

SISTEMA DE VISIÓN ARTIFICIAL PARA LA ESTIMACIÓN DEL PESO DE GANADO
BOVINO

NESTOR ALONSO ARIAS HERNANDEZ

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE FISICA
BUCARAMANGA
2004

SISTEMA DE VISIÓN ARTIFICIAL PARA LA ESTIMACIÓN DEL PESO EN GANADO
BOVINO

NESTOR ALONSO ARIAS HERNANDEZ

Trabajo de investigación presentado como requisito
Para optar al título de MAGÍSTER EN FÍSICA

Director
OSCAR GUALDRON GONZALEZ
Físico, Ph.D.

UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE CIENCIAS
ESCUELA DE FÍSICA
BUCARAMANGA
2004

AGRADECIMIENTOS

- El autor expresa sus más sinceros agradecimientos:
- A mi director Dr. Oscar Gualdrón por iniciarme en este tema. Por su guía y apoyo a lo largo del trabajo.
- A los Drs. Arturo Plata y Jaime Meneses por sus discusiones alrededor de mi tesis.
- Al Dr. Yesid Torres, por sus enseñanzas académicas, y apoyo para la culminación de la tesis.
- A la Dr. Maria del C. Lasprilla por su colaboración, apoyo y preocupación incondicional.
- A los compañeros de maestría y doctorado del GRUPO DE OPTICA Y TRATAMIENTO DE SEÑALES (UIS) por sus fructuosas discusiones y colaboración.
- Al GRUPO DE CONECTIVIDAD Y PROCESADO DE SEÑAL (Ingeniería Eléctrica y Electrónica-UIS), por las herramientas facilitadas para la implementación del mecanismo.
- Al LABORATORIO DE ÓPTICA E INFORMÁTICA de la Universidad Popular del Cesar por su colaboración y ayuda logística para el desarrollo de este proyecto.
- A los profesores de la Escuela de Física de la Universidad Industrial de Santander por los conocimientos impartidos.

*A mi madre, a mi esposa y a mi hijo,
Con todo mi amor.*

TABLA DE CONTENIDO

	Pág.
1. INTRODUCCIÓN.	1
2. SISTEMA DE VISION ARTIFICIAL.	3
2.1. VISION ARTIFICIAL.	3
2.2 PROCESADO DE IMÁGENES.	8
2.2.1 Conversión de niveles de gris.	9
2.2.2 Transformaciones geométricas.	9
2.2.3 Transformación del histograma.	9
2.2.4 Filtrado espacial y frecuencial.	10
2.2.5 Segmentación.	12
2.2.6 Extracción de rasgos característicos.	13
2.2.7 Clasificación de los objetos.	13
2.3 INTELIGENCIA ARTIFICIAL Y REDES NEURONALES.	15
2.3.1 Redes neuronales.	15
3. ADQUISICIÓN.	24
3.1 ADQUISICIÓN DE DATOS.	25
3.2 PREPROCESAMIENTO.	26
4. SEGMENTACIÓN.	28
4.1 SEGMENTACIÓN POR DIFERENCIA DE IMÁGENES (OBJETO-REFERENCIA)	29
4.2 SEGMENTACIÓN POR DIFERENCIA DE IMÁGENES DE SECUENCIA.	31
4.3 UMBRALIZACIÓN.	35
4.3.1 Umbralización Global.	37
4.3.2 Umbralización Local.	47

4.4 SELECCIÓN DEL OBJETO DE INTERÉS.	48
4.5 COMPARACIÓN DE LOS MÉTODOS DE SEGMENTACIÓN.	52
5. EXTRACCIÓN DE CARACTERÍSTICAS.	53
5.1 INVARIANZA A LA ROTACIÓN.	56
5.1.1 Transformada de Hotelling O K-L.	56
5.2 OBTENCIÓN Y SELECCIÓN DE CARACTERÍSTICAS.	60
5.2.1. Análisis de componentes principales.	64
6. ESTIMACIÓN DE PESO POR REDES NEURONALES.	67
6.1 IMPLEMENTACIÓN DE LA RED NEURONAL.	67
6.1.1 Estructura de la red neuronal.	67
6.1.2 Entrenamiento de la red.	68
7. CONCLUSIONES.	76
7.1 SUGERENCIAS PARA TRABAJOS FUTUROS.	78
8. BIBLIOGRAFÍA.	80
ANEXO 1. Datos Registrados Experimentalmente.	82
ANEXO 2. Errores por el método de segmentación por diferencia de imágenes objeto referencia.	83
ANEXO 3. Errores por el método de segmentación por diferencia de secuencias de imágenes.	85

LISTA DE FIGURAS

	Pág.
Figura 2.1. Sistema de coordenadas utilizado para la representación digital de imágenes.	6
Figura 2.2 (a) imágenes digitales de la misma escena, con una resolución espacial reducida progresivamente: 512x512, 256x256, 128x128 y 64x64. (de izq. A der. Y de arriba abajo) (b) y cuatro imágenes digitales de la misma escena, que se han codificado en 256, 16, 8 y 4 niveles de gris, respectivamente.	7
Figura 2.3. Diagrama de bloques del análisis automático de una escena.	8
Figura 2.4. (a) Imagen digital y su correspondiente histograma. (b) imagen ecualizada y su histograma.	10
Figura 2.5 Diagrama de bloques de las etapas en que puede organizarse un SVA.	14
Figura 2.6. Neurona Biológica.	16
Figura 2.7. Modelo de neurona artificial.	17
Figura 2.8. Funciones de activación.	18
Figura 2.9. Red progresiva.	20
Figura 2. 10. Red realimentada.	21
Figura 3.1. Montaje.	24
Figura 3.2. Algunas imágenes obtenidas.	25
Figura 3.3 (a) Imagen sin filtro. (b) perfil de imagen sin filtro. (c) Imagen con filtro. (d) perfil de imagen con filtro.	27
Figura 4.1 . Algunas imágenes referencia y objeto.	29
Figura 4.2 Imágenes diferencia.	30
Figura 4.3 Imágenes de distancia euclidiana en el espacio.	31
Figura 4.4 Imágenes de secuencia.	32

Figura 4.5. Diferencia de imágenes de secuencia.	33
Figura 4.6 diferencia de imágenes desplazadas. Sup. Δy , inf. Δx	34
Figura 4.7. Acumulación de planos de color para las tres diferencias.	36
Figura 4.8 Resultados de la umbralización manual para las diferencias de imágenes (Objeto-Referencia).	38
Figura 4.9. Resultados de la Umbralización manual.	39
Figura 4.10. Unión lógica de las imágenes binarizadas.	40
Figura 4.11. Comportamiento del umbral en el método iterative.	44
Figura 4.12. Imágenes binarizada con umbralización con método iterative.	44
Figura 4.13. Obtención del umbral en el método Histogram.	46
Figura 4.14. Imágenes binarizada con umbralización con método histogram.	47
Figura 4.15. Desplazamiento de las subimágenes.	48
Figura 4.16. Imágenes binarizada con umbralización local	49
Figura 4.17. Diagrama de bloques para la selección del área de interés.	49
Figura 4.18. Objeto aislado por diferencia de imágenes (objeto-referencia) (a) M. Manual. (b) M. Iterative. (c) M. Histogram. (d) M. Local.	50
Figura 4.19. Objeto aislado por diferencia de secuencias de imágenes (a) M. Manual. (b) M. local.	51
Figura. 4.20. Regiones que representan los tipos de errores.	52
Figura 4.21. Resultado de utilizar la segmentación semiautimatica.	53
Figura 5.1. Dimensiones o características extraídas de la imagen.	54
Figura 5.2. Comportamiento de las características frente al peso del animal.	63
Figura. 5.3. Ejemplos de obtención de características sobre el animal.	64
Figura 6.1. Curvas de entrenamiento de los tres grupos.	71
Figura 6.2. a. Matriz de error para la selección de el numero de capas y de neuronas de la red neuronal. b. Comportamiento del error al variar la semilla. (primer experimento)	72
Figura 6.3 a. Matriz de error para la selección de el numero de capas y de neuronas de la red neuronal. b. Comportamiento del error al variar la semilla. (segundo experimento)	73
Figura 6.4 Comportamiento de los datos de salida de la red frente a los datos esperados.	74

LISTA DE TABLAS

	Pág.
Tabla 4.1. Errores porcentuales promedios de los diferentes métodos de segmentación.	53
Tabla 4.2. Errores porcentuales promedios de los diferentes métodos de segmentación.	54
Tabla 5.1. Coeficiente de correlación de las características con el peso del animal.	62
Tabla 6.1. Configuración y error de las redes neuronales (primer experimento).	74
Tabla 6.2. Configuración y error de las redes neuronales (segundo experimento).	74

TITULO*: SISTEMA DE VISION ARTIFICIAL PARA LA ESTIMACIÓN DEL PESO EN GANADO BOVINO

AUTOR: NÉSTOR ALONSO ARIAS HERNÁNDEZ**

PALABRAS CLAVES: visión artificial, redes neuronales, procesamiento digital de imágenes, reconocimiento de patrones.

RESUMEN

Se diseñó y desarrolló un procedimiento dentro del marco de la visión artificial que estima el peso en ganado bovino. Los datos de entrada al sistema son imágenes obtenidas por una cámara de vídeo a color. Estas imágenes son digitalizadas y luego preprocesadas utilizando filtros de eliminación de ruido; posteriormente se implementan y se evalúan diferentes métodos segmentación, que permiten obtener el contorno del animal en forma semiautomática. Este proceso consta de una binarización automática en su etapa inicial y un eventual ajuste manual. Posteriormente se extraen las características que tengan una fuerte correlación con el peso del animal. Estas son extraídas de tal forma que su obtención sea independiente de la orientación del objeto dentro de la imagen. Tales características incluyen: el área superior, perímetro, ancho de abdomen, ancho de anca, ancho de escápula. La estimación del peso del animal se hace por medio de una red neuronal artificial tipo feedforward, cuyas entradas se alimentan con las características extraídas de la imagen. Finalmente el sistema es evaluado en el número y tipo de características tenidas en cuenta y en la estructura de la red neuronal utilizada.

* Trabajo de Investigación.

** Facultad de Ciencias. Maestría en Física. GUALDRÓN GONZALEZ Oscar.

TITLE*: SYSTEM OF ARTIFICIAL VISION FOR THE ESTIMATE OF THE WEIGHT IN BOVINE LIVESTOCK.

AUTHOR: NESTOR ALONSO ARIAS HERNANDEZ**

KEY WORDS: artificial vision, neural networks, digital processing images, pattern recognition.

ABSTRACT

A procedure inside the context of artificial vision that estimates the weight in bovine livestock was developed and designed. The input data to the system are images obtained by a videotape camera to color. These images are digitized and preprocessing using filters of elimination of noise; later are implemented and evaluated different segmentation methods that allow to obtain the contour of the animal in semiautomatic form. This process consists of an automatic binarización in its initial stage and an eventual manual adjustment. Later the characteristics that have a strong correlation with the weight of the animal are extracted. These are extracted in a form that is independent of the orientation of the object inside the image. Such characteristics include: the superior area, perimeter, wide of abdomen, wide of haunch, wide of scapula. The estimate of the weight of the animal is made by means of a neural network type feedforward whose inputs feed with the extracted characteristics of the image. Finally the system is evaluated in the number and type of characteristics kept in mind and in the structure of the neural network utilized.

* Research Work.

** Facultad de Ciencias. Maestría en Física. GUALDRÓN GONZALEZ Oscar.

1. INTRODUCCIÓN

En el mundo agropecuario una de las actividades más importantes es la ganadería. La tecnificación de la ganadería constituye un aspecto fundamental para incrementar su rentabilidad y de esta forma trasladar a la sociedad todos los beneficios asociados como generación de empleo, incremento del consumo de productos veterinarios o incremento de exportaciones por mencionar algunos. Sin embargo aún es muy frecuente utilizar procedimientos rudimentarios que impiden el despegue de esta próspera actividad.

Uno de los momentos de mal manejo y de peor trato del ganado bovino se presenta durante su entrada y salida a la báscula para el peso en vivo, en donde el animal es hostigado y golpeado con toda clase de objetos (palas, barras de hierro, tábano eléctrico, e incluso rociarle líquido inflamable en la cola y prenderle fuego) principalmente en las nalgas, cadera y piernas, lugar donde se encuentran las carnes de primera categoría. Esto ocasiona canales mal sangradas, carnes oscurecidas que presentan un mal aspecto y desdichan de su verdadera calidad^[1]. Además el solo hecho de realizar el pesaje del animal aumenta su ritmo cardíaco generando un elevado estrés que produce mermas en el animal^[2].

De esta forma se realiza el pesaje para el monitoreo y posterior comercialización de animales destinados a carne, salvo grandes empresas, mataderos, frigoríficos; que hacen contrataciones en las ya conocidas subastas^[1], pero en cuanto al seguimiento y monitoreo del animal continua utilizándose la báscula. Esto obliga a plantear nuevas alternativas para realizar el pesaje de ganado en vivo.

Los pequeños ganaderos que no tienen acceso a una báscula realizan la compra y venta del ganado de una manera poco traumática para el animal, pero, con baja calidad técnica, no por esto menos precisa. Estos han desarrollado la habilidad de poder determinar el peso del

animal con bastante precisión, solo con una rápida inspección visual. El pesaje del animal siempre se hace a campo abierto, en el hábitat normal del ganado, eliminando el estrés y maltrato del animal.

En la producción de ganado (bovino, porcino, etc.), el pesaje de los animales juega un papel importante en el control de factores que afectan la salida del lote para su comercialización. Estos factores incluyen la concesión de espacios de pastoreo, la regulación de la ración de comida y la detección de brotes epidémicos. Métodos para calcular el peso vivo del ganado son importantes desde este punto de vista. Normalmente, el pesaje de ganado se realiza manualmente, un proceso que a menudo necesita de más de 4 vaqueros, y toma aproximadamente de 2 a 4 minutos por animal; este procedimiento es muy estresante tanto para el animal como para el ganadero. Para obtener una alta precisión en el pesaje de comercialización, éste se realiza nuevamente antes del sacrificio del animal. Hasta el momento muchos intentos se han hecho para encontrar una alternativa al procedimiento de pesaje manual.

Principalmente dos alternativas han sido estudiadas: *la primera*, sistemas electrónicos para el pesado automáticos en combinación con equipo de identificación automática, y *la segunda*, la determinación del peso en forma indirecta desde dimensiones, medidas por una cinta métrica o por sistemas de análisis de imágenes.

Los sistemas de pesaje e identificación semi-automáticos^[3] en cerdos, fueron evaluados obteniendo buenos resultados en cuanto a la disminución del número de vaqueros y el tiempo de pesaje por cerdo, pues lograron tasas de 100 cerdos por hora. Posteriormente se implementó un sistema de identificación automática y un equipo de procesamiento de datos^[4], en donde se muestra que el peso vivo del animal tiende a obtenerse automáticamente y con más fiabilidad que en el pesaje manual, debido a la variación durante el día del peso vivo. También han sido estudiados sistemas automáticos para diferentes especies^[5] obteniéndose resultados en aves y ganado bastante aceptables.

La importancia de la correlación entre el peso vivo y las dimensiones de los cerdos ha dado pie para que muchos autores estudien la posibilidad de estimar el peso del animal a partir de las dimensiones del cuerpo de éste. Se han estudiado métodos para obtener estas dimensiones en cerdos. Los métodos para medir son: la cinta métrica, un instrumento escalado de ganado, y dimensiones obtenidas desde fotografías^[7]. El primer método es altamente preciso. Se ha encontrado que el peso de los cerdos se puede describir como una función cuadrática de las dimensiones del cuerpo^[8]. También se encontró que el área ocupada por el cerdo, depende del peso vivo, descrito por la función $\text{área-cerdo} = KxW^{0.66}$, donde K es una constante y W es el peso vivo^[9]. También se intentó usar la circunferencia como una escala de peso del cerdo^[9], pero, ésta no es conveniente debido a que éste parámetro varía mucho en el período de gestación. Muchos productos para determinar el peso vivo desde las dimensiones han sido comercializados por mucho tiempo. Estos métodos consumen menos tiempo que el pesaje manual, pero, para la obtención de las dimensiones es necesaria la inmovilización del cerdo.

El análisis de imágenes para la determinación del peso de cerdos puede ser posible determinando automáticamente las dimensiones^{[9][10][11]}. Una función de predicción puede establecer la relación entre éstas dimensiones y el peso vivo. En la búsqueda de precisión se han investigado algunos factores tales como: la raza, la edad y los métodos de alimentación en cerdos^{[14][8]}, encontrándose que estos factores deben tomarse en consideración cuando la predicción se hace desde las dimensiones. Se han evaluado varias dimensiones (área, largo cola-escápula, ancho de anca, ancho de abdomen, ancho de hombros) y aunque el coeficiente de correlación entre el peso y las dimensiones fue bastante bueno (por encima de 0.86), la dimensión con más alto coeficiente de correlación es el área (0.98)^[15]. Convirtiéndola en una candidata fuerte para establecer la función de predicción.

Este documento presenta en el primer capítulo una breve descripción de un sistema de visión artificial, detallando cada una de sus etapas. El segundo capítulo trata de la forma en que se llevo a cabo la adquisición de las imágenes y el preprocesado de estas, el tercer capítulo detalla cada uno de los métodos de segmentación implementados y su evaluación; mientras

que el cuarto capítulo describe la forma en que se obtuvieron y seleccionaron las dimensiones que tenían una fuerte correlación con el peso del animal. En el quinto capítulo se detalla la implementación y entrenamiento de la red neuronal artificial la cual es evaluada en su desempeño. Por último el sexto capítulo se presentan las conclusiones más importantes del trabajo y sugerencias para futuros trabajos.

2. SISTEMA DE VISION ARTIFICIAL

2.1. VISION ARTIFICIAL

La visión computacional (el término más similar y quizás más extendido es visión artificial) tiene como finalidad la reproducción artificial del sentido de la vista. Se trata de un objetivo excepcionalmente ambicioso y complejo, que se encuentra actualmente en una etapa muy primitiva.

Se estima que aproximadamente las tres cuartas partes de la información que maneja un ser humano es visual. Todos sabemos que existen seres vivos cuya existencia gira casi exclusivamente alrededor de la percepción visual; sin embargo, también es cierto que un ser humano desprovisto del sentido de la vista puede desarrollar completamente su capacidad mental, llevar una vida plena, aunque esto último le obligue a una dependencia estrecha de otros semejantes. No parece necesario afirmar cuán vital es la visión en los seres vivos.

Así, parece natural pensar que el objeto de dotar a las máquinas del sentido de la vista (o sucedáneos del mismo) supondrá un salto cualitativo en sus capacidades de actuación.

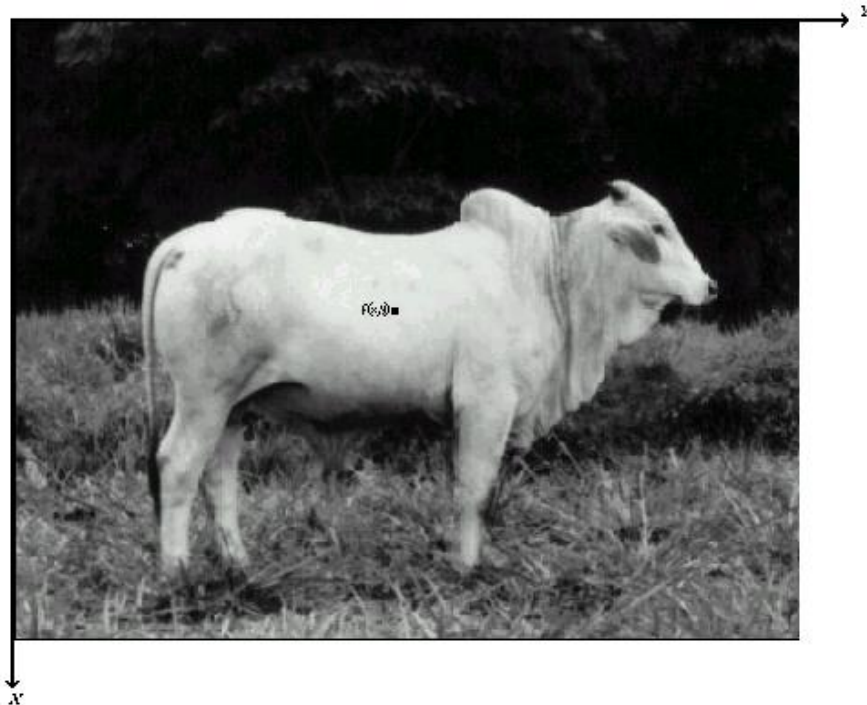
Se intenta esbozar una forma liviana y para nada exhaustiva, los pormenores de lo que hemos denominado visión artificial o visión computacional: es decir, visión basada en computador puesto que es impensable otorgar a una máquina cierto grado del sentido de la vista sin que se utilice un computador, tanto en la etapa temprana (puramente sensorial) como para la fase avanzada que demanda mayor complejidad de procesamiento.

Podría hacerse una analogía fácil entre un sistema artificial de procesamiento de imágenes y el sistema visual humano. Los ojos serían la cámara, y el computador que realiza el procesamiento, el cerebro.

La adquisición de imágenes consiste en digitalizar una señal de vídeo que contiene la información luminosa de la escena a interpretar. Naturalmente, la señal continua de vídeo deberá ser; a) muestreada y b) digitalizada para su tratamiento en computador.

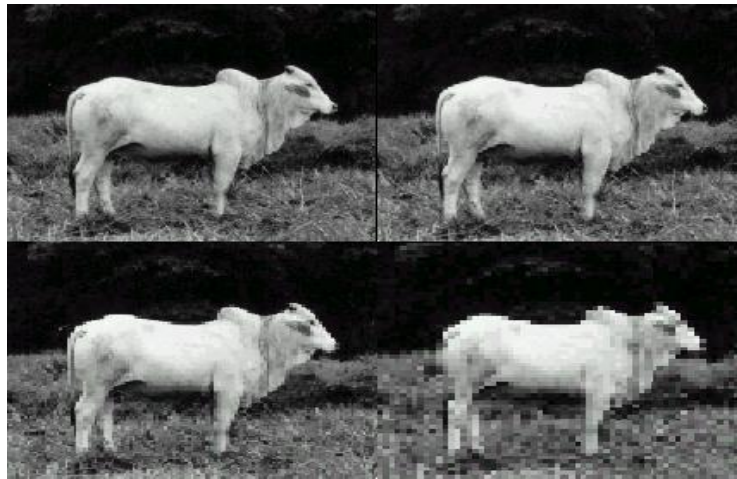
El efecto combinado del número de líneas (consustancial a la cámara de vídeo) y de la frecuencia de muestreo (es decir, el número de muestras de la señal de vídeo que se toman por cada línea) da lugar a lo que se conoce como nivel de resolución espacial de la imagen. La imagen digital (monocromática) se refiere a una función bidimensional de intensidad de luz $f(x, y)$, donde x e y representan las coordenadas espaciales y el valor de f la intensidad luminosa en un punto cualquiera (x, y) es proporcional al brillo (o nivel de gris) de la imagen de ese punto (Figura 2.1).

