



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACIONES EN INGENIERÍA Y CIENCIAS
APLICADAS

“Diseño e implementación de un sistema de procesamiento digital de imágenes para mediciones cuantitativas en rostros, para el apoyo en la rehabilitación de pacientes con parálisis facial”

TESIS PROFESIONAL PARA OBTENER EL TÍTULO DE:
DOCTORADO EN INGENIERÍA Y CIENCIAS APLICADAS CON OPCIÓN
TERMINAL EN TECNOLOGÍA ELÉCTRICA

PRESENTA:

M.I.C.A. MARCO EDUARDO BUSTOS DE LA ROSA

Asesor: Dr. J Jesús Escobedo Alatorre.

Co-Asesor: Dr. José Antonio Marbán Salgado.

CUERVAVACA MORELOS

Diciembre 2019



UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



Instituto de
Investigación en
Ciencias
Básicas y
Aplicadas

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Defensa de Tesis de Posgrado en Ingeniería y Ciencias Aplicadas



"1919-2019: en memoria del General Emiliano Zapata Salazar"

Cuernavaca, Morelos, a 09 de agosto de 2019.

DR. ROSENBERG JAVIER ROMERO DOMÍNGUEZ
COORDINADOR DEL POSGRADO EN
INGENIERÍA Y CIENCIAS APLICADAS
P R E S E N T E

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS "Diseño e implementación de un sistema de procesamiento digital de imágenes para mediciones cuantitativas en rostros, para el apoyo en la rehabilitación de pacientes con parálisis facial" que presenta el alumno **Marco Eduardo Bustos de la Rosa**, para obtener el título de **Doctorado en Ingeniería y Ciencias Aplicadas** con opción terminal en **Tecnología Eléctrica**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
DR. ÁLVARO ZAMUDIO LARA	Aprobado	
DR. DIEGO SEURET JIMÉNEZ	Aprobado	
DRA. MARGARITA TECPOYOTL TORRES	Aprobado	
DR. JOSÉ ALBERTO HERNÁNDEZ AGUILAR (FCAel)	Aprobado	
DR. JOSÉ CRISPÍN ZAVALA DÍAZ (FCAel)	Aprobado	
DR. JOSÉ ANTONIO MARBÁN SALGADO	Aprobado	Jose Antonio MS
DR. J JESÚS ESCOBEDO ALATORRE	Aprobado	

PLAZO PARA LA REVISIÓN 20 DÍAS HÁBILES (A PARTIR DE LA FECHA DE RECEPCIÓN DEL DOCUMENTO)

NOTA. POR CUESTION DE REGLAMENTACIÓN LE SOLICITAMOS NO EXCEDER EL PLAZO SEÑALADO, DE LO CONTRARIO LE AGRADECEMOS SU ATENCIÓN Y NUESTRA INVITACIÓN SERÁ CANCELADA.

AGRADECIMIENTOS.

Primero que nada quiero agradecer a dios por dejarme vivir esta etapa que es terminar un Doctorado, a mis padres que sin escatimar en desvelos me cuidaron y educaron desde que tengo uso de razón, a mis maestros que día a día en el aula compartieron sus conocimientos conmigo, a mi hermano que siempre estuvo ahí apoyándome brindándome palabras de aliento para no caer, a mis amigos que siempre estuvieron en la buenas y en las malas tendiéndome la mano, a mí mismo por ser persistente y capaz.

Agradezco de manera muy especial a mi esposa **C.R.R**, que ha estado conmigo en este proceso, agradezco cada momento, apoyo y amor que me ha brindado y por ayudarme a ser mejor persona día con día, por ser el pilar de lo mucho o poco que hago.

Con un reconocimiento especial a Conacyt por haberme financiado una beca para poder estudiar este grado de Doctorado.

Y por último a mí jurado revisor porque con sus sabios consejos mi trabajo de investigación fue mejorando poco a poco.

A todos gracias por su apoyo.

RESUMEN

La medición de la evolución de pacientes con parálisis facial, es muy importante para los especialistas que atienden este trauma, ya que el poder mostrar resultados favorables ayudan a demostrar a sus pacientes que van mejorando y con esto no deprimirse ya que la depresión podría ser una causa para retardar su evolución.

En la rama de la medicina no hay aparato, metodología, o sistema que ayude a los especialistas a medir la evolución de pacientes con este trauma y mostrar resultados favorables es muy complicado, ya que lo hacen de manera cualitativa, es decir, de acuerdo a su percepción visual.

En este trabajo de investigación se desarrolló una metodología y un sistema de procesamiento digital de imágenes, para la medición de pacientes con el trauma antes mencionado.

El desarrollo de este trabajo comienza con el procesamiento digital de imágenes, donde se segmenta el rostro, es decir solo dejar la información que ocuparemos, seguido de la reconstrucción digital del rostro en base a la parte sin daño, detectar puntos en él y así poder determinar las distancias entre estos.

Para comprobar que el sistema funciona correctamente se pidió apoyo de un terapeuta que atiende a estos pacientes.

Con los resultados obtenidos se puede determinar que el sistema podría ser implementado a futuro.

ABSTRACT

The measurement of the evolution of patients with facial paralysis is very important for the specialists who attend this trauma, since being able to show favorable results help to show their patients that they are improving, and with this they do not get depressed since this could slow their evolution.

In the field of medicine there is no apparatus, methodology, or system that helps specialists measure the evolution of patients with this trauma, and showing favorable results is very complicated since they do it qualitatively, that is, according to their visual perception.

In this research work a methodology and a digital image processing system were developed for the measurement of patients with the trauma mentioned above.

The development of this work begins in the digital image process where the face is segmented, that is, just leave the information that we will occupy, reconstruct it completely based on the part without damage, detect points in it, and thus be able to determine the distances between these.

To verify that the system works correctly, support from a therapist who cares for these patients was requested.

With results obtained it can be determined that the system can be implemented in some future.

ÍNDICE GENERAL

ÍNDICE GENERAL.....	I
ÍNDICE DE TABLAS.....	IV
ÍNDICE DE FIGURAS	V
I ANTECEDENTES.....	8
1.1 Antecedentes.	8
1.2 Planteamiento del problema.	10
1.3 Justificación.....	11
1.4 Hipótesis.	11
1.5 Objetivo.....	12
1.6 Alcance.	13
1.7 Limitaciones.	13
II MARCO TEÓRICO.....	14
2.1 Procesamiento Digital de imágenes.	14
2.1.1 Tipos de procesamiento.	16
2.1.1.1 Procesamiento de Bajo Nivel.	16
2.1.1.1.1 Filtros.	19
2.1.1.2 Procesamiento de Medio Nivel.	30
2.1.1.2.1 Segmentación por el clasificador de Bayes.	32
2.1.1.2.2 Formula de Bayes aplicada a la segmentación supervisada.	33
2.1.1.2.3 Detección de bordes.	35
2.1.1.2.4 Técnicas actuales de segmentación.....	38
2.1.1.3 Procesamiento de Alto Nivel.	38
2.1.1.3.1 Modelos en dos dimensiones.	42
2.1.1.3.2 Polilíneas.	42

2.1.1.3.3 Descripción.	42
2.1.1.3.4 Obtención de características del objeto por color.	44
2.1.1.3.5 Modelo sensorial.	45
2.1.1.3.6 Modelo perceptual.	45
2.1.1.3.7 Extracción de puntos.	47
2.1.1.3.8 Puntos característicos a partir de aristas: puntos de fuga.	47
2.1.1.3.9 Puntos característicos a partir de la propia imagen: puntos esquina.	49
2.1.1.3.10 Extracción de líneas.	50
2.1.1.3.11 Transformada de Hough para líneas.	52
2.1.1.3.12 Extracción de círculos.	53
2.1.1.3.13 Reconocimiento.	54
2.1.1.3.14 Clasificador “K Vecinos más Próximos”	54
2.1.2 Ejemplos de campos de aplicación de Procesamiento Digital de Imágenes.	56
2.1.2.1 Imágenes de rayos gamma.	57
2.1.2.2 Imágenes de rayos X.	60
2.1.2.3 Imágenes en la banda ultravioleta.	62
2.1.2.4 Imágenes en las bandas visibles e infrarrojas.	63
2.1.2.5 Imágenes en la banda de radio.	64
2.2 Lógica Difusa.	65
2.2.1 Valores Lingüísticos o Conjuntos difusos.	66
2.2.1.1 Definición de conjunto difuso.	68
2.2.1.2 Operaciones de Conjuntos Difusos.	69
2.2.1.3 Propiedades.	73
2.2.1.4 Representación de conjuntos difusos.	73
2.2.2 Variables Lingüísticas.	74

2.2.3 Aplicaciones.....	75
2.2.3.1 Ejemplos de aplicaciones.....	77
2.3 Parálisis facial.....	82
2.3.1 Tipos de parálisis.....	83
2.3.2 Tratamiento.....	87
2.3.3 Evolución.....	89
III DESARROLLO EXPERIMENTAL.....	90
3.1 Sensor Kinect.....	90
3.1.2 Especificaciones.....	91
3.1.3 Funcionamiento.....	92
3.2 Técnica de Proyección de franjas.....	98
3.2.1 Metodología.....	98
3.2.2 Obtención de imágenes Tridimensionales (3D).....	100
3.2.3 Imágenes Tridimensionales (3D) adquiridas con la técnica de proyección de franjas.....	103
3.3 Procesamiento Digital de imágenes y Lógica Difusa.....	105
3.3.1. Metodología a seguir para el Procesamiento Digital de Imágenes. ...	105
3.3.2. Lógica Difusa.....	106
IV RESULTADOS.....	110
4.1 Discusión.....	122
V CONCLUSIONES FINALES.....	124
5.1 Recomendaciones.....	125
REFERENCIAS.....	126
ANEXOS.....	133

ÍNDICE DE TABLAS

Tabla 2.1 Medidas de una Persona.....	67
Tabla 2.2 Base de Datos tradicional del Ejemplo.....	80
Tabla 2.3 Resultados de la búsqueda del ejemplo en cada registro.....	81
Tabla 3.1 Dato de profundidad en espacio de color RGBA.....	96
Tabla 3.2 Dato de profundidad en espacio de color YUV.....	96
Tabla 4.1 Posibles resultados.....	115

ÍNDICE DE FIGURAS

Figura 2.1 Histograma de una imagen.....	18
Figura 2.2 Tipos de histograma de una imagen.....	18
Figura 2.3 Histograma de una imagen digital.....	19
Figura 2.4 Imagen original y con ruido.....	21
Figura 2.5 Cambios en la señal de una imagen.....	22
Figura 2.6 Señal original y sobre-muestreada.....	23
Figura 2.7 Transformaciones lineales.....	25
Figura 2.8 Selección de propiedades.....	27
Figura 2.9 Selección individual.....	28
Figura 2.10 Esquema que representa el proceso de segmentación.....	31
Figura 2.11 Reconocimiento de caracteres en base a su codificación.....	39
Figura 2.12 Estructura de un sistema de visión basado en modelos.....	41
Figura 2.13 Obtención de características del objeto por color.....	44
Figura 2.14 Intersección de rectas.....	47
Figura 2.15 Método de la minimización del área del triángulo.....	49
Figura 2.16 Clasificador KNN.....	55
Figura 2.17 El espectro electromagnético dispuesto según la energía por fotón.....	57
Figura 2.18 Radiografía.....	58
Figura 2.19 Tomografía por emisión de positrones.....	59
Figura 2.20 Nube de gas estacionaria sobrecalentada.....	59
Figura 2.21 Radiación gamma (punto brillante) de una válvula del reactor.....	60
Figura 2.22 Radiografía de pecho.....	61
Figura 2.23 Angiografía aórtica.....	62
Figura 2.24 Maíz normal y Tizón de maíz.....	63
Figura 2.25 Taxol (agente anticancerígeno), magnificado 250x, Colesterol 40x y Microprocesador 60x.....	64

Figura 2.26 Imagen de una rodilla y una columna vertebral.....	65
Figura 2.27 a) Descripción de conjuntos crisp, y b) Fuzzy de una persona alta.....	67
Figura 2.28 Uso del modificador muy en los conjuntos bajo y alto.....	72
Figura 2.29 Estructura interna de un Controlador Difuso.....	77
Figura 2.30 Variables lingüísticas para el ejemplo.....	80
Figura 3.1 Sensor Kinect.....	90
Figura 3.2 Proyección de puntos sobre una superficie.....	93
Figura 3.3 Cada uno puede ser identificado de otro dentro del patrón proyectado.....	94
Figura 3.4 Imagen de profundidad capturada con la cámara.....	94
Figura 3.5 Δx es proporcional para el punto proyectado.....	95
Figura 3.6 a) Imagen RGB y b) Imagen profundidad.....	96
Figura 3.7 Imagen RGB encima de la de profundidad.....	97
Figura 3.8 Montaje experimental de proyección de franjas.....	99
Figura 3.9 Patrón de franjas proyectado en a) un plano de referencia y b) Una mano Humana.....	99
Figura 3.10 Patrón de franjas con corrimiento de fase $\phi_i = \pi/2$	101
Figura 3.11 Imagen de la fase (a) discontinua y (b) continúa.....	102
Figura 3.12 Imagen de un rostro 2D. (maniquí).....	103
Figura 3.13 Imagen de un rostro 3D. (maniquí).....	104
Figura 3.14 Perfil de un rostro 3D (maniquí).....	104
Figura 3.15 Valores lingüísticos que toma la $dist_1$	107
Figura 3.16 Valores lingüísticos que toma la $dist_2$	108
Figura 3.17 Valores lingüísticos que toma la $dist_3$	108
Figura 4.1 Rostro segmentado a) Imagen virtual b) Imagen real de un maniquí.....	110
Figura 4.2 Rostro reconstruido virtualmente a) Imagen virtual b) Imagen real de un maniquí.....	111

Figura 4.3 Puntos característicos a) Imagen virtual	
b) Imagen real de un maniquí.....	112
Figura 4.4 Puntos característicos detectados a) Imagen virtual	
b) Imagen real de un maniquí.....	113
Figura 4.5 Pantalla principal del Sistema.....	117
Figura 4.6 Registro de pacientes.....	118
Figura 4.7 Medición de la Evolución (terapia 1).....	119
Figura 4.8 Medición de la Evolución (terapia 2).....	120
Figura 4.9 Medición de la Evolución (terapia 3).....	121

I ANTECEDENTES.

1.1 Antecedentes.

- Se propone dividir el algoritmo de mascara Candide-3 en dos grupos uno que afecte el lado izquierdo de la máscara y otro afecte al lado derecho independientemente, donde toma en cuenta puntos clave, curvas y color de piel es así como determina la asimetría del rostro también toma la habilidad de mover ambos lados del rostro, teniendo esto en cuenta le asigna graduación y con esto mide el grado de parálisis facial siguiendo el sistema simplificado de graduación de parálisis del nervio facial de House-Brackmann donde aplica una metodología cualitativamente donde dice que grado de parálisis facial tiene ya sea: neuroplaxia, leve debilidad-axonotmesis, debilidad moderada-axonotmesis, evidente desfiguración y debilidad-neurotmesis, corte transversal parcial y corte transversar total (Gálvez 2017).
- Desarrolló un algoritmo que evalúa el grado de parálisis facial con el uso del Kinect para la extracción y la simetría de un rostro y las distancias entre las cejas a infra orbital y la distancia entre los pómulos y la parte lateral del ojo, teniendo las distancias aplica el sistema de clasificación de Nottingham donde cuantifica el grado de parálisis y da una escala de tres niveles: normal, leve y severo (Carro et al. 2016).
- Desarrolló un sistema óptico 3D automático que implica colocar puntos de reflejos especiales en la cara del paciente pidiéndole que haga ciertos movimientos en cara como sonreír, fruncir la boca, cerrar los ojos y la elevación de la frente, Las imágenes de las cámaras se procesan automáticamente con un software que genera información personalizada como puntos, ángulos, áreas vectores y velocidades entre áreas faciales para así poder medir el

movimiento facial con las distancias entre los puntos obtenidos (Hontanilla et al. 2007).

- Desarrolló un sistema computarizado para el reconocimiento de características faciales entre ellas las pupilas de los ojos y las comisuras, dicho sistema utiliza técnicas de procesamientos de imágenes para ser más específico la transformada de Hough, ya que estas características han sido localizadas se calcula la distancia entre ellas (Sucar et al. 2004).

- Presenta un nuevo marco para la medición objetiva de la parálisis facial. La Información de movimiento en las direcciones horizontal y vertical y las características de apariencia en los marcos de vértice se extraen en función de los patrones binarios locales (LBP) en el dominio temporal - espacial en cada región facial. Estas características se mejoran temporal y espacialmente mediante la aplicación de nuevos esquemas de procesamiento de bloques. Se propone una extensión multiresolución de LBP uniforme para combinar de manera eficiente los micros patrones y los patrones a gran escala en un vector de características. La simetría de los movimientos faciales se mide por la distancia promedio de la resistencia RAD) entre las características de LBP extraídas de los dos lados de la cara. La máquina de vectores de soporte se aplica para proporcionar una evaluación cuantitativa de la parálisis facial basada en la escala de House-Brackmann (He et al, 2009).

- Desarrollaron un sistema de medición simple, que usa equipos disponibles comercialmente, toma poco tiempo y proporciona mediciones significativas y precisas. Esta técnica se denomina sistema de medición de la reanimación facial. Implica colocar puntos alrededor de la boca del paciente y grabar en video al paciente realizando sonrisas de máximo esfuerzo. Usando un

programa de edición de video, se superpone un cuadro que muestra al paciente en reposo con un segundo cuadro que muestra la sonrisa del paciente. Esta imagen superpuesta se importa a Adobe PhotoShop, donde las mediciones se obtienen mediante las herramientas disponibles en el programa. Se utilizaron veinte pacientes para probar la confiabilidad entre evaluadores e intrarrateriales del sistema de medición de reanimación facial. La precisión del proceso de medición se probó comparando 10 distancias y ángulos conocidos con los obtenidos. Utilizando el sistema de medición de la reanimación facial. Tanto la fiabilidad de la distancia como del ángulo, tanto interna como externa, son altamente precisas, con correlaciones intraclase mayores a 0.9. El sistema de medición de reanimación facial tiene una precisión de 0,6 mm y 2,0 grados en comparación con una distancia y ángulo "conocidos". El sistema de medición de la reanimación facial se ha utilizado para medir los movimientos de la sonrisa de más de 200 pacientes y se ha demostrado que es valioso para detectar cambios en los movimientos faciales a lo largo del tiempo (Tomat et al. 2005).

1.2 Planteamiento del problema.

No existe una cura, tratamiento estándar o un aparato que mida la evolución de pacientes con la parálisis facial, el factor más importante es disminuir o eliminar la fuente del daño. Afecta a cada individuo de manera diferente, en algunos casos son leves y no necesitan tratamiento ya que generalmente los síntomas se disuelven por si solos en dos semanas, mientras que para otros casos se aplica un tratamiento médico y una terapia muscular, para estos casos no hay mejoría, ya que de acuerdo al especialista los pacientes se deprimen al no percibir mejoría.

1.3 Justificación.

Uno de los grandes problemas en la rehabilitación de parálisis facial es garantizar que la terapia sugerida sea la adecuada al problema particular del paciente.

Otra inquietud del ¿por qué? de este proyecto, es que en la medicina no se cuenta con un aparato que mida la evolución de este tipo de pacientes, los especialistas miden la evolución de manera cualitativa es decir: de acuerdo a su percepción visual, es por esto que surge la necesidad de crear este sistema para hacer este tipo de mediciones cuantitativas, ya que con el sistema y la metodología aplicada en este trabajo, el especialista no solo podrá determinar si de verdad el paciente está mejorando o no, sino que además también podrá darse cuenta en que área poner atención, si es que no hay una mejora, o en su defecto poder canalizarlo con otro especialista, es decir, que sea un tema psicológico, si este llegara a ser el caso el motivo principal suele ser que paciente se deprime y en su percepción no ve mejoría, pero esta herramienta puede ser de gran apoyo para los especialistas y demostrarle de manera cuantitativa su progreso, además que si el especialista lo quisiera implementar es de bajo costo ya que el sistema no requiere la implementación de muchos accesorios.

1.4 Hipótesis.

Se puede medir la evolución en el avance de la rehabilitación de un paciente con parálisis facial, mediante la comparación de datos cuantitativos de dos imágenes pertenecientes a diferentes instantes en la recuperación de un paciente.

1.5 Objetivo.

Diseñar e implementar un sistema de medición cualitativa y cuantitativa, mediante el uso de procesamiento de imágenes, para medir la evolución de pacientes con parálisis facial.

Objetivos específicos.

- Desarrollar un algoritmo que reconstruya virtualmente el rostro completo, tomando como base el lado del rostro sin daño.
- Desarrollar un algoritmo que detecte puntos característicos necesarios en la imagen del rostro de un paciente, para poder realizar correctamente las mediciones.
- Desarrollar un algoritmo que mida de forma cuantitativa la diferencia entre imágenes pertenecientes a diferentes instantes en la recuperación del paciente (medir evolución).
- Diseñar e implementar un aparato para el proceso de adquisición de imágenes.
- Probar el funcionamiento del sistema completo.

1.6 Alcance.

En la actualidad no existe un sistema que mida cuantitativamente la evolución de pacientes con parálisis facial, la implementación de un sistema como este no es alto costo y pudiera abrir una puerta en al ámbito de la medicina.

En este proyecto se desarrollará un sistema de medición de pacientes con parálisis facial, además se proporcionara una metodología para llevar a cabo dicha cuestión. Esta se basa en una serie de reglas que ayudan al especialista a poder decir si el paciente va mejorando o no, permitiendo además saber en qué puntos del rostro poner más atención para poder sugerir una terapia.

1.7 Limitaciones.

Uno de los grandes problemas que presenta este proyecto es que no se puede comprobar su funcionalidad con pacientes reales, lo que se hizo fue simulaciones en un rostro artificial y al lado de un especialista comprobar que dichas simulaciones mencionadas fueran correctas.

II MARCO TEÓRICO.

2.1 Procesamiento Digital de imágenes.

Para (González y Woods. 2008). El procesamiento digital de imágenes consiste en procesar una imagen digital por medio de un ordenador. Estos mismos autores destacan que el interés en los métodos de procesamiento de imagen digital se deriva de dos áreas de aplicación principales:

- La mejora de la información gráfica para la interpretación humana.
- El procesamiento de datos de la imagen para almacenamiento, transmisión y representación por la máquina. En este caso, la información es muy distinta a los rasgos empleados por las personas para interpretar una imagen.

Una imagen se define como una función bidimensional, $f(x,y)$, donde x e y son coordenadas espaciales del plano. Además, la amplitud de f para cualquier par de coordenadas (x,y) , se llama intensidad (color) o nivel de gris de una imagen en ese punto. Cuando los valores de intensidad de *la función* (f), así como x e y , son valores finitos y discretos (Vettorazzi, 2007), se habla de imagen digital. Así, una imagen digital está compuesta por un número finito de elementos, donde cada uno de estos elementos se llama píxel. Dicho de otro modo, una imagen digital puede representarse como una matriz donde los índices de las filas y columnas (x e y) determinan un punto concreto de la imagen y, el valor de ese punto ($f(x,y)$) representa su nivel de gris o su color (González y Woods, 2008).

Por otro lado, un píxel indica la ubicación (x,y) contando desde el origen de la imagen. La palabra píxel es una abreviatura de elemento de imagen, representando este concepto el elemento más pequeño constituyente de una Imagen digital. Además, contiene un valor numérico que es la unidad básica de información dentro de la imagen (Solomon y Breckon, 2011).

En cuanto a los tipos de imágenes, las imágenes binarias o en escala de grises (intensidad), se representan como matrices bidimensionales que asignan un valor numérico a cada píxel que es, como se ha comentado, representativo de la intensidad en ese punto. Dichas matrices utilizan un espacio de color de 2 bits (binario) o de intensidad (escala de grises). Por el contrario, las imágenes RGB (o de color verdadero) se representan como matrices tridimensionales que asignan tres valores numéricos a cada píxel, correspondiendo cada valor al componente rojo (R), verde (G) y azul (B). Este último tipo de imágenes son las que se analizan en este trabajo. Tales imágenes se pueden considerar conceptualmente como tres planos bidimensionales distintos, uno correspondiente a cada uno de los tres canales de color (Solomon y Breckon, 2011).

En los sistemas de procesamiento de imagen digital, se trata con matrices de números obtenidos mediante puntos de muestreo espacial de una imagen física. Después del procesamiento, se produce otra matriz de números y estos números, se utilizan para reconstruir una imagen continua para su visualización (Pratt, 2007).

Un aspecto fundamental del procesamiento de imágenes y que demuestra el amplio y variado campo de aplicaciones que abarca, es la capacidad de las máquinas para trabajar con imágenes generadas por fuentes que los seres humanos no asocian con imágenes tales como, ultrasonidos o microscopia electrónica. Este hecho es debido a que las máquinas de imágenes cubren casi todo el espectro electromagnético (desde ondas gamma a ondas de radio), a diferencia de los seres humanos que se limitan a la banda visual del espectro electromagnético (Gonzalez y Woods, 2008).

2.1.1 Tipos de procesamiento.

A continuación se presentan los tipos de niveles de procesamiento digital de imágenes.

2.1.1.1 Procesamiento de Bajo Nivel.

El procesamiento de bajo nivel implica operaciones primitivas como el pre-procesamiento de imágenes para reducir el ruido, la mejora del contraste y la nitidez de la imagen caracterizándose por el hecho de que sus entradas y salidas son imágenes (González y Woods, 2008), la cual no es siempre necesario llevar a cabo. Sin embargo, en muchas aplicaciones esta etapa tiene un peso importante y básicamente consiste en la atenuación del ruido presente en la imagen, el mejoramiento del contraste, así como filtrados de enfoque, entre otros (Rodríguez y Sossa, 2012).

Los procesos de realce de imágenes consisten de una serie de técnicas cuyo objetivo es mejorar la apariencia visual de una imagen, ya sea en contraste, ruido, escala de grises, distorsiones, luminosidad, falta de nitidez, etc., o bien convertir o mapear la imagen a una mejor forma para su análisis. El principal objetivo del realce de la imagen es procesar una imagen de tal manera que el resultado obtenido sea el apropiado para una aplicación específica. Los métodos de realce de imágenes se pueden dividir en dos categorías: los métodos de realce en el dominio espacial y los métodos de realce en el dominio de la frecuencia. Los métodos de la primera categoría consisten en la

manipulación directa de los píxeles de la imagen mientras que los métodos de la segunda categoría corresponden a técnicas basadas en la representación de los píxeles, a través de una transformación hacia el dominio de la frecuencia y usa como operador de mapeo o transformación a la transformada de Fourier (DFT).

Dentro de los métodos de realce espacial se pueden encontrar dos técnicas: El realce radiométrico donde las operaciones son efectuadas directamente sobre un píxel sin importar o tomar en cuenta a los píxeles vecinos. Sirve para mejorar condiciones de bajo contraste, baja luminosidad o demasiada oscuridad. Ejemplo: ecualización de histograma.

El realce con operaciones entre vecinos que, a diferencia del realce radiométrico, las operaciones son efectuadas sobre un píxel pero tomando en cuenta a los píxeles que lo rodean. Sirve para eliminar ruido o para el mejoramiento de la nitidez. Ejemplo: kernel correspondiente a un filtro paso-bajas usando la convolución para realizar un filtrado espacial.

El histograma de una imagen es la representación gráfica de la distribución que existe de las distintas tonalidades de grises con relación al número de píxeles o porcentaje de los mismos. La representación de un histograma ideal sería la de una recta horizontal, ya que eso indicaría que todos los posibles valores de grises están distribuidos de manera uniforme en la imagen (Escalante, 2006).

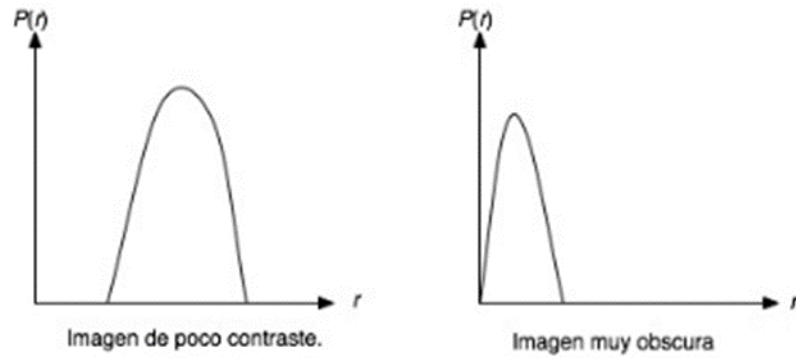


Figura 2.1 Histograma de una imagen. (Escalante, 2006)

La figura 2.1 muestra gráficamente la diferencia entre una imagen con poco contraste y una imagen oscura.

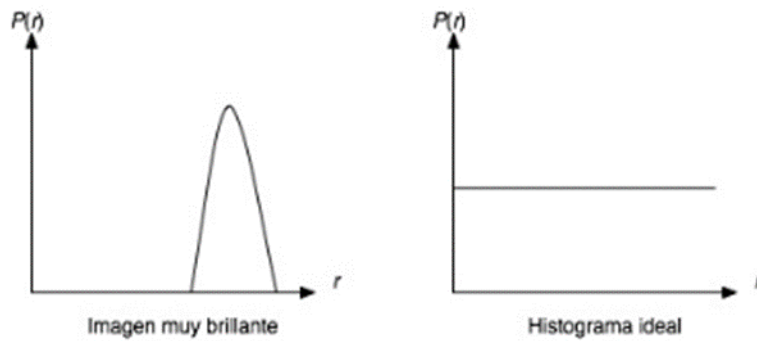


Figura 2.2 Tipos de histograma de una imagen. (Escalante, 2006)

En la figura 2.2 se muestra gráficamente la diferencia entre una imagen con mucho brillo y el histograma ideal para una imagen.

El histograma de una imagen digital con niveles de gris en el rango $[0, L - 1]$ es una función discreta $P(r) = nr$ donde r representa un nivel de gris y nr representa el número de píxeles que tienen ese valor de gris como se muestra en la figura 2.3.

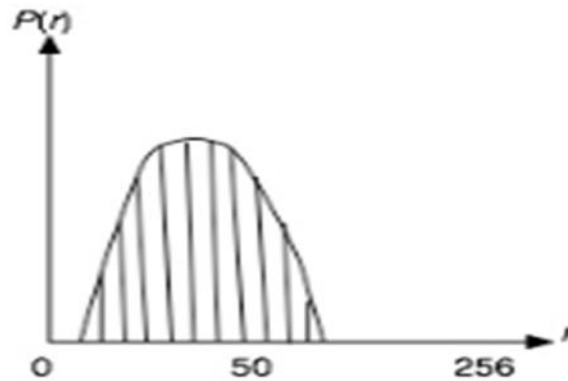


Figura 2.3 Histograma de una imagen digital (Escalante, 2006)

Es muy frecuente normalizar un histograma dividiendo cada uno de sus valores por el número total de píxeles en la imagen, denotado por n . De tal modo que el histograma nos quedaría $P(r_k) = nr_k/n$, para $k = 0, 1, \dots, L-1$. De esta manera, $P(r_k)$ da una estimación de la probabilidad de la ocurrencia del nivel de gris r_k . Hay que notar que la suma de todos los componentes de un histograma normalizado es igual a 1.

2.1.1.1.1 Filtros.

De acuerdo (González y Woods, 2002) los filtros espaciales tienen como objetivo modificar la contribución de determinados rangos de frecuencias de una imagen. El término espacial se refiere al hecho de que el filtro se aplica directamente a la imagen y no a una transformada de la misma, es decir, el nivel de gris de un píxel se obtiene directamente en función del valor de sus vecinos.

Los filtros espaciales pueden clasificarse basándose en su linealidad: filtros lineales y filtros no lineales. A su vez los filtros lineales pueden clasificarse según las frecuencias que dejen pasar: los filtros pasa bajas atenúan o eliminan las componentes de alta frecuencia a la vez que dejan inalteradas las bajas frecuencias; los filtros pasa altas atenúan o eliminan las componentes de baja frecuencia con lo que agudizan las componentes de alta frecuencia; los filtros pasa banda eliminan regiones elegidas de frecuencias intermedias.

La convolución es una operación por la cual se lleva a cabo una acción de filtrado. Como veremos más adelante y como se ha visto en las propiedades de la Transformada de Fourier, existe una relación entre el filtrado espacial y el filtrado en el dominio de la frecuencia. En general, el filtrado de señales posee tres categorías, según el resultado que se busque:

1) Filtros Pasa-Bajas: Son utilizados en la reducción de ruido; suavizan y aplanan un poco las imágenes y como consecuencia se reduce o se pierde la nitidez. En inglés son conocidos como Smoothing Spatial Filters.

2) Filtros Pasa-Altas: Estos filtros son utilizados para detectar cambios de luminosidad. Son utilizados en la detección de patrones como bordes o para resaltar detalles finos de una imagen. Son conocidos como Sharpening Spatial Filters.

