

**IDENTIFICACIÓN DE OBJETOS EN TIEMPO REAL UTILIZANDO TÉCNICAS Y
CLASIFICADORES DE VISIÓN ARTIFICIAL PARA EL RECONOCIMIENTO DE
PATRONES**

Jaime Santiago Sánchez Reinoso

Estudiante de la carrera de Ingeniería Electrónica
Facultad de ingenierías
Universidad Politécnica Salesiana

Dirigido por:

Ing. Vladimir Robles B.

Ingeniero de Sistemas
Docente de la Universidad Politécnica Salesiana
Facultad de Ingenierías
Carrera de Ingeniería De Sistemas



**CUENCA – ECUADOR
2010 – 2011**

Breve reseña del autor e información de contacto:

Jaime Santiago Sánchez Reinoso

Estudiante de la carrera de Ingeniería Electrónica
Facultad de ingenierías
Universidad Politécnica Salesiana
jsan862@hotmail.com

Ing. Vladimir Robles B.

Ingeniero De sistemas
Docente de la Universidad Politécnica Salesiana
Facultad de ingenierías
Carrera de Ingeniería De Sistemas
mroblesl@yahoo.com.ar

Todos los derechos reservados.

Queda prohibida, salvo la excepción prevista por la ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra para fines comerciales, sin contar con la autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual. Se permite la difusión de este texto con fines académicos o investigativos por cualquier medio, con la debida notificación a los autores.

DERECHOS RESERVADOS
©2011 Universidad Politécnica Salesiana
CUENCA – ECUADOR – SUDAMERICA

Ing. Vladimir Robles B.

CERTIFICA:

Haber dirigido y revisado prolijamente cada uno de los capítulos del informe de monografía, realizada por el Sr. Jaime Santiago Sánchez Reinoso, y por cumplir los requisitos, Autorizó su presentación.

Cuenca, Octubre del 2011

Ing. Vladimir Robles B
Director

DECLARACION DE RESPONSABILIDAD

Los conceptos desarrollados, análisis realizados y las conclusiones del presente trabajo, son de exclusiva responsabilidad del autor.

Cuenca, Septiembre 2011

Jaime Santiago Sánchez Reinoso

INDICE

Dedicatoria	II
Agradecimiento	III
Índice de Figuras	IV
Índice de Tablas	V

ESQUEMA DE CONTENIDOS

INTRODUCCIÓN

CAPITULO 1

VISIÓN ARTIFICIAL

- 1.1 Introducción
- 1.2 Que es la Visión artificial
- 1.3 Aplicaciones de la visión artificial
- 1.4 Elementos de una percepción visual

CAPITULO 2

ETAPAS DE LA VISIÓN ARTIFICIAL

- 2.1 Introducción
- 2.2 Adquisición de imágenes
 - 2.2.1 Usando Arrays
 - 2.2.2 Usando Sensores
 - 2.2.3 Sistema de iluminación
 - 2.2.4 Fuente de iluminación
- 2.3 Procesamiento de imágenes
 - 2.3.1 Cuantización y muestreo
 - 2.3.2 Representación de imágenes digitales
- 2.4 Segmentación
 - 2.4.1 Segmentación en pixeles
 - 2.4.2 Operaciones morfológicas

2.4.3 Etiquetado y cuenta de objetos

CAPITULO 3

TÉCNICAS DE RECONOCIMIENTO DE PATRONES Y MODELOS DE CLASIFICADORES

3.1 Introducción

3.2 Técnicas de reconocimiento de patrones

3.2.1 Momentos de HU

3.2.2 SIFT

3.2.2.1 Operación de SIFT

3.2.3 Reconocimiento estadístico de patrones

3.2.4 Redes neuronales

3.2.4.1 Características de una red neuronal

3.2.4.2 Redes Perceptron

3.3 Modelo de Clasificadores

3.3.1 Lógica difusa

3.3.1.1 Aplicaciones de la lógica difusa

3.3.1.2 Crisp Logic

3.3.1.3 Función característica

3.3.1.4 Conjunto difuso

3.3.1.5 Función de membrecía

3.3.1.6 Grado de membrecía

3.3.1.7 Variable Lingüística

3.3.1.8 Razonamiento aproximado

3.3.1.9 Controlador Difuso

3.3.1.10 Defusificación

3.3.2 Redes bayesianas

3.3.2.1 Aprendizaje de redes bayesianas

3.3.2.1 Aplicación de redes bayesianas

3.3.3 Distancia mínima

3.3.4 K vecinos

3.3.4.1 Variantes sobre el algoritmo básico

CAPITULO 4

IMPLEMENTACIÓN DEL SISTEMA DE VISIÓN ARTIFICIAL

4.1 Introducción

4.2 Implementación de Hardware

4.2.1 Cámaras

4.2.3 Sistema de iluminación

4.3 Implementación de Software

7.3.1 Captura de imágenes

4.4 Procesamiento de la imagen

4.4.1 Umbralización

4.4.2 Etiquetado

4.4.3 Filtrado

CAPITULO 5

CALCULO MATEMÁTICO DE LAS TÉCNICAS DE RECONOCIMIENTO DE PATRONES

5.1 Introducción

5.2 Cálculo de los momentos de Hu.

5.2.1 Parámetros

5.2.2 Proceso de cálculo

5.3 Resultado

CAPITULO 6

APLICACIÓN INDUSTRIAL

6.1 Introducción

6.2 Industria metálica

6.3 Identificación de piezas metálicas

6.4 Clasificación de piezas metálicas

EXPERIMENTACION Y RESULTADOS

CONCLUSIONES

BIBLIOGRAFIA

DEDICATORIA

Este trabajo lo dedico a mis padres Que fueron el pilar fundamental Para lograr terminar el trabajo con éxito. De la misma forma a los compañeros y docentes que me brindaron su apoyo incondicional.

Jaime

AGRADECIMIENTOS

Primeramente agradezco a Dios Por haber logrado con éxito la consecución de mi tesis. Después a mis padres que Fueron las personas más Importantes que estuvieron conmigo brindándome todo su apoyo en los momentos de Dificultad. Además a todas las personas que me brindaron su ayuda para la consecución de este proyecto y de manera especial a mi director Ing. Vladimir Robles B. quien me brindo su apoyo y ayuda en el transcurso del presente trabajo.

Jaime

INDICE DE FIGURAS

Fig. 1. Pintura pre renacentista. Jesús entrando a Jerusalén	15
Fig. 2. Pintura renacentista. Iglesia del espíritu santo	15
Fig. 3. Máquina de perspectiva	16
Fig. 4. Cámara oscura	16
Fig. 5. . Ojo humano	18
Fig. 6. Representación gráfica de un ojo mirando una palmera	19
Fig. 7. Rango de las sensaciones subjetivas mostradas en un nivel particular de adaptación	20
Fig. 8. Ilusiones ópticas	20
Fig. 9. Representación gráfica de una longitud de onda	21
Fig. 10. Ejemplo de un proceso de adquisición de imágenes mediante arrays	23
Fig. 11. Sensor simple	23
Fig. 12. Combinación de un sensor simple con movimiento para generar una imagen 2D	23
Fig. 13. Iluminación Difusa	24
Fig. 14. . Iluminación Frontal	25
Fig. 15. Iluminación Direccional	25
Fig. 16. Iluminación Estructurada	25
Fig. 17. Imagen Digital Generada	27
Fig. 18. Línea explorada AB	27
Fig. 19. Muestreo y Cuantización	27
Fig. 20. Línea digital explorada	27
Fig. 21. Imagen continua proyectada por un arreglo de sensores	28
Fig. 22. Resultado de la imagen muestreada y cuantizada	28
Fig. 23. Representación de una imagen digital	29
Fig. 24. Niveles de gris para distintos bits de una imagen	30
Fig. 25. Imágenes con sus respectivos histogramas	31
Fig. 26. Ejemplo de dilatación	32
Fig. 27. Imagen Original	32
Fig. 28. Imagen erosionada	32
Fig. 29. . Etiquetado y cuenta de objetos	34
Fig. 30. Imagen en forma de U etiquetada mediante el algoritmo de etiquetado	37
Fig. 31. Localización de la detección extrema	38
Fig. 32. Proceso para la obtención de características con el método SIFT	39
Fig. 33. Esquema de una red neuronal	39
Fig. 34. Red Perceptron	40
Fig. 35. Separación de objetos en un hiperplano	41
Fig. 36. Función de membresía	44
Fig. 37. Planteamiento de un problema difuso	44
Fig. 38. Grado de Membresía	45
Fig. 39. Variable lingüística	45

Fig. 40. Sistema de un Controlador Difuso	46
Fig. 41. Proceso de defusificación de un conjunto difuso	46
Fig. 42. Función de salida de un conjunto difuso	47
Fig. 43. Método de Mamdani	47
Fig. 44. Método del valor máximo	48
Fig. 45. Modelo de una red Bayesiana	49
Fig. 46. Limite de decisión de un clasificador de distancia mínima	51
Fig. 47. Ejemplo del clasificador K-NN	51
Fig. 48. Notación para el paradigma K-NN	52
Fig. 49. Ejemplo del clasificador K-NN con distancia media	53
Fig. 50 Captura de la imagen	56
Fig. 51. Binarización de la imagen	56
Fig. 52. Pieza binarizada	57
Fig. 53. Etiquetado de la Imagen	57
Fig. 54. Adquisición de Imágenes	59
Fig. 55. Binarización de la imagen	60
Fig. 56. Etiquetado de los objetos	60
Fig. 57. Momentos de Hu de los 4 objetos a reconocer	61
Fig. 58. Objeto reconocido-moneda	64
Fig. 59. Objeto reconocido-pieza amorfa2	64
Fig. 60. Objeto reconocido-Memory Stick	65
Fig. 61. Objeto reconocido-pieza amorfa 1	65
Fig. 62. Objeto reconocido- moneda	67
Fig. 63. Objeto reconocido-pieza amorfa 2	68
Fig. 64. Objeto reconocido-Memory Stick	68
Fig. 65. Objeto reconocido-Pieza amorfa 1	68
Fig. 66. Entradas difusas del momento 1 de los objetos	69
Fig. 67. Entradas difusa del momento 2 de los objetos	69
Fig. 68. Entradas difusas del momento 3 de los objetos	69
Fig. 69. Salida difusa de la moneda	70
Fig. 70. Salida difusa de la Memory Stick	70
Fig. 71. Salida difusa de la pieza amorfa 1	70
Fig. 72. Salida difusa de la pieza amorfa 2	71
Fig. 73. Definición de reglas de inferencia	71
Fig. 74. Pieza amorfa 2 reconocida	72
Fig. 75. Memory stick reconocida	72
Fig. 76. Clasificador erróneo. Confusión entre elementos	72
Fig. 77. Definición de entradas difusas de las áreas	73
Fig. 78. Definición de salidas difusas de las áreas	73
Fig. 79. Parque industrial Cuenca	75
Fig. 80. Piezas producidas en la industria metalmecánica	76
Fig 82. Aciertos y Errores del clasificador de distancia mínima para el objeto moneda	84
Fig 83. Aciertos y Errores del clasificador de distancia mínima para el objeto memory stick	85
Fig 84. Aciertos y Errores del clasificador de distancia mínima para el objeto moneda	85

Fig 85. Aciertos y Errores del clasificador de distancia mínima para el objeto amorfa 2	86
Fig 86. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto moneda	86
Fig 87. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto memory stick	87
Fig 88. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto amorfa 1	87
Fig 89. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto amorfa 2	88
Fig 90. Aciertos y Errores del clasificador de lógica difusa para el objeto moneda	88
Fig 91. Aciertos y Errores del clasificador de lógica difusa para el objeto memory stick	89
Fig 92. Aciertos y Errores del clasificador de lógica difusa para el objeto pieza amorfa 2	89
Fig 93. Aciertos y Errores del clasificador de lógica difusa para el objeto pieza amorfa2	90

INDICE DE TABLAS

Tabla 1 Pixeles de con vecindad 4	33
Tabla 2 Pixeles con vecindad 8	33
Tabla 3 Definición de un Conjunto Difuso	43

CAPITULO 1

1.1 INTRODUCCION

Dentro de la electrónica, existe un área que está en auge y se continúa desarrollando es la visión artificial. Esta área se ocupa de simular el comportamiento del ojo del ser humano para lograr realizar proyectos que requieren un alto grado de precisión visual.

Por otra parte, la visión artificial se puede aplicar en varios campos como son el médico, electrónico, mecánico, etc. Cada uno de estos campos requiere de una percepción visual artificial con un alto grado de precisión en virtud de que si se presentan errores, esto no permitirá que el sistema implementado tome acciones correctas de acuerdo a lo que ha captado.

1.2 Que es la visión artificial

La visión artificial es aquella herramienta que nos sirve para establecer la relación que existe entre un mundo que es tridimensional y tomar sus vistas bidimensionales. Si tomamos estas vistas podemos hacer una reconstrucción de este espacio y mediante esto simular una característica tridimensional y pasarla a un plano bidimensional¹.

1.2.1 Historia de la visión artificial

Los griegos llegaron a conocer muchas propiedades geométricas de la proyección. Podemos citar el caso de Tales de Mileto, que aplicó sus conocimientos de la geometría y pudo predecir un eclipse solar. Aparte de esto también pudo deducir la altura de una pirámide con solo ver la sombra proyectada de la misma. Sin embargo los griegos tenían en su pensamiento que la visión era activa, es decir, que los ojos emitían partículas de 3D al mundo exterior y no los consideraron a estos como dispositivos pasivos receptores de luz.

Por otro lado, los pintores italianos del renacimiento fueron los que primero entendieron como se formaban las imágenes y comenzaron a estudiar la geometría que les ayudaría a reproducir correctamente los efectos de la perspectiva en las imágenes de su entorno. La época anterior de la pintura fue plana, con lo que esta no mostraba ni diferenciaba la profundidad de los objetos².

Los primeros indicios de perspectiva fueron inventados por Filippo Brunelleschi. Este fue un gran arquitecto del renacimiento temprano. Además de esta profesión fue escultor y pintor. Una de sus principales obras fue la catedral santa María de Fiore que se encuentra en Florencia, en donde se encuentra la cúpula más grande del mundo.

¹ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

²<http://biometrics.org/bc2009/presentations/wednesday/Alonso-Fernandez%20MR%2016%20Wed%20930%20AM%20-%20950%20AM.pdf>



Fig 1. Pintura pre renacentista. Jesús entrando a Jerusalén [1]



Fig 2. Pintura renacentista. Iglesia del espíritu santo[2]

Avanzando en el tiempo, llegamos al siglo XVI, en donde se desarrolla la teoría de la perspectiva. Aquí se introducen las máquinas de perspectiva para ayudar a los pintores a reproducir su perspectiva sin tener que hacer cálculos matemáticos. En la figura 2 se puede observar este tipo de máquinas, en donde el ojo del dibujante se mantenía fijo y se usa un material para interceptar cada rayo visual con el plano de la imagen.



Fig 3. Máquina de perspectiva [3]

Las máquinas de perspectiva se pueden considerar como el primer intento de construir una cámara. Estas utilizaban un plano donde se formaba la imagen y un punto central que no pertenecía al plano en donde se intersecan todos los rayos que formaban la imagen.

Entre 1596 y 1650 se empieza a concebir la geometría desde un punto de vista algebraico, por esta razón las entidades geométricas son descritas como coordenadas y entidades algebraicas. Posteriormente en el año 1826 el químico francés Niepce lleva a cabo la primera fotografía colocando una superficie fotosensible dentro de una cámara oscura para fijar la imagen. Después de este año se logra realizar el primer proceso fotográfico práctico. Aquí se utilizó una placa fotográfica que se revelaba con vapor de mercurio y se fijaba con trisulfato de sodio.

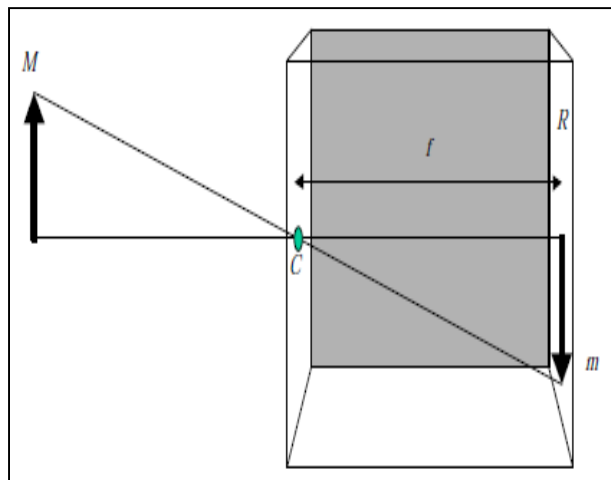


Fig 4. Cámara oscura [3]

1.3 APLICACIÓN DE LA VISION ARTIFICIAL

La visión artificial tiene un campo amplio de enfoques, ya que se puede implementar en cualquier área. Como ejemplo podemos citar las siguientes:

Médica

Procesamiento de imágenes tomadas de las radiografías de los pacientes, ya sean esto de rayos X, rayos gamma, o ultravioleta.

Fotogrametría

En donde se requiere realizar mediciones del espacio 3D a partir de fotografías tomadas de este, de tal forma que se puede medir superficies, objetos, construcciones, etc.

Electrónica

Para la construcción, diseño e implementación de componentes electrónicos en una placa electrónica. Aquí se requiere de un alto grado de precisión para que estos elementos queden ajustados si ningún grado de error.

Mecánica

Construcción de piezas con un alto grado de precisión. La percepción visual capta las imágenes de las piezas que se generan y con esto se puede ir observando que se vayan obteniendo de manera óptima las piezas.

Verificación: si es que todo el proceso que estuvo en marcha cumplió con el objetivo deseado. Por ejemplo verificar si es que las piezas se colocaron en su lugar correcto o están en la posición programada.

Detección de fallas

Involucra la detección de defectos o artefactos no deseados que tienen una forma desconocida y están en una posición no deseada. También se puede calificar a todo este proceso como un análisis cualitativo

Estimación de movimiento

Mediante una cámara que adquiere una imagen de un objeto en movimiento, se puede hacer una aproximación de este objeto a partir de ciertos puntos que tengan correspondencia con la secuencia de imágenes que se fueron adquiriendo

1.4 ELEMENTOS DE UNA PERCEPCION VISUAL

Ojo humano

Un ojo humano es cercanamente parecido a una esfera, con un diámetro promedio de alrededor de 20mm. Está constituido por tres membranas principales la cornea, la coroidea y la retina. La cornea es una capa resistente que cubre la superficie anterior del ojo. A continuación de la cornea, la esclera es una membrana opaca que se encierra y recubre el glóbulo óptico.

La coroidea se conecta directamente con la esclera. Esta membrana contiene una red de vasos sanguíneos que sirven como la mayor fuente de nutrición del ojo.

Los lentes por otro lado, son hechos de de capas concéntricas de células fibrosas y están suspendidas por fibras que unen el cuerpo ciliar. Esto contiene un 60 a 70% de agua, 6% de grasa y más proteína que otra componente en el ojo humano. Los lentes son de colores por una suave pigmentación amarillenta que se va incrementando a medida que va pasando la edad³.

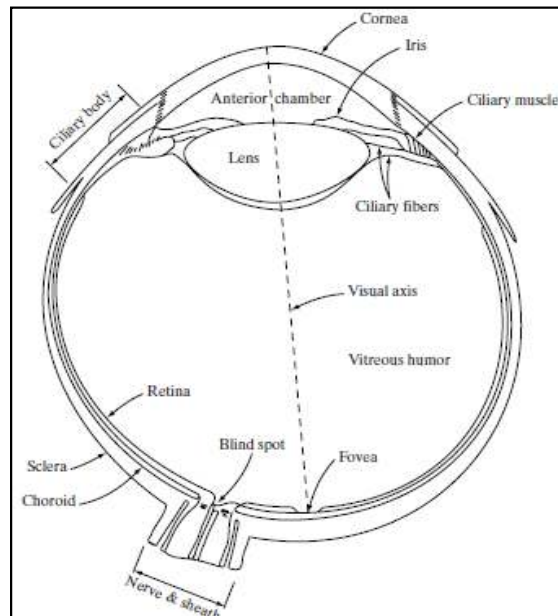


Fig 5. Ojo humano [4]

FORMACIÓN DE UNA IMAGEN EN EL OJO HUMANO

La principal diferencia entre los lentes de un ojo y unos ordinarios lentes ópticos es que la formación es flexible. La figura 4 muestra el radio de curvatura de la superficie anterior a los lentes que es mucho más grande que el radio de la superficie posterior. La forma de los lentes es controlada por la tensión en las fibras del cuerpo ciliar. Para enfocar los objetos distantes, el

³ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

control de los músculos causa que los lentes se aflojen. Similarmente, estos músculos permiten que los lentes lleguen a ser más delgados para que el enfoque de los objetos este cerca del ojo.