Figura 2.1. Sistema de coordenadas utilizado para la representación digital de imágenes

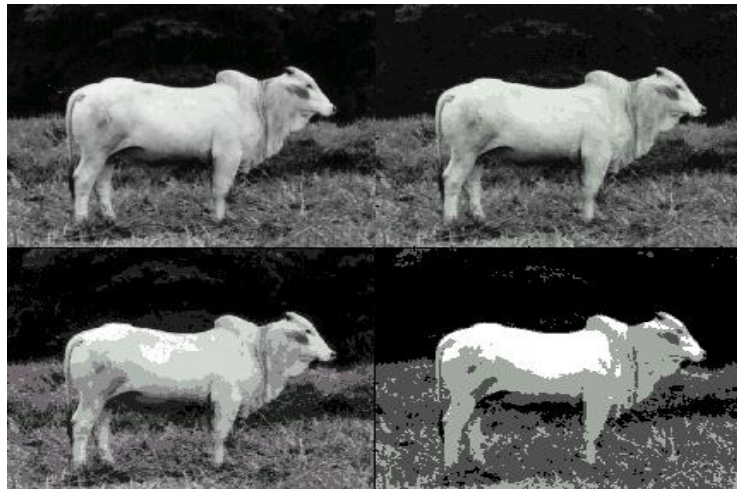


Una *imagen digital* es una función $f(x, y)$ que se ha discretizado, tanto en las coordenadas espaciales como en el brillo (nivel de gris). La imagen digital puede considerarse como una *matriz* cuyos índices de filas y columnas identifican un punto de la imagen y el valor

Figura 2.2 (a) imágenes digitales de la misma escena, con una resolución espacial reducida progresivamente: 512x512, 256x256, 128x128 y 64x64. (de izq. A der. Y de arriba abajo) (b) y cuatro imágenes digitales de la misma escena, que se han codificado en 256, 16, 8 y 4 niveles de gris, respectivamente.



(a)



(b)

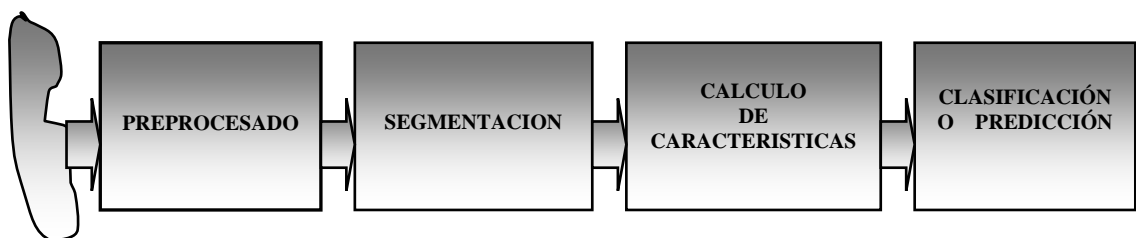
del correspondiente indica el nivel de gris en ese punto. Los elementos de esta matriz son llamados *elementos de la imagen* o más comúnmente *pixels* o *pels* contracción de los vocablos ingleses *picture elements*. Para la codificación en binario de los niveles de intensidad, se realiza por el número de bits, empleados por píxel y esto determina la resolución en niveles de grises (Figura 2.2). Así, 8 bits permiten manejar 256 niveles de gris, que es una resolución en la frontera de lo que puede discriminar el ojo humano.

Debido a cuestiones históricas, la visión artificial se ha desarrollado siempre en torno al principio del *pattern recognition*; es decir, el uso de formas. El color no es una forma, es un atributo de las formas. Sin embargo, su importancia para la discriminación o segmentación entre otras, nos permite tenerla en cuenta dentro de alguna de las etapas de la visión artificial.

2.2 PROCESADO DE IMÁGENES

El tratamiento digital de imágenes puede dividirse en algunas etapas fundamentales.

Figura 2.3. Diagrama de bloques del análisis automático de una escena



Con anterioridad al análisis automático propiamente dicho de una imagen es conveniente aplicar ciertas operaciones orientadas a mejorar su calidad.

2.2.1 Conversión de niveles de gris

Una transformación adecuada de los niveles de gris puede mejorar sensiblemente la calidad de una imagen cuando ésta es observada por un sujeto humano.

Estas técnicas se han desarrollado empleando por una parte teorías de psicología visual humana y de otros modelos de los dispositivos físicos de adquisición y presentación de imágenes. Su aplicación, hasta ahora, se ha dirigido hacia el área del procesamiento con observador humano y aunque es evidente que tiene influencia en las fases de interpretación automática, hasta ahora se está desarrollando.

2.2.2 Transformaciones geométricas

Tienen dos objetivos básicos:

- Corrección de la perspectiva o de deformaciones.
- Reconstrucción tridimensional de objetos en una escena.

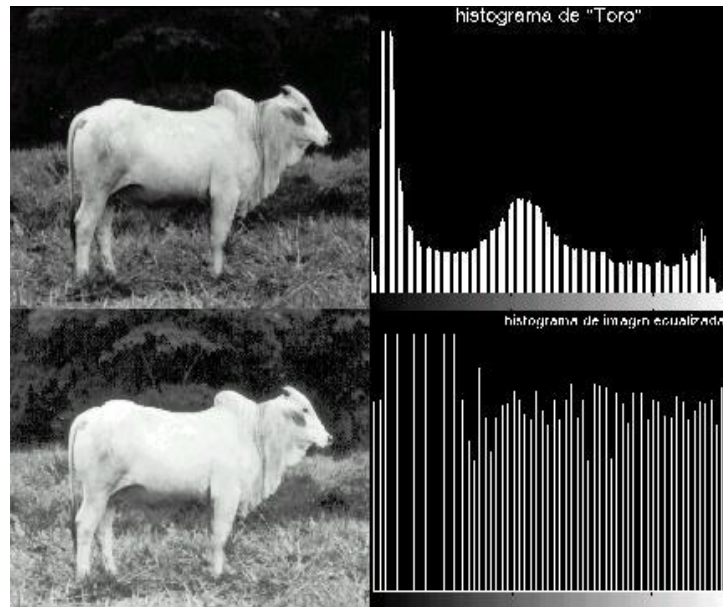
2.2.3 Transformación del histograma

En la figura 2.4. se visualiza una imagen y su correspondiente histograma. En el eje de abscisas se representan los niveles de gris y en el de ordenadas, el número de píxeles de cada nivel.

En ciertas escenas, como las que suelen aparecer en entornos industriales, el histograma puede ser suficiente para separar objetos dentro de una imagen, con lo que se facilita considerablemente la etapa de interpretación.

Sin embargo, no siempre es factible aprovechar el histograma. En muchas aplicaciones, su manipulación sólo permite aumentar/disminuir el contraste de la imagen y el margen dinámico de los niveles de grises. Sus efectos son evidentes para el observador humano.

Figura 2.4. (a) Imagen digital y su correspondiente histograma. (b) imagen ecualizada y su histograma



2.2.4 Filtrado espacial y frecuencial

No es fácil decidir si estas operaciones, absolutamente vitales en un sistema de visión por computador; pertenecen o no a la etapa de simple preprocesamiento, dada la tremenda modificación que puede suponer respecto a la imagen original. Para evitar esta ambigüedad, consideremos que el filtrado es procesamiento previo cuando la transformación operada sobre la imagen sea pequeña.

El filtrado espacial actúa sobre los niveles de gris de los píxeles de la imagen. El filtrado frecuencial se basa en trabajar en el espacio de Fourier, cuya transformación se da por la Transformada de Fourier Discreta. a la cual se le aplica un filtro bidimensional.

2.2.5 Segmentación

La segmentación puede definirse como el proceso orientado a particionar la imagen digital en zonas disyuntas con significado propio. Este significado dependerá de la imagen, así como de los objetivos que se persigan con su interpretación.

En la práctica suelen distinguirse dos segmentaciones: de regiones o zonas y de objetos. La operación es común, lo que varía es aquello que se desea segmentar. En un caso, zonas; en el otro, ya dentro de una región concreta, objetos individuales.

En general, las operaciones de segmentación de una escena, así como el número de clases de regiones y objetos a distinguir, dependerá íntimamente de la propia escena y, tal como se menciono antes, de la información que se busque dentro de la imagen.

Las técnicas de segmentación, ya estén orientadas a objetos o agrupaciones de objetos, se pueden dividir en tres grupos:

- Aplicación de umbrales a nivel de gris.
- Agrupación por rasgos comunes.
- Extracción de bordes.

La aplicación de umbrales se basa en el empleo del histograma. Su exposición detallada nos llevaría demasiado lejos. Básicamente, consiste en distinguir objetos individuales dentro de una escena a partir de sus diferentes niveles de gris.

La segmentación de imágenes mediante la agrupación (clustering) se inspira en las técnicas de reconocimiento de formas.

La extracción de bordes es el método más empleado en la segmentación de imágenes (salvo en robótica, donde con frecuencia es suficiente la aplicación de umbrales).

La efectividad, desde el punto de vista del reconocimiento de objetos, de una operación cuyo resultado sea la obtención de los bordes de un objeto, cual es la segmentación, es evidente, puesto que la mayor cantidad de información se encuentra en los bordes. De hecho, la mayoría de los objetos pueden reconocerse a partir de su forma, que está fijada por sus bordes.

Los filtrados espaciales y frecuenciales, que se comentaron en el procesamiento previo, son precisamente operaciones de extracción de bordes, aunque lo hagan con suavidad. De manera análoga, los extractores de bordes empleados en la fase de segmentación pueden clasificarse en espaciales y frecuenciales.

Tras la segmentación de la escena, aparecen dos nuevas fases: extracción de los rasgos característicos de los objetos ya diferenciados y su identificación propiamente dicha.

2.2.6 Extracción de rasgos característicos.

Dado cierto objeto físico dentro de una imagen, y a partir de su contorno (resultado de la segmentación), puede obtenerse de inmediato un conjunto de parámetros característicos, como perímetro, grado de circularidad, momentos de orden n , etc.

Estos rasgos hacen posible distinguir objetos mediante el uso de un número reducido de propiedades.

Los rasgos de un objeto forman un vector $X = (X_1 X_2 X_3 \dots X_n)^T$, que particularizado en unos valores numéricos, los del objeto a clasificar, determina un punto en el espacio n-dimensional de las características.

Las propiedades deseables en los rasgos son, principalmente: número no elevado, estar incorrelacionados entre sí al máximo, discriminación abrupta de los objetos, etc., propiedades todas ellas orientadas a realizar clasificación automática rápida y fiable.

2.2.7 Clasificación de los objetos

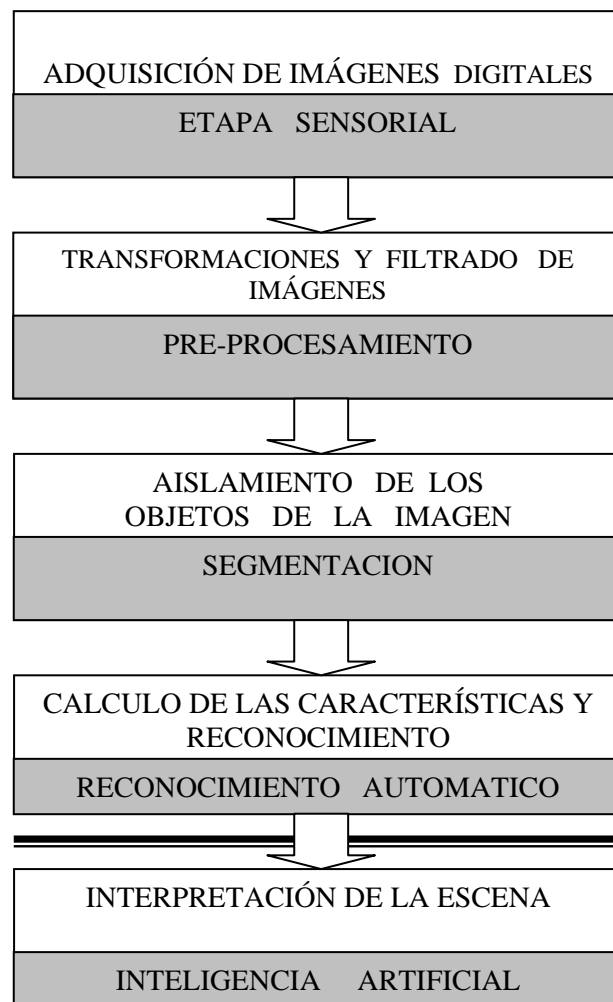
Aparte de haber sido segmentado en la escena y habersele extraído cierto número de características propias, para reconocer un objeto es imprescindible haberlo catalogado previamente; es decir, haberlo definido, dado un nombre.

Esto significa que en un sistema de visión artificial existe un conocimiento almacenado (una memoria), que ha sido elaborado por un equipo de especialistas. Este conocimiento previo es muy variado y depende de la aplicación concreta. Puede ser muy sencillo, cual es el caso de la mayoría de las aplicaciones en robótica, en donde prácticamente sólo hay que distinguir unas pocas figuras simples y, en particular, su orientación en un plano. Y puede ser un conocimiento muy complejo, cual es el caso de la detección automática de tumores (aunque en diagnóstico médico siempre existe un observador humano que es el responsable último).

Una vez adquirida la imagen, preprocesada, segmentada la correspondiente escena y aislados y reconocidos los objetos en su seno, se entraría en una fase superior que podría calificarse de cognitiva o de procesamiento del conocimiento, más cercana a lo que ha venido a denominarse *inteligencia artificial*. Aunque, todavía no se han resuelto importantes problemas en las fases previas, la segmentación en particular, para imágenes complejas. De hecho, sólo se han encontrado soluciones generales para imágenes bidimensionales

sencillas. No hablemos de temas como la percepción de la textura de los objetos, de la visión tridimensional, de las escenas en movimiento, del fenómeno de la estereopsis (visión en profundidad), etc., temas abiertos que nos recuerdan cuán lejos se está de una teoría computacional (implementable en computador) de la visión. Por todo esto, es presuntuoso hablar de percepción visual cognitiva, para escenas que no sean sencillas y bidimensionales.

Figura 2.5 Diagrama de bloques de las etapas en que puede organizarse un SVA.



2.3 INTELIGENCIA ARTIFICIAL Y REDES NEURONALES

Aunque dentro de la *inteligencia artificial* (IA) se puede situar e interconectar las *redes neuronales* (RN), existen diferencias entre las redes neuronales y lo que se entiende por inteligencia artificial, que va de la forma de representación del conocimiento (localizada, en la inteligencia artificial y distribuida, en redes neuronales), mecanismo de inferencia (lógica en IA), aprendizaje (añadir nuevos átomos, reglas, etc. en IA y en cambio en RN por refuerzo de las conexiones) y el paralelismo (características de sistemas conexionistas o distribuidos como las RN).

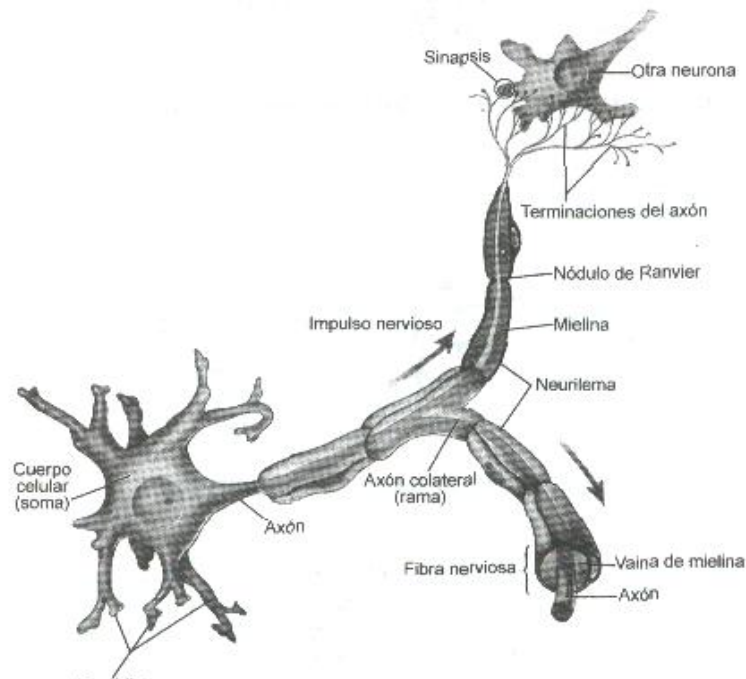
2.3.1 Redes neuronales.

2.3.1.1 Neurona biológica.

Una neurona biológica es una célula especializada en procesar información. Está compuesta por el cuerpo de la célula (soma) y dos tipos de ramificaciones: el axón y las dendritas. La neurona recibe señales (impulsos) de otras neuronas a través de sus dendritas y transmite señales generadas por el cuerpo de la célula a través del axón.

El córtex cerebral humano contiene alrededor 10^{11} neuronas (aproximadamente el mismo número de estrellas que hay en la Vía Láctea). Cada neurona está conectada en forma masiva a un número variable de entre 1000 y 10000 neuronas. Por tanto, el cerebro contiene de 10^{14} a 10^{15} interconexiones.

Figura 2.6. Neurona Biológica



Las neuronas se comunican entre sí mediante trenes de pulsos de corta duración (del orden de milisegundos). El mensaje, está modulado en la frecuencia de transmisión de pulsos. Esta frecuencia varía alrededor de los 100 Hertz. Por tanto, es más de un millón de veces inferior a la velocidad de conmutación de los circuitos electrónicos típicos de nuestros días. Este hecho se debe al paralelismo masivo que realiza el cerebro, que está extremadamente lejos de poderse implementar en circuitos electrónicos.

2.3.1.2 Modelo de neurona artificial.

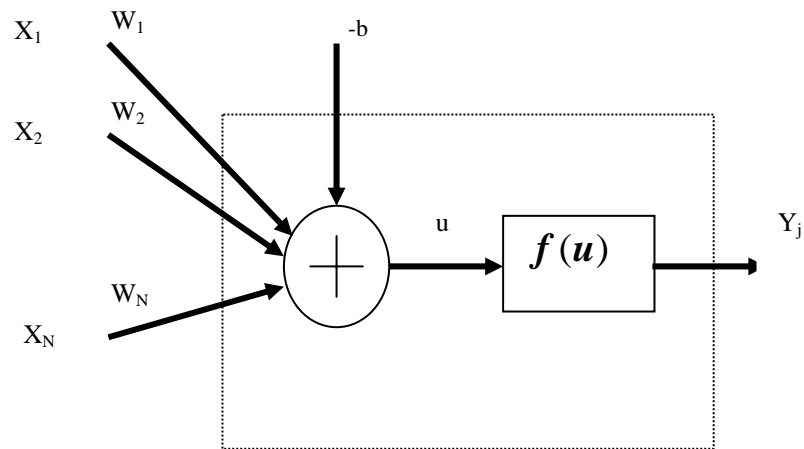
Las redes neuronales artificiales (RNA) pretenden emular la red neuronal biológica y se organizan interconectando neuronas. Para ello se supone un modelo matemático simplificado de la neurona biológica.

Este modelo es una generalización del propuesto por McCulloch y Pitts en 1943.

$$y_j = f\left(\sum_{i=1}^N w_{ji} x_i - b\right) \quad (2.1)$$

La siguiente figura representa el modelo de neurona artificial.

Figura 2.7. Modelo de neurona artificial



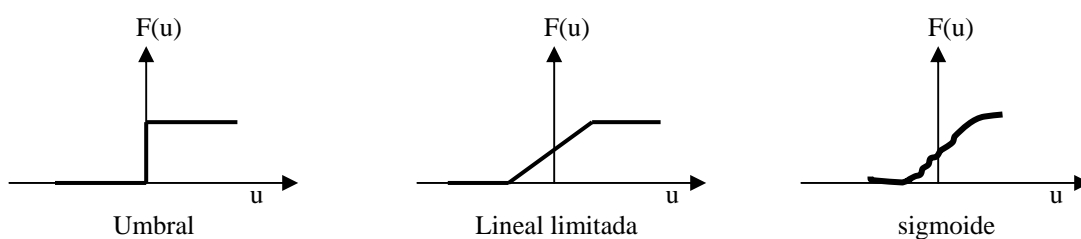
La neurona recibe del exterior un umbral b y N entradas a las que asocia un conjunto de pesos w_i ($i = 1, \dots, N$). Aplicando el producto de los pesos por la s entrada respectivas más el umbral (o desplazamiento) a una función de activación $f(u)$, se obtiene la salida y . La figura 2.8 muestra las funciones de activación más comunes.

$u(\vec{x}, \vec{w})$ recibe el nombre de función base, donde los vectores \vec{x}, \vec{w} son :

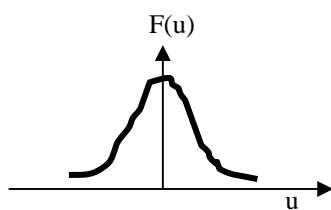
$$\vec{x} = \{x_1, x_2, \dots, x_N\} \quad \vec{w} = \{w_1, w_2, \dots, w_N\}$$

Cuando la combinación de las entradas es de tipo lineal, $u(\vec{x}, \vec{w})$ es una función de base lineal, pero son posibles otras combinaciones, entre las que se destacan las funciones de base radial, consistentes en realizar una operación no lineal sobre las entradas, con simetría radial.

Figura 2.8. Funciones de activación.

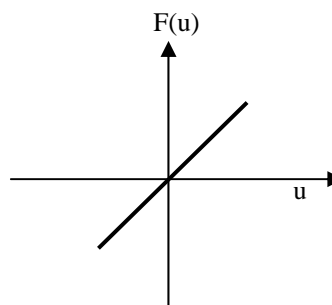


$$f(u) = \begin{cases} 0 & u < 0 \\ 1 & u \geq 0 \end{cases} \quad f(u) = \begin{cases} 0 & u \leq -\frac{1}{2} \\ u & -\frac{1}{2} < u < \frac{1}{2} \\ 1 & u \geq \frac{1}{2} \end{cases} \quad f(u) = \frac{1}{1 + e^{-\beta u}}$$



Gausiana

$$f(u) = e^{-\frac{u^2}{\sigma^2}}$$



Lineal

$$f(u) = u$$

2.3.1.3 Arquitectura de una red neuronal artificial

Las neuronas de una determinada red pertenecerán a una de las siguientes clases o conocidas también como capas:

Neuronas de entrada: Reciben las señales del exterior. Estas integran la primera capa de la red, aunque algunos autores no la consideran una capa debido a que esta no procesa solo distribuye los datos a la siguiente capa.

Neuronas ocultas: Producen resultados intermedios y no tiene comunicación con el exterior.

Neurona de salida: Su salida es observable desde el exterior. Constituye la última capa de la red.

Por otra parte, dependiendo de la forma en que se estructuren las neuronas y sus conexiones, aparecen dos grandes tipos de redes:

Redes progresivas o de conexión hacia adelante (feedforward)

Redes realimentadas o de conexión hacia atrás (feedforward/feedback).

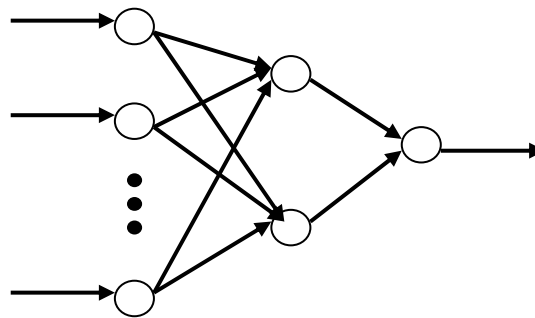
Es necesario aclarar que existen versiones evolucionadas con conexiones laterales y autorrecurrentes

2.3.1.3.1 Redes progresivas

Las conexiones de sus neuronas y, por lo tanto, el flujo de información, son siempre hacia adelante, de la entrada a la salida. No presentan memoria, por tanto son estáticas.

Las arquitecturas progresivas de mayor importancia son el perceptrón multicapa (MLP) y las redes de base radial (RBF).