3) Filtros Pasa-Banda: Son utilizados para detectar patrones de ruido. Ya que un filtro pasa-banda generalmente elimina demasiado contenido de una imagen casi no son usados. Sin embargo, los filtros pasa-banda son útiles para aislar los efectos de ciertas bandas de frecuencias seleccionadas sobre una imagen. De esta manera, estos filtros ayudan a simplificar el análisis de ruido.

Suavizado:

De acuerdo con (Escalante, 2006) los filtros paso-bajas son utilizados para difuminar y reducir ruido en las imágenes, a este proceso se le conoce en inglés como smoothing. El difuminado (blurring) es usado en etapas de pre procesamiento, desde la eliminación de pequeños detalles, hasta la extracción de objetos y relleno de pequeños huecos en líneas y curvas. La reducción de ruido puede ser completada por el difuminado usando filtros lineales o bien con un filtrado no lineal, como se muestra en la figura 2.4.

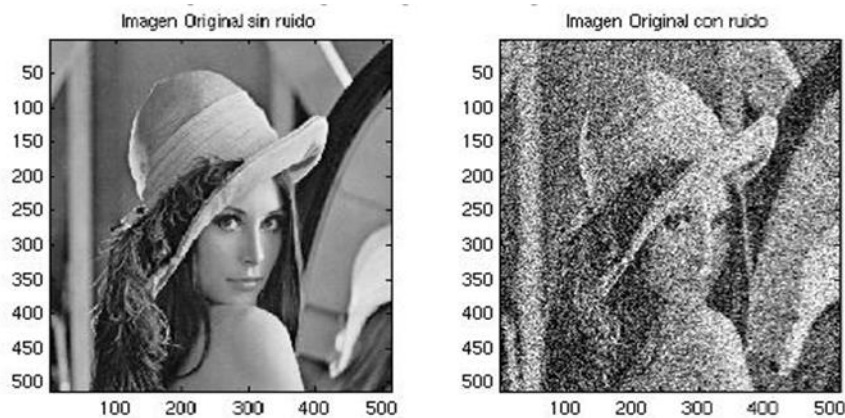


Figura 2.4 Imagen original y con ruido. (Escalante, 2006)

La salida (respuesta) de un filtro pasa-bajas lineal, simplemente es un tipo de promedio de los pixeles contenidos en la vecindad de la máscara del filtro. Estos filtros son frecuentemente llamados filtros promediadores. La idea detrás de estos filtros es la de reemplazar cada pixel en la imagen por un promedio de los niveles de gris de los vecinos definidos por la máscara del filtro.

Los filtros basados en derivadas de Gaussianas fueron en un principio obtenidos de manera heurística. Sin embargo estos filtros tienen un fundamento matemático y un comportamiento muy similar al sistema de visión humano (Human Vision System); por ello son muy importantes ya que se especializan en la detección de cambios bruscos como los bordes.

El mejoramiento de la nitidez o de la calidad visual de una imagen basado en los filtros unsharp masking, tiene bastante importancia en el procesamiento digital de imágenes. Si se quisiera traducir al español, unsharp masking podría interpretarse como enmascaramiento (masking) de imagen borrosa (unsharp). Se basa en el hecho de que se tiene una imagen borrosa (con pendiente pequeña), y se le resta una pendiente aún más pequeña. A su vez esto va multiplicado por un factor k el cual se recomienda tome valores entre 1 y 3 como se muestra en la figura 2.5.

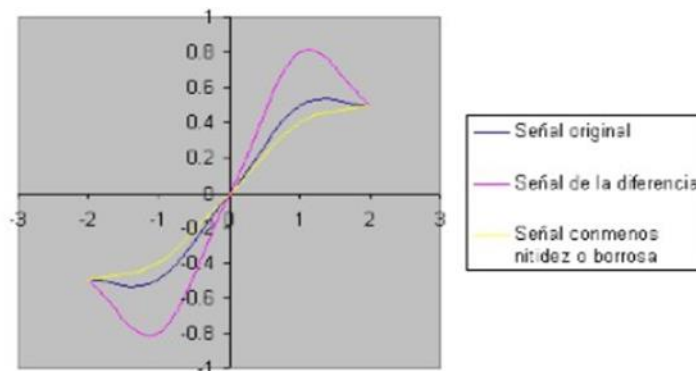


Figura 2.5 Cambios en la señal de una imagen. (Escalante, 2006)

El proceso de interpolación está relacionado con el hecho de desear cambiar la tasa de muestreo de una señal discreta, a una tasa menor o mayor con la que fue muestreada originalmente. El proceso con el cual se disminuye la tasa de muestreo de una señal discreta por un factor D se denomina decimación o submuestreo (downsampling) y el proceso con el cual se aumenta la tasa de muestreo de una señal discreta por un factor T se denomina interpolación (upsampling).

Una de las formas más simples de interpolación es elegir la amplitud de un pixel de salida como la amplitud de su vecino más cercano. Está interpolación es llamada interpolación del vecino más cercano o bien interpolación de orden cero. El factor de interpolación T indicará cuantos pixeles se desean interpolar. Por ejemplo, en una dimensión podríamos ver lo siguiente, se tiene una señal compuesta por cuatro muestras y se desea interpolar con un filtro $h(x)$ de orden cero con un factor de interpolación $T = 2$ como se muestra en la figura 2.6.

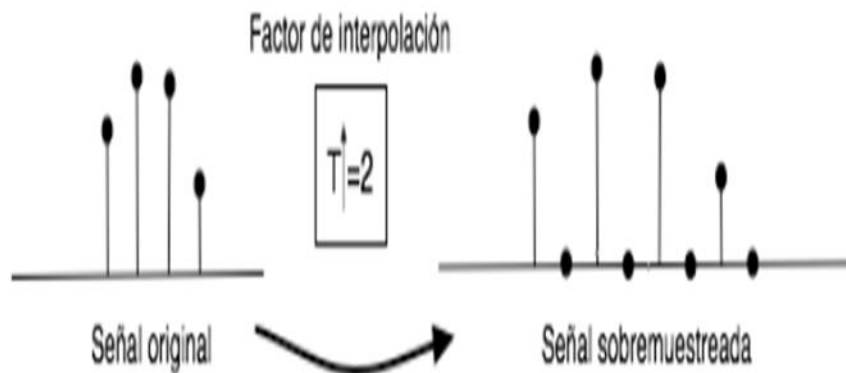


Figura 2.6 Señal original y sobre-muestreada (Escalante, 2006)

La señal interpolada será el resultado de hacer la convolución de la señal original sobre muestreada (agregando ceros), con el filtro interpolador.

Transformaciones de intensidad

De acuerdo con (Sucar y Gomez, 2016) una transformación de intensidad consiste en mapear los valores de intensidad de cada pixel a otros valores de acuerdo a cierta función de transformación. Las funciones de transformación pueden ser de dos tipos:

1. Lineales.
2. No-lineales.

En las transformaciones lineales, se tiene una relación o función lineal de los valores de intensidad de los pixeles de la imagen de salida respecto a la imagen de entrada. Los tipos de transformaciones lineales más comúnmente utilizados son:

- Obtener el negativo de una imagen.
- Aumentar o disminuir la intensidad (brillo de la imagen).
- Aumento de contraste

Las funciones de transformación para cada uno de estos tipos se especifica gráficamente en la figura 2.7. Por ejemplo, para el negativo, el pixel de entrada (eje X) de intensidad 0 se transforma en un pixel de salida (eje Y) de intensidad máxima, y el de entrada de intensidad máxima se transforma en intensidad 0.

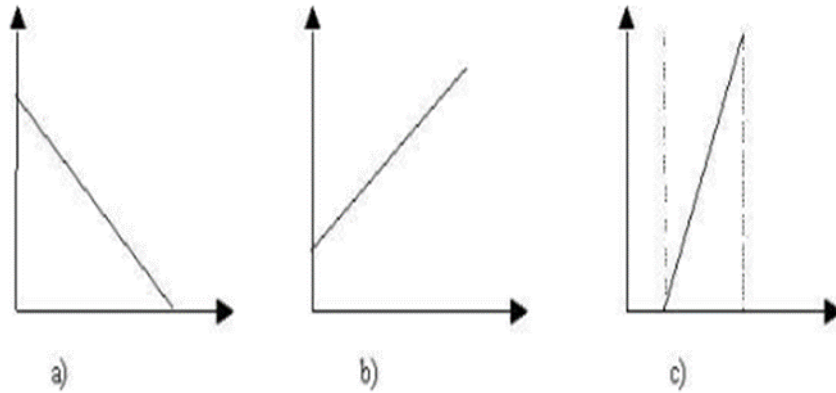


Figura 2.7 Transformaciones lineales (Sucar y Gomez, 2016).

(a) Negativo. (b) Aumento de intensidad. (c) Aumento de contraste.

Las transformaciones no-lineales normalmente son funciones monótonicas de forma que mantienen la estructura básica de la imagen. Algunos ejemplos de transformaciones no-lineales son los siguientes:

- Expansión (o aumento) de contraste. Se incrementa el contraste, en forma diferente para distintos rangos de intensidades.
- Compresión de rango dinámico. Se reduce el rango de niveles de gris o intensidades de la imagen.
- Intensificación de un rango de niveles. Se aumenta la intensidad de un rango de niveles de gris de la imagen.

Aumento lineal del contraste

Utilizando el valor de intensidad mínimo y máximo en una imagen, podemos aumentar su contraste. La idea básica es llevar el valor mínimo (min) a cero y el máximo (máx) a 255, pensando en imágenes monocromáticas (0-255). Esta transformación genera que las intensidades se espacien de acuerdo a cierto factor o pendiente; el factor para este aumento lineal de contraste es:

$$C(x, y) = \left(\frac{I(x, y) - \min}{\max - \min} * 255 \right) \quad (2.1)$$

Donde $I(x, y)$ es la imagen a procesar y $C(x, y)$ es la imagen con aumento lineal del contraste. Se puede verificar fácilmente que para $I(x, y)$ en min, $C(x, y)$ resulta cero (el numerador es cero); para $I(x, y)$ en máx, $C(x, y)$ resulta en 255 (cociente 1).

Selección de Propiedades

De acuerdo con (Ruiz, 2015) es un problema de búsqueda combinatoria. Dada una medida de calidad de un conjunto de propiedades (p. ej., el error de clasificación obtenido por un método determinado sobre un conjunto de ejemplos de prueba), tenemos que encontrar el subconjunto que posee mayor calidad. Como hay un número exponencial de subconjuntos, este tipo de selección 'global' de propiedades es computacionalmente intratable.

Si la medida de calidad es 'monótona' (esto significa que la incorporación de una nueva propiedad no empeora la calidad de un subconjunto, lo que cumplen algunos algoritmos de clasificación, pero no todos) se puede realizar una búsqueda eficiente, ordenada, basada en el algoritmo branch and bound (Se construye un árbol de búsqueda eliminando propiedades en cada rama;

no se desciende por ramas en las que se garantiza un empeoramiento de la calidad).

Una alternativa es la selección 'individual': elegimos una medida de calidad de cada propiedad por sí misma (sencilla de calcular y que indique lo mejor posible la 'separación' de las clases de interés), y nos quedamos con las mejores propiedades que sean, a la vez, estadísticamente independientes.

La selección individual de propiedades es, aparentemente, una idea muy prometedora. La calidad del conjunto será tan buena, al menos, como la del mejor componente, y, posiblemente, mejor. Si encontramos una propiedad discriminante, el problema de clasificación puede considerarse resuelto tal como se muestra en la figura 2.8.

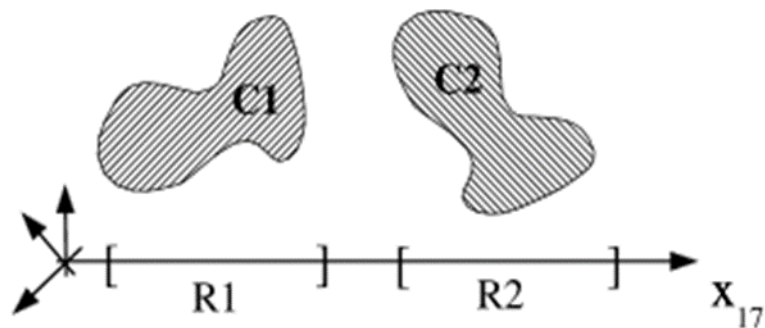


Figura 2.8 Selección de propiedades. (Ruiz, 2015)

El problema está en que la selección individual se pueden descartar propiedades poco informativas por separado pero que usadas en combinación son altamente discriminantes como se muestra en la figura 2.9.

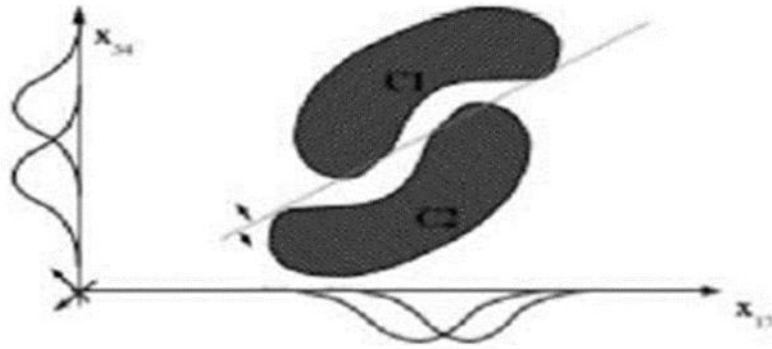


Figura 2.9 Selección individual (Ruiz, 2015)

Además, la selección individual produce a veces resultados poco intuitivos: p. ej., dadas tres propiedades p_1 , p_2 y p_3 , ordenadas de mayor a menor calidad individual, existen situaciones (debido a las dependencias entre ellas) en que el conjunto $\{p_1, p_2\}$ es peor que el conjunto $\{p_2, p_3\}$. Las medidas de calidad individual pueden ser útiles para fabricar árboles de clasificación.

Tipos de textura:

Define (Valentine y Moughamian, 2010) los siguientes conceptos:

- Difusa: Estos son los datos de píxeles actuales colocados sobre la malla de la superficie 3D. La textura puede ser una imagen o un color sólido.
- Auto iluminación: Es un color que no necesita las luces de la escena para verse. Le da al objeto el efecto de estar iluminado desde dentro, o resplandecer. Observe que si los lados del objeto son 100 por 100 opacos, esta textura no tendrá efecto.

- **Rugosidad:** Es un mapa de rugosidad que afecta a la topología local o al relieve de la superficie del modelo. Para los modelos creados como mallas a partir de capas de escala de grises, también tendrá acceso al mapa de profundidad.
- **Resplandor:** Este mapa define la cantidad de luz que refleja la superficie de un objeto. Negro es la sombra más reflectante, mientras que blanco es el menos.
- **Brillo:** Es diferente de Resplandor en que define las propiedades especulares de la superficie del objeto, o la dispersión de la luz reflejada. Con valores altos produce una reflexión pequeña de baja dispersión, mientras que valores bajos crean una reflexión dispersa más difusa.
- **Opacidad:** Controla la opacidad local de un objeto. Las regiones blancas son 100 por 100 opacas, mientras que las regiones negras son 100 por 100 transparentes.
- **Reflectividad:** Cambia la capacidad de la superficie del material para reflejar otros objetos en la escena y el mapa del entorno.
- **Entorno:** Este mapa contiene la imagen del "mundo" alrededor del objeto. Es lo que reflejara un objeto si no hay otros objetos en la escena. El mapa de entorno se hace como un panorama esférico.

2.1.1.2 Procesamiento de Medio Nivel.

El procesamiento a nivel medio supone tareas como la segmentación que consiste en la partición de una imagen en regiones u objetos, la descripción de esos objetos con el fin de reducirlos a una forma adecuada para el procesamiento informático y la clasificación o reconocimiento de objetos individuales. Este tipo de procesamiento se caracteriza por el hecho de que sus entradas generalmente son imágenes pero, sus salidas, son atributos extraídos de esas imágenes tales como bordes, contornos o la identidad de objetos individuales (González y Woods, 2008).

De acuerdo con (Rodríguez y Sossa, 2012) la segmentación es una de las etapas cruciales de cualquier sistema de análisis de imágenes, tanto por las dificultades que conlleva, como también por la importancia de sus resultados. Básicamente, la segmentación puede considerarse como la partición de una imagen, digamos $f(x, y)$ en un conjunto de regiones R no solapadas, homogéneas respecto a algún criterio cuya unión cubra la imagen completa. En otras palabras, el objetivo fundamental de la segmentación, en el proceso del análisis de imágenes, es el de separar los objetos de interés del resto no relevante el cual es considerado como fondo.

En ocasiones la segmentación suele considerarse como un proceso de clasificación de los objetos presentes en una imagen y también, en cierta medida, equivale a su reconocimiento puesto que como consecuencia de la segmentación los diferentes objetos (entendidos como realizaciones físicas de clases o patrones abstractos) se encuentran perfectamente ubicados dentro de la imagen digital. El realizar dicha clasificación no es una tarea trivial. El nivel al que se lleva a cabo esta subdivisión depende del problema a resolver y, por supuesto, de la imagen de entrada. Es por ello que se han originado diferentes técnicas de segmentación. Hasta el presente no se tiene conocimiento de una técnica única que pueda ser utilizada para segmentar

cualquier imagen. A decir de muchos autores la segmentación termina cuando se satisfacen los intereses u objetivos de la aplicación.

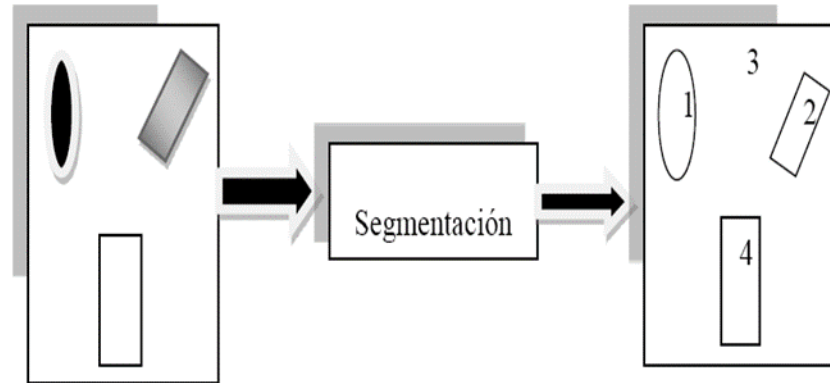


Figura 2.10 Esquema que representa el proceso de segmentación.
(Rodríguez y Sossa, 2012)

De manera gráfica en la figura 2.10 se presenta el proceso de segmentación de una imagen simple. De dicha figura se aprecia que en virtud de la segmentación se ha pasado de una imagen digital bruta, con toda su información en forma de los niveles de intensidad luminosa, a una imagen mucho más simplificada en lo que deben de estar nítidamente distinguidos entre si los diferentes objetos existentes.

En general, los métodos clásicos de segmentación se pueden categorizar como se indica a continuación:

Métodos basados en el umbralado a partir del histograma de la imagen: En este caso, a partir del histograma de una imagen es posible obtener un umbral de comparación para el agrupamiento de los píxeles.

Métodos basados en la detección de discontinuidades: En este caso, la imagen en cuestión es dividida a partir de cambios bruscos de los niveles de grises.

Métodos basados en la propiedad de similitud de los valores de los niveles de grises: En este caso se usan criterios de homogeneidad para la agrupación de los píxeles.

Método heurístico de segmentación: Estos métodos basan su operación en el conocimiento previo de la imagen a segmentar y en la experiencia del observador, e incluyen en muchas ocasiones los métodos supervisados de segmentación.

2.1.1.2.1 Segmentación por el clasificador de Bayes.

Según (Larrañaga et al. 2016) en muchas ocasiones cuando se utiliza un clasificador automático basado, por ejemplo, en la distancia euclídeana se utilizan prototipos fijos. Estos prototipos representan cada una de las clases u objetos a clasificar. Al hacer esto se está suponiendo una hipótesis determinística en el comportamiento de los elementos de las clases. Sin embargo, en aquellas situaciones en las cuales los vectores de algunas de las clases presentan una dispersión significativa con respecto a su media es conveniente utilizar un enfoque estadístico en vez de una hipótesis determinística. Así, por ejemplo, si se supone que se tienen dos tipos de frutas, por ejemplo naranjas y plátanos, las cuales pasan delante de un sistema de captación de imágenes con el objetivo de ser reconocidas para posteriormente desviarlas a diferentes lugares según las clases a la que pertenezca cada fruta.

2.1.1.2.2 Formula de Bayes aplicada a la segmentación supervisada.

Sea un elemento con un valor v y para cada clase i sea $q(i/v)$ la probabilidad de que el elemento sea de la clase i , entonces la mejor clase para asignar dicho elementos es la clase para la cual $q(i/v)$ es máxima (Díaz, Montoya, y Boulanger, 2007). Esto viene dado por la expresión siguiente:

$$\frac{q(v/i)q(i)}{q(i/v)} = q(v) \quad (2.2)$$

Dónde: $q(i)$ es la probabilidad a priori de que cualquier elemento pertenezca a la clase i . Esta es la probabilidad que se asume conocida como antes del comienzo de la segmentación o clasificación.

$q(v/i)$ es la probabilidad de que dada la clase i , el valor de la variable aleatoria tenga precisamente el valor v . Esta es la función de densidad de probabilidad condicional para la clase i , o la probabilidad que un elemento de la clase i tenga valor v .

$q(i/v)$ es la probabilidad de que un elemento con valor v este en la clase i .

$q(v)$ es la suma de las probabilidades $q(v/i)$ sobre toda las clases i . Esto se puede considerar sencillamente como un factor de normalización o un factor de escala.

La expresión anterior aporta la solución al problema de la segmentación de imágenes ya que dado el elemento con valor v , el cual también puede ser considerado como un vector descriptivo, este pertenece a la clase j si se cumpla la siguiente condición:

$$V \in \text{clase } j \text{ si } p(\text{clase } j/v) > p(i/v) \quad (2.3)$$
$$\forall i \neq j, i = 1, 2, 3, \dots, N$$

Los pasos genéricos del algoritmo para segmentar una imagen según el clasificador de Bayes son los siguientes (Díaz, Montoya, y Boulanger, 2007):

- 1.- Seleccionar las áreas de entrenamiento para cada una de las clases candidatas. Estas áreas se suponen conocidas y, en general, se debe escoger una gran cantidad de datos por clase con el fin de que estos representen lo mejor posible dicha clase.
- 2.- Calcular el histograma $H_i(v)$ de los datos de entrenamiento para cada una de las áreas seleccionadas y normalizarlos. Esta estimación es usada como una función de densidad de probabilidad condicional: $q(v/i)$. Existe una $q(v/i)$ para cada una de las clases i seleccionadas.
- 3.- Estimar la probabilidad a priori $q(i)$ la cual será usada como un factor de escala
- 4.- Clasificar cada uno de los píxeles de la imagen calculando, para la clase i , $q(i/v)$ mediante la expresión anterior. En la práctica se omite el denominador en la expresión anterior.

5.- Clasificar el valor del pixel en la clase i si el valor de $q(i/v)$ es máximo. Los pasos 1, 2 y 3 se realizan una sola vez, el paso 4 se aplica para cada uno de los pixeles de la imagen para producir la imagen segmentada.

Los métodos de segmentación supervisados han sido muy utilizados en la teledetección de los recursos naturales donde muchas veces se tiene, a partir de las pruebas de campo realizadas, algún conocimiento del terreno y de los objetos.

Otros métodos clásicos de segmentación supervisada se han reportado en la literatura. Como ejemplos se tienen el método basado en el cálculo de los centroides y el método basado en el cálculo de la mínima distancia.

Una técnica clásica opuesta a la de crecimiento de regiones es la de división y fusión de regiones. En este caso se comienza bajo el supuesto que la imagen completa es homogénea, de no cumplirse esta condición acorde algún criterio, la imagen se divide en 4 sub imágenes. En forma iterativa cada una de las sub imágenes es subdividida en otras 4 sub imágenes si no satisface dicho criterio. Al final se tiene una imagen dividida en sub imágenes cuyos elementos satisfacen el criterio de homogeneidad establecido.

2.1.1.2.3 Detección de bordes.

Según (Martin, 2013) un borde puede definirse como un cambio significativo en el valor de la intensidad de los pixeles en una región de la imagen. Durante el proceso de detección de bordes se puede obtener la siguiente información:

- Orientación local de los elementos del borde.
- Intensidad de los elementos del borde, es decir, contraste en la brillantez entre regiones vecinas.

- Ancho de los elementos del borde (puesto que los bordes no son pasos ideales, estos en general varían considerablemente en su ancho).
- Colocación de la representación del borde (puesto que pueden tener más de un pixel de ancho, es importante determinar dónde poner el punto del borde).
- Polaridad del elemento del borde (cuál de las regiones es la más brillante).
- Valor de gris del elemento del borde.
- Valor de gris de las regiones del entorno.

En la detección de bordes surgen problemas durante el proceso de selección, a saber:

- Ruido (dado por los pixeles del fondo u otros factores).
- Manchas o emborronamiento (algunas partes del contorno son más espesas o densas que otras).
- Fragmentación (algunas partes del contorno se pierden).
- Mala colocación de los puntos.

Es importante tener en cuenta que un buen detector de bordes debe contar con las propiedades siguientes (Martin, 2013):

- Debe operar localmente.
- Debe ser sensitivo a la orientación y magnitud del lado.

- Debe trabajar adecuadamente en presencia de ruido.
- Debe ser insensible al valor del umbral.
- No debe dar múltiples respuestas a un mismo borde.

En la práctica, debido a la presencia del ruido y a otros problemas en la captación de la escena, es muy poco probable que los detectores de bordes clásicos cumplan en forma eficiente con los puntos anteriormente expuestos. De ahí, la necesidad de la propuesta de nuevas técnicas de detección de bordes menos rígidas.

La importancia que tiene la determinación de los contornos de una figura dentro de una imagen radica en que muchos algoritmos requieren que se demarquen los objetos que están contenidos en una imagen. Algunas de las aplicaciones más importantes que requieren de ésta operación como paso previo son las siguientes:

- Conteo de objetos.
- Localización de objetos.
- Medición de las características métricas de los objetos.
- Determinación de las características geométricas de los objetos.
- Discriminación por tamaño o forma de objetos.
- Reconocimiento óptico de caracteres (OCR: Optical Character Recognition).

2.1.1.2.4 Técnicas actuales de segmentación.

Menciona (Martínez, 2016) la segmentación que utiliza técnicas tradicionales, tales como: el umbralado, por cálculo de gradiente, por crecimiento de regiones y otras operaciones clásicas, requieren de una considerable cantidad de iteraciones para la obtención de resultados satisfactorios. La automatización de estos modelos libre de aproximaciones es difícil debido a muchos factores, entre otros a la complejidad de los objetos en escena, los tipos de iluminación, las sombras producidas por el acomodo de los objetos y la forma en que estos son iluminados.

Además el ruido y otras variantes que pueden originar regiones incorrectas o discontinuidades en las fronteras de los objetos extraídos. Tomando en cuenta todo esto se han propuesto métodos más elaborados que, en muchos casos, superan a los métodos tradicionales. Enseguida se exponen algunos de estos métodos.

2.1.1.3 Procesamiento de Alto Nivel.

El procesamiento a alto nivel, también llamado presentación conlleva darle sentido a un conjunto de objetos reconocidos, como en el análisis de imágenes, así como realizar las funciones cognitivas normalmente asociadas con la visión.

Según (Sucar y Gomez, 2016) el objetivo principal es llevar los datos de cada uno de los objetos o regiones segmentados a formas en que la computadora pueda trabajar con ellos de manera más apropiada; una lista ligada circular que presente la información del contorno de la región de un objeto es más fácil de trabajar que un conjunto de pares de coordenadas de los pixeles del contorno de la misma región.

Una representación es “un sistema formal para hacer explicitas ciertas características o tipos de información, junto con una especificación de como el sistema hace esto”, como se muestra en la figura 2.11. Hay dos aspectos importantes de una representación para visión (Sucar y Gomez, 2016):



Figura 2.11 Reconocimiento de caracteres en base a su codificación radial. (Sucar y Gomez, 2016)

1. Representación del modelo. El tipo de estructura utilizada para modelar la representación interna del mundo.
2. Proceso de reconocimiento. La forma en que dicho modelo y la descripción de la imagen es utilizada para el reconocimiento.

Las representaciones apropiadas para reconocimiento en visión deben de buscar tener las siguientes propiedades:

- Genéricas.
- Eficientes, en espacio y tiempo.
- Invariantes, independientes de traslación, rotación y escalamiento.
- Robustas, tolerantes a ruido e información incompleta.

La estructura general de un sistema de visión basado en modelos se muestra en la figura 2.12. Tiene 3 componentes principales (Sucar y Gomez, 2016):

- Extracción de características: Obtención de información de forma de la imagen para construir una descripción geométrica.
- Modelado: Construcción de los modelos geométricos internos de los objetos de interés (a priori).
- Correspondencia o Matching: apareamiento geométrico de la descripción con el modelo interno.

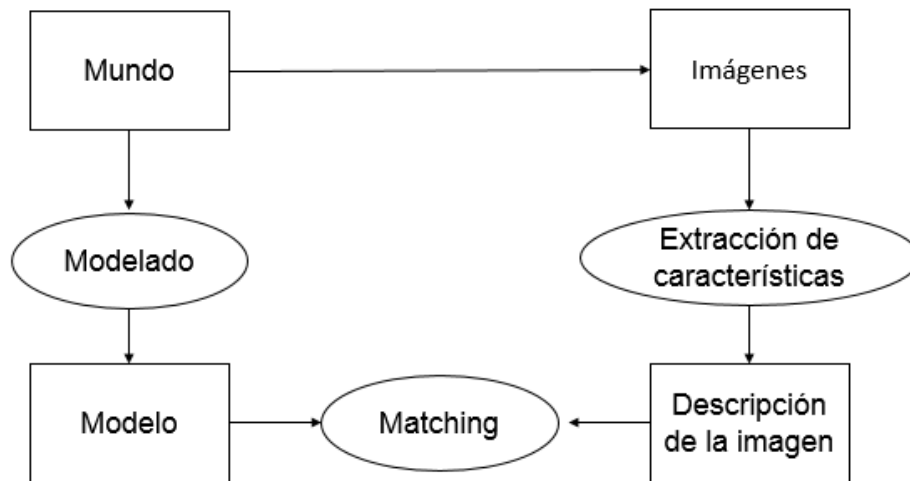


Figura 2.12 Estructura de un sistema de visión basado en modelos.

(Sucar y Gomez, 2016)

Las técnicas para correspondencia dependen del tipo de representación. Para modelos que utilizan parámetros globales se utilizan técnicas de reconocimiento estadístico de patrones. Con modelos en base graficas relacionales, se usan algoritmos de grafos (isomorfismo). En modelos paramétricos se aplican técnicas de optimización paramétrica.

Las técnicas para correspondencia dependen del tipo de representación. Para modelos que utilizan parámetros globales se utilizan técnicas de reconocimiento estadístico de patrones. Con modelos en base a graficas relacionales, se usan algoritmos de grafos (isomorfismo). En modelos paramétricos se aplican técnicas de optimización paramétrica.

2.1.1.3.1 Modelos en dos dimensiones.

De acuerdo a (Gambini, 2006) los modelos en dos dimensiones (2-D) están orientados al modelado y reconocimiento de objetos en función de su representación a nivel imagen, es decir, en dos dimensiones. Para representar un objeto en 2-D, existen básicamente dos alternativas:

- Contornos. El objeto se representa en base a su borde o contorno.
- Regiones. El objeto se representa en base a la región que define.

2.1.1.3.2 Polilíneas.

Según (Gambini, 2006) la representación de polilíneas consiste en una descripción de contornos en base a segmentos de línea, donde cada segmento (X) se especifica mediante el punto inicial y final. La concatenación de estos puntos, con el mismo punto inicial y final, describe un contorno:

$$X_1 X_2 \dots X_N, X_1 \quad (2.4)$$

Donde X_1 corresponde a las coordenadas x, y de cada punto.

2.1.1.3.3 Descripción.

Según (Quiles y Garrido, 1996) el objetivo principal consiste en capturar las diferencias esenciales entre objetos pertenecientes a clases diferentes; por supuesto se buscaría que estos mismos rasgos se mantuvieran lo más invariantes ante cambios como escalamientos, traslaciones y rotaciones. Es el proceso mediante el cual se obtienen características convenientes para diferenciar un tipo de objeto de otro, por ejemplo tamaño y forma.

La mayoría de procesamiento de imágenes se aplica directamente sobre imágenes en escala de grises, debido al bajo consumo de cómputo de éste, la mayoría de métodos matemáticos deterministas que se utilizan en su procesamiento, están basados en la diferencia de niveles de grises, por lo que no existen muchos métodos para procesamiento de imágenes en color, aun cuando estos pueden ser utilizados en este tipo de formato de imagen.

2.1.1.3.4 Obtención de características del objeto por color.

Según (Flores, 2016) existen varias representaciones o modelos de color. Estos modelos se dividen en dos clases de modelos. Unos son modelos sensoriales y otros se denominan modelos perceptuales.