La distancia entre el centro de los lentes y la retina varía aproximadamente 17mm a 14 mm, como la energía refractiva de los lentes incrementa del mínimo a su máximo. Cuando el ojo enfoca un objeto lejano más o menos unos 3 metros, los lentes muestran sus niveles más bajos de energía refractiva. Por el contrario, cuando el ojo enfoca un objeto cercano, los lentes se hacen más fuertes con la energía refractiva. Esta información hace más fácil calcular el tamaño de la imagen en la retina de cualquier objeto⁴.

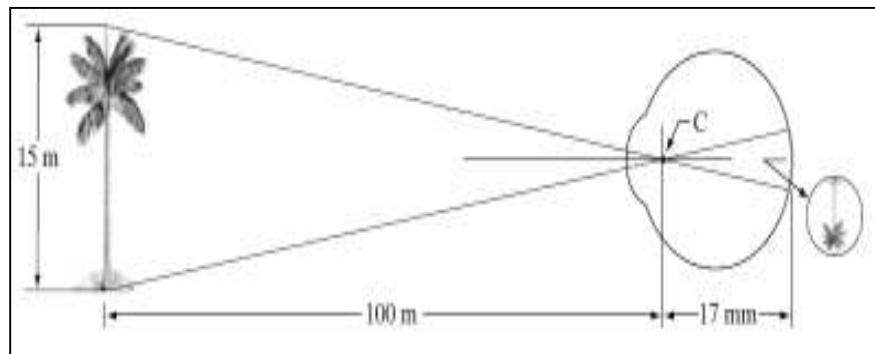


Fig 6. Representación gráfica de un ojo mirando una palmera.[3]

ADAPTACIÓN DEL BRILLO Y DISCRIMINACIÓN

Las imágenes digitales son mostradas como un conjunto de intensidades, la habilidad de los ojos para discriminar entre diferentes intensidades, es una importante consideración en la presentación del resultado de imágenes. El rango de la intensidad de la luz al cual el sistema visual humano se puede adaptar es enorme. Este rango esta en el orden de 10^{10} posibilidades. De manera experimental se indica que el brillo de la visión humana es una función logarítmica.

⁴ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

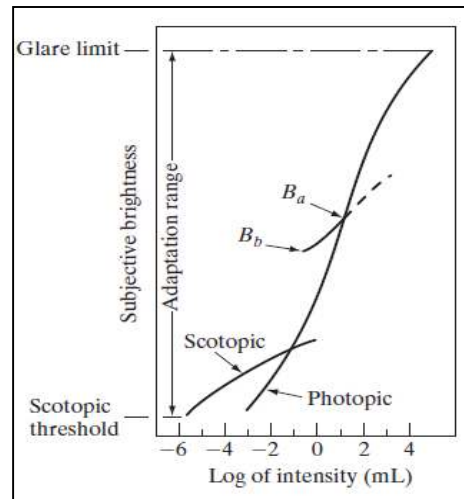


Fig 7. Rango de las sensaciones subjetivas mostradas en un nivel particular de adaptación.[4]

La curva sólida representa el rango de intensidades a las cuales el sistema visual puede adaptarse. En una visión fotópica, el rango se encuentra alrededor de 10^6 . La transición de escotópica a fotópica es gradual y esta se encuentra desde los 0.001 a los 0.1 mililamberts.

El punto esencial en interpretar el rango impresivo es que el sistema visual no puede operar sobre un rango simultáneo. Más bien, esto se enfoca a las largas variaciones por los cambios dados en la sensibilidad, un fenómeno que se conoce como adaptación al brillo.

ILUSIONES OPTICAS

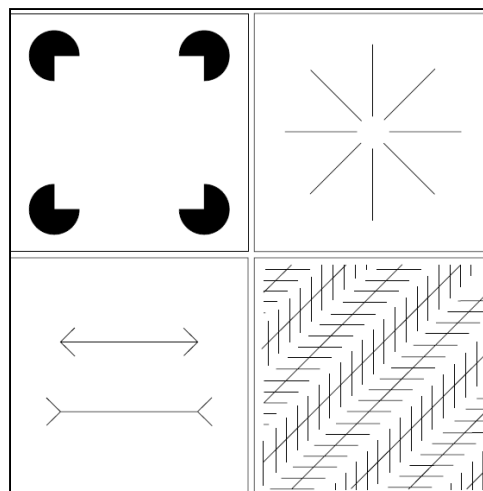


Fig. 8. Ilusiones ópticas [4]

Otros ejemplos de un fenómeno de percepción son las ilusiones ópticas en las cuales los ojos se llenan de información inexistente o erróneamente percibe propiedades geométricas de los

objetos. Un ejemplo se puede ver en la figura 8. En el primer cuadrante se pueden observar 4 círculos cortados que dan el efecto de un cuadrado. Esto mismo podemos analizar con los demás cuadrantes en donde se observa que con colocar líneas en diferente sentido se puede lograr otras percepciones visuales como es el caso de un círculo en el segundo cuadrante, las flechas más grandes en el tercero y líneas desviadas en el último.

LUZ Y ESPECTRO ELECTROMAGNETICO

En 1666, Isaac Newton descubrió que cuando se irradia un rayo de luz y este a su vez pasa por un prisma de vidrio, la irradiación emergente de luz no es blanca, en vez de esto se obtiene un espectro continuo de colores comenzando desde el violeta y terminando en rojo. En la figura 8, se puede observar que el rango de colores que nosotros percibimos representa una pequeña porción del espectro electromagnético. La finalización del espectro son ondas con una longitud que alcance millones de veces la longitud de la luz visible. Por el otro extremo del espectro están los rayos gamma con longitudes de onda millones de veces más pequeñas que de los de la luz visible. El espectro electromagnético puede ser expresado en términos de longitud de onda, frecuencia o energía. La longitud de onda λ y la frecuencia (ν) están relacionadas por la expresión:

$$\lambda = \frac{c}{\nu} \quad (1)$$

Donde c es la velocidad de la luz que es una constante que vale 2.998×10^8 m/s. La energía de varios componentes del espectro electromagnético están dados por la expresión:

$$E = h \cdot \nu \quad (2)$$

Donde h es la constante de Planck. Las unidades de la longitud de onda son metros con los prefijos micro y nano. Estos son los que se usan con frecuencia.

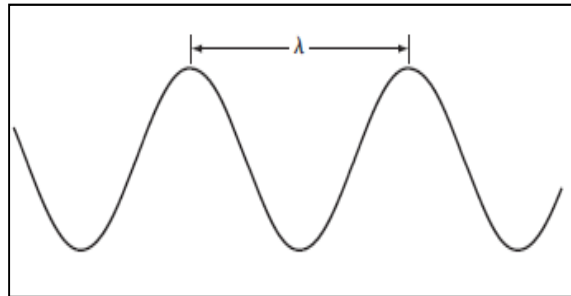


Fig. 9. Representación grafica de una longitud de onda [4]

CAPITULO 2

ETAPAS DE LA VISION ARTIFICIAL

2.1 INTRODUCCION

Dentro de un sistema de visión artificial se pueden implementar varias técnicas. Una de ellas es el procesamiento de imágenes en donde se incluyen las etapas de captura, codificación de imágenes y transformación. Otro aspecto que también se considera en un sistema de visión artificial es el reconocimiento de formas, técnica que incluye metodologías de modelado geométrico.

Por otro lado, el sistema de implementación juega un papel muy importante en el desarrollo de cualquier sistema, ya que si cualquiera de las etapas falla, el resultado no será el esperado y el proceso resultara erróneo.

2.2 ADQUISICIÓN DE IMÁGENES

Los tipos de imágenes en las cuales estamos centrados en este estudio son generados por la combinación de una fuente de iluminación y reflexión o absorción de energía de la fuente de los elementos de la escena que será capturada. Nos centraremos en la iluminación y la escena para resaltar que estos aspectos son muy importantes y por qué se deben de tomar en cuenta. Por ejemplo, la iluminación puede ser originada de una fuente de energía electromagnética tal como un radar, infrarrojo, o energía de rayos X.

De forma similar, los elementos de la escena pueden ser objetos familiares, pero estos fácilmente pueden ser moléculas, rocas sepultadas o un cerebro humano.

2.2.1 USANDO ARRAYS

Varios dispositivos electromagnéticos y ultrasónicos vienen dados en formato de arrays. Esto también es predominante en las cámaras digitales. Un tipo sensor para estas cámaras es un array CCD⁵, el cual puede ser construido con un rango amplio de propiedades de sensado y puede ser empaquetado en conjuntos de arrays de 4000X4000 elementos y hasta más⁶.

La principal manera en la cual estos sensores son usados se puede apreciar en la figura 10. Aquí se puede observar la energía de una fuente de iluminación siendo reflejada de una escena de un elemento. La primera función dada por un sistema de captura es recolectar la energía que ingresa y enfocarla en un plano donde se encuentra la imagen. Si la iluminación es luz la terminación frontal del sistema de imagen es un lente que proyecta la vista de la escena en el plano que se enfocó.

⁵ Dispositivo de carga acoplada que tienen varios capacitores capaces de transferir energía

⁶ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

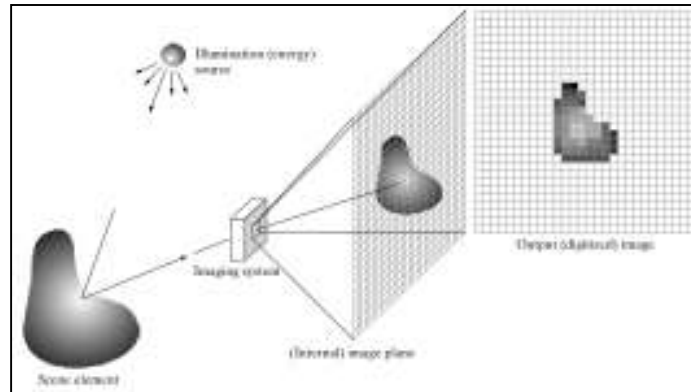


Fig. 10. Ejemplo de un proceso de adquisición de imágenes mediante arrays [4]

2.2.2 USANDO SENSORES

En la figura 10 se puede observar los componentes de un sensor simple. Quizás el tipo más familiar de este sensor es un fotodiodo, el cual es construido de un material de silicón y cuya forma de onda de voltaje de salida es proporcional a la luz. El uso de un filtro enfrente del sensor mejora la selectividad.

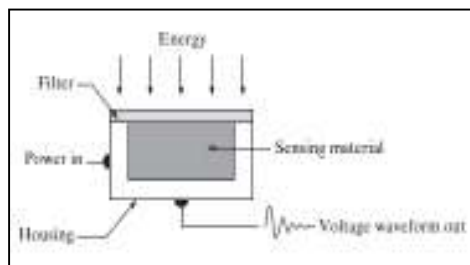


Fig. 11. Sensor simple [4]

Por otro lado, para generar una imagen en 2D usando un simple sensor, tiene que existir un desplazamiento relativo de las dos direcciones tanto x como en y entre el sensor y el área de la imagen. En la figura 11 se puede observar un arreglo usado en un escaneo de alta precisión donde la película negativa es montada sobre un tambor que tiene un mecanismo de desplazamiento que hace que esta se mueva en una dimensión.

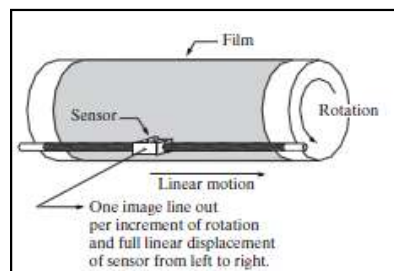


Fig 12. Combinación de un sensor simple con movimiento para generar una imagen 2D.[4]

2.2.3 SISTEMA DE ILUMINACION

La iluminación de la escena juega un papel muy importante para el desarrollo de un sistema visual. Antes de pasar mucho tiempo en implementar algoritmos complejos para arreglar este problema, es mejor prestar atención en implementar un sistema de iluminación adecuado para que cuando se capturen las imágenes sea de manera correcta. Este sistema de iluminación será adecuado si se puede corregir el problema mediante software, debido a que la velocidad de procesamiento será mayor con algoritmos más sencillos⁷.

2.2.3.1 SISTEMA DE RETROALIMENTACION DIFUSA

Este tipo de iluminación es muy adecuada para el reconocimiento o medida de una pieza solo si se necesita el contorno y es posible apoyar dicho objeto sobre una superficie transparente.

La retroalimentación difusa consiste en iluminar contra la cámara, dejando el objeto entre la cámara y la lámpara. Esta técnica entrega imágenes con un alto contraste entre el objeto y el fondo, dando como resultado una segmentación fácil ya que se puede aplicar una binarización.

Dentro de las principales aplicaciones en donde tiene un adecuado comportamiento esta técnica de iluminación son para medir el grado de porosidad de ciertas sustancias y en la inspección dimensional para calcular el tamaño de una pieza.

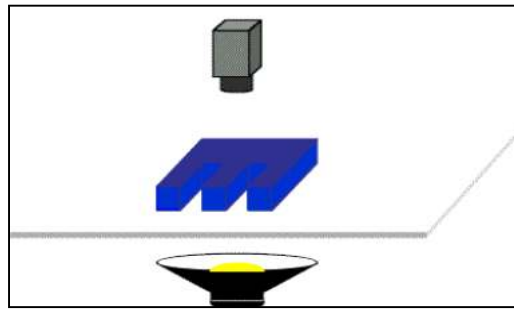


Fig 13. Iluminación Difusa [5]

2.2.3.2 ILUMINACION FRONTAL

Es la más usada. Tiene como objetivo iluminar la pieza frontalmente. La desventaja de esta técnica es que tiene problemas para obtener un buen contraste entre la pieza y el fondo, debido a que aparecen brillos y sombras que alteran las propiedades de las piezas en estudio.

Esta técnica se usa en objetos poco reflectores para evitar los brillos que molestan la captura en una iluminación difusa.

⁷ <http://tesis.pucp.edu.pe/repositorio/handle/123456789/68>

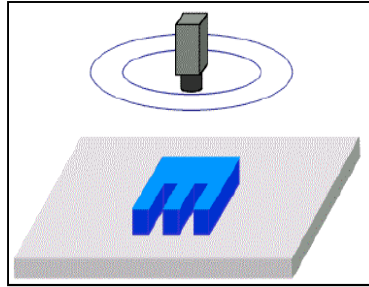


Fig 14. Iluminación Frontal [5]

2.2.3.3 LUZ DIRECCIONAL

Consiste en iluminar un objeto de manera direccionada en algún sentido del espacio para poder rescatar alguna característica del objeto. La principal característica es que se crean sombras sobre el objeto, lo que ayuda a aumentar el contraste de partes tridimensional y obtener información en 3d.

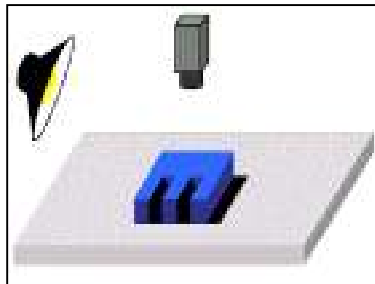


Fig 15. Iluminación Direccional [5]

2.2.3.4 LUZ ESTRUCTURADA

El objetivo de este tipo de iluminación es proyectar la pieza sobre unos patrones de luz conocidos modulados y observar la luz reflejada, que también viene modulada. De esta manera se puede obtener la información de la superficie del objeto misma que puede reconstruirse mediante triangulación.

Las fuentes de luz que se pueden emplear deben ser espaciales capaces de emitir luz estructurada que por lo general son laser.

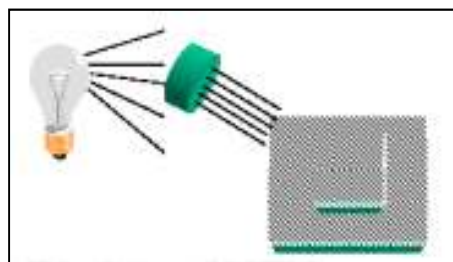


Fig 16. Iluminación Estructurada [5]

2.2.4 FUENTES DE ALIMENTACION

Entre los principales tipos de iluminación tenemos:

Lámparas incandescentes: Es la fuente de iluminación más común y está formada por un filamento de tungsteno. Su principal ventaja es que tiene una gran variedad de potencias. Por otro lado, su desventaja es que reduce la luminosidad con el tiempo, lo que ocasiona inconvenientes con el sistema de visión.

Tubos fluorescentes: son mucho mejores que las lámparas incandescentes. Estos entregan una luz más difusa que es aprovechada en objetos que son reflectores. Existe una gran variedad de este tipo de tubos.

Fibra óptica: este tipo se utiliza para iluminar zonas que realmente tienen un acceso difícil o partes que son extremadamente pequeñas.

Lasers: Estos se emplean para tener una iluminación estructurada, debido a que este es capaz de emitir luz con un control adecuado de luminosidad. La desventaja de este es que no se comporta de manera adecuada en superficies que absorben la luz.

2.3 PROCESAMIENTO DE LA IMAGEN

Para el procesamiento de la imagen, debemos considerar que una imagen en 2 dimensiones se puede representar de la siguiente forma $f(x, y)$. El valor de la amplitud de f en coordenadas espaciales (x, y) es un escalar positivo el cual tiene un significado físico que se determina de la fuente de donde se origina la imagen. Las imágenes con las que trabajaremos basan su esquema de color en una escala de grises. Cuando una imagen es generada por un proceso físico, su valor es proporcional a la energía irradiada por la fuente física, por ejemplo una onda electromagnética. Como consecuencia tenemos lo siguiente:

$$0 < f(x, y) < \infty \quad (3)$$

La función $f(x, y)$ puede ser caracterizada por 2 componentes: la primera que contiene la cantidad de una fuente de iluminación incidente en la escena que está siendo observada y la segunda que la cantidad de iluminación reflejada por los objetos que se encuentran en la escena. El nombre apropiado para definir a estas dos características son la iluminación y la reflectancia, componentes que se les denota como $i(x, y)$ y $r(x, y)$ respectivamente. Las dos funciones se combinan para obtener:

$$f(x, y) = i(x, y)r(x, y) \quad (4)$$

$$0 < i(x, y) < \infty$$

$$0 < r(x, y) < 1$$

Esta última expresión nos indica que la reflectancia es limitada a 0 cuando esta totalmente absorbida o 1 que tiene una total reflectancia. La naturaleza de $i(x, y)$ se determina por la fuente de iluminación y $r(x, y)$ se determina por la característica de los objetos fotografiados.

2.3.1 CUANTIZACION Y MUESTREO

Para tener una idea de que es el muestreo y la cuantización podemos observar las siguientes imágenes:

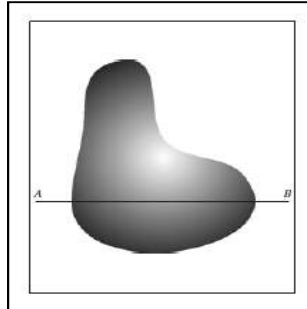


Fig. 17. Imagen Digital Generada [4]

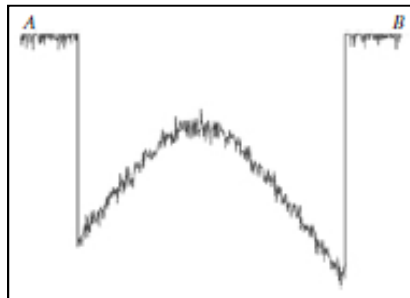


Fig. 18. Línea explorada AB [4]

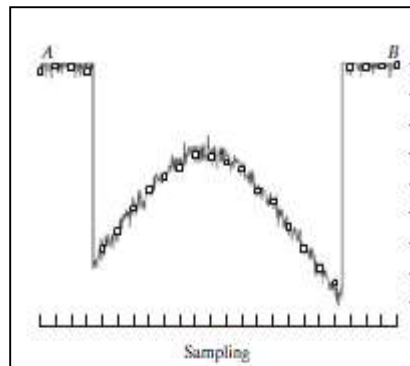


Fig. 19. Muestreo y Cuantización [4]

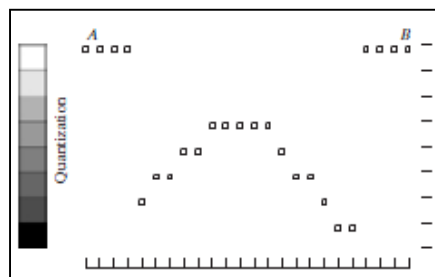


Fig. 20. Línea digital explorada [4]

En la función unidimensional de la figura 17, se grafica la amplitud de los valores continuos de la imagen a lo largo del segmento AB de la figura 18. Las variaciones aleatorias son ruidos. Para realizar el muestreo de esta función, tomamos muestras de espacios iguales a lo largo de la línea AB como se muestra en la figura 18. La localización de cada muestra está dada por una marca vertical delgada en la parte inferior de la figura. Las muestras son mostradas con cuadros blancos sobrepuestos en la función.

