



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

197

Zejeu

FACULTAD DE INGENIERIA

SEGMENTACION DE IMAGENES PARA LA IDENTIFICACION
DE OBJETOS INDUSTRIALES

T E S I S

QUE PARA OBTENER EL TITULO DE:

INGENIERO MECANICO - ELECTRICISTA
(AREA ELECTRICA - ELECTRONICA)

P R E S E N T A:

MARCO ANTONIO VILLANUEVA CASTELLANOS



DIRECTOR: DR. FRANCISCO GARCIA UGALDE

MEXICO, D. F.

1995

FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

**A MIS PADRES:
QUIENES CON SU CARIÑO Y APOYO ME MOTIVARON PARA
CUMPLIR UNA DE LAS METAS MAS IMPORTANTES DE MI VIDA,
LA CUAL SIGNIFICA LA CULMINACION DE SUS ESFUERZOS.**

**A MI HERMANO MAURICIO:
QUIEN SIEMPRE ESTUVO DISPUESTO A EXPLICARME Y AYUDARME**

**A MI HERMANA REBECA:
CON CARIÑO.**

**A MI DIRECTOR DE TESIS:
DR. FRANCISCO GARCIA UGALDE**

**A TODOS MIS FAMILIARES, AMIGOS Y PROFESORES QUIENES DE
ALGUNA FORMA HAN CONTRIBUIDO EN MI DESARROLLO PROFESIO-
NAL.**

INDICE

1	INTRODUCCION	3
2	INTRODUCCION A LOS SISTEMAS DE VISION	
21	SISTEMAS DE VISION	6
22	VISION Y AUTOMATIZACION DE FABRICAS	7
23	PARAMETROS DE FUNCIONES COMPARATIVAS	7
23.1	ADAPTABILIDAD	8
23.2	DECISIONES REALIZADAS	9
23.3	CALIDAD DE LA MEDIDA	10
23.4	DOS Y TRES DIMENSIONES	10
23.5	CONSIDERACIONES ECONOMICAS	11
24	VELOCIDAD DE RESPUESTA	11
25	RESPUESTA ESPECTRAL	12
26	COMPONENTES DE UN SISTEMA DE VISION	13
27	COMPARACION DE LA VISION HUMANA CONTRA EL APARATO DE VISION	15
3	ADQUISICION Y REPRESENTACION DE UNA IMAGEN	
31	ADQUISICION DE IMAGENES	17
31.1	ILLUMINACION	17
31.2	FORMACION DE LA IMAGEN Y ENFOQUE	20
31.3	DETECCION DE LA IMAGEN	20
32	PROCESAMIENTO	21
33	CONCEPTOS FUNDAMENTALES DEL PROCESAMIENTO DE IMAGENES	
33.1	PIXEL	22
33.2	LOCALIZACION DE PIXELES	22
33.3	ESCALA O NIVELES DE GRIS	24
4	SEGMENTACION DE IMAGENES	
4.1	INTRODUCCION	25
4.2	DETECCION DE DISCONTINUIDADES	25
4.3	DETECCION DE LINEAS	26
4.4	DETECCION DE BORDES	27
4.4.1	METODO DE UMBRAL PARA LA DETECCION DE BORDES	28
4.4.2	GRADIENTE Y NORMAL	31
4.4.3	OPERADORES GRADIENTE	33
4.4.4	GENERALIZACION DE OCHO MASCARAS DEL OPERADOR GRADIENTE	38
4.5	TRANSFORMADA HOUGH	41
4.6	TECNICAS DE SEGMENTACION	42
4.6.1	AMPLITUD DE UMBRAL	42
4.6.2	COMPONENTES ROTULADOS	43
4.6.3	PROPOSICION BASADA EN LIMITES	44
4.6.4	PROPOSICION BASADA EN REGION Y AGRUPAMIENTO	45