La idea principal de obtener características por color es sacar datos relevantes en formato RGB de objetos, para crear características que puedan ser utilizadas por otros procesos para llegar al resultado como se muestra en la figura 2.13.

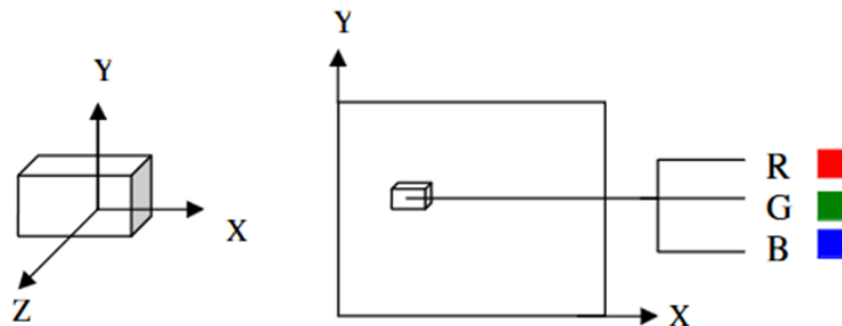


Figura 2.13 Obtención de características del objeto por color (Flores, 2016)

Por la naturaleza de la captura de la imagen, se trabaja en formato RGB, por lo cual se usa un formato sensorial en lugar de un formato de percepción como el HSI (Hue, Saturation, Intensity).

Además se pueden aplicar las mismas técnicas que se utilizan en imágenes monocromáticas. Para esto se toma cada componente (R,G,B) como una imagen monocromática y se aplica algún operador o método a cada una independientemente.

Después se combinan todas (considerando normalmente el máximo o el promedio).

2.1.1.3.5 Modelo sensorial.

De acuerdo a (Flores, 2016) el modelo RGB es el modelo básico que utiliza los componentes primarios rojo, verde y azul, normalizadas. En este modelo se basan las cámaras y receptores de televisión. Sin embargo se tienen problemas al aplicarlo al procesamiento de imágenes (ecualización).

2.1.1.3.6 Modelo perceptual.

Comenta (Martinez, 2016) el modelo HSI (Hue, Saturation, Intensity) se puede ver como una transformación del espacio RGB al espacio perceptual. En principio los modelos perceptuales deben ser mejores ya que nosotros detectamos los cambios en estos componentes. Sin embargo, es compleja la implementación de la detección de orillas en croma por no ser lineal.

Otra alternativa es definir técnicas especiales para detección de orillas en imágenes a color. El procesamiento de imágenes a color es relativamente reciente, por los altos requerimientos de memoria y cómputo necesarios.

La extracción de información de las imágenes 'Information from imagery' a través del procesamiento digital constituye hoy en día un inmenso campo de estudio e investigación en diversas disciplinas con múltiples aplicaciones. En este sentido, fotogrametristas, matemáticos, físicos, informáticos y demás, continúan investigando en cuestiones que van desde la aplicación de simples filtros lineales hasta la automatización del reconocimiento semántico de objetos. No obstante, ha sido la disciplina de la Visión Computacional donde se han conseguido los mayores logros, de hecho la detección automática de características sobre imágenes tiene aquí una dilatada tradición y cuenta multitud de métodos para tal propósito. Desafortunadamente, a pesar de la gran proliferación de métodos, no existe un 'método universal' para la

detección automática de características, sino que serán los requerimientos del propio problema los que nos obligarán a desarrollar y personalizar nuestro método.

Una imagen contiene una gran cantidad de datos la mayoría de los cuales proporciona muy poca información para interpretar la escena. Un sistema que incorpore visión artificial debe, en un primer paso, extraer de la forma más eficaz y robusta posible determinadas características que nos proporcionen la máxima información posible. Estas características deben cumplir, entre otras, las siguientes condiciones:

- Su extracción a partir de la imagen no debe suponer un coste excesivo al sistema en el cual está integrado. El tiempo total de extracción debe ser lo más pequeño posible.
- Su localización debe ser muy precisa. El error cometido en la estimación de las características también debe ser lo más pequeño posible.
- Deben ser robustas y estables. Deberían permanecer a lo largo de una secuencia.
- Contendrán la máxima información posible de la escena, es decir, debemos ser capaces de extraer información de tipo geométrico a partir de ellas.

La extracción de características más importantes en el análisis de imagen son: los puntos, las líneas y los círculos como geometrías básicas en la extracción de características.

2.1.1.3.7 Extracción de puntos.

Según (Martínez, 2016) desde el punto de vista computacional se han propuesto dos enfoques para la detección de este tipo de características geométricas:

- Métodos que obtienen los puntos como intersección de aristas o como cambio de pendiente sustancial entre dos aristas y por tanto vienen precedidos de una extracción de bordes.
- Métodos que trabajan directamente sobre imágenes de gris, es decir, y no requieren una extracción previa de aristas.

2.1.1.3.8 Puntos característicos a partir de aristas: puntos de fuga.

De acuerdo con (González, 2016) los puntos de fuga constituyen el soporte estructural y geométrico de una imagen en perspectiva u oblicua. En este sentido representan puntos de interés que vendrán determinados por la intersección de aristas o líneas de fuga. A continuación se comentan brevemente los dos métodos más sencillos para su cálculo:

Método de la intersección de rectas. Se trata del método más sencillo de todos, ya que su cálculo computacional se limita a determinar la intersección de dos rectas perspectivas.

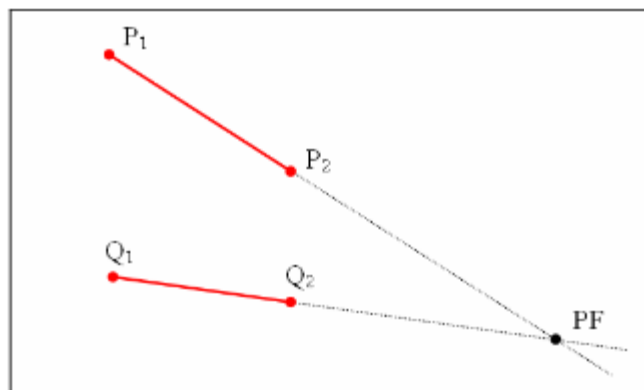


Figura 2.14 Intersección de rectas. (González, 2016)

A partir de la figura 2.14, en el que una recta queda identificada geoméricamente mediante su distancia al origen y el ángulo que dicha recta forma con los ejes cartesianos se obtiene la expresión de la misma y sus versiones lineales:

$$\begin{array}{ll} \text{Ecuación punto pendiente} & \text{Ecuación normal} \\ y = ax + b & Ax + By + 1 = 0 \end{array} \quad (2.5)$$

Con valores (a, 1, b) para los coeficientes de la X, de la Y y del término independiente, respectivamente o con valores (A, B, 1) para los mismos términos.

La resolución matemática del problema pasa por el planteamiento y resolución del siguiente sistema de ecuaciones:

$$[-a \ 1] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = b \text{ o } [A \ B] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = [-1] \quad (2.6)$$

Siendo x, y las coordenadas del punto de fuga a calcular.

Método de la minimización del área del triángulo. El método consiste en el cálculo y minimización de la superficie de los triángulos formados por cada uno de los segmentos extraídos y por el punto de fuga, como vértice opuesto de cada triángulo y común a todos ellos.

El área de cada triángulo (S) se calcula a través del determinante formado por las coordenadas de los dos puntos extremos de la línea de fuga (P1, P2) y las coordenadas del punto de fuga (PF). Como se muestra en la figura 2.15.

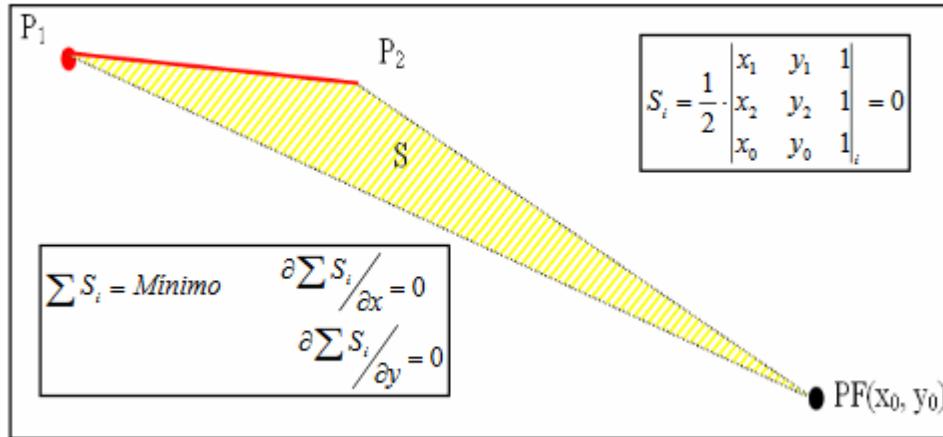


Figura 2.15 Método de la minimización del área del triángulo.
(González, 2016)

En general, la eficiencia de estos métodos dependen directamente de la calidad del método empleado para la obtención de aristas: si este último no localiza correctamente los puntos de arista difícilmente podremos detectar puntos de fuga de forma exacta. Además se añade un tiempo extra de procesamiento previo de las aristas.

2.1.1.3.9 Puntos característicos a partir de la propia imagen: puntos esquina.

Menciona (González, 2016) existen otros métodos que obtienen los puntos de máxima curvatura de una imagen utilizando directamente los valores de la imagen, sin realizar el paso previo de obtención de aristas. Estos métodos definen una medida que suelen denominar de 'esquinidad' la cual se calcula para todos los puntos de la imagen. Cuando esta medida supera un cierto umbral se considera que el punto es una esquina. La mayoría de estos métodos utilizan operadores diferenciales.

2.1.1.3.10 Extracción de líneas.

De acuerdo con (Gómez y Guerrero, 2016) los métodos de gradiente y laplaciano vistos hasta ahora no proporcionan por si solos una solución de calidad para la extracción de líneas, de ahí que tengamos que recurrir a estrategias más complejas. A continuación, se describen tres metodologías para la extracción automática de líneas:

- Canny + Burns
- RANSAC + MMCC
- Transformada de Hough

Canny & Burns. Extracción de líneas siguiendo un proceso multifase jerárquico que se basa en la extracción de bordes mediante el algoritmo de Canny y en la segmentación posterior de dichos bordes mediante el algoritmo de Burns.

Detección de bordes: Filtro de Canny. El detector de bordes de Canny (Canny, 1986) resulta el más idóneo para la detección de bordes en imágenes donde existe la presencia de geometrías regulares, ya que mantiene tres criterios vitales para nuestros propósitos:

- Precisión en la localización del contorno, garantizando la mayor proximidad de los bordes detectados a los bordes verdaderos.
- Fiabilidad en la detección de los puntos de contorno, minimizando la probabilidad de detección de falsos bordes causados por el ruido, así como también la pérdida de bordes reales.
- Unicidad en la obtención de un único borde para el contorno, garantizando bordes con una achura máxima de un píxel.

RANSAC+MMCC.

RANSAC (RANmdom SAmple Consensus), es un estimador robusto desarrollado por (Fischler and Bolles, 1981) que se basa en la aplicación de una técnica de votación resultante de un muestreo aleatorio, con el objetivo de determinar el número de observaciones válidas '*inliers*' y el número de observaciones erróneas '*outliers*'. En nuestro caso la aplicación de *RANSAC* servirá para determinar posibles errores groseros que de ser incluidos en el proceso mínimo cuadrático depararía gravísimos errores en la extracción de las líneas.

Los pasos a considerar por parte de *RANSAC* serían los siguientes:

1. Selección aleatoria de dos puntos aleatorios para constituir una posible recta candidata.
2. Validación de la recta, en función del número de puntos que más menos una cierta tolerancia tengan una variación mínima de su distancia ortogonal a la recta candidata.
3. Repetición del paso 1 y 2 un determinado número de veces.
4. El máximo resultante del proceso de votación después de un número aleatorio de combinaciones se corresponderá con la recta candidata y por consiguiente los votos favorables '*inliers*' del máximo contendrán los puntos favorables a constituir una recta. Los segmentos que superen la tolerancia en los tres grupos generados serán considerados como segmentos erróneos '*outliers*' y serán eliminados.

2.1.1.3.11 Transformada de Hough para líneas.

De acuerdo con (Gomez & Guerrero, 2016) la transformada de Hough está diseñada especialmente para encontrar líneas. Definimos una línea como una colección de puntos de borde que son adyacentes y que tienen la misma dirección. La transformada de Hough es un algoritmo que tomará una colección de puntos de borde, encontrados mediante un detector de bordes y buscará todas las líneas sobre las cuales estos puntos de borde se encuentran.

La idea básica es convertir los puntos de bordes al espacio de parámetros.

La estrategia para la extracción de líneas mediante Hough es la siguiente:

- Detección de los píxeles de borde de las rectas mediante un filtro de bordes.
- Establecimiento de un espacio de parámetros de dimensiones el espacio de búsqueda y una cuantización suficientemente precisa.
- Se barre la imagen de manera que cada pixel de borde da lugar a una recta: las celdas por las que "pasa" esta recta reciben un "voto".
- En teoría todos los píxeles que pertenecen a una misma recta (en la representación espacial) son rectas (en el dominio de parámetros) que se cortan en una misma celda (en la representación de los parámetros): la recta resultará ser la celda más votada.

2.1.1.3.12 Extracción de círculos.

Según (González, 2016) transformada de Hough para círculos. Al igual que se explotaba la dualidad punto-línea en la transformada de Hough para líneas se podrá hacer algo similar en el caso de los círculos.

Espacio de parámetros. Si el radio es conocido, el dominio de parámetros de cada círculo es bidimensional: coordenadas del centro de cada círculo. En este dominio, cada círculo del espacio se representa con un punto y simétricamente, un punto del dominio espacial se representa en el dominio de parámetros mediante un círculo formado por todos los puntos (dominio de parámetros) que representan a todos los círculos (dominio espacial) que pueden pasar por el punto (dominio espacial).

Menciona (García P., 2012) después de procesar adecuadamente la imagen se obtienen las características más significativas que permita diferenciar los objetos a clasificar. Si las imágenes fueran adquiridas mediante un escáner, la inclinación podría ser muy pequeña, incluso tanto que podría despreciarse. Sin embargo, cuando las imágenes son adquiridas con un móvil o una cámara fotográfica, la inclinación entra a ser un factor importante. Por esta razón es necesario encontrar características que sean invariantes a la rotación y a la escala, es decir, que al rotar o escalar una imagen, el valor numérico de las características sea similar.

Para lograr un cierto grado de independencia de las características a la rotación y a la escala se pueden tomar dos enfoques: un primer enfoque es transformar la imagen, esto significa rotarla y escalarla para hacerla lo más similar posible a un patrón estándar, después de esto se pueden extraer características numéricas de la imagen. Evidentemente en cualquiera de los dos casos, las características buscadas deben ser lo más similares posibles para objetos de la misma clase y lo más diferentes posibles para objetos de distintas clases. Un segundo enfoque consiste en trabajar con la imagen tal

cual es tomada por la cámara, y utilizar características en las que se obtengan valores numéricos similares para una misma imagen rotada y escalada, es decir que sean invariantes a la rotación y a la escala.

2.1.1.3.13 Reconocimiento.

Según (Rodríguez & Sossa, 2012) el reconocimiento es el procedimiento que consiste en decir si una instancia de un objeto se encuentra presente en una imagen.

1.- Análisis: Es el procedimiento que asigna una determinada etiqueta a un objeto a partir de la información proporcionada por sus descriptores.

2.- Interpretación de la imagen: Es la que determina el significado de un conjunto de objetos previamente reconocidos.

Hay diversos métodos para llevar a cabo dicha etapa, una de ellas es el clasificador “K vecinos más próximos”.

2.1.1.3.14 Clasificador “K Vecinos más Próximos”.

Según (García E. , 2008) la clasificación por los “K” vecinos más próximos (K Nearest Neighbor o K-NN) es ampliamente utilizada en el reconocimiento de formas. Dado un vector a clasificar (rasgos característicos del objeto a clasificar) y un conjunto de vectores prototipo asignados a las diversas clases existentes (base de conocimiento). La regla consiste en calcular la distancia del primero a cada uno de los segundos, seleccionar los “K” vecinos más próximos y decidir por la clase más votada entre los mismos.

Sea x un vector de dimensión “n” a clasificar, sea M una base de datos de referencia construida de N vectores de dimensión “n” y se conoce la clase C_i a la cual pertenecen los vectores de la base de referencia M . Como se muestra en la figura 2.16 el clasificador K vecinos más próximos está basado en la estimación local de la densidad de probabilidad de la muestra x a partir de los “ K próximos vecinos” de la base de referencia.

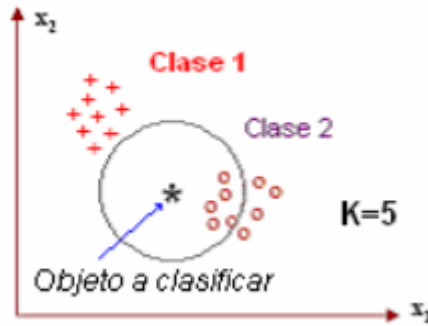


Figura 2.16 Clasificador KNN. (Garcia E. , 2008)

Sea $p(\vec{x}/C_i)$ la densidad de probabilidad. A partir de esta estimación, la regla de BAYES nos permite expresarlo en términos de la probabilidad a posteriori que la muestra x pertenezca a la clase C_i , tal que:

$$p_r(C_i/\vec{x}) = \frac{p(\vec{x}/C_i) * p_r(C_i)}{p(\vec{x})} = \frac{p_r(\vec{x}/C_i) * p_r(C_i)}{\sum_{k=1}^C p_r(\vec{x}/C_k) * p_r(C_k)} \quad (2.7)$$

Dónde:

$p_r(C_i)$ = probabilidad de aparición de la clase C_i .

$p_r(\vec{x})$ = probabilidad de que la muestra \vec{x} pertenezca a la clase C_i .

$p_r(C_i/\vec{x})$ = densidad de probabilidad condicional de la muestra \vec{x} conociendo la clase C_i .

2.1.2 Ejemplos de campos de aplicación de Procesamiento Digital de Imágenes.

Hoy en día, son pocas las áreas de esfuerzo técnico que no utilicen el procesamiento de imágenes digitales. Las áreas de aplicación son tan diversas que es deseable alguna forma de organización al intentar capturar la amplitud de este campo. Una de las formas más simples para una comprensión básica del alcance de las aplicaciones donde es utilizado el procesamiento de imágenes es categorizar las imágenes de acuerdo con su fuente (por ejemplo, Visual, Rayos X, etc.). Hoy en día el espectro de energía electromagnética es la principal fuente de energía para las imágenes en uso. Otras fuentes importantes de energía incluyen acústica, ultrasónica y electrónica (en forma de haces de electrones utilizados en microscopía electrónica). Las imágenes sintéticas, utilizadas para modelar y visualizar, son generadas por computadora. Las imágenes basadas en la radiación del espectro EM son las más familiares, especialmente las imágenes de rayos X y las bandas visuales del espectro. Las ondas electromagnéticas pueden conceptualizarse como ondas sinusoidales de propagación de diferentes longitudes de onda, o pueden considerarse como una corriente de partículas sin masa, cada una de las cuales viajan en un patrón ondulatorio y se mueve a la velocidad de la luz. Cada partícula sin masa contiene una cierta cantidad (o haz) de energía. Cada paquete de energía se llama fotón. Si las bandas espectrales se agrupan según la energía por fotón, obtenemos el espectro tal y como se muestra en la figura 2.17, que va desde los rayos gamma (energía más alta) en un extremo hasta las ondas de radio (energía más baja) en el otro.

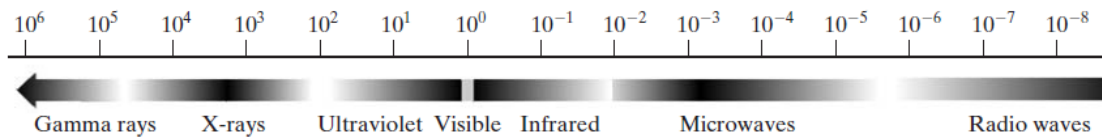


Figura 2.17 El espectro electromagnético dispuesto según la energía por fotón. (González y Woods, 2008)

Las bandas se muestran sombreadas para transmitir el hecho de que las bandas del espectro EM no son distintas, sino que pasan sin problemas de una a otra (González y Woods, 2008).

2.1.2.1 Imágenes de rayos gamma.

De acuerdo a (González y Woods, 2008) los principales usos de las imágenes basadas en rayos gamma están presentes en la medicina nuclear y las observaciones astronómicas. En medicina nuclear, el sentido consiste en inyectar a un paciente con un isótopo radiactivo que emite rayos gamma a medida que se descompone. Las imágenes se producen a partir de las emisiones recogidas por los detectores de rayos gamma. La figura 2.18 muestra una imagen de una exploración ósea completa obtenida mediante imágenes de rayos gamma. Las imágenes de este tipo se utilizan para localizar sitios de patología ósea, como infecciones o tumores.

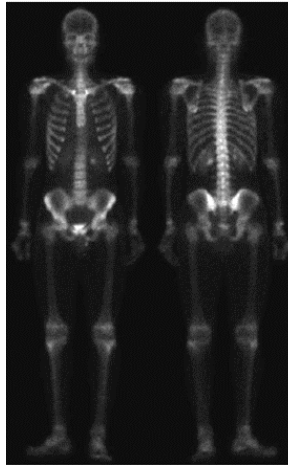


Figura 2.18 Radiografía. (González y Woods, 2008)

Otra modalidad importante de imagen nuclear llamada tomografía por emisión de positrones (PET). El principio es el mismo que con la tomografía de rayos X, como ya se había mencionado. Sin embargo, en lugar de utilizar una fuente externa de energía de rayos X, el paciente recibe un isótopo radioactivo que emite positrones a medida que se descompone. Cuando un positrón se encuentra con un electrón, ambos son aniquilados y se emiten dos rayos gamma. Estos son detectados y se crea una imagen tomográfica utilizando los principios básicos de la tomografía. La imagen que se muestra en la figura 2.19 es una muestra de una secuencia que constituye una representación tridimensional del paciente. Esta imagen muestra un tumor en el cerebro y uno en el pulmón, fácilmente visibles como pequeñas masas blancas.



Figura 2.19 Tomografía por emisión de positrones (PET). (González y Woods, 2008)

Una estrella en la constelación de Cygnus explotó hace unos 15,000 años, generando una nube de gas estacionaria sobrecalentada (conocida como Cygnus Loop) que brilla en una espectacular variedad de colores. La figura 2.20 muestra una imagen del Cygnus Loop en la banda de rayos gamma, a diferencia de los dos ejemplos en las figuras 2.18 y 2.19, esta imagen se obtuvo utilizando la radiación natural del objeto que se está fotografiando.



Figura 2.20 Nube de gas estacionaria sobrecalentada. (Cygnus Loop). (González y Woods, 2008)

En la figura 2.21 se muestra una imagen de radiación gamma de una válvula en un reactor nuclear. Donde se observa un área de radiación fuerte en el lado inferior izquierdo de la imagen.

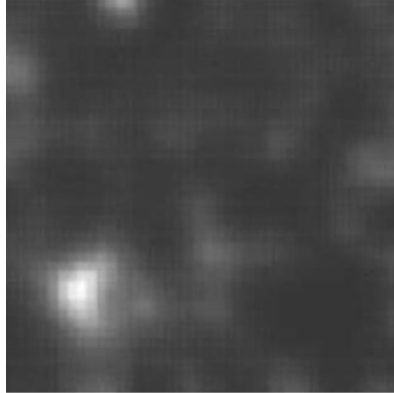


Figura 2.21 Radiación gamma (punto brillante) de una válvula del reactor.
(González y Woods, 2008)

2.1.2.2 Imágenes de rayos X.

Los rayos X se encuentran entre las fuentes más antiguas de radiación EM utilizadas para obtener imágenes. El uso más conocido de los rayos X es el diagnóstico médico, pero también se usan ampliamente en la industria y otras áreas, como la astronomía. Los rayos X para imágenes médicas e industriales se generan utilizando un tubo de rayos X, que es un tubo de vacío con un cátodo y un ánodo. El cátodo se calienta y se liberan electrones libres. Estos electrones fluyen a alta velocidad hacia el ánodo cargado positivamente. Cuando los electrones golpean un núcleo, se libera energía en forma de radiación de rayos X. La energía (poder de penetración) de los rayos X se controla mediante un voltaje aplicado a través del ánodo y una corriente aplicada al filamento en el cátodo. La figura 2.22 se muestra una radiografía de tórax familiar generada simplemente colocando al paciente entre una fuente de rayos X y una película sensible a la energía de los rayos X. La intensidad de los rayos X se modifica por absorción a medida que pasan a través del paciente, y la energía resultante que cae sobre la película la desarrolla, de la

misma manera que la luz desarrolla una película fotográfica. En radiografía digital, las imágenes digitales se obtienen por uno de dos métodos: (1) digitalizando películas de rayos X; o (2) haciendo que los rayos X que pasan a través del paciente caigan directamente sobre dispositivos (como una pantalla de fósforo) que convierten los rayos X en luz. La señal de luz a su vez es capturada por un sistema de digitalización sensible a la luz.



Figura 2.22 Radiografía de pecho. (González y Woods, 2008)

La angiografía es otra aplicación importante en un área llamada radiografía de contraste. Este procedimiento se usa para obtener imágenes (llamadas angiogramas) de los vasos sanguíneos. Se inserta un catéter (un tubo pequeño, flexible y hueco), por ejemplo, en una arteria o vena en la ingle. El catéter se inserta en el vaso sanguíneo y se guía al área a estudiar. Cuando el catéter llega al sitio debajo En la investigación, se inyecta un medio de contraste de rayos X a través del tubo. Esto mejora el contraste de los vasos sanguíneos y permite al radiólogo ver cualquier irregularidad u obstrucción. La figura 2.23 muestra un ejemplo de angiografía aórtica. Se puede ver el catéter insertado en el vaso sanguíneo grande en la parte inferior izquierda de la imagen. Tenga en cuenta el alto contraste del vaso grande a medida que el medio de contraste fluye hacia arriba en la dirección de los riñones, que también son visibles en la imagen. Como se discutió en el Capítulo 2, la angiografía es un área importante del procesamiento de imágenes digitales,

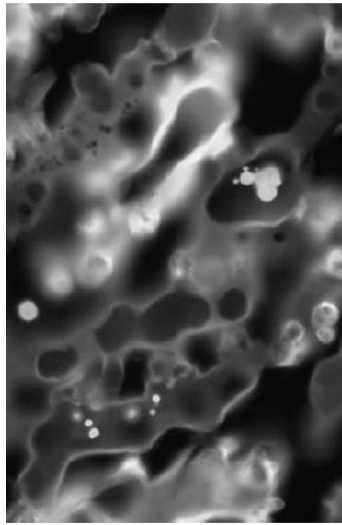
donde la sustracción de imágenes se usa para mejorar aún más los vasos sanguíneos que se estudian.



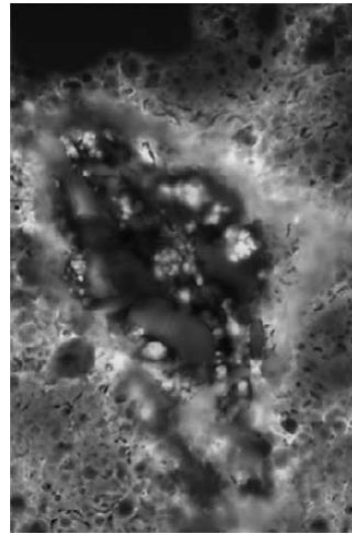
Figura 2.23 Angiografía aórtica. (González y Woods, 2008)

2.1.2.3 Imágenes en la banda ultravioleta.

Las aplicaciones de la "luz" ultravioleta son variadas. Incluyen litografía, inspección industrial, microscopía, láser, imágenes biológicas y observaciones astronómicas. Ilustramos imágenes en esta banda con ejemplos de microscopía y astronomía. La luz ultravioleta se usa en la microscopía de fluorescencia, una de las áreas de microscopía de más rápido crecimiento. La luz ultravioleta en sí no es visible, pero cuando un fotón de radiación ultravioleta choca con un electrón en un átomo de un material fluorescente, eleva el electrón a un nivel de energía más alto. Posteriormente, el electrón excitado se relaja a un nivel inferior y emite luz en forma de un fotón de menor energía en la región de luz visible (roja). La microscopía de fluorescencia es un método excelente para estudiar materiales que pueden hacerse fluorescentes, ya sea en su forma natural (fluorescencia primaria) o cuando se tratan con productos químicos capaces de fluorescencia (fluorescencia secundaria). Las figuras 2.24 (a) y (b) muestran resultados típicos de la capacidad de la microscopía de fluorescencia.



a)



b)

Figura 2.24 Maíz normal y Tizón de maíz. (González y Woods, 2008)

2.1.2.4 Imágenes en las bandas visibles e infrarrojas.

Teniendo en cuenta que la banda visual del espectro electromagnético es la más familiar en todas nuestras actividades, no es sorprendente que las imágenes en esta banda superen con creces a todas las demás en términos de amplitud de aplicación. La banda infrarroja a menudo se usa junto con imágenes visuales, por lo que hemos agrupado las bandas visibles e infrarrojas en esta sección con fines ilustrativos. Consideramos en la siguiente discusión aplicaciones en microscopía de luz, astronomía, teledetección, industria y aplicación de la ley. La figura 2.25 muestra varios ejemplos de imágenes obtenidas con un microscopio óptico. Los ejemplos van desde productos farmacéuticos y micro inspección hasta caracterización de materiales. Incluso solo en microscopía, las áreas de aplicación son demasiado numerosas para detallarlas aquí. No es difícil conceptualizar los

tipos de procesos que uno podría aplicar a estas imágenes, que van desde la mejora hasta las mediciones.

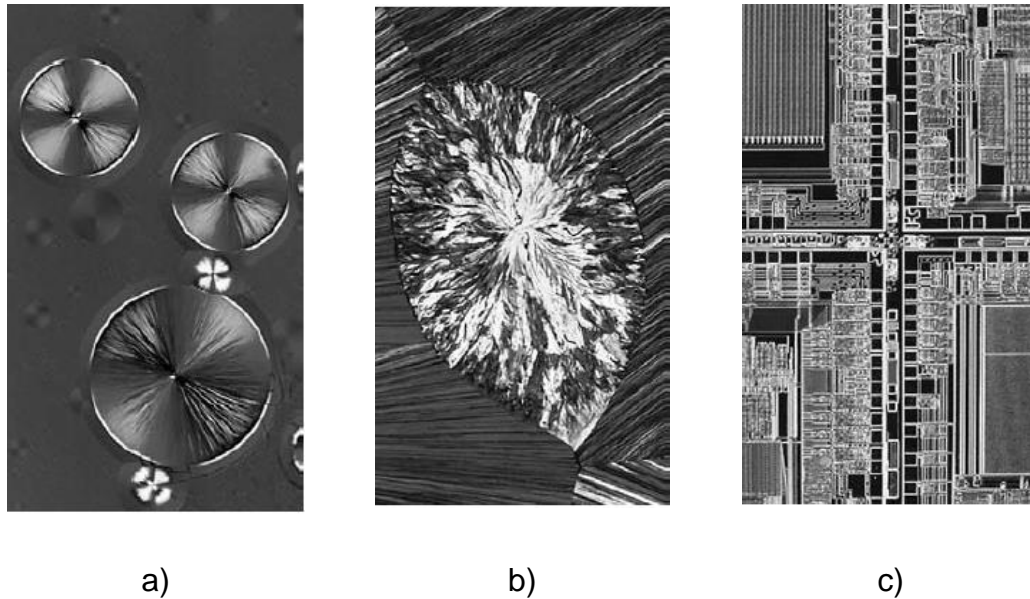


Figura 2.25 Taxol (agente anticancerígeno), magnificado 250x, Colesterol 40x y Microprocesador 60x. (González y Woods, 2008)

2.1.2.5 Imágenes en la banda de radio.

Como en el caso de las imágenes en el otro extremo del espectro (rayos gamma), las principales aplicaciones de las imágenes en la banda de radio son en medicina y astronomía. En medicina, las ondas de radio se utilizan en la resonancia magnética (MRI). Esta técnica coloca al paciente en un poderoso imán y transmite ondas de radio a través de su cuerpo en pulsos cortos. Cada pulso hace que los tejidos del paciente emitan un pulso de respuesta de ondas de radio. Una computadora determina la ubicación desde la cual se originan estas señales y su fuerza, que produce una imagen bidimensional de una sección del paciente. La resonancia magnética puede producir imágenes en cualquier plano.

La figura 2.26 muestra imágenes de resonancia magnética de una rodilla y columna vertebral humanas.



Figura 2.26 Imagen de una rodilla y una columna vertebral. (González y Woods, 2008)

2.2 Lógica Difusa.

Básicamente la Lógica Difusa es una lógica multivaluada que permite representar matemáticamente la incertidumbre y la vaguedad, proporcionando herramientas formales para su tratamiento.