Figura 2.9. red progresiva



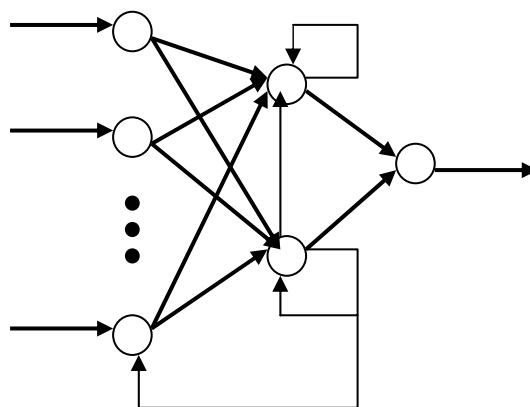
2.3.1.3.2 Redes realimentadas

A diferencia de las redes progresivas, presentan como mínimo un camino de realimentación, ya sea entre neuronas de una misma capa (conexiones laterales o conexiones hacia atrás).

Las arquitecturas más destacadas son:

- Redes competitivas
- Mapas autoorganizativos de Kohonen.
- Red Hopfield
- Red de Elman
- Modelos ART.

Figura 2. 10. Red realimentada



2.3.1.4 Aprendizaje

Una vez fijada la arquitectura y el número de neuronas de la red, hay que calcular el valor de los pesos para que realice la función deseada. Usualmente, la red debe aprender los pesos de las conexiones a partir de un conjunto de patrones de entrenamiento. Por tanto, no se calculan los pesos de forma analítica sino que es necesario un proceso de aprendizaje automático a partir de una serie de ejemplos.

En vez de una serie de reglas explícitas, es la red neuronal la que aprende las reglas explícitas que relacionan la salida con la entrada, a partir de una colección de ejemplos representativos. Esta es una de las mayores ventajas de las redes neuronales sobre los sistemas tradicionales.

Para diseñar un proceso de aprendizaje es necesario disponer de:

- Un conjunto de información disponible para el aprendizaje.

- Un proceso para actualizar los pesos, o algoritmo de aprendizaje.

Existen tres métodos básicos de aprendizajes:

Aprendizaje supervisado: La red conoce la salida correcta para cada patrón de entrada. Por tanto, Se ajustan los pesos de forma que las salidas de la red sean lo más parecidas posibles a la salida correcta.

Aprendizaje por refuerzo: Es una variante del método anterior, en el cual la red únicamente conoce una crítica sobre la exactitud de las salidas, pero no la respuesta correcta.

Utiliza un mecanismo de penalización y recompensa, mediante el cual la red resulta recompensada por una decisión correcta y castigada por una equivocada. De esta forma, el sistema tiende a actuar de la forma favorecida y debilita la tendencia a proporcionar la respuesta penalizada.

Aprendizaje autoorganizativo: No requiere el conocimiento de la salida correcta y debe aprender la estructura implícita en los datos. Generalmente requiere un número mayor de datos de entrenamiento.

También son posibles aprendizajes híbridos, combinando aprendizaje supervisado y no supervisado. En este caso, parte de los pesos se calcula mediante un método y la otra parte mediante el otro método.

En el aprendizaje existen tres cuestiones importantes:

Capacidad de la red para implementar funciones y aprender patrones: dependerá de su estructura, tamaño, etc.

Complejidad de los datos. Determina el número de patrones de entrenamiento necesario para entrenar la red de forma que se garantice una buena capacidad de generalización (capacidad

de procesar correctamente datos no utilizados en el aprendizaje). Si el número de patrones es pequeño, la red se comportará bien con aquellos patrones utilizados en el entrenamiento, pero presentará un comportamiento pobre con aquellos datos no utilizados.

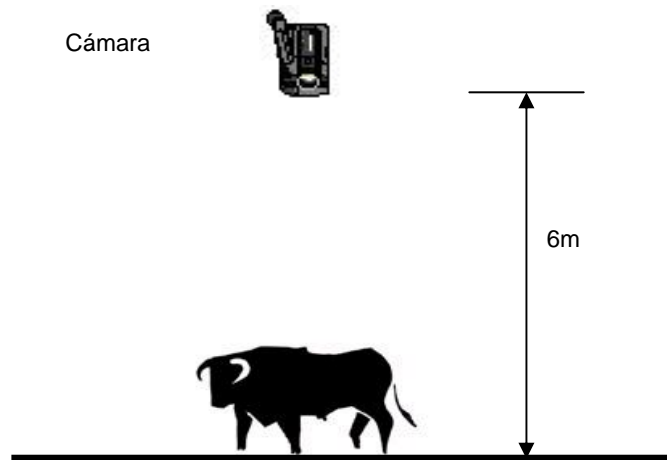
Complejidad computacional: Condiciona el tiempo necesario para que el algoritmo de aprendizaje proporcione una solución ajustada.

El conjunto de patrones utilizados en el proceso de aprendizaje recibe el nombre de *secuencia de entrenamiento*. Cada vez que se ha presentado una secuencia de entrenamiento a la red, se dice que se ha transcurrido una época. Normalmente los algoritmos que consumen poco tiempo en cada época requieren muchas épocas y viceversa.

3. ADQUISICIÓN

En un sistema de visión artificial (SVA) la adquisición de las imágenes digitales constituye una parte fundamental e ineludible. Además, su implementación define en gran parte la problemática inicial del tratamiento de la imagen. Esta puede facilitar o por el contrario dificultar las etapas posteriores de un SVA. Por eso, se hace necesario definir los elementos para hacerlo; tales como el tipo de dispositivo de adquisición, el rango de sensibilidad espectral del sensor y el control del ambiente de trabajo. Es muy posible en la mayoría de los casos que algunos de estos elementos no ofrezcan la posibilidad de ajustarse a las condiciones requeridas, en este caso las etapas posteriores del sistema deben resolverlo.

Figura 3.1. Montaje



Para este caso se utilizó una cámara de video convencional a color (f:3.6mm-72mm) para realizar un registro analógico de los datos en continuo; se adecuó el ambiente de trabajo de tal forma que el fondo de la imagen capturada tuviese un mismo color (verde) de poca posibilidad

de presencia en el objeto. La cámara se ubicó a una altura de 6m de tal manera que permite registrar las imágenes con una vista superior del objeto (figura 3.1).

3.1 ADQUISICIÓN DE DATOS.

En la toma de datos se seleccionó una población adecuada. Se tomaron 100 imágenes de especímenes de raza Cebú con aproximadamente 98% de pureza, con sexo y edades diferentes (ver Anexo 1), donde estos tienen un mismo régimen alimenticio (animales de exposición).

Figura 3.2. Algunas imágenes obtenidas.



Las imágenes fueron adquiridas mediante una cámara de vídeo a color convencional y posteriormente digitalizadas utilizando una tarjeta (MATROX meteoro) y los valores de peso

(en realidad valores de masa en kilogramos) por medio de una báscula en el mismo instante en que estas fueron tomadas. Posteriormente el video se digitalizó (la digitalización no fue realizada en campo, debido a la limitación del equipo) en imágenes de 480x640x3 píxeles. Las imágenes obtenidas se organizaron de dos maneras: La primera teniendo en cuenta una imagen referencia (sin el espécimen) y una imagen objeto (con el espécimen), la segunda tomando secuencias de video del espécimen en momento de pasar por campo de visión de la cámara. La figura 3.2 muestra algunas imágenes adquiridas mediante el montaje descrito.

3.2 PREPROCESAMIENTO.

Otra etapa importante es la de mejorar la calidad de la imagen de tal forma que aumente las posibilidades de obtener mejores resultados en etapas posteriores. En la mayoría de los casos esas mejoras atacan problemas como bajo contraste, eliminación de ruido entre otras.

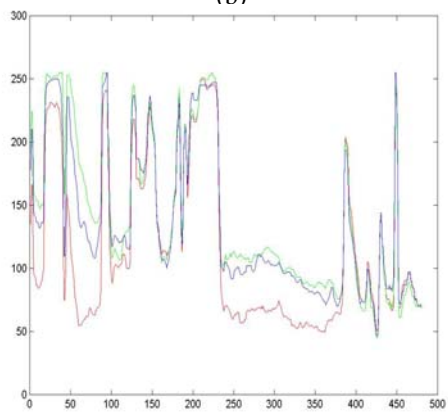
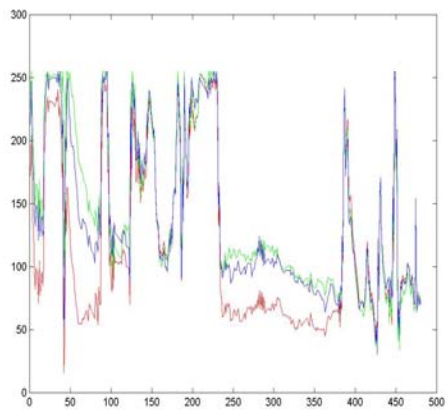
En el proceso de digitalización de video se introduce ruido de tipo electrónico y también errores en el entrelazado de la imagen, ambas originadas por la tarjeta digitalizadora. Este ruido se atenúa utilizando un filtro mediana en cada plano de color.

El filtro mediana es no lineal y consiste en el ordenamiento numérico del valor del píxel y de su entorno asignando el valor central al píxel.

$$f_f(x, y)_{RGB} = \mathit{filmed} \{ f(x, y)_{RGB} \} \quad (3.1)$$

Este filtro tiene la característica de eliminar el ruido tipo impulsional o de forma de pico y mantener las transiciones bruscas propias del objeto. Los efectos de este filtro para 9 entradas en cada plano de color se muestran en la figura 3.3. En ella se aprecia los perfiles de cada plano de color antes y después del filtro; es notorio el suavizamiento o reducción de los picos propios del ruido en los perfiles después de aplicar el filtro en cada plano de color.

**Figura 3.3 (a) Imagen sin filtro. (b) perfil de imagen sin filtro.
(c) Imagen con filtro. (d) perfil de imagen con filtro.**



4. SEGMENTACIÓN

En el análisis de imágenes digitales es importante determinar, cual o cuales partes de la imagen son de interés para su posterior separación, esto se conoce como segmentación, aunque esta se define comúnmente como la división de la imagen en partes disyuntas y con significado. En nuestro caso es mejor verla como el aislamiento de una zona o región específica. Dentro de las diferentes etapas de un SVA esta es definitivamente la más crítica, debido a que el éxito del análisis es determinado por la calidad de la segmentación. En la mayoría de los casos no se logra una solución requerida de tal forma que se hace necesario que esta tenga una robustez bastante alta.

En el momento del diseño del sistema de visión artificial se pone especial cuidado en la posibilidad de controlar el entorno de trabajo con el fin de facilitar la segmentación, para ello se utiliza con frecuencia fondos e iluminaciones que eleven el contraste. Sin embargo esto no siempre es posible, por tanto se hace necesario la utilización de sensores que permitan realzar los objetos de estudio tales como infrarrojos o de color. Cuando no se tiene acceso a ninguna de las opciones anteriores se utiliza normalmente análisis de los valores de discontinuidad o similitud de los valores de los píxeles. La primera la técnica consiste en dividir la imagen de acuerdo a los cambios bruscos que esta sufre. Dentro las principales líneas de esta categoría se encuentran: la detección de puntos, detección de líneas y bordes de la imagen. Por otro lado la segunda se basa principalmente en la homogeneidad de los píxeles pertenecientes a regiones de la imagen, a esta pertenecen: la umbralización, crecimiento de regiones y división y fusión de regiones. Sin embargo este concepto se puede aplicar a escenas estáticas como a dinámicas; en esta última el movimiento del objeto puede ser de mucha importancia a la hora de segmentar.

En nuestro caso las imágenes obtenidas ofrecen varias posibilidades, debido a que la captura de los datos se hizo en continuo. Una de ellas es la de tener información de imágenes de secuencia y la otra tener acceso a imágenes antes y después de la captura del objeto en la imagen a los cuales llamamos imágenes referencia.

A continuación mostramos una descripción detallada de los diferentes métodos utilizados para la segmentación del objeto y una aproximación de cuantización del error para cada método de segmentación.

4.1 SEGMENTACIÓN POR DIFERENCIA DE IMÁGENES (OBJETO-REFERENCIA)

Este método consiste en tomar dos imágenes una antes o después de pasar el objeto (imagen referencia) y otra justamente cuando este el objeto (imagen objeto).

Figura 4.1 . Algunas imágenes referencia y objeto.

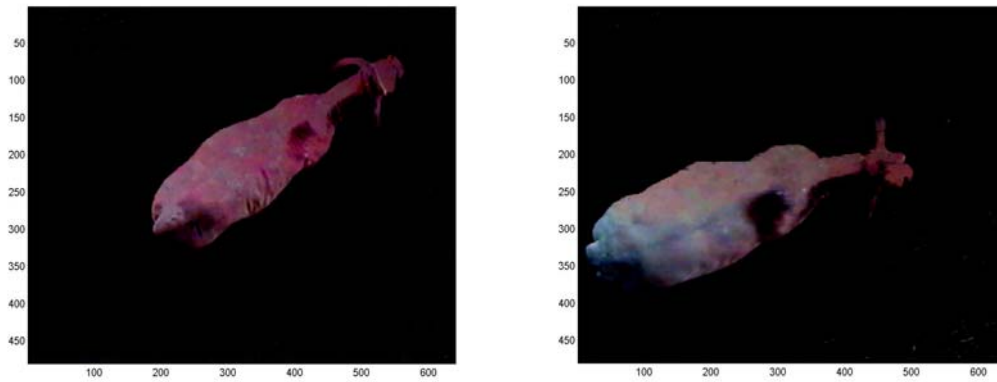


El propósito de aplicar este método es el de eliminar regiones de las imágenes (referencia – objeto) que guardan alta similitud y mantener aquellas que han sufrido cambios considerables, es decir donde está presente el objeto. Esto se puede lograr con una diferencia entre las imágenes.

$$d(x, y)_{RGB} = f(x, y)_{RGB}^{obj} - f(x, y)_{RGB}^{ref} \quad (4.1),$$

donde $f(x, y)_{RGB}^{obj}$ y $f(x, y)_{RGB}^{ref}$ corresponde a la imagen objeto y referencia respectivamente.

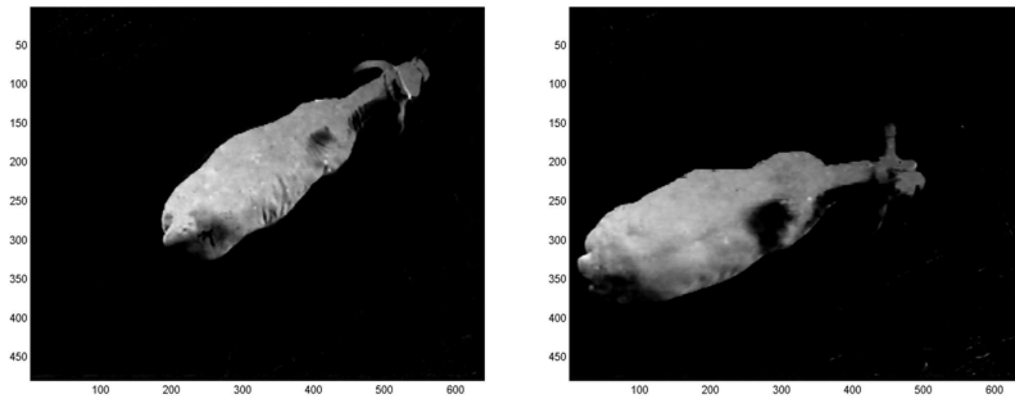
Figura 4.2 Imágenes diferencia



Como las imágenes de diferencias son a color (en el espacio RGB), es decir cada color esta representado en un espacio de coordenadas cartesianas; en donde en los ejes es posible representar todos los valores que pueda tomar cada color. En estas imágenes cada plano de color puede tomar 256 valores, para imágenes de 24 bits. Es posible obtener imágenes cuyo valor de píxel corresponda a la distancia euclidiana entre los puntos representativos de color entre la imagen referencia y objeto.

$$d(x, y)_{euc} = \sqrt{(d(x, y)_R)^2 + (d(x, y)_G)^2 + (d(x, y)_B)^2} \quad (4.2)$$

Figura 4.3 Imágenes de distancia euclidiana en el espacio



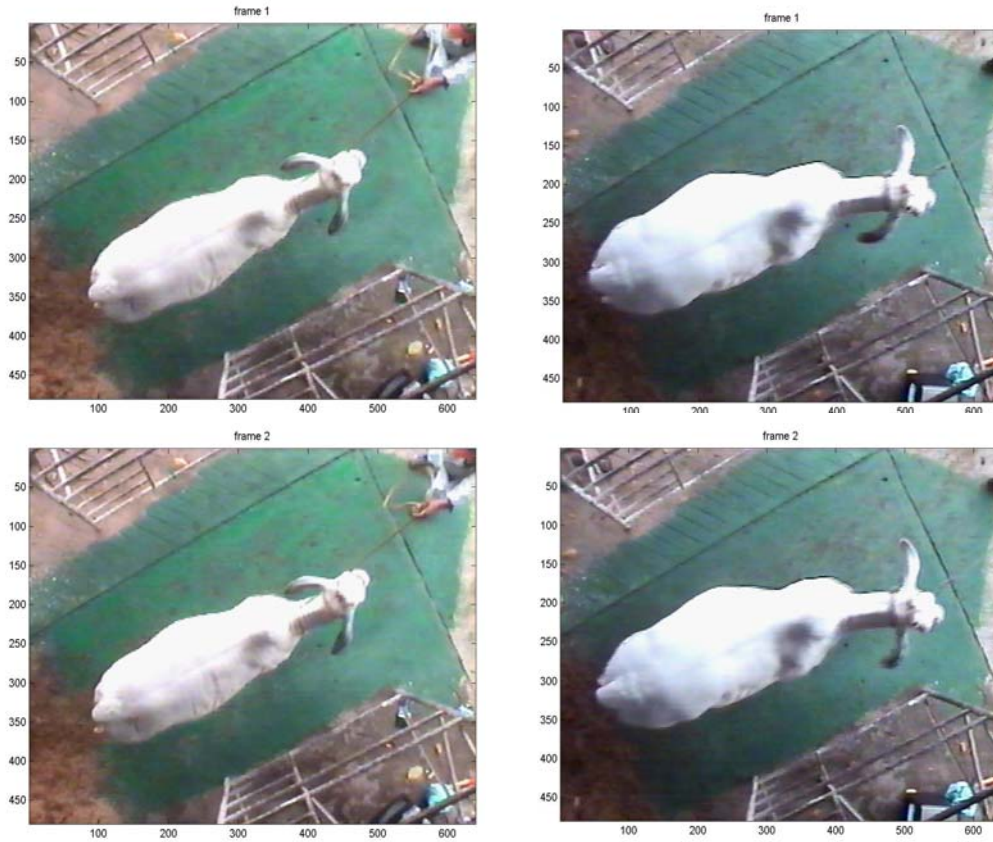
Este criterio brinda la posibilidad de utilizar un método sencillo para discernir cuales pixeles pertenecen al objeto, debido a que los puntos que pertenecen a regiones de alta similitud su distancia euclidiana serán casi cero y por el contrario las que no, tendrán una distancia euclidiana de valor mayor que cero.

4.2 SEGMENTACIÓN POR DIFERENCIA DE IMÁGENES DE SECUENCIA.

Este método consiste en obtener dos imágenes consecutivas de la secuencia de video mientras el objeto se mantiene dentro del cuadro. Inicialmente se digitaliza el video con ayuda de una tarjeta MATROX o en su defecto una tarjeta TV-VGA y el software *Virtual Dub* que permite obtener de 10 a 15 cuadros de videos (frames) por segundo (fps). De los cuales se toman dos al azar para iniciar el proceso (figura 4.4).

El propósito de este método es el de eliminar regiones que se mantienen constante en el tiempo y rescatar aquellas que han sufrido fuertes variaciones debido al movimiento del objeto.

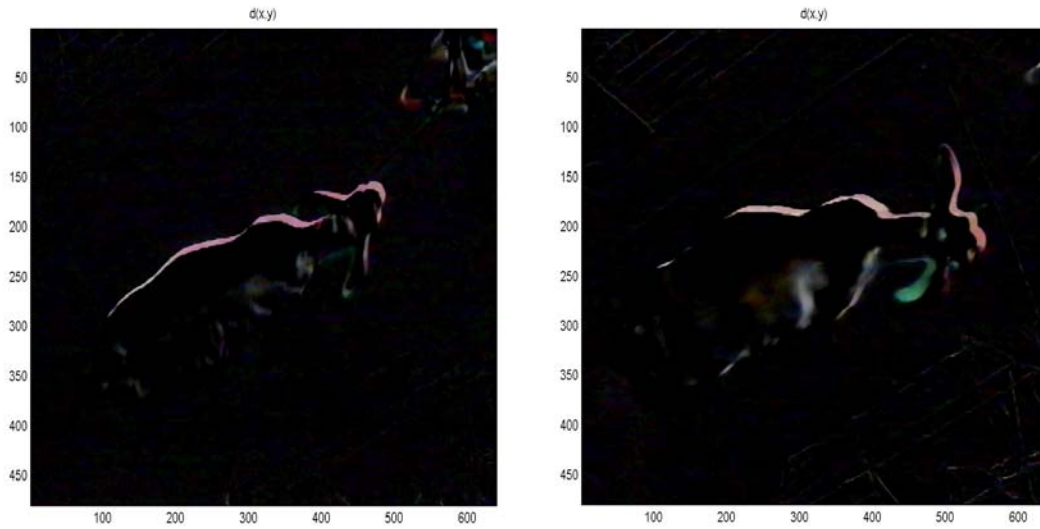
Figura 4.4 Imágenes de secuencia.



$$d(x, y, \Delta t)_{RGB} = f(x, y, t_2)_{RGB} - f(x, y, t_1)_{RGB} \quad (4.4)$$

donde $f(x, y, t_2)_{RGB}$ y $f(x, y, t_1)_{RGB}$ corresponden a frames en tiempos consecutivos, y $d(x, y, \Delta t)_{RGB}$ la imagen diferencia en ese intervalo de tiempo.

Figura 4.5. Diferencia de imágenes de secuencia.



La imagen diferencia corresponde a una imagen de borde, que será de diferentes espesores dependiendo si el objeto tiene alta o baja velocidad, teniéndose en cuenta que se toman los frames a un mismo Δt (figura 4.5).