Por otro lado, el muestreo se lo hace asumiendo que tenemos una imagen continua en ambas direcciones como también en amplitud. En la práctica el método de muestreo se determina por un arreglo de sensores usados para generar la imagen. Cuando la imagen es generada por un sensor simple y combinada con un movimiento mecánico, la salida del sensor es cuantizada en la forma que se describió arriba.

Cuando se usa un censado de banda para adquirir imágenes, el número de sensores en la banda establece las limitaciones del muestreo. El movimiento mecánico en la otra dirección puede ser controlado más exactamente, pero esto hace que pequeñas variaciones intenten muestrear toda la densidad en esa dirección lo que hace que se exceda los límites de muestreo establecido por el número de sensores en la dirección opuesta.

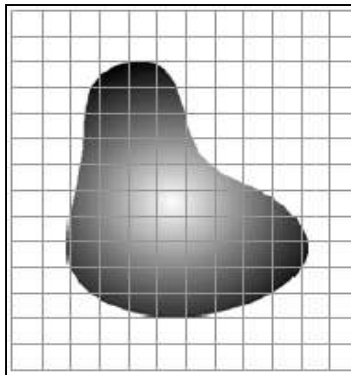


Fig. 21. Imagen continúa proyectada por un arreglo de sensores [4]

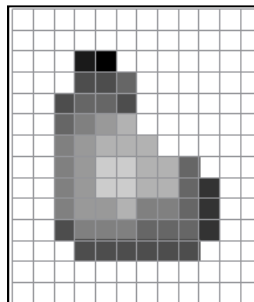


Fig. 22. Resultado de la imagen muestreada y cuantizada [4]

2.3.2 REPRESENTACION DE IMÁGENES DIGITALES

El resultado del muestreo y cuantización es una matriz de números reales. Si asumimos que una imagen $F(x, y)$ es muestreada entonces el resultado de la imagen tendrá filas y columnas. El valor de las coordenadas (x, y) serán ahora cantidades discretas. Para tener una notación clara y de conveniencia se denota con valores enteros estas coordenadas discretas. Los valores que se tiene en el origen son $(x,y)=(0,0)$. La próxima coordenada en el eje de las abscisas se representará como $(x,y)=(0,1)$. Es importante tener en cuenta que la notación $(0,1)$ se usa para representar la segunda muestra a lo largo de la primera columna⁸.

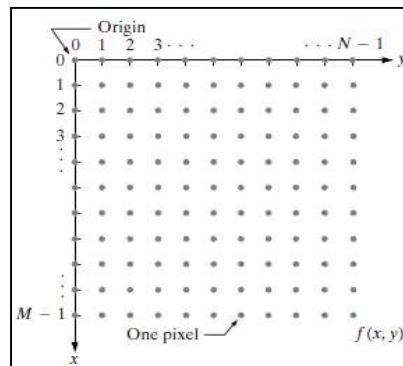


Fig. 23. Representación de una imagen digital [4]

Una vez descrita la notación de una imagen digital, podemos describirla como una matriz de la siguiente manera:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1) & \dots & f(1, N-1) \\ \dots & \dots & \dots & \dots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix} \quad (5)$$

El lado derecho de la ecuación es la definición de una imagen digital. Cada elemento de la matriz se denomina elemento imagen, elemento gráfico o pixel.

En algunos casos es mejor denotar como una matriz con sus elementos así:

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,N-1} \\ \dots & \dots & \dots & \dots \\ a_{M-1,0} & a_{M-1,1} & \dots & a_{M-1,N-1} \end{bmatrix} \quad (6)$$

Si denotamos Z y R como un conjunto de enteros reales y un conjunto de números reales respectivamente, el proceso de muestreo puede ser visto como una partición del plano xy en una malla, con las coordenadas del centro de cada malla siendo un par de elementos de un producto cartesiano Z^2 , el cual es un conjunto de pares ordenados (z_i, z_j) con z_i y z_j siendo enteros. De aquí

⁸ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

$F(x, y)$ es una imagen digital si (x, y) son enteros de Z^2 y f es una función que asigne valores de niveles de grises.

Por otro lado, el proceso de digitalización requiere de decidir valores para filas, columnas y para L que es un nivel discreto de gris para cada pixel. No existen reglas para las filas y columnas solamente que los enteros tengan valores enteros positivos.

A la escala de grises se la puede representar de la siguiente manera:

$$L=2^k$$

Asumimos que los niveles discretos están igualmente espaciados y que son enteros en un intervalo de $[0, L-1]$. Algunas veces el rango de valores que abarca la escala de grises es llamada rango dinámico de una imagen, y se refiere a aquellas imágenes cuyos niveles de grises tiene una porción significativa en la escala dinámica.

N/k	1 ($L = 2$)	2 ($L = 4$)	3 ($L = 8$)	4 ($L = 16$)	5 ($L = 32$)	6 ($L = 64$)	7 ($L = 128$)	8 ($L = 256$)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

Fig. 24. Niveles de gris para distintos bits de una imagen. [4]

2.4 SEGMENTACION

La segmentación es aquel proceso mediante el cual una imagen se descompone en regiones o elementos que pueden corresponder a objetos o parte de objetos. Este proceso es el encargado de evaluar si cada pixel de la imagen pertenece o no a dicho objeto que se está analizando. Esta técnica de procesamiento de imágenes genera una imagen binaria, donde se representa con 1 los pixeles que pertenecen a dicho objeto y 0 a los que no.

2.4.1 SEGMENTACION BASADA EN PIXELES

Este método de segmentación considera solo el valor de gris de un pixel, para decidir si el mismo es o no parte del objeto. Para determinar esto, se debe encontrar el rango de valores de gris que son característicos del objeto. Mediante un histograma se localiza estos valores característicos.

El objetivo de este método, es el de encontrar de manera óptima los valores característicos del objeto que se está analizando con respecto a las regiones que no pertenecen a este. En un primer caso si los valores de gris del objeto y el resto de la imagen difiere claramente, tendremos un histograma con una distribución bimodal con dos valores máximos distintos que no representan

los valores característicos y que idealmente fueran igual a cero, con lo que se lograría una separación entre el objeto y la imagen.

Por otro lado, si el valor del histograma está ubicado entre los dos máximos es diferente de cero, las funciones de probabilidad de valores de gris del objeto y la región restante se sobrepondrán, lo que hace que algunos píxeles del objeto deberán ser tomados en cuenta como pertenecientes a la región.

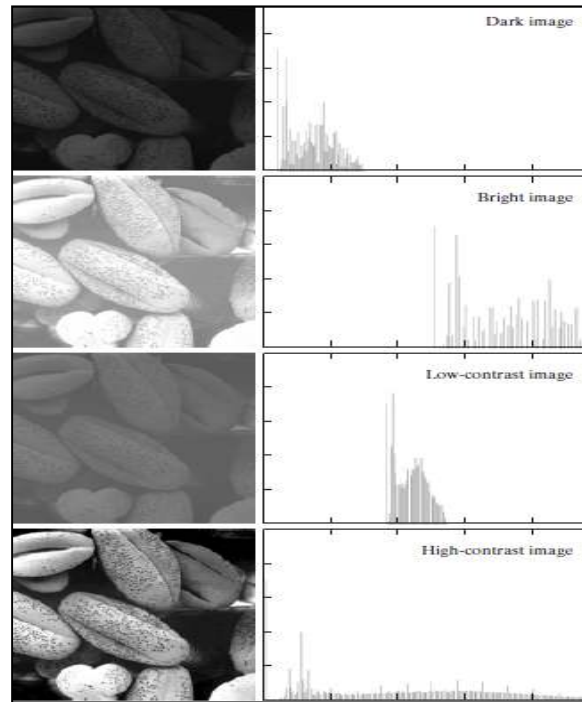


Fig. 25. Imágenes con sus respectivos histogramas [4]

2.4.2 OPERACIONES MORFOLOGICAS

Las operaciones morfológicas son métodos para procesar imágenes binarias basadas en formas. Estas operaciones toman una imagen binaria como entrada y a la salida obtiene también una imagen binaria. El valor de cada pixel en la imagen de salida se basa en cada pixel de la imagen de entrada y de sus respectivos vecinos.

Dilatación: si consideramos A y B como conjuntos de Z^2 y \emptyset representando al conjunto vacío, la dilatación de A por B se representa como:

$$A \oplus B = \{x / [(B)_x] \cap A\} \subseteq A \quad (6)$$

Por lo tanto el proceso de dilatación consiste en obtener la reflexión de B sobre su origen y después cambiarla por la reflexión en x. La dilatación de A por B es entonces, el conjunto de todos los desplazamientos x tales que B y A se sobremonten en al menos algún elemento que sea distinto de cero.

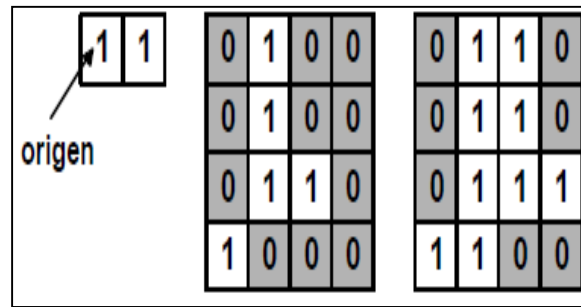


Fig. 26. Ejemplo de dilatación [6]

Erosión: Para los conjuntos de A y B de Z^2 , la erosión de A por B representada por $A \ominus B$ se define como:

$$A \ominus B = \{x / (B)_x \subseteq A\} \quad (7)$$

Esta expresión nos dice que la erosión de A por B es el conjunto de todos los puntos x tal que B trasladado por x, está contenido en A.

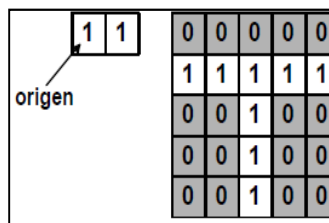


Fig. 27 Imagen Original [6]

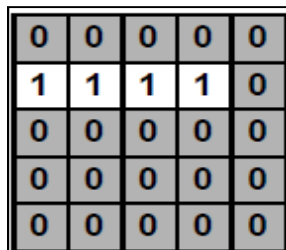


Fig. 28. Imagen erosionada [6]

2.4.3 ETIQUETADO

Muchas veces no es suficiente el proceso de binarización de una imagen, por lo que el proceso de segmentación se expande con nuevas técnicas para el análisis de imágenes y una de ellas es el

etiquetado y cuenta de objetos⁹. Para poder analizar este aspecto debemos entender conceptos importantes como:

Conectividad y vecindad: el primer inconveniente que se genera en el análisis es el de diferenciar el objeto de la imagen de fondo. Para esto se utiliza el criterio de vecindad de un pixel. Para esto tenemos dos casos:

Pixeles 4- vecinos de A0

	A3	
A5	A0	A1
	A7	

Tabla 1. Pixeles de con vecindad 4

Los cuales se representan como: $N_4(A_0) = \{A_1, A_3, A_5, A_7\}$

Pixeles 8- vecinos de A0

A4	A3	A2
A5	A0	A1
A6	A7	A8

Tabla 2. Pixeles con vecindad 8

Los cuales se representan como: $N_8(A_0) = \{A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8\}$

Se dice que dos puntos p y q están 4-conectados si p pertenece a los 4-vecinos de q y viceversa. Y si están 8-conectados si p pertenece a 8-vecinos o viceversa.

Una vez entendido este concepto, podemos ahora si definir la técnica de etiquetado y cuenta de objetos. El etiquetado es una técnica que teniendo una imagen binaria, nos permite etiquetar cada uno de los objetos conectados presentes en la imagen. Con esto podemos:

- Distinguir los objetos respecto del fondo
- Distinguir un objeto respecto a los demás objetos
- Conocer el número de objetos de una imagen.

⁹<http://biometrics.org/bc2009/presentations/wednesday/Alonso-Fernandez%20MR%2016%20Wed%20930%20AM%20-%20950%20AM.pdf>

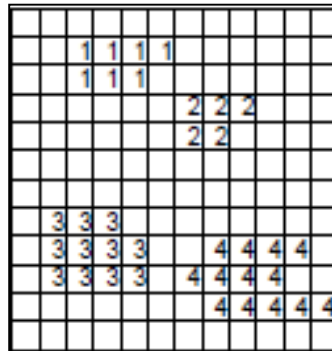


Fig. 29. Etiquetado y cuenta de objetos [5]

Los algoritmos que se utilizan para etiquetar tienen buenos resultados con objeto convexos, pero presentan problemas cuando aparecen objetos que tienen concavidades. Considerando este aspecto, no se podría ocupar este algoritmo cuando un objeto tiene forma espiral.

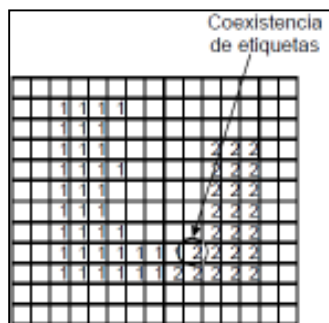


Fig. 30. Imagen en forma de U etiquetada mediante el algoritmo de etiquetado. [5]

CAPITULO 3

TÉCNICAS DE RECONOCIMIENTO DE PATRONES Y MODELOS DE CLASIFICADORES

3.1 INTRODUCCIÓN

Cuando ya se tiene la imagen segmentada en regiones, los pixeles segmentados se encuentran representados y descritos de manera que estén listos para un procesamiento posterior. Cuando hablamos de representar regiones estamos tocando dos ámbitos, el primero que es posible representar la región en términos de sus características externas, y la segunda representarla en función de sus características internas.

La etapa de reconocimiento de patrones denominada también lectura de patrones, identificación de figuras y reconocimiento de formas, consiste en el reconocimiento de patrones de características y descripción donde cada objeto queda representado por un conjunto de descriptores. El sistema de reconocimiento debe asignar a cada objeto su categoría o clase.

3.2 TÉCNICAS DE RECONOCIMIENTO DE PATRONES

Existen diversas técnicas para el reconocimiento de patrones, que están divididas en dos categorías: Decisiones teóricas y estructurales. En las decisiones teóricas la base para el procesamiento son descriptores cuantitativos como son el área, longitud, textura, etc. Por otro lado, las técnicas de decisión estructurales se basan en descriptores cualitativos.

3.2.1 MOMENTOS DE HU

Los momentos de Hu y los relacionados con la invariancia han sido extensivamente analizados en el reconocimiento de patrones de imágenes en una gran variedad de aplicaciones.

Los momentos invariantes fueron introducidos por primero vez por Hu. HU derivo seis momentos ortogonales y un momentos asimétrico basados todos estos en el algebra, los mismos que no son dependientes de la rotación, tamaño y posición pero si son dependientes de una proyección paralela. Los momentos invariantes han sido probados para ser medidos adecuadamente para la coincidencia de imágenes a pesar de su rotación, escalamiento y rotación. Los momentos invariantes han sido extensivamente aplicados al reconocimiento de imágenes como ya se explico anteriormente, al registro de imágenes y a la reconstrucción de las mismas. Sin embargo, las imágenes digitales en las aplicaciones prácticas no son continuas y tampoco están libres de ruido, debido a que estas son cuantificadas con una precisión finita de pixeles y sus coordenadas discretizadas¹⁰.

Por otro lado, el ruido puede ser introducido por varios factores, uno de ellos puede ser la cámara lo que hace que exista errores al momento de realizar el cálculo computacional. En otras palabras, los momentos invariantes pueden variar dependiendo de la geometría de la imagen.

¹⁰ Object Detection using Geometric Invariant Moment, American Journal of Applied Sciences

La traslación, rotación y escalamiento son las transformaciones más simples de las coordenadas espaciales, que a veces son denominadas transformada de similitud, en donde influyen 4 parámetros que son:

$$x' = sR \cdot x + t \quad (8)$$

Donde t es un vector de traslación, s es un factor de escalamiento positivo y R es una matriz de rotación.

$$R = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \quad (9)$$

Donde α es un ángulo de rotación.

La invarianza con respecto a la traslación, rotación y escalamiento se requiere en casi todas las aplicaciones prácticas, debido a que los objetos deben ser reconocidos correctamente, independientemente de su posición y orientación en la escena y la distancia del objeto a la cámara.

Por otro lado, el modelo TRS (transformación, rotación y escalamiento) es una aproximación suficiente de la deformación de la imagen actual si la escena es lisa y perpendicular al eje de la óptica.

DEFINICION DE LOS MOMENTOS

Un momento general $M_{pq}^{(f)}$ de una imagen $F(x, y)$, en donde p y q son enteros no negativos y $r=p+q$ que se denomina el orden del momento. Este se puede definir de la siguiente manera.

$$M_{pq}^{(f)} = \iint_D p_{pq}(x, y) f(x, y) dx, dy \quad (10)$$

En donde $p_{00}(x, y)$ $p_{10}(x, y)$... $p_{kj}(x, y)$ son bases de polinomios definidos sobre la región D .

La definición más común de los momentos geométricos puede ser dada, tomando como referencia la ecuación 10, de la siguiente forma:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^r y^s f(x, y) dx, dy \quad (11)$$

Escribiendo de otra forma, los momentos estadísticos se pueden definir como:

$$m_{rs} = \sum_{ij \in \mathcal{R}} i^r j^s \quad \text{Para } r, s \in N \quad (12)$$

En donde \mathcal{R} es el conjunto de pixeles de la región. El parámetro $r+s$ denota el orden del momento. El momento de orden cero m_{00} corresponde al área de la región A . el centro de gravedad de una región queda definido por:

$$k = \frac{m_{10}}{m_{00}} \quad l = \frac{m_{01}}{m_{00}} \quad (13)$$

En donde el momento de orden cero m_{00} corresponde al área de la región a analizarse, m_{10} es el centroide de la coordenada en x y m_{01} es el centroide de la coordenada en y .

Con la ayuda de las coordenadas del centro de gravedad se definen los momentos que son invariantes al desplazamiento de la región en la imagen.

$$u_{rs} = \sum_{ij \in \mathcal{R}} (i - k)^r (j - l)^s \quad (14)$$

Los parámetros r y s denota el orden del momento.

En la teoría de reconocimiento de patrones son las características derivadas de los momentos centrales, denominados momentos de Hu.

$$\phi_1 = \eta_{20} + \eta_{02} \quad (15)$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \quad (16)$$

$$\phi_3 = (\eta_{30} + \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (17)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (18)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \quad (19)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} - \eta_{12})(\eta_{21} + \eta_{03}) \quad (20)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (20)$$

Con

$$\eta_{rs} = \frac{\mu_{rs}}{\mu_{00}^t} \quad t = \frac{r+s}{2} + 1 \quad (21)$$

El parámetro t es aquella constante que determina la constante del centro de masa del objeto en análisis. Para cada uno de los 7 momentos esta varía con lo que se obtiene su valor característico.

Los momentos de Hu son invariantes a la traslación, rotación y escalamiento. Es decir que dos regiones que tengan la misma forma pero que sean de distinto tamaño y que estén ubicadas en posiciones y orientaciones distintas en la imagen tendrán momentos de Hu iguales.

3.2.2 SIFT

Este algoritmo fue desarrollado para el reconocimiento de objetos. SIFT detecta un punto característico de un objeto el mismo que puede ser reconocido con la invariancia de la iluminación, escala rotación y transformaciones afines.

Las ventajas que ofrece este método son:

- No se necesita transformar a coordenadas polares, para una precisión alta en la segmentación de imágenes.
- Debido a su invarianza a la iluminación, escala y rotación se espera que sea confiable cuando se use en la adquisición de imágenes con las siguientes condiciones:
 - Conveniencia del usuario
 - Aplicabilidad a los ambientes no cooperativos¹¹.

3.2.2.1 Operación de SIFT

Para poder aplicar SIFT y verificar si los objetos son identificados de manera correcta, debemos considerar dos aspectos.