5 IMPLEMENTACION DE LOS METODOS DE IDENTIFICACION DE FORMAS	
5.1 INTRODUCCION	48
5.2 ANALISIS Y SELECCION DE LOS METODOS OPERADORES DETECTORES DE BORDES	48
5.2.1 ALGORITMOS DE PROGRAMACION	51
5.3 IMPLEMENTACION DE LOS METODOS PARA LA IDENTIFICACION DE FORMAS	67
5.3.1 ALGORITMOS DE PROGRAMACION	68
5.4 RESULTADOS DE LOS METODOS DE IDENTIFICACION DE FIGURAS	70
5.5 PRUEBAS DE ROBUSTEZ DE LOS METODOS DE IDENTIFICACION DE FIGURAS	71
5.6 APLICACION A OBJETOS INDUSTRIALES	79
6 ANALISIS DE LA IMPLANTACION DE UN METODO DE IDENTIFICACION DE FORMAS SOBRE UNA ESTRUCTURA DE DATOS MULTIRESOLUCION	
6.1 INTRODUCCION	95
6.2 ALGORITMOS DE DESCOMPOSICION Y RECONSTRUCCION EN DOS DIMENSIONES	96
6.3 INTRODUCCION A UNA ESTRUCTURA DE DATOS MULTIRESOLUCION	97
6.3.1 METODO PIRAMIDAL PASA BAJAS	98
6.3.2 METODO PIRAMIDAL PASA BANDA ORTOGONAL	101
6.4 IMPLANTACION DE UN METODO SOBRE UNA ESTRUCTURA DE DATOS MULTIRESOLUCION	103
6.5 COMPARACION DE LOS RESULTADOS OBTENIDOS EN MONORESOLUCION Y MULTIRESOLUCION	104
CONCLUSIONES	112
REFERENCIAS BIBLIOGRAFICAS	116

1 INTRODUCCION

El objetivo de esta tesis es la identificación de objetos industriales mediante la segmentación de imágenes. La localización de objetos es importante en muchas áreas del análisis de imágenes, pero es especialmente importante en aplicaciones de la industria tal como inspección automática y ensamblado. En general todas las partes necesitan ser chequeadas, ensambladas y medidas con alta precisión. Para la localización de objetos se puede aplicar un sistema de visión automático el cual proporciona una imagen digital a la que se le aplicaran algoritmos eficientes para realizar dicha localización e identificación. Desafortunadamente, la mayoría de los algoritmos no son lo suficientemente robustos en muchas aplicaciones, por ejemplo, las imágenes pueden ser "mutiladas" por agentes externos como fantasmas y ruido.

Un sistema de visión comprende todos los elementos necesarios para obtener la representación digital de una imagen visual, para modificar los datos y presentar una imagen digital de datos al mundo externo. Un sistema de visión comprende tres componentes principales: adquisición de imágenes, procesamiento y despliegue. El equipo para la adquisición de datos incluye la iluminación, la cámara y posiblemente el marco de recolección. El equipo de procesamiento incluye "hardware" y "software" en la unidad de visión y el equipo de salida es un sistema de interfase electrónica.

El primer paso en el análisis de imágenes es frecuentemente segmentar la imagen. El objetivo de la segmentación es separar los componentes de la imagen hasta subconjuntos que correspondan a los objetos físicos de la escena. Los componentes segmentados son usados en procesos superiores como son la interpretación y reconocimiento de las imágenes. La mayor dificultad en la segmentación es debida a que el interés es encontrar los límites de los objetos físicos. Los métodos de segmentación de imágenes asumen que los objetos tienen una superficie de suavidad homogénea que responde a regiones de intensidad constante o variación suave.

Para encontrar los límites de los objetos físicos (bordes), el método más simple es el de umbral, el cual proporciona una imagen binaria (negra y blanca) donde la imagen será blanca o negra si cada pixel excede un umbral T . Existen otros métodos de detección de bordes como son el operador gradiente y el operador compass, los cuales miden el valor del gradiente a lo largo de la imagen. Otro método pero para la detección de líneas consiste en aplicar la transformada

Hough a toda la imagen.

El desarrollo constante de las tecnologías de punta hacen necesario que las generaciones de profesionistas que año con año egresan de las distintas universidades, no solo lleven consigo los conocimientos generales impartidos en una aula de clase, sino que también se formen un criterio propio para así enfrentar y poder resolver cualquier situación que se presente, con los avances tecnológicos. En el caso de los ingenieros, en ellos recae una parte importante de dichos avances.

La finalidad de esta tesis es proporcionar las herramientas necesarias para realizar el reconocimiento e identificación de objetos industriales. En el presente trabajo, con el cual se culmina un primer ciclo en el desarrollo de mi vida profesional, no pretendo ser un erudito en la materia, sino aplicar conocimientos adquiridos durante mi estancia en la Facultad de Ingeniería, en un proyecto práctico para ser aplicado en la industria, así como proporcionar una referencia valiosa para estudiantes e investigadores en el campo de procesamiento de imágenes, en el área de la segmentación de imágenes.

El tema de la tesis se titula "Segmentación de imágenes para el reconocimiento de objetos industriales" y consta de 7 capítulos, los cuales se estructuran de la siguiente manera:

El capítulo 1 es esta introducción la cual proporciona el objetivo del proyecto así como una introducción al problema de segmentación de imágenes.