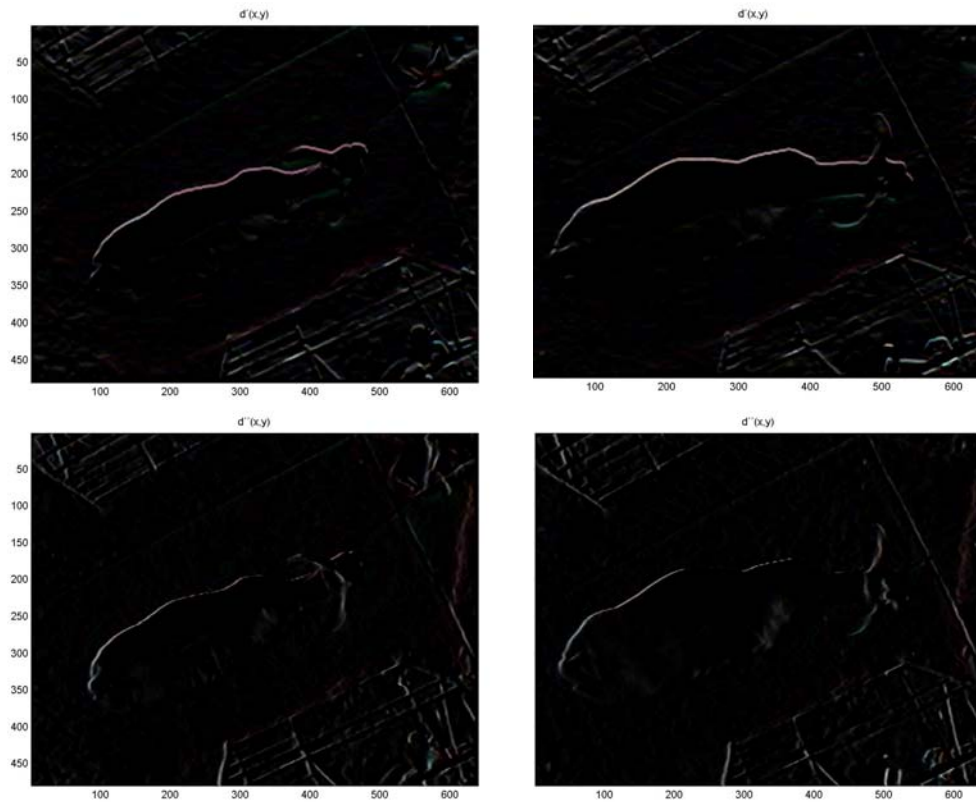
Sin embargo se observa que estas diferencias que se esperaban marcadas en todo el contorno del objeto no lo fueron así, debido a la similitud de ciertas regiones, ya sea por igualdad en el color o por baja reflectividad. De tal forma que es necesario implementar un procedimiento de contención que permita suplir esta deficiencia. Este mecanismo consiste en desplazar la imagen correspondiente al frame1 tanto verticalmente como horizontalmente y posteriormente obtener la diferencia entre ellas. De esta forma aumentamos la probabilidad de obtener más información del borde del animal o por lo menos de regiones donde esta diferencia no es notoria en la primera diferencia (figura 4.6).

$$d'(x, y)_{RGB} = f(x + \Delta x, y)_{RGB} - f(x, y)_{RGB} \quad (4.5)$$

y

$$d''(x, y)_{RGB} = f(x, y + \Delta y)_{RGB} - f(x, y)_{RGB} \quad (4.6)$$

Figura 4.6 diferencia de imágenes desplazadas. Sup. Δy , inf. Δx



Como las imágenes de diferencia tienen información en los tres planos de color el rojo, verde y azul (R,G,B), de tal manera que cada plano de color de la imagen diferencia puede contribuir a resaltar los bordes de la imagen, por tanto se suman sus absolutos de cada plano. Este proceso se aplica a los distintos métodos de diferencia explicados anteriormente.

$$d(x, y)_{\Sigma} = \sum_{R,G,B} |d(x, y)| \quad (4.7),$$

$$d'(x, y)_{\Sigma} = \sum_{R,G,B} |d'(x, y)| \quad (4.8)$$

$$d''(x, y)_{\Sigma} = \sum_{R,G,B} |d''(x, y)| \quad (4.9)$$

donde $d(x, y)_\Sigma$, $d'(x, y)_\Sigma$ y $d''(x, y)_\Sigma$ son imágenes monocromáticas que corresponden a la acumulación de los planos absolutos de color de las imágenes diferencia.

4.3 UMBRALIZACIÓN

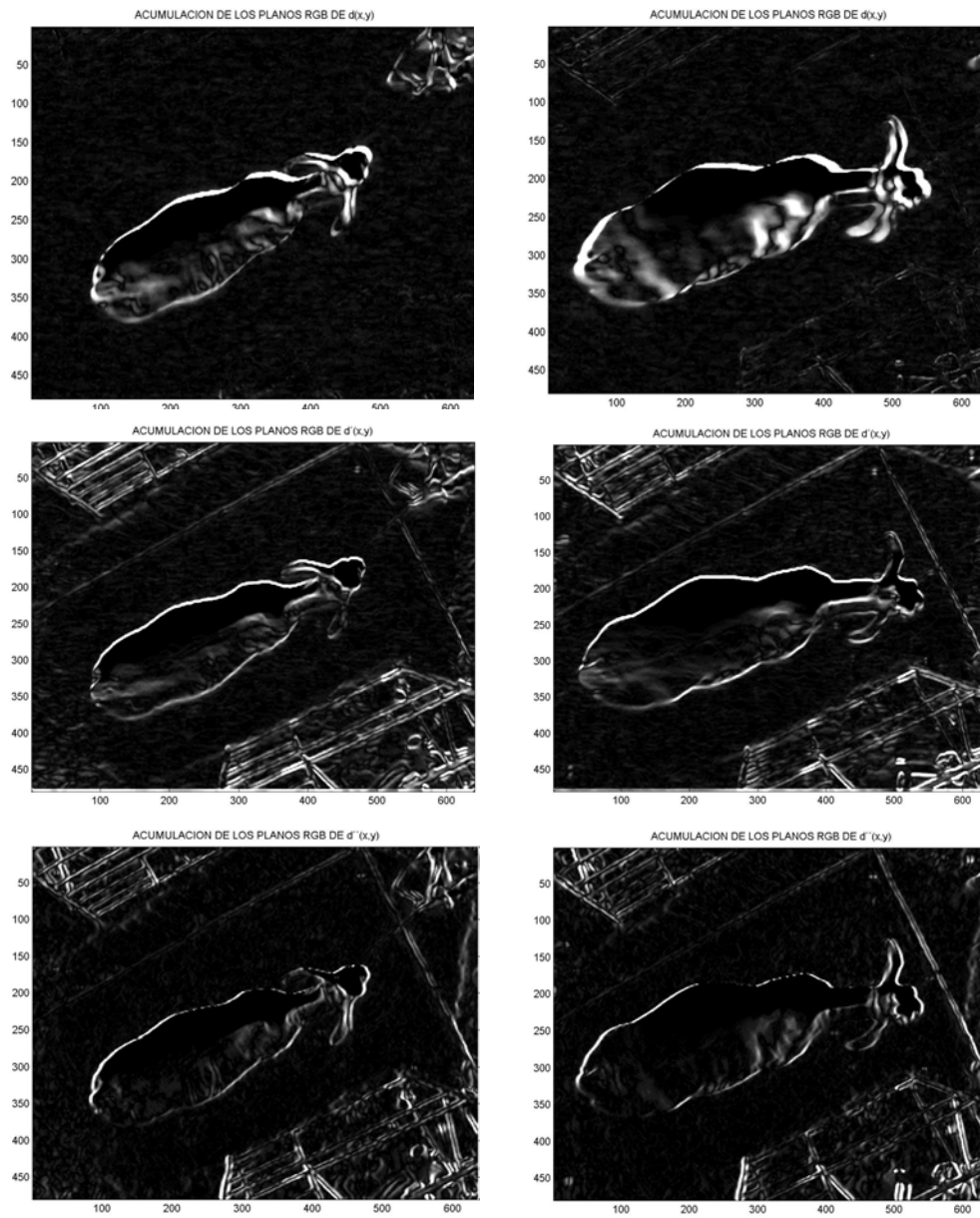
Uno de los métodos más importantes y más utilizados de la segmentación es la umbralización. Consideremos el histograma de niveles de gris que corresponda a imagen monocromática $f(x, y)$, compuesta de objetos luminosos sobre un fondo oscuro, de tal manera que los píxeles del objeto y el fondo tienen los niveles de gris agrupados en dos modos dominantes o bimodal. Una forma de extraer los objetos del fondo es elegir un umbral T que separe dichos nodos. Entonces cualquier punto (x, y) para el que $f(x, y) > T$ se denomina punto del objeto ; en el caso contrario , se denomina punto del fondo.

Basándose en lo anteriormente expuesto, la umbralización se puede definir como una operación que implica realizar comprobaciones frente a una función T de la forma:

$$T = T[x, y, p(x, y), f(x, y)] \quad (4.10)$$

donde $f(x, y)$ es el nivel de gris de coordenadas, (x, y) y $p(x, y)$ representa una propiedad local de este punto (por ejemplo la media del nivel de gris de una vecindad centrada en (x, y)). Una imagen umbralizada $g(x, y)$ se define como:

Figura 4.7. Acumulación de planos de color para las tres diferencias



$$g(x, y) = \begin{cases} 1 & \text{si } f(x, y) > T \\ 0 & \text{si } f(x, y) \leq T \end{cases} \quad (4.11)$$

De este modo los píxeles marcados con 1 corresponden a objetos, mientras los píxeles marcados con 0 corresponden al fondo. Cuando T depende solamente de $f(x,y)$, el umbral se denomina *global*. Si T depende de $f(x,y)$ y $p(x,y)$, el umbral se denomina *local*. Si, además, T depende de las coordenadas espaciales x e y , el umbral se denomina *dinámico*.

En la sección anterior se ha logrado aislar considerablemente el objeto. Ahora se busca obtener una imagen binaria que permita realizar operaciones morfológicas para su mejoramiento y además la obtención de parámetros con mayor facilidad.

Para tal caso se implementó una binarización basada en umbral global y local. En donde el umbral global se obtuvo de tres formas: una manual y dos de manera automática. Y el umbral local se logró en forma automática. A continuación describiremos más detalladamente cada uno de estos procedimientos.

4.3.1 Umbralización Global

Para obtener una imagen binaria debemos encontrar un valor de umbral, que este caso corresponde al mismo en toda la imagen, es decir, depende de $f(x,y)$, como lo habíamos mencionado anteriormente. Para la obtención de este umbral se utilizaron tres métodos; uno manual y dos que seleccionan el umbral en forma automática.

4.3.1.1 Umbral Manual

Este método consiste en colocar valores diferentes T y evaluar el error de la segmentación tanto subjetiva (determinado por la apreciación visual del autor) como objetivamente (en apartes posteriores se hará una descripción más detallada de la forma en que el error es evaluado cuantitativamente). La forma en que estos valores son colocados es haciendo un

barrido manual en forma creciente de estos valores. El umbral T seleccionado será aquel que menor error presente, éste es evaluado al final del proceso de segmentación. (ver tabla 4.1 y 4.2).

4.3.1.1.1 Para diferencia de imágenes (Objeto - Referencia)

La umbralización se aplicó a la imagen obtenida por distancia euclidiana de los puntos representativos de color en el espacio RGB de la imagen objeto y la imagen referencia, obteniéndose así una imagen binaria, en donde se espera que el objeto esté suficientemente aislado. (Ver figura 4.8). . El umbral utilizado es de valor 35 para imágenes de 256 valores de gris.

Figura 4.8 Resultados de la umbralización manual para las diferencias de imágenes (Objeto-Referencia)

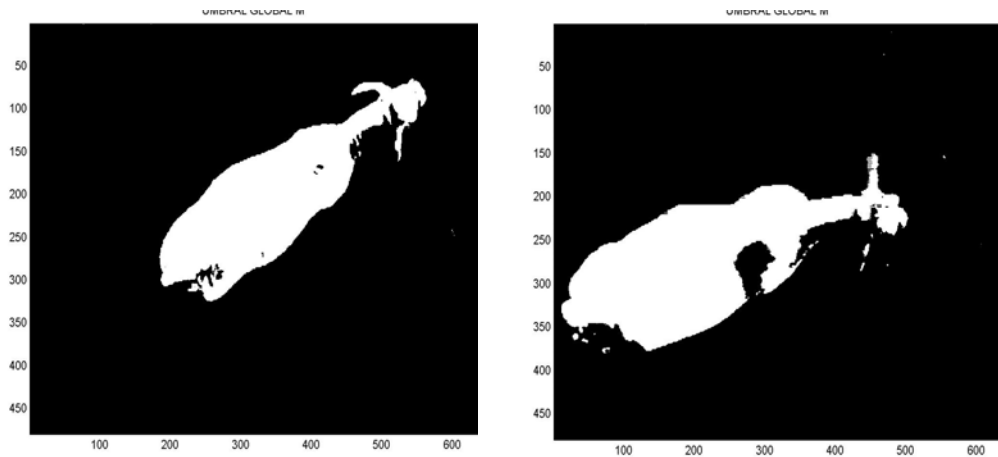
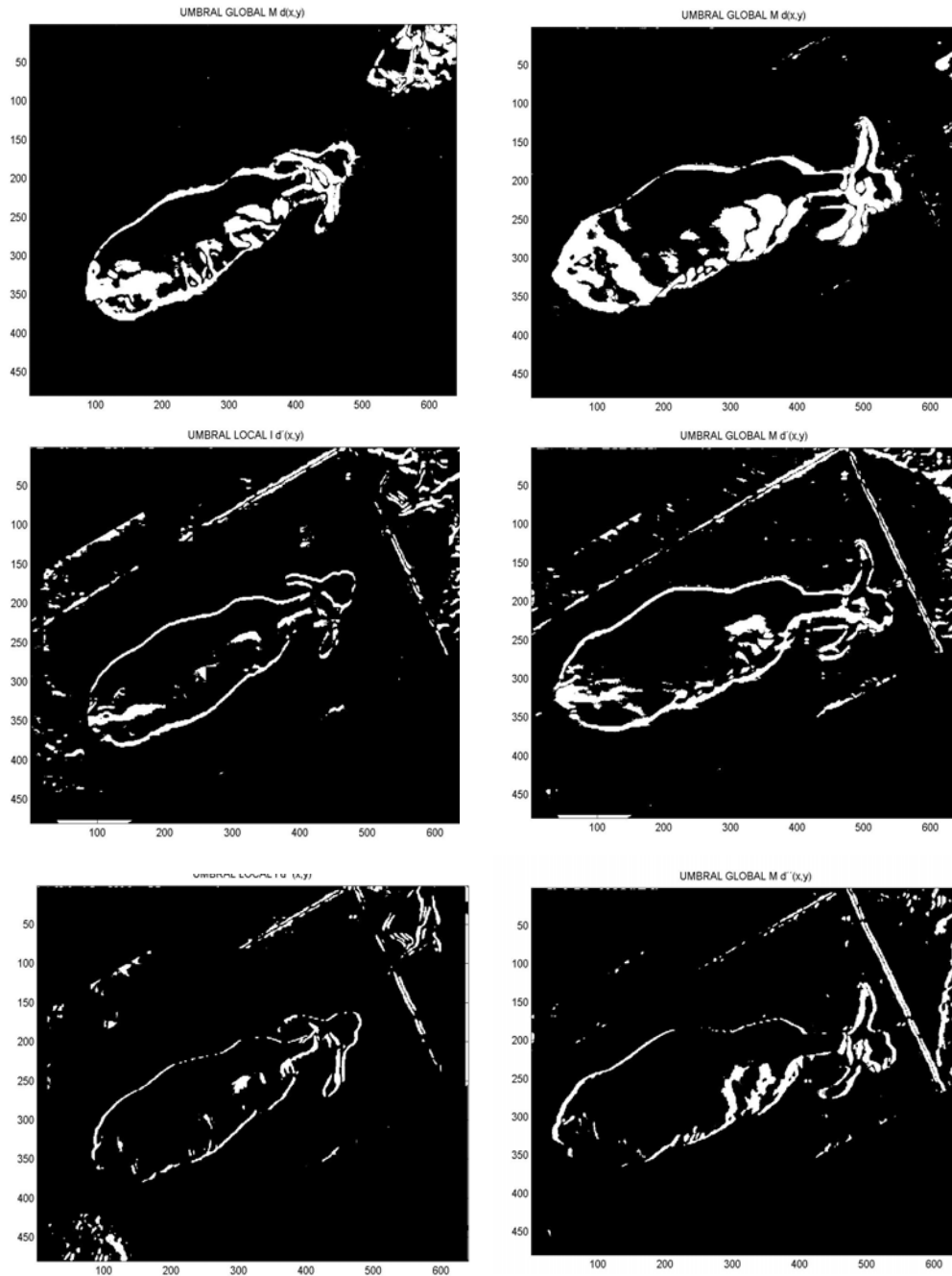


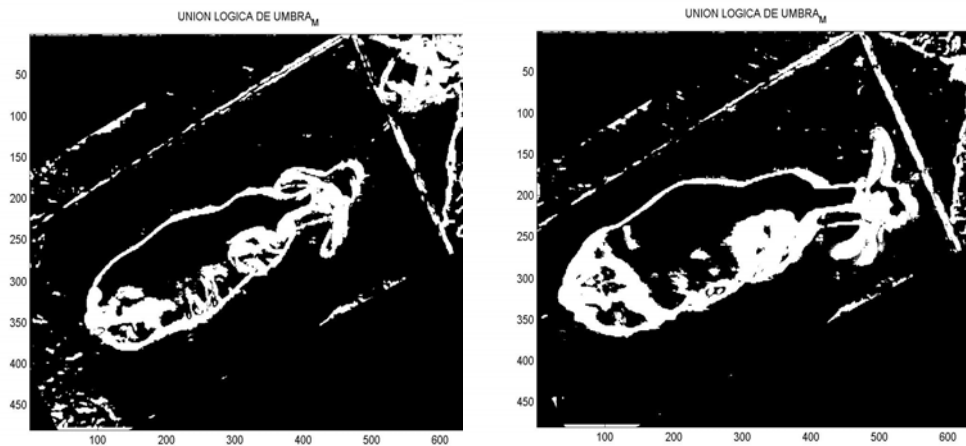
Figura 4.9. Resultados de la Umbralización manual



4.3.1.1.2 Para diferencia de imágenes de secuencia.

La umbralización en este caso se le aplico a las tres imágenes de acumulación de planos de color de las diferencias absolutas $d(x, y)_{\Sigma}$, $d'(x, y)_{\Sigma}$ y $d''(x, y)_{\Sigma}$, mencionada anteriormente, de tal forma que obtenemos tres imágenes binarias (ver figura 4.9). Como las imágenes ofrecen información complementaria unas a otras, se hace necesario una unión de los resultados, que es implementada con una operación OR entre las imágenes. (ver figura 4.10).

Figura 4.10. Unión lógica de las imágenes binarizadas



4.3.1.2 Umbral automático

Encontrar el umbral óptimo en forma manual es en la mayoría de las ocasiones un proceso largo y tedioso. Un método que permita obtener el umbral óptimo en forma automática es entonces la solución rápida al problema, pero nace la pregunta ¿qué método puede ser tan generoso?, es muy posible que ningún método lo sea, sobre todo por que existen muchas parámetros que influyen, tales como la iluminación, la reflectividad, etc. Sin embargo en nuestro caso se ha realizado un preprocesado a la imagen para poder tener una imagen cuyo histograma sea en lo posible muy cercano a bimodal, en donde podemos encontrar un umbral que separe las dos regiones de brillo. A continuación se describen dos métodos que fueron utilizados para encontrar el umbral óptimo.

Método Iterativo

Podemos suponer que nuestra imagen contiene dos regiones de brillo principales. El histograma de una imagen de este tipo se puede considerar como una estimación de la función densidad de probabilidad de brillo, $p(z)$. Esta función de densidad total es la suma o mezcla de dos densidades unimodales. Una para las regiones claras de la imagen y la otra para las oscuras. Además, los parámetros de la mezcla son proporcionales a las superficies de cada región de brillo. Si se conoce o se supone las formas de las densidades, es posible determinar un umbral óptimo.

Supongamos que una imagen contiene dos valores combinados con ruido aditivo gaussiano (que en nuestro caso simularía las dos regiones de brillo). La función de probabilidad de la mezcla es:

$$p(z) = P_1 p_1(z) + P_2 p_2(z) \quad (4.11)$$

que, para el caso gaussiano, es:

$$p(z) = \frac{P_1}{\sqrt{2\pi}\sigma_1} \exp\left[-\frac{(z-\mu_1)^2}{2\sigma_1^2}\right] + \frac{P_2}{\sqrt{2\pi}\sigma_2} \exp\left[-\frac{(z-\mu_2)^2}{2\sigma_2^2}\right] \quad (4.12)$$

donde μ_1 y μ_2 son los valores medios de los dos niveles de brillo, σ_1 y σ_2 dos las desviaciones estándar respecto a la media, y P_1 y P_2 son la probabilidades a priori de los dos niveles. Se debe satisfacer la restricción:

$$P_1 + P_2 = 1 \quad (4.13)$$

por lo que la densidad de la mezcla tiene cinco parámetros desconocidos, Si se conocen todos los parámetros, se puede determinar con facilidad el umbral óptimo.

Supóngase que las regiones oscuras corresponden al fondo y que las claras corresponden a los objetos (puede ser al contrario). En este caso $\mu_1 > \mu_2$, se puede definir un umbral T de forma que todos los píxeles con nivel con nivel de gris superior a T se consideren puntos del objeto. La probabilidad de clasificar (erróneamente) un punto del objeto como un punto del fondo es:

$$E_1(T) = \int_{-\infty}^T p_2(z) dz \quad (4.14)$$

De forma similar, la probabilidad de clasificar un punto del fondo como un punto del objeto es:

$$E_2(T) = \int_T^{\infty} p_1(z) dz \quad (4.15)$$

por lo tanto, la probabilidad total de error es:

$$E(T) = P_2 E_1(T) + P_1 E_2(T) \quad (4.16)$$

Para encontrar el valor umbral para el que el error es mínimo se necesita derivar $E(T)$ con respecto a T (utilizando la regla de Leibnitz) e igualar el resultado a 0. Así:

$$P_1 p_1(T) = P_2 p_2(T) \quad (4.17)$$

Aplicando este resultado a la densidad gaussiana, tomando logaritmos, y simplificando, se obtiene la ecuación cuadrática:

$$AT^2 + BT + C = 0 \quad (4.18)$$

donde:

$$A = \sigma_1^2 - \sigma_2^2 \quad (4.19)$$

$$B = 2(\mu_1\sigma_2^2 - \mu_2\sigma_1^2) \quad (4.21)$$

$$C = \sigma_1^2 \mu_2^2 - \sigma_2^2 \mu_1^2 + 2\sigma_1^2 \sigma_2^2 \ln(\sigma_2 P_1 / \sigma_1 P_2) \quad (4.22)$$

La posibilidad de tener dos soluciones indica que se pueden necesitar dos valores de umbral para obtener una solución óptima.

Si las varianzas son iguales, $\sigma^2 = \sigma_1^2 = \sigma_2^2$, es suficiente con un umbral único:

$$T = \frac{\mu_1 + \mu_2}{2} + \frac{\sigma^2}{\mu_1 - \mu_2} \ln\left(\frac{P_2}{P_1}\right) \quad (4.23)$$

Si las probabilidades anteriores son iguales, $P_1 = P_2$, el umbral óptimo es promedio de las medias. Lo mismo es valido para $\sigma = 0$.

Se puede utilizar un método de error medio por mínimos cuadrados para estimar los parámetros de una imagen a partir de un histograma. En general, la determinación analítica de los parámetros que minimiza el error cuadrático medio no es un tema sencillo. No obstante, para el caso gaussiano, el calculo directo de igualar las derivadas parciales a 0 conducen a un conjunto de ecuaciones simultáneas trascendentales que normalmente solo se pueden resolver por procedimientos numéricos (gradiente descendente, gradiente conjugado o un método de Newton, entre otros); puede ser suficiente considerando a priori que todas las probabilidades son iguales. También los valores de media y varianza se pueden determinar simplemente dividiendo el histograma en dos partes respecto a su valor medio y calculando las medias y varianzas de las dos partes como valores de salida.