1. Características de extracción

- ✓ Construcción de la escala espacial
- ✓ Localización de la detección extrema
- ✓ Descriptor de la extracción del punto clave

Construcción de la escala espacial: la primera etapa es construir una pirámide Gausiana o también llamada escala espacial por la imagen de ingreso. Este está formado por la convolución de la imagen original con una función gaussiana de ancho variable. La diferencia de un gaussiano (DOG), $D(x,y,\sigma)$ es calcular como diferencia entre dos imágenes filtradas, una con una constante de escalamiento K y la otra si esta:

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (22)$$

Estas imágenes, $L(x, y, \sigma)$, son producidas por la convolución de funciones gaussianas, $G(x, y, k\sigma)$ con una imagen de entrada, $l(x, y)$

$$L(x, y, k\sigma) = G(x, y, \sigma) * l(x, y) \quad (23)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\{-(x^2 + y^2)/2\sigma^2\} \quad (24)$$

Localización de la detección extrema: En esta etapa se debe encontrar los puntos DOG de la pirámide. Para detectar un máximo y un mínimo local de $D(x,y,\sigma)$, cada punto se compara con los píxeles de todos los vecinos. Si el valor es el mínimo o máximo, este punto es una extrema. Para

¹¹ http://www.cs.st-andrews.ac.uk/~yumeng/yumeng-SIFTreport-5.18_bpt.pdf

mejorar la localización de este punto, se puede usar las series de Taylor, que está definido de la siguiente manera:

$$z = -\left(\frac{\partial^2 D}{\partial x^2}\right)^{-1} \frac{\partial D}{\partial x} \quad (25)$$

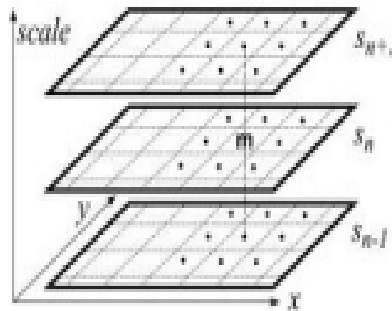


Fig 31. Localización de la detección extrema

Descriptor de la extracción del punto clave: en esta etapa se intenta eliminar algunos puntos de la lista de candidatos de puntos clave para encontrar a los que tengan bajo contraste y tengan una localización muy pobre. El valor del punto clave en la pirámide DoG está definido por:

$$D(z) = D + \frac{1}{2} \frac{\partial D^{-1}}{\partial x} z \quad (26)$$

2. Características de reconocimiento

- ✓ Punto clave del reconocimiento
- ✓ Recorte de falsas concordancias

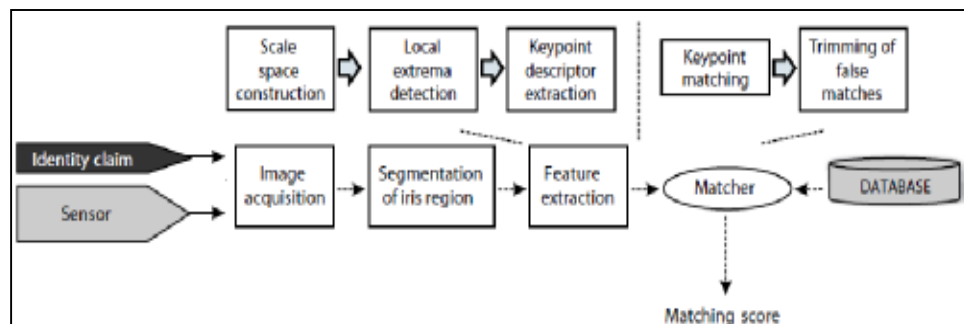


Fig. 32. Proceso para la obtención de características con el método SIFT [7]

3.2.4 REDES NEURONALES

Las redes neuronales son estructuras pensadas para simular las estructuras y el comportamiento del sistema nervioso, por esta razón una neurona artificial posee entradas, salidas y un estado. Además de estar conectadas a otras neuronas siendo las salidas de estas las entradas de las demás¹².

Cada neurona realiza una operación matemática, y además de agrupar en capas lo que constituye una RNA. Un sistema neuronal o conexionista se compone de lo siguiente:

- Un conjunto de procesadores elementales
- Un patrón de conectividad o arquitectura
- Una dinámica de activaciones
- Una regla o dinámica de aprendizaje
- En entorno donde opera

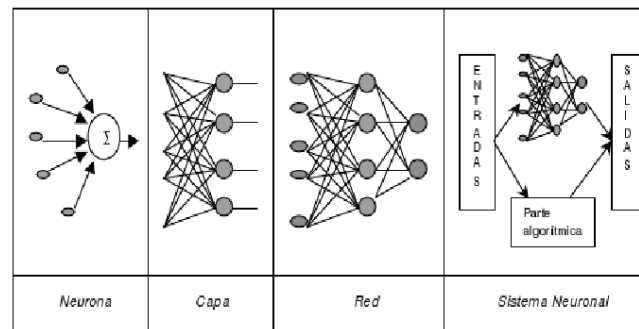


Fig. 33 Esquema de una red neuronal [4]

3.2.4.1 CARACTERÍSTICAS DE UNA RED NEURONAL

Una red neuronal consta de las siguientes características:

Procesamiento paralelo: Si queremos procesar una imagen, un ordenador tarda algunos minutos en recorrer toda la imagen. Por otro lado el cerebro necesita unos pocos segundos en procesar millones de píxeles, ya que las neuronas que interviene en el sistema trabajan en paralelo.

Memoria distribuida: en un ordenador la información se almacena en posiciones continuas bien definidas, en una red neuronal se almacena en las sinapsis de la red. Si una se daña perdemos la información.

Aprendizaje adaptativo: la capacidad de aprender tareas en base al entrenamiento elimina la necesidad de elaborar modelos a priori o indicar funciones de distribución de probabilidad.

Auto organización: Las redes neuronales utilizan su capacidad de aprendizaje adaptativo para organizar la información que reciben durante el aprendizaje/operación.

3.2.4.2 REDES PERCEPTRON

¹² Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

En la forma más básica, la red perceptron, aprende una decisión lineal entre dos tipos de opciones. La respuesta de este dispositivo está basada en la suma de los pesos de las entradas.

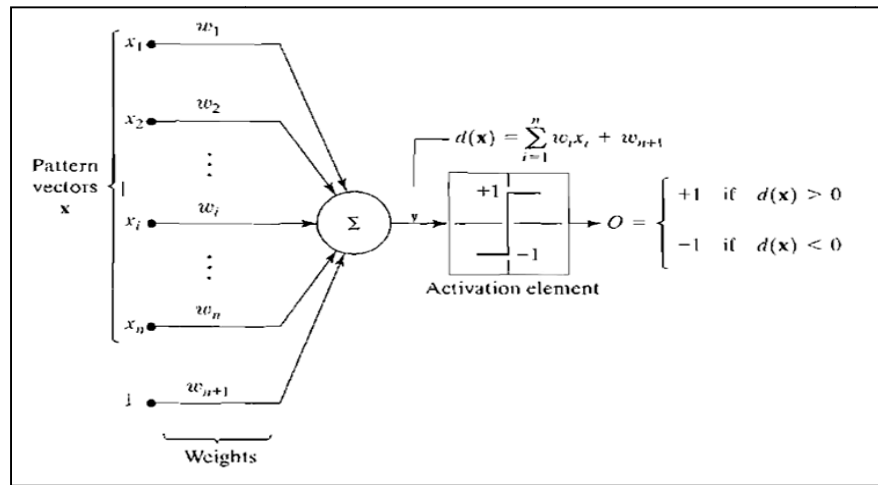


Fig. 34. Red Perceptron [4]

$$d(x) = \sum_{i=1}^n w_i x_i + w_{n+1} \quad (27)$$

Esta función representa la decisión de una función de lineal con respecto a los componentes de los vectores patrones. Los coeficientes $w_i = 1, 2, \dots, n, n+1$ son llamados pesos que modifican los ingresos antes que ellos sean sumados y vueltos a ingresar por el mismo elemento. En este sentido los pesos son análogos a la sinapsis en un el sistema neurológico de un ser humano.

Cuando $d(x) > 0$ el elemento de disparo causa la salida de la red perceptron en +1 indicando que el patrón x fue reconocido y que pertenece a la clase w_1 . Lo contrario ocurre cuando $d(x) < 0$. Este modo de operación concuerda con la acotación de la ecuación anterior. Cuando $d(x) = 0$ el objeto x no sabe a qué clase pertenece lo que da una condición de indeterminación¹³.

$$d(x) = \sum_{i=1}^n w_i x_i + w_{n+1} = 0 \quad (28)$$

Donde la ecuación del hiperplano es un patrón de n dimensiones en el espacio. Geométricamente los primeros n coeficientes establecen la orientación del hiperplano, y los últimos coeficientes $w_{n+1} = 0$ es proporcional a la distancia perpendicular del origen hacia el hiperplano.

¹³ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

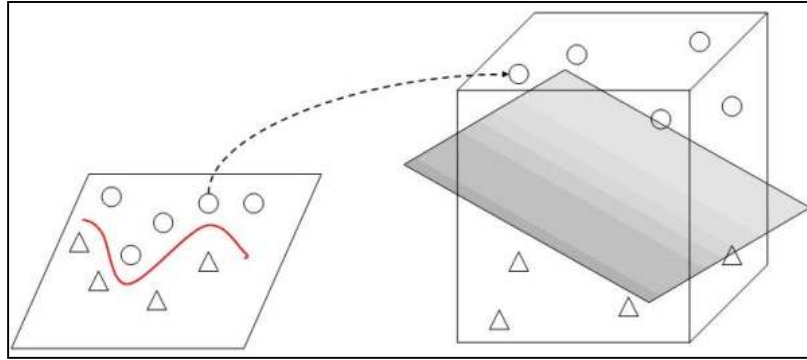


Fig 35. Separación de objetos en un hiperplano [15]

La salida del elemento depende del signo de $d(x)$. En lugar de probar una función entera para determinar si es positivo o negativo, podríamos probar con la sumatoria de la ecuación anterior, donde podemos definir lo siguiente:

$$0 = \left\{ \begin{array}{ll} +1 & \text{si } \sum_{i=1}^n w_i x_i > -w_{n+1} \\ -1 & \text{si } \sum_{i=1}^n w_i x_i < -w_{n+1} \end{array} \right\} \quad (29)$$

3.3 MODELO DE CLASIFICADORES

3.3.1 LÓGICA DIFUSA

La lógica difusa nació en 1965 siendo Lofti Zadeh su creador. A este científico se le atribuyen muchas contribuciones en el desarrollo de la transformada Z para sistemas lineales discretos. La lógica difusa difiere de la lógica clásica debido a que esta última maneja de forma imprecisa la clasificación de valores. Esta además permite una clasificación más amplia donde la pertenencia a un grupo es gradual.

La lógica difusa como tal se refiere a toda la serie de teorías para modelación, control, algoritmos computacionales e inteligencia artificial que se basan en un razonamiento de incertidumbre o imprecisión, donde la palabra clave se basa en límites que no son cortantes¹⁴.

3.3.1.1 APLICACIONES DE LA LOGIA DIFUSA

SECTOR AUTOMOTRIZ

- Control de la transmisión

¹⁴ Maestría en control y automatización-Control digital avanzado, Fundamentos del Control Difuso, Ing. Ismael Minchala, Msc

- Control del motor
- Control del frenado

SECTOR INDUSTRIAL

- Hornos y calderas
- Columnas de destilación
- Producción del cemento
- Refinería y petroquímicas
- Industria química farmacéutica y alimenticia

3.3.1.2 CRISP LOGIC

La teoría de conjuntos clásica o Crisp es la base de los conjuntos difusos o Fuzzy sets, ya que la teoría difusa tiene operaciones similares a la Crisp Logic.

En la teoría clásica las operaciones de unión, intersección y complemento están definidas por:

Sean A y B subconjuntos del universo X, la unión, intersección y complemento se definen:

Unión de un conjunto Clásico A y B

$$A \cup B = \{x | x \in A \vee x \in B\} \quad (30)$$

Intersección de un conjunto clásico de A y B

$$A \cap B = \{x | x \in A \wedge x \in B\} \quad (31)$$

Complemento de un conjunto clásico A

$$A^c = \{x | x \notin A\} \quad (32)$$

3.3.1.3 FUNCION CARACTERISTICA

Si A representa un conjunto del universo X, su función característica γ_A puede definirse como:

$$\gamma_A = X \rightarrow \{0, 1\} \quad (33)$$

Es decir que si "x" pertenece a "A" entonces la función vale 1, y si no pertenece vale 0.

3.3.1.4 CONJUNTO DIFUSO

Mientras los conjuntos clásicos pueden definirse por funciones características, los conjuntos difusos se definen por funciones de membrecía.

Por ejemplo un conjunto clásico es:

X= Integrantes

A=Grupo de hombres

B= Estudiantes de Música

Un conjunto difuso es:

\tilde{A} = *Personas moderadamente Altas*

\tilde{B} = *Personas medianamente Altas*

Para estos conjuntos no podemos utilizar diagramas de Venn porque moderadamente altas y medianamente blancas varían de persona a persona y dependen de la situación. Por ejemplo, asignaremos un grado entre 0 y 1 para denotar el moderadamente altas y medianamente blancas, donde se obtiene:

Conjunto	Paty	Ana	Danny	Eric	Juan	Linda
Moderadamente altas	0.5	0.7	0.3	0.4	0.8	0.3
Medianamente blancas	0.3	0.1	0.5	0.8	0.9	0.2

Tabla. 3. Definición de un Conjunto Difuso

Donde 1 representa una total pertenencia a cada grupo y 0 una no pertenencia al grupo. En un conjunto clásico tendríamos que decidir entre 0 y 1 exclusivamente, pero en un difuso puede ser cualquier valor entre este rango [0,1]

3.3.1.5 Función de Membrecía

Si A representa un conjunto del universo X, su función de membresía $\mu_A(x)$ para el conjunto difuso A se le llama valor de membresía o grado de membresía. El grado de membresía es la pertenencia de x al conjunto difuso A.

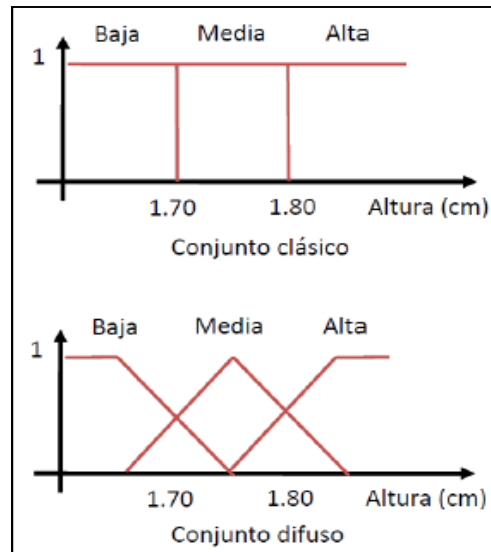


Fig. 36. Función de membresía [8]

Es importante mencionar que los conjuntos difusos deben definirse correctamente para reflejar la situación. Por ejemplo si nos encontramos en una carretera libre o en una autopista.

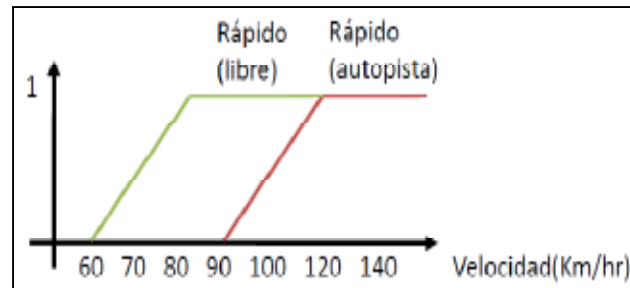


Fig. 37. Planteamiento de un problema difuso. [8]

Los conjuntos difusos con aquellos que poseen límites suaves y están definidos por una función que mapea objetos en un dominio de acuerdo a su grado de pertenencia o membresía en el conjunto.

3.3.1.6 GRADO DE MEMBRESIA

El grado de membresía es el valor numérico que se asigna como pertenencia a un conjunto difuso para una variable lingüística.

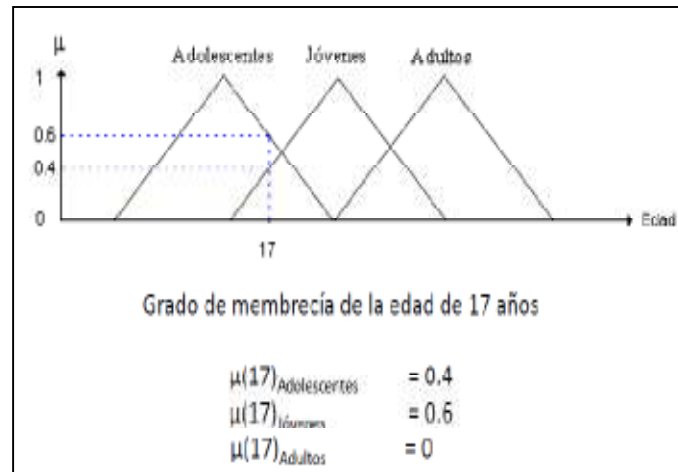


Fig. 38 Grado de Membrecía [8]

3.3.1.7 VARIABLE LINGÜÍSTICA

La variable lingüística es aquella que toma valores con palabras en lugar de valores numéricos. En una función de membrecía los conjuntos son asociados con variables lingüísticas.

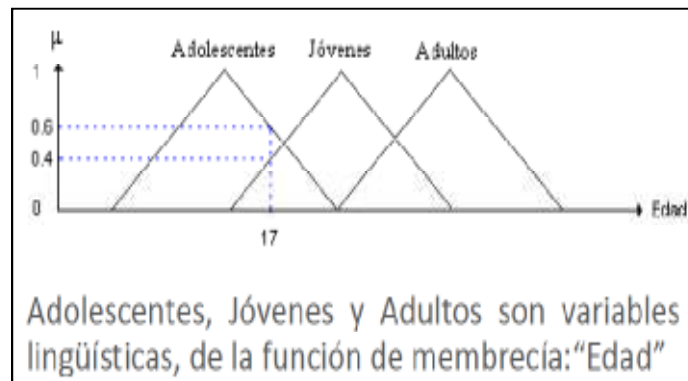


Fig. 39 Variable lingüística [8]

3.3.1.8 RAZONAMIENTO APROXIMADO

El razonamiento aproximado es la forma más conocida de la lógica difusa que incluye una variedad de reglas de inferencia. Este en contraste con la inferencia de la lógica clásica, establece que el consecuente de un conjunto de proposiciones difusas depende esencialmente de la interpretación dada a estas proposiciones. Este es usado para representar y razonar con conocimientos expresados en forma primitiva, esto es, expresados en una forma de lenguaje natural¹⁵.

¹⁵ Maestría en control y automatización-Control digital avanzado, Fundamentos del Control Difuso, Ing. Ismael Minchala, Msc

3.3.1.9 CONTROLADOR DIFUSO

Fuzificación: convierte los valores leídos de las variables del proceso necesarias para el control del proceso a grados de membresía o pertenencia a un conjunto difuso.

Mecanismo de inferencia: calcula el valor de la manipulación o salida del controlador con base en las reglas de inferencia difusas.

Defuzificación: convierte el valor difuso de la manipulación a una señal de control que se pueda escribir en el proceso.

Normalmente en un sistema de control difuso en una entrada-una salida, las variables de error (e) y la derivada del error son las que procesas para emitir la señal de control $m(t)$.

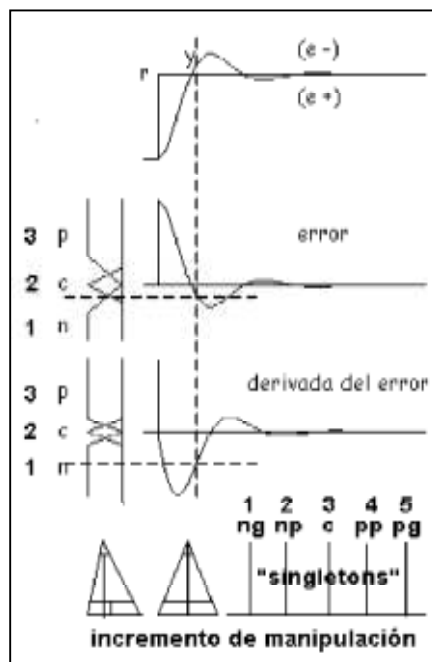


Fig. 40 Sistema de un Controlador Difuso

3.3.1.10 DEFUSIFICACION

El proceso de defuzificación también es conocido como Fuzzy Reasoning. Este proceso consiste en calcular a partir de las reglas de inferencia difusas, el valor crisp que va de salida al proceso.

