asegura es el estado emocional en el instante de la toma y así se etiqueta la fotografía. Esta base no incluye lo que los observadores humanos creen sentir del individuo que ha expresado su emoción en la toma de la imagen. Este punto es medular, y pudiera ser considerada como un defecto de la base de datos si el sujeto está solamente “actuando” y no es sincero en lo que “siente” dada una emoción. Este hecho es evidente en la Figura 9, y nosotros haremos algunos discernimientos de las consecuencias de esto en la sección siguiente.



Figura 9: Todas estas expresiones faciales parecieran estar en el estado de “felicidad”, pero en la base de datos algunas no se etiquetan como tal.

Las imágenes han sido tomadas desde 12 diferentes ángulos con respecto a la cámara. Cada imagen tiene el formato RGB bitmap con un tamaño de 320×240 píxeles. La base de datos contiene individuos con lentes oscuros en algunas de las tomas. Nosotros extraemos de la base de datos 33 que denotan *sorpresa*, 26 *felicidad* y 33 *enojo*, para un total de 92 imágenes de prueba y entrenamiento, las cuales usamos en nuestro algoritmo de aprendizaje visual. Adicionalmente, analizamos únicamente a la región rectangular que comprende la cara. Para ello, se utilizó el algoritmo que propusimos en Trujillo et al. (2005). Este algoritmo busca la posición exacta de nariz del sujeto de prueba, mediante el cálculo del centroide geométrico de todos los píxeles de la imagen que sobrepasan cierto umbral. Seguido se hacen ajustes finos sobre los dos ejes buscando el pixel que representa la menor información térmica

y se extrae el marco de referencia, ver Figura 10. Esta área es reducida a una serie de imágenes de tamaño 32×32 píxeles en tonos de gris en formato bitmap.



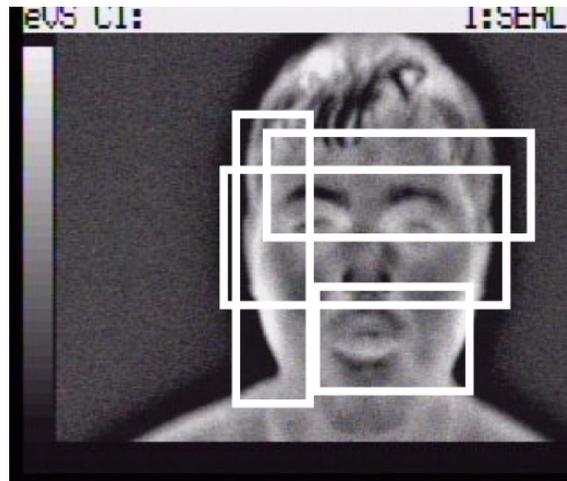
Figura 10: Calcula de la región más prominente de la expresión facial y su escalamiento a una marco de 32×32 píxeles.

II.6.2. Resultados

Es importante puntualizar que el número de imágenes que empleamos en la evaluación de nuestro algoritmo y el cálculo de la calidad de clasificación es limitado. Como mencionamos en la sección II.4.2.1, las técnicas de aprendizaje de máquina que trabajan con un número reducido de imágenes para el entrenamiento, emplean la técnica de validación-cruzada para evitar el sobreajuste, y así obtener resultados estadísticos relevantes. Esto se debe a que cuando se realiza validación-cruzada, se asegura que todas las imágenes son usadas para validar la clasificación. Los datos del conjunto de validación nunca son usados para el proceso de entrenamiento.

Para nuestras pruebas nosotros utilizamos 10 diferentes corridas del algoritmo de aprendizaje visual genético. En este trabajo mostramos el mejor superindividuo, es decir el que obtuvo el mejor valor en la función de aptitud, de estos 10 experimentos. Este individuo obtuvo una calidad del 77 % en la validación-cruzada. Asimismo, este cromosoma extrae 4 regiones diferentes en la imagen con cierto grado de traslape como se observa en la Figura 11. La matriz de confusión de los resultados de la relación-cruzada se muestran en la Tabla II.1.

A partir de los datos de la Tabla II.1, nosotros observamos que la clase *felicidad* es la que



ROI

Figura 11: Ubicación de las 4 regiones de interés del mejor cromosoma de nuestro experimento.

presenta una degradación en la clasificación. Sólo el 70 % de los individuos de esta clase se clasifican correctamente y los restantes se confunden con *enojo*, en un 18 %, y con *sorpresas* en un 12 %. Más aún, el 20 % de los individuos de la clase *sorpresas*, son confundidos con la clase *felicidad* y el 16 % de integrantes de la clase *enojo* también son confundidos con la clase *felicidad*. Esto es causado por dos razones principalmente: la primera, el traslape de las expresiones modeladas por los sujetos de prueba, vea la Figura 9, y la segunda, la forma cuestionable en la cual el proveedor de la base de datos ha etiquetado cada imagen. Esto es un ejemplo de como los sujetos de prueba pueden estar solo actuando y no mostrar una expresión sincera de lo están sintiendo. Lo que redundo en que los datos pueden presentar ruidos o perturbaciones, si los diseñadores de base de verdad no toman el cuidado suficiente en su construcción.

Debido a este problema de la base de entrenamiento, nosotros usamos 10 imágenes adicionales como nuestro conjunto de prueba que no fueron empleadas durante el proceso de aprendizaje. Nuestro algoritmo, clasifico correctamente 8 de las 10 imágenes. La matriz de confusión se muestra en la Tabla II.2

Adicionalmente, se elaboró un experimento estadístico con seres humanos para validar el

Tabla II.1: Matriz de confusión con los valores de relación-cruzada.

	<i>Sorpresa</i>	<i>Felicidad</i>	<i>Enojo</i>
<i>Sorpresa</i>	77 %	20 %	3 %
<i>Felicidad</i>	12 %	70 %	18 %
<i>Enojo</i>	0 %	16 %	84 %

Tabla II.2: Matriz de confusión de los sujetos de prueba.

	<i>Sorpresa</i>	<i>Felicidad</i>	<i>Enojo</i>
<i>Sorpresa</i>	4	0	0
<i>Felicidad</i>	0	2	1
<i>Enojo</i>	0	1	2

hecho de que la base de datos que usamos establece un problema difícil de clasificación. Se seleccionó en forma aleatoria a 30 imágenes del conjunto de entrenamiento. Se les solicitó a 100 personas que clasificaran a dicha muestra. El experimento fue repetido usando la imagen correspondiente en el espectro visible. El objetivo es tener una medida de la dificultad que presenta el uso de imágenes térmicas para discernir entre una y otra expresión facial. Este tipo de experimentación está reportada también por Sugimoto et al. (2000), que pone de manifiesto también, que los problemas en la construcción de las bases de verdad públicas provocan una complejidad adicional en la elaboración de la validación de los algoritmos.

Los resultados de la clasificación humana de las imágenes térmicas y las correspondientes al espectro visible son mostradas en las Tablas II.3 y II.4, respectivamente. Este experimento pone en evidencia la dificultad extrema del problema FER presentado en este documento. En las tablas anteriores podemos observar que nuestro algoritmo de aprendizaje visual supera a la clasificación hecha por los seres humanos; en la Tabla II.3, el porcentaje de éxito para las expresiones *sorpresa*, *felicidad* y *enojo*, la clasificación humana tiene valores de 56 %, 48 % y 79 % respectivamente, mientras que nuestro algoritmo de aprendizaje visual alcanza los valores de 77 %, 70 % y 84 % respectivamente. Adicionalmente, nuestro algoritmo es competitivo

con la clasificación humana realizada en las imágenes correspondientes al espectro visible; el promedio de acierto de las 3 clases es del orden del 77 %, ver tablas II.1 y II.4. Es por estas razones que el resultado obtenido es prometedor. Nosotros sugerimos que el porcentaje de acierto tan bajo de estas clasificaciones se deben básicamente a: diferencias culturales entre los sujetos de prueba y las personas que las clasificaron y lo laxo con que fue establecido y diseñada esta base de verdad en particular.

Tabla II.3: Clasificación Humana de imágenes térmicas.

	<i>Sorpresa</i>	<i>Felicidad</i>	<i>Enojo</i>
<i>Sorpresa</i>	56 %	33 %	11 %
<i>Felicidad</i>	23 %	48 %	29 %
<i>Enojo</i>	7 %	14 %	79 %

Tabla II.4: Clasificación Humana de imágenes en el espectro visible.

	<i>Sorpresa</i>	<i>Felicidad</i>	<i>Enojo</i>
<i>Sorpresa</i>	78 %	21 %	1 %
<i>Felicidad</i>	16 %	82 %	2 %
<i>Enojo</i>	5 %	25 %	70 %

Es importante contrastar estos resultados con nuestro trabajo previo Trujillo et al. (2005), con el propósito de enriquecer nuestras conclusiones. En él, presentamos un algoritmo basado en el análisis de componentes principales para realizar el proceso de extracción de características, mientras que en la selección de la región de interés se utilizaron ventanas en la cercanía de los ojos y la boca. En este trabajo, la principal contribución es el diseño de un proceso automático para localizar y extraer los rasgos faciales y como el comité de votación SVM con esquema ponderado por la calidad de la clasificación es utilizado. El nivel de calidad obtenido en Trujillo et al. (2005) para un conjunto de 30 imágenes de prueba fue del 76,6 %. Esto es comprable con los resultados obtenidos por la validación-cruzada y la calidad de las distintas pruebas presentadas con antelación. Más aún, en el trabajo previo empleamos

50 *eigen-características* por cada región, lo que excede al vector empleado en este trabajo. En el ejemplo de nuestro mejor cromosoma, ver Figura 11, se utilizaron 35 descriptores provenientes de la GLCM, mismos que se emplearon en nuestro proceso de clasificación.

II.7. Conclusiones

Este trabajo propone un algoritmo nuevo para resolver el problema FER en imágenes más allá del espectro visible. El aprendizaje visual de nuestro algoritmo es realizado con imágenes térmicas que contienen tres diferentes tipos de expresiones faciales. La técnica propuesta efectúa el aprendizaje visual a través de un algoritmo evolutivo, con el propósito de resolver automáticamente *dos* de las tres tareas principales del problema FER: *selección de las ROI* y *extracción de características*. Dado que nosotros empleamos un solo algoritmo para realizar estas dos tareas, entonces el aprendizaje se realiza en procesos paralelos. Además, la búsqueda de la solución conjunta de ambas tareas, provoca que nuestro algoritmo explote de una manera óptima la dependencia entre ellas. Esto no se toma en cuenta en forma común en la mayoría de las publicaciones relacionadas con el problema FER. La clasificación final es realizada por un comité de SVM que utilizan un esquema de votación. El comité SVM aumenta la robustez del proceso de clasificación. Esto es observable al incrementarse el grado de calidad de clasificación, cuando se compara con la obtenida en una región aislada de la imagen.

Los resultados experimentales muestran que nuestro algoritmo es capaz de separar y reconocer expresiones faciales utilizando imágenes térmicas. Esta clase de imágenes implican un problema con alto grado de dificultad. Nuestro algoritmo de aprendizaje visual, es superior hasta en un 16 % en promedio, del obtenido por la clasificación humana. El algoritmo reporta buenos resultados para el problema FER; sin embargo, la limitada muestra estadística con que contamos nos limita a presentar conclusiones más hipotéticas de las que podríamos aspirar. Nosotros proponemos que la base de datos presentada en Avidi (OTCBVS) debe ser expandida para que incluya un número mayor de muestras en los grupos que se presentan.

En contraste, la principal aportación de este trabajo es mostrar la viabilidad de hacer reconocimiento de expresiones faciales con imágenes térmicas. La información proveniente de las imágenes térmicas es normalmente considerada por los investigadores como imágenes de bajo perfil en relación con las obtenidas en el espectro visible, basándose únicamente en el prejuicio, de que las imágenes térmicas tiene una calidad pobre. Este trabajo muestra que es posible obtener buenos resultados, entre un 60 – 80 % de certeza, en la clasificación de las expresiones faciales, lo que es comparable con la calidad de discernimiento del ser humano en imágenes en el espectro visible.

Finalmente, hemos propuesto un método eficiente para resolver los dos primeros pasos esenciales en el problema de reconocimiento de patrones: Selección de las regiones de interés y la extracción de características. Nuestro algoritmo genético de aprendizaje visual es independiente de como las regiones son definidas y que tipo de característica es la relevante.

Capítulo III

Reconocimiento Morfo-Cinemático de Nebulosas Planetarias

III.1. Introducción

Las nebulosa planetarias (NPs) son uno de los objetos más hermosos en el cielo nocturno, debido a su gran diversidad de colores y formas. Las NPs son estructuras gaseosas que rodean a las estrellas después de acabarse su combustible nuclear (Chaisson y McMillan, 2001). El estudio de las NPs nos ayuda a entender la evolución de las estrellas, la composición química del medio interestelar y los efectos magneto-hidrodinámicos producidos en las estrellas. Ellas proveen, primordialmente, al medio interestelar de los elementos químicos pesados sintetizados por fusión nuclear en su interior. La cantidad de elementos pesados en las nubes de gas en las cuales se forman nuevas estrellas, son decisivos en la evolución futura de estas nuevas estrellas. Esta cáscara nebulosa se ve alterada en su forma debido a la interacción de vientos estelares consecutivos con propiedades cambiantes (López, 2000; Zhang y Kwok, 1998; Soker y Hadar, 2002). Entonces, estos vientos producen una amplia gama de estructuras a diferentes escalas y formas globales en las NPs y son las responsables de las morfologías observables en ellas.

Una NP es una estructura tridimensional (3D). Determinar la forma 3D de estos objetos es una tarea compleja, debido a que las observaciones desde la tierra proveen sólo de un pun-

to de vista o proyección sobre el plano del cielo. Desde la primera clasificación morfológica de Curtis (1918), muchos trabajos se han publicado tratando de entender y clasificar las diversas morfologías observadas en estos objetos. En general, las NPs son clasificadas acorde a su estructura global (Manchado et al., 1996; Manchado, 2004) que se ha visto enriquecida con el descubrimiento de micro estructuras tales como flujos colimados (López, 2000; Kwok et al., 2000) y los nudos de gas de baja excitación (Balick et al., 1998). Existen 4 morfologías básicas de las NPs a partir de sus patrones de simetría: 1) esféricas, 2) elípticas, 3) bipolares y 4) simetría de punto. A partir de estas cuatro formas básicas y su combinación, se han elaborado una gran variedad de esquemas de clasificación. Sin embargo, estos esquemas de clasificación dependen de la sensibilidad del instrumento con el cual es observada: exposiciones profundas frecuentemente han revelado estructuras más tenues, las cuales cambiaron la clasificación de la NP. Por otro lado, la imagen de la NP es una vista bidimensional de los efectos de proyección sobre el plano del cielo, en donde la morfología aparente debe ser complementada con observaciones cinemáticas del objeto para obtener su morfología 3D.

Para recuperar la estructura 3D de una nebulosa planetaria a partir de imágenes bidimensionales, nosotros hemos utilizado un catálogo morfológico que incluye diferentes proyecciones sobre el plano del cielo de las principales clases de nebulosa planetarias, así como una descripción detallada de su cinemática. Debido a que la expansión de las partículas de la NP es aproximadamente homóloga, conservando su forma, existe un mapeo entre la velocidad y la posición de cada una de ellas. Esto nos permite la reconstrucción 3D de la estructura a partir de observaciones espectrales. Las proyecciones son producidas a través de modelos tridimensionales, incluyendo la simulación del espectro de posición-velocidad (P-V) de alta resolución de una rendija larga. Al obtener observaciones espectrales sobre todo el objeto se puede calcular el comportamiento cinemático del gas en todo el objeto y de ahí reconstruir su estructura en base a un mapeo de posición-velocidad perpendicular al plano del cielo. Entonces, nuestro problema consiste en reconocer los patrones estructurales de la nebulosa y sus espectros P-V de un objeto cualquiera y poder clasificarlo acorde a los modelos sintéticos producidos. Es así, que este trabajo está orientado a mostrar un algoritmo de clasificación

automático que hemos desarrollado para tal efecto. El algoritmo inspecciona 4 regiones de interés en el espectro P-V y una en la nebulosa. Se extraen sus *eigencaracterísticas* aplicando el análisis de componentes principales (PCA¹). Seguido, se intenta separar las *eigencaracterísticas* buscando una función de decisión en el hiperplano que mejor separe las clases, aplicando el estadístico *máquina de vector de soporte* (SVM²). Finalmente se realiza un proceso de votación en las regiones de interés, que toma en consideración la probabilidad de pertenencia del objeto para cada clase. Se calcula la clase con la máxima probabilidad y la clase con la segunda máxima probabilidad de cada región de interés del objeto. A partir de estas dos “opiniones estadísticas”, el objeto es clasificado. En particular, hemos encontrado que nuestro algoritmo clasifica correctamente el 98 % de los objetos de prueba a los cuales fue sometido. Así mismo se ha probado con 5 nebulosas reales en diferentes filtros y posiciones de la rendija.

El trabajo está estructurado de la forma siguiente. En la sección de Trabajos Previos se muestra una revisión breve de los trabajos publicados en clasificación automática de nebulosas planetarias. Seguido se plantea el problema de la identificación de la estructura tridimensional de una nebulosa a través del estudio cinemático de la misma. En la sección III.4 se describe la forma de generar modelos tridimensionales de las nebulosas y la integración a lo largo de línea de visión de ella y su información cinemática. Asimismo, se describe el instrumento utilizado en la obtención del espectro cinemático bidimensional y su correspondiente nebulosa. En la sección III.5, se detalla el algoritmo computacional para reconocer y clasificar las nebulosas planetarias de forma automática. Finalmente se presenta el resultado de la experimentación sobre imágenes sintéticas y una NP real, así como las conclusiones del trabajo.

¹por sus siglas en inglés *Principal Component Analysis*.

²por sus siglas en inglés *Support Vector Machine*.

III.2. Trabajos Previos

Son escasos los artículos publicados en el rubro de clasificación automática de nebulosas planetarias. No obstante, las técnicas computacionales de reconocimiento han sido empleadas ampliamente para la clasificación de otros objetos como las galaxias y los cuásares. Existe un número reducido de NPs en relación con el número de galaxias y estrellas observadas. La literatura reporta que existen aproximadamente 2500 NPs observadas en nuestra galaxia (Miszalski et al., 2008). Sin embargo, con el desarrollo de nuevos telescopios, terrestres y espaciales, se espera encontrar un número sustancial de NPs, en nuestra galaxia y en otras, y los procesos automáticos para su eventual clasificación serán de gran impacto en esta comunidad. Por ejemplo, el satélite **GAIA**, construido por la Agencia Espacia Europea (*ESA*) tiene como objetivo efectuar un mapa tridimensional de nuestra galaxia, *la Vía Láctea*, con más de 1000 millones de estrellas y descubrir un número “astronómico” de otros objetos, tiene planeado ser lanzado en el 2012. El telescopio de 30 mts., proyectado por una asociación de Universidades Estadounidenses y Canadienses, es un telescopio óptico terrestre cuya tecnología alcanzará a “ver más profundo” que cualquier otro telescopio y se espera que su “primera luz” sea en el 2017. Es aquí, en donde radica la importancia de los procesos automáticos de reconocimiento de patrones y en especial los de reconocimiento y clasificación de NPs. Dichos procesos coadyuvarán en el desarrollo científico que se espera adquirir.

Faúndez-Abans et al. (1996), clasifican a las NPs utilizando el esquema de abundancias químicas propuesto por Peimbert (1990) para un universo de 192 objetos conocidos. Ellos utilizan las abundancias químicas de la razón de los elementos He/H, O/H, N/H, S/H y Ar/H de los objetos para producir una clasificación de 5 clases, llamadas Tipo I, IIa, IIb, III y IV. A partir de estos datos, emplean la técnica de *análisis de conglomerados* (CA³) para producir 5 grupos con una muestra de 159 NPs bien conocidas. CA es una técnica estadística que intenta asignar a cada objeto dentro de n diferentes grupos, basándose en la similitud entre los diferentes integrantes de cada grupo. Aquí la similitud es una medida de distancia entre los

³por sus siglas en inglés *Cluster Analysis*.

objetos; distancia Euclidiana, distancia Manhattan, norma máxima, distancia Mahalanobis, producto punto, distancia Hamming. Con estos grupos, utiliza el paradigma de *Redes Neuronales Artificiales* (ANN⁴) para clasificar los objetos restantes. Faúndez-Abans et al. (1996), reportan que su algoritmo coincide con el 74 % de las clasificaciones realizadas previamente por otros autores.

Es importante hacer notar que Faúndez-Abans et al. (1996) es el único trabajo relacionado con clasificación automática de la morfología de NPs. Sin embargo, ellos emplean la composición química de estos objetos para realizar su clasificación. En oposición, nosotros empleamos parámetros cinemáticos a partir de modelos sintéticos tridimensionales del comportamiento de las nebulosas planetarias para obtener una clasificación morfológica.

Philip et al. (2002) emplean el paradigma de ANN para separar a las estrellas de las galaxias utilizando la banda *R* de las imágenes publicadas por el National Optical Astronomy Observatory (NOAO) Deep Wide Field Survey. Ellos utilizan el algoritmo de “confianza diferencial” de redes neuronales (DBNN⁵) para realizar una clasificación binaria; separación de galaxias y estrellas. DBNN consiste de 3^n redes neuronales, generalmente $n = 1$, que se ensamblan y se establece una forma de composición del clasificador generalmente dada por una compuerta *XOR*. Philip et al. (2002) experimentan con tres catálogos distintos de 402, 304 y 408 objetos y reportan que su algoritmo tiene un éxito del $\approx 98\%$.

Naim et al. (1995) clasifican 16 clases de galaxias basada en la morfología que se observa en el catálogo “APM Equatorial Catalogue of Galaxies”. Ellos emplean ANN y PCA para reducir el número de parámetros, en una muestra de 820 objetos. Reportan, que la dispersión del error de su algoritmo es de $\sigma_{RMS} = 1,8$ y es comparable con la clasificación realizada por los expertos. En particular reporta que el 9 % de las galaxias están clasificadas erróneamente ($\sigma_{RMS} \geq 2$) y el 28 % se considera totalmente acertados ($\sigma_{RMS} < 0,5$).