El método implementado consiste en suponer a priori que las varianzas y las probabilidades son iguales, lo que nos lleva a:

Figura 4.11. Comportamiento del umbral en el método iterativo

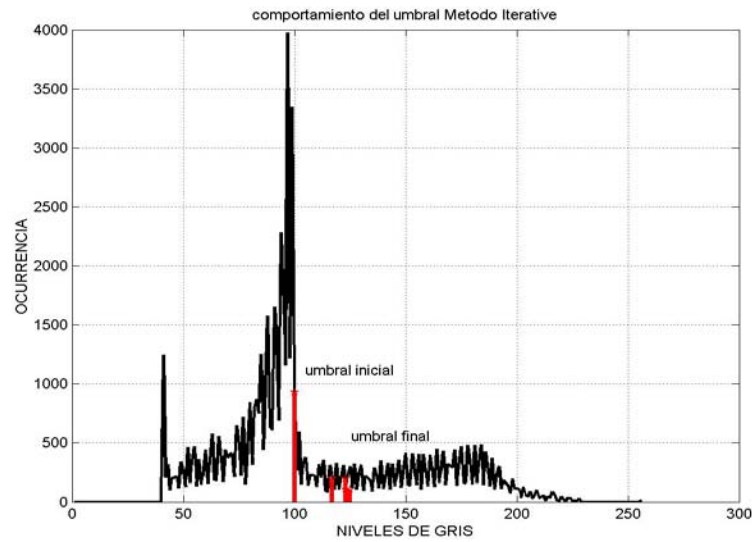
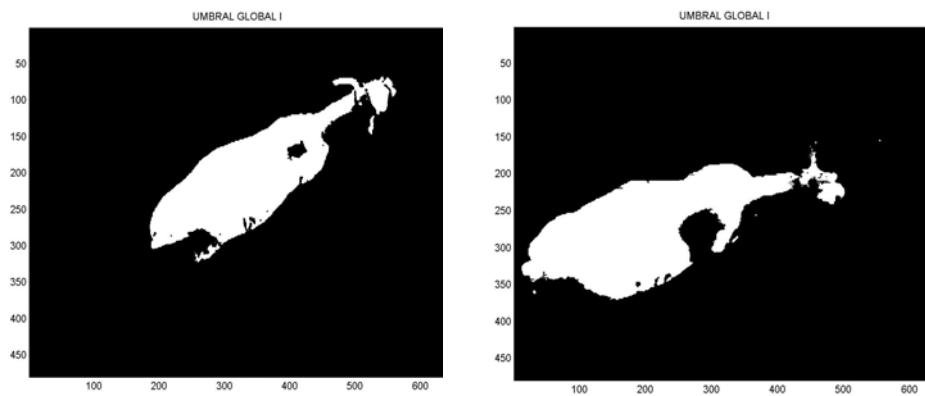


Figura 4.12. Imágenes binarizada con umbralización con método iterativo



$$T = \frac{\mu_1 + \mu_2}{2} \quad (4.24)$$

Posteriormente se divide el histograma en dos partes, cuyo valor constituye el umbral inicial, el cual se asume divide de manera aproximada las dos modas, luego se calcula la media a las dos partes y se aplica el criterio establecido en la ecuación inmediatamente anterior, el cual nos permite obtener el nuevo umbral. Seguidamente se compara el umbral inicial (anterior o viejo) con el actual (nuevo), este proceso se realiza de manera iterativa hasta que la diferencia entre el umbral inmediatamente anterior y actual sea mínima. De esta forma se obtiene el umbral para la binarización de la imagen (Ver figura 4.11 y 4.12).

Método Histograma

Este método consiste en la manipulación del histograma para lograr obtener el umbral óptimo, partiendo de la premisa que el histograma es bimodal. Este consiste en encontrar el mínimo entre los dos máximos del histograma.

El proceso se inicia aplicando al histograma $P(z)$ un filtro pasabajo fuerte, específicamente un promediado, con una ventana de tamaño 9 obteniéndose un histograma suavizado $P(z)_s$

Posteriormente se calcula la derivada del histograma suavizado $\frac{\partial P(z)_s}{\partial r}$ y a ésta se le aplica

también un filtro pasabajo fuerte $\left[\frac{\partial P(z)_s}{\partial r} \right]_s$. Seguidamente se obtienen los dos máximos del

histograma buscando los ceros de la derivada, $\left[\frac{\partial P(z)_s}{\partial r} \right]_s = 0$, obtenido sus niveles de

grises correspondientes $z_{\max 1}$ y $z_{\max 2}$ y evaluando su valor $P(z_{\max 1})$ y $P(z_{\max 2})$. Por

último se localiza el nivel de gris correspondiente al mínimo del histograma en medio de los dos máximos obtenidos

Figura 4.13. Obtención del umbral en el método Histogram

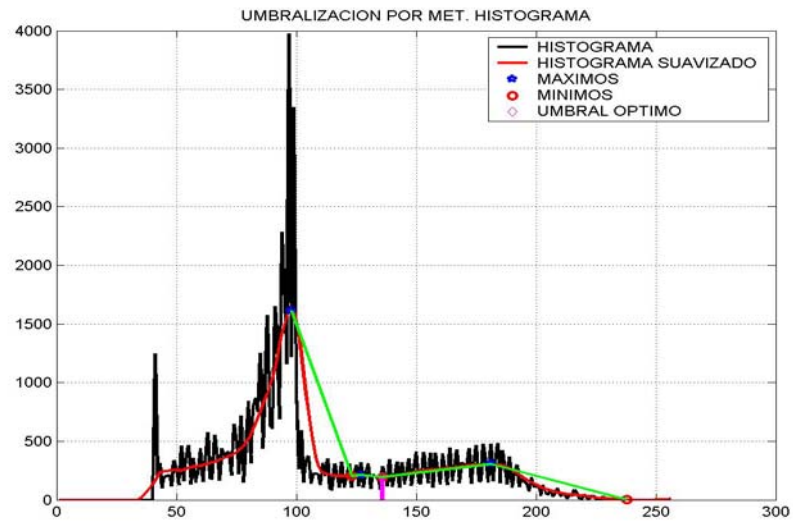
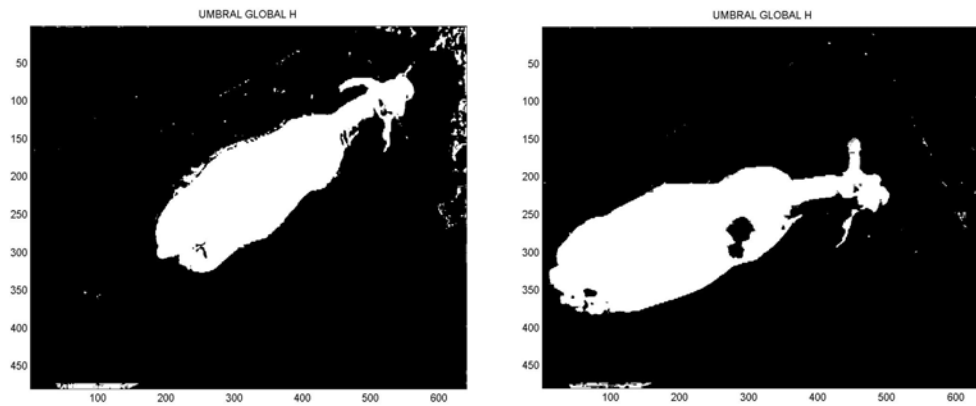


Figura 4.14. Imágenes binarizada con umbralización con método histogram



anteriormente $z_{\max 1} < z_{\min} < z_{\max 2}$, donde este valor será el umbral óptimo por este método

$$T = z_{\min} .$$

4.3.2 Umbralización Local

Con el propósito de mejorar la segmentación de la imagen, se ha realizado binarización local. En este caso se definen ventanas, de tamaño $M \times M$ pixeles; para cada una de estas se establece su propio umbral.

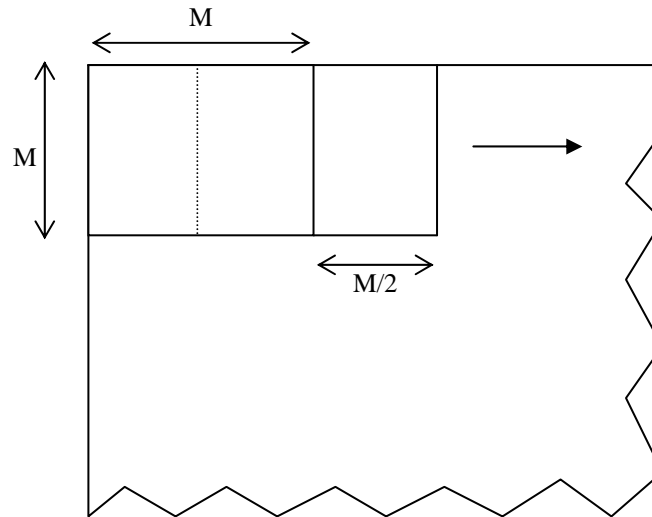
Para cada ventana i se calcula un umbral T_i por el método iterativo, con la particularidad que cada ventana tendrá su umbral dependiendo de su histograma. Así, el umbral local se define como:

$$T_i = \frac{\mu_{1i} + \mu_{2i}}{2} \quad (4.25)$$

donde μ_{1i} y μ_{2i} son las desviaciones estándar de cada moda extraídas del histograma local. El uso de ventanas se propone para mejorar la estimación del umbral, adaptándose de acuerdo a las características particulares de la subimagen en la ventana. Sin embargo el tamaño de las ventanas debe ser lo suficientemente grande para tratar de mantener el comportamiento estadístico del histograma bimodal local.

La binarización de toda la imagen utilizando binarización de subimágenes, se obtiene realizando desplazamientos de ventana de $M/2$ pixeles en sentido horizontal y vertical. Esto implica que habrá zonas de solapamiento (figura 4.15). Para decidir en las zonas de solapamiento se aplica la operación AND, de tal forma un pixel se clasifica como perteneciente al objeto si supera el umbral T_i de todas las i -ventanas que lo incluyan.

Figura 4.15. Desplazamiento de las subimágenes.

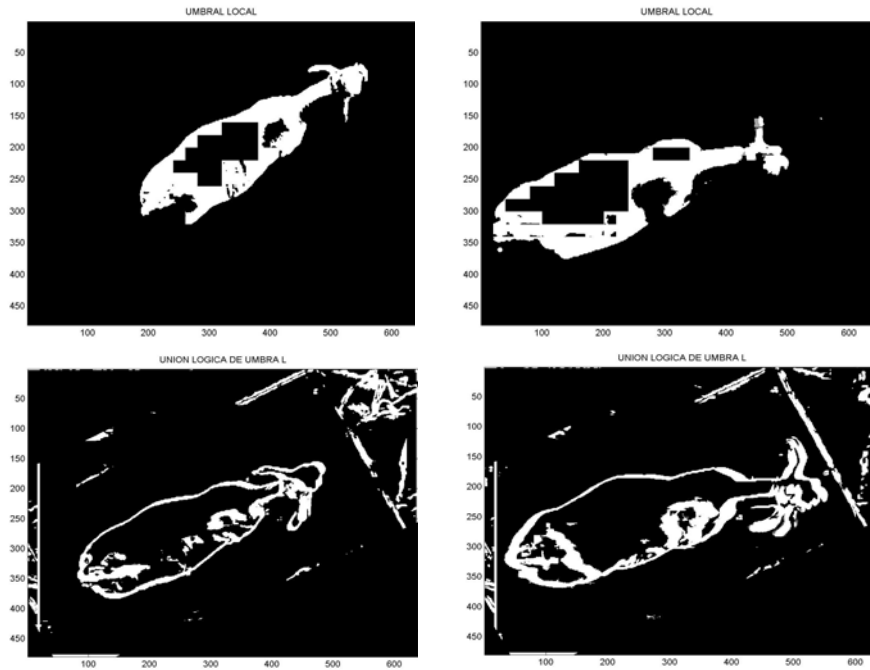


El solapamiento de ventanas adyacentes en la binarización local asegura una transición suave entre umbrales de ventanas vecinas. La figura 4.16 muestra los resultados de la binarización local con un tamaño de ventana de $M = 64$, para los dos tipos de segmentación (diferencia objeto-referencia y diferencia de secuencia).

4.4 SELECCIÓN DEL OBJETO DE INTERÉS.

En las imágenes binarias obtenidas, hay pequeños objetos binarios que no pertenecen al espécimen (objeto de interés), se puede considerar que es ruido binario la cual conviene eliminar y otras regiones que por lo general no son tan pequeñas, que si pertenecen a la espécimen y por tanto es conveniente conservar.

Figura 4.16. Imágenes binarizada con umbralización local



Suponemos que el área del objeto es mucho mayor que la del ruido binario, pero, a su vez los objetos pertenecientes al espécimen aislados de este, son de área mayor que el ruido y menor que la de este. Para aislar el objeto de interés, primero se remueven los pixeles aislados, posteriormente se realiza una dilatación de cinco iteraciones, para lograr conectar las regiones pertenecientes a un mismo objeto, y rellenamos las regiones, para luego hacer una erosión de cinco iteraciones, de tal manera que no divida las regiones nuevamente. Por último se elimina definitivamente el ruido binario, calculando el área de cada uno de las regiones en la imagen y seleccionando el objeto de área mayor, que en este caso es el espécimen u objeto de interés (figura 4.17). Los resultados de este proceso se pueden apreciar en la figura 4.18 y 4.19.

Figura 4.17. Diagrama de bloques para la selección del área de interés.

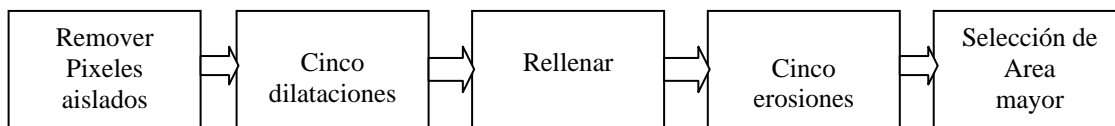


Figura 4.18. Objeto de interés por diferencia de imágenes (objeto-referencia) (a) M. Manual. (b) M. Iterative. (c) M. Histogram. (d) M. Local.

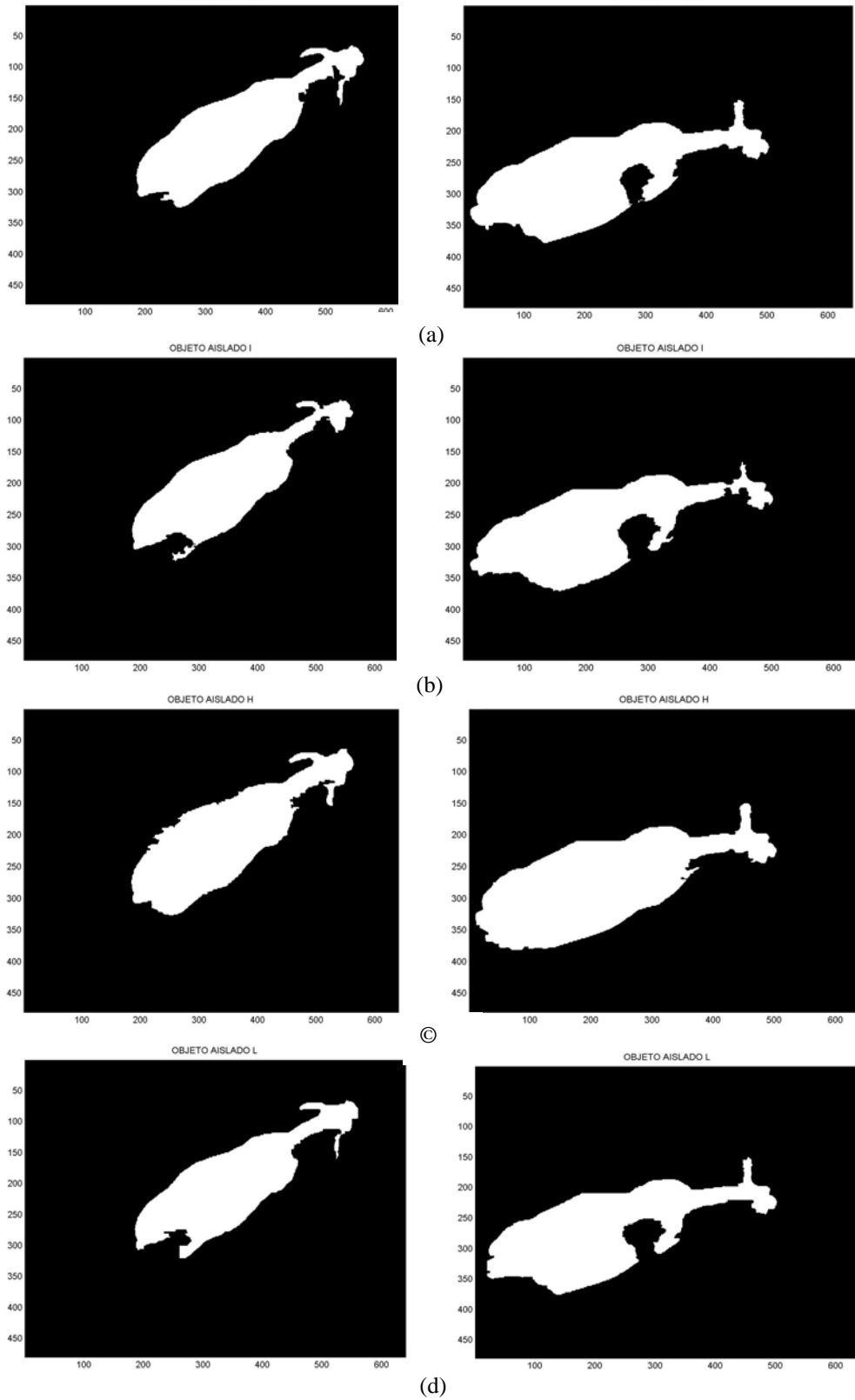
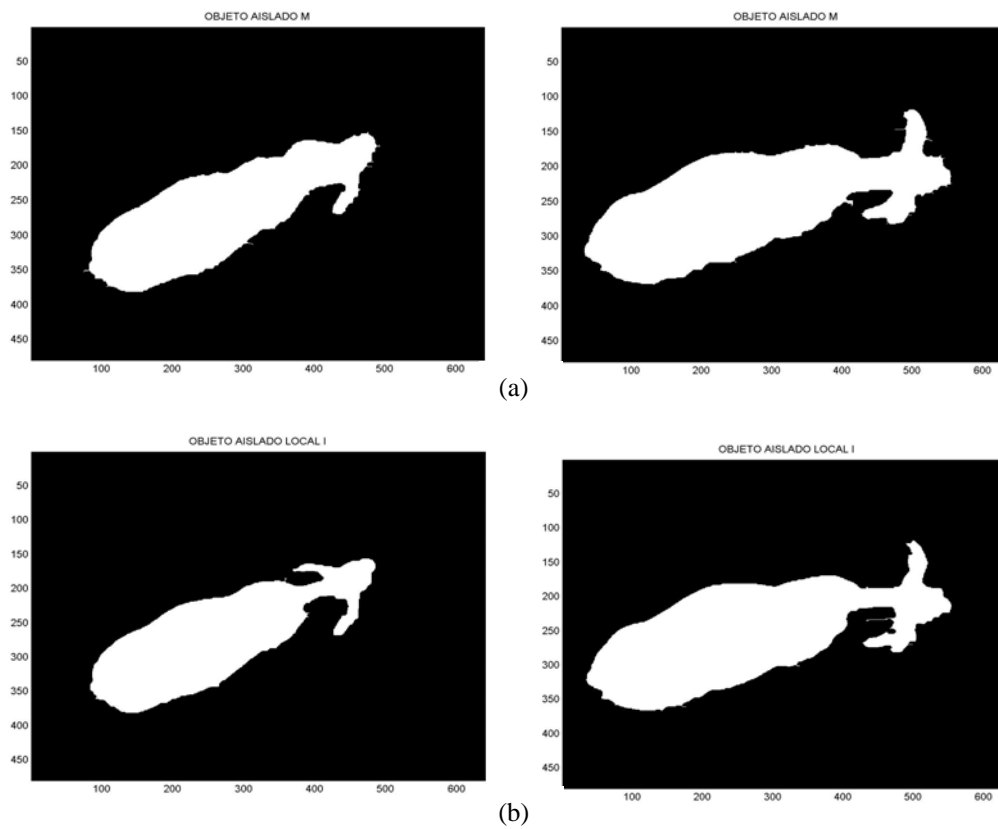


Figura 4.19. Objeto de interés por diferencia de secuencias de imágenes (a) M. Manual..
(b) M. local.



4.5 COMPARACIÓN DE LOS MÉTODOS DE SEGMENTACIÓN

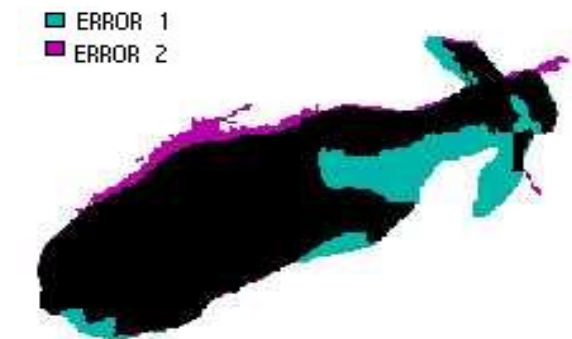
Para la comparación de los métodos de segmentación utilizados, se implementó un criterio para medir el error basado en la diferencia y la razón porcentual con respecto a un valor aceptado. El valor aceptado se obtiene demarcando manualmente el contorno del animal.

Se midió el error de la segmentación para cada uno de los métodos de segmentación aplicados, para lo cual se tuvo en cuenta el error por defecto (región interna del objeto faltante) y por exceso (región externa del objeto sobrante). Al error por defecto le llamaremos *Error 1* y al error por exceso *Error 2*.

$$ERROR\ 1 = \left| \frac{V_e - V_{o1}}{V_e} \times 100 \right| , \quad ERROR\ 2 = \frac{V_{o2}}{V_e} \times 100 , \quad (4.26)$$

donde, V_e , V_{o1} y V_{o2} son el valor esperado, el valor obtenido interno y el valor obtenido externo respectivamente. Los errores obtenidos para las 100 muestras en cada uno de los métodos de segmentación aplicado se presentan en la siguiente tabla en sus valores promedios.

Figura. 4.20. Regiones que representan los tipos de errores



Una versión más ampliada se puede apreciar en el Anexo 2 y el Anexo 3.

Tabla 4.1. Errores porcentuales promedios de los diferentes métodos de segmentación.

SEGMENTACIÓN POR DIFERENCIA DE IMÁGENES OBJETO-REFERENCIA				
UMBRAL GLOBAL				UMBRAL LOCAL
	MANUAL	M. ITERATIVE	M. HISTOGRAM	M. ITERATIVE
ERROR 1	24.03 %	36.8 %	14.0 %	38.9 %
ERROR 2	17.7 %	19.1 %	44.5 %	3.9 %

Tabla 4.2. Errores porcentuales promedios de los diferentes métodos de segmentación.