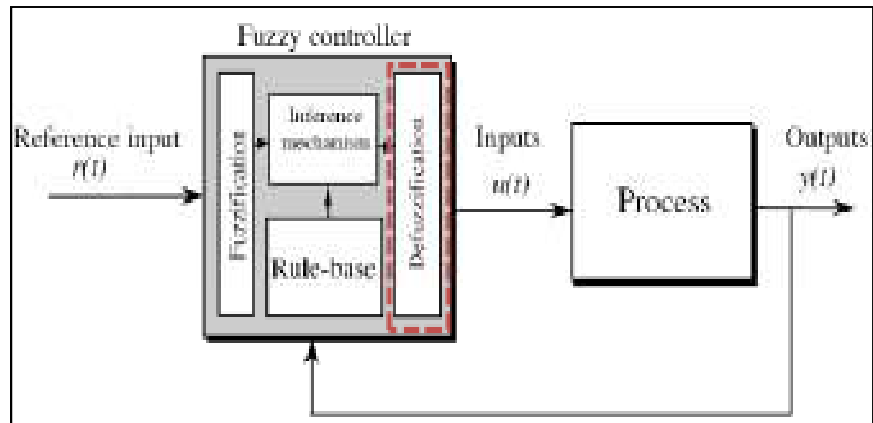


Fig. 41 Proceso de defusificación de un conjunto difuso [8]

Existen varios métodos para defusificar, entre los más importantes se encuentran:

- El método Mamdani
- El método del máximo valor de membresía
- El método ponderado o simplificado

Por otro lado los métodos de defusificación se basan en las reglas de inferencia del tipo:

Regla 1: Si x es A_1 y y es B_1 , entonces z es C_1
 Regla 2: Si x es A_2 y y es B_2 , entonces z es C_1
 · · · · ·
 · · · · ·
 · · · · ·
 Regla 3: Si x es A_n y y es B_n , entonces z es C_n

Donde la función de salida se expresa como la unión o el máximo de los conjuntos de salida, en caso de que la acción sea función, es decir:

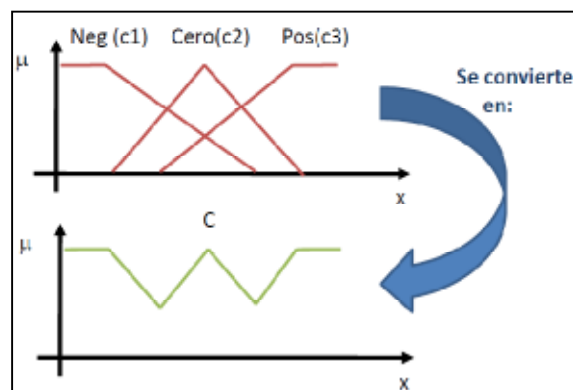


Fig. 42. Función de salida de un conjunto difuso

EL MÉTODO MAMDANI

El método del centroide aplica para sistemas continuos básicamente la defusificación consiste en encontrar el centro de masa de la función de salida.

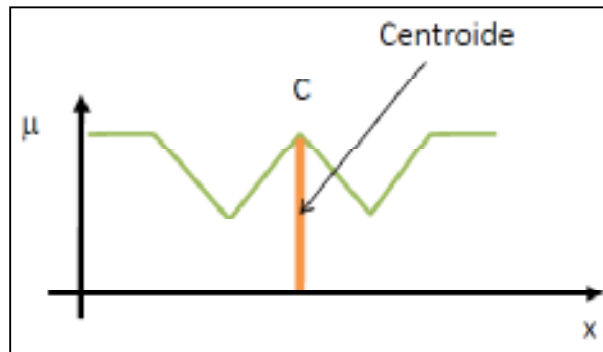


Fig. 43 Método de Mamdani

El centroide o centro de masa se calcula como:

$$z = \int \frac{\mu_c(z)z \cdot dz}{\mu_c(z) \cdot dz} \quad (34)$$

El símbolo de la integral es una algebraica ya que la función o variable de salida es ahora crisp.

EL MÉTODO DEL VALOR MÁXIMO

El método del valor máximo es muy sencillo, ya que solo se debe escoger el pico máximo de la función generada de salida.

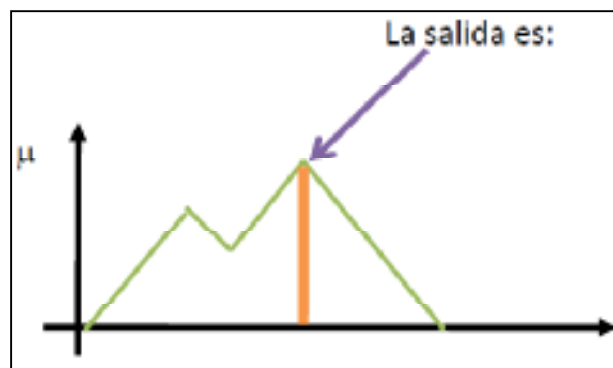


Fig. 44. Método del valor máximo

EL MÉTODO DEL PONDERADO

El método ponderado es el más utilizado en aplicaciones digitales o en control de procesos. Sin embargo es deseable que la función de salida sea simétrica para su mejor desempeño. La salida puede calcularse como:

$$z = \frac{\sum_{i=1}^n w_i c_i}{\sum_{i=1}^n w_i} \quad (35)$$

3.3.2 REDES BAYESIANAS

Las redes bayesianas o probabilísticas son una representación gráfica de dependencias para el razonamiento probabilístico en sistemas expertos.

Formalmente las redes bayesianas son gráficos a cíclicos dirigidos cuyos nodos representan variables y los arcos que los unen codifican dependencias condicionales entre las variables.

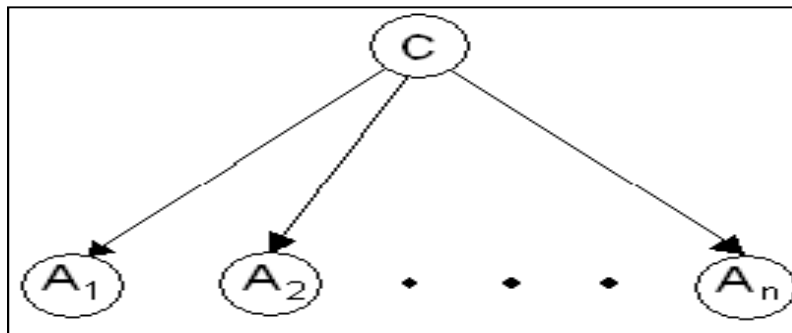


Fig. 45. Modelo de una red Bayesiana.

Sea $U=\{x_1,x_2,..X_n\}$ un conjunto de variables aleatorias. Formalmente una red bayesiana para U es un par $B=<G, T>$ en el que:

G es un gráfico a cíclico dirigido en el que cada nodo representa una de las variables $X_1,X_2,..X_n$ y cada arco representa relaciones de dependencia directas entre las variables. La dirección de los arcos indica que la variable apuntada por el arco depende de la variable situada en su origen¹⁶.

T es un conjunto de parámetros que cuantifica la red. Contiene las probabilidades para cada posible valor de x_i de cada variable X_i y cada posible valor P_{x_i} de P_{x_i} , donde este último denota al conjunto de padres de X_i en G . así una red bayesiana B de una distribución de probabilidad conjunta sobre U dada por:

$$P_B(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_B(X_i | \prod_{X_j \in \text{Pa}(X_i)} X_j) \quad (36)$$

En donde $X_1,X_2,..X_n$ son los posibles valores que la red puede tomar.

3.3.2.1 APRENDIZAJE DE REDES BAYESIANAS

El problema del aprendizaje bayesiano puede describirse informalmente como: dado un conjunto de entrenamiento $D=\{u_1,u_2,..u_n\}$ de instancias de U , encuéntrase la red B que se ajuste mejor a D .

¹⁶ <http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t9knn.pdf>

Típicamente este problema tiene dos partes:

Aprendizaje estructural: obtener la estructura de la red

Aprendizaje paramétrico: conocida la estructura del grafo, obtener las probabilidades correspondientes a cada nodo.

3.3.2.2 APLICACIONES

Las redes bayesianas son un tipo de modelos de minería de datos que pueden ser utilizados en cualquiera de las siguientes actividades:

- ✓ Prevención del fraude
- ✓ Prevención del abandono de clientes
- ✓ Blanqueo de dinero
- ✓ Marketing personalizado
- ✓ Scoring de clientes
- ✓ Clasificación de datos estelares

3.3.3 CLASIFICADORES DE DISTANCIA MINIMA

Las técnicas de reconocimiento basadas en coincidencias, representan a cada clase en un vector patrón prototipo. Un patrón desconocido es asignado a la clase en la cual existe mayor coincidencia en términos de algunas métricas predefinidas. Uno de los enfoques más simples pero robustos, es el clasificador de distancia mínima Euclidiana entre los vectores de los objetos desconocidos y los vectores de los patrones conocidos. Se elige la distancia más pequeña para realizar la decisión¹⁷.

Supongamos que definimos el prototipo de cada clase como la medida de su vector:

$$m_j = \frac{1}{N_j} \sum x_j \quad j = 1, 2, \dots, W \quad (37)$$

Donde N representa el tamaño de los vectores. Una forma de determinar la clase a la cual un objeto desconocido es miembro, consiste en asignarlo a la clase prototipo más cercano. Utilizando la distancia Euclidiana para determinar la cercanía, el problema de medición de distancias se reduce a calcular la siguiente ecuación:

$$D_j(x) = |x - m_j| \quad j = 1, 2, \dots, W \quad (38)$$

Donde $|a| = \sqrt{a^T a}$ es la norma euclidiana. Luego de analizar las distancias euclidianas del objeto desconocido con los patrones, asignamos este a la clase W_i si $D(x)$ es la distancia más pequeña. Es decir la distancia más pequeña implica la mejor coincidencia en esta ecuación. Resulta sencillo

¹⁷ Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition

probar con algunas manipulaciones algebraicas en matrices que la distancia más pequeña es equivalente a evaluar las funciones:

$$d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \quad j = 1, 2, \dots, w \quad (39)$$

Y asignar x a la clase W_i implica analizar el resultado numérico $d_i(x)$ máximo.

El límite de decisión que separa dos clases w_i de w_j está dado por los valores de x para los cuales $d_i(x) = d_j(x)$ o de manera equivalente para los valores de x que:

$$d_i(x) - d_j(x) = 0 \quad (40)$$

A partir de las ecuaciones anteriores, el límite de decisión entre dos clases w_i y w_j para un clasificador de distancia mínima es:

$$d_{ij}(x) = d_i(x) - d_j(x) \quad (41)$$

$$d_{ij}(x) = x^T (m_i - m_j) - \frac{1}{2} (m_i - m_j)^T (m_i + m_j) = 0 \quad (42)$$

La superficie dada por la ecuación 33 es la línea perpendicular que bisecta la línea que une m_i con m_j . Para $n=2$, la superficie se representa por una línea perpendicular, para $n=3$ es un plano y para $n>3$ resulta un hiperplano.

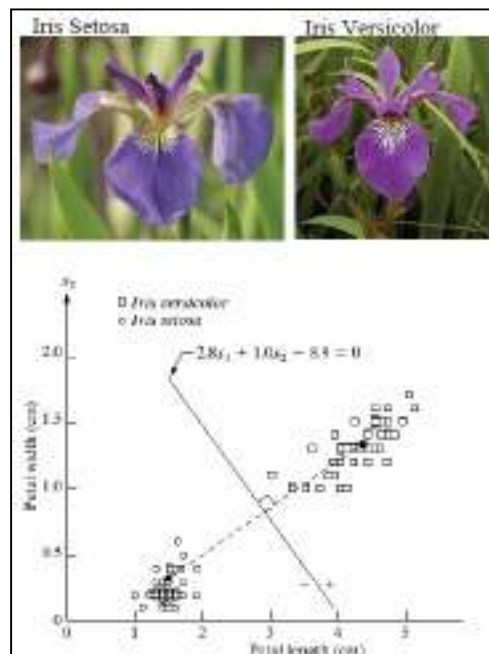


Fig. 46. Límite de decisión de un clasificador de distancia mínima [9]

3.3.4 CLASIFICADORES DE K VECINOS

K-NN (K vecinos más cercanos) es uno de los clasificadores más utilizados por su simplicidad. El proceso de aprendizaje de este clasificador consiste en almacenar una tabla con los ejemplos disponibles, junto a la clase asociada a cada uno de ellos¹⁸.

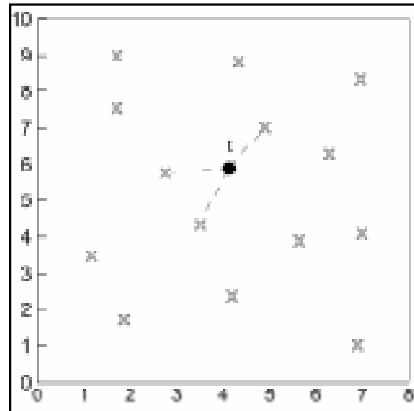


Fig. 47. Ejemplo del clasificador K-nn [10]

Algoritmo K-NN básico

La notación que se utiliza para representar este algoritmo es:

		X_1	...	X_j	...	X_n	C
(x_1, c_1)	1	x_{11}	...	x_{1j}	...	x_{1n}	c_1
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots
(x_i, c_i)	i	x_{i1}	...	x_{ij}	...	x_{in}	c_i
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots
(x_N, c_N)	N	x_{N1}	...	x_{Nj}	...	x_{Nn}	c_N
x	$N+1$	$x_{N+1,1}$...	$x_{N+1,j}$...	$x_{N+1,n}$?

Fig.48. Notación para el paradigma K-NN [10]

D indica un fichero de N casos, cada uno de los cuales está caracterizado por n variables predictoras, X_1, \dots, X_n y una variable a predecir, la clase C.

Los N casos se denotan por:

$$(x_1, c_1), \dots, (x_n, c_n) \quad (41)$$

donde

$$x_i = (x_{i,1}, \dots, x_{i,n}) \quad \text{para todo } i = 1, \dots, N \quad (43)$$

$$c_i \in \{c^1, \dots, c^m\} \quad \text{para todo } i = 1, \dots, N \quad (44)$$

¹⁸ <http://es.scribd.com/doc/58302048/85/Clasificador-de-k-vecinos-mas-cercanos>

3.3.4.1 VARIANTES SOBRE EL ALGORITMO BÁSICO

En este apartado vamos a introducir algunas variantes sobre el algoritmo básico.

- **K-NN CON RECHAZO**

La idea subyacente al K-NN con rechazo es que para poder clasificar un caso debo tener ciertas garantías. Es por ello por lo que puede ocurrir que un caso quede sin clasificar, si no existen ciertas garantías de que la clase a asignar sea la correcta.

- **K-NN CON DISTANCIA MEDIA**

En el K-NN con distancia media, la idea es asignar un nuevo caso a la clase cuya distancia media sea menor.

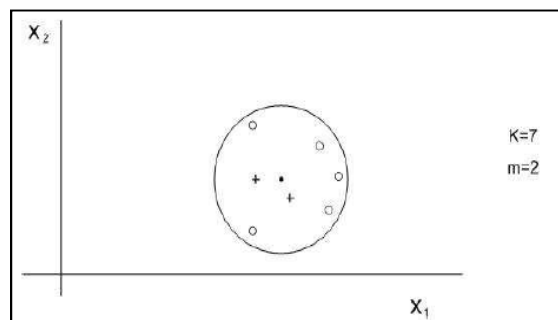


Fig 49. Ejemplo del clasificador K-NN con distancia media [10]

- **K-NN CON DISTANCIA MINIMA**

En el K-NN con distancia mínima se comienza seleccionando un caso por clase, normalmente el caso más cercano al baricentro de todos los elementos de dicha clase. En este paso se reduce la dimensión del fichero de casos a almacenar de N a m . A continuación se asigna el nuevo caso a la clase cuyo representante este más cercano.

- **K-NN CON PESADO DE CASOS SELECCIONADOS**

La idea en el K-NN con el que efectúan un pesado de los casos seleccionados es que los K casos seleccionados no se contabilicen de igual forma, sino que se tenga en cuenta la distancia de cada caso seleccionado al nuevo caso que se pretende seleccionar.

- **K-NN CON PESADO DE VARIABLES**

En todas las aproximaciones presentadas hasta el momento, la distancia entre el nuevo caso que se pretende clasificar, x , y cada uno de los casos X_r , $r=1... N$ ya clasificados pertenecientes al

fichero de casos D , da el mismo peso a todas y cada una de las n variables, X_1, \dots, X_n . Es decir, la distancia $d(x, x_r)$ entre x y x_r , se calcula por medio de la distancia euclidiana.

$$d(x, x_r) = \sum_{j=1}^n (x_j, x_{rj})^2 \quad (45)$$

Esta manera de calcular la distancia, es decir, dando el mismo nivel de importancia a todas las variables, puede resultar peligrosa para el paradigma K-NN en el caso de que algunas variable sean irrelevantes para la variable clase C . este es el motivo por el que resulta interesante el utilizar distancias entre casis que ponderen cada variable de una manera adecuada. Es decir, la distancia entre x y x_r , se calcularía:

$$d(x, x_r) = \sum_{j=1}^n w_j (x_j, x_{rj})^2 \quad (46)$$

CAPITULO 4

IMPLEMENTACIÓN DEL SISTEMA DE VISIÓN ARTIFICIAL

4.1 INTRODUCCIÓN

Para realizar todo el procedimiento del cálculo de los momentos de Hu y los clasificadores se ha utilizado el software Labview, el cual nos permite obtener varios parámetros necesarios para realizar la clasificación respectiva.

4.2 IMPLEMENTACIÓN DE HARDWARE

4.2.1 Cámaras

La cámara que se utilizó fue una Basler industrial. Esta cámara permite capturar imágenes con una alta claridad, con lo que ayuda que procesamiento de las imágenes sea lo más óptimo posible. La colocación de la misma también juega un papel muy importante, ya que al tener una buena ubicación tendremos una mejor captación de las imágenes. Por el contrario, si no la colocamos en un lugar adecuado, las imágenes que se obtendrán serán de mala calidad ocasionando que nuestro procesamiento sea erróneo.

4.2.2 Sistema de iluminación

Como mencionamos anteriormente, la iluminación es un factor muy importante en el proceso de visión para la obtención de resultados óptimos. Una deficiente iluminación implica obtener imágenes de baja calidad, lo que a su vez conlleva a resultados erróneos o un procesamiento excesivo. En nuestro caso no tuvimos mayor inconveniente en cuanto a la iluminación por lo que el sistema va a ser implementado en un lugar de bajo iluminación y poca reflectancia. Esto ayuda mucho a la cámara ya que la misma solo enfocara a los objetos en cuestión y no habrá problemas en cuanto a reconocer sus características se refiere.

4.3 Implementación de Software

4.3.1 Captura de imágenes

La captura de las imágenes se la realiza mediante el siguiente bloque de programa

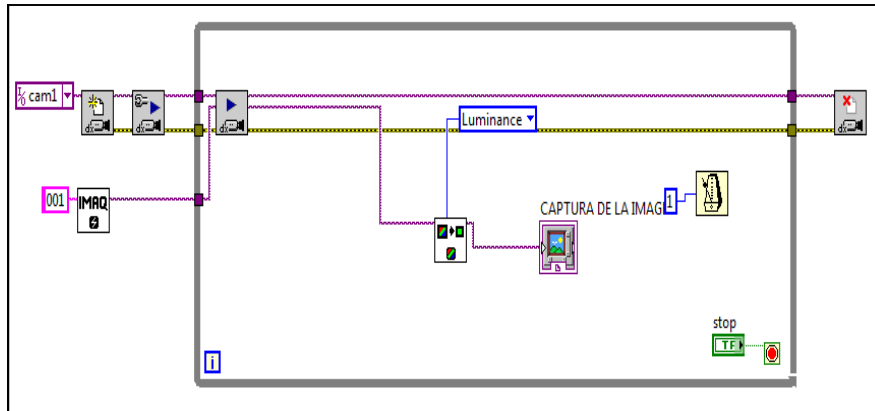


Fig. 50. Captura de la imagen

Mediante este bloque podemos realizar la captura de nuestras imágenes para en su posterior trabajo poder procesarlas.

4.4 Procesamiento de la imagen

4.4.1 Umbralización

La umbralización es un paso muy importante ya que este nos permite binarizar a nuestra imagen para que el procesamiento sea mucho más fácil. Al binarizar nuestra imagen trabajamos con valores que van de 0 a 255. Esto hace que nosotros tengamos un 0 o un 1 en nuestra imagen y según esto saber cuántos píxeles están activados y cuantos no. El bloque que nos permite hacer esto es el siguiente:

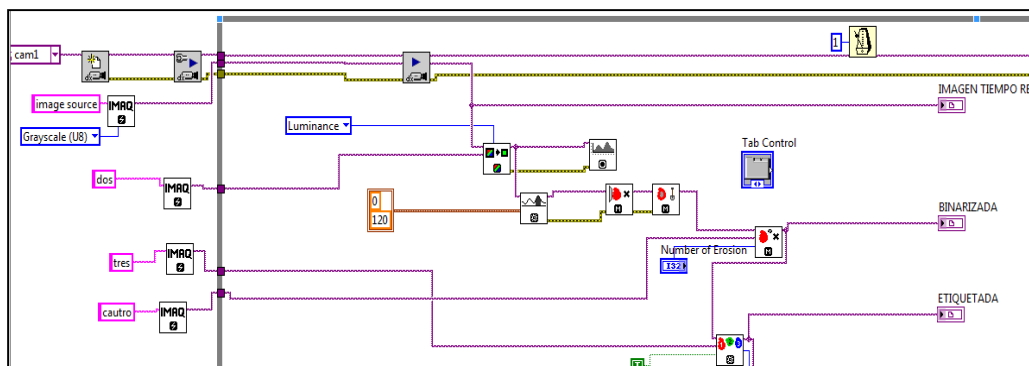


Fig. 51. Binarización de la imagen

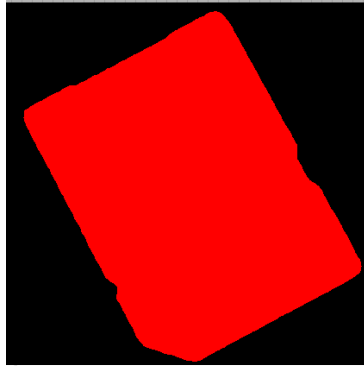


Fig. 52. Pieza binarizada

El resultado de realizar la binarización es como se puede observar en la figura anterior. Al tener de esta forma, podemos calcular los momentos de Hu de manera mucho más fácil.