De acuerdo a (Zadeh, 1973), cuando aumenta la complejidad, los enunciados precisos pierden su significado y los enunciados útiles pierden precisión. Básicamente, cualquier problema del mundo puede resolverse como dado un conjunto de variables de entrada (espacio de entrada), obtener un valor adecuado de variables de salida (espacio de salida). La lógica difusa permite establecer este mapeo de una forma adecuada, atendiendo a criterios de significado (y no de precisión).

Le término Lógica Difusa fue utilizado por primera vez en 1974. Actualmente se utiliza en un amplio sentido, agrupando la teoría de conjuntos difusos, reglas si-entonces, aritmética difusa, cuantificadores, etc.

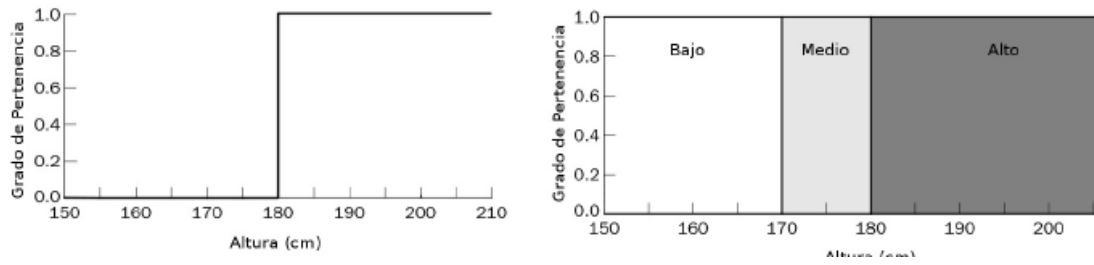
2.2.1 Valores Lingüísticos o Conjuntos difusos.

Como lógica difusa es multivaluada, es decir que a un elemento x le pueden corresponder uno o más elementos de y , en la definición de grados de pertenencia, la lógica difusa emplea valores continuos entre 0 (que representa hechos totalmente falsos) y 1 (totalmente ciertos). Así, la lógica binaria clásica puede verse como un caso particular de la lógica difusa.

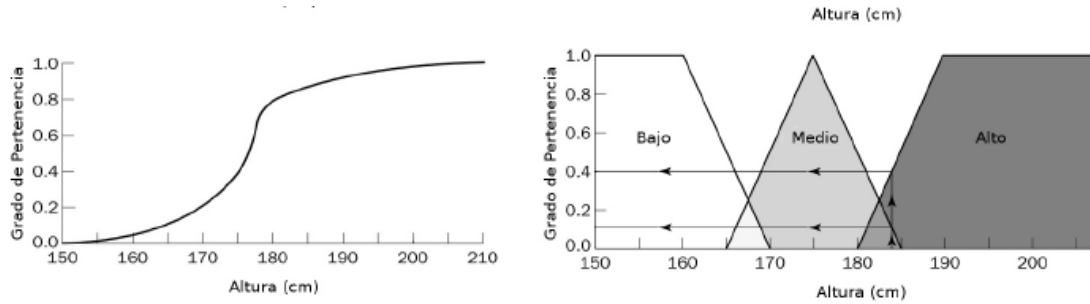
De acuerdo a (Zadeh, 1965), los conceptos se asocian a conjuntos difusos (asociando los valores de pertenencia) en un proceso llamado fuzzificación. Una vez que tenemos los valores fuzzificados podemos trabajar con reglas lingüísticas y obtener una salida, que podrá seguir siendo difusa o defuzzificada para obtener un valor discreto crisp.

De este modo, a diferencia de la teoría clásica de conjuntos que se basa en el principio básico de la lógica, de forma que un individuo pertenece o no pertenece a un conjunto, la idea básica de un conjunto difuso es que un elemento forma parte de un conjunto con un determinado **grado de pertenencia**. De este modo una proposición no es totalmente (sino parcialmente) cierta o falsa. Este grado se expresa mediante un entero en el intervalo $[0; 1]$.

A continuación se muestra la representación de la altura de una población de individuos como se muestra en la figura 2.27 a y b.



a)



b)

Figura 2.27 a) Descripción de conjuntos crisp, y b) Fuzzy de una “persona alta”. (Zadeh, 1965)

Tabla 2.1 Medidas de una Persona			
Nombre	Altura	Crisp	Fuzzy
Paco	2.05	1	1.0
Juan	1.95	1	1.0
Tomás	1.87	1	0.95
Carlos	1.80	1	0.82
Pedro	1.79	0	0.71
Andrés	1.60	0	0.36

En la representación crisp, se dibuja una línea que separa claramente en 1.8m los individuos que son altos de los que no lo son, asociando un valor de pertenencia estricto, al conjunto de los altos a aquellos que superan esa altura. Sin embargo, el conjunto difuso permite expresar que Carlos tiene un grado de pertenencia al conjunto de los altos en $\mu_A(Altura) = 0.82$.

Así, un conjunto difuso proporciona una transición suave entre los límites de lo que sería un conjunto crisp. El **Universo del discurso** se define como todos los posibles valores que puede tomar una determinada variable (en el caso de la imagen anterior se correspondería con el eje horizontal de las gráficas, desde 150 a 210cm).

2.2.1.1 Definición de conjunto difuso.

La teoría de conjuntos difusos es un intento de desarrollar una serie de conceptos para tratar de un modo sistemático el tipo de imprecisión que aparece cuando los límites de las clases de objetos no están claramente definidos. Un conjunto difuso puede definirse como una clase en la que hay una progresión gradual desde la pertenencia al conjunto hasta la no pertenencia; o visto de otra forma, en la que un objeto puede tener un grado de pertenencia definido entre la pertenencia total (valor uno) o no pertenencia (valor cero). Desde esta perspectiva, los conjuntos convencionales (o conjuntos crisp) pueden verse como un caso particular de conjuntos difusos; un conjunto difuso que sólo admite dos grados de pertenencia (uno y cero).

Un conjunto difuso puede definirse de forma general como un conjunto con límites difusos. Sea **X** el Universo del discurso, y sus elementos se denotan como **x**. En la teoría clásica de conjuntos crisp se define un conjunto **C** se define sobre **X** mediante la función característica de **C** como **fc**.

$$f_c(x) = \begin{cases} 1 & \text{cuando } x \in C \\ 0 & \text{cuando } x \text{ no } \in C \end{cases} \quad (2.8)$$

Este conjunto mapea el universo X en un conjunto de dos elementos, donde la función $fc(x)$ es 1 si el elemento x pertenece al conjunto C y 0 si el elemento x no pertenece al conjunto C . Si generalizamos esta función para que los valores asignados a los elementos del conjunto caigan en un rango particular y así indicar el grado de pertenencia de los elementos a ese conjunto, tendremos una **función de pertenencia** de un determinado conjunto difuso. La función de pertenencia μ_A por la que se define un conjunto difuso A, sería:

$$\mu_A = X \rightarrow [0,1] \quad (2.9)$$

Donde $\mu_A = 1$ si x está totalmente en A, $\mu_A = 0$ si x no está en A y $0 < \mu_A < 1$ si x está parcialmente en A. Este valor entre 0 y 1 representa el grado de pertenencia (también llamado valor de pertenencia de un elemento x a un conjunto A. Así, el intervalo de la ecuación anterior es de números reales e incluye los extremos. Aunque $[0,1]$ es el rango de valores más utilizado para representar funciones de pertenencia, cualquier conjunto arbitrario con alguna ordenación total o parcial podría ser utilizado.

2.2.1.2 Operaciones de Conjuntos Difusos.

Las tres operaciones básicas que se definen sobre conjuntos crisp (complemento, unión e intersección), pueden generalizarse de varias formas en conjuntos difusos. No obstante, existe una generalización particular que tiene especial importancia. Cuando se restringe el rango de pertenencia al conjunto $[0; 1]$, estas operaciones “estándar” sobre conjuntos difusos se comportan de igual modo que las operaciones sobre conjuntos crisp.

$$\begin{aligned} \mu_{\bar{A}}(x) &= 1 - \mu_A(x) \\ \mu_{A \cap B}(x) &= \perp [\mu_A(x), \mu_B(x)] \\ \mu_{A \cup B}(x) &= T [\mu_A(x), \mu_B(x)] \end{aligned} \quad (2.10)$$

Unión

La forma generalizada de la unión es la T-conorma. Podemos definirla con la siguiente función:

$$\begin{aligned} \perp: [0,1] \times [0,1] &\rightarrow [0,1] \\ \mu_A \cup B(x) &= \perp [\mu_A(x), \mu_B(x)] \end{aligned} \quad (2.11)$$

Para que una función se pueda considerar como una unión difusa, debe satisfacer los siguientes axiomas $\forall a, b, c \in [0,1]$:

- U1) Elemento Neutro: $\perp (a, 0) = a$
- U2) Conmutatividad: $\perp (a, 0) = \perp (b, a)$
- U3) Monotonicidad: si $a \leq c$ y $b \leq d$ entonces $\perp (a, b) = \perp (c, d)$
- U4) Asociatividad: $\perp (\perp (a, b), c) = \perp (a, \perp (b, c))$

Algunas T-conormas ampliamente utilizadas son:

- Máximo: $\perp (a, b) = \max(a, b)$
- Producto: $\perp (a, b) = (a + b) - (a \times b)$
- Suma limitada (o de Lukasiewick): $\perp (a, b) = \min(a + b, 1)$

Intersección

La forma generalizada de la intersección se denomina T-norma. Es una función de la forma:

$$\begin{aligned} T: [0,1] \times [0,1] &\rightarrow [0,1] \\ \mu_A \cap B(x) &= T [\mu_A(x), \mu_B(x)] \end{aligned} \quad (2.12)$$

Una T-norma satisface los siguientes axiomas $\forall a, b, c \in [0,1]$

I1) Elemento unidad: $T(a, 1) = a$

I2) Conmutatividad: $T(a, b) = T(b, a)$

I3) Monotonicidad: Si $a \leq c$ y $b \leq d$ entonces $T(a, b) = T(c, d)$

I4) Asociatividad: $T(T(a, b), c) = T(a, T(b, c))$

Algunas T-normas ampliamente utilizadas son:

Mínimo: $T(a, b) = \min(a, b)$

Producto algebraico: $T(a, b) = ab$

Diferencia limitada (o de Lukasiewick): $T(a, b) = \max(0, a + b - 1)$

Complemento

El complemento \bar{A} de un conjunto difuso A, se denota por cA ; está definido por una función del tipo $c : [0,1] \rightarrow [0,1]$. Tiene que satisfacer los siguientes axiomas:

C1) Condiciones límite o frontera: $c(0) = 1$ y $c(1) = 0$.

C2) Monotonicidad: $\forall a, b \in [0,1]$ si $a < b$ entonces $c(a) \geq c(b)$.

C3) c es una función continua.

C4) c es involutiva $\forall a \in [0,1]$ tenemos $c(c(a)) = a$.

Al igual que sucedía con los operadores de unión y de intersección, también para el complemento existen gran variedad de clases. Uno de los más utilizados, además del complemento clásico ($\mu_{\bar{A}}(x) = c(a) = 1 - a$), es el λ -complemento de Sugeno, que viene definido por la siguiente expresión:

$$\mu_{\bar{A}^\lambda}(x) = \frac{1 - \mu_A(x)}{1 + \lambda \mu_A(x)} \text{ con } \lambda \in (-1, \infty) \quad (2.13)$$

Como se puede observar, si $\lambda = 0$, la función se comporta como el complemento clásico. Además, para cada valor de λ , obtenemos una expresión particular para el complemento. Otro tipo de complemento borroso muy utilizado es el de Yager, que se define con la siguiente expresión:

$$\mu_{A^{\lambda}}(x) = (1 - \mu_A(x)^w)^{1/w} \text{ con } w \in (0, \infty) \quad (2.14)$$

Al igual que con el complemento de Sugeno, cambiando el valor de w obtenemos distintos tipos de complemento. Si $w = 1$ tenemos el complemento clásico como se muestra en la figura 2.28.

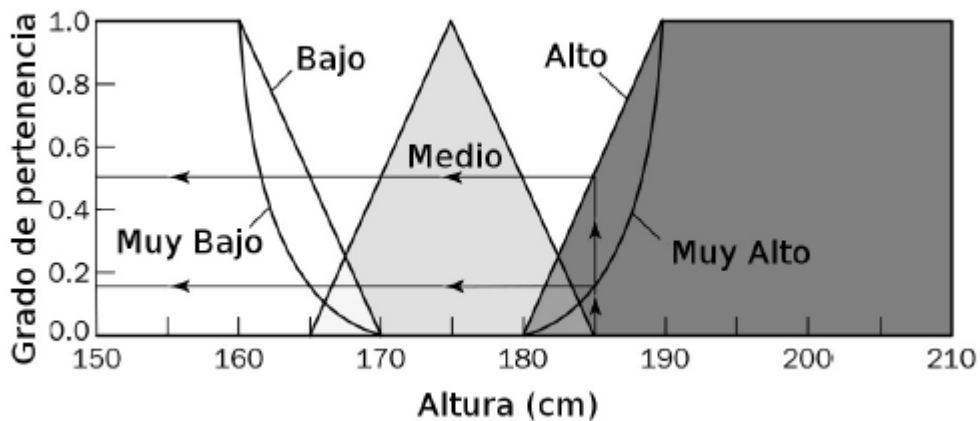


Figura 2.28 Uso del modificador muy en los conjuntos bajo y alto. (Zadeh, 1965)

2.2.1.3 Propiedades.

Los conjuntos Crisp y los difusos tienen las mismas propiedades (en realidad los conjuntos crisp pueden verse como un subconjunto de los conjuntos difusos).

Conmutativa: $A \cap B = B \cap A$

Asociativa: $A \cup (B \cap C) = (A \cup B) \cap C$

Distributiva: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Idempotencia: $A \cup A = A$ y $A \cap A = A$

Involución: $\neg(\neg A) = A$

Transitiva: $\text{If } (A \subset B) \cap (B \subset C) \text{ then } A \subset C^1$

Leyes de Morgan: $\neg(A \cap B) = \neg A \cup \neg B$ y $\neg(A \cup B) = \neg A \cap \neg B$

Empleando estas operaciones, propiedades y modificadores se pueden obtener gran variedad de expresiones. Por ejemplo, siendo **A** el conjunto alto y **B** bajo, podemos derivar el conjunto **C** como no muy alto y no muy bajo como

$$\mu_C(x) = [1 - \mu_A(x)^2] \cap [1 - \mu_B(x)^2] \quad (2.15)$$

2.2.1.4 Representación de conjuntos difusos.

Los conjuntos crisp son útiles pero presentan problemas en muchas situaciones. Examinando el Universo del discurso de la altura, tendríamos la representación gráfica de la figura 2.27. Para definir un conjunto difuso hay que definir su función de pertenencia. Un método habitual es preguntar a un experto sobre el dominio del problema y representarlo mediante diferentes funciones (típicamente triangulares y trapezoidales). También se pueden utilizar, como veremos más adelante, funciones curvas o la función singleton. Para representar un conjunto difuso continuo en un ordenador necesitamos

expresar esa función de pertenencia y mapear los elementos del conjunto con su grado de pertenencia. Aunque puede usarse a priori cualquier tipo de función, en la práctica se emplean funciones lineales con una descripción de su **vector de ajuste**, como:

$$\text{hombre-medio} = (0/165, 1/175, 0/185)$$

Es decir que un hombre de estatura mediana toma valor 0, en la estatura 165, valor 1 en la estatura 175 y valor 0 nuevamente en la estatura 185.

Esta representación se corresponde con el conjunto difuso Medio de la Figura 2.28, donde para la altura 165 se asocia el grado de pertenencia 0, a la altura 175 el grado de pertenencia 1, y de nuevo a la altura 185 el grado de pertenencia 0

2.2.2 Variables Lingüísticas.

Para representar el conocimiento en razonamiento aproximado tenemos que utilizar variables lingüísticas. Una variable lingüística (Zadeh, 1975) es aquella cuyos valores son palabras o sentencias en un lenguaje natural o artificial. De esta forma, una variable lingüística sirve para representar cualquier elemento que sea demasiado complejo, o del cual no tengamos una definición concreta; es decir, lo que no podemos describir en términos numéricos.

Así, una variable lingüística está caracterizada por una quintupla

$$(X, T(X), U, G, M)$$

X: es el nombre de la variable.

T(X): es el conjunto de términos de X; es decir, la colección de sus valores lingüísticos (o etiquetas lingüísticas).

U: es el universo del discurso (o dominio subyacente). Por ejemplo, si la hablamos de temperatura “Cálida” o “Aproximadamente 25°”, el dominio subyacente es un dominio numérico (los grados centígrados).

G: es una gramática libre de contexto mediante la que se generan los términos en $T(X)$, como podrían ser “muy alto”, “no muy bajo”, etc.

M: es una regla semántica que asocia a cada valor lingüístico de X su significado $M(X)$ ($M(X)$ denota un subconjunto difuso en U).

Los símbolos terminales de las gramáticas incluyen:

Términos primarios: “bajo”, “alto”, etc.

Modificadores: “Muy”, “más”, “menos”, “cerca de”, etc.

Conectores lógicos: Normalmente NOT, AND y OR.

Normalmente se definen los conjuntos difusos de los términos primarios y, a partir de éstos, se calculan los conjuntos difusos de los términos compuestos (por ejemplo, con “muy” y “alto” construimos el término compuesto “muy alto”). Una etiqueta lingüística se forma como una sucesión de los símbolos terminales de la gramática: “Muy alto, no muy bajo, etc”.

Un uso habitual de las variables lingüísticas es en reglas difusas. Ejemplo: IF duración-examen IS larga THEN probabilidad-aprobar IS small. Por ejemplo, la variable lingüística velocidad podría incluir conjuntos difusos como muy lento, lento, medio, rápido, muy-rápido. Naturalmente cada uno de estos conjuntos representa un valor lingüístico que puede tomar la variable.

2.2.3 Aplicaciones.

Desde mediados de los años 70, la lógica difusa se ha utilizado ampliamente debido a varios factores. Uno de ellos es que el uso de conocimiento experto

permite la **automatización de tareas**. En muchas áreas de aplicación se reduce considerablemente la necesidad de operadores que basan su conocimiento en la experiencia (y que difícilmente podría ser expresado con ecuaciones diferenciales). De este modo, si existe un conocimiento del proceso, es posible modelarlo mediante lógica difusa. Los sistemas basados en lógica difusa son **fáciles de diseñar, modificar y mantener**. Pese a la pérdida de precisión, la reducción de tiempo de desarrollo y mantenimiento es muy relevante para su uso industrial.

Otro factor a tener en cuenta es que el control difuso permite diseñar soluciones de alta calidad, que eviten las patentes existentes en otros sistemas de control. En Japón este tipo de controladores se asocia a modernidad, alta calidad y tecnológicamente potente. En Europa sin embargo se trata de ocultar el término “difuso” por su significado negativo. En la actualidad multitud de productos de electrónica de consumo emplean lógica difusa.

Por citar algunos ejemplos de uso, la empresa Japonesa Matsuhita utiliza en sus lavadoras un sistema de control que determina automáticamente el ciclo de lavado según el tipo, cantidad de suciedad y tamaño de la colada. Los estabilizadores de imágenes en sus cámaras digitales incorporan reglas que eliminan las vibraciones involuntarias de la mano del operario, comparando la imagen actual con las imágenes anteriores de la memoria. En el ámbito de la automoción, Mitsubishi y General Motors emplean sistemas de transmisión automática y control de temperatura basados en lógica difusa.

Otro caso de éxito es el metro de Sendai (Japón), que cuenta con 16 estaciones. El sistema de control difuso está dividido en dos módulos, uno para el control de la velocidad (similar al de [1]) y otro para la parada automática. Este controlador difuso ofrece importantes ventajas sobre los controladores convencionales, como el mayor confort en el viaje para los pasajeros y menor consumo de energía.

2.2.3.1 Ejemplos de aplicaciones.

Menciona (Duarte, 2000) los siguientes ejemplos de aplicación:

1. CONTROLADOR DIFUSO ADAPTATIVO.

Quizás la aplicación en que la Lógica Difusa ha conseguido un éxito mayor, y por ende un mayor número de seguidores, se encuentra en el Control Industrial. Aun cuando existen numerosas versiones de controladores que emplean lógica difusa, suele asignarse el término Controlador Difuso a un sistema de control cuya estructura interna corresponde a la de la figura 2.29.

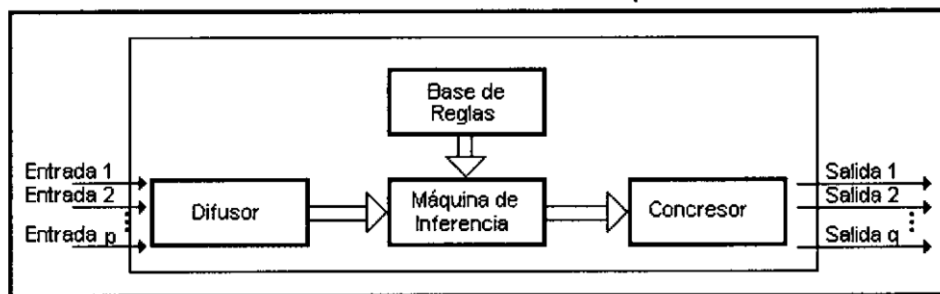


Figura 2.29 Estructura interna de un Controlador Difuso. (Duarte, 2000)

Una de las ventajas que suelen mencionarse de los Controladores Difusos, frente a otro tipo de controladores, es que pueden diseñarse aunque no se tenga un modelo matemático exacto de la Planta a controlar, gracias a que están basados en reglas. Esta situación, sin embargo, no es del todo sencilla: el no tener un modelo matemático de la Planta implica no poder realizar simulaciones sobre la misma, de tal manera que los ajustes del controlador deben realizarse en línea.

Este hecho ha promovido la aparición de controladores autoajustables, que cuentan con algún algoritmo que les permite evaluar su desempeño, y de acuerdo con cierto criterio ajustar su diseño. Estos algoritmos de autoajuste (también conocidos como algoritmos de auto sintonía) son muy útiles también en situaciones en las que la Planta a controlar varía en el tiempo, lo que implica que el controlador deba ajustarse a dichos cambios.

2. IDENTIFICADOR DE IMÁGENES AÉREAS.

El problema aquí planteado es el siguiente: Se tienen tres imágenes aéreas de la misma zona; las imágenes se han captado empleando cámaras a blanco y negro, de tal forma que muestran en cada pixel un cierto nivel de gris (nivel de luminancia); las tres fotografías no son iguales, porque cada una de ellas se ha tomado anteponiendo a la cámara un filtro que sólo permite captar una franja de colores (unas ciertas longitudes de onda), y los filtros para cada fotografía han sido diferentes. En estas imágenes se ha captado un área extensa con zonas que se pueden clasificar así:

- Zonas de río.
- Zonas de construcciones humanas.
- Zonas dedicadas a la agricultura.
- Zonas boscosas.

El problema consiste en diseñar un algoritmo que, conociendo los niveles de luminancia para un cierto pixel en las tres fotografías, decida a cuál de las zonas anteriores corresponde ese pixel.

No existe una única combinación de luminancias que identifique a cada una de las zonas (no hay un único verde en la zona de bosques), y por tanto no se conoce el conjunto de todas las combinaciones de luminancias posibles asociadas a cada zona.

Nótese que el problema puede replantearse en términos de Conjuntos Difusos: Se puede definir como Universo de Discurso el conjunto de todas las combinaciones posibles de luminancias, con lo cual la tarea consiste en encontrar cuatro Conjuntos Difusos definidos sobre dicho universo, cada uno de los cuales debe representar a una de las zonas de las imágenes.

3. BASE DE DATOS DIFUSA.

La Lógica Difusa busca desarrollar un conjunto de procedimientos para manejar la información precisa y/o vaga. Ahora bien, los Sistemas de Bases de Datos tienen por propósito, hablando en términos muy generales, la organización de la información; por tanto no es de sorprendente que se haya intentado incorporar las técnicas de Lógica Difusa en el diseño de Bases de Datos.

De acuerdo a (Miyamoto y Umanodistinguen, 1995) dos tipos de técnicas difusas en las Bases de Datos:

Bases de Datos Difusas.

Técnicas Difusas para la recuperación de la información.

En la primera de estas técnicas el concepto de Conjunto Difuso se incorpora en la estructura misma de la Base de Datos, mientras que en la segunda se emplea en las estrategias de recuperación de la información. A continuación se presenta una de las técnicas del primer tipo. La técnica en cuestión se denomina Búsqueda Difusa, y fue propuesta inicialmente por (Tahani, 1977). En esta técnica la Base de Datos tiene dos componentes": por una parte se tiene una Base de Datos tradicional, como la representada en el tabla 2, y por otra parte se tiene una definición Difusa de las variables cuantificables de dicha cuadro, como la que se muestra en la figura 2.30.

Tabla 2.2 Base de Datos tradicional del Ejemplo			
Nombre	Edad	Salario	Año de Ingreso
Anderson	30	20.000	1995
Brown	30	15.000	1995
Long	25	40.000	1993
Nelson	55	20.000	1980
Smith	25	23000	1996

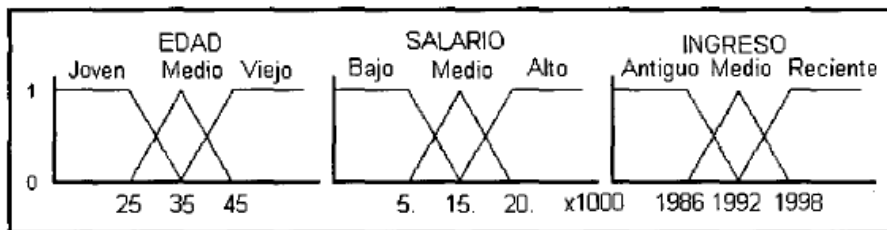


Figura 2.30 Variables lingüísticas para el ejemplo. (Miyamoto y Umanodistinguen, 1995)

Ante una consulta a la Base de Datos de la forma:

¿"Cuáles son los nombres de las personas jóvenes o recientemente empleadas pero con sueldo alto?"

Propone (Tahani, 1977) evaluar la función de pertenencia de cada registro a cada uno de los valores lingüísticos involucrados en la consulta, y entregar como resultado de la búsqueda un conjunto difuso con funciones de pertenencia obtenidas mediante la utilización de operadores AND, OR YNOT difusos, los operadores empleados son el mínimo para el AND, el máximo para el OR y el complemento para el NOT.

Con esta metodología la consulta anterior podría representarse por la operación:

$$\mu_{busueda}(x, y, z) = \{\mu_{joven}(x) \text{ OR } \mu_{reciente}(y)\} \text{ AND } \mu_{alto}(z) \quad (2.16)$$

En donde **x** es la edad, **y** el año de ingreso, **z** el salario, $\mu_{joven}(x)$ es la función de pertenencia al conjunto Joven definido sobre la variable edad, $\mu_{reciente}(y)$ es la función de pertenencia al conjunto Reciente definido sobre la variable Año de Ingreso, $\mu_{alto}(z)$ es la función de pertenencia al conjunto Alto definido sobre la variable salario, y $\mu_{busueda}(x, y, z)$ es la función de pertenencia al conjunto resultante de la búsqueda.

Empleando la información de la tabla 2 y de la figura 2.29 se puede construir la tabla 2.3 que muestra los resultados de la búsqueda para cada registro.

Tabla 2.3 Resultados de la búsqueda del ejemplo en cada registro				
Nombre	$\mu_{joven}(x)$	$\mu_{reciente}(y)$	$\mu_{alto}(z)$	$\mu_{busueda}(x, y, z)$
Anderson	0.5	0.6	0.5	0.5
Brown	0.5	0.6	0.0	0.0
Long	1.0	0.2	1.0	1.0
Nelson	0.0	0.0	0.5	0.0
Smith	1.0	0.8	0.8	0.8

El resultado final del ejemplo sería el conjunto:

Búsqueda = {0,5/Anderson, 1,0/Long, 0,8/Smith}

2.3 Parálisis facial.

La parálisis facial es un trastorno del nervio que controla el movimiento de los músculos, este daño al nervio causa la debilidad o parálisis esto significa que se puede usar los músculos (MedlinePlus, 2019).

En México la parálisis facial es una enfermedad más común en mujeres que en hombres y en personas de entre 15 y 50. Cada año se diagnostican aproximadamente 20 mil nuevos casos en el país. La secretaria de salud federal informo que el 20 por ciento de los casos de parálisis tienen secuelas permanentes entre ellas incapacidad de cerrar los ojos para sonreír, o la falta de movilidad de una área de la cara, lo que percute en su simetría y calidad de vida. El trastorno antes mencionado se diagnostica, por medio de una presentación con un especialista, que revisa el aspecto facial deformado, buscando debilidad facial superior e inferior y la incapacidad de mover los músculos del lado afectado de la cara. En la mayoría de los casos solo se presenta en una parte de la cara, la frente, el parpado o la boca. No obstante no existe un análisis específico de laboratorio para poder confirmar el diagnóstico. Existe una prueba llamada electromiografía (EMG) puede confirmar la presencia de daño nervioso y poder determinar la gravedad y el alcance del trauma nervioso (Diagnostican al año 20 mil casos de parálisis facial en México, 2015).

2.3.1 Tipos de parálisis.

De acuerdo a (González, 2001) presenta los siguientes tipos de Parálisis Facial:

Parálisis de Bell (idiopática)

Se define como aquella parálisis en la que no se puede encontrar la causa (idiopática). Es un diagnóstico de exclusión que debe realizarse ya que se descartaron con estudios adecuados todas las causas conocidas de parálisis facial. Es la causa más frecuente de parálisis facial la cual representa del 50 al 80% de todas las parálisis faciales (Hoffer, 1995).

No tiene predilección de sexo, ni edad, ni grupo geográfico. Sin embargo, aumenta su incidencia en menores de 13 años y mayores de 65. Generalmente se presenta de forma súbita (57%), sin otra sintomatología sistémica o neurológica acompañante, siendo en la mayoría de las veces parcial o incompleta. Puede encontrarse dolor o parestesias en oído, cara, cuello o lengua hasta en el 50% de los casos. Y hasta en un 14% de los casos se encuentra una historia familiar.

En la parálisis de Bell el reflejo estapedial se encuentra ausente, el ipsilateral del oído del lado de la lesión en un 90% de los casos. Tiene un nivel pronóstico, pues en pacientes que conservan el reflejo después de las 2 semanas se espera una recuperación completa. Antes de las 2 primeras semanas del comienzo de la parálisis el reflejo estapedial no tiene ninguna significancia en el pronóstico.

De las múltiples teorías que se han formulado, dos son las que tiene mayor aceptación: un origen viral por infección del virus herpes simple tipo 1 (HSV-1) y un origen isquémico por disminución en la irrigación del séptimo par por la vasa nervorum.

Esta entidad es un diagnóstico de exclusión por lo que el galeno tendrá que mantenerla como última posibilidad diagnóstica después de haber descartado otras causas más ominosas (Stankiewicz, 1991).

Parálisis traumática

Es la segunda causa más frecuente, la cual comprende del 8-22% de los casos. Una porción significativa ocurre durante el nacimiento. Otras causas, aunque menos frecuentes, son: el barotrauma, fracturas faciales, fracturas del hueso temporal y heridas penetrantes en región parotídea. Se debe evaluar por un equipo multidisciplinario que conste de médicos otorrinolaringólogos, neurocirujanos y neurorradiólogos. Dentro de las primeras horas se debe pedir al menos una tomografía computada de alta resolución de región temporal comparativa (Fisch, 1981).

Parálisis infecciosa

Considerada por algunos autores como la causa más común de parálisis facial, debido a que incluyen bajo este rubro a la parálisis de Bell. Siendo el porcentaje de presentación entre el 5 y 15.3%. Después de la parálisis de Bell la etiología más frecuente de la parálisis facial infecciosa es la causada por el virus de la Varicela Zoster, el cual causa el síndrome de Ramsay y Hunt causante de un 4.5% a un 8.9% de los casos.

El síndrome de Ramsay Hunt (herpes Zoster ótico) el cual se identifica por parálisis facial acompañada de otalgia y erupciones vesiculares en el pabellón auricular, en el conducto auditivo externo, en la membrana timpánica y ocasionalmente en cavidad oral y cuello. Las lesiones vesiculares pueden no aparecer sino hasta 3-7 días después del inicio de la parálisis y la preceden en un 25% de los casos. El octavo par craneal frecuentemente también es afectado por lo que se presenta una hipoacusia acompañante. Sólo el 10% de los pacientes con parálisis total presentan recuperación completa y 66% de los

pacientes con parálisis incompleta presentan una recuperación total (Mattox, 1998).

Otras causas infecciosas incluyen el virus de polio-mielitis tipo 1, el virus de las parotiditis, virus de la influenza, virus Coxsackie, virus de Epstein Barr, virus de inmunodeficiencia humana (HIV), mononucleosis infecciosa. También se han reportado casos de parálisis faciales posteriores a vacunación con vacunas anti polio, rabia o influenza (Adour, 2000).