Hakkila et al. (2000), estudian las emanaciones de Rayos Gama, utilizando la base de objetos del catálogo BATSE 3B, para analizar dos tipos de clases, “largas-blandas-brillantes”

⁴por sus siglas en inglés *Artificial Neural Network*.

⁵por sus siglas en inglés *Difference Boosting Neural Network*.

y “cortas-duras-débil”, y proponen una tercera clase que llamaron “intermedia”. Ellos, utilizan las técnicas de CA, árboles de decisión y lógica difusa en un conjunto de 14 reglas y 13 atributos en su clasificación. Para un conjunto de 778 objetos su algoritmo reporta un 99 % de acierto. Adicionalmente, experimentan también con el catálogo BATSE 4Br y disminuyeron el número de parámetros utilizados; 4 reglas y 3 atributos. Para un conjunto de 623 objetos y el problema de clasificar las 2 primeras clases, reportan que su algoritmo tiene el mismo porcentaje de acierto.

Odehahan (1995) separa estrellas y galaxias a partir de una reducción fotométrica de imágenes obtenidas en el “Minnesota Automated Plate Scanner”, utilizando las bandas “O” y la banda “E”. Él emplea ANN en su modalidad de “propagación hacia atrás” para realizar una clasificación de 8 y 16 clases para un universo de 2258 galaxias. Reporta un error $1,6 \leq RMS \leq 1,9$, en donde el 20 % son erróneas con $\sigma_{rms} = 2,1$.

Recientemente, Gupta et al. (2004) emplean también propagación hacia atrás para clasificar 2000 “Fuentes Brillantes” de la base de datos de Calgary del *Infrared Astronomical Satellite*. Para un conjunto de 17 clases su algoritmo reporta una certeza del 80 %.

En resumen, Faúndez-Abans et al. (1996), Philip et al. (2002), Naim et al. (1995), Hakikila et al. (2000), Odehahan (1995) y Gupta et al. (2004) emplean el paradigma de redes neuronales y análisis de conglomerados para realizar las clasificaciones automáticas de sus objetos.

El análisis de componentes principales también ha sido utilizado por la comunidad astronómica para clasificar objetos de interés. A continuación citaremos algunos trabajos.

Folkes et al. (1999), utilizan la base de datos de “2-degree-field (2dF) Galaxy Redshift Survey” extraída del 2dF espectrógrafo multifibra del “Anglo-Australian Telescope”. Ellos, utilizan PCA para clasificar 5 clases de espectros en una muestra de 5869 espectros de galaxias, utilizando los dos primeros eigenespectros más significativos. Cada una de las clases de espectros corresponden aproximadamente a las morfologías E/SO, Sa, Sb, Scd y Irr de galaxias.

Vanden-Berk (2006) separan quásares y núcleos activos (AGN⁶) de la galaxia huésped utilizando PCA. Ellos emplean un universo de 7664 espectros obtenidos del “Sloan Digital Sky Survey” con un corrimiento hacia el rojo (redshift⁷) entre 0,0 y 0,75. En su separación, utilizan 3 eigenespectros para separar las AGNs y 5 eigenespectros para la detección de quásares. En particular, ellos logran detectar la galaxia huésped de 4666 AGNs.

El aprendizaje a partir de SVM, es una técnica utilizada recientemente para el análisis de imágenes astronómicas. Citaremos algunos trabajos que hacen uso de este método.

Wadadekar (2005), aplica ésta técnica para determinar el redshift (z) usando muestras de galaxias provenientes del *Sloan Digital Sky Survey (SDSS) Data Release 2* para los objetos cercanos comprendidos entre $0,01 < z < 0,5$ para un universo de 10000 objetos. Del mismo modo experimenta con el modelo sintético *Galaxies in Cosmological Simulations* para generar 6965 objetos. Para el caso de los modelos sintéticos, emplean $\sim 50\%$ para entrenamiento y prueba. En ambos casos reporta un error $\sigma_{rms} = 0,03$ en su clasificación.

Willemsen et al. (2005) aplican SVM, para clasificar estrellas sintéticas en dos clases, binarias y simples, a partir de la simulación de la banda fotométrica media del satélite Gaia. Utiliza un universo de 30,000 objetos, en donde 6,000 de ellos son usados para entrenamiento y los restantes para prueba. Ellos reportan que su algoritmo tiene una certeza del 70%.

Huertas-Company et al. (2008) utiliza un catálogo de 1472 objetos provenientes del SDDS en la banda I para clasificar galaxias lejanas, $z \sim 1$, en dos grupos, tempranas y tardías. Para entrenar a la máquina de vector de soporte, ellos utilizan 500 galaxias y las restantes para prueba. Reportan que el error en la separación es del orden del 20%.

⁶por sus siglas en inglés *Active Galactic Nuclei*.

⁷en inglés conocido como *Redshift*.

III.3. Nebulosas Planetarias

Una nebulosa planetaria es una cáscara gaseosa alrededor de una estrella en una fase especial de su evolución. Para explicar el comportamiento del gas difuso que rodea a la estrella, los astrónomos analizan la evolución estelar de la composición química del núcleo de la estrella, su medio interestelar y cuáles podrían ser los efectos magneto-hidrodinámicos que producen las diferentes formas observadas. En la Figura 12 se muestran 6 diferentes NPs con distintas formas, producto de relaciones complejas entre los fenómenos antes citados. Por ejemplo, las nebulosas Esquimal y Helix tiene una forma general esférica, M 2-9 y MyCn18 tienen una forma bipolar, He3-1475 es un ejemplo clásico de una nebulosa con simetría de punto y NGC 6891 es en esencia una elipse. Es así, que el estudio de la forma de la cáscara difusa que rodea a estas estrellas es un fenómeno relevante para la comunidad astronómica. Los astrónomos se basan en las observaciones de estos objetos con instrumentación especializada, telescopios e instrumentos en una gama grande del espectro electromagnético que va desde los rayos X hasta el milimétrico, y han desarrollado distintos modelos teóricos para la explicación de estas formas. Una buena referencia fotográfica y de los fenómenos básicos de las NPs se puede consultar en Kwok (2001).

Históricamente el nombre de nebulosa planetaria fue dado por William Herschel, quien al observar un objeto difuso en tono verdoso le recordó el tono de un planeta. Sin embargo, no existe relación entre un planeta y estas nebulosas. Con la construcción de nuevos y mejores telescopios se pudo distinguir el material gaseoso que rodea a una estrella, las NPs, de las áreas difusas que correspondían a las galaxias. Actualmente, una estimación basada en las observaciones, propone que la población de NPs en nuestra galaxia es de $28,000 \pm 5000$ (Frew y Parker, 2006). Moe y De-Marco (2006) estima que la población asciende a $46,000 \pm 13,000$, basándose en sus modelos de síntesis de población estelar. Sin embargo, el número de NPs que han sido observadas e identificadas es del orden ~ 2500 . Recientemente, Miszalski et al. (2008) han reportado el descubrimiento de 240 NPs basado en las observaciones del “AAO/UKST $H\alpha$ Survey” del plano galáctico sur.

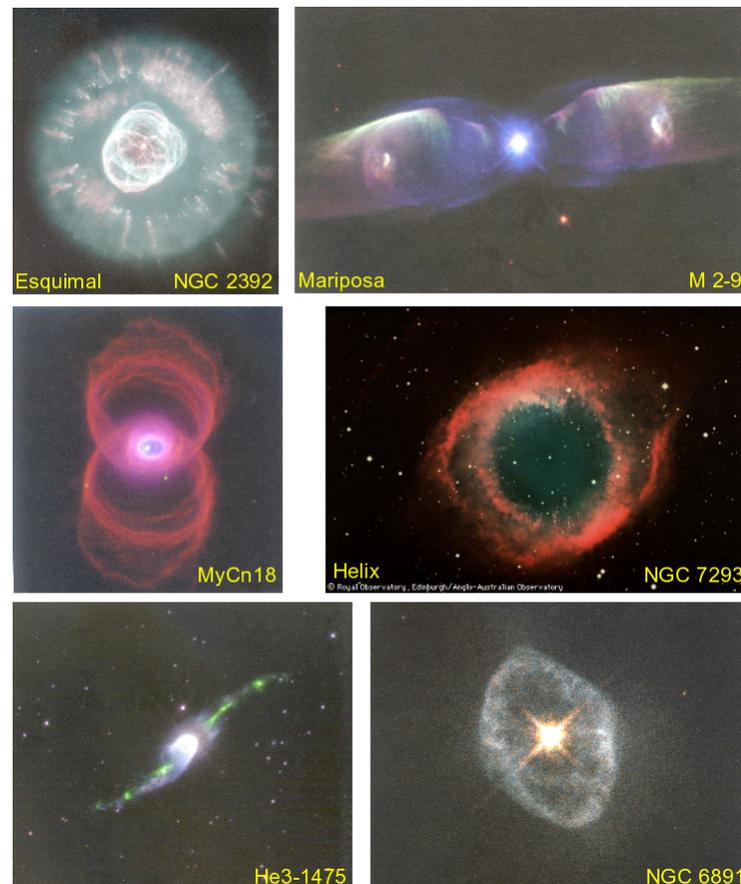


Figura 12: Diferentes morfologías de NPs. La imagen de Helix fue tomada por el Anglo-Australiano Telescope, las restantes fueron adquiridas por el Hubble Space Telescope (HST).

Sin embargo, no todas las estrellas evolucionan como una NP. Se cree, que solo las estrellas que tiene entre 0.5 masas solares (M_{\odot}) y $8M_{\odot}$ pueden evolucionar hacia una nebulosa planetaria. El proceso de evolución de una estrella, parecida a nuestro sol, es el siguiente, ver Figura 13:

- Cuando una estrella está formada, se encuentra quemando hidrógeno (H) en su núcleo y dura en este proceso alrededor del 90 % de su vida. La estrella es estable y para el caso de estrellas entre 1 y $3M_{\odot}$ este proceso tarda $\sim 10^{10}$ años.
- Una vez que en el núcleo el H se agota, se queda quemado únicamente helio (He). Esta combustión produce elementos más pesados como el carbón (C) y el oxígeno (O). En este momento, se genera un cascarón delgado alrededor del núcleo en donde se quema

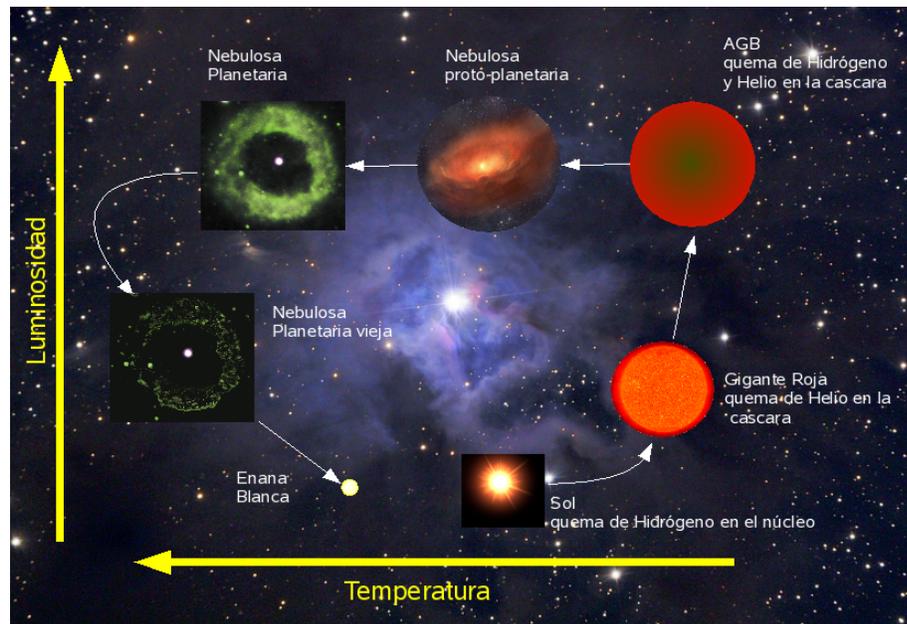


Figura 13: Diagrama esquemático Hertzsprung-Russell de la evolución de una estrella cuya masa inicial está entre 0.5 y $8 M_{\odot}$.

hidrógeno. Esta cáscara rica en H aumenta su tamaño enormemente y la temperatura superficial de la estrella desciende, lo que produce que la estrella se observe roja. Se dice que la estrella es una *gigante roja*. La mayor parte del carbono, la base de la vida, y alguna otra materia es creada y dispersada en el universo por las estrellas gigantes rojas.

- El He en el núcleo se agota y en su lugar queda carbono y oxígeno. En ese momento se generan dos cascarones; uno que quema He y otro que quema H. En la capa más interior se quema helio para producir carbono y en la siguiente se quema hidrógeno para producir helio. La estrella se encuentra en la *Rama Asintótica de las Gigantes* (AGB⁸) que son la progenitoras de las NPs. La energía producida, genera vientos estelares densos por la presencia de polvo y gas, a una velocidad de $\sim 20 \text{ km s}^{-1}$. Dicho viento está compuesto mayormente de hidrógeno, pero puede contener otros elementos previamente sintetizados en el núcleo. Esto produce una envoltura o halo alrededor de la estrella. En la Figura 14 se muestra en forma esquemática la estructura de una estrella AGB. En

⁸por sus siglas en inglés *Asymptotic Giant Branch*.

este estado la estrella pierde $\sim 80\%$ de su masa inicial.

- La envoltura de la estrella es removida por pérdida de masa. La estrella se contrae y se calienta rápidamente. El núcleo queda expuesto y produce fotones de alta energía ultra violeta (UV) y un viento poco denso pero rápido, del orden de $\sim 2000\text{km s}^{-1}$. Estos fotones ionizan la envoltura y la interacción compleja de los vientos la expanden y es cuando la **nebulosa planetaria** es “visible” y produce las líneas espectrales que observamos.
- La temperatura del núcleo expuesto decrece considerablemente y la estrella se convierte en una *enana blanca*.

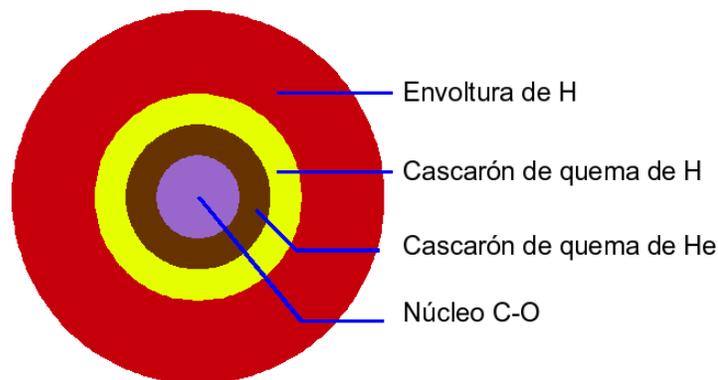


Figura 14: Diagrama de una estrella AGB, progenitora de un NP. El núcleo está compuesto por carbón y oxígeno, C-O. Reacciones de fusión nuclear ocurren en dos capas separadas. En la capa más interior se quema helio para producir carbón y en la siguiente se quema hidrógeno para producir helio. La capa exterior de hidrógeno es la que ocupa el mayor volumen de la estrella.

Se han realizado varios esfuerzos para explicar y clasificar la gran variedad de morfologías aparentes de las NPs. Correlacionando éstas con propiedades tales como, temperatura de la estrella central, masa, posición en la galaxia y muchas otras (Khromov y Kohoutek, 1968; Aaquist y Kwok, 1996; Balick et al., 1998; Masson, 1990; Zhang y Kwok, 1998; Manchado, 2004). De la misma forma se han realizado trabajos teóricos empleando conceptos tales como, modelado de gas y polvo, campos magnéticos y sistemas centrales binarios. A

pesar de esto, no se ha alcanzado un acuerdo general entre los modelos empleados y sus propiedades. Además de lo anterior, el efecto de proyección de la nebulosa planetaria la sido un problema abierto en esta área.

En la siguiente sección introduciremos dos conceptos básicos: 1) el espectrómetro utilizado en la adquisición de las imágenes reales a las cuales fue sometido nuestro algoritmo de clasificación y la generación de un espectro de posición - velocidad. Seguido, profundizaremos en los problemas de proyección sobre el plano del cielo de las nebulosas planetarias.

III.3.1. Espectrómetro

En astronomía, y en la actualidad en cualquier dispositivo óptico digital, se emplean generalmente arreglos de sensores fotosensibles bidimensionales (CCD ⁹) en la adquisición de una imagen. Éstos colectan y cuentan el número de fotones que inciden en cada posición del arreglo. Por otro lado, un *fotón* es una partícula de energía que viaja a cierta longitud de onda. Si tomamos en cuenta el modelo planetario de un átomo, un fotón es emitido o absorbido cuando un electrón pasa de un nivel a otro. Así, una estrella o fuente puntual, emite fotones dependiendo de su composición química y su grado de excitación (temperatura). Esta energía es absorbida por los elementos químicos de su atmósfera o bien por la nebulosa. En el caso de la NP, los fotones, producto de las emanaciones *UV* de alta energía de su núcleo, son absorbidos por la nube de gas de la nebulosa. Esto es, los electrones de los átomos del material en la NP son transferidos a niveles de más alta energía. Dado que la materia tiende a estar en su mínima energía, éstos pasan a un nivel inferior de energía, entonces las nubes de material emiten fotones con la energía equivalente al salto que da el electrón para su estado de reposo. Es así como las NP producen espectros de *emisión*. Después los fotones viajan por el medio interestelar (lo que está entre las estrellas) hasta llegar al telescopio. El resultado final, es que el CCD acoplado en el punto focal del telescopio, colecta fotones con una cierta longitud de onda. La cantidad de fotones colectados son proporcionales a la calidad de la óptica

⁹por sus siglas en inglés *Charge Couple Device*.

e instrumentación asociada al telescopio. En la astronomía es de especial interés conocer la descomposición del haz de luz que inciden en el CCD, su espectro, ya sea para estudios de abundancia química del objeto o su cinemática. Nuestro interés es la cinemática del objeto observado.

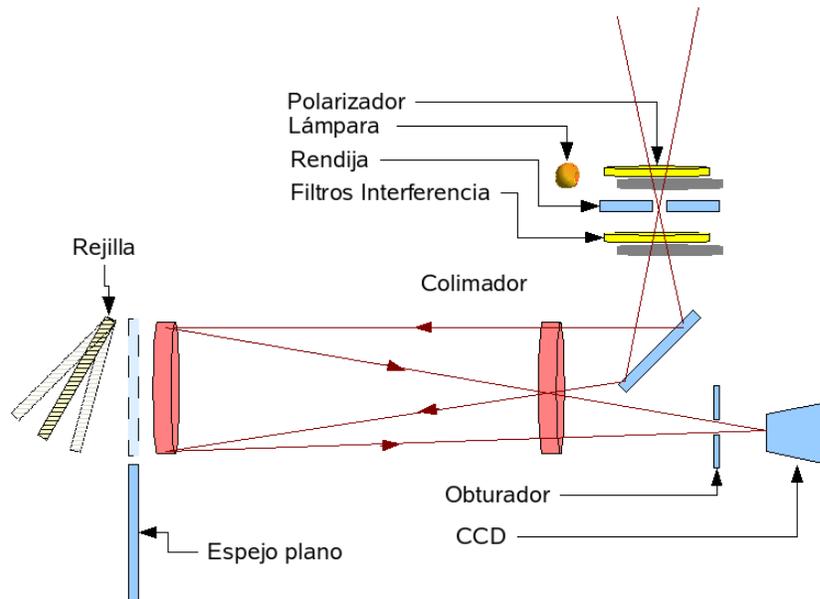


Figura 15: Esquema óptico del espectrómetro MEZCAL.

Para descomponer un haz de luz se utiliza generalmente un espectrógrafo. Si el espectrógrafo es utilizado como instrumento de medición se le llama *espectrómetro*. Un espectrómetro de *rendija* es un instrumento óptico diseñado para aislar una franja delgada de un haz de luz, llamado *beam* en inglés, que proviene del punto focal del telescopio y pasa por una *rendija*. El haz de luz de la rendija pasa a través de uno o más elementos dispersores y genera una nueva imagen en el detector, esta imagen es el haz disperso en sus diferentes longitudes de onda. Entonces, un espectro bidimensional captado por el detector consiste en una rebanada delgada del cielo en uno de los ejes y una secuencia de imágenes de la rendija a diferentes longitudes de onda expandida a lo largo del otro eje. Los espectros de imágenes astronómicas reales mostradas en este trabajo fueron adquiridas en el Observatorio Astronómico Nacional ubicado en la Sierra de San Pedro Mártir (OAN-SPM) con el instrumento *Manchester Echelle Spectrometer MES-SPM*.

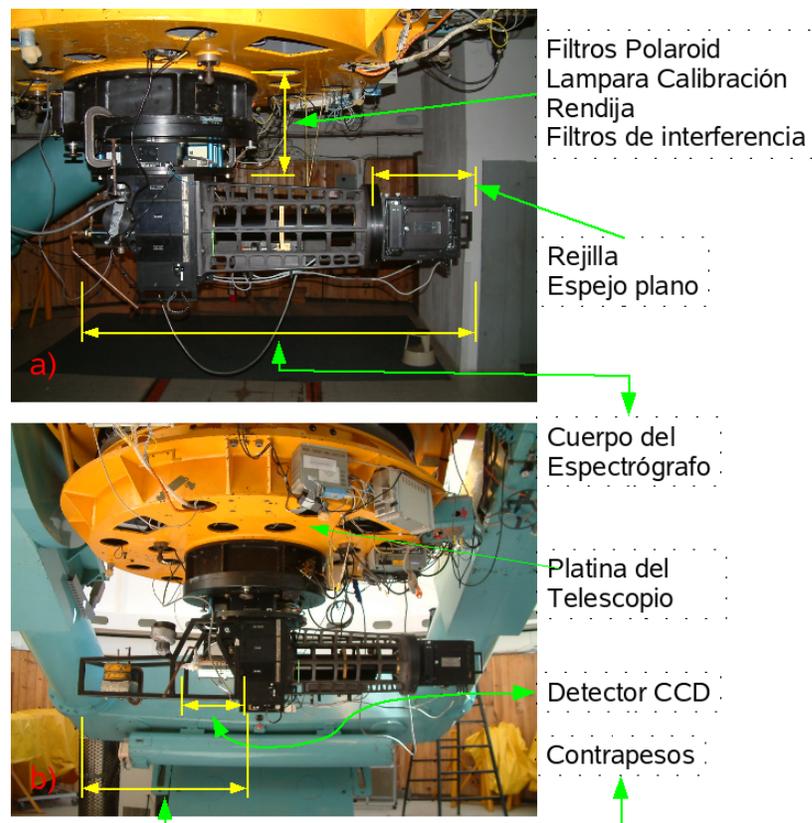


Figura 16: Ubicación de los elementos básicos del espectrómetro “MES-SPM”. La posición de rendija es ortogonal al brazo del instrumento. En el gráfico a) el brazo está orientado norte \leftrightarrow sur, entonces la rendija este \leftrightarrow oeste. El gráfico b) muestra la orientación inversa.