SEGMENTACIÓN POR DIFERENCIA DE SECUENCIAS DE IMÁGENES		
UMBRAL GLOBAL		UMBRAL LOCAL
	MANUAL	M. ITERATIVE
ERROR 1	15.67 %	36.0 %
ERROR 2	18.8 %	11.5 %

De acuerdo a los errores obtenidos para la segmentación por diferencia de imágenes objeto-referencia, se puede decir que los métodos de umbralización automáticos se comportan de manera similar, aunque con algunas diferencias. El menor error por defecto (*Error 1*) se obtuvo para el *método histogram* de 14.0%, quiere decir, que este método recupera mejor la información perteneciente al objeto que los demás métodos de umbralización. Sin embargo no

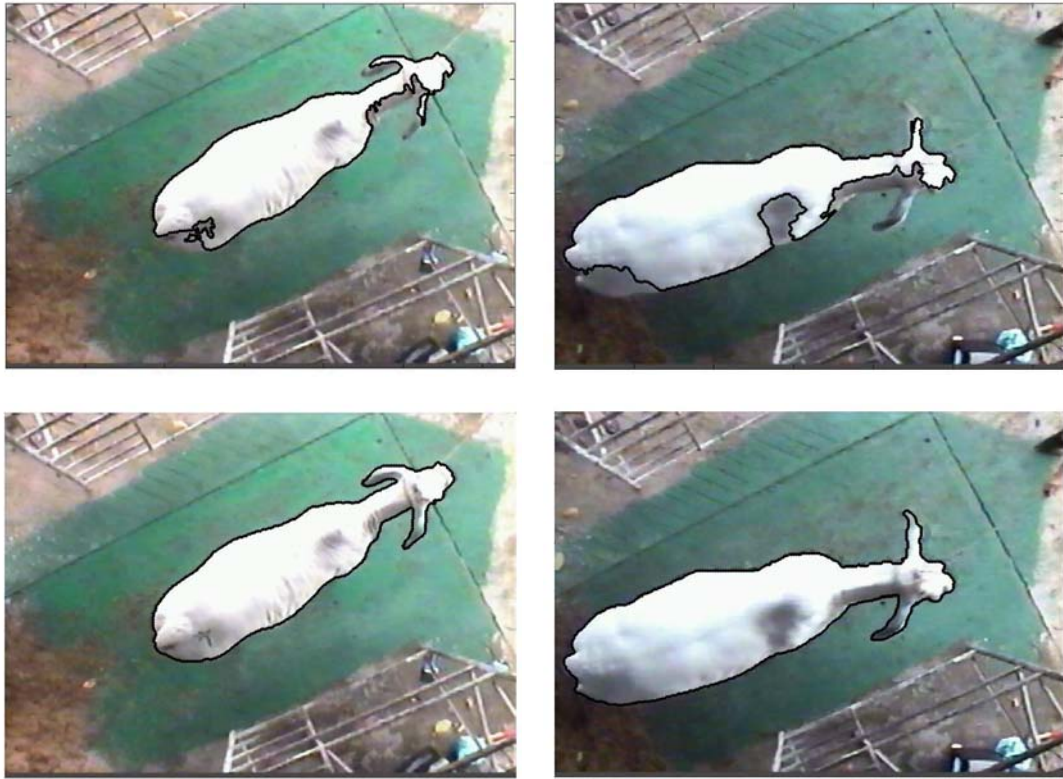
los hace nada bien en el error por exceso (*Error 2*) de 44.5%, en este caso obtuvo el mayor error promedio. Debido a lo anterior, este método no solo es bueno recuperando información perteneciente al objeto, también lo hace recuperando información que no pertenece, consecuencia del umbral global. Sin embargo el umbral local se comporta muy bien obteniendo el menor error por exceso de 3.9%, pero a su vez el mayor error por defecto de 38.9%.

En el caso de la segmentación de diferencias de secuencias de imágenes solo se utilizaron dos métodos, uno de umbral global (manual) y el otro umbral local. Aunque el comportamiento de los errores es parecido para los dos tipos, el menor en el *Error 1* 15.67%, y en cambio para la umbralización local se obtiene menor error en el *Error 2* 11.5%.

Es claro que para solucionar el problema propuesto en esta investigación el error en la segmentación para los dos tipos (1 y 2) debe ser cero. Esto debido a que las dimensiones que guardan alta correlación con la masa del animal se miden directamente de la imagen, y si esta no es realizada con error cero estas dimensiones serán alteradas, ocasionando errores en la estimación del peso.

Por tal razón el método que se implementa en este es logrado ajustando manualmente el contorno resultante al aplicar una de los métodos automáticos anteriores (por diferencia de imágenes objeto-referencia, por el M. Iterative), Se Obtiene así un método semiautomático de segmentación, cuyo error es cero (*la medida del error es subjetiva, pero válida en cuanto a segmentación se refiere*). Resultados de la aplicación de este método se pueden apreciar en la figura 4.21.

Figura 4.21. Resultado de utilizar la segmentación semiautomática.



5. EXTRACCIÓN DE CARACTERÍSTICAS

En todo sistema de visión artificial, la etapa de selección de rasgos característicos constituye una etapa fundamental y difícil de formalizar en el diseño. Los criterios para la selección de estas dependen en gran parte del ingenio y de la experiencia del diseñador, sin embargo existen criterios objetivos que ayudan en gran parte en esa selección. En adelante expondremos como afrontamos este problema de acuerdo a las dificultades que se presentaron.

5.1 INVARIANZA A LA ROTACIÓN

Las imágenes de los objetos binarios obtenidos en la escena de entrada presentan orientaciones diversas, la cual implica cierta dificultad en el momento obtener las dimensiones (características), por tal razón se hace necesario utilizar un método que permita acceder a estas en forma automática e independiente de la orientación del objeto. Para tal caso se utiliza las bondades de la transformada de Hotelling o *karhunen-loève* (K-L).

5.1.1 Transformada de Hotelling O K-L

La transformada que se describe a continuación es la conocida transformada *karhunen-loève* (K-L), también llamada como la transformada del autovector, de la componente principal o de Hotelling (Gonzalez y Wood 1992), se basa en las propiedades estadísticas de las representaciones vectoriales. La transformada K-L tiene algunas propiedades que la hacen una herramienta útil para muchas aplicaciones de procesado de imágenes.(Levy y Rubinstein 1999; Charraut et al 1996; sajovi y Burger 1985; Fukunaga y Koontz).

Considérese una población de vectores aleatorios de la forma

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (5.1)$$

el vector media de esta población se define como

$$\bar{X} = E\{X\}, \quad (5.2)$$

Donde $\bar{X} = E\{\text{arg}\}$ es el valor esperado del argumento. Además, hay que tener en cuenta que el valor esperado de un vector o de una matriz se obtiene tomando el valor esperado de cada elemento.

La matriz de covarianza de la población de vectores se define como

$$C_x = E\left\{(X - \bar{X})(X - \bar{X})^T\right\}, \quad (5.3)$$

Donde T indica el vector traspuesto. Dado que X es n dimensional, C_x y $(X - \bar{X})(X - \bar{X})^T$ son matrices de orden $n \times n$. La matriz C_x es real y simétrica.

Para M muestras vectoriales de una población aleatoria, el vector media y la matriz de covarianza se pueden aproximar por:

$$\bar{X} = \frac{1}{M} \sum_{k=1}^M X_k, \quad (5.4)$$

y

$$C_X = \frac{1}{M} \sum_{K=1}^M (x_K x_K^T - \bar{X} \bar{X}^T), \quad (5.5)$$

como C_X es real y simétrica, siempre es posible hallar un conjunto de n valores propios ortonormales. Por definición, los vectores y valores propios de una Matriz $n \times n$, de C_X , satisface la relación

$$C_X V_i = \lambda_i V_i, \text{ para } i=1,2,3,\dots,n, \quad (5.6)$$

donde V_i y λ_i son los $i=1,2,3,\dots,n$ valores y vectores propios correspondientes de C_X , que ordenados en forma decreciente tal que $\lambda_j \geq \lambda_{j+1}$, para $j=1,2,\dots,n-1$.

Por tal razón podemos definir, la transformada K-L mediante la ecuación

$$Y = A(X - \bar{X}), \quad (5.7)$$

donde A es una matriz de transformación cuyas filas están formadas por los vectores propios de C_X , ordenadas de tal forma que la primera fila de A es la del vector propio correspondiente al mayor valor propio, y la última fila es el vector propio correspondiente al menor valor propio.

Una característica de los vectores transformados Y es que su media es cero y la matriz de covarianza de Y puede obtenerse en términos de A y C_X mediante

$$C_Y = A C_X A^T. \quad (5.8)$$

Además, C_Y es una matriz diagonal cuyos elementos de la diagonal principal son los valores propios de C_X , es decir

$$C_Y = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}. \quad (5.9)$$

Los elementos que están fuera de la diagonal de la matriz de covarianza son 0, por tanto los elementos de los vectores Y no están correlacionados.

Al aplicar esta transformada a imágenes es preciso prepararla en forma de vectores. Esta preparación dependerá del tipo de problema que se quiera solucionar. La primera posibilidad es que la imagen puede ser puesta en forma de vector; otra es que en imágenes multidimensionales (multiespectrales) los vectores se organizan con los valores de los píxeles correspondientes en cada espectro, y una más es que en una imagen binaria el vector se obtenga de las coordenadas del objeto, la cual en tal caso sería un vector bidimensional

$$X = (a, b)^T, \quad (5.10)$$

donde a y b son los valores de las coordenadas de ese punto. Esos vectores se emplean para calcular el vector medio y la matriz de covarianza de la población (objeto).

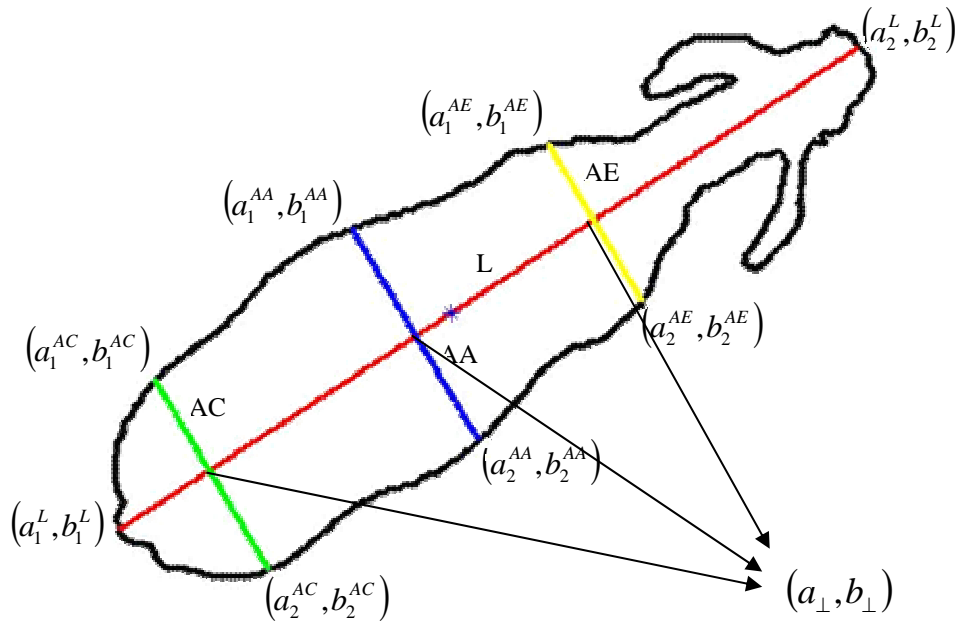
El efecto, es tomar los vectores de entrada, que en nuestro caso son las coordenadas de posición de los píxeles de valor 1 (pertenecientes al objeto binario) y alinearlos con su autovector principal, es decir, es una transformación de rotación que alinea los datos con los autovectores. Esto nos permite tener acceso al ángulo del autovector principal (inclinación del

objeto), aquel, en la que la mayoría de los datos están alineados y poder así tomar las dimensiones independientemente de la inclinación del objeto.

5.2 OBTENCIÓN Y SELECCIÓN DE CARACTERÍSTICAS.

El procedimiento para la estimación del peso consiste en encontrar ciertas dimensiones o características que se puedan obtener sobre la imagen y guarden una alta correlación con el peso. Se conoce varias dimensiones que tienen este comportamiento, todas ellas han sido evaluadas sobre el animal mostrando una alta correlación con su peso ^[16] (realmente su masa en kilogramo). En nuestro caso tomamos solo algunas de ellas, pero estas medidas sobre la imagen adquirida (en pixeles). Algunas dimensiones no se consideraron, debido a la vista elegida (vista superior). Las dimensiones (ver figura 5.1) obtenidas son :

Figura 5.1. Dimensiones o características extraídas de la imagen.



- **Área (A).** Es el número de píxeles pertenecientes al objeto, es decir, aquellos que tienen valor 1.

$$\text{Area} = \text{card}\{I_a: (a,b)=1\}, \quad (5.10)$$

Donde, *card* denota la cardinalidad, es decir, el número de elementos y I_a la imagen segmentada.

- **Perímetro (P).** Es el número de píxeles pertenecientes al perímetro del objeto.

$$\text{Perímetro} = \text{card}\{I_p: (a,b)=1\} \quad (5.11)$$

- **Largo (L).** Es la distancia a lo largo del objeto, medido desde la cola hasta la cabeza. Para obtener coordenadas de los extremos se hace uso del ángulo α obtenido de la inclinación del autovector (con ayuda de la transformada K-L) y de las coordenadas del centroide del objeto. La expresión para obtener el largo del objeto es.

$$\text{Largo} = \sqrt{(a_2^L - a_1^L)^2 + (b_2^L - b_1^L)^2} \quad (5.12)$$

- **Ancho de Cadera (AC).** Es la distancia entre los extremos del objeto a la altura de la cadera, medido a lo largo de una línea perpendicular al Largo del animal y a una fracción correspondiente a la 8.33 parte del largo del objeto ($L/8.33$). Este valor se obtiene de promediar la ubicación de la cadera (al igual para las siguientes anchos) en las muestras.

Para obtener las coordenadas de ubicación de la recta perpendicular en la recta del largo del animal, utilizamos la expresión,

$$a_{\perp} = \frac{(f-1)a_1^L + a_2^L}{f}, \quad b_{\perp} = \frac{(f-1)b_1^L + b_2^L}{f} \quad (5.13)$$

donde f corresponde a la fracción considerada sobre la recta del largo del objeto. De esta forma se puede obtener el ancho de cadera con la siguiente expresión.

$$\text{AC} = \sqrt{(a_2^{AC} - a_1^{AC})^2 + (b_2^{AC} - b_1^{AC})^2} \quad (5.14)$$

Ancho de Abdomen (AA). Es igualmente la distancia entre los extremos del objeto a la altura del abdomen, medida a lo largo de la línea recta perpendicular al largo del animal, pero a una fracción de la 2.5 parte del largo del objeto.

$$AA = \sqrt{(a_2^{AA} - a_1^{AA})^2 + (b_2^{AA} - b_1^{AA})^2} \quad (5.15)$$

- **Ancho de Escápula (AE).** Al igual que las anteriores se obtiene de medir la distancia entre sus extremos a la altura de la escápula, a una fracción del largo de 1.55 parte del animal.

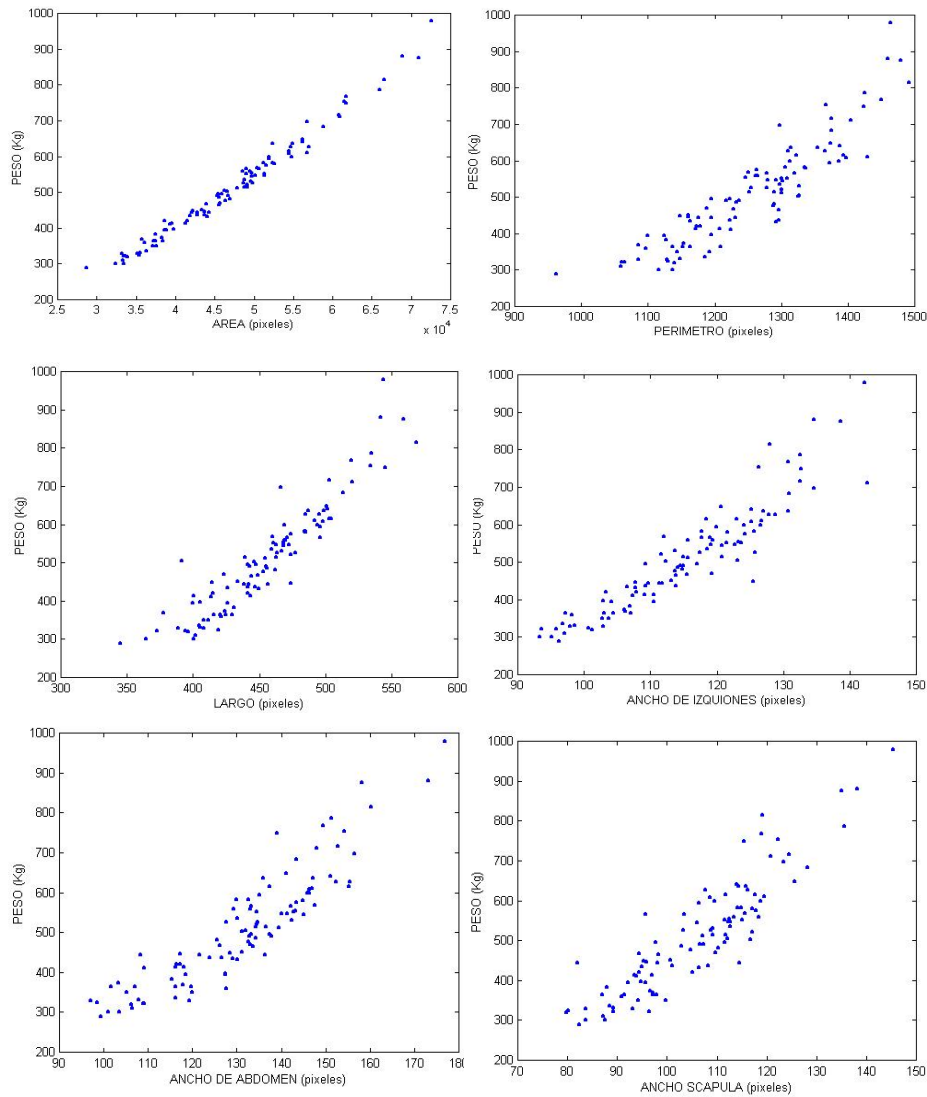
$$AE = \sqrt{(a_2^{AE} - a_1^{AE})^2 + (b_2^{AE} - b_1^{AE})^2} \quad (5.16)$$

Posteriormente se evaluó la correlación existente entre las dimensiones obtenidas sobre el objeto y el peso en kilogramos en vivo, buscando descartar aquellas que presentasen una baja correlación; sin embargo todas las características presentaron una fuerte correlación; tal como se presenta en la tabla 5.1 y en la figura 5.2.

Tabla 5.1. Coeficiente de correlación de las características con el peso del animal

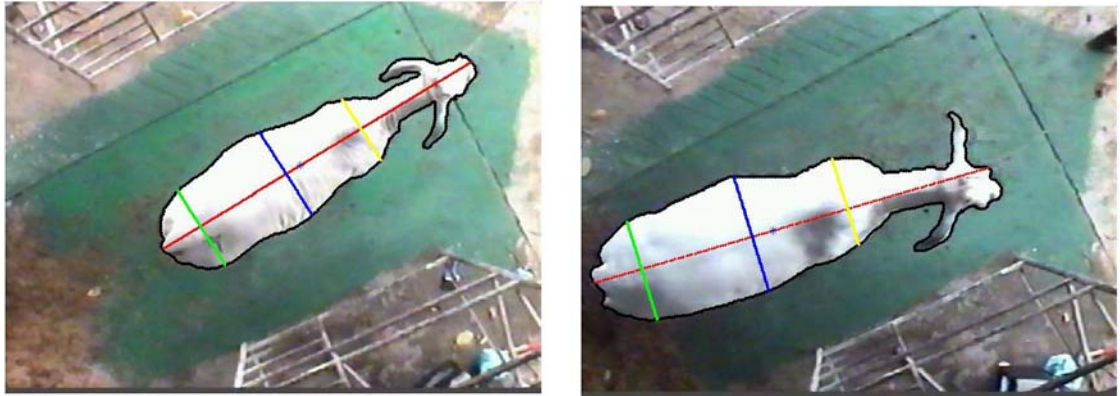
CARACTERÍSTICAS	COEFICIENTE DE CORRELACIÓN
Área (A)	0.9899
Perímetro (P)	0.9083
Largo (L)	0.9300
Ancho de Cadera (AC)	0.9247
Ancho de Abdomen (AA)	0.9050
Ancho de Escápula (AE)	0.8996

Figura 5.2. Comportamiento de las características frente al peso del animal.



Se puede observar que el coeficiente de correlación más alto, es el área (0.9899), la más baja el ancho de la escápula (0.8996), sin embargo debido a su alto valor, todas son consideradas características importantes, que con gran posibilidad permiten encontrar una función de estimación entre estas y el peso del animal en vivo.

Figura. 5.3. Ejemplos de obtención de características sobre el animal.



5.2.1. Análisis de componentes principales.

La selección de las características está basada en el comportamiento que han tenido estas dimensiones frente al peso pero, tomadas directamente sobre el animal. Por tal razón se evaluaron las mismas en donde estas son tomadas directamente de la imagen y presentaron un comportamiento satisfactorio. Sin embargo a pesar del cuidado en la selección de las características su calidad puede ser mejorada notablemente. Esto puede ser logrado utilizando un aparato matemático de cierta complejidad que permite transformar el vector de características en otro que sea *mejor* que el antiguo. Esta herramienta matemática denominada análisis de componentes principales (propio del análisis multivariable) permite una transformación cuyo objetivo es independizar estadísticamente las características, es decir, decorrelacionar estas y también reducir la dimensión del espacio de las características.

El análisis de componentes principales actúa evaluando la importancia de un conjunto de variables independientes X_1, X_2, \dots, X_n en el comportamiento de cierta población de individuos, para lo cual se parte de la diagonalización de la matriz de covarianza C de la citada población.

$$X^* = AX, \quad C^* = ACA^T = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \quad (5.17)$$

donde C^* , A y λ corresponden a la matriz de covarianza de los vectores transformados X^* , la matriz de transformación (compuesta por los autovectores) y los autovalores de la matriz C respectivamente. Las técnicas del análisis de componentes principales se pueden aplicar a la disminución del vector de características si la diagonalización de la matriz de covarianza se realiza sobre el conjunto global de todas las clases existentes.

Es decir, al estimarse la matriz de covarianza global de todas las muestras disponibles de cada una de las clases y posteriormente estas se diagonalicen, de tal forma que se cumpla que $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, en donde las componentes del nuevo vector de características X^* , cuyos autovalores sean mayores corresponderán a las características con mayor dispersión estadísticas (los autovalores de C corresponden a la varianza de las características) y al haberse agrupado todas las clases en una sola población es evidente que las características con dispersión elevada presentarán, mayor capacidad de discriminante o lo que es lo mismo las característica con poca dispersión tenderán a presentar valores parecidos entre las diferentes clases, es decir con poca capacidad discriminante.

Por lo tanto, se puede introducir un umbral de corte λ_h en la secuencia $\lambda_1 > \lambda_2 > \cdots > \lambda_h > \cdots > \lambda_n$, de tal manera que los autovalores (varianzas de las características transformadas) que estén por debajo de dicho umbral corresponderán a las características eliminadas (con poca capacidad discriminante), de tal modo que las características $X_1^*, X_2^*, \dots, X_h^*$ se mantendrán. Consecuentemente la matriz de transformación sufre una reducción en sus dimensiones, A_R conformada ahora por los autovectores correspondientes, de tal forma que $X_R^* = A_R X$, donde X_R^* y A_R corresponden al nuevo vector de características reducido y mejorado (en su capacidad de discriminación e dependencia).

En nuestro vector de características compuesto por $X = [A P L A C A A A E]$ de seis características, se utilizó el análisis de componentes principales con el propósito de mejorarlo y reducirlo (aunque el número de características no es alto, puede contribuir a su optimización), de ese modo las características eliminadas corresponde a aquellas que contribuyen en menos del 2% a la varianza total del conjunto de datos. Si inicialmente se tenía una matriz de características de 6×100 , la reducción permite obtener una matriz de 4×100 .