También infecciones que envuelvan al oído medio tales como otitis media aguda (en el 0.6% de los pacientes), otitis media crónica (en el 6.3% de los pacientes), mastoiditis, el tratamiento de estas patologías se basa en antibiótico específico y drenaje del oído medio por medio de una miringotomía o mastoidectomía según el caso por lo cual su manejo deberá ser por un especialista.

Infecciones sistémicas como enfermedad de Lyme, tuberculosis, sífilis, tétanos, botulismo, lepra, malaria y enfermedad por arañazo de gato, al igual que infecciones micóticas como mucormicosis o aspergilosis que en su trayecto pueden involucrar al nervio facial (Devriese, 1988).

Parálisis neoplásica

Se representa del 5% al 13% de las parálisis. Generalmente se trata de una parálisis de comienzo insidioso, progresiva, aunque puede presentar un patrón intermitente con recurrencias, y de larga evolución. Una masa en parótida, glándula submaxilar o cuello, historia de cáncer en otras partes más comúnmente (mama, pulmón, tiroides y aparato genitourinario) nos debe hacer sospechar el diagnóstico y se pedirá una tomografía y/o una resonancia magnética para comprobar el diagnóstico (Telisch, 2000).

Las dos principales causas son por tumores primarios de ángulo pontocerebeloso (nerinoma del acústico) y tumores de región parotídea. (Carcinoma adenoideo quisítico). Otras causas son debidas a metástasis con primarios en mama, pulmón, tiroides, riñón, ovario, melanoma y próstata. El manejo es según la histología tumoral.

Parálisis neurológica

Ésta puede ser la causa en el 15% al 30% de los casos, son varias las patologías que están involucradas principalmente por su frecuencia, daremos mayor explicación a dos de ellas:

Síndrome de Melkersson-Rosenthal. Se compone de una tríada: Edema recurrente orofacial, fisura lingual, y parálisis facial recurrente. Siendo el signo más característico para el diagnóstico el edema orofacial. La tríada completa sólo se presenta en el 25% de los casos. La parálisis facial es de inicio súbito y es común que presente una historia de parálisis bilaterales secuenciales y recurrentes.

Puede representar una forma de sarcoidosis. Síndrome de Guillain-Barré (polirradiculoneuropatía aguda inflamatoria ascendente) es la principal causa de parálisis bilateral simultánea, el síntoma principal es la debilidad motora que puede llegar a la parálisis de inicio en miembros inferiores con un ascenso en horas a días. La parálisis facial se presenta en un 50% de los casos. Algunos autores reportan una recuperación espontánea tan alta como en la parálisis de Bell (Marenda, 1991).

Otras causas neurológicas pueden ser la esclerosis múltiple, la miastenia gravis, la neuropatía hereditaria hipertrófica, el síndrome de Möbius (parálisis congénita del nervio facial que puede estar acompañado de parálisis de otros pares craneales como el III, IV, VI, IX, X, XI, XII), y accidentes cerebro vasculares.

La combinación de la alteración de los pares VI y VII debe sugerir una lesión del tallo cerebral, la alteración de los nervios craneales VII, IX, X y XI sugieren patología a nivel de la fosa posterior o de la base de cráneo y la afección del apex petroso incluye la alteración de los pares V, VI, VII y VIII. Cuando se encuentran anomalías craneales salteadas se debe pensar en una polineuropatía omieloma múltiple (May, 1991).

Parálisis sistémica metabólica

Las enfermedades implicadas en la parálisis facial incluyen la diabetes mellitus, el hipertiroidismo, la hipertensión arterial, el embarazo, la porfiria aguda, los síndromes autoinmunes, la sarcoidosis, la amiloidosis, la toxicidad por monóxido de carbono, toxicidad por metales pesados, la deficiencia de vitamina A, la ingestión de etilenglicol, y el alcoholismo. Algunos autores refieren que existe 4.5 veces más de posibilidad de presentar parálisis de Bell en los diabéticos que en la población normal, al igual que durante el embarazo es un aumento de 3.3 veces más frecuente la presentación de ésta (Schaitkin, 2000).

2.3.2 Tratamiento.

El manejo de la parálisis facial se plantea con un enfoque multidisciplinario y una estrategia basada en las necesidades individuales de cada paciente. Esto involucra la participación de múltiples especialidades médicas tales como oftalmología, otorrinolaringología, genética, fisioterapia, cirugía plástica y neurología, entre otras. Además se debe contar con el apoyo y expertise de otros profesionales de la salud como kinesiólogos, fonoaudiólogos y psicólogos, que son claves para la obtención de buenos resultados para los pacientes.

Desde un inicio se debe comprometer al paciente en la toma de decisiones junto con establecer una dinámica que permita una adherencia adecuada al tratamiento, especialmente en relación a la rehabilitación que será fundamental para un resultado óptimo (Benítez, 2016).

Ejemplos de los tratamientos para algunas de las diferentes tipos de parálisis son:

Parálisis de Bell (idiopática)

Su tratamiento, aunque controversial, ya que se trata por lo general de una parálisis incompleta, se basa en Prednisona (2 mg/kg/día) y aciclovir (200-400 mg cada 4 h) ambos por vía oral, durante 10 días, teniendo especial cuidado con pacientes diabéticos. Acompañándose de terapia física de rehabilitación (son sugeridos los ejercicios faciales en frente del espejo, aislando cada grupo muscular cuando no hay suficiente fuerza para la realización activa de estos ejercicios el paciente debe ayudarse con sus dedos). Además no se debe olvidar la importancia en cualquier caso de parálisis facial que altera el cierre palpebral, la sustitución de lágrimas con metilcelulosa o cualquier otra solución estéril que mantenga hidratada la conjuntiva y la córnea para evitar infecciones y ulceraciones que pueden dañar la visión (Stankiewicz, 1991).

Parálisis traumática

En general se acepta que no se debe de realizar un procedimiento quirúrgico si la parálisis es incompleta. Si la parálisis es completa e inmediata en un paciente, por lo demás estable, la descompresión quirúrgica del séptimo par está indicada. Si la parálisis facial es completa pero tardía el papel de la cirugía es controversial. El manejo de la parálisis facial por heridas penetrantes sigue el mismo principio (Fisch, 1981).

Parálisis infecciosa

El tratamiento es a base de esteroides (Prednisona 1 a 2 mg/kg/día) y aciclovir (400 a 800 mg/dosis/cada 4 h). El 15% de los pacientes presentan criterios quirúrgicos de descompresión (Mattox. 1998).

Parálisis neoplásica

Sin embargo, en el 47% de los casos, los pacientes con neoplasia comienzan con un inicio brusco de la parálisis similar a la parálisis de Bell respondiendo a esteroides, recurriendo al poco tiempo después (Telisch, 2000).

Independientemente de estos tratamientos también se recomienda también poner atención a la protección ocular ya que esto puede dejar expuesto el ojo dejándolo expuesto a la irritación y la resequedad, por ello es importante mantener el ojo húmedo y protegerlo de lesiones (Parálisis de Bell, 2016).

2.3.3 Evolución.

El pronóstico para los individuos con parálisis generalmente es bueno, pero en ocasiones el tiempo de recuperación varia, la mayoría comienzan mejorando dentro de las dos semanas del inicio de los síntomas recuperándose completamente, regresando a su función normal dentro de los tres a seis meses, mientras que para otros este periodo se puede extender ya que de acuerdo a especialistas dicen que a la percepción de los pacientes no ven una mejoría y pueden deprimirse y hacer más tardado la rehabilitación (Parálisis de Bell, 2016).

III DESARROLLO EXPERIMENTAL.

Como primera propuesta se presentó la adquisición de imágenes de profundidad por medio del sensor Kinect ya que se pensó que la variante Z podría darnos mejor precisión.

3.1 Sensor Kinect.

El sensor Kinect es un controlador creado por Microsoft, inicialmente fue desarrollado para la consola de juegos Xbox 360 posteriormente se desarrolló una versión para PC. Lo innovador de este sensor es que los usuarios pueden interactuar con los juegos sin tener la necesidad de contacto físico, el sensor únicamente trabaja mediante reconocimiento de gestos, movimientos y comandos de voz. Este sensor se compone de una cámara RGB, un sensor de profundidad, un micrófono multi-array y un procesador personalizado para la ejecución del software que permite el reconocimiento del movimiento del cuerpo. En la figura 3.1 se puede apreciar este sensor.



Figura 3.1 Sensor Kinect.

3.1.2 Especificaciones.

Las especificaciones dadas por (ESPECIFICACIONES TÉCNICAS DE KINECT, 2010):

Sensores

- Lentes de color y sensores de profundidad
- Micrófono multi-arreglo
- Ajuste de sensor con su motor de inclinación

Campo de visión

- Campo de visión horizontal: 57 grados
- Campo de visión vertical: 43 grados
- Rango de inclinación física: ± 27 grados
- Rango de profundidad: 1.2 –3.5 metros

Data Streams (Flujo de datos)

- 320 x 240 a 16 bits de profundidad @ 30fps
- 640 x 480 32-bit de color @30fps
- Audio de 16-bit @ 16 kHz

Sistema de Seguimiento

- Rastrea hasta 6 personas, incluyendo 2 jugadores activos.
- Rastrea 20 articulaciones por jugador activo
- Chat en vivo y voz dentro del juego (requiere Xbox Live Gold).
- Sistema de cancelación de eco que aumenta la entrada de voz
- Reconocimiento de voz múltiple.

3.1.3 Funcionamiento.

El sensor Kinect basa su funcionamiento en la combinación de la información que obtiene de la cámara RGB y del sensor de profundidad. Básicamente, adquiere una imagen RGB-D con una resolución de 320x240 píxeles, a través de la cámara RGB y del sensor de profundidad, asignando información de color y profundidad a cada píxel. A continuación se describe el funcionamiento de cada una de las partes que conforman este dispositivo.

Sensor de profundidad

Basa su funcionamiento en una técnica llamada “luz estructurada” la cual consiste en la proyección de puntos infrarrojos sobre una superficie, al hacer corresponder la imagen captada con el patrón original, es posible triangular la posición de cada píxel y determinar su profundidad con respecto al plano perpendicular de la cámara (Kinect docs, 2012).

Emisión del patrón de puntos infrarrojos

Cuando una superficie es iluminada por una onda de luz, cada punto actúa como una fuente de onda esférica. La luz en cualquier punto en el campo de luz dispersa, se compone de ondas que han sido emitidas desde cada punto de la superficie iluminada. Si la superficie es suficiente rugosa para crear longitud de trayectoria superior a la longitud de onda del punto, se da lugar a cambio de fase mayor que la amplitud, y por lo tanto la intensidad.

La tecnología usada por el Kinect para adquirir la imagen de profundidad “LightCoding”, la cual utiliza un emisor infrarrojo para la proyección del patrón de puntos, los cuales son invisibles para el ojo humano. Cada punto del patrón está hecho para que sea diferente a los demás, de esta forma, la cámara infrarroja lee la luz codificada y ejecutan su sofisticado algoritmo de procesamiento paralelo para descifrar el código de la luz y poder generar una imagen de profundidad. El método es sensible en la luz ambiental (Kinect docs, 2012).

Triangulación

La visión por computadora, la triangulación se refiere al proceso para determinar la distancia de un punto en un espacio 3D a partir de la visión de dos o más imágenes. La figura muestra la geometría epipolar de un par de cámaras de visión estéreo.

Un punto P es proyectado en un espacio tridimensional, que pasa por el punto focal de cada cámara, lo que resulta en dos puntos correspondientes de imagen, si estos puntos son conocidos y la geometría de la imagen es conocida, la proyección de esas dos líneas puede ser determinada. Usando álgebra lineal se puede determinar la distancia de ese punto. Esta técnica es empleada por el sensor kinect para calcular la profundidad de los objetos y obtener una reconstrucción 3D, la triangulación es realizada para cada punto entre una imagen virtual y el patrón de observación (Kinect docs, 2012).

El proceso de triangulación para obtener la profundidad de los objetos es mostrado a continuación:

1.- Patrones de luz son proyectados en la escena, como se muestra en la figura 3.2.

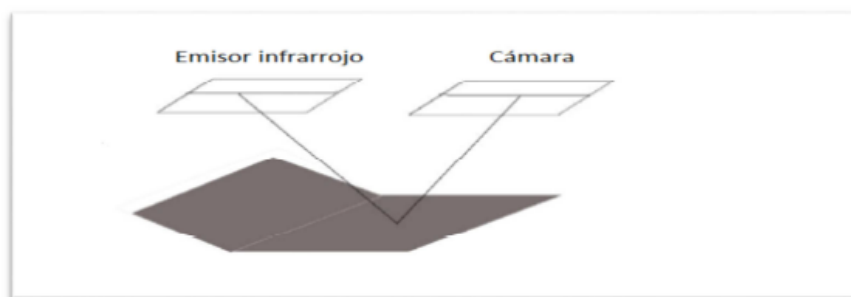


Figura 3.2 Proyección de puntos sobre una superficie (Kinectdoc, 2012).

2.- Los patrones son distinguibles uno del otro, como se muestra en la figura 3.3.

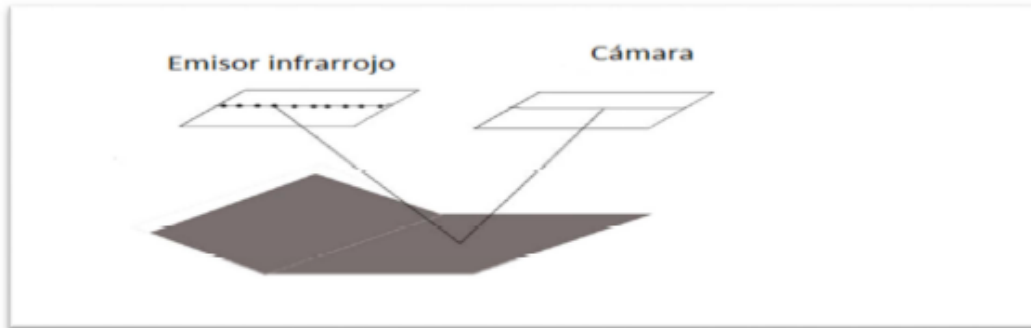


Figura 3.3 Cada uno puede ser identificado de otro dentro del patrón proyectado (Kinectdoc, 2012).

3.- Una imagen de profundidad es capturada para usarse como referencia, como se muestra en la figura 3.4.

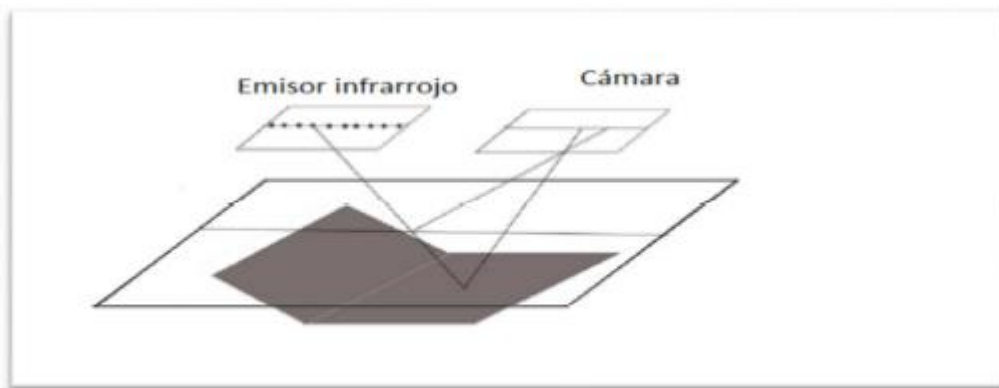


Figura 3.4 Imagen de profundidad capturada con la cámara (Kinectdoc, 2012).

4.- Δx es proporcional para la profundidad del objeto , como se muestra en la figura 3.5.

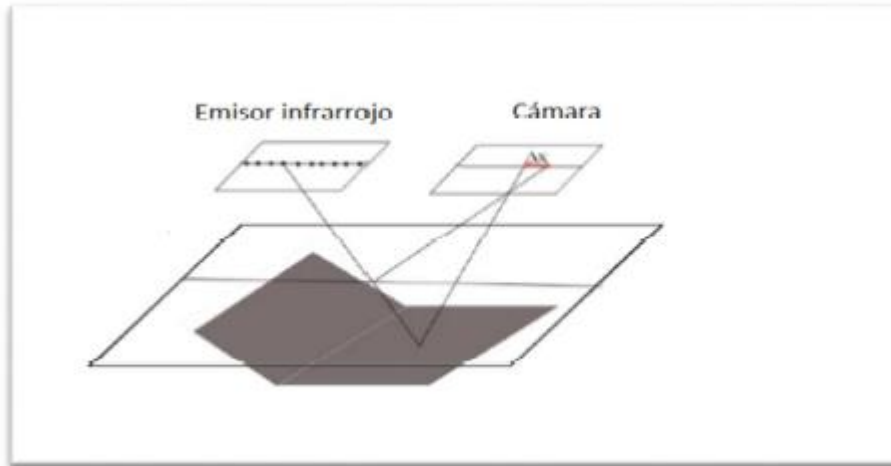


Figura 3.5 Δx es proporcional para el punto proyectado (Kinectdoc, 2012).

Imágenes de profundidad tomadas con el sensor Kinect.

El sensor Kinect imágenes de profundidad a una velocidad de treinta tramas por segundo, con una resolución de 640*480 píxeles. Los datos de profundidad son representados en un espacio de color rgb o yuv (dependiendo la configuración), en 16 bits de datos. Para el caso del espacio rgb se utilizan 16 bits para establecer los datos de profundidad, los cuales corresponden al canal R (rojo) y al canal G (verde), el canal B (azul) se utiliza para establecer datos e usuario, por ejemplo cuando un píxel corresponde a los datos de usuario, este es establecido en el octeto del canal azul, por último el canal A es establecido en 255.

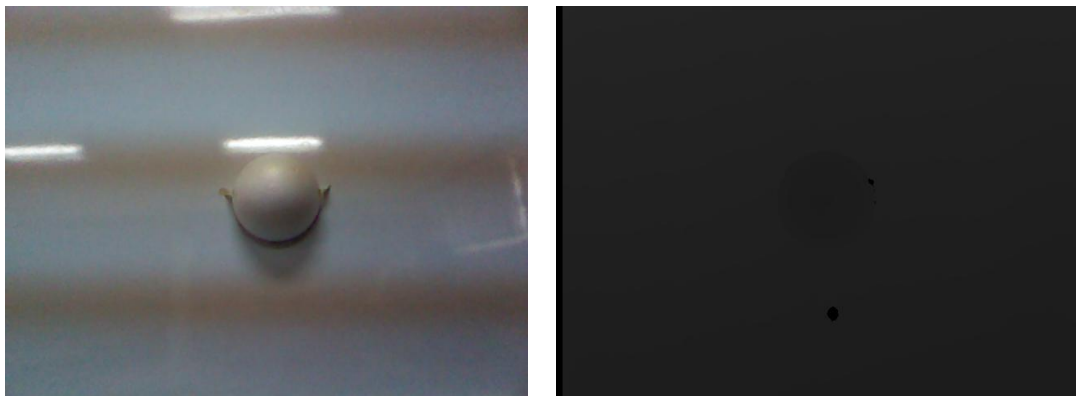
Cuando se utiliza el espacio con YUV (espacio en frecuencia) se utilizan 4 bytes por cada dos píxeles, por lo tanto se pueden almacenar 16 bits en datos de profundidad, pero valores adicionales como el número de usuario detectado, no pueden ser incluidos.

Las tablas 3.1 y 3.2 muestran cómo se distribuyen los datos de profundidad utilizando los espacios de color RGBA y YUV.

Tabla 3.1 Dato de profundidad en espacio de color RGBA			
R	G	B	A
3 bits de 8	8 bits	0 ó ID de usuario	255

Tabla 3.2 Dato de profundidad en espacio de color YUV
11 bits/16 bits valores

A continuación en la figura 3.6 se muestra una imagen de profundidad obtenida con el Kinect:



a)

b)

Figura 3.6 a) Imagen RGB y b) Imagen profundidad

En la figura 3.6 se puede observar la imagen en RGB y la imagen de profundidad tomada con el Kinect donde cada nivel de gris equivale a la profundidad de la imagen (es decir lo más cerca o más lejos que se encuentra el objeto del sensor).

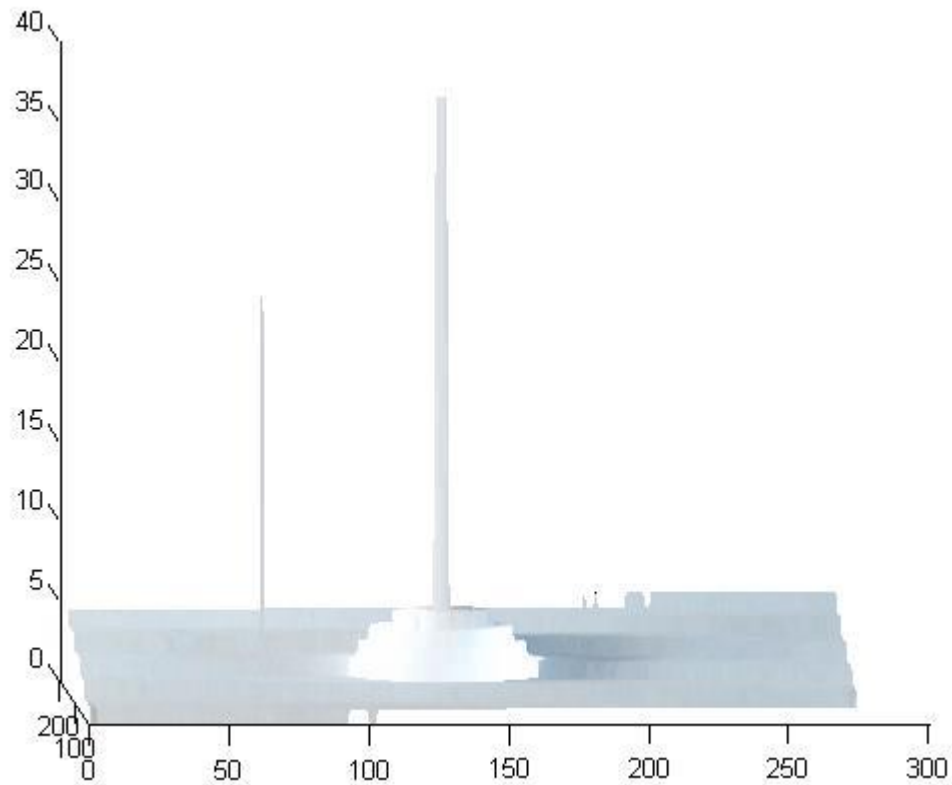


Figura 3.7 Imagen RGB encima de la de profundidad.

En la imagen 3.7 se observa una imagen de profundidad (3D) y encima la imagen RGB que le corresponde, tomadas con el sensor Kinect, se alcanza a distinguir que los saltos de niveles de grises son muy grandes y como se ve en la imagen pareciera que fuera como en escalera.

Ya que para esta investigación se requería hacer cálculos de distancias milimétricas se descartó la propuesta de utilizar el sensor Kinect para obtener imágenes 3D.

3.2 Técnica de Proyección de franjas.

Como se mencionó en el punto anterior se descartó la obtención de imágenes de profundidad (3D) con el sensor Kinect porque se necesitaba calcular distancias milimétricas, ya que las distancias se mueven de apoco.

Es por esto que se propone la reconstrucción 3D por medio de la técnica proyección de franjas.

La técnica de proyección de franjas permite, en aproximadamente dos segundos, adquirir las imágenes necesarias para obtener una reconstrucción 3D con resolución micrométrica y precisión sub píxel. Esta es una de las principales razones por las que al consultar la literatura especializada, los antecedentes evidencian una amplia preferencia por esta técnica.

3.2.1 Metodología.

La técnica de proyección de franjas consiste en proyectar sobre el objeto que se va a reconstruir franjas paralelas blancas y negras, que se deforman por la topografía superficial del objeto (Salvi, 2001). La figura 3.8 muestra un montaje típico empleado para ejecutar la técnica. El sistema consta de una unidad de proyección/adquisición y una unidad de procesamiento. La unidad de proyección/adquisición está conformada por un dispositivo de proyección (tradicionalmente un video proyector) para proyectar las franjas y un dispositivo de observación para la adquisición de las imágenes (cámara CCD). Las imágenes adquiridas son procesadas por la unidad de control, que permite adicionalmente sincronizar las adquisiciones y calibrar el dispositivo.

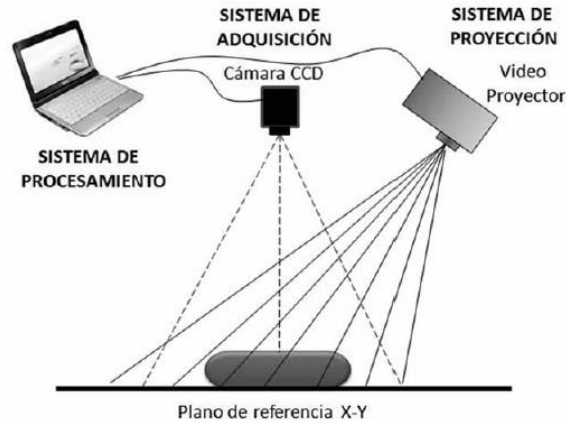


Figura 3.8 Montaje experimental de proyección de franjas. (González, 2012)

Es ideal que el eje óptico de la cámara se ubique de forma perpendicular al plano de proyección donde se focalizan las franjas y la cámara; este plano se llama plano de referencia.

El proceso físico de codificación de la información topográfica se realiza de la siguiente manera: Cuando la proyección se realiza sobre una superficie plana se observan franjas paralelas y de paso constante; al proyectarlas sobre un objeto, las irregularidades de la superficie (relacionadas directamente con la topografía superficial del objeto) quedan evidenciadas con la deformación de las franjas (Pérez, 2006), como se muestra en la figura 3.9.



Figura 3.9 Patrón de franjas proyectado en a) un plano de referencia y b) Una mano Humana (Pérez, 2006)

Considerando un sistema formador de imágenes no-telecéntrico, el ángulo entre los ejes de proyección y observación y la influencia de las aberraciones geométricas, la distribución en intensidad de las imágenes obtenidas con la cámara CCD sobre el plano de referencia tiene la forma dada por la siguiente ecuación:

$$I(x_0, y_0) = I_0(x_0, y_0) + A(x_0, y_0) * \text{Cos} (2\pi f_0 x_0 + \phi_0) \quad (3.1)$$

Donde I_0 y A corresponden a la intensidad del fondo continuo y el contraste de las franjas, respectivamente. El término f_0 es la frecuencia espacial media de las franjas y es la fase inicial de las franjas, que corresponde a la deformación inicial sufrida por las franjas. Al ubicar sobre el plano de referencia un objeto, la ecuación se modifica de acuerdo:

$$I(x_0, y_0) = I_0(x_0, y_0) + A(x_0, y_0) * \text{Cos} (2\pi f_0 x_0 + \phi_0 + \Delta\phi) \quad (3.2)$$

Donde el término $\Delta\phi$ es la fase introducida por la topografía del objeto.

3.2.2 Obtención de imágenes Tridimensionales (3D).

Teniendo en cuenta que ϕ corresponde a la fase del patrón de las franjas, de este modo, la deformación que genera en un objeto alter directamente a la fase y para la reconstrucción 3D es necesario recuperar la fase. Para recuperar la fase es necesario aplicar el método corrimiento de fase (Phase Shifting) (Larkin et al, 1995). Para este método se determina la fase del objeto a partir de la intensidad luminosa de n imágenes digitalizadas, con su respectiva fase modificada en valores ϕ_i conocidos dentro de un intervalo de 2π , como se muestra en:

$$\phi_i = \frac{2\pi}{n} (i - 1) \quad (3.3)$$

$$i = 1, 2, 3 \dots n$$

El procedimiento es comúnmente utilizado con $n = 4$ imágenes (Larkin et al, 1995), lo cual implica un desplazamiento de fase igual a $\pi/2$. Así, adquiriendo las cuatro imágenes se puede calcular el valor de A y ϕ para cada punto (x_0, y_0) , usando:

$$\phi = \tan^{-1} \frac{I_4 - I_2}{I_3 - I_1} \quad (3.4)$$

$$2A(x, y) = \sqrt{(I_4 - I_2)^2 + (I_3 - I_1)^2} \quad (3.5)$$

Donde I_1 es la imagen correspondiente a un desplazamiento de fase $\phi_1 = 0$, I_2 para un $\phi_2 = \pi/2$, I_3 para un $\phi_3 = \pi$, e I_4 para un $\phi_4 = 3\pi/2$. La figura 3.10 muestra un ejemplo de estas imágenes.

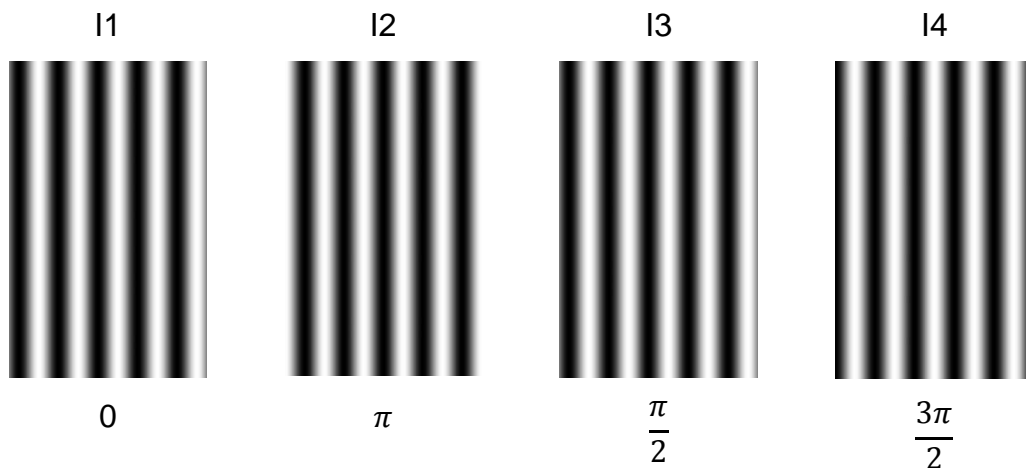


Figura 3.10 Patrón de franjas con corrimiento de fase $\phi = \pi/2$

Como se observa en la ecuación 3.4, el uso de la función arco-tangente para el cálculo de la fase genera puntos de discontinuidad entre $-\pi$ y π , razón por la cual es necesario convertir dicha fase discontinua en una fase continua. En la figura 3.11 (a) el valor de $-\pi$ corresponde al nivel de gris negro y $+\pi$ al nivel de gris blanco. Matemáticamente se demuestra que la transición entre $\pm\pi$ ocurre en el centro de una franja negra. Con el fin de obtener la fase continua es necesario identificar las discontinuidades y adicionar valores de $2\pi N$, siendo N una función escalón entera apropiada para eliminar las discontinuidades. Este procedimiento se conoce como “Unwrapping” y se usa para obtener la fase continua del patrón de franjas proyectado sobre el plano de referencia y la fase continua del mismo patrón de franjas proyectado sobre el objeto, como se muestra en la figura 3.11 b (Meneses, 2005).

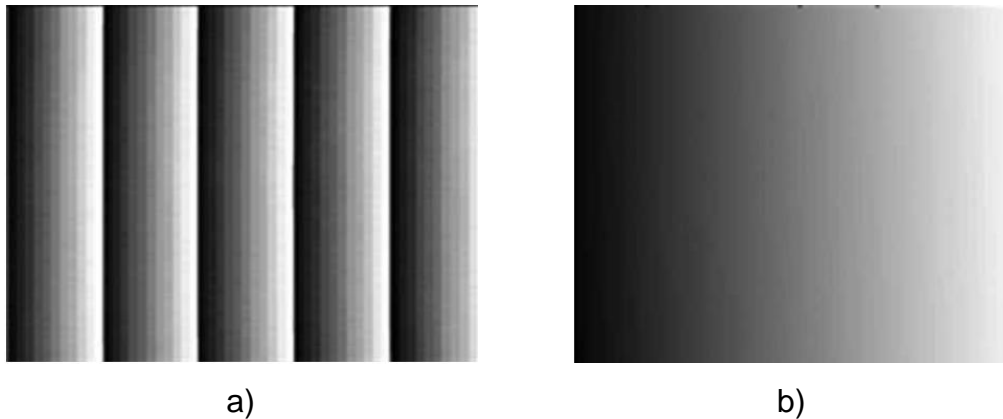


Figura 3.11 Imagen de la fase (a) discontinua y (b) continua. (González, 2012)

El proceso de reconstrucción 3D se completa al convertir la distribución de fase $\Delta\phi$ en valores de coordenadas Z para cada posición (X,Y) digitalizada por la cámara.