Los elementos de MES-SPM en la dirección del flujo óptico son los siguientes, ver Figura 15 y Figura 16: 1) un conjunto de filtros excluyentes que polarizan el haz de luz a 0° , 60° , 120° y *vacío*, en el cual el filtro 0° está alineado con la rendija, 2) un conjunto de rendijas con un ancho de $70\mu\text{m}$, $150\mu\text{m}$, $300\mu\text{m}$ y *vacía* que concentra el haz de luz en cierta columna del objeto, 3) un arreglo de filtros de interferencia H_α , $[OIII]$, $[SII]$ y *vacío*, 4) un espejo plano que permite ubicar espacialmente a la rendija con respecto al objeto observado, 5) una rejilla como principal elemento dispersor que contiene 31.6 ranuras mm^{-1} que puede oscilar $\sim \pm 1,8^\circ$. MES-SPM es un espectrómetro especializado para el estudio de un número limitado de problemas astrofísicos, v.g., en donde se requieran estudiar perfiles de líneas resueltas espacialmente provenientes de fuentes extendidas y débiles con alta relación señal a ruido (Meaburn et al., 2003).

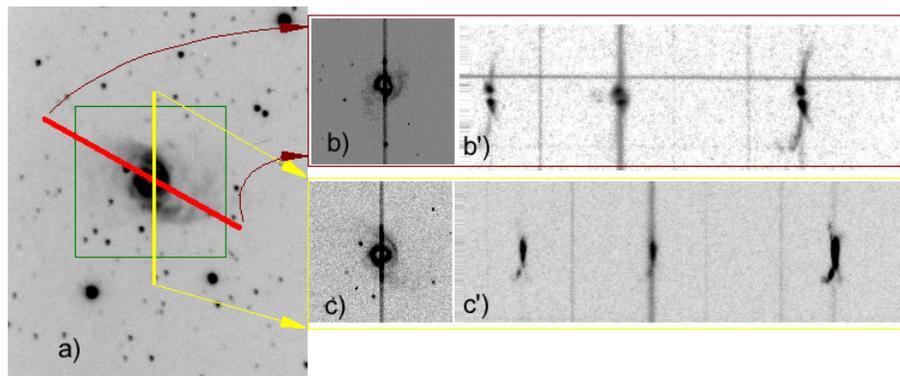


Figura 17: Imágenes adquiridas con el espectrómetro MES-SPM. Las líneas roja y amarilla muestran 2 posiciones distintas de la rendija. Los gráficos b) y c) son imágenes de la posición de la rendija girados con respecto al gráfico a). Los gráficos b') y c') son los espectros bidimensionales del haz de luz que pasa por la rendija. Cortesía de H. Tirado, A. López; Octubre 2004; Observatorio Astronómico Nacional, San Pedro Mártir B.C. México.

MES-SPM puede proporcionar una tupla de 3 imágenes en una observación, dependiendo del arreglo de los lentes y filtros que se seleccionen. La primera, puede ser una imagen del campo observado aplicando un filtro de interferencia y sea lo suficientemente profunda para capturar la morfología general de la nebulosa, que llamaremos imagen de *cámara directa* o *imagen directa*, en donde la rendija está en posición *vacío* y el espejo plano en el eje óptico del instrumento. La Figura 17.a) es una imagen directa del objeto K4-55 utilizando el filtro H_{α} y un tiempo de integración de 10 min. La segunda, una imagen menos profunda del objeto donde se ubique espacialmente la rendija, que llamaremos imagen de *posición*, ver los gráficos 17.b) y 17.c) que se adquirieron con el mismo filtro y con 100 segundos de tiempo de exposición. Este tipo de imágenes se obtienen a través de una exposición corta del objeto con la configuración anterior del instrumento y se toma otra exposición con la rendija adentro, el obturador se cierra automáticamente mientras se desplaza la rendija a su posición. La tercera imagen es el espectro bidimensional. El espejo plano está fuera del flujo óptico y los filtros, rendija y rejilla están operando, ver gráficos 17.b') y 17.c'), el instrumento está adquiriendo el espectro P-V.

III.3.2. Espectro de Posición-Velocidad

En la literatura, existen diferentes métodos para interpretar o reproducir la estructura 3D de una nebulosa. Khromov y Kohoutek (1968); Balick et al. (1998); Aaquist y Kwok (1996); Zhang et al. (2003); Kwok et al. (2000); Mellema (1995), reproducen las formas observadas a través de modelos elipsoidales, de un cilindro abierto o estructuras bipolares, que son evolucionados basados en la interacción de vientos estelares consecutivos (ISW¹⁰). Monteiro et al. (2000); Morisset et al. (2000) utilizan modelos de fotoionización para explicar las formas o bien reproducir la estructura de una nebulosa planetaria en particular. Magnor et al. (2005) aplican técnicas de renderizado¹¹ para ajustar sus modelos a las observaciones reales. (Steffen y López, 2006) generan modelos basados en la geometría observada y la descripción completa de la cinemática de un objeto. En particular, nosotros utilizaremos el vector de velocidades y la proyección de la nebulosa planetaria, en nuestro algoritmo de clasificación. Para ello es necesario explicar como se calcula la cinemática de un objeto.

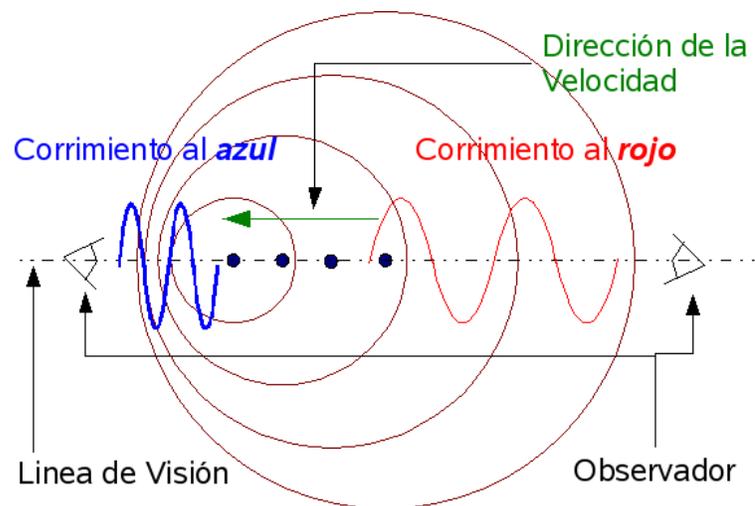


Figura 18: Efecto Doppler.

El *efecto Doppler*, propuesto por Cristian Andreas Doppler, señala que la percepción que se tiene de las ondas sonoras depende de la velocidad relativa entre el observador y la fuente

¹⁰por sus siglas en inglés *Interacting Wind Model*.

¹¹anglicismo de *rendering*.

emisora. Hippolyte Armand Louis Fizeau, en 1848, extiende este concepto y comprueba que las ondas electromagnéticas se comportan de la misma manera que el descrito Doppler. En astronomía, cuando un objeto se acerca, el espectro de la longitud de onda del elemento en reposo λ_r , es detectado en una longitud de onda inferior λ_s , entonces se dice que el objeto tiene un *corrimiento hacia el azul*. Por el contrario, cuando el objeto se aleja, λ_s tiene valores superiores a λ_r , entonces se dice que el objeto tiene un *corrimiento hacia el rojo*, ver Figura 18. La velocidad detectada por el efecto Doppler, es la componente radial de la dirección del vector de velocidad del objeto, en la línea de visión del observador. Esto es, si un objeto se mueve con una velocidad $\vec{v} = (v_r, v_t)$, en donde v_r es la componente radial en la dirección del observador y v_t es su componente tangencial, entonces el efecto Doppler corresponde a la magnitud de v_r . En casos extremos, si el objeto se desplaza perpendicularmente a la línea de visión entonces $v_r = 0$ y no existirá corrimiento hacia el azul o el rojo. La ecuación siguiente muestra el cálculo de la velocidad a través del efecto Doppler.

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda_r - \lambda_s}{\lambda_r} = \frac{v_r}{c} \quad (\text{III.1})$$

en donde λ_r es longitud de onda en reposo; λ_s es la longitud de onda detectada en el espectrómetro; v_r es la componente radial del vector de velocidad y $c = 3 \times 10^5 \text{ km s}^{-1}$ es la velocidad de la luz.

En el párrafo anterior, se ha descrito el cálculo de la velocidad en relación a un objeto puntual. Por ejemplo, si desea estudiar la velocidad radial de una estrella, objeto puntual, se pueden obtener espectros de los elementos mas abundantes de la misma (carbón, azufre, oxígeno, argón, nitrógeno, entre otros), se calcula el corrimiento y se obtiene su velocidad. Si son más de un espectro se reporta el promedio de las velocidades. Esto implica que el espectro es *unidimensional* y generalmente se le denomina *perfil de velocidad*. Para el caso de una nebulosa, que es un objeto *extendido*, es deseable conocer el comportamiento del gas en diferentes posiciones de la nebulosa y en los elementos más abundantes. Ésto, para darnos una idea de la forma del volumen del gas. Entonces es posible obtener perfiles en posiciones discretas aleatorias. En este punto es en donde los espectros de *posición - velocidad*, también

llamados en la literatura *diagramas P-V*, nos proveen de una herramienta efectiva para inferir la estructura 3D de la nebulosa, ya que nos ayuda a ordenar los vectores de velocidad.

Un diagrama P-V, en un conjunto ordenado de perfiles de velocidad radial que corresponden a la posición exacta por donde pasa la rendija, en una imagen o modelo proyectado en el plano ortogonal a la línea de visión del observador. Un diagrama P-V tiene dos ejes, el eje de las ordenadas corresponde a la posición en el objeto en donde se ubica la rendija y el eje de las abscisas corresponde su velocidad radial. En el caso de imágenes reales, el espectrómetro de rendija larga nos proporciona un espectro bidimensional. En donde un eje de las ordenadas corresponde a posición y el de las abscisas a la expansión de las diferentes longitudes de las ondas observadas. El cálculo de la velocidad se realiza *aposteriori* mediante la Ecuación III.1. Tomando en cuenta lo anterior, el diagrama P-V y el espectro P-V serán considerados equivalentes. En la Figura 19 se muestra una proyección de un modelo sintético tridimensional de una nebulosa, con densidades complejas, basado en una forma elipsoidal. En el gráfico 19.a), se presenta la proyección en el plano del cielo (imagen directa), y la ubicación de la rendija (Δ_s). Los ejes horizontales (x) y verticales (y) están en pixeles. En el gráfico 19.b), se simula el espectro de P-V que muestra sólo la velocidad radial del modelo en la porción en donde está ubicada la rendija. Esto es, para cada pixel (Δ_s, y) en la imagen directa, se calcula la magnitud de la velocidad radial, o se adquiere el corrimiento λ_r en el caso de un espectrómetro, a lo largo de y en toda la rendija. El resultado final, es que el espectro P-V es bidimensional. En consecuencia, nuestros algoritmos de reconocimiento siempre opera sobre imágenes.

En este trabajo, nosotros utilizamos diferentes proyecciones sobre la línea de visión del observador o plano del cielo, de modelos 3D de nebulosas planetarias y sus correspondientes espectros de P-V para entrenar a nuestro algoritmo de clasificación, así como imágenes astronómicas reales con sus correspondientes espectros P-V para probar a nuestro algoritmo.

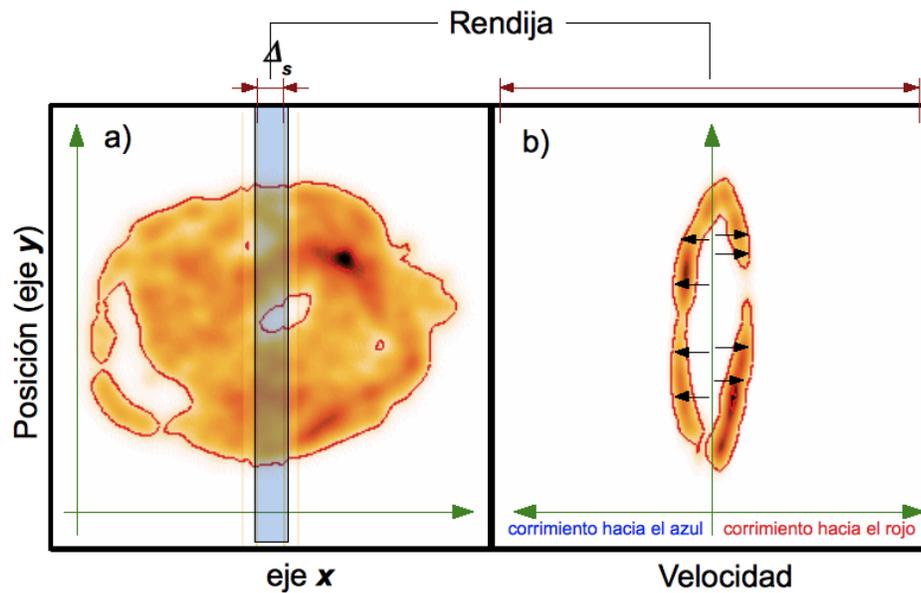


Figura 19: Espectro P-V.

III.3.3. Efectos de Proyección de una NP

Una nebulosa planetaria es una estructura 3D y la morfología que observamos de ella en la imagen directa, es solo el efecto de su proyección sobre el plano del cielo. Esta proyección juega definitivamente un papel preponderante para definir *que es lo que vemos*. Monteiro et al. (2000), muestran que algunos objetos pueden tener una clasificación errónea, si se basa únicamente en su morfología proyectada. Ellos construyen un modelo tridimensional de fotoionización de la nebulosa NG 3132, en donde pueden reproducir los perfiles de las observaciones y concluyen que esta nebulosa pertenece a una estructura bipolar y no a una elíptica. De la misma forma, Manchado (2004) muestra que el 17% de las NPs que pertenecen a la clase “redondas” en su catálogo, para nuestro trabajo esféricas, podrían realmente pertenecer a la clase “elíptica”. Más aún, los efectos de proyección pudieran ser tan ambiguos que nebulosas totalmente distintas podrían ser producto de un mismo modelo morfológico. En la Figura 20 mostramos a la nebulosa “Esquimal” (NGC 2392) y a la “Hormiga” (Mz-3) obtenidas por el HST. La serie de gráficos inferiores muestran diferentes perspectivas de un modelo tridimensional producido por Wolfgang Steffen. El primer gráfico de la secuencia muestra el

modelo visto desde el polo. Dicha perspectiva es similar a la NP “Esquimal”. Sin embargo, si se cambia de perspectiva, rotando convenientemente el modelo, podemos observar como el gráfico final es similar a la nebulosa “Hormiga”. Con esta secuencia, enfatizamos la importancia de la perspectiva en las observaciones de morfologías aparentes que puede tener una nebulosa planetaria dada.

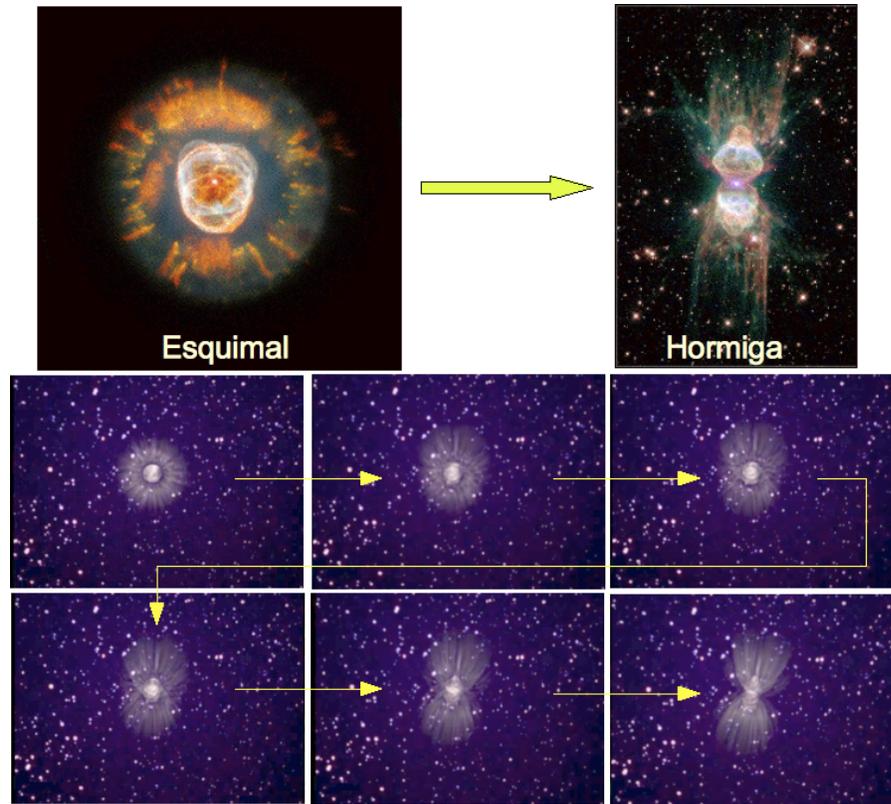


Figura 20: Importancia de los efectos de proyección de la morfología aparente de una nebulosa planetaria. Vea texto.

En general, una imagen proyectada sobre el plano de cielo provee información espacial de sus estructuras. Por otro lado, el campo de velocidades provee información de la componente radial del vector de velocidades a lo largo de la línea de visión y ofrece información de la profundidad del objeto. Esta información es limitada, ya que la componente tangencial solo es posible si se observan cambios en la imagen directa. Esto lleva muchos años entre cada una de las observaciones y depende también de la distancia del objeto. Entonces, para conocer la estructura 3D completa, se requiere de la información de todo el campo de velocidades. Esta

situación no ocurre, no obstante la información morfológica y su simetría aparente nos ayuda a resolver estas ambigüedades. Un caso simple ocurre cuando la velocidad de un elemento de volumen es constante a través del tiempo de expansión de la NP. En objetos complejos, esta distribución de velocidades puede esperarse si la envoltura de la nebulosa tiene eventos de pérdida de masa relativamente pequeños y se mueven en forma balística (Zijlstra et al., 2001), o bien mediante una continua interacción de viento con estructuras relativamente pequeñas (Steffen y López, 2004). Para estos casos, después de un tiempo suficiente, el patrón de velocidades es proporcional a la distancia de su centro (ley de velocidades tipo hubble u homólogo). La forma global de la nebulosa se comporta de manera similar, conserva su forma en el tiempo. En una gran mayoría de NPs, una expansión de velocidades proporcional a la distancia desde la estrella central es una aproximación razonable, al menos para la regiones más brillantes Wilson (1950); Sabbadin et al. (2000).

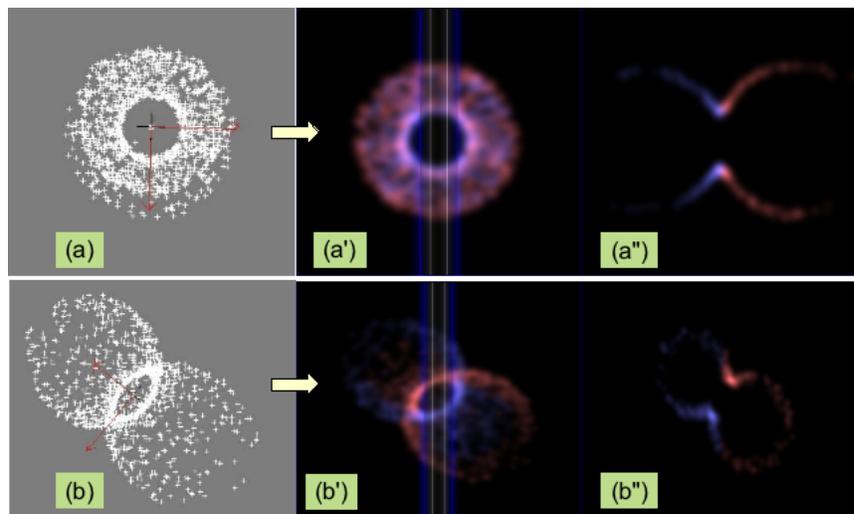


Figura 21: Modelo tridimensional de una NP bipolar con cintura. (a) y (b) distribución de partículas vistas desde el polo y a 45° sobre dos ejes. Las proyecciones en el plano del cielo se observan distintos en su forma debido a la rotación del modelo; (a') puede ser clasificada como una NP esférica, mientras que (b') se observa bipolar. Los espectros (a'') y (b'') corresponde a la posición de la rendija dada en (a') y (b') respectivamente.

Bajo estas condiciones, el vector de velocidades es proporcional al vector de posición para cada elemento dentro de la nebulosa. Así, la forma de la nebulosa a lo largo de la línea de visión es transformada linealmente en el correspondiente componente del vector de veloci-

des. Midiendo el corrimiento del efecto Doppler, podemos inferir la estructura que se pierde en la imagen bidimensional de la nebulosa. Este caso se aprecia en las nebulosas bipolares con simetría axial, en donde si la nebulosa es vista desde el polo, la imagen bidimensional se observa como si perteneciera a una morfología esférica, ver Figura 21. Sin embargo, en los perfiles de velocidad, la profundidad del objeto, se observan claramente bipolares, gráficos (a'') y (b'') de la Figura 21.

III.4. Modelado Morfo-Cinemático

La fuente de información de nuestro algoritmo de clasificación es el modelado tridimensional de la nebulosa, su proyección a lo largo de la línea de visión y el cálculo del diagrama P-V a diferentes posiciones de la rendija. Para ello se utiliza el software llamado *SHAPE*¹² (Steffen y López, 2006) desarrollado por Steffen y Koning (2006-2009). A través de *SHAPE*, se generó un catálogo de nebulosas planetarias que contiene las principales proyecciones de 11 clases de morfologías, proporcionado por Karla Escalante. A dicho catálogo lo llamaremos *Atlas Morfo-Cinemático de Nebulosas Planetarias* (AMCNP).