6. ESTIMACIÓN DE PESO POR REDES NEURONALES.

Para la estimación del peso es necesario utilizar un mecanismo que permita asociar los parámetros extraídos de la imagen y el peso en kilogramos del espécimen. En este caso se utiliza un red neuronal artificial (RNA) que permita realizar dicha asociación. Para la implementación y entrenamiento de la RNA utilizamos la herramienta Matlab para Windows^[17]. En este capítulo se hace una descripción del proceso realizado para obtener la RNA de mejor desempeño.

6.1 IMPLEMENTACIÓN DE LA RED NEURONAL.

Para la construcción de la red neuronal se utilizó como datos de entrada, en un primer experimento la matriz inicial (6x100), y luego en otro la matriz transformada (por componentes principales), en la cual se consiguió reducir el número de entradas de 6 a 4 y como datos de salida los pesos (en kilogramos) del animal. El propósito es diseñar un instrumento que permita determinar el peso de una animal a partir de parámetros o característica obtenidas de la imagen de este. Este constituye un problema típico de aproximación de funciones, en el cual las variable independientes serán las característica y la dependiente el peso del animal.

6.1.1 Estructura de la red neuronal.

Este tipo de problema puede ser solucionado en forma apropiada por una estructura de red neuronal artificial conocida como: red neuronal con propagación hacia adelante (feedforward), dentro de la cual sobresale una; la Perceptron Multicapa. Esta red neuronal esta constituida

como mínimo por tres bloques de procesamiento, todos con sus salidas conectadas con las neuronas de las capas siguientes: uno bloque de entrada, que realmente se forma de una capa, con un número de neuronas determinado por el número de parámetros de entrada, uno de salida constituido también por una capa cuyo número de neuronas está determinado por el número de salidas requerido y uno oculto, con un número de capas y de neuronas no definido previamente sino por el grado de dificultad del problema. Este tipo de estructura de red neuronal junto a funciones de transferencias diferenciables (lineales, sigmoideas, gaussianas) permiten resolver problemas de aproximación de funciones con bastante versatilidad.

Pues bien, de acuerdo a lo anteriormente expuesto pretenderemos solucionar este problema con una estructura de red neuronal de propagación hacia adelante, conocida como Perceptrón Multicapa. Esta estructura tiene 6 neuronas de entrada para el primer experimento y 4 neuronas de entrada (las reducidas) para el segundo en la primera capa (capa de entrada) y 1 neurona en su capa de salida, encargada de dar el peso del animal, el número de capas y de neuronas ocultas en cada experimento se determinará más adelante. Puede pensarse que la relativa linealidad entre las características y el peso del animal indica que la red puede ser sencilla.

6.1.2 Entrenamiento de la red.

En este momento cuando ya conocemos la matriz de características (de entrada) y la matriz de salida (el peso en kilogramos del animal), solo falta utilizar un método mediante el cual le enseñemos a la red a responder un peso adecuado a cada vector de características en la entrada. El método que más aplicaciones prácticas tiene se conoce como *backpropagation* o algoritmo de propagación del error hacia atrás, que está basado en la generalización de la regla delta propuesta por Widrow-Hoff. Este algoritmo de entrenamiento tiene la ventaja de responder adecuadamente a entradas que no estuvieron presentes en la fase de entrenamiento, capacidad conocida como generalización. Sin embargo han surgido varias

modificaciones basadas en técnicas de optimización estándar tales como de Gradiente Conjugado y métodos de Newton^{[18][19]}.

Estas variaciones del algoritmo de retropropagación han surgido con el propósito de minimizar el tiempo de entrenamiento (en épocas), debido a que en el algoritmo básico éste es bastante grande. Sin embargo estas nuevas técnicas de optimización (algoritmo de gradiente conjugado y cuasi-Newton) muy a pesar de lograr converger más rápidamente, tienen la dificultad de aumentar carga computacional y requerimiento de memoria en cada iteración, por tanto nace la necesidad de crear un puente entre estos dos algoritmos y es el algoritmo de *one step secant*^[20] que requiere menos almacenamiento y carga computacional debido básicamente a evitar almacenar la matriz Hessiana y el cálculo de su inversa en cada iteración.

Otra de las exigencias en el comportamiento de una red neuronal es mejorar la generalización. Esta mejora radica en lograr que dicha red genere resultados adecuados a entradas que no estuvieron presentes en la etapa de entrenamiento. Una de las formas en que se logra mejorar la generalización es la implementación de redes suficientemente grandes, de tal forma que se pueda obtener un adecuado ajuste sin importar la complejidad de la función; por el contrario si usamos una red suficiente pequeña no tendría el suficiente poder para sobreajustarse a los datos, lo que implica una metodología de disminución del tamaño de la red hasta prevenir el sobreajuste a los datos. Sin embargo es de gran dificultad conocer con anterioridad el tamaño de la red para un aplicación específica.

Existen otras dos técnicas para mejorar la generalización que eliminan la dificultad anteriormente expuesta y son: la regularización automática y *early stopping* (parada anticipada); la primera consiste en la modificación de la función de desempeño, que en las redes neuronales normalmente se utiliza el error cuadrático medio^[21]. Esta modificación consiste en adicionarle un término que consiste en el promedio de los cuadrados de los pesos (propios de la red neuronal), de esta forma:

$$ECM_{REG} = \frac{\gamma}{N} \sum_{i=1}^N (t_i - a_i)^2 + \frac{(1-\gamma)}{n} \sum_{j=1}^n w_j^2, \quad (6.1)$$

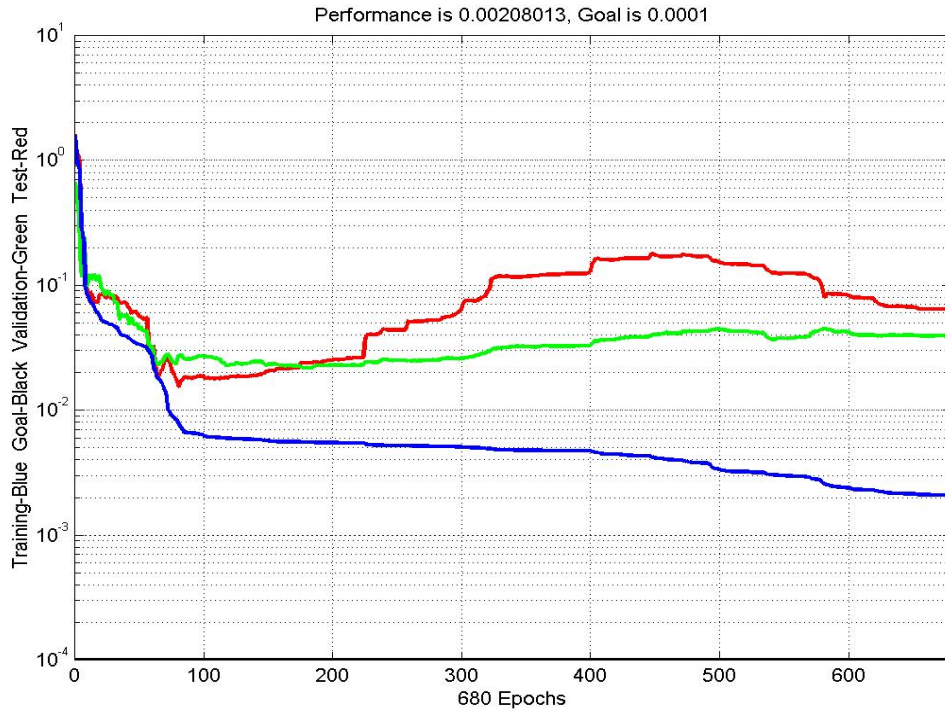
donde el primer termino corresponde al error cuadrático medio y, t_i y a_i son el i-esimo valor esperado y la i-esima salida de la red respectivamente, y el segundo termino corresponde a la media del cuadrado de los pesos de la red y, w y γ son los pesos de la red y un coeficiente de razón de desempeño que es controlado manualmente y su función es regular el sobreajuste. Sin embargo este parámetro puede encontrarse de forma automática (de igual forma los pesos de la red), en donde se consideran variables aleatorias desconocidas que se encuentran asociadas a distribución especifica. Este método se conoce como regularización Bayesiana^{[22][23]}.

La segunda técnica de mejora de regularización llamada *early stopping* consiste en dividir los datos en tres grupos. Uno para el entrenamiento normal de la red, otro para la validación cuya función es el de comparar en cada iteración el comportamiento del error entre estos dos grupos de tal forma que cuando el error de este grupo comience a elevarse por encima del error del grupo de entrenamiento inmediatamente se detiene (para un diferencia prefijada), de ahí el nombre; y por último el grupo llamado de prueba, que no es usado durante el entrenamiento y cuyo propósito es evaluar la generalización.

En este trabajo se seleccionaron tres grupos en forma aleatoria, de tal manera que al grupo de entrenamiento corresponde al 60% del total de la matriz de datos, o sea, que la matriz de entrenamiento tiene dimensiones de 6x60 para el primer experimento y 4x60 para el segundo. Por otro lado los grupos de validación y prueba corresponden al 20% del total de la matriz de datos cada uno, cuyas dimensiones serán de 6x20 y 4x20 para el primer y segundo experimento respectivamente.

También se realizó un estudio pretendiendo evaluar la importancia del sexo del animal en la precisión del sistema. De tal manera que se implementaron redes neuronales en donde los datos de entrada y salida corresponden a especímenes del mismo sexo.

Figura 6.1. Curvas de entrenamiento de los tres grupos



Para poder obtener la configuración completa de todas las redes, hace falta saber cual es el número de capas y neuronas ocultas. No existe una metodología estándar para encontrar de forma precisa el número de capas y neuronas requerido para la solución de determinado problema, sin embargo, se implementó una metodología (no quiere decir que sea la mejor) de barrido en busca de la configuración de red neuronal que generara menor error. Para este caso utilizamos como instrumento de medición del error, la desviación estándar de las diferencias entre el valor esperado y el valor generado por la red, definida como:

$$S_d = \sqrt{\frac{\sum_{i=1}^N [(t_i - a_i) - (\overline{t_i - a_i})]^2}{N - 1}} \quad (6.2)$$

Donde S_d es la desviación estándar y N el número de datos considerados.

La metodología de buscar una estructura en el bloque oculto de la red, se desarrolla de la siguiente forma:

El primer paso es considerar como número máximo de capas ocultas dos, posteriormente se evalúa el número de neuronas en cada capa, incluso la posibilidad de que la segunda capa oculta no exista (cero neuronas), el barrido se hace evaluando el error mientras cambia el número de neuronas de 5 a 100 y de 0 a 100 a intervalos de 5 en la primera y segunda capa respectivamente, para seleccionar aquella configuración que presente menor error.

En el segundo paso hacemos un nuevo barrido, mucho más fino alrededor de la configuración obtenida en el paso anterior, este lo hacemos a paso de uno, en un intervalo de 20 o 30 elementos. Nuevamente en este escogemos la configuración con menor error.

Figura 6.2. a. Matriz de error para la selección de el numero de capas y de neuronas de la red neuronal. b. Comportamiento del error al variar la semilla (primer experimento)

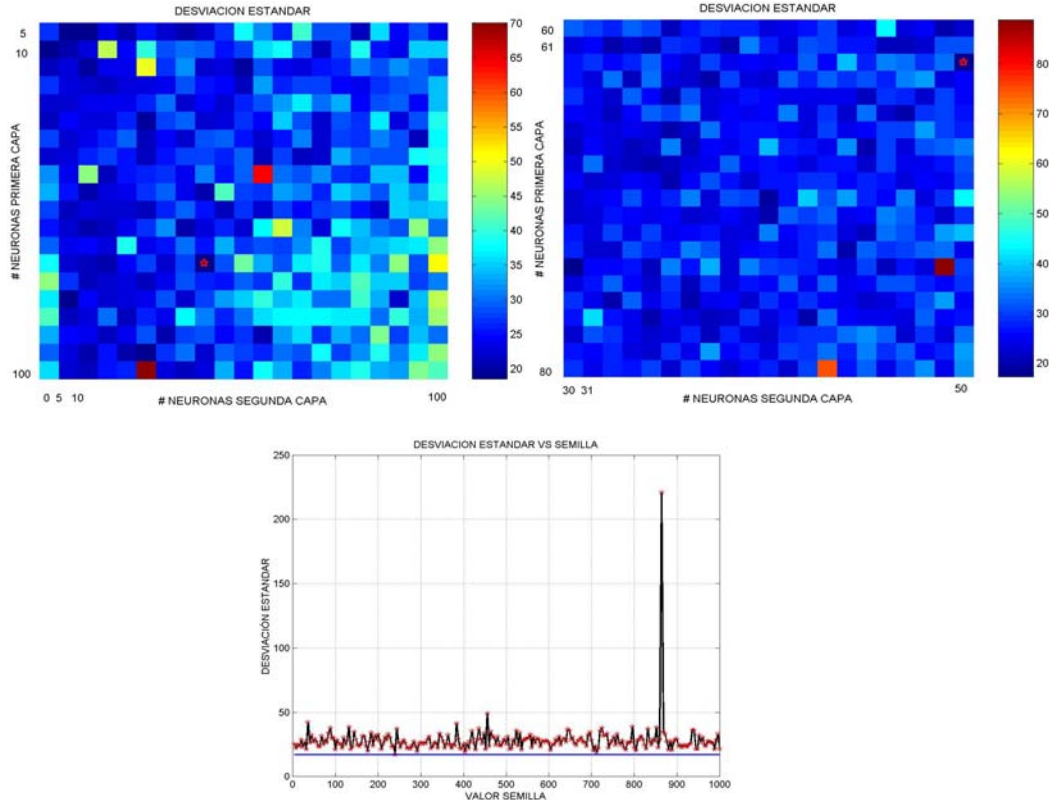
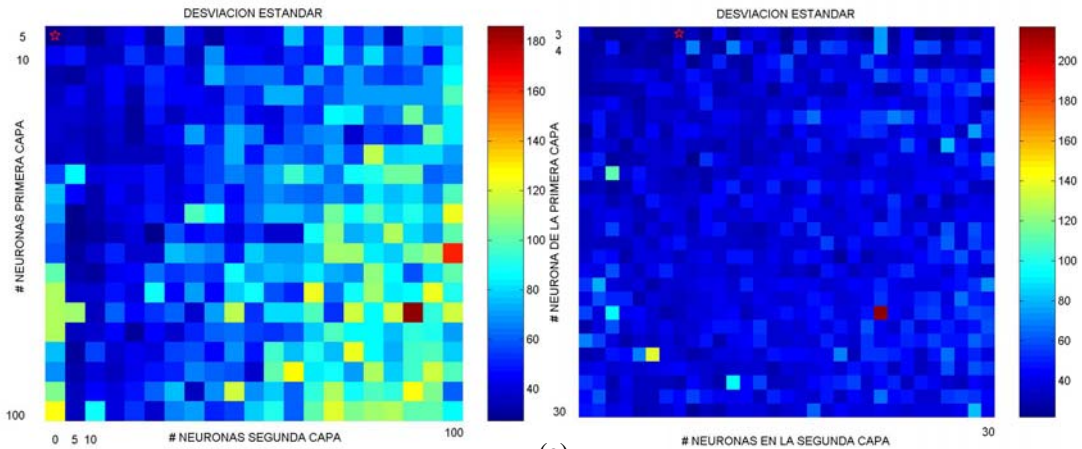
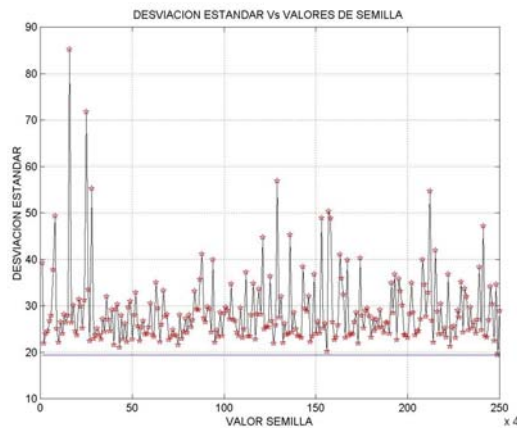


Figura 6.3. a. Matriz de error para la selección de el numero de capas y de neuronas de la red neuronal. b. Comportamiento del error al variar la semilla (segundo experimento)



(a)



(b)

En el último paso hacemos un barrido de los valores de semilla, la cual se encarga de generar los valores aleatorios de los pesos de la red. Esto con el propósito de posicionarnos en diferentes puntos de la superficie de error, evitando caer en mínimos locales. Lógicamente en este paso solo nos permite llegar a una configuración de pesos de la red que nos proporcione menor error, más no el cambio en su estructura oculta (figura 6.2 y 6.3).

De esta forma obtenemos la red que mejor responde a la solución del problema, que es determinar el peso de ganado bovino, de raza cebú con alto porcentaje de pureza, en la

forma más precisa. Un resumen de los resultados obtenidos en la implementación y entrenamiento de las redes neuronales artificiales se presenta en la tabla 6.1 y 6.2.

Tabla 6.1. Configuración y error de las redes neuronales. (primer experimento).

	ESTRUCTURA Y DESVIACIÓN ESTANDAR DE LAS REDES				
	L1: 5 a 100 L2: 0 a 100		L1: 60 a 80 L2: 30 a 50		
	Configuración	Error (Kg)	Configuración	Error (AS) (Kg)	Error (DS) (Kg)
MACHOS/HEMBRAS	6:70:40:1	18.55	6:62:50:1	17.25	17.21
MACHOS	6:60:25:1	19.95	6:60:25:1	19.95	14.55
HEMBRAS	6:100:10:1	14.34	6:90:13:1	13.59	12.75

Tabla 6.2. Configuración y error de las redes neuronales. (segundo experimento).

	ESTRUCTURA Y DESVIACIÓN ESTANDAR DE LAS REDES				
	L1: 5 a 100 L2: 0 a 100		L1: 3 a 30 L2: 0 a 30		
	Configuración	Error (Kg)	Configuración	Error (AS) (Kg)	Error (DS) (Kg)
MACHOS/HEMBRAS	4:5:1	27.16	4:3:7:1	21.94	19.35
MACHOS	4:5:5:1	32.32	4:5:2:1	23.65	23.65
HEMBRAS	4:5:5:1	17.98	4:3:15:1	14.06	14.06

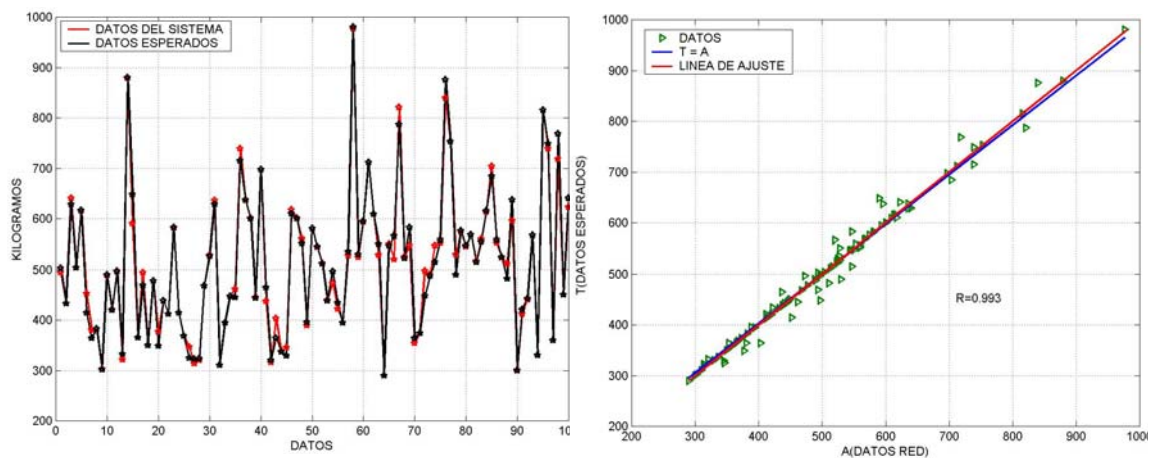
Podemos apreciar en las tablas que las redes que mejor comportamiento tienen frente al error (desviación estándar) son las que fueron obtenidas en el segundo barrido, al mismo tiempo el proceso de búsqueda de la semilla (DS) en la mayoría de los casos logra con éxito disminuir el error. Con respecto al sexo de los especímenes, podemos decir que a pesar de presentarse menor error en la red entrenada para las hembras, la red entrenada para todos los datos (Machos/Hembras) tiene un buen comportamiento tanto en el primer como en el segundo experimento, permitiendo al sistema ser invariante al sexo del animal. Sin embargo en el segundo experimento a pesar de transformar las entradas con el propósito de mejorarlas no se logró superar los resultados presentados por las redes cuyas entradas no fueron

transformadas. En este caso el proceso de transformación de las entradas por componentes principales no cumplió con las expectativas. Este método es exitoso cuando el número de entradas de la red es elevado, la cual no era nuestro caso.

Por tal razón la red neuronal seleccionada para la solución de este problema es la que presenta la configuración 6:62:50:1 y semilla 240. Vale la pena aclarar que sería una red neuronal con 4 capas dos ocultas una de entrada y una de salida, con 6 entrada en la capa de entrada, 62 y 50 neuronas con función de transferencia sigmoideal en las capas ocultas y una neurona en la salida con función de transferencia lineal.

Esta red se comporta de una manera bastante adecuada a todos los datos, se puede medir su comportamiento no solo con la desviación estándar, si no con el coeficiente de correlación entre las salidas de la red y los valores esperados, cuyo valor es de 0.993; un coeficiente de correlación superior al obtenido si usáramos un estimador lineal a partir del área (0.9899). Es claro que si el funcionamiento de la red fuera perfecto el coeficiente de correlación tendría valor 1.

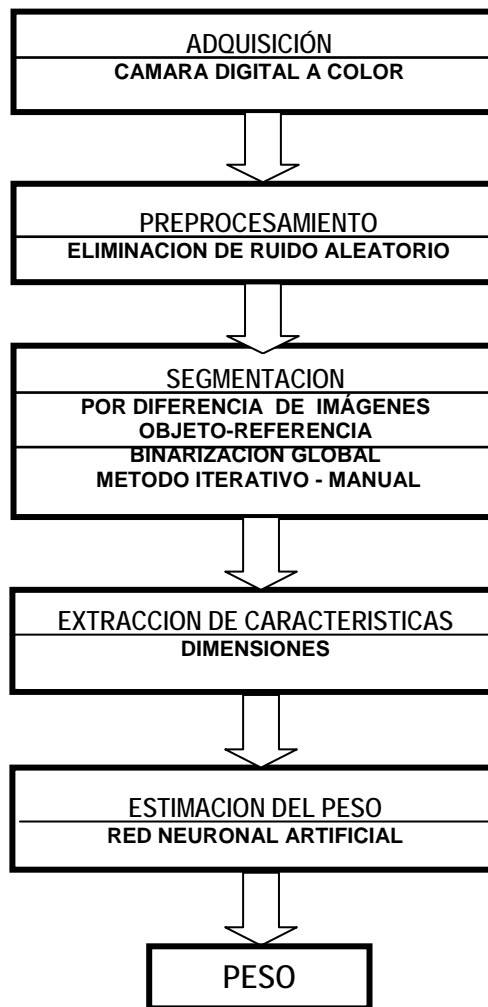
Figura 6.4 Comportamiento de los datos de salida de la red frente a los datos esperados.



7. CONCLUSIONES

En este trabajo se implementó un mecanismo que permite medir el peso de ganado bovino (raza cebú con alto grado de pureza) mediante un procedimiento basado en la visión artificial.