4.4.2 Etiquetado

El etiquetado por su parte es otro de los aspectos que debemos considerar ya que esto nos ayuda al tener varios objetos a clasificarlos según su color y con esto obtener características independientes de cada uno de ellos evitando así la confusión de los mismos.

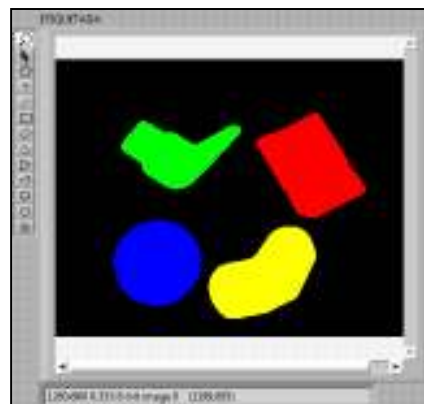


Fig. 53. Etiquetado de la Imagen

4.4.3 Filtrado del tamaño

El filtrado de la imagen nos ayuda a eliminar bordes y contornos no deseados de una imagen, con el fin de que el proceso sea lo más puro posible y que no existan valores extras que perjudiquen a este proceso.

CAPITULO 5

CALCULO MATEMÁTICO Y PROCESAMIENTO DE LAS TÉCNICAS DE RECONOCIMIENTO DE PATRONES

5.1 INTRODUCCIÓN

Dentro las técnicas de reconocimiento de las técnicas para el reconocimiento de patrones, se ha escogido para su desarrollo los momentos de Hu. Los momentos de Hu como ya se vio anteriormente, nos entregan información del objeto que estamos procesando, con el fin de que con estos valores se pueda clasificar a los objetos según sea su característica.

5.2 CALCULO DE LOS MOMENTOS DE HU.

Para poder calcular los momentos de Hu, debemos primero hacer un pre procesamiento de la imagen, es decir aplicar una serie de operaciones que permitan realizar el posterior cálculo. Todo este procesamiento se lo realizó en el software Labview.

5.2.1 PARAMETROS

Los parámetros que se consideraron para el desarrollo del software para calcular los 7 momentos fueron:

Adquisición: en esta etapa se configura la sesión de la cámara y como la imagen se la captura en RGB se incluye un paso a imagen monocromática, extrayendo a la imagen original el plano de intensidad.



Fig. 54. Adquisición de Imágenes

Preprocesamiento: En esta etapa se realiza un enmascaramiento de la imagen a través de coordenadas ingresadas por el usuario para obtener el ROI (región de interés).

Segmentación: El objetivo de la etapa de segmentación de la imagen es ingresar una imagen al sistema y obtener características como resultado, en este caso de manera particular imágenes

etiquetadas con información de centros de masa y área, para lo cual se realizan los siguientes pasos:

Umbralización: Se utiliza para obtener imágenes binarias. Para elegir los límites de umbralización debemos considerar los siguientes aspectos.

- *Erosión:* es una operación morfológica que permitirá eliminar objetos de un tamaño menor a los objetos que estamos analizando, de forma que el etiquetado de la etapa siguiente sea relativamente más sencillo de realizar.
- *Etiquetado:* se agrupan los pixeles vecinos considerando vecindad de 8.

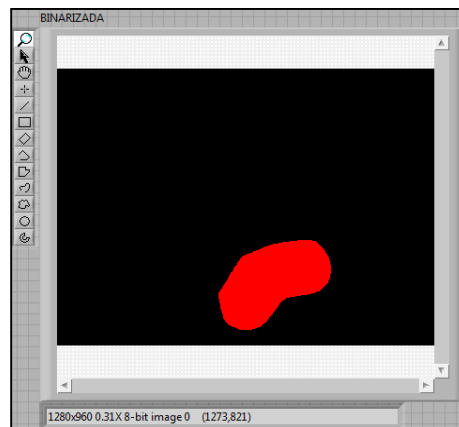


Fig.55. Binarización de la imagen



Fig.56. Etiquetado de los objetos

Caracterización: en esta etapa se convierte la imagen etiquetada en una matriz de vectores columna, conteniendo los resultados del cálculo de los momentos de Hu.

Particle Measurements (Pixels)

0	0,174529	0,002911	0,000001	0,000000	0,000000	-0,000000	0,000000	885,968767	372,106942
0	0,236672	0,016219	0,005958	0,000969	0,000001	0,000019	0,000002	415,529936	325,575302
	0,159298	0,000034	0,000001	0,000000	0,000000	0,000000	0,000000	353,148364	704,029576
	0,199231	0,012178	0,000407	0,000031	-0,000000	-0,000003	0,000000	724,552675	748,631741

Fig.57. Momentos de Hu de los 4 objetos a reconocer

Clasificación: Aplicando el criterio del clasificador de distancia mínima, k vecinos más cercanos y lógica difusa, se asigna el objeto capturado por la cámara a algún patrón previamente entrenado en el sistema.

5.2.1 PROCESO DE CÁLCULO

Una vez hecho el proceso anterior, procedemos al cálculo de los momentos de Hu. Para esto se utilizaron las siguientes fórmulas:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (11)$$

$$\phi_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \quad (12)$$

$$\phi_3 = (\eta_{30} + \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (13)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (14)$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} - \eta_{03})^2] + 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \quad (14)$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} - \eta_{12})(\eta_{21} + \eta_{03}) \quad (15)$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (16)$$

Para poder calcular, debemos obtener el centro de masa de cada objeto, y el número total de pixeles que lo conforman, ya que esto nos ayudara a definir cuál es el valor de cada uno de los momentos.

$$\eta_{rs} = \frac{\mu_{rs}}{\mu_{00}^t} \quad t = \frac{r+s}{2} + 1 \quad (17)$$

Donde r y s representan el momento a calcularse, y η_{rs} el dato de cada momento proveniente del número de pixeles de la imagen tanto en la coordenada x como y . cabe recalcar que estos se calculan utilizando el centroide y el área de la imagen.

Considerando las distintas formulas que nos ayudan a obtener los 7 momentos de Hu, se recopilaron los siguientes datos para cada uno de los objetos con los que se trabajó.

Por ejemplo para calcular uno de los momentos se siguió el siguiente proceso:

$$\eta_{rs} = \frac{\mu_{rs}}{\mu_{00}^t} \quad (47)$$

$$t = \frac{r+s}{2} + 1 \quad (48)$$

Donde r y s son el momento que va a calcularse.

$$t = \frac{0+2}{2} + 1 = 2 \quad (49)$$

$$\eta_{02} = \frac{\mu_{02}}{\mu_{00}^2} = \frac{1}{\mu_{00}^2} \sum_x^n \sum_y^m (x - \bar{x})^0 (y - \bar{y})^2 \quad (50)$$

$$\eta_{02} = \frac{1}{\mu_{00}^2} \sum_y^m (y - \bar{y})^2 \quad (51)$$

Los datos de y y \bar{y} son obtenidos de cada figura para el calcula. Estos significa que y son todos los pixeles del objeto en la coordenada y y \bar{y} es el centro de masa del mismo. Por otro lado, Pero para el primer momento necesitamos ahora el parámetro η_{20} . Siguiendo el mismo proceso que lo explicado anteriormente obtenemos:

$$\eta_{20} = \frac{1}{\mu_{00}^2} \sum_x^m (x - \bar{x})^2 \quad (52)$$

Ahora bien como en el caso anterior x significa todos los pixeles del objeto en la coordenada x y \bar{x} es el centro de masa del mismo.

De esta forma es como se van obteniendo los momentos para cada objeto. A continuación se describe el valor que se obtuvieron en los objetos de cada momento de Hu.

Moneda

$$\begin{aligned} \phi_1 &= 0,159745 \\ \phi_2 &= 0,000161714 \\ \phi_3 &= 0,0000024753 \\ \phi_4 &= 0,00000000142066 \\ \phi_5 &= 0,0000000000000000627111 \\ \phi_6 &= 0,000000000000761959 \\ \phi_7 &= 0,0000000000000000562556 \end{aligned}$$

Memory Stick

$$\begin{aligned} \phi_1 &= 0,174548 \\ \phi_2 &= 0,00287047 \\ \phi_3 &= 0,0000263223 \\ \phi_4 &= 0,00000111915 \\ \phi_5 &= 0,000000000000376112 \\ \phi_6 &= 0,0000000583927 \\ \phi_7 &= 0,00000000000476976 \end{aligned}$$

Pieza amorfa 1

$$\begin{aligned}\phi_1 &= 0,203677 \\ \phi_2 &= 0,0141866 \\ \phi_3 &= 0,000386307 \\ \phi_4 &= 0,0000291578 \\ \phi_5 &= 0,00000000308993 \\ \phi_6 &= 0,00000346912 \\ \phi_7 &= 0,000000000169138\end{aligned}$$

Pieza amorfa 2

$$\begin{aligned}\phi_1 &= 0,237502 \\ \phi_2 &= 0,015327 \\ \phi_3 &= 0,00611976 \\ \phi_4 &= 0,000907733 \\ \phi_5 &= 0,00000102021 \\ \phi_6 &= 0,00000103566 \\ \phi_7 &= 0,00000188055\end{aligned}$$

Ahora bien, los momentos de Hu como tal no hacen nada o más bien no entregan ninguna información, ya que no tenemos ningún clasificador que permita determinar parámetros que definan el reconocimiento de los objetos. Para que estos tengan utilidad se necesita de un clasificador. Existen varios tipos de clasificadores que se pueden emplear para clasificar los datos. Los que se utilizaron fueron el de distancia euclidiana o distancia mínima, el de los k vecinos más cercanos y el de lógica difusa. Estos tres clasificadores se implementaron y se contrastaron entre sí para ver cuál de ellos era el más fiable.

CLASIFICADOR DE DISTANCIA MINIMA

El clasificador de distancia mínima, está definido por la siguiente fórmula:

$$d_{ij}^{(k)} = \sqrt{\sum_{n=1}^{N_{\text{momentos}}} (X_{ij}^n - \bar{X}_{ij}^{(k)})^2} \quad (53)$$

En donde X^n representa el momento obtenido de objeto en analisis y X^k el momento del objeto patrón.

Con la formula anteriormente descrita, se pueden obtener las distancias correspondientes a cada objeto. Los resultados obtenidos fueron:

Moneda:

D1=0,0129378
D2= 0,0443367
D3= 0,136743
D4= 0,000806814

Memory stick

D1= 0,0151617
D2= 0,0285317
D3=0,15203
D4= 0,0030566

Pieza amorfa 1

D1= 0,0391771
D2=0,0272015
D3= 0,0285922
D4= 0,00440241

Pieza amorfa 2

D1= 0,0746991
D2= 0,0626122
D3= 0,0321515
D4= 0,00789217

Las distancias anteriormente descritas son las distancias medidas con respecto a los objetos en análisis. Por ejemplo la distancias de la moneda son medidas respecto a la memory stick, pieza amorfa 1 y pieza amorfa 2. Esas distancias a su vez sin contrastadas y se ve cuál de ellas es la menor para así saber si es o no la moneda que se esta clasificando.



Fig.58. Objeto reconocido-moneda

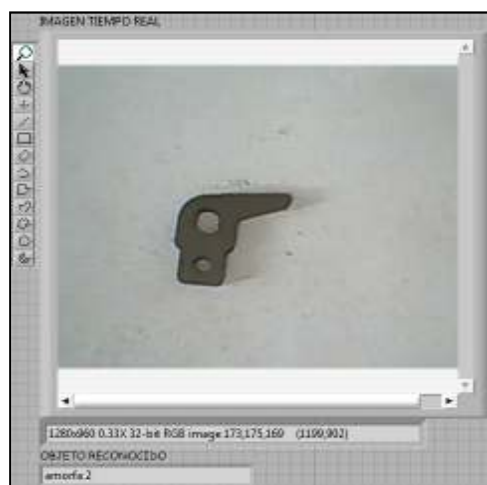


Fig.59. Objeto reconocido-pieza amorfa2



Fig.60. Objeto reconocido-memory Stick



Fig.61 Objeto reconocido-pieza amorfa 1

CLASIFICADOR K VECINOS

Aplicando el mismo concepto de la distancia euclidiana, podemos obtener las distancias más cercanas de los objetos pero aplicando k vecinos. Esta técnica a diferencia de la distancia mínima, nos permite obtener una mayor información de las distancias de los objetos que estamos calculando. Esto hace que esta técnica sea un poco más fiable que la anterior. Está definida por la siguiente ecuación:

$$d_{ij}^{(k)} = \sqrt{\sum_{n=1}^{N_{\text{momentos}}} (X_{ij}^n - \bar{X}_{ij}^{(k)})^2} \quad (54)$$

Moneda:

D1=0, 0129378
 D2= 0, 0443367
 D3= 0,136743
 D4= 0,000806814

D1.1=0, 0156778
 D2.1= 0, 0442257
 D3.1= 0,1477843
 D4.1= 0,0090681

D2.1=0, 0179378
 D2.2= 0, 041257
 D2.3= 0,133843
 D2.4= 0,000868814

Memory stick

D1= 0, 0151617
D2= 0, 0285317
D3=0, 15203
D4= 0, 0030566

D1.1=0, 0179378
D2.1= 0, 02987
D3.1= 0, 14307
D4.1= 0,003376

D2.1=0, 0129378
D2.2= 0, 03017
D2.3= 0, 14779
D2.4= 0,003576

Pieza amorfa 1

D1= 0,0391771
D2=0,0272015
D3= 0,0285922
D4= 0,00440241

D1.1=0, 039278
D2.1= 0, 0274026
D3.1= 0,029677
D4.1= 0,004671

D2.1=0, 040183
D2.2= 0, 028432
D2.3= 0, 02976
D2.4= 0,004236

Pieza amorfa 2

D1= 0,0746991
D2= 0,0626122
D3= 0,0321515
D4= 0,00789217

D1.1= 0,075791
D2.1= 0,0735233
D3.1= 0,0422735
D4.1= 0,00838317

D2.1=0, 0129378

D2.2= 0,0443367
D2.3= 0,0521616
D2.4= 0,00989317

Como se puede observar, en este caso al igual que la distancia mínima, se va comparando las distancias entre cada objeto para luego elegir la distancia más corta. En este caso se obtienen más valores con el fin de obtener un poco más de criterios de elección.



Fig. 62. Objeto reconocido- moneda



Fig. 63. Objeto reconocido-pieza amorfa 2



Fig. 64. Objeto reconocido-Memory stick



Fig. 65. Objeto reconocido-Pieza amorfa 1

CLASIFICADORES DE LOGICA DIFUSA

Los clasificadores de lógica difusa están basados en la definición de reglas de inferencia. A diferencia de la lógica tradicional estos no tienen solamente valores de 0 o 1 si no valores intermedios que hacen que la clasificación tenga un alto rendimiento.

Para lograr esto, se comienza por definir las entradas y salidas difusas del sistema, esto con el fin de que posteriormente podemos definir las reglas de inferencia que harán posible la clasificación del objeto.

Las entradas quedan definidas de la siguiente manera:

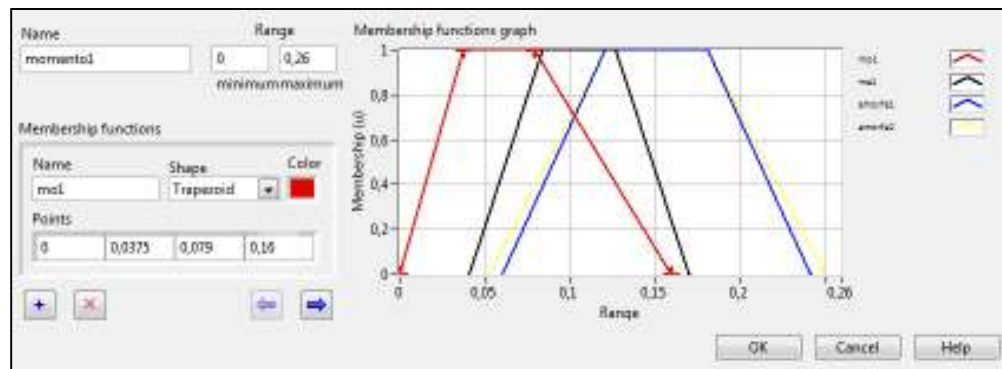


Fig. 66. Entradas difusas del momento 1 de los objetos

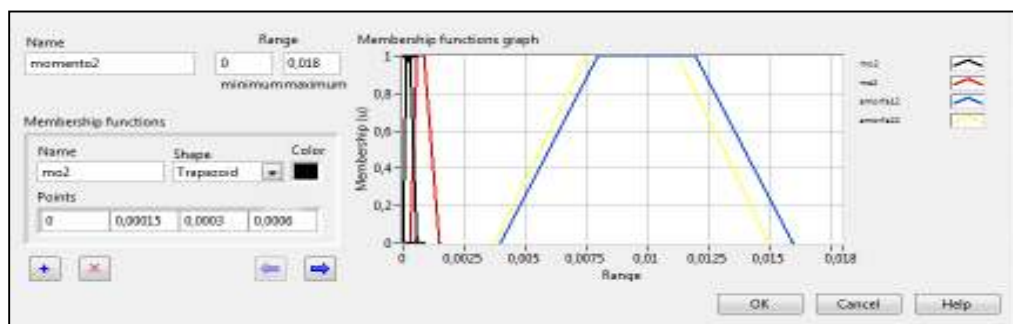


Fig.67. Entradas difusa del momento 2 de los objetos

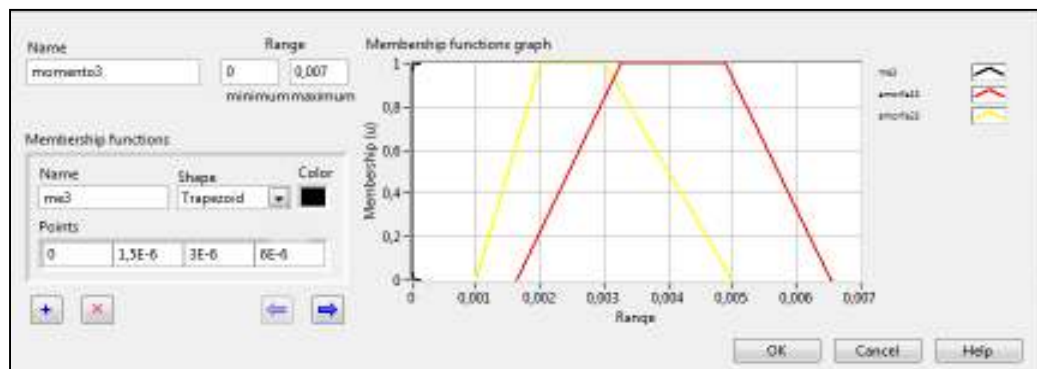


Fig.68. Entradas difusas del momento 3 de los objetos

Una vez definidas las entradas difusas, procedemos a definir las salidas para su clasificación.