3.2.3 Imágenes Tridimensionales (3D) adquiridas con la técnica de proyección de franjas.

A continuación se muestra algunas imágenes tridimensionales obtenidas por medio de la técnica explicada en esta sección. :

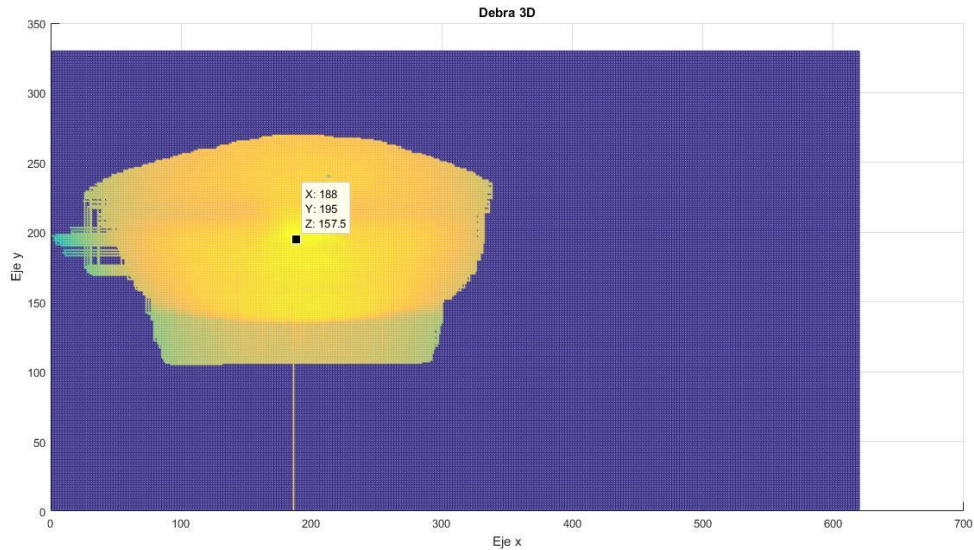


Figura 3.12 Imagen de un rostro 2D. (maniquí)

En la figura 3.12 se observa una imagen en 2D, es decir, una imagen plana de un rostro (maniquí), pero con información tridimensional ya que el punto señala las 3 coordenadas: x, y, z.

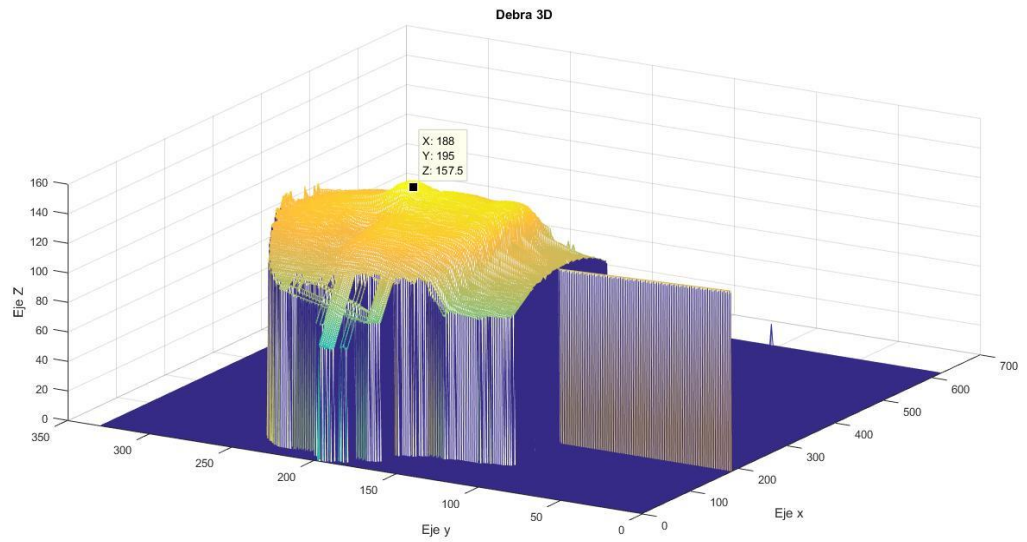


Figura 3.13 Imagen de un rostro 3D. (maniquí)

En la figura 3.13 se observa una imagen tridimensional, de un rostro (maniquí).

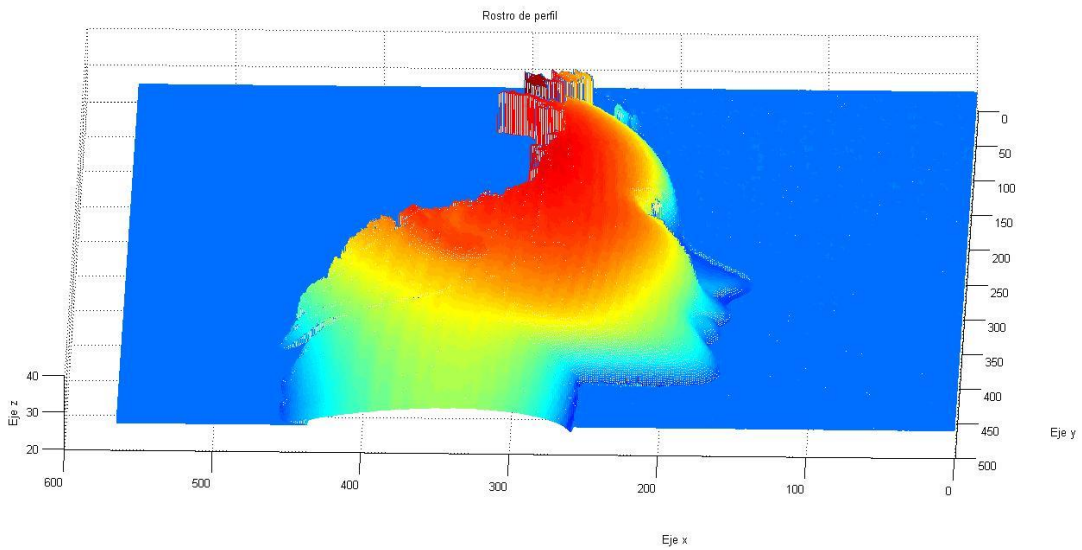


Figura 3.14 Perfil de un rostro 3D (maniquí).

En la figura 3.14 se observa el perfil de una imagen tridimensional de un rostro (maniquí).

Cabe mencionar que aunque los resultados que se obtuvieron con esta técnica son realmente buenos ya que se alcanza a distinguir que reconstruía muy bien el rostro (maniquí), se tuvo que dejar de lado esta propuesta, ya que se implementó el uso de una técnica en conjunto, que se llama Lógica Difusa, y para utilizarla solo bastaba con medir distancias y aumentaran o disminuyeran y eso se podía obtener con una imagen plana (2D).

3.3. Procesamiento Digital de imágenes y Lógica Difusa.

Como ya se había mencionado el procesamiento digital de imágenes consiste en procesar una imagen digital por medio de un ordenador.

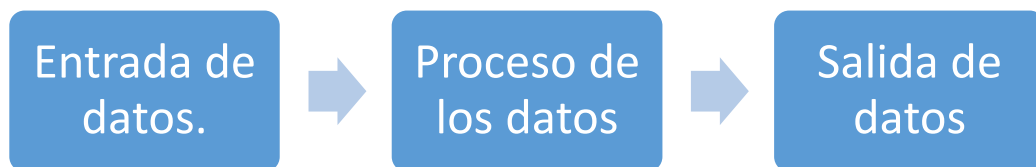
3.3.1. Metodología a seguir para el Procesamiento Digital de Imágenes.

- 1.- Se implementó un aparato para que la toma de las imágenes tenga la misma distancia.
- 2.- Se implementó un algoritmo de pre-procesamiento digital de imágenes para dejar únicamente la información de interés (Segmentar el rostro).
- 3.- En la etapa de reconstrucción virtual del rostro completo, en base a la parte sin daño (para darle una idea al terapeuta del posible resultado).
- 4.- Se implementó un algoritmo que detecte los puntos necesarios para poder realizar correctamente mediciones de las imágenes del mismo paciente.
- 5.- Se determinó de forma cuantitativa la diferencia entre imágenes y a partir de esta información medir la evolución.

3.3.2. Lógica Difusa.

Como ya se había mencionado la Lógica Difusa es una lógica multivaluada que permite representar matemáticamente la incertidumbre y la vaguedad, proporcionando herramientas formales para su tratamiento.

A grandes rasgos puede ser representada como lo muestra el siguiente diagrama de bloques:



La entrada de datos sería, variables lingüísticas explicadas anteriormente, que para nuestro caso se definieron en conjunto con el terapeuta, se tomaron 3 distancias muy representativas en el lado del rostro afectado a la hora que se presenta una parálisis facial:

Comisura a la oreja, esta distancia tiene que disminuir para que se pueda observar una mejoría ya que la comisura va regresando a su lugar de origen.

Comisura al mentón, esta distancia tiene que aumentar para que se pueda observar una mejoría ya que la comisura va regresando a su lugar de origen.

Parpado a la ceja, esta distancia tiene que aumentar para que se pueda observar una mejoría ya que el parpado tiene que cerrar completamente.

Cada una de las distancias pertenecerán a un valor de acuerdo a como se vayan moviendo durante el tratamiento los cuales son mejorando normal y empeorando.

A las tres distancias las nombraremos:

Dist₁ = Comisura a la oreja.

Dist₂ = Comisura al mentón.

Dist₃ = Parpado a la ceja.

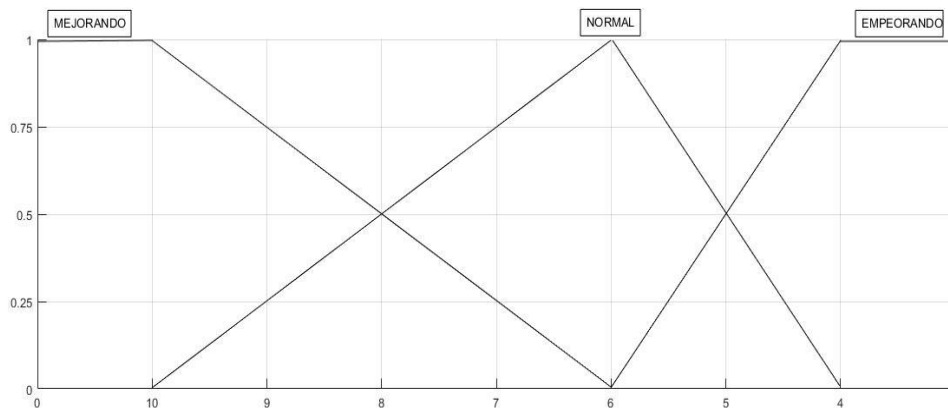


Figura 3.15 Valores lingüísticos que toma la dist₁.

En la figura 3.15 se puede observar los tres valores que puede tomar la distancia de la comisura a la oreja que van de 0 a 1, el especialista explico que un valor normal de esta distancia es de 6 cm y se puede desfigurar hasta 4 cm, es decir que la distancia de un rostro afectado puede llegar a medir 10 cm aproximadamente, en el eje X de la figura el valor está representado en cm y en el eje Y el valor que toma dependiendo del movimiento de las distancias para así decidir en cuál de los tres valores se encuentra, cabe mencionar que disminuyendo de 10 cm va a tomar el valor de mejorando o normal y si disminuye más de 6 cm entonces va a tomar el valor de empeorando, el valor del eje Y lo tomara evaluara el sistema automáticamente.

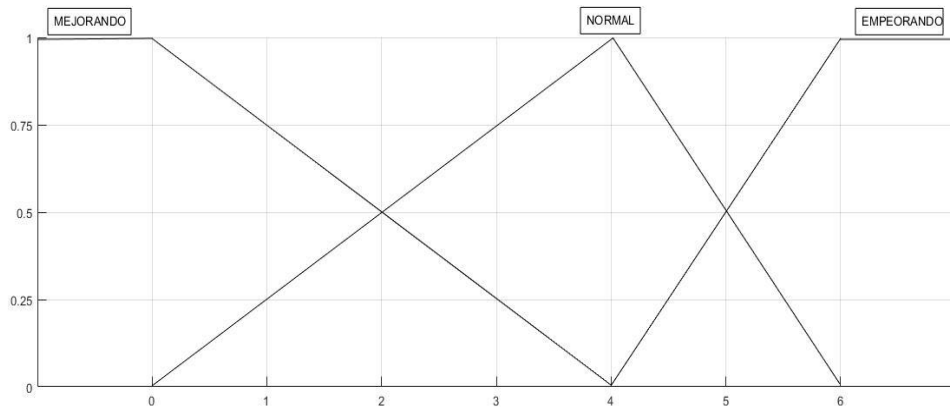


Figura 3.16 Valores lingüísticos que toma la $dist_2$.

En la figura 3.16 se puede observar los tres valores que puede tomar la distancia de la comisura al mentón que van de 0 a 1, el especialista explica que un valor normal de esta distancia es de 4 cm y se puede desfigurar hasta 3.5 cm, es decir que la distancia de un rostro afectado puede llegar a medir 0.5 cm aproximadamente, en el eje X de la figura el valor está representado en cm y en el eje Y el valor que toma dependiendo del movimiento de las distancias para así decidir en cuál de los tres valores se encuentra, cabe mencionar que aumentando de 0.5 cm va tomando el valor de mejorando o normal si aumenta de 4 cm entonces tomara el valor de empeorando, el valor del eje Y lo tomara y evaluara el sistema automáticamente.

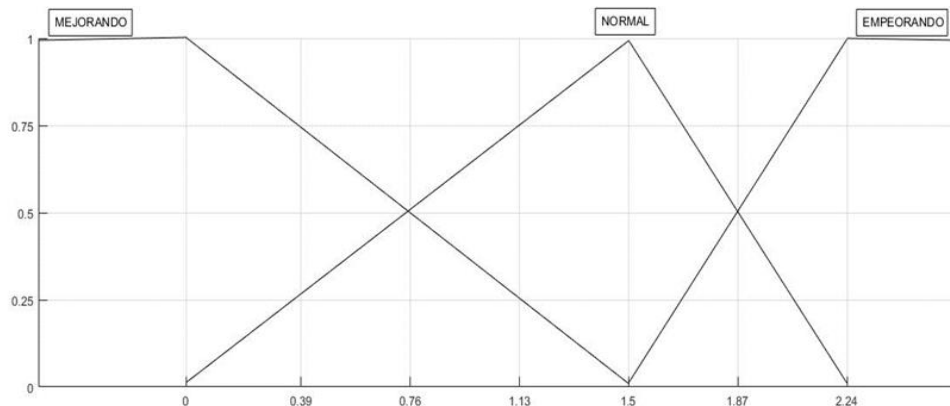


Figura 3.17 Valores lingüísticos que toma la $dist_3$.

En la figura 3.17 se puede observar los tres valores que puede tomar la distancia del parpado a la ceja que van de 0 a 1, el especialista explico que un valor normal de esta distancia es de 1.5 cm y se puede desfigurar hasta 1.25 cm, es decir que la distancia de un rostro afectado puede llegar a medir 0.25 cm, en el eje X de la figura es el valor en cm y el en el eje Y el valor que toma dependiendo del movimiento de las distancias y así decidir en cuál de los tres valores se encuentra, cabe mencionar que aumentando de 0.25 cm va tomando el valor de mejorando o normal si aumenta de 1.5 entonces tomara el valor de empeorando, el valor del eje Y lo tomara evaluara el sistema automáticamente.

IV RESULTADOS.

En este capítulo se presentan los resultados obtenidos durante la investigación los cuales fueron: una metodología para poder realizar el procesamiento digital de imágenes, una serie de reglas y en conjunto, un sistema computacional para poder medir la evolución de pacientes con parálisis facial.

I. Un algoritmo de procesamiento digital de imágenes para la segmentación de un rostro.



Figura 4.1 Rostro segmentado a) Imagen virtual b) Imagen real de un maniquí.

En la figura 4.1 a) se puede observar una imagen virtual segmentada de un rostro y en la figura 4.1 b) una imagen real de un maniquí, donde se discrimina todo el fondo azul de la imagen, que como parte de la metodología se estableció que este color es más fácil discriminarlo ya que en un rostro predomina color rojo y amarillo, así con esto dejando solo la parte de interés que es el rostro.

II. Un algoritmo de procesamiento digital de imágenes que reconstruya virtualmente el rostro completo, tomando como base el lado del rostro sin daño.



a)



b)

Figura 4.2 Rostro reconstruido virtualmente a) Imagen virtual b) Imagen real de un maniquí.

Como parte de la reconstrucción del rostro completo en base a la parte sin daño, el algoritmo toma la parte dañada, le toma un espejo, y junta las dos partes dando como resultado un rostro reconstruido virtualmente como se muestra en la figura 4.2 a) y 4.2 b), cabe mencionar que este resultado es muy importante para poder mostrarle al paciente un posible resultado de como pudiera quedar después de las terapias.

III. Un algoritmo de procesamiento digital de imágenes que detecte puntos característicos necesarios en la imagen del rostro de un paciente, para poder realizar correctamente las mediciones.

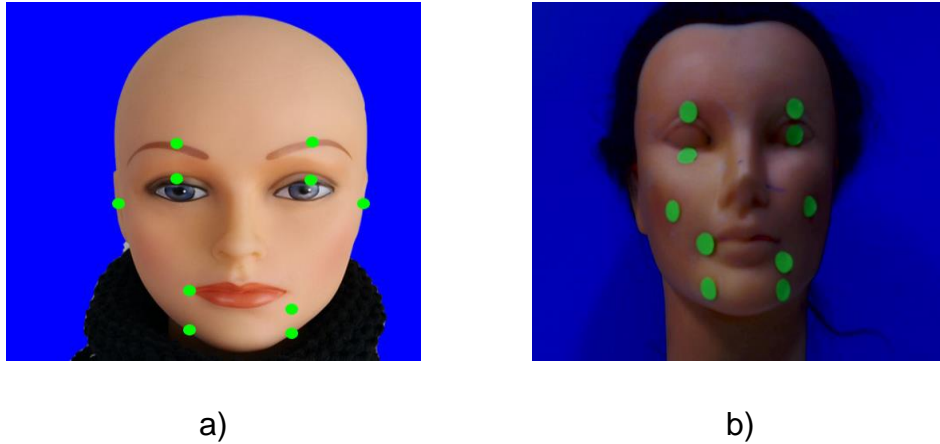


Figura 4.3 Puntos característicos a) Imagen virtual b) Imagen real de un maniquí.

En la figura 4.3 a) y 4.3 b) se observan 10 puntos en la cara, que se establecieron con la ayuda de un experto (terapeuta) el cual indico que esos puntos son los más representativos cuando se presenta una parálisis facial, que el algoritmo detecta en la imagen automáticamente, como se muestra en la figura 4.4 a) y 4.4 b).

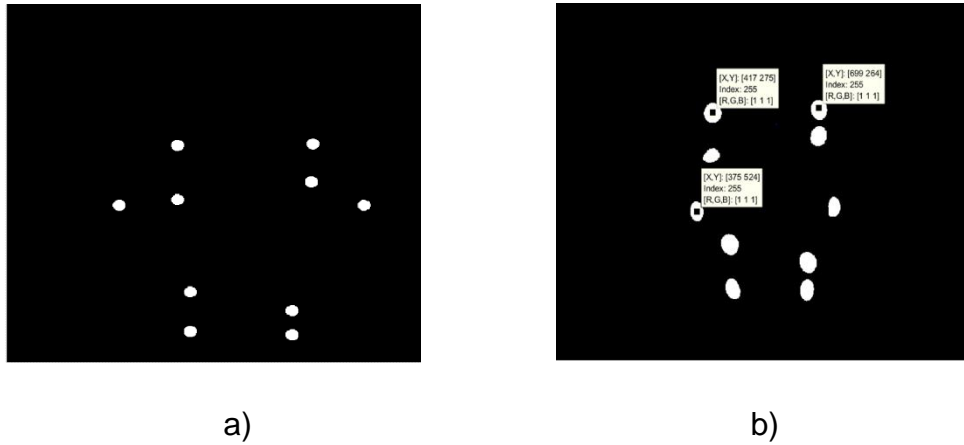


Figura 4.4 Puntos característicos detectados a) Imagen virtual b) Imagen real de un maniquí.

En la imagen 4.4 a) y 4.4 b) se puede observar solamente los puntos detectados automáticamente en la imagen, y así poder hacer las mediciones pertinentes entre cada distancia que se requiere para medir la evolución.

Cabe mencionar que se presentaron los resultados de una imagen virtual de un rostro y una imagen de un maniquí, esto para comprobar que el sistema es funcional de manera real.

IV. La metodología que se mencionó en el desarrollo experimental para poder medir la evolución de pacientes con parálisis facial de manera cualitativa.

V. Como resultante de la metodología se obtuvieron las siguientes reglas:

Con ayuda del terapeuta, se definieron las reglas con las cuales se determinará uno de los posibles resultados que puede tener un paciente, en este caso en particular son 27, tal como se muestra en la tabla 2, Solo son 27 ya que solo se tomaron 3 variables lingüísticas y 3 valores lingüísticos para no convertir tan robusto a los posibles resultados ya que si aumentamos el valor de variables lingüísticas o valores lingüísticos incrementan los resultados y además por otra parte, el experto nos mencionó que esas distancias son las más visibles y representativas.

Donde la primer la columna indica a qué valor lingüístico pertenece la distancia 1, la segunda columna indica un operador AND (Y), la tercer columna a qué valor lingüístico pertenece la distancia 2, la cuarta columna un operador AND (Y), la quinta columna a qué valor lingüístico pertenece la distancia 3, la sexta columna un operador THEN (Entonces) y la séptima columna el resultado de los valores del grado de avance después de un periodo de terapia y estos se definen después de la tabla, donde además, el color de la celda está relacionado al nivel de mejoría del paciente: donde amarillo indica una ligera mejoría, azul indica una mejoría modera o aceptable y rojo indica que no hay mejoría.

Expresado de otra manera por ejemplo haciendo referencia a la primera fila de la tabla 4.1:

Distancia 1 pertenece al valor lingüístico Mejorando y distancia 2 pertenece a Mejorando y distancia 3 pertenece a Mejorando, entonces el resultado es: A.

Tabla 4.1 Posibles resultados.						
dist1∈Me	AND	dist2∈Me	AND	dist3∈Me	THEN	A
dist1∈Me	AND	dist2∈Me	AND	dist3∈No	THEN	A
dist1∈Me	AND	dist2∈No	AND	dist3∈Me	THEN	A
dist1∈Me	AND	dist2∈No	AND	dist3∈No	THEN	A
dist1∈Me	AND	dist2∈Me	AND	dist3∈Em	THEN	B
dist1∈Me	AND	dist2∈No	AND	dist3∈Em	THEN	B
dist1∈Me	AND	dist2∈Em	AND	dist3∈Me	THEN	C
dist1∈Me	AND	dist2∈Em	AND	dist3∈No	THEN	C
dist1∈No	AND	dist2∈Em	AND	dist3∈Me	THEN	C
dist1∈No	AND	dist2∈Em	AND	dist3∈No	THEN	C
dist1∈Em	AND	dist2∈Me	AND	dist3∈Me	THEN	C
dist1∈Em	AND	dist2∈Me	AND	dist3∈No	THEN	C
dist1∈Em	AND	dist2∈No	AND	dist3∈Me	THEN	C
dist1∈Em	AND	dist2∈No	AND	dist3∈No	THEN	C
dist1∈No	AND	dist2∈No	AND	dist3∈Em	THEN	D
dist1∈No	AND	dist2∈Me	AND	dist3∈Em	THEN	D
dist1∈No	AND	dist2∈Me	AND	dist3∈Me	THEN	E
dist1∈No	AND	dist2∈Me	AND	dist3∈No	THEN	E
dist1∈No	AND	dist2∈No	AND	dist3∈Me	THEN	E
dist1∈No	AND	dist2∈No	AND	dist3∈No	THEN	F
dist1∈Me	AND	dist2∈Em	AND	dist3∈Em	THEN	G
dist1∈No	AND	dist2∈Em	AND	dist3∈Em	THEN	G
dist1∈Em	AND	dist2∈Me	AND	dist3∈Em	THEN	G
dist1∈Em	AND	dist2∈No	AND	dist3∈Em	THEN	G
dist1∈Em	AND	dist2∈Em	AND	dist3∈Me	THEN	H
dist1∈Em	AND	dist2∈Em	AND	dist3∈No	THEN	H
dist1∈Em	AND	dist2∈Em	AND	dist3∈Em	THEN	I

A = Ligera.

B = Ligera mejoría a nivel rama inferior facial.

C = Ligera (atención en la boca).

D = Ligera mejoría (atención en rama superior).

E = Moderada.

F = Aceptable.

G = No mejora (revisar tratamiento).

H = No mejora (revisar tratamiento, rama inferior).

I = No mejora (revisar tratamiento, cita médica con especialista).

VI. Un sistema completo implementado tal como lo descrito antes en el interior del sistema para poder realizar la medición de pacientes con parálisis facial, además se integraron al sistema herramientas de apoyo al terapeuta, tal como un formulario para el registro y gestión de datos para cada paciente como también cada terapeuta.

Alta y baja de Terapeutas Actualizar datos de Terapeutas Registro de Pacientes Medir Evolución

Usuario:

Contraseña:

Tipo:

Entrar

Figura 4.5 Pantalla principal del Sistema.

En la figura 4.5 se observa la pantalla inicial del sistema de medición de pacientes con parálisis facial, donde pide que introduzca sus datos personales del terapeuta como su usuario contraseña y el rol de usuario, cabe mencionar que dependiendo el rol que tenga cada terapeuta son las opciones que se habilitaran.

registro_pacientes

Alta de terapeutas Medir Evolución

Registro de Pacientes

Nombre: Foto Fecha

Apellido:

Terapia

Distancia 1 : Lado sin daño

Distancia 2 :

Distancia 3 : Rostro Reconstruido

Distancia 4 :

Distancia 5 :

Distancia 6 :

Figura 4.6 Registro de pacientes.

En la figura 4.6 se observa la pantalla donde se registra los datos generales de cada paciente y los necesarios para la medición, por ejemplo nombre apellidos el valor de las distancias del lado normal y las del lado desfigurado, también muestra el rostro reconstruido virtualmente para poder mostrar al paciente un posible resultado al termino de las terapias.

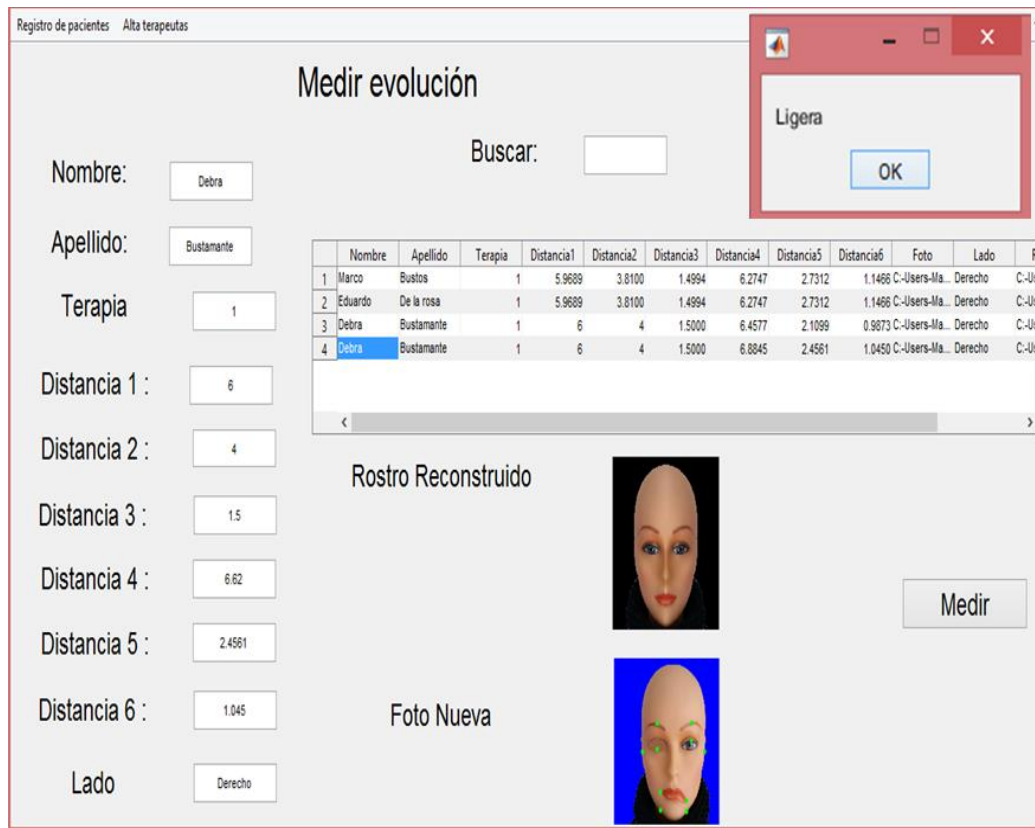


Figura 4.7 Medición de la Evolución (terapia 1).

En la figura 4.7 se observa la pantalla donde mide la evolución de pacientes, donde se está simulando las mediciones en la imagen virtual del rostro la terapia número 1, muestra el rostro desfigurado simulando mejoría al ir moviendo los puntos característicos, también para la comodidad del terapeuta se implementó un campo para que pueda hacer una búsqueda y encontrar rápidamente al paciente en cuestión, donde también muestra los datos generales tales como: nombre, apellido, numero de terapia, las distancias del lado sin parálisis, y al presionar el botón de medir calcula las nuevas distancias del lado con parálisis después de las terapias sugeridas e indica a cuál de los 27 resultados presenta la evolución, que para este caso se obtuvo el resultado que corresponde a **“Ligera”**.

Registro de pacientes Alta terapeutas

Medir evolución

Buscar:

Nombre:

Apellido:

Terapia:

Distancia 1 :

Distancia 2 :

Distancia 3 :

Distancia 4 :

Distancia 5 :

Distancia 6 :

Lado:

	Nombre	Apellido	Terapia	Distancia1	Distancia2	Distancia3	Distancia4	Distancia5	Distancia6	Foto	Lado	R
1	Marco	Bustos	1	5.9689	3.8100	1.4994	6.2747	2.7312	1.1466	C-Users-Ma...	Derecho	C-Ue
2	Eduardo	De la rosa	1	5.9689	3.8100	1.4994	6.2747	2.7312	1.1466	C-Users-Ma...	Derecho	C-Ue
3	Debra	Bustamante	1	6	4	1.5000	6.4577	2.1099	0.9873	C-Users-Ma...	Derecho	C-Ue
4	Debra	Bustamante	1	6	4	1.5000	6.8845	2.4561	1.0450	C-Users-Ma...	Derecho	C-Ue

Moderada

Rostro Reconstruido




Foto Nueva




Figura 4.8 Medición de la Evolución (terapia 2).

En la figura 4.8 se observa la pantalla idéntica a la 4.7 solo que en esta figura simula la mejoría al ir moviendo los puntos característicos, se obtuvo el resultado que corresponde a **“Moderada”**.

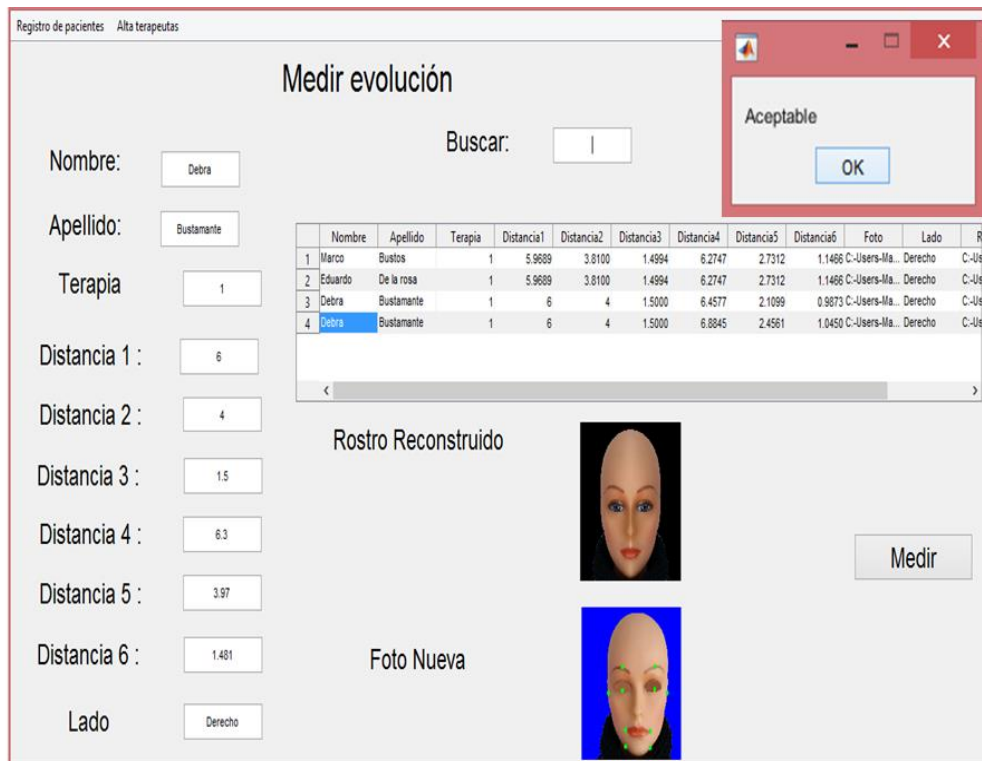


Figura 4.9 Medición de la Evolución (terapia 3).