El proceso general de modelado es el siguiente. Mediante la inspección de las observaciones disponibles de un objeto, se obtiene una idea general de la forma de la clase de NP. Esta morfología es reproducida mediante la creación de una distribución de partículas en el espacio, con su correspondiente campo de velocidades. Las partículas pueden ser distribuidas sobre una superficie topológicamente compleja o en un volumen. El resultado de la emisión de cada partícula es integrado a lo largo de la línea de visión. El procesamiento de datos incluye principalmente, la dirección del observador, las radiaciones emitidas, la estructura interna y la resolución espacial y espectral. En la Figura 22, se presenta un sistema de partículas de una nebulosa bipolar abierta. *SHAPE* calcula la imagen proyectada de la nebulosa y su espectro P-V, desde los datos de posición y velocidad del modelo 3D. La luminosidad y el color que emite cada partícula depende de la posición en el objeto o en la subestructura del

¹²<http://www.astrosen.unam.mx/shape>

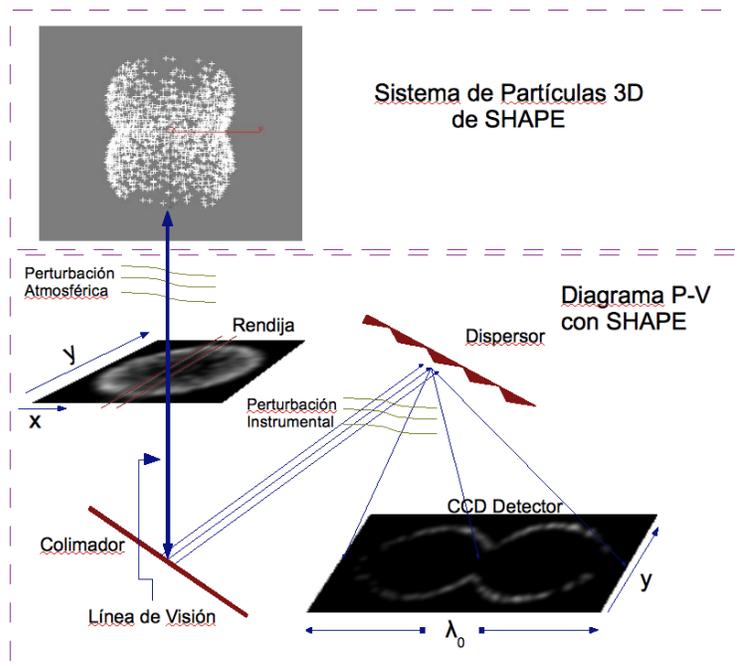


Figura 22: Simulación del proceso de adquisición de una nebulosa planetaria utilizando SHAPE.

cual es parte. Entonces, la distribución de emisiones de las subestructuras y las partículas son convolucionados con los parámetros de las perturbaciones atmosféricas, llamado *seeing* en inglés. La imagen difuminada por las perturbaciones atmosféricas es modelada mediante la convolución de una función Gaussiana PSF (point spread function). La resolución instrumental es incluida en la imagen y en el espectro P-V, también con una convolución de núcleos Gaussianos. Después de la convolución, se obtiene un diagrama unidimensional de velocidad para la estructura contenida en la rendija en una posición del eje y , vea la parte inferior de la Figura 22. Finalmente se repite este proceso a largo del eje y para obtener un espectro bidimensional de posición y velocidad.

Por ejemplo, para modelar una estructura bipolar se necesita que la expansión de las partículas se mueva más rápidamente sobre el eje de simetría de la estructura. Mientras que velocidades más lentas son asignadas a lo largo del eje ecuatorial. En la Figura 23 se muestra el modelo 3D de una nebulosa *bipolar cerrada con cintura* para las proyecciones siguientes

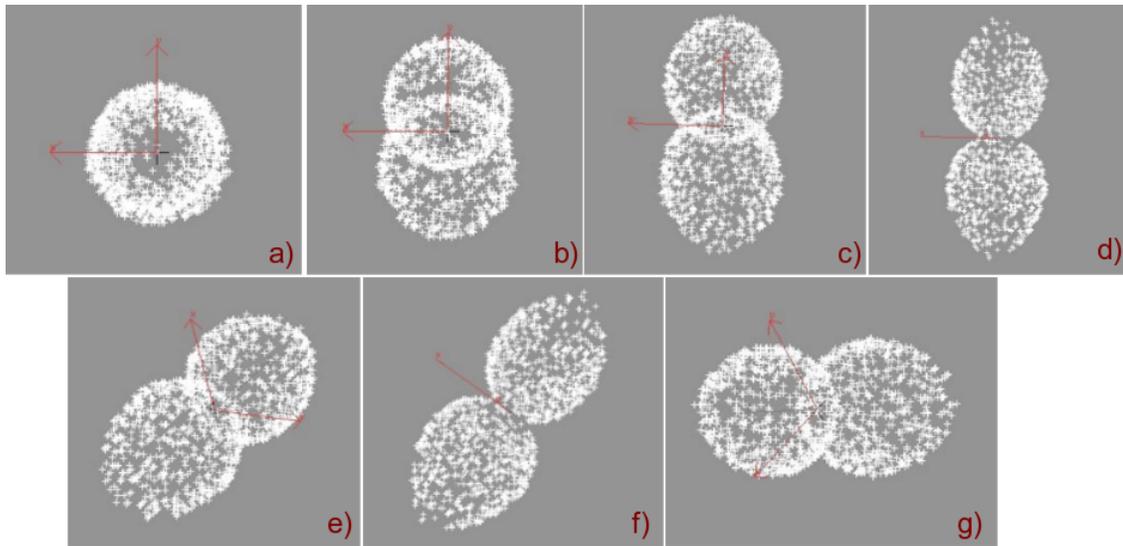


Figura 23: Puntos de visión del modelo 3D de una nebulosa planetaria bipolar cerrada con cintura del AMCNP.

a lo largo de la línea de visión: el gráfico a) muestra una orientación $x = 0^\circ$, $y = 0^\circ$, $z = 0^\circ$, denominada “vista desde el polo”, b) $x = 25^\circ$, $y = 0^\circ$, $z = 0^\circ$ un giro relativamente pequeño sobre el eje x , c) $x = 45^\circ$, $y = 0^\circ$, $z = 0^\circ$, posición media de la estructuras sobre el mismo eje, d) $x = 90^\circ$, $y = 0^\circ$, $z = 0^\circ$, vista lateral de la nebulosa, e) $x = 45^\circ$, $y = 0^\circ$, $z = 45^\circ$, rotación media de los ejes x y z , f) $x = 90^\circ$, $y = 0^\circ$, $z = 45^\circ$ y g) $x = 0^\circ$, $y = 45^\circ$, $z = 0^\circ$. Del mismo modo se calcula el espectro de P-V para cada uno de los puntos de visión en siete posiciones de la rendija. Uno en la parte central, 3 a la izquierda y 3 a la derecha del centro. En la Figura 24 se muestra el espectro cinemático de una nebulosa bipolar cerrada orientada $x = 90^\circ$, $y = 0^\circ$, $z = 45^\circ$. Los gráficos en la columna izquierda representan a 3 posiciones distintas de la rendija en el modelo 3D. De la misma forma, los de la columna derecha son 3 posiciones hacia la derecha del modelo. En cada uno de estos gráficos, la sección derecha corresponde al espectro P-V y la izquierda a la ubicación por donde pasa la rendija en el modelo 3D.

Una característica importante de *SHAPE* es que tiene la posibilidad de aplicar diferentes colores a secciones distintas de un objeto complejo. En un esquema rojo-azul, como el que se muestra en la Figura 25, nos facilita distinguir e interpretar el espectro P-V, así como las estructuras embebidas en la imagen proyectada de la nebulosa.

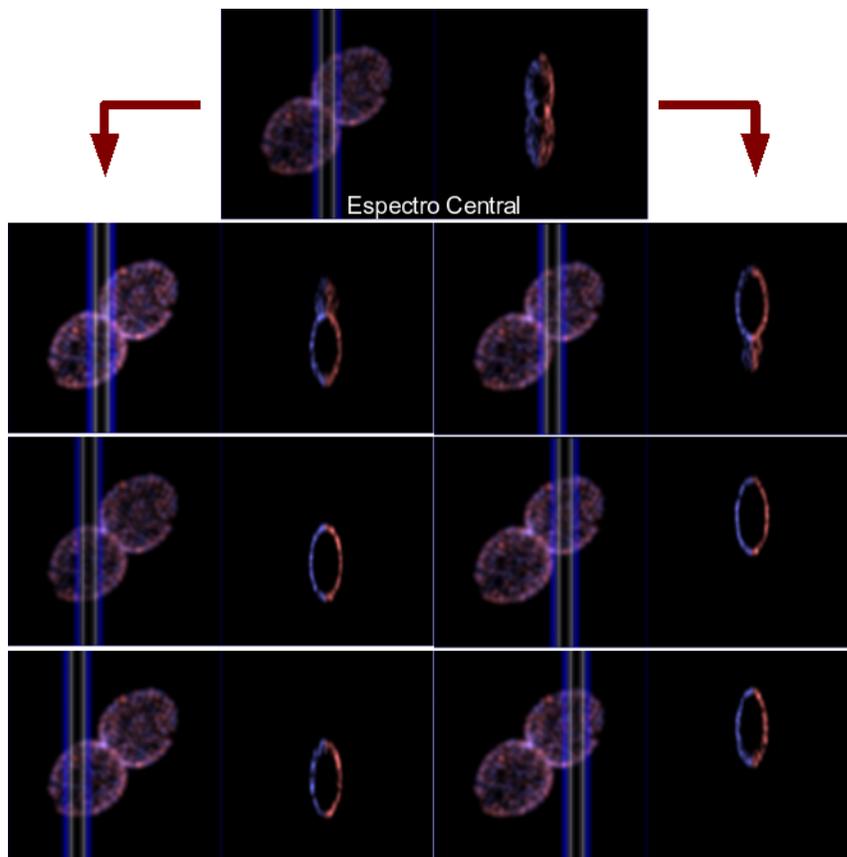


Figura 24: La base de datos esta compuesta por las proyecciones del modelo 3D de la nebulosa y el cálculo del espectro P-V en 7 posiciones distintas de la rendija.

Existe una gran diversidad de esquemas de clasificación desde la primera clasificación Curtis (1918). Algunos esfuerzos intentan reproducir las morfologías de las NPs como resultado de diferentes vistas de una estructura unificada tridimensional básica. Khromov y Kohoutek (1968), explican la morfología en términos de un cilindro abierto y su proyección sobre el cielo. Aaquist y Kwok (1996); Zhang y Kwok (1998), emplean un modelo elipsoidal. Otras aproximaciones se basan en la interacción de vientos estelares consecutivos, modelo ISW, los cuales producen las diversas morfologías (Mellema, 1995). Uno de los trabajos más citados de morfologías de NP es el catálogo del Instituto de Astrofísica de Canarias (Manchado et al., 1996)) en el cual son cuatro las morfologías básicas de las NPs a partir de sus patrones globales: 1) esféricas, 2) elípticas, 3) bipolares y 4) simetría de punto. Sin embargo, estas clasificaciones no incluyen información cinemática que permita entender



Figura 25: Muestra del AMCNP.

suficientemente la estructura tridimensional de la proyección sobre el cielo de la nebulosa planetaria.

En este trabajo, se emplean 11 clases distintas de nebulosas con diferentes integraciones sobre la línea de visión y su correspondiente comportamiento en el espectro P-V, muestreando la rendija a lo ancho del patrón de la nebulosa. Estas clases son:

- 1 Bipolar Abierta.
- 2 Bipolar Abierta con Cintura.
- 3 Bipolar Cerrada.
- 4 Bipolar Cerrada con Cintura.

- 5 Bipolar con Simetría de Punto.
- 6 Elipse.
- 7 Elipse con Densidad Ecuatorial.
- 8 Esfera con Anillo fuera de la Cáscara.
- 9 Esfera con Anillo en la Cáscara.
- 10 Esfera.
- 11 Simetría de Punto.

Para el reconocimiento y clasificación se han etiquetado con un número entero del 1 al 11 en forma respectiva. En la Figura 25 se presenta una muestra del catálogo morfológico utilizado. En la columna izquierda se presenta la etiqueta de la clase. Seguido, se muestra el volumen tridimensional para las clases 6 a 10 y el sistema de partículas para las clases restantes. Las otras columnas son la proyección bidimensional sobre el plano del cielo de la imagen directa y su espectro P-V. El espectro P-V es muestreado en 7 posiciones consecutivas excluyentes a lo ancho de la imagen. Una en la posición central de la nebulosa y 3 muestras a la derecha y a la izquierda de dicha posición. Las clase 1 al 6 y 11 corresponden a una rotación $(90^\circ, 0, 0)$ sobre los ejes (x, y, z) del modelo tridimensional de la estructura. La posición $(0, 0, 0)$ corresponde a la vista desde el polo del modelo. La elipse con densidad ecuatorial está orientada $(45^\circ, 0, 0)$. Las clases 8 y 9 están orientadas $(45^\circ, 45^\circ, 0)$. De esta forma se construyó un catálogo de 539 imágenes con las posiciones que muestran estructuras relevantes en su espectro P-V, ver cuadro III.2. Ahora, el problema es reconocer y clasificar automáticamente las imágenes acorde a dicho catálogo.

III.5. Esquema de Reconocimiento

El reconocimiento de objetos embebidos en imágenes digitales es un problema abierto en el campo de visión por computadora. La digitalización de un objeto de una escena real

y su posible reconocimiento depende primordialmente de dos factores principalmente: a) las características físicas del sistema de adquisición y b) los parámetros o mediciones que se emplean para discriminar entre uno y otro objeto. Los primeros factores causan que siempre exista un grado de incertidumbre en las posibles mediciones que puedan ser extraídas de una imagen digital. Estos factores primordialmente son: 1) la posición relativa de la cámara con respecto al objeto, 2) la orientación interior de la cámara, calibración, 3) calidad de cielo, iluminación en fotometría de rango corto, 4) oclusión de los objetos ya sea por otros o modificación de su apariencia debido a elementos traslúcidos en la línea de visión de los mismos, e 5) instrumentación asociada, óptica y electrónica, del dispositivo sensor (Olague y Hernández, 2005). Los segundos, se relacionan con el conocimiento experto de los objetos que se pretende reconocer y la capacidad de modelar dicho conocimiento en términos de algoritmos computacionales e inteligencia artificial. Existen dos grandes categorías de métodos para reconocimiento de objetos. Los primeros utilizan descriptores estadísticos globales calculados desde la imagen. Estas técnicas determinan el objeto más probable en la escena, mediante el emparejamiento de la estadística de una región de la imagen con la estadística de un objeto previamente muestreado (Schiele y Pentland, 1999). Los segundos se basan en descriptores locales y su relación espacial entre ellos. Estos métodos se basan en la construcción de vectores de características locales, como derivadas Gaussianas, los cuales contienen información discriminante relevante. Estos vectores se emplean para representar a los objetos y su reconocimiento. Esta categoría tiene la característica de poder reconocer a los objetos si están rotados, escalados o visto desde diferentes perspectivas (Schmid et al., 2000). Una buena referencia de algoritmos de clasificación puede ser consultada en Duda et al. (2001) y en Schölkopf y Smola (2001). Entonces, nuestro problema consiste en reconocer de un conjunto, $(\Gamma_1, \Gamma'_1, \Gamma'_2, \dots, \Gamma'_l, y_1), (\Gamma_2, \Gamma'_1, \Gamma'_2, \dots, \Gamma'_l, y_2), \dots, (\Gamma_m, \Gamma'_1, \Gamma'_2, \dots, \Gamma'_l, y_m)$, cual es el mapa correcto $(\Gamma_i, \Gamma'_j) \rightarrow y_i$; en donde Γ_i son imágenes que pertenecen a la proyección sobre el cielo de las nebulosas, Γ'_j sus correspondientes espectros de velocidades y y_i las etiquetas o clase asociada de Γ_i, Γ'_j .

La base de datos consta de dos objetos principales; la morfología general de la NP mode-

lada en una imagen y su espectro de posición-velocidad. Es importante enfatizar, que existe una relación $1 - a - 1$ en el eje de las ordenadas de la rendija Δ_s de la imagen directa y la imagen de su espectro P-V. De manera similar, el diagrama P-V representa el campo bidimensional de velocidades sólo de los píxeles contenidos en Δ_s , ver Figura 19. Nuestro algoritmo utiliza un subconjunto de las posibles proyecciones de los modelos tridimensionales de las NPs, así como un subconjunto de posiciones de la rendija. Tomando en cuenta estas consideraciones, la base de datos se divide en dos conjuntos disjuntos: el primero de entrenamiento que permite establecer las bases del reconocimiento y el segundo de prueba para validar el algoritmo. En forma general nuestro algoritmo consiste en, ver Figura 26:

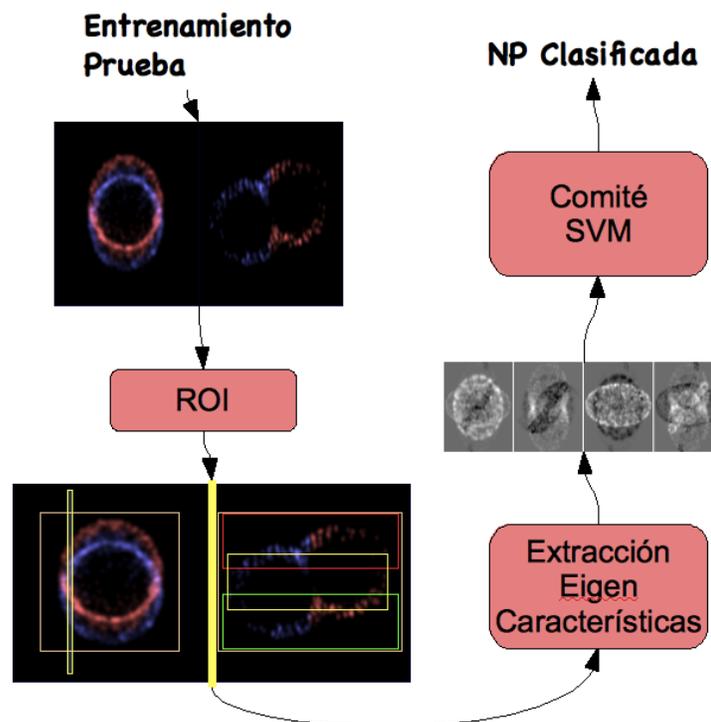


Figura 26: Esquema general de reconocimiento y clasificación de las NP.

1. Selección de la región de interés. Debido a que nuestras imágenes de entrenamiento y prueba están bien registradas, el centro de masa del espectro y de la imagen directa corresponde al centro geométrico de sus respectivas imágenes. A partir de estos centros analizamos 5 regiones de interés definidas como, ver Figura 27:

- a) Ω_d es la proyección a largo de la línea de visión del observador del modelo tridimensional de la NP. En general una ventana de 201×201 píxeles de ancho y alto respectivamente, es suficiente para incluir todas las clases de imágenes del catálogo.
- b) Ω_w es el espectro bidimensional P-V completo con una ventana de 251×201 píxeles.
- c) $\Omega_n, \Omega_s, \Omega_c$ que cubren el 40 % del espectro P-V cada una, ubicadas en la región polar superior e inferior y en la parte ecuatorial, respectivamente. El tamaño de la ventana es de 251×81 píxeles.

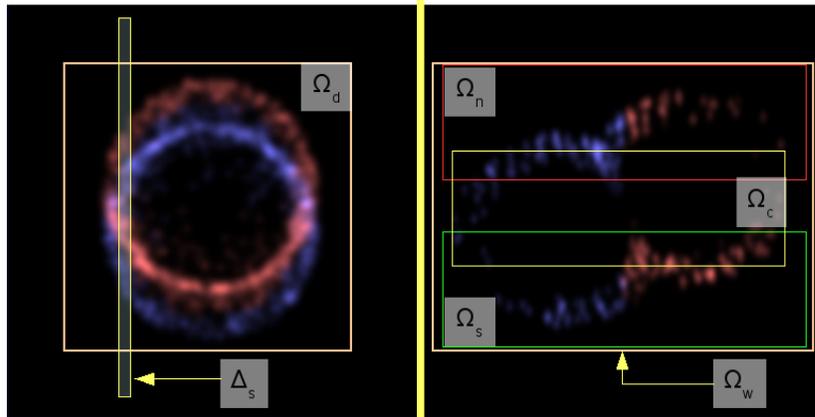


Figura 27: Regiones de interés $\{\Omega_d, \Omega_w, \Omega_n, \Omega_s, \Omega_c\}$. Las regiones sobre el espectro P-V se muestran de diferente tamaño en el eje v para facilitar su ubicación.

Ω_d nos proporciona una primera aproximación de la morfología general de la NP, sin embargo no puede discernir entre dos clases distintas que proyectadas sobre el plano del observador parecen similares, por ejemplo una nebulosa esférica tendrá la misma forma que una NP elíptica si esta última se observa desde el polo. Por otro lado, los espectros P-V (Ω_w) por si mismos dependen de la posición de la rendija donde se han adquirido, entonces es posible encontrar también morfologías similares. En la Figura 28 los gráficos (a) y (b) muestran los espectros para una NP elíptica vista desde el polo y una NP cerrada con cintura, donde se puede observar que los espectros P-V son similares.

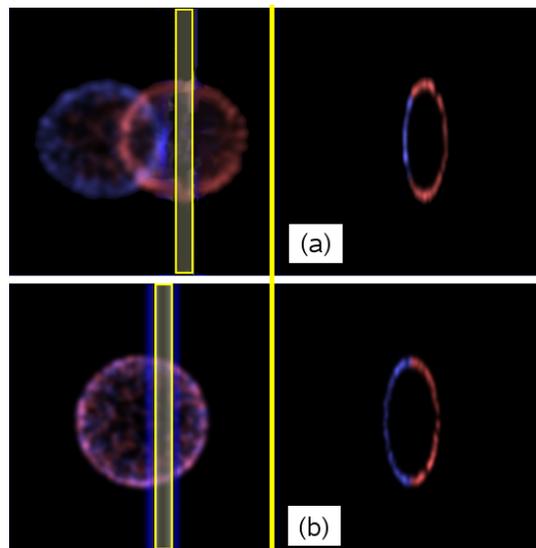


Figura 28: Espectros similares (a) y (b) pueden ocurrir en NP con morfología distinta dependiendo de la posición de la rendija.