El mecanismo implementado está constituido por las etapas mostradas en el siguiente diagrama:



El sistema consta de: una etapa de adquisición realizado por una cámara a color, una etapa de preprocesamiento en la cual se atenúa el ruido aleatorio introducido en gran parte por la tarjeta digitalizadora de video, una de segmentación en donde se aísla el objeto de interés, otra de extracción de características o dimensiones que tenga alta correlación con el peso, y por última una de estimación que se encarga de estimar el valor del peso en kilogramos del animal presente en la escena de entrada.

Se implementaron y evaluaron dos metodologías de segmentación, una de diferencia de imágenes objeto-referencia y otra de diferencia de imágenes de secuencia. También se implementaron y se evaluaron varios métodos de binarización automática a nivel global y local (en éste solo se implementó búsqueda de umbral por método iterativo), basados en la búsqueda de umbral manual, por método iterativo, y método de manipulación de histograma.

El método de segmentación que mejor comportamiento frente al error tiene es una estrategia combinada semiautomática, que consiste en binarizar la imagen en una etapa inicial con el método de diferencia de imágenes objeto referencia y método de binarización iterativo, finalizando con el ajuste manual del contorno.

Se implementó un procedimiento de extracción de características o dimensiones del animal desde de la imagen, que es independiente de la orientación del objeto. Este proceso está basado en la transformada de Hotelling. El número de características extraídas es de 6, las cuales fueron reducidas por análisis de componentes principales (4), pero en esta configuración el desempeño del sistema fue inferior al alcanzado cuando se utilizaron las características no transformadas (6).

El proceso de estimación del peso del animal en kilogramos esta basado en la implementación de una red neuronal artificial tipo feedforward, con una configuración 6:62:50:1, con funciones de transferencia sigmoideal en las capas ocultas (62:50) y lineal (1) a la salida. Esta red permite asociar las características obtenidas sobre la imagen al peso del animal en kilogramos con una desviación estándar de 17.21 Kg y un coeficiente de correlación 0.993.

7.1 SUGERENCIAS PARA TRABAJOS FUTUROS

- El número de datos utilizado no contiene un número de objetos estadísticamente representativos. Los resultados podrían ser más confiables si se contara con un mayor número de muestras adquiridas. Los datos podrían ser adquiridos por réplicas sobre un número pequeño de animales; quiere decir, que a cada animal se le puede tomar varias veces medidas (imágenes y peso en kilogramos) en distinta épocas (edades del animal) aumentando así el número de datos.
- Es claro que la segmentación es uno de los problemas que más dificultad ofrece en esta investigación. Por tal razón este problema se convierte en una posibilidad para implementar nuevas alternativas de solución. Lo importante es que el ambiente de trabajo en la captura de las imágenes debe ser el adecuado (dificultad que se presentó en esta investigación debido a disposición y disponibilidad del campo de trabajo, que impedía el control de las variables para tener un ambiente propicio) de tal forma que aumente el contraste entre el fondo y el objeto, por ejemplo con iluminación controlada, o cambiando la sensibilidad de la cámara al infrarrojo (cámara térmica), o implementando técnicas de segmentación aun más adecuadas.
- Considerar la posibilidad de implementar al sistema otras vistas (por ejemplo la vista lateral) para determinar su contribución en el rendimiento del sistema.
- La introducción y evaluación de nuevas dimensiones o características que permitan aumentar la precisión del sistema en la determinación del peso en kilogramos del animal.
- Evaluar la posibilidad de implementar este procedimiento en tiempo real, ya que este tipo de proceso, entraría a formar parte de una posibilidad más para la obtención del peso en

kilogramos, con disminución radical de estrés, maltrato y posiblemente tiempo durante la pesa del animal.

- El número de referencias sobre trabajos afines son bastante escasos, solo se presentan algunas en Europa. Lo que convierte este trabajo en pionero en Latinoamérica, en cuanto a su aplicación se refiere y a nivel mundial debido al número de características usadas en la función de estimación y el espécimen utilizado.

8. BIBLIOGRAFIA

- [1] Gonzalo Téllez Iregui. Sistema de producción pecuaria. Ed. McGraw-Hill. 1990.
- [2] <http://grandin.com/spanish/Acechar.como.predador.html>
- [3] Smith, R. A. and Turner, M.J.B. Electronic aids for use en fat stock weighing. J. Agric. Eng. Res., 19:299-311. 1990.
- [4] Filby, D.E. and Turner, M.J.B. A walk-through weigher for dairy cows. J. Agric. Eng. Res., 24: 67-78. 1979.
- [5] Turner. M.J.B. Performance monitoring of animals using on-line computer in animal production, Occas. Publ. Br. Soc. Anim. Prod., Anim. Prod., 5: 41-46. 1981.
- [6] Slader. R. W. and Gregory. A.M. An automatic feeding and weighing system for ad lib. Fed pigs. Comput. Electron. Agric. 3: 157-170. 1974.
- [7] Phillips, R.W. and Dawson, W.M. A study of methods for obtaining measurements of swine. Anim. Prod. 29: 93-99. 1936.
- [8] Klant, G. And Glende, P. Tierkörpermabe bei Schweinein als Grundlage für tiergerechte Standund Buchtenkonstruktionen. Tierzucht, 29:420-422. 1975.
- [9] Schofield, C.P. Evaluation de image analysis as a means of estimating the weight of pigs. J. Agric. Eng. Res., 47:287-296. 1990.
- [10] Minagawa, H. And lechikawaT. Measurement of pig weights by an image analysis. ASAE Paper 92-0000. 1992.
- [11] Schofield, C.P., Addicott, A.L., Alloway, L., Bryant, F., Cameron, M. And Dows, C. "Finding a way of obtaining the mass of pig without disturbing the pigs". Div. Note DN/1520, Engineering Education Scheme Student Project Report, Silsoe Research Institute, Silsoe, Bedford, approx. 100 pp. 1989.

- [12] Tillet. R. And Marchant, J.A. Model-based image procesing for characterizing pigs in scenes. SPIE. Optics Agric., 1379: 201-208.
- [13] Marchant, J.A. and Shofield, C.P. Pigs, Snakes and Maggots. Div. Note, Silsoe Research institute, Silsoe, Bedford. 1990.
- [14] Delate. J.J. and Babu, R. Détermination d'équations baryométriques sur des porcs rustiques en milieu tropical. J. Rech. Porcine Fr., 22:35-42. 1990.
- [15] Nabil Brandl, Erik Jorgensen. Determination of live weight of pigs from dimensions measured using image analysis. Comput. Electron. Agric.. 15: 57-72. 1996.
- [16] Acebedo A., Luis I., Ramos., B., Luis F. Estimación del peso vivo en base a medidas corporales en ganado Cebú comercial.
- [17] Neural Network toolbox, For use with MATLAB, howard Demuth, Mark Beale. 1998.
- [18] Hagan, M.T. , H. B. Demuth. And M. H. Beale. Neural Network Design. Boston. MA: PWS publishing. 1996.
- [19] Dennis, J.E., and R. B. Schnabel, Numerical Methods for Unconstrained optimization and Nonlinear Equations, Englewood Cliffs. N. J. Prentice-Hall. 1993.
- [20] Battiti, R. " Firts and Second order methods for learning: Between steepest decent and Newton's method", Neural computation, Vol. 4, no. 2, pp. 141-166, 1992.
- [21] Foresee, F. D., and M.T. hagan, "Gauss.neuton aproximation to Bayesian Regularization, Proceeding of the 1997 International joint Conference on Neural Networks, Pages 1930-1935, 1997.
- [22] Mackay, D. J. C., "Bayesian Interpolation" , Neural computation, Vol 4, no. 3, pp. 415-447, 1992.
- [23] Hagan, M.T. and M. Menhaj, "Training feedforward networks with the Marquardt algorithm" IEEE transactions on Neural Networks. Vol. 5 no. Pp. 989-993. 1994.

ANEXO 1. Datos Registrados Experimentalmente

Nº	PESO (Kg)	EDAD (Dias)	SEXO	Nº	PESO (Kg)	EDAD (Dias)	SEXO
1	503	595	F	51	662	663	F
2	433	495	F	52	644	644	F
3	628	759	F	53	495	495	F
4	504	498	F	54	496	471	F
5	616	704	F	55	434	329	M
6	414	439	F	56	395	338	M
7	364	340	F	57	535	447	M
8	382	341	F	58	980	1037	M
9	302	282	F	59	530	426	M
10	490	419	M	60	594	552	M
11	420	338	M	61	712	656	M
12	496	405	M	62	609	538	M
13	332	298	M	63	551	600	F
14	880	991	M	64	290	295	F
15	648	591	M	65	547	712	F
16	365	360	F	66	566	556	M
17	469	539	F	67	787	991	M
18	350	357	F	68	522	519	M
19	478	566	F	69	582	561	M
20	349	338	F	70	364	278	M
21	438	399	F	71	374	302	F
22	412	373	F	72	448	482	F
23	582	527	M	73	487	534	F
24	414	327	M	74	515	588	F
25	368	279	M	75	559	609	F
26	325	286	F	76	875	857	M
27	322	316	F	77	753	723	M
28	323	306	F	78	490	399	M
29	468	632	F	79	576	624	F
30	527	589	F	80	547	524	F
31	628	865	F	81	568	659	F
32	310	365	F	82	515	612	F
33	395	371	M	83	555	670	F
34	447	367	M	84	615	827	F
35	445	485	M	85	684	547	M
36	715	680	M	86	554	423	M
37	637	630	M	87	525	412	M
38	600	607	M	88	482	322	M
39	444	421	F	89	637	538	M
40	697	697	M	90	300	304	F
41	465	449	F	91	421	464	F
42	320	297	F	92	443	487	F
43	364	271	M	93	567	914	F
44	337	276	M	94	330	356	F
45	329	322	F	95	815	811	M
46	610	779	F	96	749	890	M
47	600	801	F	97	360	287	M
48	552	667	F	98	768	1066	M
49	396	435	F	99	450	611	F
50	580	662	F	100	640	770	F

ANEXO 2. Errores por el método de segmentación por diferencia de imágenes objeto referencia.

	UMBRAL GLOBAL						UMBRAL LOCAL	
	M. MANUAL		M. ITERATIVE		M. HISTOGRAM		M. ITERATIVE	
	ERROR 1	ERROR 2	ERROR 1	ERROR 2	ERROR 1	ERROR 2	ERROR 1	ERROR 2
1	6,7	0,7	21,2	0,1	2,8	9,0	15,6	0,1
2	10,3	1,8	28,4	0,6	4,2	8,1	34,0	0,5
3	13,4	1,4	27,8	0,9	8,0	2,0	29,4	1,0
4	25,8	1,5	39,4	0,6	8,3	5,6	56,9	0,7
5	15,3	1,0	23,1	0,5	7,7	1,8	20,6	0,6
6	8,3	2,3	21,3	0,8	4,7	7,2	17,3	1,5
7	3,2	3,8	16,1	1,3	3,3	7,9	11,6	2,7
8	10,9	1,0	19,9	0,6	7,4	4,0	17,0	0,7
9	13,6	2,8	39,5	1,4	3,4	30,7	37,2	1,6
10	16,4	2,5	29,4	1,7	7,8	6,8	22,2	1,8
11	17,6	6,7	31,4	1,2	12,2	80,5	26,4	1,2
12	9,5	0,4	26,9	0,1	5,5	3,5	14,3	0,2
13	6,9	1,7	19,9	0,0	2,7	9,1	40,2	0,1
14	22,3	0,3	50,6	0,0	12,7	0,8	39,9	0,4
15	14,1	0,5	48,4	0,0	6,4	10,7	37,0	0,0
16	8,9	0,6	25,2	0,3	4,0	8,2	17,3	0,3
17	24,7	0,1	38,9	0,1	16,4	8,8	67,3	0,1
18	10,8	0,6	47,0	0,1	4,5	4,4	38,4	0,3
19	6,8	1,0	14,2	0,1	3,0	6,3	11,0	0,1
20	11,2	0,1	35,8	0,0	8,6	1,7	33,7	0,3
21	14,1	0,0	36,7	0,0	8,1	0,3	31,6	0,3
22	12,6	0,4	23,6	0,0	7,2	3,9	20,9	0,0
23	11,0	0,2	31,9	0,0	8,1	2,8	70,9	0,1
24	10,5	0,3	25,9	0,0	5,9	1,8	38,1	0,1
25	7,5	0,1	19,7	0,0	4,4	1,5	9,9	0,0
26	10,2	0,1	23,6	0,0	6,3	2,0	13,5	0,2
27	4,6	0,9	14,5	0,4	1,2	13,3	11,3	0,5
28	68,9	3,1	74,1	0,6	40,8	264,9	75,6	0,5
29	76,8	145,4	73,7	189,0	67,1	318,8	98,2	1,5
30	21,0	0,0	39,5	0,0	8,2	0,7	27,4	0,2
31	10,1	0,6	18,6	0,1	15,0	1,3	53,7	0,5
32	5,3	0,1	27,6	0,0	3,0	0,6	13,3	0,1
33	9,2	0,1	21,0	0,0	6,0	0,7	37,4	0,1
34	14,6	0,1	26,1	0,0	21,0	0,4	32,2	0,1
35	58,8	11,6	58,4	13,6	24,1	268,8	96,1	5,0
36	41,6	0,0	45,8	0,0	38,8	0,1	76,4	0,0
37	23,8	0,1	44,4	0,0	15,9	0,4	44,0	0,1
38	12,8	1,6	29,4	0,0	9,6	4,0	21,7	39,0
39	21,6	0,0	34,2	0,0	9,9	0,5	23,2	0,2
40	17,3	0,1	38,2	0,0	12,4	0,9	41,5	0,1
41	11,6	0,4	27,1	0,0	4,5	2,5	23,2	0,0
42	18,9	0,0	30,6	0,0	14,8	0,6	27,7	0,0
43	13,2	0,1	25,1	0,0	9,9	0,8	21,5	0,1
44	14,7	0,6	32,9	0,0	11,5	0,9	26,6	0,0
45	7,5	0,0	15,3	0,0	5,8	0,6	35,3	0,1
46	22,7	0,0	30,3	0,0	18,6	0,4	48,4	0,0
47	16,7	0,2	24,4	0,0	6,9	0,9	21,1	0,1
48	11,9	0,1	24,1	0,0	6,8	0,7	18,3	0,1
49	9,7	0,1	18,6	0,0	6,4	0,9	13,1	0,1
50	18,7	0,0	25,8	0,0	7,0	1,1	21,0	0,1

	UMBRAL GLOBAL						UMBRAL LOCAL	
	M. MANUAL		M. ITERATIVE		M. HISTOGRAM		M. ITERATIVE	
	ERROR 1	ERROR 2	ERROR 1	ERROR 2	ERROR 1	ERROR 2	ERROR 1	ERROR 2
51	10,3	0,1	25,5	0,0	5,5	0,7	44,5	0,1
52	11,4	0,1	23,1	0,0	7,8	0,4	16,7	0,1
53	8,8	0,2	20,2	0,0	6,1	1,0	10,9	0,2
54	11,8	0,0	25,5	0,0	7,0	0,6	16,3	0,0
55	17,9	0,0	23,1	0,0	14,8	0,4	16,5	0,1
56	12,1	0,1	22,2	0,0	9,5	0,7	16,2	0,0
57	11,7	0,0	33,9	0,0	7,8	0,6	43,4	0,0
58	50,8	0,0	75,5	0,0	42,6	0,5	73,9	0,0
59	59,9	0,0	67,8	0,0	51,8	0,6	70,2	0,0
60	49,1	0,2	55,9	0,0	37,5	0,1	59,1	0,0
61	56,6	0,0	61,7	0,0	52,8	0,3	63,4	0,0
62	12,8	0,4	26	0,0	8,9	1,0	24,9	0,0
63	20,8	0,3	25,4	0,0	18,9	0,7	22,9	0,2
64	97,7	19,4	96,9	202,1	96,2	416,0	98,2	3,3
65	8,0	0,1	19,1	0,0	6,1	0,6	10,6	0,0
66	42,1	0,0	51,1	0,0	20,9	1,3	54,7	0,0
67	22,7	0,0	30,0	0,0	17,3	0,3	22,9	0,2
68	15,1	0,0	19,4	0,0	13,1	0,8	20,9	0,0
69	27,8	0,0	40,0	0,0	19,1	0,7	35,2	0,0
70	15,2	0,0	22,1	0,0	14,0	0,7	9,6	0,1
71	9,6	0,2	25,9	0,0	16,8	1,8	16,3	0,0
72	28,6	0,1	42,4	0,0	15,0	9,1	55,7	0,1
73	18,9	0,1	29,2	0,0	16,1	0,5	23,0	0,0
74	18,8	0,0	24,3	0,0	10,2	0,3	23,5	0,0
75	7,4	0,1	15,6	0,0	5,7	0,6	10,5	0,0
76	14,4	0,0	29,1	0,0	11,7	0,3	18,6	0,0
77	34,5	0,0	41,3	0,0	32,9	0,4	37,0	0,1
78	8,1	0,0	20,8	0,0	14,1	0,6	52,9	0,0
79	3,4	0,1	9,7	0,0	1,9	1,1	24,2	0,1
80	23,4	0,1	64,6	0,0	6,1	8,8	85,4	0,1
81	6,3	0,0	19,4	0,0	4,3	0,5	33,6	0,1
82	10,6	0,1	20,9	0,0	7,0	0,8	17,2	0,1
83	2,3	0,2	6,5	0,0	1,4	1,9	5,9	0,7
84	8,6	0,1	16,7	0,0	4,6	0,7	14,9	0,1
85	11,8	0,1	27,2	0,0	5,5	0,5	21,6	0,0
86	15,4	0,1	29,8	0,0	7,4	1,4	27,5	0,0
87	16,3	0,2	32,3	0,0	6,3	2,5	26,7	0,4
88	16,1	0,0	26,3	0,0	8,5	14,1	40,6	0,0
89	32,4	0,0	45,1	0,0	16,0	0,7	47,7	0,0
90	28,1	1,7	37,3	0,6	3,8	9,2	51,3	0,0
91	100,0	39,2	99,7	110,1	0,0	429,0	100,0	3,9
92	99,9	106,0	99,4	245,7	92,8	360,0	100,0	1,0
93	88,4	134,2	99,6	0,0	6,2	303,0	99,0	1,0
94	70,8	391,3	71,6	365,9	0,0	528,0	97,0	39,5
95	0,0	213,0	61,7	180,7	0,0	225,0	86,0	101,8
96	97,5	162,0	98,8	106,5	84,2	247,0	98,0	68,5
97	93,7	311,0	93,6	313,7	0,0	498,0	100,0	17,0
98	80,7	188,0	86,4	166,0	0,0	251,0	88,1	84,0
99	27,8	0,2	52,4	0,0	16,1	6,6	57,3	0,0
100	11,8	0,2	18,2	0,0	9,7	0,8	16,8	0,0
med	24,03	17,7	36,8	19,1	14,0	44,5	38,9	3,9

ANEXO 3. Errores por el método de segmentación por diferencia de secuencias de imágenes.

	UMBRAL GLOBAL		UMBRAL LOCAL	
	M. MANUAL		M. ITERATIVE	
	ERROR 1	ERROR 2	ERROR 1	ERROR 2
1	72.4	31,8	67,6	20,4
2	2.3	45,3	91,1	25,8
3	0.4	11,9	3,5	8,1
4	83.3	3,2	80,2	1,7
5	3.4	5,9	2,7	3,0
6	3.0	21,2	77,7	9,8
7	1.2	19,3	76,3	19,0
8	2.8	3,9	7,2	3,2
9	3.8	5,4	42,1	5,3
10	0.3	10,0	4,1	9,1
11	0.8	7,5	3,8	6,3
12	1.4	2,0	8,9	1,7
13	2.8	23,8	11,8	14,3
14	66.4	8,2	87,4	2,8
15	58.3	42,7	65,1	35,3
16	56.8	81,7	76,2	53,1
17	1.0	71,3	71,4	56,1
18	1.3	89,6	66,1	70,9
19	2.2	12,6	4,9	15,8
20	2.3	64,5	5,5	43,8
21	3.5	8,6	6,8	7,1
22	1.7	29,9	13,4	18,2
23	3.0	27,2	6,0	7,3
24	0.5	41,3	9,7	9,9
25	1.1	48,5	4,4	9,1
26	3.6	8,6	6,1	7,7
27	1.3	20,9	4,5	28,5
28	43.3	21,7	50,7	10,7
29	60.3	17,7	60,1	12,9
30	0.9	6,6	9,7	5,1
31	0.5	15,9	4,5	10,0
32	2,2	10,2	5,3	8,3
33	1.5	14,8	3,8	11,0
34	1.0	13,3	5,6	8,1
35	69.6	9,5	74,1	9,5
36	54.6	5,5	74,0	3,6
37	0.5	31,6	4,6	5,7
38	0.2	39,7	2,4	16,1
39	1.2	11,5	7,4	9,3
40	1.3	28,1	9,3	15,8
41	0.1	18,6	53,0	14,2
42	1.2	11,9	10,3	9,7
43	1.8	26,4	4,7	9,1
44	1.6	11,1	7,4	8,6
45	1.7	9,6	4,2	8,7
46	1.8	18,7	3,7	12,8
47	2.4	6,2	4,7	4,6
48	2.1	7,4	6,0	6,0
49	0.6	15,5	5,7	10,2
50	0.3	9,2	3,8	6,6

	UMBRAL GLOBAL		UMBRAL LOCAL	
	M. MANUAL		M. ITERATIVE	
	ERROR 1	ERROR 2	ERROR 1	ERROR 2
51	2,7	7,7	75,2	4,7
52	1,1	11,3	4,6	7,8
53	0,9	13,3	3,4	9,9
54	1,8	14,0	79,4	6,3
55	0,8	7,5	4,1	6,6
56	0,3	21,4	6,1	16,5
57	0,9	8,1	67,9	5,5
58	4,8	6,4	65,5	2,4
59	9,4	13,2	39,9	9,2
60	13,7	8,3	72,7	3,3
61	19,1	7,2	55,5	5,3
62	0,2	19,1	6,4	6,8
63	0,9	7,9	3,2	6,4
64	53,6	31,7	55,8	18,4
65	0,4	10,4	1,7	8,3
66	3,4	5,7	13,9	3,0
67	0,8	8,5	73,8	6,0
68	0,3	33,2	4,1	4,6
69	1,2	9,9	9,1	7,6
70	1,1	8,5	5,7	6,0
71	0,9	53,2	2,9	44,2
72	0,3	12,7	1,6	11,6
73	1,9	27,3	3,3	10,6
74	0,9	7,5	78,8	5,6
75	0,8	6,1	2,8	4,8
76	1,0	8,8	5,7	6,3
77	2,0	10,9	60,2	7,6
78	0,6	15,2	4,4	11,4
79	0,1	13,9	4,3	4,5
80	7,3	7,8	67,1	4,3
81	0,7	6,5	93,0	4,6
82	0,9	16,1	81,9	10,7
83	0,1	15,6	85,2	14,4
84	1,1	7,9	6,1	6,5
85	0,9	7,6	3,3	5,5
86	3,0	12,0	77,8	5,6
87	61,1	12,3	68,8	5,9
88	53,9	16,8	69,1	12,9
89	0,9	6,7	63,1	4,1
90	3,6	6,0	87,5	3,8
91	70,8	12,9	77,5	5,4
92	64,6	10,6	90,2	4,9
93	65,8	10,7	71,6	6,1
94	63,8	22,9	83,5	15,7
95	57,6	34,9	77,6	6,4
96	67,4	29,7	94,3	5,9
97	64,4	39,0	88,3	25,8
98	66,6	27,8	99,3	13,2
99	1,4	43,0	76,4	17,3
100	1,7	19,8	6,7	13,6
med	15,67	18,8	36,0	11,5