En la figura 4.9 se observa la pantalla idéntica a la 4.7 y 4.8 solo que en esta figura simula la mejoría al ir moviendo los puntos característicos, se obtuvo el resultado que corresponde a **“Aceptable”**.

4.1 Discusión.

Para llegar a la solución planteada que es medir la evolución de pacientes con parálisis facial, se tuvo que analizar cómo se iba a tratar la información es por eso que se segmenta la imagen, es decir se deja solo el rostro para poder tratar más fácilmente la imagen, se reconstruye virtualmente el rostro para mostrar un posible resultado al paciente, se detectaron los puntos necesarios para la medición, y se calcula la distancia entre los puntos característicos, todo planteado por el terapeuta consultado.

En base a los antecedentes que se consultaron referentes al tema, hay algunos que se asemejan ya que proponen algoritmos que miden el grado de parálisis facial, tal es el caso del algoritmo propuesto por (Gálvez 2017) donde después de tomar en cuenta puntos clave le asigna graduación y con esto mide el grado de parálisis facial siguiendo el sistema simplificado de graduación de parálisis del nervio facial de House-Brackmann donde aplica una metodología cualitativamente donde dice a qué grado de parálisis facial pertenece.

Estos autores (Carro et al. 2016) desarrollaron un algoritmo que evalúa el grado de parálisis facial con el uso del Kinect para la extracción y la simetría de un rostro, determinan ciertas distancias en el rostro y aplica el sistema de clasificación de Nottingham donde cuantifica el grado de parálisis y da una escala de tres niveles: normal, leve y severo.

Entre otros utilizan metodología similar detectan características y calculan las distancias entre ellas.

Lo que diferencia este trabajo de los demás es que permite al terapeuta poner atención a los lugares donde no hay mejoría, esto gracias a los 27 posibles resultados que nos permite obtener con la técnica llamada "Lógica Difusa".

V CONCLUSIONES FINALES.

El principal problema en la medición de pacientes con parálisis facial, es la falta de objetividad a la hora de determinar si la terapia está funcionando. En la gran mayoría de los casos los pacientes, se puede lograr una recuperación casi total entre 3 y 5 sesiones de terapia, las cuales suelen ser cada 15 días, pero hay casos en los pacientes que a pesar de que hay mejoría, esta es imperceptible y los pacientes se depriman, lo que puede ocasionar que el tiempo de recuperación sea mayor, esto de acuerdo con el especialista. Con la ayuda del especialista, se analizaron tres distancias significativas que se mueven al observar parálisis facial, donde los pacientes pueden presentar desfiguración de la comisura y no cerrar un ojo completamente.

Analizando los resultados de la presente investigación, se llegó a la conclusión que sí es posible medir la evolución de pacientes con parálisis facial y también poner atención en las partes del rostro donde no hay mejoría, ya que uno de los propósitos es medir cuantitativa y cualitativamente, la aportación más relevante de este trabajo es el sistema desarrollado que le permitirá al terapeuta llevar un expediente de sus pacientes, mostrar un posible resultado terminando las terapias, medir el desplazamiento de las distancias en el rostro y dar el resultado que arroja el sistema. El terapeuta puede tomar este sistema para apoyarse en los resultados, y así poder informar al paciente si está evolucionando o no, también le permitirá detectar dónde poner atención en dado caso que el resultado que se obtenga no sea favorable, y así cambiar la terapia o canalizarlo con otro especialista y ver porque no hay mejoría.

5.1 Recomendaciones.

Se recomienda optimizar el software de detección de puntos característicos en el rostro para poder detectarlos automáticamente ya que en este trabajo se apoyó de unas calcomanías en forma de círculo pegados en el rostro.

Como recomendación fundamental es probarlo con pacientes de parálisis facial humanos, ya que en este trabajo se realizaron simulaciones con el rostro de un maniquí, moviendo los puntos para poder simular la medición de la evolución de los pacientes.

REFERENCIAS.

Adour K.K. Inflammatory disorders of the facial nerve: Bell's palsy, Ramsay Hunt Syndrome, Otitis media and Lyme disease. Canalis RF, Lambert PR. The Ear Comprehensive Otolaryngology. Filadelfia, (2000) 705-718.

Benítez S., Danilla S., Troncoso E., Moya A., Mahn J., MANEJO INTEGRAL DE LA PARÁLISIS FACIAL, Elsevier Revista Médica Clínica Las Condes. (2016) 22-28.

Carro R. C., Huerta E. B., Caporal R. M., Hernández J. C., Cruz F. R., Facial Expression Analysis with Kinect for the Diagnosis of Paralysis Using Nottingham System, IEEE LATIN AMERICA TRANSACTIONS. 14 (2016) 3418- 3426.

DEPT, H. E. Kinect docs. es.scribd.com. (2012) es.scribd.com, es.scribd.com: <https://es.scribd.com/document/86437713/kinectdoc>.

Devriese P.P. y Moesker W.H. The natural history of facial paralysis in herpes Zoster. Clin Otolaryngol 13 (1988) 289-298.

Diagnostican al año 20 mil casos de parálisis facial en México. (2015). www.eluniversal.com.mx. México, www.eluniversal.com.mx: <https://www.eluniversal.com.mx/articulo/nacion/sociedad/2015/10/27/diagnostican-al-ano-20-mil-casos-de-paralisis-facial-en-mexico>.

Duarte, Oscar., Aplicaciones de la Lógica Difusa, Revista Ingeniería e Investigación. 45 (2000) 7-12.

Escalante, R. (2006). <http://verona.fi-p.unam.mx>. México, <http://verona.fi-p.unam.mx>: <http://verona.fi-p.unam.mx/boris/teachingnotes/Capitulo4.pdf>.

ESPECIFICACIONES TÉCNICAS DE KINECT. (2010) [www.fayerwayer, www.fayerwayer:](http://www.fayerwayer.com/2010/06/especificaciones-tecnicas-de-kinect) <http://www.fayerwayer.com/2010/06/especificaciones-tecnicas-de-kinect>.

Florencia. (2016). <http://catarina.udlap.mx/>. Mexico, <http://catarina.udlap.mx/>: http://catarina.udlap.mx/udla/tales/documentos/msp/florencia_y_an/capitulo3.pdf.

Fisch U. Surgery for Bell's palsy. Arch Otolaryngol Head Neck Surg. 107 (1981) 1-11.

Gálvez R. E., Cuantificación del grado de parálisis facial mediante algoritmos basados en tracking facial, Tesis de Maestría, Chile, 2017.

Gambini, M. (2006). Modelos de Segmentación basados en Regiones y Contornos Activos aplicados a Imágenes de Radar de Apertura Sintética, Tesis Doctoral. Universidad de Buenos Aires, Argentina.

García, P. (2012). <http://eprints.sim.ucm.es/>. Madrid España, <http://eprints.sim.ucm.es/>: <http://eprints.sim.ucm.es/23444/1/ProyectoFinMasterPedroPablo.pdf>.

González A., Meneses J., León J. Proyección de franjas en metrología óptica facial. Revista INGE CUC, 8 (2012) 191-206.

González, E. (Mayo de 2008). Detección y clasificación de objetos dentro de un salón de clases empleando técnicas de procesamiento digital de imágenes (Tesis de Maestría). Universidad Autónoma Metropolitana, Mexico.

González M., Parálisis facial. Hospital General "Dr. Manuel Gea González". 4 (2001) 21-26.

Gómez, V., y Guerrero, A. (Mayo de 2016). utp.edu.co. Pereira Colombia, utp.edu.co:

<http://repositorio.utp.edu.co/dspace/bitstream/handle/11059/6494/00642G633.pdf?sequence=1>.

González, R.C. y Woods, R.E. Digital Image Processing. (2ª ed.) New Jersey, Estados Unidos 2002, Pearson Prentice Hall.

González, R.C. y Woods, R.E. Digital Image Processing. (3ª ed.) New Jersey, Estados Unidos 2008, Pearson Prentice Hall.

He S., Soraghan J. J., O'Reilly B. F., Xing D., Quantitative Analysis of Facial Paralysis Using Local Binary Patterns in Biomedical Videos, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING. 56 (2009) 1864-1870.

Hoffer E. Annual meeting of the facial nerve study group. ArchOtolaryngol Head Neck Surg (1995) 121: 240.

Hontanilla B., Aubá., Automatic three-dimensional quantitative analysis for evaluation of facial movement, Elsevier, Reconstructive & Aesthetic Surgery. 61 (2008) 18-30.

Kinectdoc. es.scribd.com. (2012) es.scribd.com:
<https://es.scribd.com/document/86437713/kinectdoc>.

Larkin k., Hibino K., Oreb B., Farrant I. Phase shifting for nonsinusoidal waveforms with phase-shift errors. *J. Opt Soc. Am. A*, 12 (1995).

Larrañaga, P., Inza, I., Moujahid, A. (2016). sc.ehu.es. Paris, sc.ehu.es:
<http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t6bayesianos.pdf>.

Mattox D.E. Clinical disorders of the facial nerve. CummingsCW, Missouri Mosby, (1998) 2767-2784.

Marenda S, Olsson J. E. The evaluation of facial paralysis. Otolaryngologic Clin, North Am 30 (1997) 669-682.

Martin, M. (Enero de 2013). <http://mmartin.cs.buap.mx/>. México, <http://mmartin.cs.buap.mx/>: <http://mmartin.cs.buap.mx/notas/PDI-MM-Rev.2013.pdf>.

Martínez, E. (2016). turing.iimas.unam.m. Mexico, turing.iimas.unam.m: turing.iimas.unam.mx/elena/PDI-Mast/Tema_6_C.pp.

Martínez, J. (2016). esi2.us.es. España, esi2.us.es:
http://www.esi2.us.es/~jdedios/asignaturas/Master_1.pdf.

May M, Klein S.R. Differential diagnosis of facial nerve palsy. Otolaryngol Clin North Am (1991) 24: 613.

MedlinePlus. (31 julio 2019). medlineplus.gov. Estados Unidos.medlineplus.gov:
<https://medlineplus.gov/spanish/ency/article/000773.htm>.

Meneses J., Gharbi T., Philippe H., Phase-unwrapping algorithm for images with high noise content based on a local histogram. *Applied Optics*. 44 (2005) 1207-1215.

MIYAMOTO, S. and UMANO, M. Recent developments in fuzzy databases and applications. Fuzzy logic and its applications to engineering information sciences and intelligent systems, Korea 1995 Editado por Bien, Z. & Min, K. C. Klumer Academic Publishers.

Parálisis de Bell (2016). www.espanol.ninds.nih.gov. Estados Unidos, www.espanol.ninds.nih.gov:

https://espanol.ninds.nih.gov/trastornos/paralisis_de_bell.htm.

Pérez Z. (2006) Aproximación Espacio-Temporal para la medida absoluta de la forma 3D de un objeto por proyección de franjas (Trabajo de grado). Escuela de Física. Universidad Industrial de Santander, Colombia.

Pratt, W.K. (2007). Digital Image Processing. (4^a ed.) New Jersey, Estados Unidos: Wiley 2014.

Quiles, F. y Garrido, A. Computadores paralelos y evaluación de prestaciones. Ciudad Real, España 1996, Ediciones de la Universidad de Castilla-La Mancha.

Rodríguez, R., y Sossa, J. Procesamiento y análisis digital de imágenes. Madrid, España 2012, Alfaomega.

Ruiz, A. (2 de Febrero de 2015). <http://dis.um.es/>. España, <http://dis.um.es/>: <http://dis.um.es/~alberto/material/percep.pdf>.

Salvi J., Pagès P., Batlle J. Pattern codification strategies in structured light systems. Instituto de Informática, Universidad de Girona, España. 2001.

Schaitkin B.M., May M. Office management of the patient with acute facial palsy. The Facial Nerve. New York Thieme, (2000) 295-300.

Solomon, C. y Breckon, T. (2011). Fundamentals of Digital Image Processing. A Practical Approach with Examples in Matlab. Noida 2011, India: Wiley-Blackwell.

Stankiewicz J.A. Steroids and idiopathic facial paralysis. Otolaryngol Head Neck Surg (1991) 672-677.

Sucar, E., y Gomez, G. (2016). <http://ccc.inaoep.mx/>. México, <http://ccc.inaoep.mx/>: <http://ccc.inaoep.mx/~esucar/Libros/vision-sucar-gomez.pdf>.

Sucar L. E., Vargas A., Hidalgo J. L., Martínez A., Leder R. S., Hernández J., Bach-y-Rita P., Paredes O., A prototype for automatic image analysis to quantify rehabilitation of chronic facial paralysis, REVISTA MEXICANA DE INGENIERÍA BIOMÉDICA. 25 (2004) 109-113.

TAHANI, V. A conceptual framework for fuzzy querying processing - a step toward very intelligent database system, Information Processing Management 13 (1977) 289 – 303.

Telischki F.F, Chandler J.R, May M. Infection: Otitis media, cholesteatoma, necrotizing external otitis, and other inflammatory disorders. The Facial Nerve. New York Thieme, (2000) 383- 397.

Tomat L. R., Manktelow M. S., Manktelow R. T., Evaluation of a New Measurement Tool for Facial Paralysis Reconstruction, *Plast. Reconstr. Surg.* (2005) 115-696.

Valentine, S., y Moughamian, D. *Fotomontaje con photoshop cs4.* (1ª ed.) Madrid España 2010, Anaya multimedia.

VETTORAZZI, J. (2007). Restauración de imágenes distorsionadas mediante técnicas de procesamiento digital y comparación entre dos métodos de restauración (Tesis Doctoral). Universidad de San Carlos, Guatemala.

Zadeh, L.A., Fuzzy set. *Information and Control*, 8 (1965) 338–353.

Zadeh, L.A. Outline of a new approach to the analysis of complex system. *IEEE Transaction on System Man and Cybernetics*, 1 (1973) 28–44.

Zadeh, L.A., The concept of a linguistic variable and its applications to approximate reasoning, *Information Science*. 8 (1975) 199–249.

ANEXOS.

A continuación se presentan los borradores de un artículo y de una patente.

Artículo

Sistema de ayuda para medir la evolución de pacientes con parálisis facial, utilizando procesamiento digital de imágenes y lógica difusa.

Simulación de medición de la evolución de pacientes con parálisis facial por medio de computadora.

Resumen: La parálisis facial ocurre cuando una persona ya no es capaz de mover algunos o ninguno de los músculos en un lado de la cara. Uno de los grandes problemas en la rehabilitación de parálisis facial, es la falta de un instrumento que permita medir si la terapia sugerida es adecuada al problema particular del paciente, es decir, en la actualidad no se cuenta con un aparato que permita a los terapeutas medir la evolución de pacientes con parálisis facial, es por eso que se desarrolló un sistema de visión artificial que aporte una solución a este problema. **Objetivo:** Desarrollar un sistema de adquisición y procesamiento de imágenes, así como una metodología, que permita medir la evolución de un paciente de parálisis facial. **Materiales y Metodología:** Desarrollamos un sistema de adquisición y procesamiento digital de imágenes para identificar puntos característicos en el rostro del paciente, una vez identificados los puntos, se midieron las distancias entre ellos, estos datos se procesaron y por medio de un algoritmo de lógica difusa, se hicieron las estimaciones de la evolución del paciente. **Resultados:** Una vez implementado el sistema, se midió la evolución de un paciente de manera cuantitativa (simulación) y se tradujeron los resultados a una escala fácil de interpretar por el terapeuta, además se integraron al sistema herramientas de apoyo al terapeuta, tal como un formulario para el registro y gestión de datos para cada paciente. **Conclusiones:** El sistema de adquisición y

procesamiento de imágenes, junto con el algoritmo de lógica difusa es un método efectivo para medir la evolución de pacientes con parálisis facial, lo que nos permite ofrecer un instrumento de medición que actualmente no existe en el mercado.

Keywords: Procesamiento digital de imágenes, lógica difusa, parálisis facial

Introducción

En México la parálisis facial es una enfermedad que suele presentarse en personas entre los 15 y 50 años de edad y es más común en mujeres que en hombres. Cada año se diagnostican aproximadamente 20 mil nuevos casos en el país¹. La secretaría de salud federal informó que el 20 por ciento de los casos de parálisis tienen secuelas permanentes¹, entre ellas, la incapacidad de cerrar los ojos para sonreír, o la falta de movilidad de un área de la cara, lo que repercute en la apariencia física y por lo tanto, en la calidad de vida del paciente. Uno de los grandes problemas en la rehabilitación de parálisis facial, es la falta de un instrumento que permita medir cuantitativamente la evolución del paciente de manera sistematizada. Actualmente los especialistas miden la evolución de manera cualitativa es decir: de acuerdo a su percepción visual. En la literatura podemos encontrar trabajos basados en métodos de medición de ángulos en el rostro, utilizando la escala de House-Brackmann para saber la gravedad de la parálisis², por otra parte otros autores miden la evolución mediante imágenes capturadas con sensores ópticos para la detección de características y las distancias entre ellas³ y otros utilizando la misma metodología de capturar imágenes para la detección de características y distancias entre ellas y además implementando la clasificación de Nottingham⁴. En este trabajo se desarrolló un sistema de adquisición de

imágenes mediante sensores ópticos, para la detección de puntos de interés, con la ventaja que además utilizamos un algoritmo basado en lógica difusa, el cual nos ayuda determinar el grado de evolución en base a las mediciones realizadas con las imágenes adquiridas, esto último nos permite obtener una escala cualitativa en el grado de evolución (Mejorando, Sin Evolución y Empeorando), ofreciendo al especialista un parámetro para determinar si la terapia aplicada es la adecuada, o es necesario realizar ajustes. Los resultados de la técnica que presentamos están basados en la percepción de un especialista, es decir, el sistema detecta los puntos de interés, mide las diferentes distancias entre ellos y determina si hay evolución positiva, caso contrario, el sistema indica que la terapia no está funcionando de manera adecuada.

El objetivo de este estudio fue desarrollar una metodología y las herramientas necesarias para poder medir la evolución de pacientes con parálisis facial.

Material y métodos

Se desarrolló e implementó un sistema para el Procesamiento Digital de Imágenes a nivel de software, además se implementó un algoritmo para el procesamiento de datos, basado en Lógica Difusa. Como elemento de pruebas para la captura de las imágenes se utilizó el rostro de un maniquí. El sistema de adquisición mostrado en la figura 1 tiene las siguientes funciones: I. Tomar las imágenes a una distancia constante, esto con ayuda de un aparato parecido al que utilizan los oftalmólogos para recargar el mentón del paciente llamado "lámpara de hendidura". II. Reconstruir virtualmente el rostro completo con base al lado sin daño para poder mostrar un resultado estimado al paciente. III. Basado en las imágenes previas, determinar la posición exacta de los puntos característicos mínimos necesarios en el rostro para poder realizar mediciones pertinentes entre estos; dichos puntos se definieron con la ayuda de un terapeuta experto. IV. Calcular las distancias entre los puntos basados en aritmética euclidiana. V. Determina el grado de evolución del

paciente, mediante la aplicación de algoritmos instrumentados en lógica difusa.



Figura 1. Sistema utilizado para la adquisición de imágenes.

De forma general el grado de evolución se determina mediante la siguiente metodología:

- 1.- Entrada de datos: hace referencia a las distancias calculadas por el sistema.
- 2.- Procesado de datos: hace referencia a la serie de reglas que se obtuvieron con ayuda del terapeuta experto en el tema y que se implementaron mediante lógica difusa.
- 3.- Salida de datos: hace referencia al resultado que nos arroja el sistema y que establece de manera puntual en que parte del rostro el paciente está mejorando o no.

Algoritmo de Lógica Difusa: La lógica difusa es una técnica de inteligencia computacional que permite trabajar con información con alto grado de imprecisión, a diferencia de la lógica binaria que trabaja con información definida y precisa ya que solo presenta dos posibilidades: falso o verdadero. La lógica difusa es multivaluada, ya que permite valores intermedios para poder definir evaluaciones entre: si/no, verdadero/falso, negro/blanco, caliente/frío etc⁵.

En nuestro caso estamos haciendo uso de la lógica difusa para determinar de manera cualitativa la evolución de un paciente. Para la implementación de esta técnica es necesario definir las llamadas variables lingüísticas, que consisten en identificar tres distancias significativas cuando se presenta una parálisis facial, además de los valores lingüísticos que toman cada una de las variables lingüísticas, en nuestro caso, estas variables toman los siguientes valores: mejorando, normal y empeorando.

Como se mencionó previamente, nuestra aplicación está basada en medir la distancia entre los puntos característicos identificados por el terapeuta, donde a partir de los cuales podemos calcular: $dist_1$ = comisura-oreja, $dist_2$ = mentón-comisura y $dist_3$ = parpado-ceja, tal como se muestra en la figura 2. En función de estas distancias el sistema determina a que valor lingüístico pertenecen, para poder así establecer el grado de evolución.

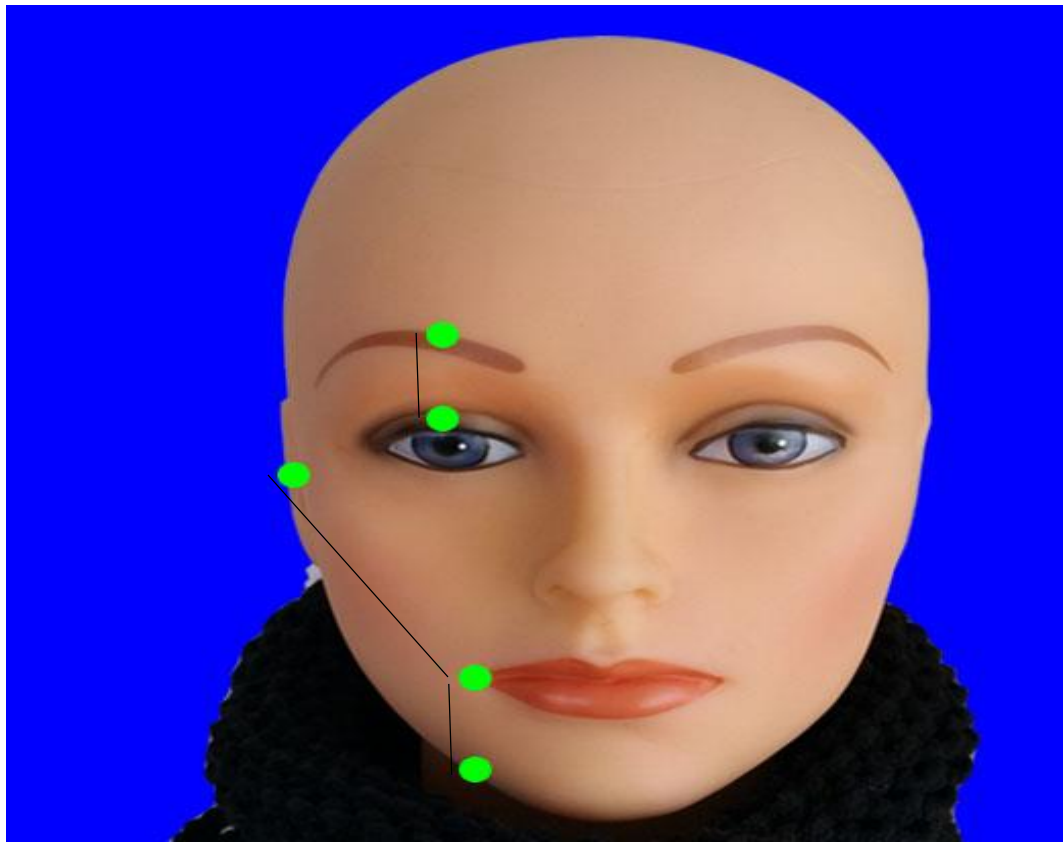


Figura 2. Distancias.

Para calcular las distancias se utiliza la siguiente ecuación:

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Se aplica la distancia euclidiana para calcular la distancia entre dos puntos.

En la tabla 1 se muestra los parámetros para poder aplicar la técnica llamada Lógica Difusa:

Variables lingüísticas	Valor lingüístico
dist ₁	Me = Mejorando
dist ₂	No = Normal
dist ₃	Em = Empeorando

Tabla 1. Parámetros Logia Difusa.

Con ayuda del terapeuta, definimos las reglas con las cuales se determinará uno de los posibles resultados que puedan presentarse en un paciente. En este caso en particular se tienen 27 posibilidades, tal como se muestra en la tabla 2. Este número viene determinado ya que solo se establecieron 3 variables lingüísticas y 3 valores lingüísticos ya que si se incrementa el número de variables lingüísticas o valores lingüísticos crecen de forma exponencial los posibles resultados y por otra parte, el experto mencionó que esas distancias son las más visibles y representativas.

En el uso de lógica difusa además de las variables y valores lingüísticos se hace uso de operadores AND y THEN que son utilizados para combinar las condiciones descritas en la tabla 2 donde podemos observar que la primer columna indica a que valor lingüístico pertenece la distancia 1, la segunda columna indica un operador AND (Y), la tercer columna a que valor lingüístico pertenece la distancia 2, la cuarta columna un operador AND (Y), la quinta columna a que valor lingüístico pertenece la distancia 3, la sexta columna un operador THEN (entonces) y la séptima columna el resultado de los valores del grado de avance después de un periodo de terapia y estos se definen

después de la tabla, donde además, el color de la celda está relacionado al nivel de mejoría del paciente: donde amarillo indica una ligera mejoría, azul indica una mejoría moderada o aceptable y rojo indica que no hay mejoría.

Expresado de otra manera por ejemplo haciendo referencia a la primera fila de la tabla 2:

Distancia 1 pertenece al valor lingüístico Mejorando y distancia 2 pertenece a Mejorando y distancia 3 pertenece a Mejorando, entonces el resultado es: A.

Tabla 2 Posibles resultados.						
dist1∈Me	AND	dist2∈Me	AND	dist3∈Me	THEN	A
dist1∈Me	AND	dist2∈Me	AND	dist3∈No	THEN	A
dist1∈Me	AND	dist2∈No	AND	dist3∈Me	THEN	A
dist1∈Me	AND	dist2∈No	AND	dist3∈No	THEN	A
dist1∈Me	AND	dist2∈Me	AND	dist3∈Em	THEN	B
dist1∈Me	AND	dist2∈No	AND	dist3∈Em	THEN	B
dist1∈Me	AND	dist2∈Em	AND	dist3∈Me	THEN	C
dist1∈Me	AND	dist2∈Em	AND	dist3∈No	THEN	C
dist1∈No	AND	dist2∈Em	AND	dist3∈Me	THEN	C
dist1∈No	AND	dist2∈Em	AND	dist3∈No	THEN	C
dist1∈Em	AND	dist2∈Me	AND	dist3∈Me	THEN	C
dist1∈Em	AND	dist2∈Me	AND	dist3∈No	THEN	C
dist1∈Em	AND	dist2∈No	AND	dist3∈Me	THEN	C
dist1∈Em	AND	dist2∈No	AND	dist3∈No	THEN	C
dist1∈No	AND	dist2∈No	AND	dist3∈Em	THEN	D
dist1∈No	AND	dist2∈Me	AND	dist3∈Em	THEN	D
dist1∈No	AND	dist2∈Me	AND	dist3∈Me	THEN	E
dist1∈No	AND	dist2∈Me	AND	dist3∈No	THEN	E
dist1∈No	AND	dist2∈No	AND	dist3∈Me	THEN	E
dist1∈No	AND	dist2∈No	AND	dist3∈No	THEN	F
dist1∈Me	AND	dist2∈Em	AND	dist3∈Em	THEN	G
dist1∈No	AND	dist2∈Em	AND	dist3∈Em	THEN	G
dist1∈Em	AND	dist2∈Me	AND	dist3∈Em	THEN	G
dist1∈Em	AND	dist2∈No	AND	dist3∈Em	THEN	G

dist1∈Em	AND	dist2∈Em	AND	dist3∈Me	THEN	H
dist1∈Em	AND	dist2∈Em	AND	dist3∈No	THEN	H
dist1∈Em	AND	dist2∈Em	AND	dist3∈Em	THEN	I

A = Ligera.

B = Ligera mejoría a nivel rama inferior facial.

C = Ligera (atención en la boca).

D = Ligera mejoría (atención en rama superior).

E = Moderada.

F = Aceptable.

G = No mejora (revisar tratamiento).

H = No mejora (revisar tratamiento, rama inferior).

I = No mejora (revisar tratamiento, cita médica con especialista).

Resultados

Un ves que se implementa el hardware y software y se establecieron las condiciones a las que operar el sistema de medición basado en los puntos que se fijaran el rostro y las reglas de operación de la lógica difusa se procesa a las pruebas con un maniquí el cual fue forzado a una deformación severa.

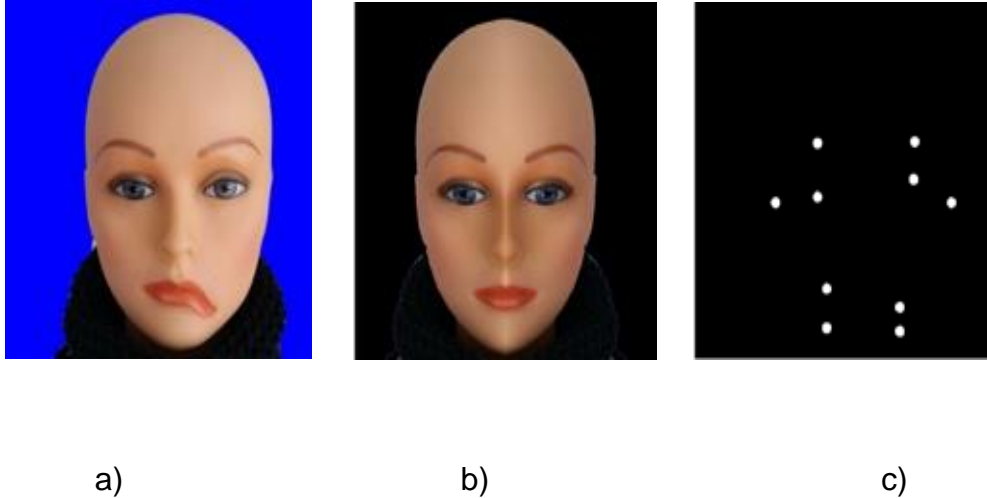


Figura 3 a) Rostro desfigurado.

En la figura 3 a) se puede observar el rostro desfigurado del maniquí donde simularemos la evolución de parálisis facial.

Figura 3 b) Rostro reconstruido virtualmente.

En la figura 3 b) se puede observar el rostro reconstruido virtualmente tomando como base la parte sin daño (imagen espejo).

Figura 3 c) Puntos característicos.

En la figura 3 c) se puede observar la detección de los puntos característicos aislados de la imagen que se toma, y de esta forma poder calcular su desplazamiento en base a las sus distancias tomando como referencia la parte de la cara que no ha sufrido daño.

El sistema que se esta proponiendo es el presente trabajo, captura las imágenes y realiza el respectivo procesameit digitalal, tal como se presenta en el siguiente ejercicio:

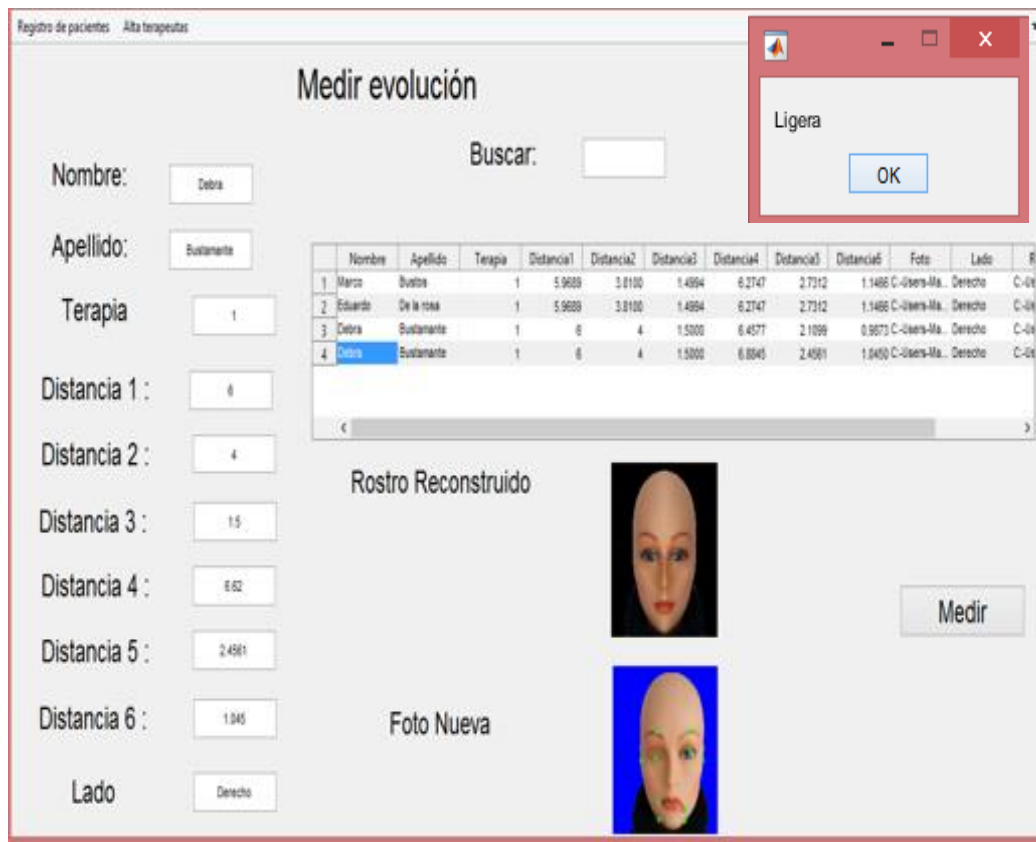


Figura 4. Primer resultado de la medición.