Hemos decidido tomar 3 muestras adicionales, Ω_n , Ω_s , Ω_c , del espectro P-V debido a que el espectro presenta poca información de textura con respecto a la imagen directa y existe sólo una muestra en cada posición de la rendija y para cada orientación de las distintas clases de NP. Adicionalmente, estas regiones inspeccionan por separado la simetría observada en la NP, la cual presenta generalmente cierta simetría axial en su cuerpo principal (Manchado, 2004; Stanghelline et al., 1993), cuyos efectos se observan también en el espectro P-V. Finalmente, para reforzar la textura de los espectros se agrega el patrón correspondiente al ancho de la rendija Δ_s de la imagen directa al final del espectro.

2. Se extraen las eigencaracterísticas de cada región con el fin de reducir las dimensiones del problema. Éstas representan una proyección lineal sobre un espacio no correlacionado de las imágenes directas y de los espectros P-V. El método que se emplea, es el análisis de componentes principales que se discute más a detalle en la siguiente sección.
3. Se clasifican las 11 clases sobre las 5 regiones independientemente, mediante la búsqueda de una función de decisión que relacione las eigencaracterísticas con su respectiva clase. Para ellos se utiliza una máquina de vector de soporte, ver sección III.5.2. Una

vez que las 5 regiones están clasificadas se lleva a cabo un proceso de votación. Para cada región se calcula la probabilidad de pertenencia para cada clase. Seguido se toma la máxima y la segunda máxima probabilidad. Con estas dos opiniones estadísticas se clasifica el objeto.

III.5.1. Extracción de Eigencaracterísticas

La representación de imágenes a partir de vectores característicos es una técnica empleada en reconocimiento de patrones y procesamiento de señales. La idea principal es representar imágenes o rasgos de las imágenes en un espacio transformado, en el cual cada una de las características no están correlacionadas. Para ello se utiliza el proceso estadístico conocido como *análisis de componentes principales* (PCA), *transformada Karhunen Loève* o *transformada Hotelling*, (Hotelling, Hotelling (1933)). Esta técnica ha sido empleada ampliamente en el problema de reconocimiento de expresiones faciales (FER), (Trujillo et al., 2005; Turk y Pentland, 1991) y clasificación espectral de galaxias (Folkes et al., 1999; Vanden-Berk, 2006). En ambos campos, PCA reduce la complejidad de un conjunto de datos o características mediante la construcción de un conjunto pequeño de vectores ortogonales, que aglutinan las variaciones principales de los objetos observados. El método consiste en:

Dado un conjunto de L imágenes de entrenamiento $\mathbf{\Gamma} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_L\}$, donde Γ_i es la representación vectorial de una imagen o subregión de $m \times n$ píxeles. Se calcula el vector pixel promedio mediante

$$\Psi = \frac{1}{L} \sum_{i=1}^L \Gamma_i \quad (\text{III.2})$$

Seguido se calcula la desviación sobre el promedio para cada una de las imágenes de entrenamiento $\mathbf{\Phi} = \{\Phi_1, \Phi_2, \dots, \Phi_L\}$

$$\Phi_i = \Gamma_i - \Psi \quad (\text{III.3})$$

para obtener la matriz de covarianza

$$\mathbf{C} = \frac{1}{L} \sum_{i=1}^L \Phi_i \Phi_i^T = \frac{1}{L} \mathbf{\Phi} \mathbf{\Phi}^T \quad (\text{III.4})$$

\mathbf{C} es la matriz que describe la variación sobre la media del i -ésimo pixel, para todo el conjunto de las L imágenes de entrenamiento, con respecto a los j -ésimos pixeles distintos y sobre si mismo, en el conjunto de datos analizados. Así, la matriz de covarianza tiene un tamaño de $[(n \times m) \times (n \times m)]$ y es simétrica y positiva semidefinida. PCA busca entonces los vectores característicos μ_i y valores característicos λ_i que describen las variaciones significantes de $\mathbf{\Gamma}$. Esto es, cumplen con:

$$\mathbf{C} \vec{\mu}_i = \lambda_i \vec{\mu}_i \quad (\text{III.5})$$

dado que \mathbf{C} es una matriz simétrica positiva semidefinida, entonces el espacio formado por $\vec{\mu}_i$ es ortonormal. En la literatura $\vec{\mu}_i$ es conocida como *Eigenimágenes* y el escalar λ_i determina la importancia de cada una de nuestras eigenimágenes en este espacio transformado. Nótese que para una región Γ_i de 128×128 pixeles, \mathbf{C} es de dimensión 20244×20244 , lo que equivale a calcular 20244 eigenvectores con el mismo número de elementos. Si el propósito principal de PCA es reducir la dimensión del problema, es evidente que calcular de esta forma \mathbf{C} está en oposición de dicho propósito, además de que es una operación computacionalmente costosa. Turk y Pentland (1991), proponen calcular los eigenvectores \vec{v}_i de $\mathbf{\Phi}^T \mathbf{\Phi}$ y sus correspondientes eigenvalores b_i . Así, se obtienen L eigenvectores \vec{v}_i con L elementos cada uno, donde $L \ll (n \times m)$.

Es importante hacer notar que la nueva matriz $\mathbf{\Phi}^T \mathbf{\Phi}$ representa la variación sobre la media de cada imagen de entrenamiento Γ_i consigo misma y con el universo de entrenamiento $\mathbf{\Gamma}$, no sobre los valores de los pixeles. Sin embargo es posible calcular los primeros L eigenvectores $\vec{\mu}_i$ en función de \vec{v}_i mediante

$$\vec{\mu}_i = \mathbf{\Phi} \vec{v}_i, \quad \lambda_i = \frac{1}{L} b_i \quad (\text{III.6})$$

Ordenando los eigenvectores en forma decreciente a través de λ_i (los eigenvalores representan la varianza a lo largo del nuevo eje) es posible tomar en el cálculo los $R < L$ eigenvectores más significativos, mediante

$$\frac{\sum_{i=1}^R \lambda_i}{\sum_{j=1}^L \lambda_j} > t \quad (\text{III.7})$$

en donde t es el porcentaje de reconstrucción que se desea. En general, sólo se requiere un número reducido de eigenimágenes para el análisis. Finalmente se calculan los coeficientes de proyección w_k , $k = 1, 2, \dots, R$ de todas las R eigenimágenes más significativas $\vec{\mu}_k$, sobre una imagen Γ_i

$$\omega_k = \vec{\mu}_k^T (\Gamma_i - \Psi) \quad (\text{III.8})$$

entonces el vector que representa la proyección lineal sobre la i -ésima imagen Γ_i es

$$\vec{\omega}_i = (\omega_1, \omega_2, \dots, \omega_R), \quad i = 1, 2, \dots, L \quad (\text{III.9})$$

Al vector $\vec{\omega}_i$, lo llamaremos *Eigencaracterísticas*, el cual será sujeto a clasificación a partir de SVM y nuestro algoritmo de votación. Es posible reconstruir cualquier imagen de entrenamiento Γ_{rec} utilizando las primeros R eigenimágenes con mayor varianza a partir de (III.8), de la forma

$$\Gamma_{rec} = \Psi + \sum_{i=1}^R \omega_i \vec{\mu}_i \quad (\text{III.10})$$

Vanden-Berk (2006), utiliza PCA para separar núcleos activos y quásares de la galaxia huésped en una muestra de 7664 espectros adquiridos en el telescopio Sloan Digital Sky Survey. Ellos muestran, que es suficiente 3 eigenespectros de galaxias y 5 eigenespectros de

quásares para su clasificación. Sin embargo, ellos emplean en sus cálculos 5 eigenespectros de galaxias y 10 eigenespectros de pulsares y para algunos tipos especiales de galaxias. Folkes et al. (1999), usan los dos primeros eigenespectros para clasificar 5 clases de espectros, en una muestra de 5869 espectros de galaxias adquiridos con el 2dF espectrógrafo multifibra en el Anglo-Australian Telescope.

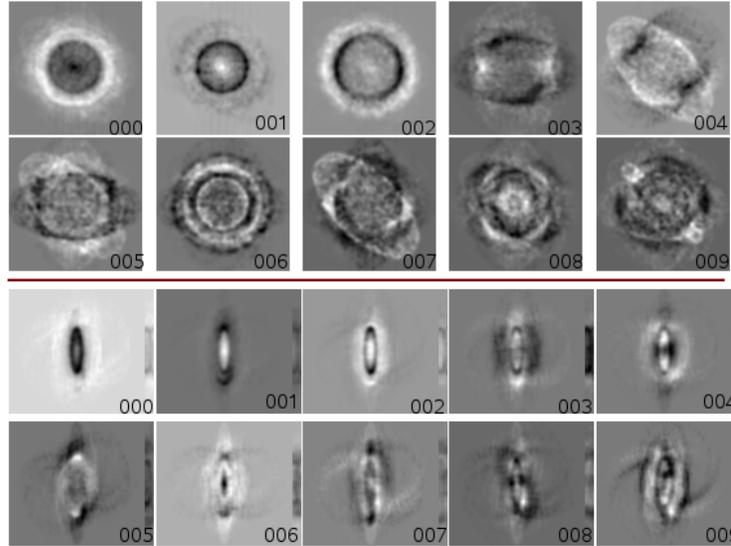


Figura 29: Los 10 gráficos superiores muestran las eigenimágenes con mayor varianza para la región Ω_d . Los 10 restantes pertenecen a Ω_w .

En nuestro problema, calculamos las eigencaracterísticas correspondientes a las regiones:

- 1) Ω_d $E_d = \{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_L\}$ que representa el patrón de la nebulosa en la imagen directa,
- 2) Ω_w $E_w = \{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_L\}$ corresponde al espectro P-V,
- 3) Ω_n $E_n = \{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_L\}$ corresponde a la sección nórdica de espectro P-V,
- 4) Ω_s $E_s = \{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_L\}$ corresponde a la sección sur de espectro P-V y
- 5) Ω_c $E_c = \{\vec{\omega}_1, \vec{\omega}_2, \dots, \vec{\omega}_L\}$ corresponde a la sección central de espectro P-V.

Encontramos que el 95 % de la varianza, $t = 0,96$ en la Ecuación (III.7), es representada por 48 eigenimágenes para Ω_d , 174 para Ω_w , 135 para Ω_n , 134 para Ω_s y 151 para Ω_c en las 432 imágenes directas y espectros de nuestra muestra de entrenamiento, en la Figura 29 se muestran las primeras 10 eigenimágenes para Ω_d y Ω_w . Es importante hacer notar que el reconocimiento y separación de las 11 clases de nuestros objetos se realiza utilizando SVM. PCA proporciona las eigencaracterísticas que alimentan a SVM. Por lo tanto,

las eigenimágenes representan el espacio transformado de todas las clases y no es posible visualizar los eigenvectores que representan una clase en particular. Obsérvese como Φ_{000} , Φ_{001} y $\Phi_{003} \in \Omega_d$ indican que existe una máxima variabilidad de las 11 clases en la región central con una radio de aproximadamente 100 píxeles y no depende de la morfología de la nebulosa planetaria. Lo mismo ocurre Φ_{000} , $\Phi_{001} \in \Omega_w$ donde la forma es una elipsoide. En oposición Vanden-Berk (2006) y Folkes et al. (1999), calculan las eigenimágenes (eigenespectros) para cada una de sus clases y la separación de clase se realiza en este punto. En nuestro método, aplicando la Ecuación (III.10), es posible reconstruir cualquier imagen Γ_i no importando a que clase pertenece.

III.5.2. Máquina de Vector de Soporte

SVM es un estadístico particular de la teoría de aprendizaje. Este método es conocido como la minimización del riesgo estructural, tal que minimiza el riesgo de una función de pérdida, en términos del riesgo empírico y de un intervalo de confianza Vapnik (1998); Schölkopf y Smola (2001). SVM es usado en una amplia gama de aplicaciones en el problema de reconocimiento de patrones, tales como: reconocimiento de texto (Joachims, 2002), reconocimiento digital de escritura con la mano (Cortes y Vapnik, 1995), reconocimiento de expresiones faciales en imágenes en el espectro visibles o infrarrojo (Hernández et al., 2007; Shih y Liu, 2006) y clasificación astronómica (Wadadekar, 2005; Willemsen et al., 2005; Huertas-Company et al., 2008).

SVM parte del principio general de que las observaciones de un espacio de rasgos o características \vec{x}_i es generado independientemente desde una función de distribución de probabilidad fija y conocida, comúnmente referida como independiente e idénticamente distribuida. Entonces, nosotros tenemos un conjunto de muestras $(\vec{x}_1, y_1), (\vec{x}_2, y_2), \dots, (\vec{x}_m, y_m)$, $\vec{x}_i \in R^m$, donde $y_i \in Y$, algunas veces llamado espacio entrada o etiquetas, relaciona la clase al miembro \vec{x}_i . Para clasificaciones binarias $Y = \{\pm 1\}$. El algoritmo intenta encontrar el mapa correcto $\vec{x}_i \rightarrow y_i$, por medio de la búsqueda del hiperplano que represente el máximo margen de sepa-

ración entre las clases (Howley y Madden, 2005). En concreto, sea ϕ un mapa no lineal en el espacio de características, tal que $\vec{x} \rightarrow \phi(\vec{x})$, entonces el hiperplano óptimo es definido como:

$$w \cdot \phi(\vec{x}) + b = 0 \quad (\text{III.11})$$

donde el vector de peso w es una combinación lineal de los vectores de soporte, b es un valor de tendencia, \vec{x}_i está sujeto a $y_i(w \cdot \phi(\vec{x}) + b) = 1$ y

$$w = \sum_{i=1}^m y_i \alpha_i \phi(\vec{x}_i) \quad (\text{III.12})$$

Las dos ecuaciones anteriores definen la función de decisión como

$$f(\vec{x}) = \text{sign} \left(\sum_{i=1}^m y_i \alpha_i \phi(\vec{x}) \cdot \phi(\vec{x}_i) + b \right) \quad (\text{III.13})$$

donde α_i es determinada maximizando la formulación dual

$$W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \phi(\vec{x}_i) \cdot \phi(\vec{x}_j) \quad (\text{III.14})$$

sujeto a

$$\sum_{i=1}^m y_i \alpha_i = 0, \quad \alpha_i \geq 0 \quad (\text{III.15})$$

La función de decisión es definida como el producto interno en el espacio de rasgos de dimensión superior vía algún mapa no lineal $\phi(\mathbf{x})$. Esto produce una frontera de decisión no lineal en el espacio de entrada. Para evitar el cálculo de la separación del hiperplano sin el transporte explícito del mapa dentro del espacio de rasgos se asigna una función núcleo $K(x_i, x_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j)$. Existen tres núcleos SVM usados comúnmente: 1) núcleo polinomial $K(x_i, x_j) = (\vec{x}_i \cdot \vec{x}_j)^d$, 2) núcleo Gaussiano (RBF) $K(x_i, x_j) = e^{\left(\frac{-\|\vec{x}_i - \vec{x}_j\|^2}{2\sigma^2}\right)}$ y 3) núcleo sigmoideal $K(x_i, x_j) = \tanh(\kappa(\vec{x}_i \cdot \vec{x}_j) - \theta)$. Para nuestros cálculos nosotros usamos el núcleo Gaussiano.

III.5.3. Comité de Clasificación SVM

Nuestro problema consiste en clasificar 5 regiones distintas, localizadas en la imagen directa y su espectro P-V. Para ello, nosotros usamos el esquema de votación basado en la probabilidad de espacio de entrada para cada una de estas regiones. Debido a que la función de decisión, ecuación (III.13), nos proporciona únicamente la etiqueta de la predicción pero no información probabilística, hemos usado el paquete LIBSVM (Chang y Lin, 2001) para obtenerla. Esta librería calcula la probabilidad $p_i = p(y = i|\vec{x})$, $i = 1, \dots, l$ de las clases $l = 11$ en el espacio de rasgos \vec{x} . Entonces, calculamos la probabilidad para $k = 5$ regiones de interés. Nuestro esquema de votación toma la contribución de la clase con la máxima probabilidad $\max(p_i)$ y el segundo mejor valor $2\max(p_i)$, acorde a lo siguiente:

Primero, se construye una matriz de decisión de la forma

$$D_{r,s} = \begin{cases} p_s & \text{Si}(p_s = \max(p)) & \forall p \in \Omega_r \text{ ó} \\ p_s & \text{Si}(p_s = 2\max(p)) \\ 0 & \text{otro} & \text{valor} \end{cases} \quad (\text{III.16})$$

para cada región de interés $r = 1, \dots, k$ y para las $s = 1 \dots l$ clases.

Segundo, se construye un vector de decisión d_s como la suma de las contribuciones probabilísticas de cada una de la regiones y clases. d_s es la suma de la probabilidad marginal $\max(p)$ y $2\max(p)$ de Ω

$$d_s = \frac{1}{k} \sum_{r=1}^k D_{rs} \quad (\text{III.17})$$

Finalmente, la clase a la que pertenece el objeto es calculada como el índice del valor máximo del vector de decisión d_s . Si dos clases distintas tienen el mismo valor en d_s , entonces se toma la de menor índice.

En nuestro esquema de votación la decisión de la clase correcta es conducida por el peso probabilístico, producto de un aprendizaje estadístico (SVM). Este comité está basado en

Tabla III.1: Matriz de decisión de una nebulosa bipolar, orientada $(20^\circ, 0, 0)$, con un ancho de rendija de $\Delta_s = 19$ centrada en el pixel $x = 82$. La etiqueta real es 1.

Región	p_{max}	y_{max}	p_{2max}	y_{2max}
Ω_d	0.90299	1	0.0161928	6
Ω_w	0.514536	1	0.374965	5
Ω_n	0.311214	5	0.301078	1
Ω_s	0.685975	1	0.22081	5
Ω_c	0.420049	5	0.372772	1

dos opiniones estadísticas, el máximo y segundo máximo valor de probabilidad, con objeto de intentar un mejor reconocimiento y clasificación. El cuadro III.1 muestra la probabilidad estimada del máximo, p_{max} y y_{max} , y el segundo máximo, p_{2max} y y_{2max} , para todas las regiones de interés (ver Figura 27) de una nebulosa bipolar de nuestro conjunto de prueba. Observe, que la región norte Ω_n y la central Ω_c del espectro P-V está clasificada erróneamente (bipolar con simetría de punto) pero el segundo máximo sugiere que existe una cierta probabilidad de que el objeto pertenezca a la clase bipolar abierta. El caso opuesto ocurre en la región sur Ω_s y en la región completa Ω_w del mismo espectro. Donde la primera opinión es correcta y la segunda apunta a la clase de Ω_n y Ω_c . Entonces el problema es calcular el peso de las contribuciones de las dos opiniones. Para ello, nosotros proponemos nuestro esquema de votación para efectos de reforzar el proceso de aprendizaje.

III.6. Experimentación y Discusión

Una vez planteado nuestro algoritmo de clasificación las pruebas se han dividido en 2 grandes secciones. La primera, se refiere al comportamiento de nuestro algoritmo sobre el catálogo sintético de prueba, en donde formamos 2 conjuntos; uno de prueba y otro de entrenamiento. En la segunda parte, se genera la función de decisión con todo nuestro catálogo completo y las imágenes de prueba son producto de observaciones de NPs reales.

III.6.1. Catálogo general de NPs

El catálogo de NPs está compuesto de 539 imágenes en el visible y sus correspondientes espectros. Cada espectro se adquirió con un ancho de rendija de 19 píxeles. La primera rendija está ubicada en el píxel central de la nebulosa, cuyo centro de la rendija corresponde al centro de la NP. A partir de este punto se tomaron adicionalmente 3 muestras excluyentes hacia la derecha y hacia la izquierda de la rendija central. Estas 7 muestras representan aproximadamente el 66 % de la nebulosa completa. En esta primera muestra se tomaron de 5 a 9 proyecciones distintas del modelo tridimensional de la NP, tomando en consideración la riqueza de las estructuras obtenidas, ver cuadro III.2. En el caso particular de la NP esférica no se considera la rotación de los ejes debido a que la eyección del material de la cáscara se considera isotrópica. El catálogo se dividió en dos conjuntos disjuntos: el primero para entrenamiento y el segundo de prueba. Así, el número de imágenes y espectros de prueba corresponde al 20 % de la base de datos, seleccionados en forma aleatoria para cada una de las clases morfológicas. Finalmente, el tamaño de las regiones de interés del experimento en píxeles es: $\Omega_d = 201 \times 201$, $\Omega_w = 251 + \Delta_s \times 201$, $\Omega_n = \Omega_s = \Omega_c = 251 + \Delta_s \times 81$ píxeles.

A partir de estos valores se calcularon las eigencaracterísticas E_d, E_w, E_n, E_s, E_c para todo el conjunto de entrenamiento; así, como la función de decisión propuesta por el clasificador SVM. De este modo utilizamos la Ecuación III.10, estimamos el número de eigencaracterísticas que representan el 96 % de la reconstrucción como sigue: $E_d = 43, E_w = 179, E_n = E_s = 138$ y $E_c = 147$. Finalmente se emplea nuestro algoritmo de votación para obtener la clase final.