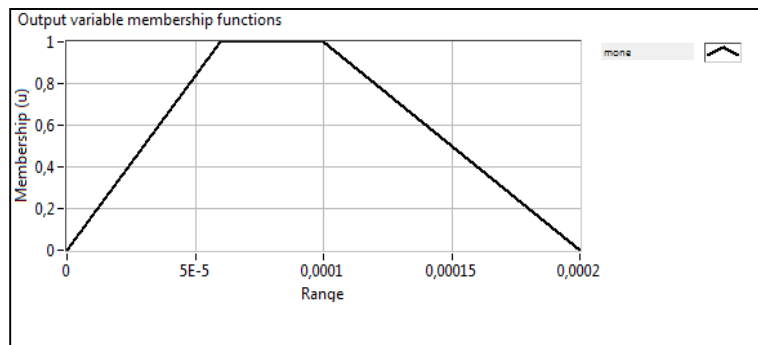


Fig 69. Salida difusa de la moneda

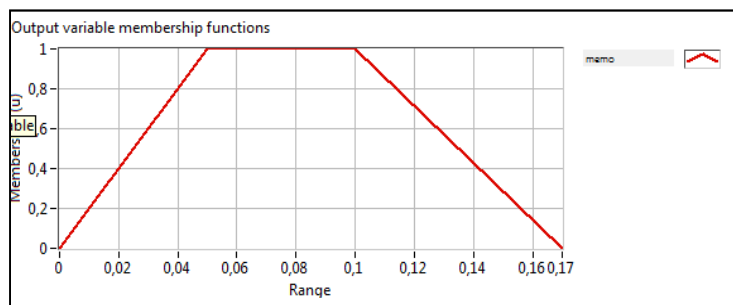


Fig. 70. Salida difusa de la memory stick

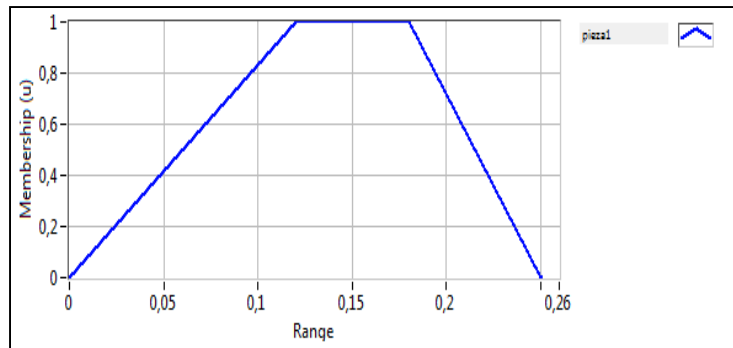


Fig. 71. Salida difusa de la pieza amorfa 1

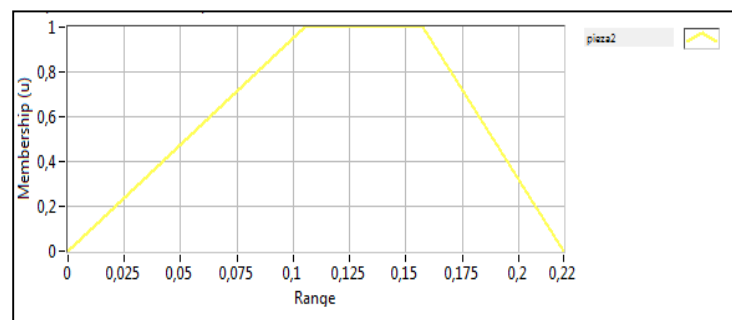


Fig. 72. Salida difusa de la pieza amorfa 2

Definidas las entradas y salidas difusas se procede a la implementación de reglas de inferencia con el fin de clasificar a los objetos.

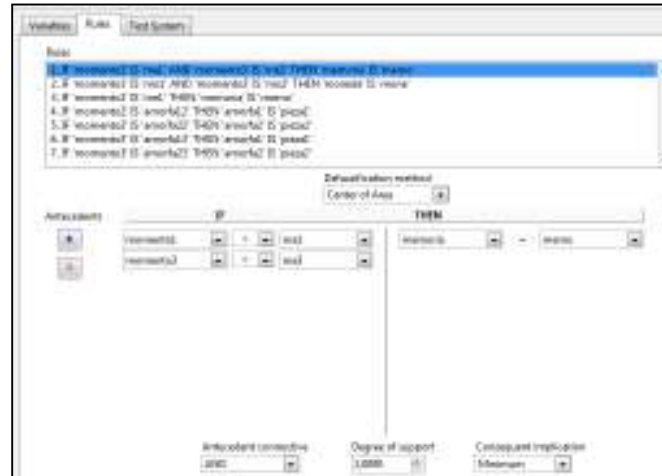


Fig. 73. Definición de reglas de inferencia

Todas la entradas dadas por los momentos de Hu se sacaron logaritmos para obtener un mayor rango de valores con el fin de que la clasificación se realice de manera adecuada.

Ahora bien, al momento de realizar la clasificación, los objetos no fueron clasificados como se esperaba. La clasificación resultó de la siguiente manera:



Fig 74 Pieza amorfa 2 reconocida



Fig 75. Memory stick reconocida



Fig 76. Clasificador erróneo. Confusión entre elementos

Como se puede observar, este clasificador no está identificando de manera correcta a los objetos por lo que podemos decir que este no es el adecuado para utilizarlo. Esto es debido a que los valores de los momentos de Hu son muy pequeños y al definir las reglas de inferencia se confunden. Se trabajó sacando el logaritmo de esto momentos para poder ver si en algo mejoraba la clasificación, sin embargo el resultado era el mismo y no clasificaba de manera adecuada. Se trabajó una tercera opción que era el emplear las áreas de las figuras, y al final se obtuvieron los mismos resultados. Con esto se puede concluir que este clasificador no es el adecuado para reconocer e identificar objetos.

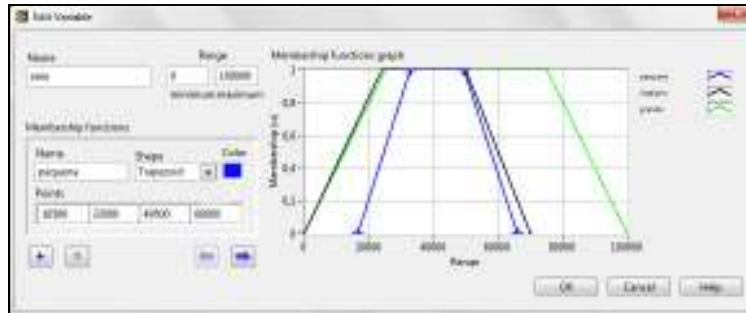


Fig. 77 Definición de entradas difusas de las áreas

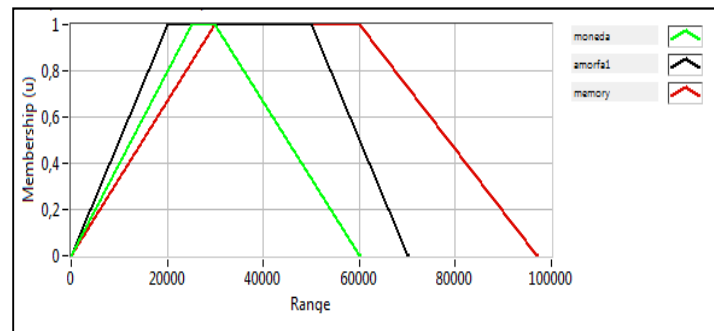


Fig.78. Definición de salidas difusas de las áreas

Como se puede observar, con las áreas tampoco se logró una mejora sustancial ya que se seguían confundiendo al momento de realizar la clasificación. Con esta opción queda definitivamente claro que el clasificador no es el adecuado.

5.3 RESULTADO

La clasificación de los objetos mediante la técnica de los momentos de Hu tiene un mejor rendimiento cuando se utiliza técnicas de clasificación de distancia mínima y la de vecinos más cercanos. Como se pudo observar estos clasificaban sin confundirse y no existían inconvenientes como ocurrió en el caso del clasificador de lógica difusa que clasificaba pero con confusión. Lo que se puede decir que no es una técnica adecuada para trabajar con valores muy pequeños que en este caso nos entregaba los 7 momentos de Hu.

CAPITULO 6

APLICACIÓN INDUSTRIAL

6.1 Introducción

La visión artificial es muy utilizada dentro de la industria metálica debido a que las aplicaciones que esta área ofrece son muy amplias. Esta industria posee un alto grado de aplicaciones con visión artificial ya que las mismas se las emplea en el reconocimiento de piezas, calibración de objetos, cortado y soldado, clasificación, etc. Cada una de las aplicaciones antes mencionadas debe tener un alto grado de precisión ya que al ser una industria en donde las medidas deben ser lo más exactas posible entonces se requiere de un sistema también que sea altamente preciso. Para esto se ha llevado una aplicación de las técnicas de los momentos de Hu y sus clasificadores a la industria metálica.



Fig. 79. Parque industrial Cuenca [16]

6.2 Industria metálica

Es el sector industrial que incluye las actividades relacionadas con el procesamiento de metales para la fabricación de piezas, máquinas y herramientas que se necesitan en la industria y en otros sectores de la economía. Por ejemplo, actividades como la fabricación de láminas de acero, y planchas para hornos, ingeniería de precisión, construcción de maquinaria eléctrica, fabricación de equipos electrónicos, y materiales, para transporte aéreo, terrestre, astilleros, etc.



Fig 80. Piezas producidas en la industria metalmecánica[12].

La industria que tiene esta clasificación manual se llama INSOMET. Esta industria además de fabricar productos como son los textiles, se dedica a la fabricación de piezas metálicas para la distribución a las fabricas que forma parte de la empresa, una de ellas es TELARTEC, que se encarga de la reconstrucción de maquinas en mal estado, y a su vez necesita de piezas que esta empresa les distribuye. Sin embargo el trabajo se lo hace de manera manual y es tedioso por la confusión y la cantidad de piezas que hay que identificar. Por esta razón se hizo el estudio a fin de analizar si es o no factible implementar en la fábrica. Según los resultados obtenidos los resultados fueron bastante buenos por lo que se hizo varias pruebas con 4 objetos que nos facilitaron en la industria para ver si con estos se llegaba a clasificar de manera adecuada. El prototipo que se construyó cumple con las normas que requiere la industria, por lo que se puede decir que es muy factible su implementación.

6.3 Identificación y clasificación de piezas metálicas

Varias son las formas que se pueden obtener de un tratamiento de acero. Clasificarlas de manera manual todos estos tipos resultaría un poco medio tedioso ya que tomaría tiempo y muchas veces hasta abría confusión ya que se tienen muchas piezas de distinto tamaño y muy parecidas en su dimensión, que en algún momento nos llegaríamos a confundir. Para esto se ha aplicado una técnica de visión artificial denominado los momentos de Hu. Estos momentos son invariantes a la rotación, movimiento, traslación y escalamiento. Con los datos obtenidos que nos entregan estos momentos dependiendo de la figura que estemos analizando, se procede a aplicar un clasificador el cual con son ciertas consideraciones permitirá que identifiquemos a los objetos sin problema alguno. Como ya se ha explicado en los capítulos anteriores, los datos deben ser clasificados y ordenados de manera tal que el clasificador no tenga confusión alguna y pueda saber cuál es la pieza que deseamos identificar.



Fig 81. Reconocimiento de una pieza metálica

En la figura anterior se puede observar la clasificación de una pieza metálica con una forma amorfa. Lo interesante de este clasificador es que no solo sirve para clasificar figuras con un estándar de forma definida, si no que no interesa la forma, posición o el tamaño que tenga, el clasificador lo identifica sin ninguna manera. La consideración que se debe tener es que el sistema debe ser implementado en un lugar en donde la luz no intervenga, ya que la misma puede incidir en el cálculo y procesamiento de las imágenes.

6.4 ANALISIS DEL PROTOTIPO EN LA INDUSTRIA INSOMET

Una vez terminado el prototipo con sus distintas pruebas, se llevó a la industria solicitante del proyecto. Aquí se realizaron distintas pruebas de los objetos que la misma nos proporcionó. Los resultados fueron bastante buenos ya que clasificaba las distintas piezas sin cometer errores.

Por otro lado, otra consideración que se debe tener presente es la ubicación de la cámara, ya que si esta no está enfocando a las imágenes o esta con un nivel de intensidad que no es la adecuada el resultado será erróneo y la clasificación de igual forma. Esto ocurrió en un principio que la cámara estuvo mal enfocada y los momentos no calculaban los valores adecuados.

EXPERIMENTACION Y RESULTADOS

Una vez obtenido los distintos resultados aplicando las técnicas de clasificadores se pudieron observar cosas interesantes al momento de realizar la clasificación. La primera observación que se pudo hacer es que en la obtención de los momentos de Hu, se obtuvieron parámetros muy pequeños que al principio parecían está mal, debido a que todos eran muy parecidos y en algunos momentos se obtuvieron valores de 0. A medida que se fueron realizando las distintas pruebas se fue entiendo el por qué de estos momentos. La cuestión es que las figuras con las que estamos trabajando no tienen un área muy grande por lo que el valor de los momentos es pequeño. Por otro lado, también debemos considerar trabajar de 6 a 7 decimales con estos, ya que al comienzo se trabajaron solo con 2 lo que ocasionaba que la clasificación no se haga de manera adecuada.

La segunda observación que se pudo realizar es que la clasificación de los objetos con los datos de los 7 momentos se la puede hacer de distinta manera, ya sea utilizando un clasificador de distancia mínima, de K vecinos más cercanos o la de lógica difusa. Este último clasificador se implementó para hacer un contraste entre estos 3 para ver cuál era la fiabilidad de cada uno de ellos. Con los dos primeros clasificadores no se tuvo ningún inconveniente ya que estos entregan valores de acorde al cálculo de la distancia y los objetos eran reconocidos sin ningún inconveniente. Con el tercer clasificador sucedió algo bastante extraño, al implementar un clasificador difuso se tenía la idea que este iba a funcionar mucho mejor que los anteriores debido a su gran nivel de aplicabilidad y fiabilidad que este tiene en procesos industriales, sin embargo al hacer las pruebas este fue el que peor reaccionó, ya que los objetos no se clasificaban de manera adecuada. Esto se debió a que al momento de definir las entradas y salidas difusas los valores de los momentos de Hu eran muy pequeños y parecidos lo que había un solape entre estas. Esto se dio a que fue imposible clasificar de esta manera. Se trato de sacar los logaritmos de cada momento, sin embargo no funciono. También se probó obteniendo las áreas de los objetos y tampoco funciono. Lo que se llegó a la conclusión que este clasificador no es el adecuado para este proceso cuando se tiene valores pequeños y sobre todo muy parecidos. Todo lo explicado anteriormente se puede ver en las siguientes graficas de precisión y cobertura.



Fig 82. Aciertos y Errores del clasificador de distancia mínima para el objeto moneda

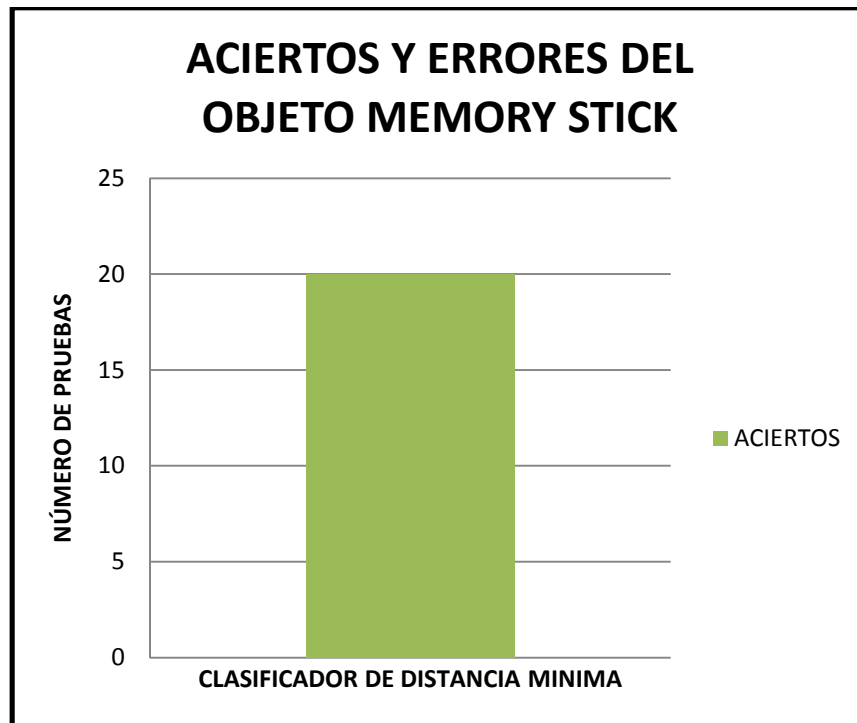


Fig 83. Aciertos y Errores del clasificador de distancia mínima para el objeto memory stick

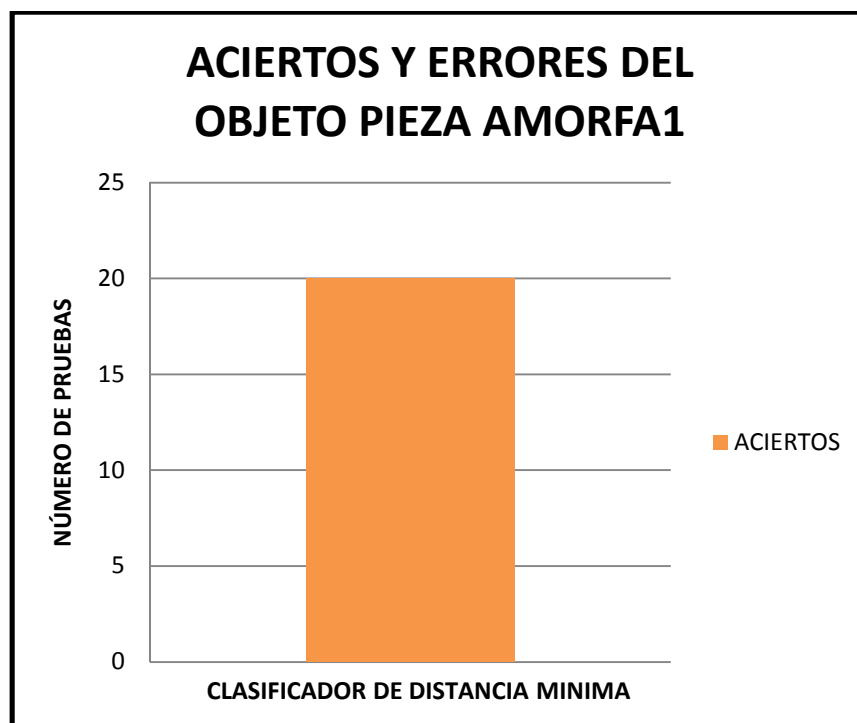


Fig 84. Aciertos y Errores del clasificador de distancia mínima para el objeto moneda

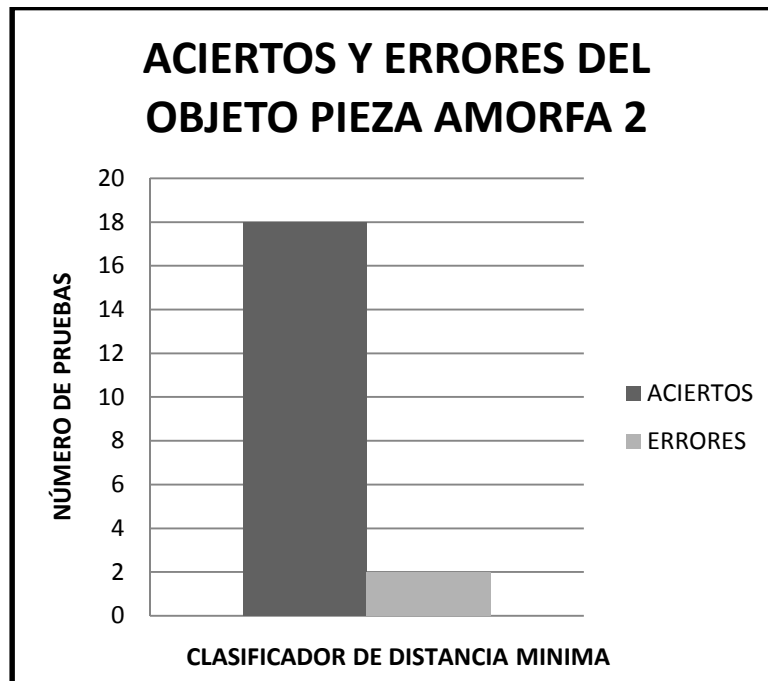


Fig 85. Aciertos y Errores del clasificador de distancia mínima para el objeto amorfa 2

De la misma manera se trabajo con pruebas con el clasificador de los vecinos más cercanos. Los resultados obtenidos fueron muy favorables. A continuación se muestran los resultados:

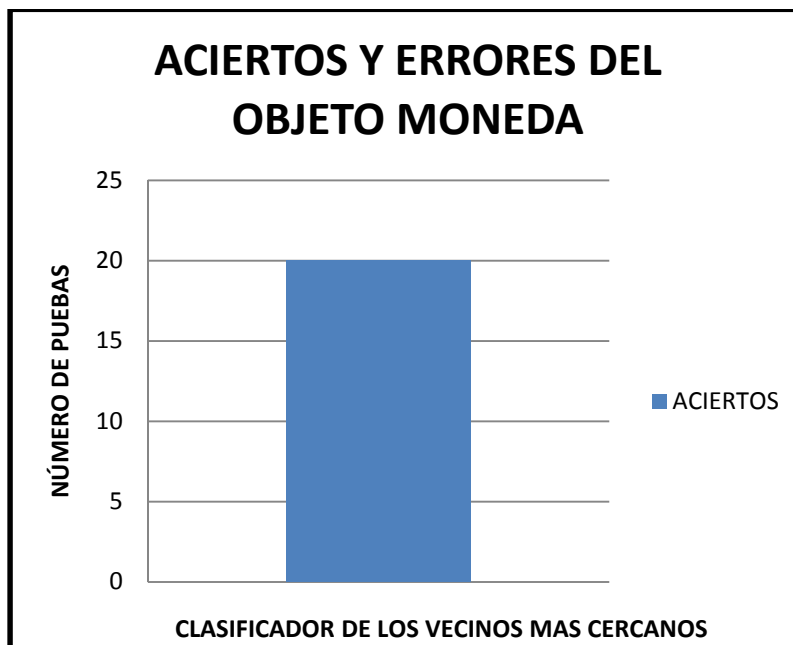


Fig 86. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto moneda

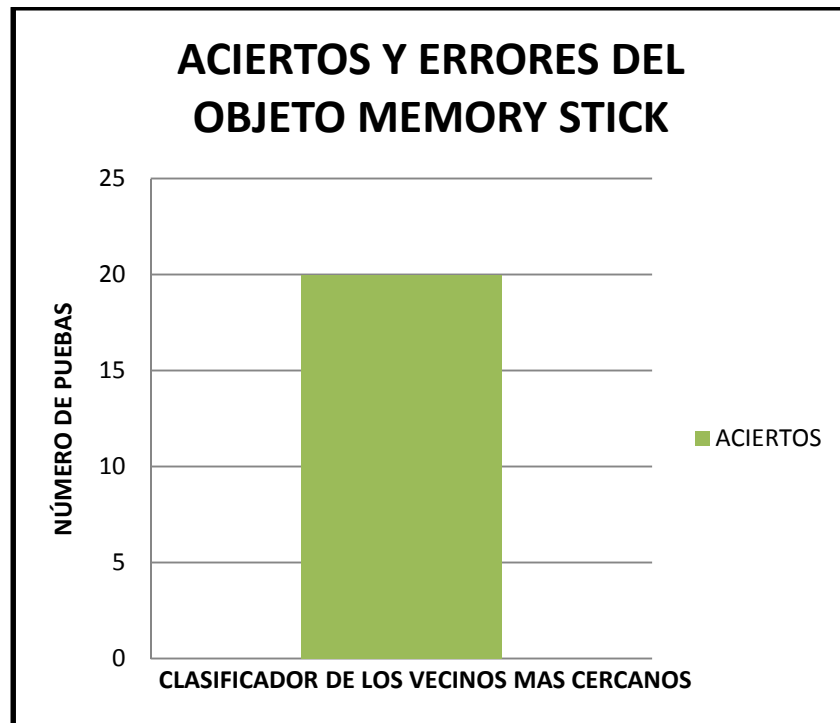


Fig 87. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto memory stick

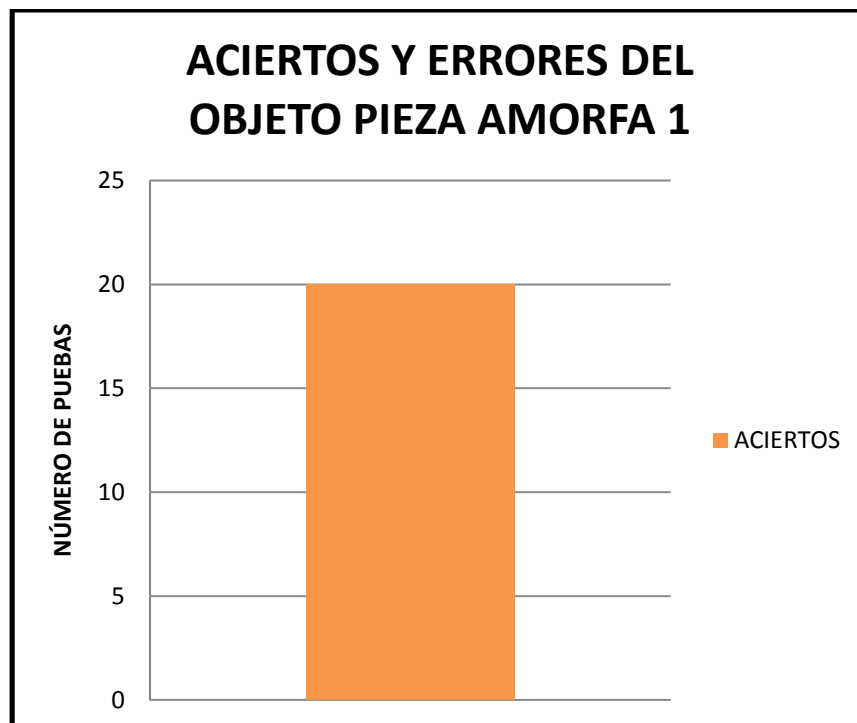


Fig 88. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto amorfa 1

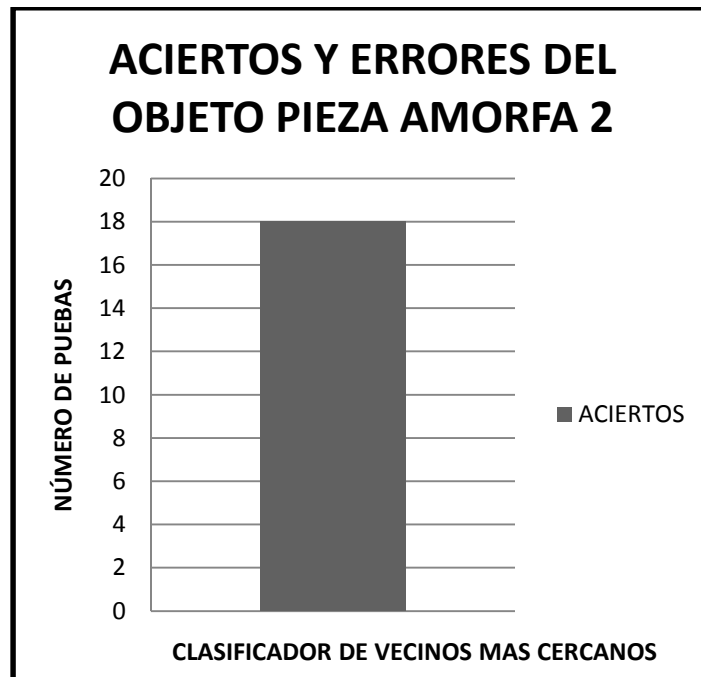


Fig 89. Aciertos y Errores del clasificador de los vecinos más cercanos para el objeto amorfa 2

Las graficas que se obtiene con el clasificador de lógica difusa tienen una gran diferencia con los vecinos más cercanos y distancia mínima. La mayor parte de las pruebas este se equivocaba al identificar los objetos. Estos se puede observar en los siguientes cuadros comparativos.

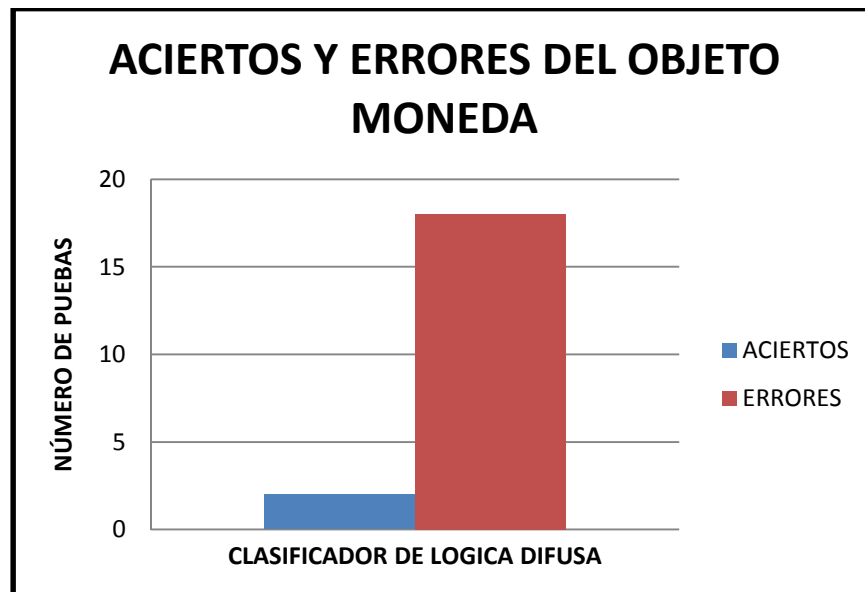


Fig 90. Aciertos y Errores del clasificador de lógica difusa para el objeto moneda

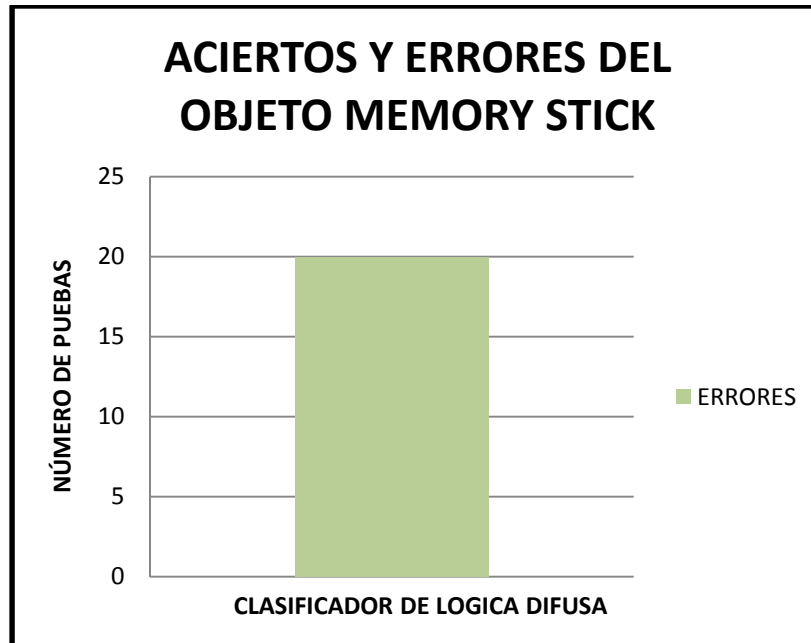


Fig 91. Aciertos y Errores del clasificador de lógica difusa para el objeto memory stick

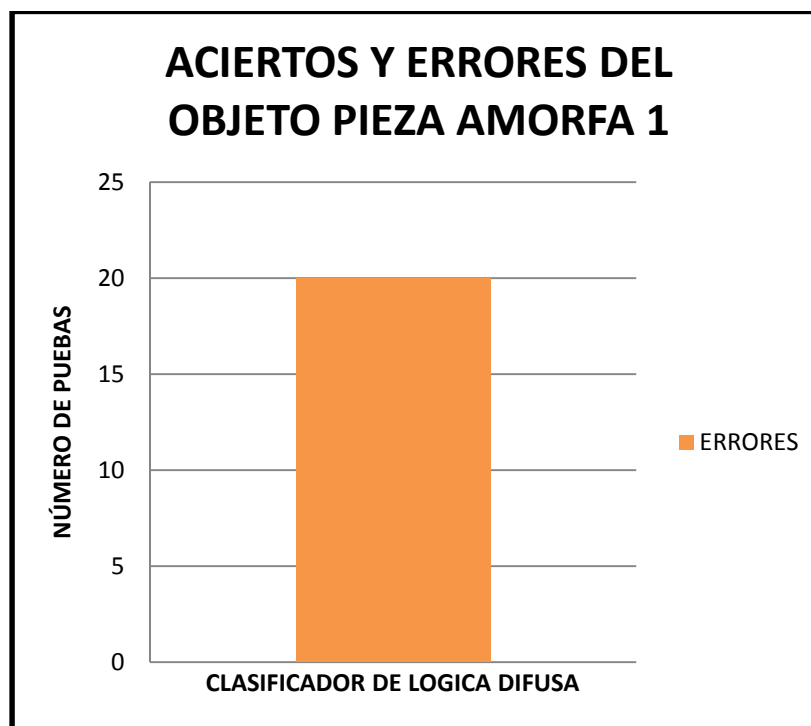


Fig 92. Aciertos y Errores del clasificador de lógica difusa para el objeto pieza amorfa 1

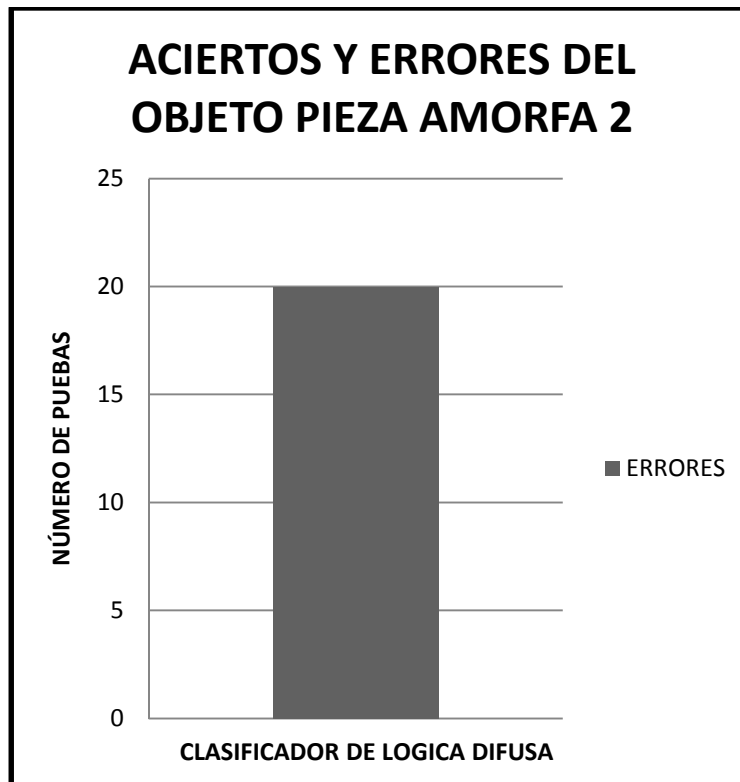


Fig 93. Aciertos y Errores del clasificador de lógica difusa para el objeto pieza amorfa2

De esta experimentación se puede decir que los momentos de Hu es una técnica bastante útil y sencilla de implementar ya que utilizando clasificadores adecuadas se puede lograr un trabajo bastante aceptable en cuanto al reconocimiento de patrones se refiere. Acotando también que el clasificador difuso no es bueno en algunos casos siendo este descartado cuando se vaya a realizar un trabajo de reconocimiento de objetos.

Estos serian los resultados obtenidos en la experimentación y estudio del presente trabajo.

CONCLUSIONES

Dentro de la electrónica tenemos un campo amplio denominada visión artificial. Este campo es muy útil para poder realizar distintas aplicaciones ya sea a nivel domestico como a nivel industrial, como es el control y clasificación de productos, control de calidad, verificación de medidas en las distintas piezas que se fabrica, etc. Tomando como referencia este aspecto hemos propuesto en esta tesis investigar un poco más a fondo la aplicación que se le puede dar a este campo implementando un sistema de clasificación basado en la técnica de los momentos de Hu más clasificadores que permitan realizar el objetivo planteado.

Como se pudo observar, varios fueron los aspectos que se tuvieron que considerar para lograr implementar el sistema, uno de ellos fue el saber cuáles son las etapas necesarios para poder obtener la correcta información de los objetos que queremos clasificar. Uno de las primeras etapas era la adquisición. Esta etapa constituye en un factor muy importante ya que si desde aquí no obtenemos la calidad necesaria de los objetos o la cámara o el factor de iluminación no está bien definido, no podremos obtener valores correctos lo que el procesamiento nos resultara falloso. Por otro parte una vez solucionada este factor procedemos a procesar la imagen como es la binarización de la misma, ya que así podremos obtener valores de 1 a 255 para un mejor y más rápido procesamiento.

Por otro lado, con todas las características descritas se pudo obtener los momentos de Hu. Estos momentos de Hu son muy útiles para la caracterización de los objetos ya que los mismos permiten obtener valores para luego clasificarlos y de esta forma saber que objeto es el que estamos analizando. Además de esto se debe tener muy en cuenta que estos momentos en nuestro caso son muy pequeños por lo que puede darse la confusión que no existe valores para ciertos de ellos. Lo que se debe hacer o lo que se hizo en nuestro caso es trabajar con 7 decimales para saber el valor exacto de los mismos. Con esto podríamos implementarlos en los respectivos clasificadores. Cabe además recalcar que si los valores en los demás momentos salían cero era por el tamaño del objeto analizado, ya que según esto

En cuanto a la parte de los clasificadores, se tuvo un muy buen resultado ya que el de distancia mínima y el de K vecinos no daba ningún error y clasificaba sin ningún problema. Pero en cuanto al clasificador de lógica difusa, se tuvo un problema que los objetos no eran reconocidos de manera adecuada. Esto se debió a que al ser los valores de los momentos de Hu muy pequeños no se podría definir la reglas de inferencia de manera adecuada por lo que había la confusión en su clasificación. Se trato de sacar el logaritmo de los valores y definir nuevas reglas, pero el resultado siguió siendo el mismo. Luego se intento trabajar con las áreas de las figuras para redefinir nuevamente las reglas, sin embargo no se obtuvo nada. Lo que nos lleva a concluir que este clasificador no es el adecuado cuanto se trabaja con valores pequeños y sobre todo muy parecidos.

Podemos además recalcar que los momentos de Hu y los clasificadores pueden ser aplicados en cualquier área, ya sea de la medicina, industria alimentaria, metálica, de cerámicas etc, ya que al ser una técnica que nos entrega valores y posteriormente podemos clasificarlos, podemos implementar en cualquiera de estas áreas sin problema alguno.

Un trabajo a futuro que se podría implementar es trabajar con los momentos de Hu aplicando clasificadores de redes bayesianas. Este tipo de clasificadores al tener un alto nivel de cálculo matemático y sobre todo de probabilidad y estadística, nos puede dar un porcentaje de precisión bastante alto, lo que sería aun más robusto el sistema. Además se podría dar nuevas aplicaciones a esta técnica, como es el caso de la clasificación de figuras pero tomando en cuanto su representación en 3d. Este tipo de aplicación sería muy bueno ya que aquí se podría considerar absolutamente todas las características del objeto en cuestión que se desee analizar.

Se puede concluir acotando que esta técnica nueva de visión artificial es robusta, a pesar de que existan otras técnicas, esta ofrece fiabilidad y su fácil implementación

BIBLIOGRAFIA

- [1]http://3.bp.blogspot.com/_uXlJEleijms/TAC_Lt-xhl/AAAAAAAAAEeQ/YPJz2lmkmDU/s1600/bernhard-plockhorst-entrada-de-jesus-en-jerusalen.jpg
- [2]http://bcs.fltr.ucl.ac.be/FE/18/ARIST/Aristote_fichiers/filippino_lippi.jpg
- [3]http://eisc.univalle.edu.co/materias/Vision_Computacional/material/Libro%20Vision%20Artificial/visiona.pdf
- [4] Digital Image Processing, Rafael C. Gonzalez, Richard E. Woods, Second edition
- [5]<http://tesis.pucp.edu.pe/repositorio/handle/123456789/68>
- [6]<http://biometrics.org/bc2009/presentations/wednesday/Alonso-Fernandez%20MR%2016%20Wed%20930%20AM%20-%20950%20AM.pdf>
- [7] Maestria en control y automatizacion-Control digital avanzado, Fundamentos del Control Difuso, Ing. Ismael Minchala, Msc
- [8] Reconocimiento de objetos con visión artificial a través de Vectores de características y distancias mínimas, Ismael Minchala Ávila.
- [9]<http://es.scribd.com/doc/58302048/85/Clasificador-de-k-vecinos-mas-cercanos>
- [10]<http://sci2s.ugr.es/docencia/algoritmica/Practicas%20-%20SI.pdf>
- [11]<http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t9knn.pdf>
- [12]<http://www.jmcpri.net/GLOSARIO/METALURGIA.HTM>
- [13]http://www.cs.st-andrews.ac.uk/~yumeng/yumeng-SIFTreport-5.18_bpt.pdf
- [14] Object Detection using Geometric Invariant Moment, American Journal of Applied Sciences
- [15]<http://www.neuronet.cl/egroupware/sitemgr/sitemgr-site/templates/neuronet/images/SeparacionClasesHiperPlano.JPG>
- [16] http://www.eltiempo.com.ec/fotos-cuenca-ecuador/ecuador/t1_1234404033.jpg