En la figura 4 se observa el sistema de medición de pacientes con parálisis facial donde muestra el rostro desfigurado simulando mejoría al ir moviendo los puntos característicos, el rostro reconstruido, datos del paciente y el cálculo de las distancias dadas originalmente en pixeles pero convertidas mediante un

algoritmo a centímetros, para poder indicar que el primer resultado es: **“Ligera”**.

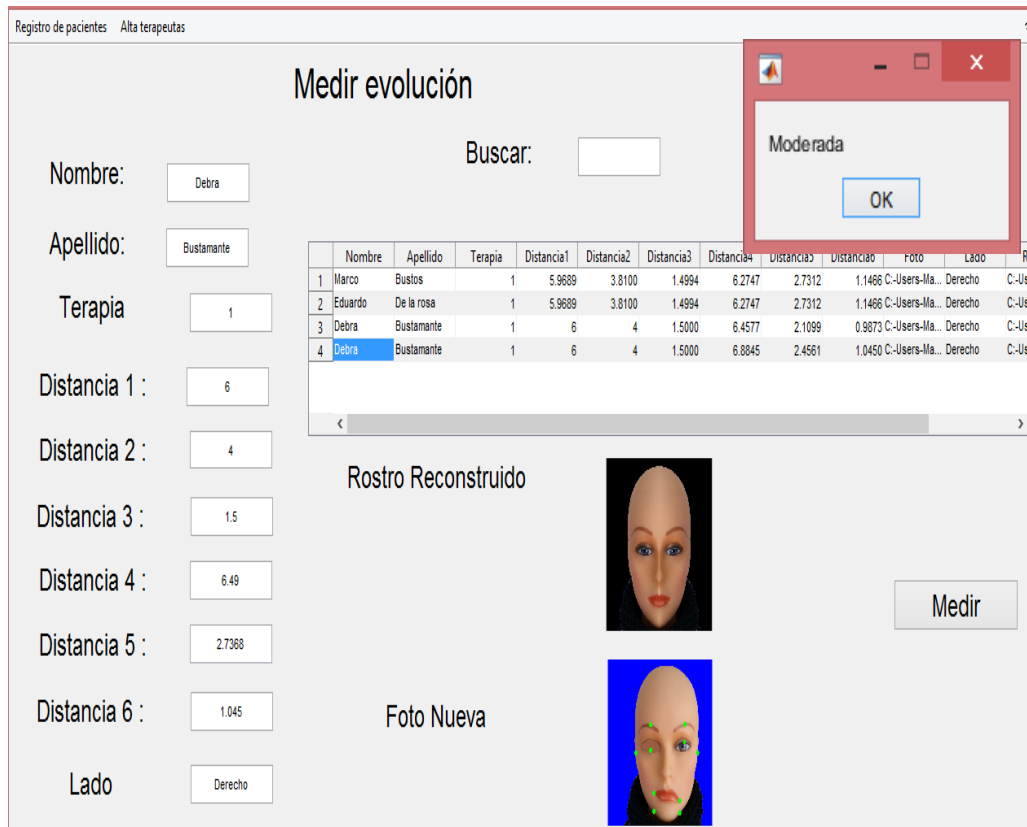


Figura 5. Segundo resultado de la medición.

Como un segundo ejercicio, en la figura 5 se observa el sistema de medición de pacientes con parálisis facial donde muestra el rostro desfigurado simulando mejoría al ir moviendo los puntos característicos, el rostro reconstruido, datos del paciente y el cálculo de las distancias dadas en pixeles pero convertidas a centímetros, para poder indicar que el segundo resultado es: **“Moderada”**.

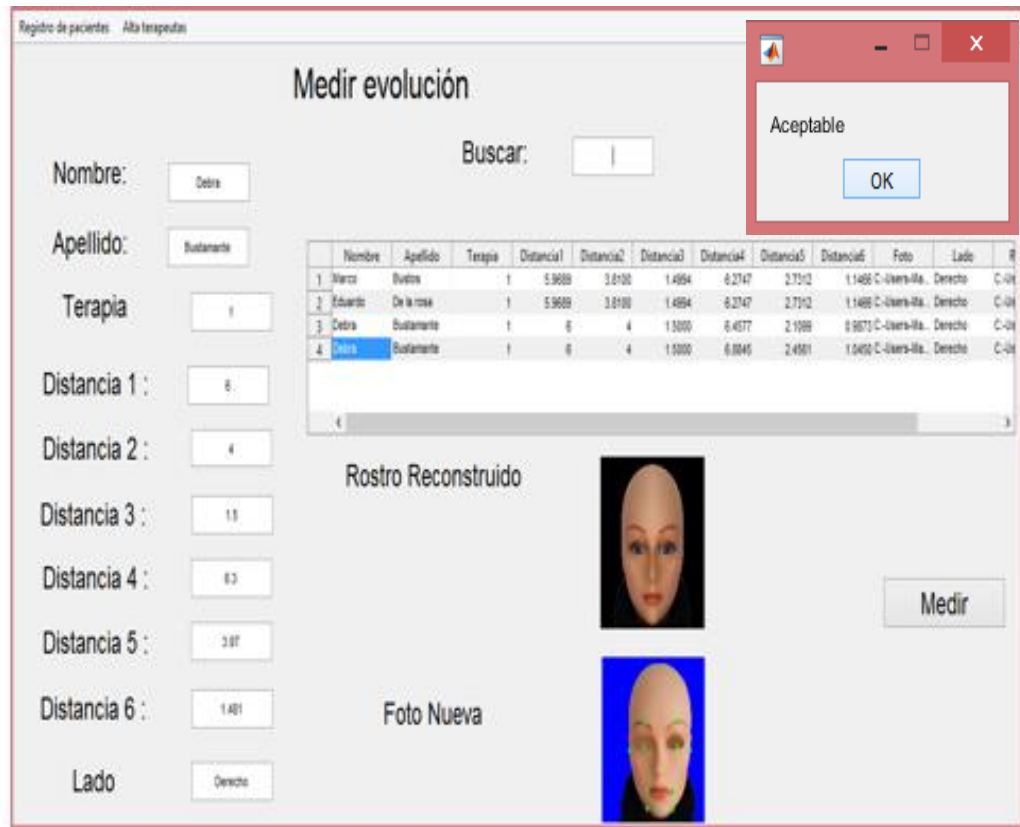


Figura 6. Tercer resultado de la medición.

En este último ejercicio de la figura 6 se observa el sistema de medición de pacientes con parálisis facial donde muestra el rostro desfigurado simulando mejoría al ir moviendo los puntos característicos, el rostro reconstruido, datos del paciente y el cálculo de las distancias dadas en pixeles pero convertidas a centímetros, para poder indicar que el tercer resultado es: **“Aceptable”**.

Tabla 3 Resumen del proceso de evolución.				
Terapia	dist ₁	dist ₂	dist ₃	Resultado
1	6.62	2.4561	1.045	Ligera
2	6.49	2.7368	1.045	Moderada
3	6.3	3.97	1.481	Aceptable

En la tabla 3 podemos observar el resumen de simulación de la evolución de un paciente con parálisis facial donde las distancias se van moviendo y van tomando su valor lingüístico para poder arrojar un posible resultado de los antes mencionados.

Discusión

El principal problema en la medición de pacientes con parálisis facial, es la falta de objetividad a la hora de determinar si la terapia está funcionando. En la gran mayoría de los casos los pacientes, se puede lograr una recuperación casi total entre 3 y 5 sesiones de terapia, las cuales suelen ser cada 15 días, pero hay casos en los pacientes que a pesar de que hay mejoría, esta es imperceptible (de forma visual) y los pacientes se depriman, lo que puede ocasionar que el tiempo de recuperación sea mayor, esto de acuerdo con el especialista. Con la ayuda del especialista, se analizaron tres distancias significativas que se mueven al observar parálisis facial, donde los pacientes pueden presentar desfiguración de la comisura y no cerrar un ojo completamente.

Analizando los resultados de la presente investigación, se llegó a la conclusión que si es posible medir la evolución de pacientes con parálisis facial y también poner atención en las partes del rostro donde no hay mejoría, ya que uno de los propósitos es medir cuantitativa y cualitativamente, la aportación más relevante de este trabajo es el sistema desarrollado que le permitirá al terapeuta llevar un expediente de sus pacientes, mostrar un posible resultado terminando las terapias, medir el desplazamiento de las distancias en el rostro y dar el resultado que arroja el sistema. El terapeuta puede tomar este sistema para apoyarse en los resultados, y así poder informar al paciente si está evolucionando o no, también le permitirá detectar dónde poner atención en dado caso que el resultado que se obtenga no sea favorable, y así cambiar la terapia o canalizarlo con otro especialista y ver porque no hay mejoría.

Patente

SISTEMA DE ADQUISICION Y PROCESAMIENTO DIGITAL DE IMÁGENES PARA LA MEDICION DE LA EVOLUCIÓN EN PACIENTES CON PARÁLISIS FACIAL.

ANTECEDENTES

En la actualidad el principal problema en la medición de la evolución de pacientes con parálisis facial, es la falta de objetividad a la hora de determinar si la terapia está funcionando o no. En la gran mayoría de los casos, los pacientes pueden lograr una recuperación casi total después de 3 a 5 sesiones de terapia, las cuales suelen ser cada 15 días, pero hay casos en que a pesar de haber mejoría, esta es imperceptible, lo que puede ocasionar que los pacientes sufran depresión y que el tiempo de recuperación sea mayor.

Se han propuesto solicitudes de patentes que buscan proteger inventos relacionados con la parálisis facial, por ejemplo: CN101341515 A, propone un sistema informático para clasificar la parálisis facial, que incluye un sistema de cámara para recopilar una serie de imágenes de la mirada del paciente para su análisis por un sistema informático, y en base a esto el sistema le sugiere al paciente realizar movimientos específicos o ejercicios faciales, tales como sonreír.

Otro caso es el de la solicitud CN109508644, donde proponen un sistema de evaluación del grado de parálisis facial basado en análisis de datos de video en profundidad, el cual en base a procesamiento digital, determina y le informa a una persona el grado de parálisis que sufre.

Observando la necesidad de informar al paciente no solo de su padecimiento, sino de la evolución en su parálisis facial, se desarrolló en un sistema que a continuación se describe. Se trata de un sistema de adquisición de imágenes que mide la evolución de pacientes con parálisis facial, que, a diferencia de las solicitudes de patente ya mencionadas, esta indica al paciente si está mejorando o empeorando, y en caso de esta última, le informa al paciente en que parte de la cara debe prestar atención.

Las características y detalles del sistema se describen claramente en en los dibujos que acompañan esta solicitud.

BREVE DESCRIPCION DE LAS FIGURAS

Fig. 1 Montaje del sistema completo.

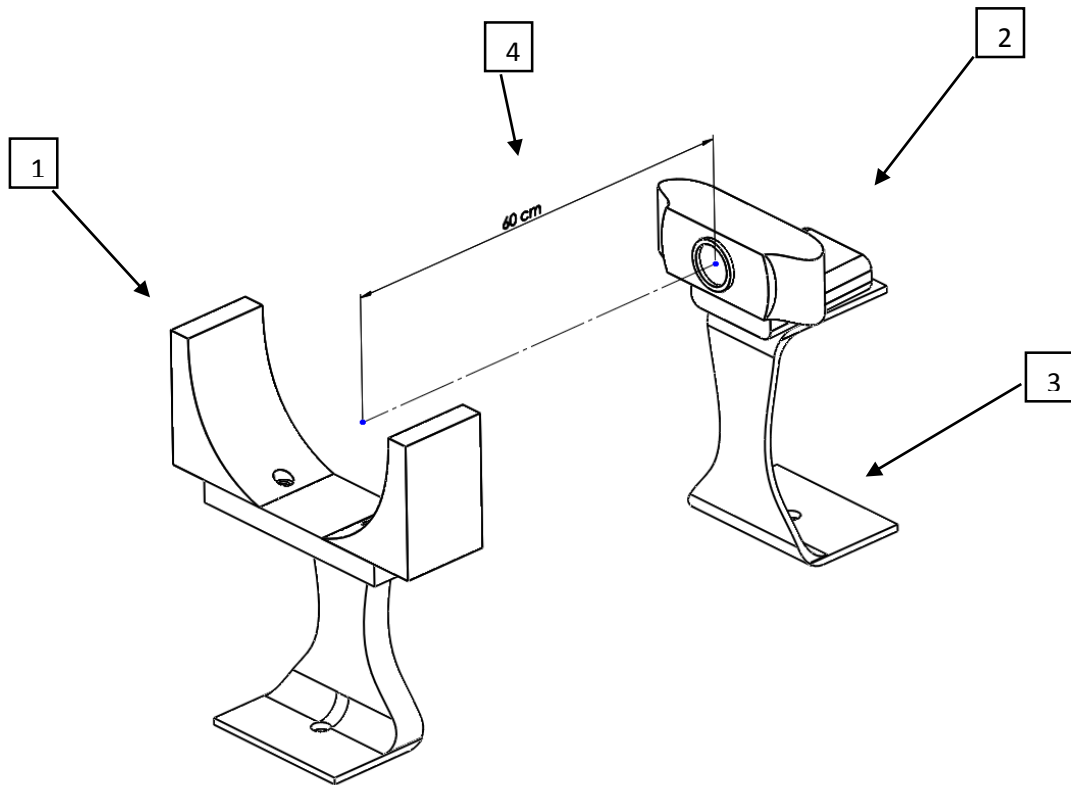


Fig. 2 Base ajustable que sostiene la cara del paciente.

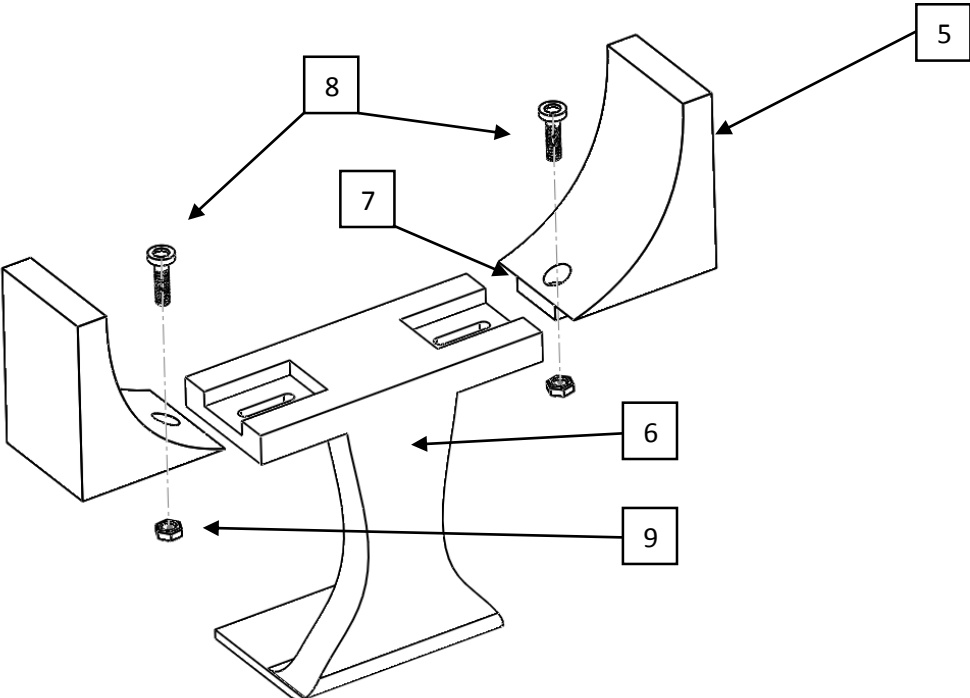


Fig. 3 Base que sostiene el sensor óptico.

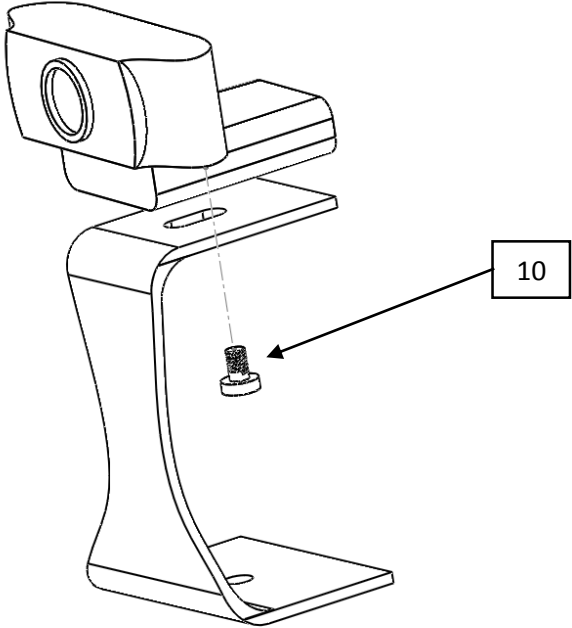


Fig. 4

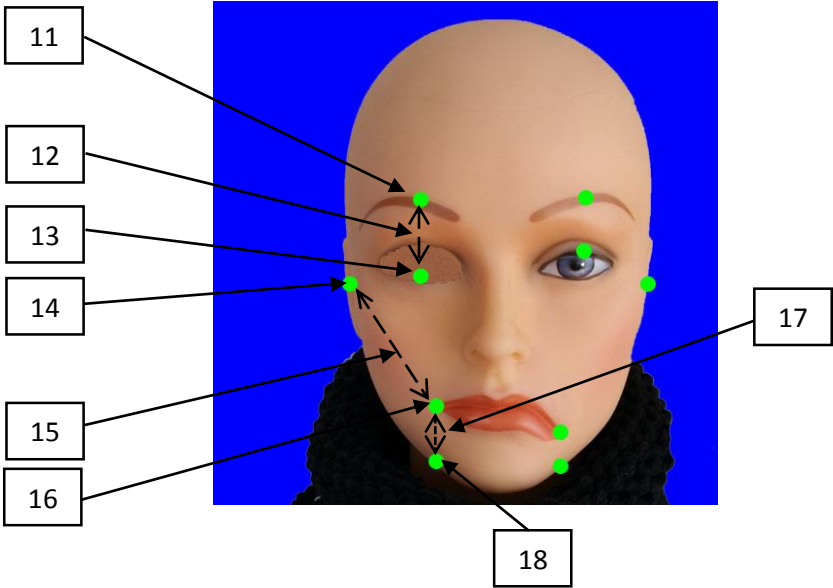


Fig. 5 Interfaz gráfica para el registro de los pacientes.

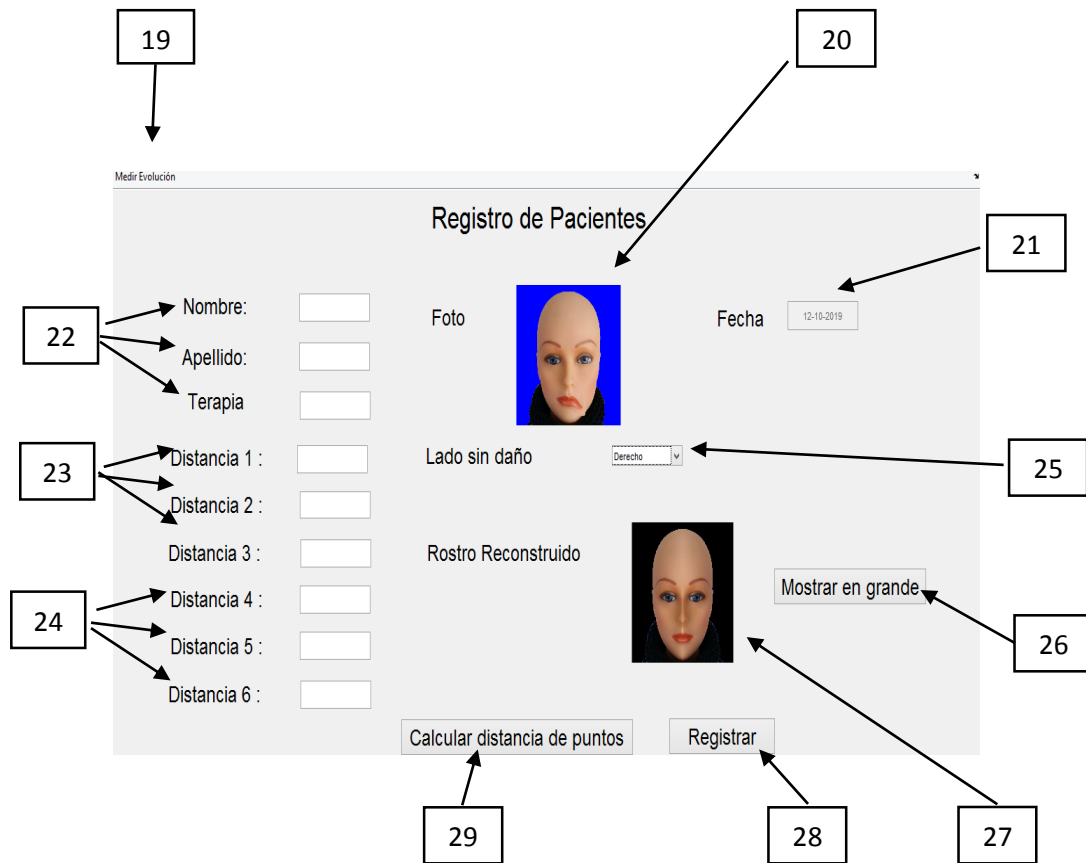
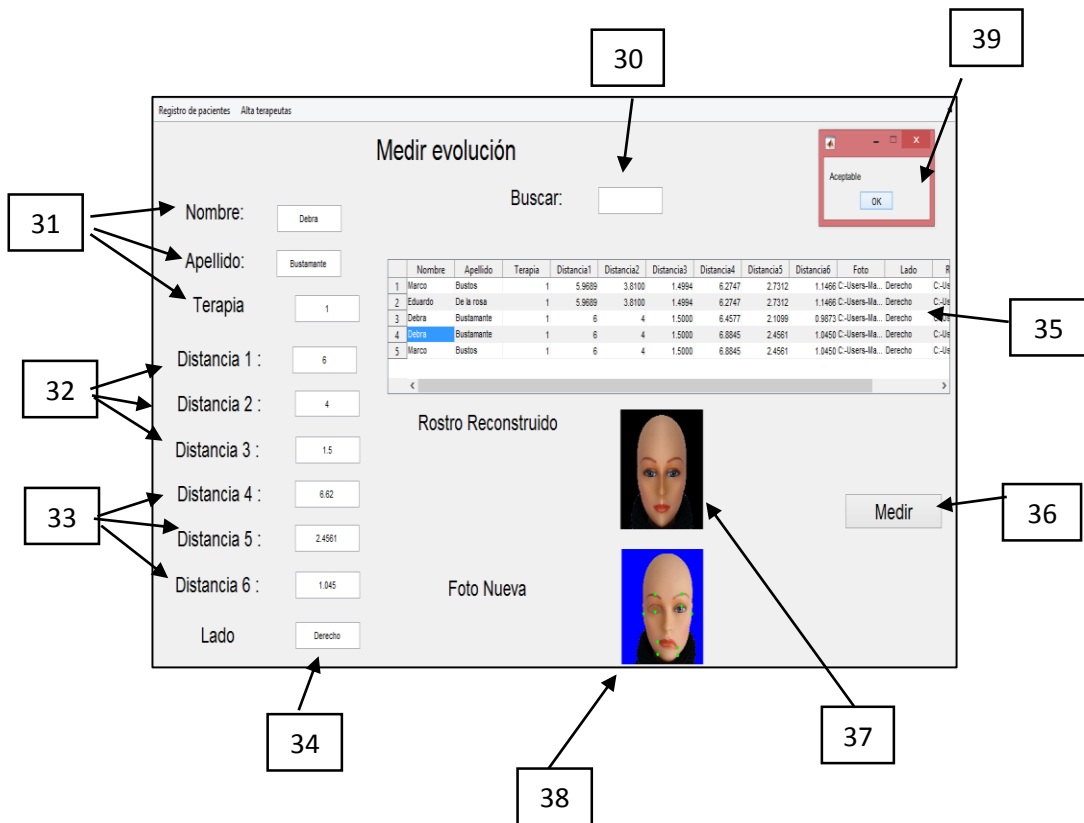


Fig. 6 Interfaz gráfica para medir la evolución del paciente.



SISTEMA DE ADQUISICION Y PROCESAMIENTO DIGITAL DE IMÁGENES PARA LA MEDICION DE LA EVOLUCIÓN EN PACIENTES CON PARÁLISIS FACIAL.

Esta invención se refiere a un sistema de adquisición y procesamiento digital de imágenes para medir la evolución de pacientes con parálisis facial, el sistema está constituido por un montaje (Ver fig. 1), que consta de una base (1), la cual cuenta con un par de mordazas (5) ajustables en un rango de separación entre mordaza y mordaza desde 12 cm y hasta 20 cm, cada mordaza cuenta en la parte inferior con una saliente (7) la cual se desliza sobre una base principal (6) permitiendo ajustarlas a una distancia determinada, las mordazas se fijan a la base (6) mediante dos tornillos (8) y dos tuercas (9), dejándolas se puede fijar en una posición, las mordazas se ajustan de acuerdo al tamaño del mentón de cada paciente, para que este se coloque de tal manera que su cara quede de frente al sensor óptico (2), el cual está fijado a una base (3) mediante un tornillo (10), la distancia que separa a ambas bases es de 60 cm (4), la cual se estableció en base a las pruebas experimentales, a esta distancia existe suficiente información de píxeles para hacer mediciones milimétricas en un rostro, el sensor óptico es una cámara web con resolución Full HD 1080p, utilizado para capturar imágenes RGB_1920x1080, una vez que el mentón del paciente queda fijo a la base (1), y con base a la recomendación de un terapeuta, se le deben colocar algunos puntos de referencia al paciente (Ver fig. 4) en comisura (16), oreja (14), parte inferior del mentón (18), ceja (11) y párpado (13), para poder obtener las distancias entre la comisura-oreja (15), comisura-mentón (17) y ceja-párpado (12) en ambas partes del rostro, ya que son las más representativas cuando se presenta una parálisis facial, después se procede a capturar una imagen para su procesamiento digital mediante software, la imagen (20) se muestra en la interfaz "Registro de pacientes" (Ver fig. 5), la cual también muestra la fecha de captura (21), una vez capturada la imagen se procede al registro del paciente, ingresando nombre, apellido y número de terapia (22),

posteriormente la interfaz permite seleccionar mediante un menú desplegable (25), cual es el lado del rostro del paciente sin daño, “derecho” o “izquierdo”, para que mediante procesamiento digital se discrimine la parte dañada, después, el lado sin daño se utiliza para obtener una nueva imagen en espejo, ambas imágenes se unen para crear una nueva imagen (27) y mostrar cómo se vería el paciente sin daño alguno, la interfaz cuenta con un botón “Mostrar en grande” (26) que le permite mostrar la imagen (27) de manera más a detalle. Al presionar el botón “Calcular distancia entre puntos” (29) el sistema aplica un algoritmo de procesamiento digital el cual consiste en discriminar el rostro para dejar solamente los puntos de referencia y calcular las distancias (12) (15) (17) de la siguiente manera: Dado que la imagen está compuesta por pixeles, el algoritmo encuentra el pixel del centro de cada punto de referencia (11)(13)(14)(16)(18) en el plano cartesiano, y aplica la ecuación trigonométrica del teorema de Pitágoras para el cálculo de las distancias(12) (15) (17) entre los centros de los puntos, dado que las distancias están dadas en pixeles, es necesario convertirlas a milímetros o centímetros de la siguiente manera: sabiendo el diámetro real de alguno de los puntos de referencia, el algoritmo calcula a cuantos pixeles equivale en la imagen tomada y con una regla de tres se determina a cuantos cm o mm equivale un pixel, una vez hecha la conversión de pixeles a milímetros, se muestra el resultado de las distancias calculadas de la parte sin daño (23) y el resultado de la parte con daño (24), para finalizar se presiona el botón de “Registrar” (28), para que se registre en una base de datos la información de cada paciente, la interfaz también cuenta con la pestaña (19) para dirigirse a la interfaz, “Medir Evolución”, (Ver fig. 6) la cual cuenta con un buscador (30) para encontrar rápidamente ya sea al paciente que se acaba de registrar o algún otro previamente registrado, en la tabla (35) se elige al paciente y la interfaz muestra el nombre, apellido y el número de terapia (31), la interfaz también indica el lado de la cara afectada (34), las distancias del lado sin daño (32), también muestra la imagen que pertenece a la sesión de terapia en que se encuentra el paciente (38), se muestra nuevamente el rostro reconstruido (37), al presionar el botón de medir,

se calculan las nuevas distancias del lado con daño del paciente, aplicando el mismo algoritmo de procesamiento digital del botón “Calcular distancias” (29) de la interfaz “Registro de pacientes” (Ver fig. 5), después estas distancias calculadas alimentan un algoritmo de lógica difusa para poder medir de manera cualitativa la evolución. Para la implementación de esta técnica, se definieron las llamadas variables lingüísticas, que apoyados de un terapeuta especialista, consisten en identificar tres distancias (12) (15) (17) significativas cuando se presenta una parálisis facial, además de los valores lingüísticos que toman cada una de las variables lingüísticas, en nuestro caso, se toman los siguientes valores: mejorando, normal y empeorando.

También, con ayuda del terapeuta, se definieron las reglas del algoritmo de lógica difusa, con las cuales, el sistema determina cada uno de los posibles resultados que puedan presentarse en un paciente, en nuestro caso, debido a que se establecieron 3 variables lingüísticas y 3 valores lingüísticos, el sistema arrojará uno de 27 posibilidades, pudiendo incrementarse de manera exponencial al incrementar el número de valores y variables lingüísticas. Se llegó a la determinación de que fueran únicamente 3 variables debido a la recomendación del terapeuta, quien argumentó que esas distancias (12)(15)(17) son las más visibles y representativas al momento de determinar la evolución en parálisis facial. El resultado es mostrado en una ventana emergente (39). En dado caso que el resultado sea desfavorable para el paciente, el sistema es capaz de determinar en qué parte de la cara debe el terapeuta prestar mayor atención o si es necesario canalizarlo con otro tipo de especialista.

REIVINDICACIONES

1.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial, que comprende:

Una base ajustable desde 12 cm y hasta 20 cm de ancho, donde el paciente puede colocar su mentón;

Una base colocada a 60 cm de la primera, que soporta a una cámara web con resolución Full HD 1080p la cual captura una imagen del paciente; y

Un sistema de software caracterizado por medir la evolución de un paciente con parálisis facial, el sistema mide el cambio de 3 distancias representativas en el rostro del paciente mediante procesamiento digital de imágenes y mediante un algoritmo de lógica difusa con 3 variables lingüísticas y 3 valores lingüísticos programadas con ayuda de un terapeuta especialista. El sistema arroja una de 27 posibilidades, indicando al paciente si está mejorando o no, y en que parte de la cara prestar mayor atención.

2.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial de acuerdo a la reivindicación 1, caracterizada por tener mayor o menor longitud de distancia entre la base donde el paciente coloca su mentón y la cámara, cambiando únicamente la relación de pixeles/longitud con la que se hace el procesamiento digital.

3.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial de acuerdo a la reivindicación 1, caracterizada por contar con cualquier tipo de cámara digital para realizar la captura de la imagen del paciente, cambiando únicamente el número de pixeles capturados.

4.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial de acuerdo a la reivindicación 1, caracterizada por medir el cambio en “N” distancias representativas.

5.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial de acuerdo a la reivindicación 1, caracterizada por que algoritmo de lógica difusa cuenta con “N” variables lingüísticas y “N” valores lingüísticos programadas.

6.- Sistema de adquisición y procesamiento digital de imágenes para la medición de la evolución de pacientes con parálisis facial de acuerdo a la reivindicación 1, caracterizada por arrojar una respuesta de entre “N” cantidad de posibilidades.

RESUMEN

Esta invención se refiere a un sistema que mide de manera cuantitativa y cualitativa la evolución de pacientes con parálisis facial. El sistema consta de tres partes, 1.- base ajustable diseñada para que el paciente pueda colocar su mentón, 2.- base para sostener un sensor óptico a 60 cm de distancia frente al rostro del paciente, y 3.- un sistema digital, el cual se encarga de procesar digitalmente las imágenes y medir la evolución del paciente mediante la implementación de un algoritmo de lógica difusa, determinando así la evolución del paciente y realizando recomendaciones en base a ello.