Nuestro algoritmo clasifica correctamente el 98.1 % del conjunto de prueba. El comportamiento del error se presenta en la matriz de confusión; ver Tabla III.3, en donde los renglones son las clases correctas y las columnas representan el número de imágenes y espectros clasificados. En dicha tabla podemos observar que dos objetos de la clase bipolar abierta son confundidos con la clase bipolar con simetría de punto. En particular las imágenes corresponden a una rotación de ejes $(90^\circ, 0^\circ, 0^\circ)$ para una posición Δ_s centrada en los píxeles 70 y

Tabla III.2: Composición de NPDB $\Delta_s = 17$

Tipo de Morfología	Etiqueta y	N. Pos.	N. Ent.	N. Pba.	Total
Bipolar abierta	1	7	39	10	49
Bipolar abierta con cintura	2	7	40	9	49
Bipolar cerrada	3	8	45	11	56
Bipolar cerrada con cintura	4	8	45	11	56
Bipolar con simetría de punto	5	9	50	13	63
Elipse	6	9	50	13	63
Elipse con densidad ecuatorial	7	5	28	7	35
Esfera anillo fuera de la cáscara	8	9	51	12	63
Esfera anillo en la cáscara	9	9	50	13	63
Esfera	10	1	6	1	7
Simetría de Punto	11	5	28	7	35
Totales	-	-	432	107	539

184; ver Figura 30. Obsérvese como el espectro P-V de estas estructuras es muy similar. Esto se debe a que la posición de sus rendijas están a la misma distancia del centro de la imagen (57 píxeles). Por consiguiente, si una de ellas es clasificada incorrectamente, es claro que la otra imagen mantiene el mismo error. La razón de la falla se descubrió de la siguiente manera. Para el caso de la clase bipolar abierta, cada proyección contiene 7 imágenes, de las cuales 4 son usadas para el entrenamiento y 3 para la prueba, seleccionadas en forma aleatoria. Para este caso, la clase bipolar abierta muestra una simetría tanto en la imagen directa como en el espectro P-V y el conjunto de entrenamiento corresponde a las posiciones de la rendija $\Delta_s = \{-1, 0, 1, 2\}$ y las de prueba a las posiciones $\Delta_s = \{-3, -2, 3\}$, ver Figura 31. Nótese como las imágenes mal clasificadas corresponde a las posiciones $\Delta_s = \{-3, 3\}$ y por simetría corresponden al mismo patrón. Entonces podemos concluir que no existen suficientes muestras en nuestro entrenamiento para hacer una clasificación correcta. Adicionalmente, también podemos decir que nuestro algoritmo es consistente ya que clasifica de la misma manera el mismo patrón.

El comportamiento individual de cada región de interés, nos proporciona el grado de relevancia de cada una de ellas. Nos indica la habilidad de nuestro algoritmo para separar clases con escasa información. Si suponemos que tenemos únicamente la información del

Tabla III.3: Matriz de confusión

	Tipo de Morfología	1	2	3	4	5	6	7	8	9	10	11
1	Bipolar abierta	8	0	0	0	2	0	0	0	0	0	0
2	Bipolar abierta con cintura	0	9	0	0	0	0	0	0	0	0	0
3	Bipolar cerrada	0	0	11	0	0	0	0	0	0	0	0
4	Bipolar cerrada con cintura	0	0	0	11	0	0	0	0	0	0	0
5	Bipolar con simetría de punto	0	0	0	0	13	0	0	0	0	0	0
6	Elipse	0	0	0	0	0	13	0	0	0	0	0
7	Elipse con densidad ecuatorial	0	0	0	0	0	0	7	0	0	0	0
8	Esfera anillo fuera de la cáscara	0	0	0	0	0	0	0	12	0	0	0
9	Esfera anillo en la cáscara	0	0	0	0	0	0	0	0	13	0	0
10	Esfera	0	0	0	0	0	0	0	0	0	1	0
11	Simetría de Punto	0	0	0	0	0	0	0	0	0	0	7

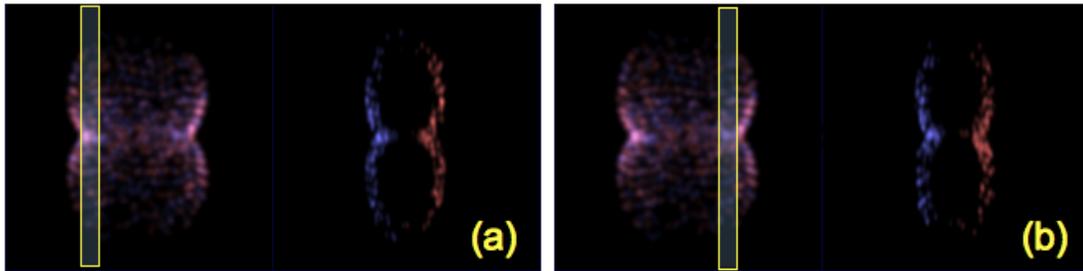


Figura 30: Nebulosas mal clasificadas.

campo de velocidades y queremos conocer su morfología; entonces, basándonos en cada región aislada, encontramos que los porcentajes de acierto son $\Omega_w = 86,7 \%$, $\Omega_n = 84,11 \%$, $\Omega_s = 81,3 \%$ y $\Omega_c = 82,242 \%$. Así, la región completa del espectro es la que presenta una mejor estadística. Esto obedece a que se tiene una mejor separación de clases a medida que se presente mas información relevante. En el caso de las otras regiones, Ω_n , Ω_s y Ω_c , la sección central es la que nos proporciona información más relevante. Por lo que, si aplicamos un esquema simple de votación al espectro P-V, en donde se cuenten el número de votos no pesados; entonces, podemos reconocer al 87,8 % de las imágenes. Por otro lado, si aplicamos nuestro esquema de votación ponderado con las dos mejores opiniones estadísticas; entonces, estamos en la posibilidad de clasificar correctamente el 89,7 % de las estructuras de prueba sin ocupar la proyección sobre el plano del cielo de la nebulosa. Finalmente, el porcentaje de

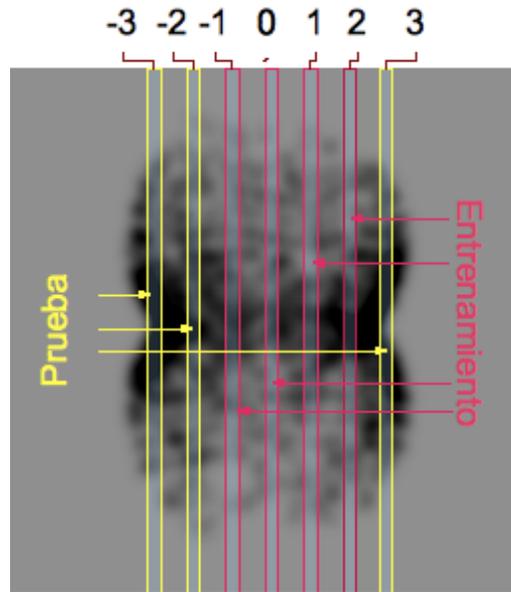


Figura 31: Ubicación de la rendija para el conjunto de entrenamiento y prueba para esta proyección.

acierto de la imagen directa es de $\Omega_d = 100\%$, debido a la forma de construcción del catálogo de prueba, en donde siempre existe una muestra idéntica a la estructura completa que la produjo, por lo tanto siempre hay un Ω_d idéntico en el conjunto de entrenamiento. Esto pone de manifiesto la importancia de seleccionar adecuadamente el conjunto de entrenamiento.

III.6.2. Objetos reales

El algoritmo de clasificación junto con nuestro esquema de votación ha mostrado gran eficiencia sobre imágenes sintéticas. Las imágenes sintéticas de nuestro catálogo está formada por una dupla de 2 imágenes; la imagen de cámara directa y la imagen del espectro. Para probar el algoritmo en un escenario real, se tiene un conjunto de imágenes adquiridas con el espectrómetro MES-SPM. En esta sección se discutirán los algoritmos utilizados para registrar las imágenes reales y hacerlas compatibles con el formato del atlas morfo-cinemático de nebulosas planetarias, discutido anteriormente. Es importante enfatizar que el espectrómetro nos proporciona un conjunto de 3 imágenes, formado por las 2 anteriores y una tercera para ubicar la posición de la rendija; ver Figura 17. En el conjunto de imágenes adquiridas

con el MES-SPM la relación de escala y orientación entre la imagen directa y su espectro no necesariamente son los mismos, ver sección III.3.1. En general se requiere de un proceso de eliminación de ruido, de un proceso de ubicación de la rendija y de transformaciones de similitud entre el espectro y la imagen directa; así como, de un proceso de escalamiento en el rango dinámico de la imagen para obtener la dupla *imagen directa* \longleftrightarrow *espectro P-V*, lo cual explicaremos a continuación.

III.6.2.1. Histograma en imágenes astronómicas

Una imagen astronómica producida por un telescopio es captada por un arreglo de celdas fotosensibles, CCD, en donde cada celda “cuenta” el número de fotones que inciden en él. Una imagen generalmente lleva un tiempo de integración de varios minutos y en ocasiones puede ser de hasta varias horas, dependiendo de lo lejano y débil del objeto o del grado de profundidad que se quiere observar. Es así, que el concepto general de colores y tonos de grises no es aplicable a este tipo de imágenes, entonces se habla de *cuentas* (número de fotones) o valor de la celda. Generalmente una imagen puede tener de 8 a 16 bits que equivalen a un rango dinámico de 256 a 64000 cuentas. Otro aspecto diferente de este tipo de imágenes es que la información de las estructuras en estudio se encuentran más o menos distinguibles en un rango acotado de valores al observarse su histograma. La Figura 32, muestra el histograma y la imagen de una NP típica, en donde el eje horizontal representa a las cuentas y el eje vertical su frecuencia en una escala logarítmica. Nótese que el primer máximo se encuentra en valores cercanos a cero, lo que representa la estructura del fondo del cielo; vea la sección (1) de la Figura 32. Después del máximo, descendiendo por la curva de distribución se encuentra la región nebulosa (Berry y Burnell, 2001) que corresponde a nuestra región de interés, ver sección (2) de la misma figura. En particular, cuando una nebulosa está bien definida se observa un segundo máximo, el cual corresponde normalmente al disco nebuloso formado alrededor de la estrella. Aproximadamente el 10 % de las cuentas corresponden a la NP. Después de la sección de la nebulosa, se encuentra a las estrellas, que se muestran como una distribución uniforme discontinua, ver sección (3). Finalmente, se encuentra la región de

saturación del instrumento, ubicado en la vecindad de la cuenta máxima.

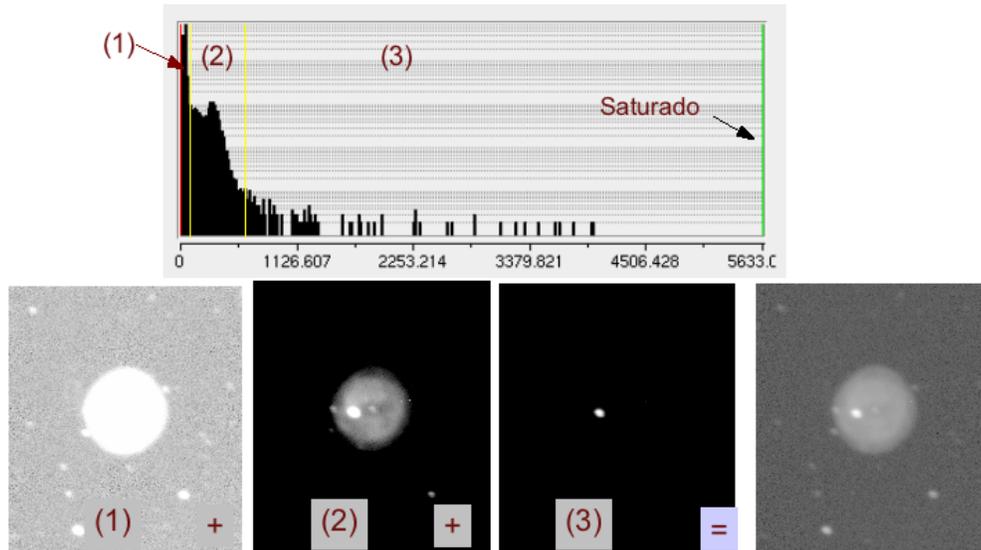


Figura 32: Ubicación de estructuras en el histograma de una nebulosa planetaria. Los gráficos inferiores muestran a la imagen que corresponde a cada estructura así como su composición final. Las estructuras son: (1) fondo del cielo, (2) nebulosa, (3) estrellas. Objeto He1-5, Hor-tensia Riesgo, *et al.* SPM 2001.

A partir de la inspección del histograma de la NP y al hecho de que es posible identificar cual es el rango de valores de pixeles donde se encuentra la nebulosa, se ha desarrollado un algoritmo que elimina el ruido de fondo de la imagen, el cielo, y los pixeles con un valor de saturación. El algoritmo funciona como sigue: sea $H(p) = \{p_1 = \min(I), \dots, p_i = \max(I)\}$ el vector de la frecuencia de ocurrencia del valor del pixel p en una imagen I . Seguido se obtiene el valor del pixel máximo del histograma $p_{max} = \max(H)$. Por lo que se calcula la frecuencia acumulada $H_a(p_1) = \frac{\sum_{i=\min(I)}^{p_1} H(i)}{n \times r}$. Entonces, se descartan los pixeles cuya frecuencia acumulada sea inferior al 90 %, $H_a(p_1) < 0,90$. Asimismo, se descartan los pixeles alrededor de la región de saturación. Es así, que la imagen resultante, sección (2) de la Figura 32, contiene $\sim 10\%$ de la información original y corresponde a la nebulosa planetaria sin el ruido de fondo.

III.6.2.2. Generación de la imagen de prueba

Una vez que se ha eliminado el cielo de fondo en la NP se requiere que se forme una dupla de imágenes reales en el mismo formato que los objetos dentro de nuestro catálogo,

v.g. como el mostrado en la Figura 27. Es importante enfatizar que el espectrómetro genera 3 imágenes, ver Figura 17, que presentan los problemas siguientes: 1) la imagen de posición del espectro es una imagen poco profunda, por lo que la NP es poco distinguible, ver Figura 17.b). 2) la imagen directa tiene una escala y puede tener una orientación distinta a la imagen de posición de la rendija, ver Figura 17.a). Finalmente, 3) estas imágenes tienen rangos dinámicos distintos a los del catálogo, esto es, las imágenes reales pueden llegar hasta un máximo de 16 bits, mientras que las del catálogo sólo contienen 8 bits.

En términos computacionales y debido al hecho de que el objeto se encuentra prácticamente en el infinito, el problema consiste en encontrar la transformación de similitud entre la imagen directa y la de la posición de la rendija. Aplicar dicha transformación a la imagen directa y añadir el espectro a esta nueva imagen transformada, ver Figura 33. El algoritmo es el siguiente. Sea I_D , I_P e I_S , una imagen directa, una de posición de la rendija y otra del espectro respectivamente, después de que se ha eliminado el fondo del cielo se procede a lo siguiente:

- Se calcula el centro de masa (\bar{x}, \bar{y}) de al menos 2 regiones de dimensión $m \times n$, empezando en el pixel (a, b) , para I_D e I_P aplicando

$$\begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \frac{1}{\sum_{x=a}^m \sum_{y=b}^n I(x,y)} \begin{pmatrix} \sum_{x=a}^m \sum_{y=b}^n x \cdot I(x,y) \\ \sum_{x=a}^m \sum_{y=b}^n y \cdot I(x,y) \end{pmatrix} \quad (\text{III.18})$$

estos puntos, se utilizan como puntos de correspondencia entre la imagen directa y su posición de la rendija, definidos por el usuario. En la Figura 33, están marcados como $P_1, P_2, P_3 \in I_D$ y $P'_1, P'_2, P'_3 \in I_P$

- Después, se calcula la transformación de similitud de I_D y I_P con al menos 2 puntos de correspondencia de la siguiente forma

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s \cos \theta & -s \sin \theta & t_x \\ s \sin \theta & -s \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (\text{III.19})$$

en donde s es el factor de escala isotrópico y $t_{x,y}$, θ es una transformación isométrica. El ángulo es calculado siempre en sentido contrario de las manecillas del reloj desde el eje y , que es la posición natural de la rendija. En la Figura 33 se utilizan las relaciones $P1 \leftrightarrow P1'$, $P2 \leftrightarrow P2'$ y $P3 \leftrightarrow P3'$ para determinar los parámetros de la transformación.

- Con los valores de la similitud se escala y se rota I_D . Nótese que la imagen directa es la que debe ser acoplada a la ventana que define la rendija. En particular, para suavizar la imagen y determinar el valor final del pixel transformado se utilizó una interpolación Bicúbica B-Spline, de la siguiente forma:

$$\begin{aligned} I'(x', y') &= \sum_{m=-1}^2 \sum_{n=-1}^2 I(x+m, y+n) \\ &\quad R(m-dx)R(n-dy) \\ R(z) &= \frac{1}{6} [P(z+2)^3 - 4P(z+1)^3 + \\ &\quad 6P(z)^3 - 4P(z-1)^3] \\ P(z) &= \begin{cases} z & z > 0 \\ 0 & z \leq 0 \end{cases} \end{aligned} \quad (\text{III.20})$$

- Finalmente se calcula el centro de masa de la nebulosa en I_D y en esa posición se une con el espectro. El usuario selecciona cual es la línea más significativa para el estudio, generalmente se utiliza la línea de $[NII]$ o H_α . En la Figura 33 se seleccionó la línea de $[NII]$ de I_S para formar la imagen que será clasificada por nuestro algoritmo.

III.6.2.3. Prueba

Nuestro algoritmo se sometió a prueba con 13 imágenes reales, que corresponden a 5 nebulosas planetarias distintas, obtenidas en el Observatorio Astronómico Nacional de San

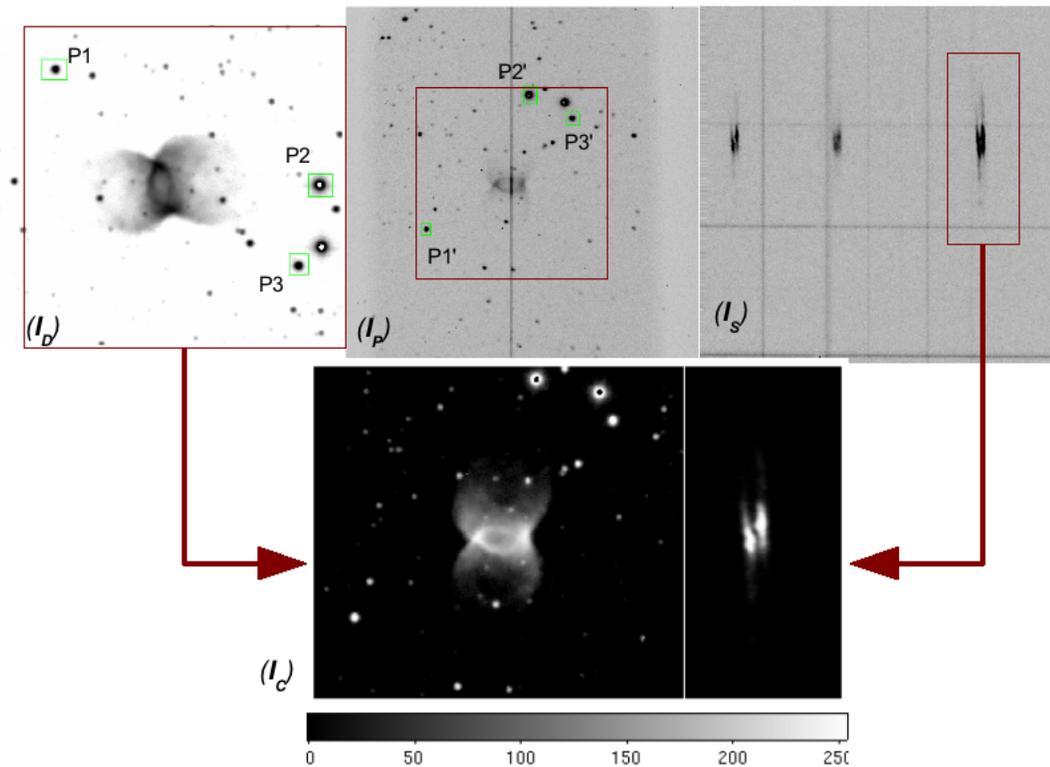


Figura 33: Generación de una imagen de prueba a partir de observaciones reales. La imagen directa I_D sufre una transformación de similitud dada por la posición de la rendija I_P . Una sección del espectro I_S es seleccionado por el usuario y es agregado a la transformación anterior para obtener una imagen en el formato de nuestro catálogo (I_C).

Pedro Mártir, utilizando el telescopio de 2 mts y el espectrómetro MES-SPM. Estas imágenes corresponden a diferentes temporadas (periodos de observación) por el equipo integrado por A. López, M. Richer y H. Riesgo. En la Figura 34 se muestran estas imágenes. Cada gráfico corresponde a 2 imágenes: los gráficos de la columna izquierda corresponden a la imagen obtenida con la cámara directa y los de la derecha a la cinemática del objeto. Cada una de las imágenes de la cámara directa están etiquetadas con el nombre de la secuencia de prueba y el nombre del objeto, ubicado en la parte inferior de estos gráficos. Los espectros cinemáticos están etiquetados en la parte superior con el nombre de línea espectral. Para nuestro caso estas líneas corresponden a la línea de Hidrógeno de la serie de Balmer H_α y a la línea de Nitrógeno [NII]. La ubicación de la rendija se muestra en el rectángulo rojo sobre la imagen directa. La base de prueba está construida como sigue:

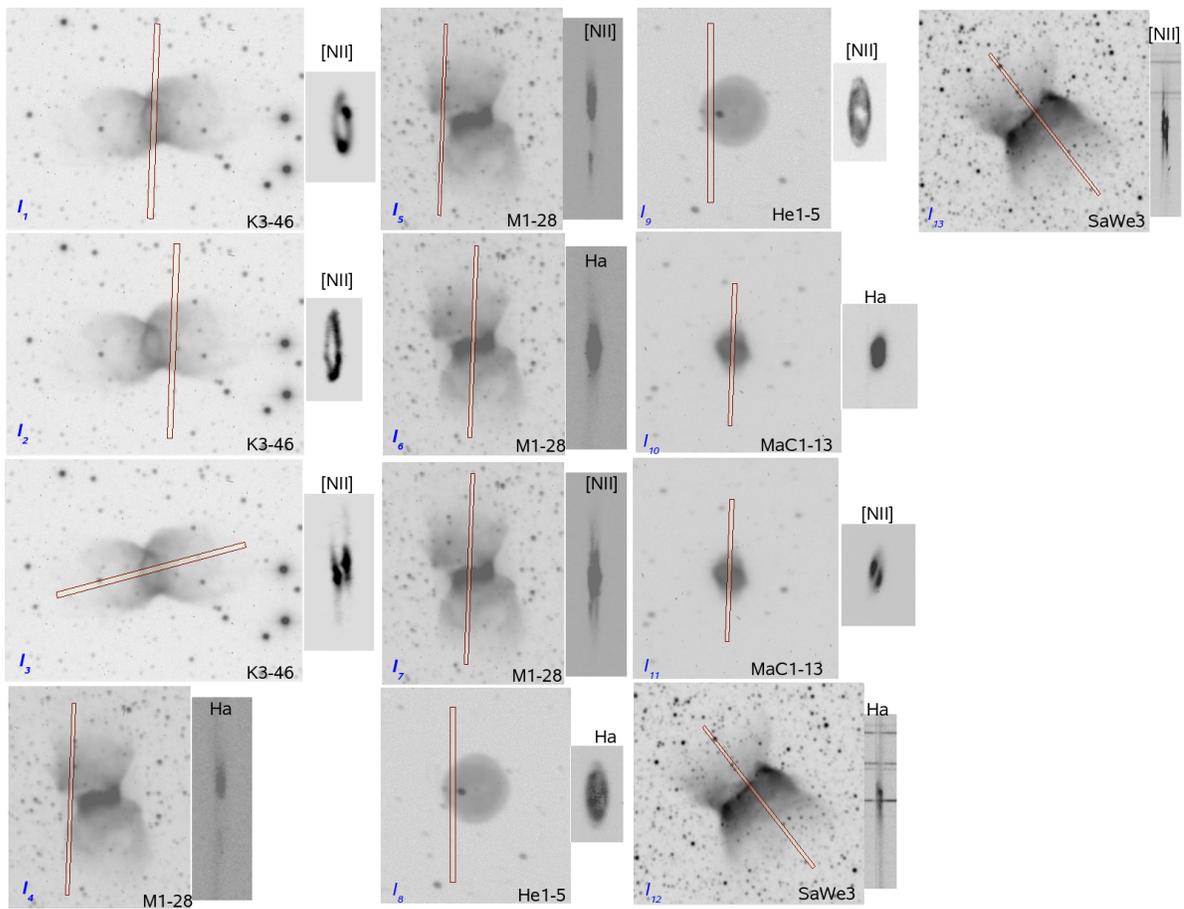


Figura 34: Imágenes de prueba obtenidas en el OAN-SPM.

- Tres imágenes de la nebulosa K3-46. Los espectros P-V corresponden a la línea de nitrógeno $[NII]$, para tres posiciones distintas de la rendija. Estas imágenes son nombradas I_1, I_2, I_3 .
- Cuatro imágenes del objeto M1-28. Las imágenes I_4, I_5 corresponden a una misma posición de la rendija pero en dos líneas espectrales distintas. Las imágenes I_6, I_7 corresponden a otro movimiento de la rendija en dos líneas espectrales distintas.
- Dos imágenes de la nebulosa He1-5 (I_8, I_9), para una posición de la rendija y dos líneas espectrales.
- Dos imágenes de la nebulosa MaC1-13 (I_{10}, I_{11}), para una posición de la rendija y dos líneas espectrales.

- Dos imágenes de la nebulosa SaWe3 (I_{12}, I_{13}), para una posición de la rendija y dos líneas espectrales.

Para cada una de estas imágenes se calculó su similitud y se ajustaron los tamaños a las características de la base sintética de prueba. Por ejemplo, para el caso del objeto K3-46, observado en el 2001 y 2004, se calculó la similitud entre la imagen directa y la imagen de posición de rendija, encontrándose que los ángulos toman los valores siguientes: $\theta_1 = -1,99^\circ$, $\theta_2 = -1,99^\circ$, $\theta_3 = -76,25^\circ$, para los factores de escala $s_1 = 2,00$, $s_2 = 2,01$, $s_3 = 2,00$ en las imágenes I_1, I_2, I_3 respectivamente.

Por otra parte, se generó la función de decisión con las 512 imágenes sintéticas de nuestro catálogo astronómico. Se utilizó un tamaño de ventana de 201×201 , con el 95 % de los eigenimágenes y eigenespectros más significativos. El número de eigenespectros para las regiones de interés esta dado por: $\Omega_d = 39$, $\Omega_w = 165$, $\Omega_n = 124$, $\Omega_s = 123$ y $\Omega_c = 133$. El tamaño de la rendija se calculó para 17 píxeles.

Bajo estas condiciones se generó la matriz de decisión mostrada en la Tabla III.4. La primera columna corresponde al nombre de imagen de prueba. La segunda columna muestra la clase final del objeto calculado a través de nuestro comité de votación SVM. El formato de esta columna es: **C : pc**, en donde **C** es la etiqueta de la clase y **pc** es la probabilidad marginal acumulada normalizada para **C** en las Ω_i regiones de interés. La siguiente tupla de columnas contiene el valor de probabilidad máxima p_m y la segunda máxima p_{2m} para la misma región. El formato es: **C : p** en donde **C** es la clase y **p** es la probabilidad de dicha clase. Las columnas restantes obedecen a esta misma lógica.

Para el caso de la nebulosa K3-46, las imágenes I_1 a I_3 , en todas las posiciones de la rendija y en su línea espectral [*NII*] el objeto es clasificado como *bipolar abierta con cintura*. Este resultado es congruente con la opinión de los expertos. Esto también ocurre para el caso del objeto M1-28 (imágenes I_4 a I_7) y el objeto SaWe3 en su línea de [*NII*] (imagen I_{13}) en las cuales nuestro clasificador las etiquetó también como *bipolar abierta con cintura*. Entre los expertos existe un consenso en dicha clasificación. El objeto He1-5, nebulosa *esférica*, fue

Tabla III.4: Matriz de decisión para las 13 imágenes de prueba.

I	Clase:pc	$\Omega_d : p_m$ C : p_m	$\Omega_d : p_{2m}$ C : p_{2m}	$\Omega_w : p_m$ C : p_m	$\Omega_w : p_{2m}$ C : p_{2m}	$\Omega_n : p_m$ C : p_m	$\Omega_n : p_{2m}$ C : p_{2m}	$\Omega_s : p_m$ C : p_m	$\Omega_s : p_{2m}$ C : p_{2m}	$\Omega_c : p_m$ C : p_m	$\Omega_c : p_{2m}$ C : p_{2m}
I_1	2:0.411	6:0.455	8:0.350	2:0.880	1:0.102	1:0.963	2:0.016	2:0.808	1:0.087	2:0.352	7:0.227
I_2	2:0.382	6:0.455	8:0.350	2:0.865	1:0.113	7:0.284	1:0.225	2:0.787	1:0.097	7:0.342	2:0.258
I_3	2:0.393	6:0.444	8:0.362	2:0.860	1:0.114	7:0.391	4:0.277	2:0.778	1:0.091	2:0.326	7:0.235
I_4	2:0.580	6:0.459	8:0.345	2:0.890	1:0.093	2:0.751	1:0.247	2:0.875	1:0.055	2:0.383	7:0.204
I_5	2:0.574	6:0.459	8:0.345	2:0.890	1:0.093	2:0.729	1:0.270	2:0.881	1:0.054	2:0.373	7:0.207
I_6	2:0.575	6:0.459	8:0.345	2:0.428	3:0.236	2:0.694	8:0.130	2:0.887	8:0.045	2:0.864	7:0.053
I_7	2:0.347	6:0.459	8:0.345	2:0.836	1:0.123	4:0.487	8:0.419	2:0.715	1:0.098	7:0.305	2:0.184
I_8	10:0.289	8:0.564	6:0.196	10:0.748	1:0.176	10:0.916	6:0.041	10:0.524	1:0.155	7:0.409	10:0.175
I_9	10:0.446	8:0.564	6:0.196	10:0.435	7:0.284	10:0.962	5:0.025	10:0.252	7:0.174	10:0.811	5:0.061
I_{10}	6:0.259	6:0.443	8:0.363	1:0.299	3:0.270	6:0.856	4:0.059	4:0.397	8:0.321	4:0.501	7:0.332
I_{11}	7:0.389	6:0.443	8:0.363	7:0.727	1:0.212	7:0.486	4:0.143	7:0.987	8:0.007	7:0.475	4:0.362
I_{12}	3:0.674	6:0.454	8:0.350	3:0.919	1:0.069	3:0.992	1:0.007	3:0.924	1:0.0461	3:0.536	7:0.121
I_{13}	2:0.290	6:0.454	8:0.350	8:0.341	2:0.316	2:0.640	8:0.125	8:0.565	9:0.244	2:0.495	9:0.212

también clasificada exitosamente, ver imágenes I_8 e I_9 . El objeto MaC1-13, en la literatura es reportado como una *ellipse*. Sin embargo, nuestro clasificador provee dos clasificaciones distintas. Para el caso del espectro en la línea H_α , ver imagen I_{10} , existe poca resolución de la estructura; más aún, el clasificador toma su decisión a través del peso probabilístico de la región de la imagen directa Ω_d y la porción norte del espectro Ω_n . En ningún otro caso, ya sea para el valor máximo de probabilidad o el segundo máximo, el clasificador repite esta clase de objeto. En la imagen I_{11} , la línea de nitrógeno ionizado $[NII]$ muestra más riqueza de estructura. Entonces, nuestro clasificador cambia su clase a una *ellipse con densidad ecuatorial*. La última imagen, objeto SaWe3 en H_α mostrado en I_{12} , el algoritmo clasifica a esta imagen como *bipolar cerrada*. Esto representa una clasificación errónea. Nótese, como el espectro de esta imagen en donde son evidentes las líneas del continuo de las estrellas que están en el campo de la rendija, barras horizontales, las cuales “contaminan” al espectro y nuestro clasificador falla. En resumen, nuestro clasificador se confunde con 2 imágenes de las 13 de prueba a que fue sometido.

Es interesante observar como nuestro clasificador funciona. Si centramos la atención exclusivamente en la etiqueta de la imagen directa, observamos la clase reportada para las NP bipolares son etiquetadas como *elipses* en su primera “opinión” y como *esferas con anillo fuera de la cáscara* en su segunda “opinión”. Esto se debe primordialmente a que no se ha podido extraer en su totalidad, a las estrellas que están en el campo de visión de la imagen di-

recta de la nebulosa, lo que provoca la confusión en esa región. Entonces, el peso específico de la clasificación recae en el análisis de las regiones espectrales del campo de posición-velocidad. Así, para el caso de I_1 la clase correcta aparece en el primer o segundo máximo de probabilidad. Esto pone de manifiesto la importancia de haber considerado distintas regiones de interés y tomar en cuenta el peso probabilístico de cada una de ellas, para tener una buena clasificación. En consecuencia, nuestro algoritmo puede alcanzar un buen éxito en la clasificación de imágenes astronómicas reales, si la sección espectral del objeto ofrece la suficiente riqueza de textura.

III.7. Conclusiones

Se ha mostrado que utilizar las eigenimágenes como un vector de rasgos y la utilización de la máquina de vector de soporte aunado a nuestro esquema de votación probabilístico tiene una gran eficiencia, 98 %, sobre el conjunto de imágenes sintéticas generadas por nuestro modelo. Asimismo, nuestro trabajo comprobó que los modelos 11 modelos sintéticos, sus proyecciones sobre el plano del cielo y el cálculo del campo de velocidades propuestas por SHAPE, son separables al menos en el mismo porcentaje de eficiencia en nuestro clasificador. Los resultados experimentales, han probado la factibilidad de nuestro algoritmo para clasificar en forma fiable imágenes reales. Del mismo modo, el catálogo de formas de nebulosas planetarias, a partir de SHAPE, puede ser de gran utilidad a la comunidad y éste contiene las variantes geométricas más significativas a lo largo de la línea de visión.

Si omitimos en nuestro análisis la imagen directa, entonces la habilidad de nuestro esquema de clasificación con la información del espectro es del $\sim 90\%$. Esto es importante ya que generalmente las confusiones en la clasificación morfológica de las NPs provienen de la inspección de la imagen directa; entonces, una correcta interpretación del espectro enriquece y disminuye el grado de incertidumbre de la interpretación. Es así, que nuestro algoritmo es sensible a la calidad del espectro cinemático al que es sometido, primordialmente en imágenes reales. Del mismo modo, debe tenerse cuidado al escoger la posición de la rendija sobre

el objeto, para efectos de una clasificación confiable.

Otra característica importante de nuestro algoritmo, es que se pueden incorporar más movimientos geométricos de los modelos, así como más posiciones de la rendija, o bien nuevas morfologías generales; sin necesidad de modificar su código. Lo que lo hace más robusto y escalable.

Finalmente, con el desarrollo de nuevos telescopios espaciales y terrestres, se espera que este tipo de algoritmos automáticos de reconocimiento, sean un instrumento esencial que ayuden a ordenar y explotar la gran cantidad de datos disponibles en un futuro cercano.

Capítulo IV

Conclusiones

Se han presentado dos trabajos de reconocimiento de objetos, que emplean imágenes digitales no visibles para el ojo humano. En el primero se abordó el problema reconocimiento de expresiones faciales, en el cual nuestro sistema de aprendizaje visual emplea imágenes térmicas. En él, se utilizaron los descriptores de textura - obtenidos a través de una estadística de segundo orden- como el vector de rasgos que puede captar la información más relevante del objeto en estudio. Estos rasgos son buscados en diferentes regiones de interés. La selección de las regiones más promisorias y de los mejores descriptores que representan el comportamiento de este problema, fue realizada en forma autónoma y en un solo paso por nuestro *algoritmo visual de aprendizaje*. El clasificador se estableció a través de un comité de votación SVM simple que decidiera a que clase pertenece dicho objeto. En el segundo trabajo, se generó un algoritmo de clasificación para reconocer y clasificar a las estructuras llamadas nebulosas planetarias. En éste, se emplean las eigenimágenes y los eigenespectros como un vector de rasgos que nos representan las principales características de estos objetos. En él, se seleccionaron cinco regiones de interés. A partir del peso probabilístico de la primera y la segunda “opinión” de un comité de votación SVM, se estableció una forma distinta para obtener la clase a la que un objeto pertenece. En ambos casos las imágenes presentan poca información de textura y los patrones morfológicos a clasificar son tenues y están “contaminados” con ruido cualitativo en extremo difícil de medir. Esto es, en el caso de la NPs las estrellas que están en el campo de adquisición del objeto contaminan de manera significativa

la estructura de la nebulosa, ver el objeto SaWe3 de la Figura 34. Es importante enfatizar que el aprendizaje fue realizado con imágenes sintéticas, en las cuales no existe ningún tipo de ruido. En el caso del reconocimiento de expresiones faciales, las imágenes de entrenamiento, están contaminadas por el hecho de que los sujetos de prueba debieran expresar en su rostro, su estado emocional con veracidad. Esto es evidente en los resultados mostrados en la Tabla II.4, en la cual la clasificación realizada por los humanos para las imágenes en el espectro visible presentan confusiones del orden del 25 %. En ambos casos las imágenes a clasificar son en extremo complejas.

Los resultados experimentales nos muestran que, para el caso de este tipo de imágenes, generar un comité de votación y encontrar un esquema de votación adecuado es de suma importancia para lograr clasificaciones confiables. En el caso de las NPs, cuando se comparan con imágenes reales las regiones que corresponden al espectro de velocidad son las que determinan la clase final del espectro. En estas regiones es en donde existe una mayor coincidencia entre la primera y la segunda opinión probabilística referente a pertenecer a una u otra clase. Para el caso del problema FER, este proceso es automático, en donde el mejor individuo genera 4 regiones de interés o regiones promisorias, en las cuales se pueden extraer los rasgos para clasificar el objeto. Es así, que consideramos que es relevante incluir heurísticas adicionales y complementarias que identifiquen de mejor manera los rasgos característicos del objeto, en imágenes no visible.

IV.1. Trabajos futuros

Una de las características principales de la computación evolutiva y de una gran cantidad de técnicas de reconocimiento de patrones, es el tiempo que invierten estos algoritmos en el aprendizaje, cuando se trata de la clasificación de los objetos. Por ejemplo, nuestro clasificador SVM, invierte 3 minutos, en una máquina Intel Core Duo 2.1 Mhz con 2 GB de memoria RAM, para evaluar a cada individuo. Si se espera que este algoritmo intente dar una primera aproximación de la morfología general de una nebulosa a un conjunto de $\sim 30,000$ NPs,

podríamos tener una respuesta después de 4 meses. Nuestro algoritmo evolutivo, que es en esencia diferente al de las NPs, tarda del orden de 15 días para su evolución utilizando un conjunto de ~ 100 imágenes. Entonces, resulta poco práctico intentar aplicar este algoritmo en su estado actual a un conjunto de prueba del orden de las nebulosas que se esperan clasificar. Es aquí, en donde se presentan retos interesantes desde el punto de vista computacional, sobre los que podríamos mencionar:

- Desarrollar nuevos algoritmos evolutivos que hagan uso de los paradigmas de cómputo paralelo. Hoy en día, el desarrollo de la masiva integración de circuitos en las unidades centrales de proceso (CPU), ha llevado a que existan más de un CPU por circuito integrado en el mercado a un bajo costo. Esto ha traído como consecuencia que la instalación de “granjas de computadoras” de bajo costo, sean un instrumento para implementar laboratorios de cómputo que permitan investigar nuevos procesos o técnicas evolutivas para resolver problemas de aprendizaje de máquina.
- Otra técnica interesante de explorar es programar algoritmos evolutivos o de reconocimiento de patrones utilizando las *Unidades de Procesamiento Gráfico (GPU¹)* altamente paralelas. Las GPU están actualmente dejando su estado inicial, que era un acelerador especializado para el cálculo de vértices y píxeles para convertirse en un procesador de punto flotante con alto rendimiento y paralelizable de aplicación general. La empresa líder de esta tecnología es la compañía NVIDIA. La NVIDIA Tesla, es una tarjeta de vídeo que contiene 240 GPUs en paralelo y actualmente es posible programarlos con el software CUDA de la misma compañía. Una computadora basada en 4 procesadores Tesla, 960 GPUs, tiene un precio de $\sim 10,000$ USD en el mercado y reporta un rendimiento de hasta 3.7 TeraFlops. Es por ésto, que resulta interesante la adaptación y la creación de nuevos algoritmos de aprendizaje que exploten esta tecnología.

¹por sus siglas en inglés *Graphics Processing Unit*

En el campo astronómico, las aplicaciones y desarrollo de nuevos algoritmos de reconocimiento de patrones es un campo amplio y fértil. En el problema específico de nebulosas planetarias podemos vislumbrar dos desarrollos inmediatos:

- El primero es el desarrollo de un algoritmo que aisle la región de la nebulosa del resto de las estrellas que aparecen en el campo de adquisición de la imagen proveniente del telescopio. Este desarrollo sería de gran utilidad para el reconocimiento automático o autónomo de la morfología general de la nebulosa y serviría para evaluar de mejor manera el desarrollo presentado en este trabajo. Asimismo, algunas investigaciones requieren del análisis de las estrellas, en donde la región nebulosa es la causante del ruido en las mediciones. En términos computacionales, se requiere de algoritmos de segmentación especializado que aislen a las NPs. Por otra parte, detectar cuáles son los factores de escala en donde se encuentran esas estructuras y poder reproducirlas lo más fielmente es un problema abierto.
- El segundo, es la detección de microestructuras que aparecen en las NPs tales como: los “Fliers” que son pares de pequeños nudos brillantes de baja excitación y los “Jets” que son estructuras lineales. La existencia de estas estructuras sugiere que el flujo podría ser colimado en vez de esférico. Estructuras de “punto simétricas” que son pares de nodos simétricos en forma de “S”, referidos como BRETs, los cuales implican flujos colimados bipolares de alta velocidad con pérdida de masa de carácter episódica. Otra microestructura son los “Flujos colimados” que producen la forma casi cilíndrica de una bipolar. Así mismo, los “Anillos” y los “Arcos” circulares y concéntricos con separación uniforme han sido observados en algunas NPs y sugieren que son proyecciones de cascarones esféricos sin perturbar el cielo.

El grupo de *Visión Evolutiva* encabezado por el Dr. Gustavo Olague ha desarrollado interesantes descriptores que podrían ser aplicables a diferentes problemas astronómicos. En particular, el descriptor evolucionado con programación genética descrito por Cinthya Pérez, que se basa en la “medida-F” y es invariante ante transformaciones afines, de iluminación

y de difuminado, *puede ser aplicable para detectar estas microestructuras en la NP*. Asimismo, los descriptores que responden en tiempo real, descritos por Leonardo Trujillo también invariante a la mayoría de las transformaciones antes descritas *podrían detectar y responder más eficientemente ante estas mismas microestructuras*. Finalmente, el desarrollo de nuevos procesos de reconocimiento de patrones, programados en base a las nuevas tecnologías emergentes de cómputo masivo, coadyuvarán en el desarrollo de frontera que se pretenda hacer con el advenimiento de los nuevos telescopios.

Bibliografía del Autor

Artículos

- G. Olague, R. Hammoud, L. Trujillo, B. Hernández y E. Romero. (2009). Facial Expression Recognition in Nonvisual Imagery. En *Augmented Vision Perception in Infrared, Algorithms and Applied Systems*. Raid I, Hammoud (Eds.), Springer-London, Advanced Pattern Recognition Series, Capítulo 10, 213-239 pp.
- G. Olague y B. Hernández. (2008). Autonomous Model-based Corner Detection using Affine Evolutionary Algorithms. En *Genetic and Evolutionary Computation for Image Processing and Analysis*. S. Cagnoni, E. Lutton y G. Olague (Eds), Hindawi Publishing Corporation, Vol. 8, 135-155 pp.
- B. Hernández, G. Olague, R. Hammoud, L. Trujillo y E. Romero. (2007). Visual Learning of Texture Descriptors for Facial Expression Recognition in Thermal Imagery. *Computer Vision and Image Understanding*, **106**(2-3):258-269.
- G. Olague y B. Hernández. (2005). A New Accurate and Flexible Model Based Multi-corner Detector for Measurement and Recognition. *Pattern Recognition Letters*, **26**(1):27-41.

Conferencias Internacionales con Comité de Lectura

- L. Trujillo, G. Olague, R. Hammoud y B. Hernández. (2005). Automatic Feature Localization in Thermal Images for Facial Expression Recognition. En *2nd. Joint IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum*. OTCBVS 2005. En conjunción con el CVPR 2005.
- G. Olague, B. Hernández y E. (2003). Accurate L-Corner Measurement using USEF Functions and Evolutionary Algorithms. En *Applications of Evolutionary Computing*. 420-421 pp. EvoWorkshops 2003, Springer ISBN 0302-9743.

- G. Olague, B. Hernández y E. Dunn. (2003). Hybrid Evolutionary Ridge Regression Approach for High-Accurate Corner Extraction. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. June 16-22, Madison, USA.
- G. Olague y B. Hernández. (2003). Detección precisa de esquinas usando funciones FUB y Algoritmos Evolutivos. En *Segundo Congreso Español de Metaheurísticas, Algoritmos Evolutivos y Bioinspirados*. 460-467 pp. Febrero 5-7, 2003, Gijón, España.
- G. Olague y B. Hernández. (2002). Flexible Model-based Multi-corner Detector for Accurate Measurements and Recognition. En *16th International Conference on Pattern Recognition*. IEEE Computer Society Press. 578-583 pp., Vol. 2. 11-15 August. Québec, Canada.
- G. Olague y B. Hernández. (2001). Autonomous Model Based Corner Detection using Evolutionary Algorithms. En *American Society for Photogrammetry and Remote Sensing*. ASPRS Annual Conference 2001.

Conferencias Nacionales con Comité de Lectura

- B. Hernández y G. Olague. (2001). Un Nuevo Detector Sub-Pixel Paramétrico de Esquinas Múltiples. En *VI Taller Iberoamericano de Reconocimiento de Patrones TIARP 2001*. 37-49 pp. México D.F. Este artículo fue elegido dentro de los 5 mejores trabajos en aportaciones científicas y tecnológicas.

Referencias

- O. B. Aaquist y S. Kwok. 1996. Radio morphologies of planetary nebulae. *Astrophysical Journal*, 462(2):813–824.
- B. Avidi, OTCBVS. IEEE OTCBVS WS Series Bench; DOE University Research Program in Robotics under grant DOE-DE-FG02-86NE37968; DOD/TACOM/NAC/ARC Program under grant R01-1344-18; FAA/NSSA grant R01-1344-48/49; Office of Naval Research under grant No. N000143010022. <http://imaging.utk.edu>. Diciembre 2008.
- J. Bala, K. DeJong, J. Huang, H. Vafaie, y H. rry Wechsler. 1996. Using learning to facilitate the evolution of features for recognizing visual concepts. *Evolutionary Computation*, 4(3): 297–311.
- B. Balick, J. Alexander, A. R. Hajian, Y. Terzian, M. Perinotto, y P. Patriarchi. 1998. Fliers and other microstructures in planetary nebulae. iv. images of elliptical pns from the hubble space telescope. *The Astronomical Journal*, 116(1):360–371.
- R. Berry y J. Burnell. 2001. *The Handbook of Astronomical Image Processing*. Willmann-Bell Inc. 624 pp.
- J. E. Besag. 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, 36(2):192–236.
- B. Bhanu y Y. Lin. 2003. Genetic algorithm based feature selection for target detection in sar images. *Image Vision Computing*, 21(7):591–608.
- F. Bourel, C. C. Chibelushi, y A. A. Low. 2002. Robust facial expression recognition using a state-based model of spatially-localised facial dynamics. In *Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, 113-118 pp.
- F. W. Campbell y J. G. Robson. 1968. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 197:551–566.
- E. Chaisson y S. McMillan. 2001. *Astronomy Today*. Prentice Hall, forth edition. ISBN 0130935719. 576 pp.
- C.-C. Chang y C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001.

- J. M. Coggins y A. K. Jain. 1985. A spatial filtering approach to texture analysis. *Pattern Recognition Letters*, 3:195–203.
- C. Cortes y V. Vapnik. 1995. Support vector network. *Machine Learning*, 20(3):273–297.
- O. G. Cula y K. J. Dana. 2001. Compact representation of bidirectional texture functions. In *CVPR*, 1041-1047 pp.
- H. D. Curtis. 1918. The planetary nebulae. *Publ. Lick Observatory*, 13(3):57–74.
- G. Donato, M. S. Bartlett, J. C. Hager, P. aul Ekman, y T. J. Sejnowski. 1999. Classifying facial actions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10): 974–989. ISSN 0162-8828.
- S. Dubuisson, F. Davoine, y J. P. Cocquerez. 2001. Automatic facial feature extraction and facial expression recognition. In *AVBPA '01: Proceedings of the Third International Conference on Audio and Video-Based Biometric Person Authentication*, 121-126 pp, London, UK. Springer-Verlag. ISBN 3-540-42216-1.
- R. O. Duda, P. E. Hart, y D. G. Stork. 2001. *Pattern Classification*. Wiley Interscience, second edition. ISBN 0-471-05669-3. 680 pp.
- B. Fasel y J. Luetin. 2003. Automatic facial expression analysis: A survey. *Pattern Recognition*, 36(1):259–275.
- M. Faúndez-Abans, M. Ormeño, y M. Oliveira-Abans. 1996. Classification of planetary nebulae by cluster analysis and artificial neural networks. *Astronomy & Astrophysics Supplement Series*, 116:395–402.
- L. Fogel, A. Owens, y M. Walsh. 1996. *Artificial Intelligence Through Simulated Evolution*. John Wiley & Sons, Inc. 170 pp.
- S. Folkes, S. Ronen, I. Price, O. Lahav, M. Colless, S. Maddox, K. Deeley, K. Glazebrook, J. Bland-Hawthorn, R. Cannon, S. Cole, C. Collins, W. Couch, S. P. Driver, G. Dalton, G. Efstathiou, R. S. Ellis, C. S. Frenk, N. Kaiser, I. Lewis, S. Lumsden, J. Peacock, B. A. Peterson, W. Sutherland, y K. Taylor⁴. 1999. The 2df galaxy redshift survey: spectral types and luminosity functions. *Monthly Notices of the Royal Astronomical Society*, 308: 459–472.
- D. J. Frew y Q. A. Parker. 2006. Towards a new distance scale and luminosity function for nearby planetary nebulae. In *IAU Symposium 234. Planetary Nebulae in our Galaxy and Beyond*, 49-54 pp.
- M. Gen y R. Cheng. 1999. *Genetic Algorithms and Engineering Optimization*. John Wiley & Son, Inc. 512 pp.

- G. Gimel'farb, L. V. Gool, y A. Zalesny. 2003. To frame or not to frame in probabilistic texture modelling? In *17th International Conference on Pattern Recognition (ICPR'04)*, volume II, 691-692 pp.
- D. E. Goldberg. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional. ISBN 0201157675. 432 pp.
- C. Goutte. 1997. Note on free lunches and cross-validation. *Neural Computation*, 9(6): 1245–1249.
- R. Gupta, H. P. Singh, K. Volk, y S. Kwok. 2004. Automated classification of 2000 bright iras sources. *The Astrophysical Journal Supplement Series*, 152:201–209.
- J. Hakkila, D. J. Haglin, G.Ñ. Pendleton, R. S. Mallozzi, C. A. Meegan, y R. J. Roiger. 2000. Gamma-ray burst classification. *The Astrophysical Journal*, 538(1):165–180.
- R. M. Haralick. 1979. Statistical and structural approaches to texture. In *IEEE*, volume 67, 786-804 pp.
- R. M. Haralick, K. Shanmugam, y I. Dinstein. 1973. Textural features for image classification. *IEEE Transactions On Systems, Man, and Cybernetics*, 3(6):610–621.
- M. Hassner y J. Sklansky. 1980. The use of markov random fields as models of textures. *Computer Graphics and Image Processing*, 12:357–370.
- B. Hernández, G. Olague, R. Hammound, L. Trujillo, y E. Romero. 2007. Visual learning of texture descriptors for facial expression recognition in thermal imagery. *Computer Vision and Image Understanding*, 106(2-3):258–269.
- J. Holland. 1992. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. University of Michigan Press. 228 pp.
- H. Hotelling. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Education Psychology*, 24:417–441: 448–520.
- D. Howard, S. C. Roberts, y R. Brankin. 1999. Target detection in imagery by genetic programming. *Advances in Engineering Software*, 30(5):303–311.
- P. Howarth y S. M. Rüger. 2004. Evaluation of texture features for content-based image retrieval. In *Conference of Image and Video Retrieval*, 326-334 pp.
- T. Howley y M. G. Madden. 2005. The genetic kernel support vector machine: Description and evaluation. *Artificial Intelligence Review*, 24:379–395.
- M. Huertas-Company, D. Rouan, L. Tasca, G. Soucail, y O. L. Fèvre. 2008. A robust morphological classification of high-redshift galaxies using support vector machines on seeing limited images. *Astronomy & Astrophysics*, 478(3):971–980.

- T. Joachims. 2002. *Learning to Classify Text using Support Vector Machines: Methods, Theory and Algorithms*. Kluwer Academic Publishers / Springer. 228 pp.
- J. M. Keller, R. M. Crownovre, y R. Y. Chen. 1987. Characteristic of natural scenes related to the fractal dimension. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:621–627.
- G. S. Khromov y L. Kohoutek. 1968. Morphological study of planetary nebulae. In *IAU Symposium 34. Planetary Nebulae*, 227-235 pp. de. D. E. Osterbrock & C. R. O’Dell (Dordrecht: Reidel).
- J. Kjell. 1994. Comparative study of noise-tolerant texture classification. *IEEE International Conference on Systems, Man, and Cybernetics. 'Humans, Information and Technology'*, 3:2431–2436.
- S. G. Kong, J. Heo, B. R. Abidi, J. Paik, y M.ñgi A. Abidi. 2005. Recent advances in visual and infrared face recognition: a review. *Computer Vision and Image Understanding*, 97(1):103–135. ISSN 1077-3142.
- J. R. Koza. 1992. *Genetic Programming*. MIT Press, Cambridge, MA. ISBN 0262111705. 840 pp.
- S. Kwok. 2001. *Cosmic Butterflies: THE COLORFUL MYSTERIES OF PLANET*. Cambridge University Press. ISBN 0521791359. 190 pp.
- S. Kwok, B. J. Hrivnak, y K. Y. L. Su. 2000. Discovery of a disk-collimated bipolar outflow in the proto-planetary nebula iras 17106-3046. *The Astrophysical Journal*, 544(2):149–152.
- T. K. Leung y J. Malik. 2001. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43(1): 29–44.
- Y. li Tian, T. Kanade, y J. F. Cohn. 2001. Recognizing action units for facial expression analysis. *IEEE Transaction Pattern Analysis and Machine Intelligence*, 23(2):97–115. ISSN 0162-8828.
- M. J. Lyons, J. Budynek, y S. Akamatsu. 1999. Automatic classification of single facial images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 21(12):1357–1362. ISSN 0162-8828.
- J. A. López. 2000. Collimated outflows in planetary nebulae. *Revista Mexicana de Astronomía y Astrofísica*, 9:201–209.
- M. Magnor, G. Kindlmann, C. Hansen, y N. Duric. 2005. Reconstruction and visualization of planetary nebulae. *IEEE Transactions on Visualization and Computer Graphics*, 11(5): 485–496.

- J. Malik, S. Belongie, J. Shi, y T. K. Leung. 1999. Textons, contours and regions: Cue integration in image segmentation. In *International Conference on Computer Vision*, 918-925 pp.
- A. Manchado. 2004. Correlation of pn morphologies and nebular parameters. In *Astronomical Society of the Pacific Conference Series*, volume 313, 3-9 pp.
- A. Manchado, M. A. Guerrero, L. Stanghellini, y M. Serra-Ricart. 1996. *The IAC Morphological Catalog of Northern Galactic Planetary Nebulae*. Instituto de Astrofísica de Canarias, Tenerife, España. 150 pp.
- C. R. Masson. 1990. On the structure of ionization-bounded planetary nebulae. *Astrophysical Journal*, 384:580-587.
- J. Meaburn, J. A., L. Guitérrez, F. Quiróz, J. M. Murillo, J. Valdéz, y M. Pedrayes. 2003. The manchester echelle spectrometer at the san pedro mártir observatory (mes-spm). *Revista Mexicana de Astronomía y Astrofísica*, 39:185-195.
- G. Mellema. 1995. Radiation gasdynamics of planetary nebulae - vi. the evolution of aspherical planetary nebulae. *Monthly Notices of the Royal Astronomical Society*, 277:173-192.
- Z. Michalewicz. 1998. *Genetic Algorithm + Data Structure = Evolution Programs*. Springer-Verlag, New York. 387 pp.
- B. Miszalski, Q. A. Parker, A. Acker, J. L. Bieky, D. J. Frew, y A. Kovacevic. 2008. Mash-ii: more planetary nebulae from the aao/ukst *ha* survey. *Monthly Notices of the Royal Astronomical Society*, 384:525-534.
- M. Moe y O. De-Marco. 2006. Do most planetary nebulae derive from binaries? i. population synthesis model of the galactic planetary nebula population produced by single stars and binaries. *The Astrophysical Journal*, 650(2):919-932.
- H. Monteiro, C. Morisset, R. Gruenwald, y S. M. Viegas. 2000. Morphology and kinematics of planetary nebulae. ii. a diabolito model for ngc 3132. *The Astrophysical Journal*, 537(2): 853-860.
- C. Morisset, H. Monteiro, y S. M. Viegas. 2000. Morphology and kinematics of planetary nebulae. i. a new modeling tool. *The Astrophysical Journal*, 533:931-937.
- A. Ñaim, O. Lahav, L. Sodré, y M. C. Storrie-Lombardi. 1995. Automated morphological classification of apm galaxies by supervised artificial neural networks. *Monthly Notices of the Royal Astronomical Society*, 275:567-590.
- S. C. Odewahan. 1995. Automated classification of astronomical images. *Astronomical Society of the Pacific*, 107:770-775.
- P. P. Ohanian y R. C. Dubes. 1992. Performance evaluation for four classes of textural features. *Pattern Recognition*, 25(8):819-833.

- T. Ojala, M. Pietikainen, y D. Harwood. 1996. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition*, 29(1):51–59.
- G. Olague. Desarrollo de un filtro autónomo, empleando computación evolutiva. Master's thesis, Instituto Tecnológico de Chihuahua, México, 1994.
- G. Olague y B. Hernández. 2005. A new accurate and flexible model-based multi-corner detector for measurement and recognition. *Pattern Recognition Letters*, 26(1):27–41.
- C. Padgett y G. Cottrell. 1997. Representing face images for emotion classification. *Advances in Neural Information Processing Systems*, 9.
- M. Pantic y L. J. M. Rothkrantz. 2000a. An expert system for recognition of facial actions and their intensity. *Image and Vision Computing*, 18:881–905.
- M. Pantic y L. J. M. Rothkrantz. 2000b. Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1424–1445.
- I. Pavlidis, J. Levine, y P. Baukol. 2001. Thermal image analysis for anxiety detection. In *International Conference on Image Processing*, volume 2, 315-318 pp.
- M. Peimbert. 1990. Planetary nebulae. *Reports on Progress in Physics*, 53(12):1559–1619.
- A. P. Pentland. 1984. Fractal-base description of natural scenes. *PAMI*, 6(6):661–674.
- N. S. Philip, A. Wadadekar, A. Kembhavi, y K. B. Joseph. 2002. A difference boosting neural network for automated star-galaxy classification. *Astronomy & Astrophysics*, 385:1119–1126.
- J. Portilla y E. Simoncelli. 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. In *International Journal of Computer Vision*, volume 40, 49-70 pp.
- F. Prokoski. 2000. History, current status, and future of infrared identification. In *CVBVS '00: Proceedings of the IEEE Workshop on Computer Vision Beyond the Visible Spectrum: Methods and Applications*, 5 p, Washington, DC, USA. IEEE Computer Society. ISBN 0-7695-0640-2.
- I. Rechenberg. 1973. *Evolutiostrategie: Optimieruung Technischer Systeme nach Prinzipien der Biologischen Evolution*. Frommann-Holzboog, Stuttgart, Germany. 176 pp.
- T. R. Reed y H. Wechsler. 1990. Segmentation of textured images and gestalt organization using spatial/spatial-frequency representation. *IEEE transaction on Pattern Analysis and Machine Intelligence*, 12:1–21.
- W. Richards y A. Polit. 1974. Texture matching. *Kybernetcs*, 16:155–162.

- F. Sabbadin, E. Cappellaro, S. Benetti, M. Turatto, y C. Zanin. 2000. Tomography of the low excitation planetary nebula ngc 40. *Astronomy & Astrophysics*, 355:688–698.
- B. Schiele y A. Pentland. 1999. Probabilistic object recognition and localization. In *International Conference on Computer Vision*, 177-182 pp.
- C. Schmid, R. Mohr, y C. Bauckhage. 2000. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172.
- B. Schölkopf y A. J. Smola. 2001. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge MIT Press. 644 pp.
- H. P. Schwefel. 1995. *Evolution and Optimum Seeking*. John Wiley & Sons, Inc. 456 pp.
- P. Shih y C. Liu. 2006. Face detection using discriminating feature analysis and support vector machine. *Pattern Recognition*, 39(2):260–276.
- P. Silapachote, D. R. Karuppiah, y A. Hanson. September 2004. Feature selection using adaboost for face expression recognition. In *Proceedings of the International Conference on Visualization, Imaging, and Image Processing*, 84-89 pp, Marbella, Spain.
- J. Sklansky. 1978. Image segmentation and feature extraction. *IEEE Transactions on System, Man and Cybernetics*, SMC-8:237–247.
- N. Soker y R. Hadar. 2002. Classification of planetary nebulae by their departure from axisymmetry. *Monthly Notices of the Royal Astronomical Society*, 331:731–735.
- L. Stanghelline, R. L. M. Corradi, y H. E. Schwarz. 1993. The correlations between planetary nebula morphology and central star evolution. *Astronomy & Astrophysics*, 279:521–528.
- W. Steffen y N. Koning, 2006-2009. SHAPE a morpho-kinematic modeling tool for astrophysical nebulae. <http://www.astrosen.unam.mx/shape>. Mayo 2009.
- W. Steffen y J. A. López. 2004. On the velocity structure in clumpy planetary nebulae. *Astrophysical Journal*, 612(1):319–331.
- W. Steffen y J. A. López. 2006. Morpho-kinematic modeling of gaseous nebulae with shape. *Revista Mexicana de Astronomía y Astrofísica*, 42:99–105.
- Y. Sugimoto, Y. Yoshitomi, y S. Tomita. 2000. A method for detecting transitions of emotional states using a thermal facial image based on a synthesis of facial expressions. *Journal of Robotics and Autonomous Systems*, 31(3):147–160.
- Z. Sun, G. Bebis, y R. Miller. 2004. Object detection using feature subset selection. *Pattern Recognition*, 37(11):2165–2176.
- H. Tamura, S. Mori, y Y. Yamawaki. 1978. Textural features corresponding to visual perception. *IEEE Transactions on System, Man and Cybernetics*, SMC-8:460–473.

- A. Teller y M. Veloso. PADO: A new learning architecture for object recognition. In K. Ikeuchi y M. Veloso, editors, *Symbolic Visual Learning*, 81–116 pp. Oxford University Press, 1996.
- L. Trujillo, G. Olague, R. Hammoud, y B. Hernández. 2005. Automatic feature localization in thermal images for facial expression recognition. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 14 p, Washington, DC, USA. IEEE Computer Society. ISBN 0-7695-2372-2-3.
- M. Turk y A. P. Pentland. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86.
- D. E. Vanden-Berk. 2006. Spectral decomposition of broad-line agns and host galaxies. *Astronomical Journal*, 131:84–99.
- V. Vapnik. 1998. *Statistical Learning Theory*. Wiley Interscience. 736 pp.
- M. Varma y A. Zisserman. 2002. Classifying images of materials: Achieving viewpoint and illumination independence. In *European Conference on Computer Vision*, volume 3, 255-271 pp.
- M. Varma y A. Zisserman. 2003. Texture classification: are filter banks necessary? In *CVPR*, volume II, 691-692 pp.
- P. Viola y M. Jones. 2004. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.
- Y. Wadadekar. 2005. Estimating photometric redshifts using support vector machines. *the Astronomical Society of the Pacific*, 117:79–85.
- J. Wilder, P. J. Phillips, C. Jiang, y S. Wiener. 1996. Comparison of visible and infrared imagery for face recognition. In *2nd International Conference on Automatic Face and Gesture Recognition (FG '96), October 14-16, 1996, Killington, Vermont, USA*, 182-191 pp.
- P. G. Willemsen, T. A. Kaempf, C. A. L. Dailer-Jones, y K. S. d.-B. . 2005. Automated identification of unresolved binaries using medium band photometry. In *Gaia Symposium: The three-dimensional Universe with Gaia*, 479-482 pp. Paris-Meudon observatory, France.
- O. C. Wilson. 1950. A survey of internal motions in the planetary nebulae. *Astrophysical Journal*, 111:279–305.
- Y. Yoshitomi, S. Kim, T. Kawano, y T. Kitazoe. 2000. Effect of sensor fusion for recognition of emotional states using voice, face image and thermal image of face. In *Proceedings of ROMAN*, 178-183 pp.
- J. Yu y B. Bhanu. 2006. Evolutionary feature synthesis for facial expression recognition. *Pattern Recognition Letters*, 27(11):1289–1298.

- C. Y. Zhang y S. Kwok. 1998. A morphological study of planetary nebulae. *The Astrophysical Journal Supplement Series*, 117:341–359.
- M. Zhang, V. B. Ciesielski, y P. Andrae. July 2003. A domain-independent window approach to multiclass object detection using genetic programming. *EURASIP Journal on Applied Signal Processing*, 2003(8):841–859. ISSN 1110-8657. Special Issue on Genetic and Evolutionary Computation for Signal Processing and Image Analysis.
- Z.-Q. Zhao, D.-S. Huang, y B.-Y. Sun. 2004. Human face recognition based on multi-features using neural networks committee. *Pattern Recognition Letters*, 25(12):1351–1358.
- S. C. Zhu, Y. Wu, y D. Mumford. 1998. Filters, random fields and maximum entropy (frame): To a unified theory for texture modeling. *International Journal of Computer Vision*, 27(2): 107–126.
- A. A. Zijlstra, J. W. Chapman, P. te Lintel Hekkert, L. Likkell, F. Comeron, R. P. Norris, F. Molster, y R. J. Cohen. 2001. Bipolar outflows from oh/ir stars. *Monthly Notices of the Royal Astronomical Society*, 322:280–308.
- S. W. Zucker y K. Kant. 1981. Multiple-level representations for texture discrimination. In *IEEE Conference on Pattern Recognition and Image Processing*, 609-614 pp.