El capítulo 2 proporciona la información de un sistema de visión, esto con la finalidad de mostrar las características principales de un sistema de visión que pudiera existir en la industria o que pudiera ser aplicado a la misma.

En el capítulo 3 se analiza la adquisición y la representación de una imagen, el cual proporciona la forma en que se adquiere una imagen y los conceptos fundamentales de procesamiento de imágenes.

El capítulo 4 analiza los métodos de detección de bordes y las técnicas de segmentación de imágenes, (Amplitud de umbral, Pixel rotulado, contorno siguiente y proposición basada en región y agrupamiento).

En el capítulo 5 se implementan métodos para la identificación de objetos aplicando métodos de detección de bordes y segmentación de imágenes, en este

mismo capítulo se muestran los resultados de dichos métodos.

El capítulo 6 proporciona la teoría de datos en multiresolución para implantar los mismos métodos del capítulo anterior pero en multiresolución, en este mismo capítulo se muestran los resultados en multiresolución y se realiza la comparación de los resultados en monoresolución y multiresolución. (En este capítulo se aplican los programas realizados por anteriores tesis para obtener la imagen en multiresolución)

En el capítulo 7 se muestran las conclusiones de este trabajo.

Resumiendo, el objetivo de esta tesis es implantar unos métodos de identificación de figuras en base a los métodos de detección de bordes y técnicas de segmentación de imágenes digitales. Así como también proporcionar una valiosa información para estudiantes e investigadores en el área de procesamiento de imágenes.

La identificación de objetos en la industria es muy importante porque se puede aislar o identificar ciertas piezas industriales entre otras. En esta tesis se explicará primero que es un sistema de visión, para poder ser aplicado en la industria, posteriormente se explicará la adquisición de imágenes así como también los conceptos fundamentales del procesamiento de imágenes, luego se analizarán los métodos de detección de bordes así como también las técnicas de segmentación de imágenes. En seguida se implantarán algunos métodos para la identificación y localización de figuras utilizando los métodos de detección de bordes y las técnicas de segmentación de imágenes, la implantación de estos métodos de identificación será hecha para sistemas en monoresolución y multiresolución. Por último se harán comparaciones de los sistemas en monoresolución y multiresolución y se darán las conclusiones a este trabajo.

Para la localización de figuras geométricas se realizará una exhaustiva identificación, ya que se podrá localizar cualquier figura entre otras, estas figuras podrán tener una rotación o no. Como una muestra de la aplicación del proyecto a la industria, se identificarán algunas piezas mecánicas como pueden ser: candados, pinzas, desarmadores y martillos.

2 INTRODUCCION A LOS SISTEMAS DE VISION

2.1 SISTEMAS DE VISION

La tecnología de los sistemas de visión e imágenes digitales es una multidisciplina en el sentido de que su campo usa conocimientos tradicionales de ingeniería y programación de computadoras para diferentes partes de los procesos. Los procesos pueden ser subdivididos dentro de las tres siguientes actividades:

- a) Obtención de la representación digital de una imagen.
- b) Empleo computacional de técnicas para procesos o modificaciones de imágenes de datos, y
- c) Análisis y uso de los resultados de los procesos para propósitos de dirección de robots o equipos de control automático, asegurando un nivel de calidad en la manufactura de los procesos, o soportando análisis estáticos en sistemas de Manufactura Asistida por Computadora (CAM).

Tres condiciones sobresalen necesariamente para extensas aplicaciones de una nueva tecnología, las cuales son:

- a) Hardware confiable a costo razonable.
- b) Hardware individual y conocimientos de programación aplicados a la tecnología, y
- c) Una necesidad o un problema que requieren una solución.

El amplio uso de robots y de tecnología CAD/CAM en el sector industrial, impone la necesidad de automatizar los procesos de adquisición visual en forma digital. Al mismo tiempo los cambios social, industrial y económico demandan la necesidad de incrementar la productividad en el sector de manufactura de la economía, mejorando los niveles de calidad y confianza en el producto final.

Las operaciones de manufactura fueron desarrolladas usando la capacidad de visión incluyendo un humano "en el ciclo o loop". La disponibilidad de visión para toda actividad del ser humano en el sector industrial fue dado por descontado. La introducción de robots en las fábricas hizo necesario la automatización de la función de "visión", desde entonces el humano fue eliminado del ciclo.

2.2 VISION Y AUTOMATIZACION DE FABRICAS

Las principales funciones que fueron desarrolladas por los sistemas de visión se encuentran dentro de las tres principales categorías:

- Control
- Inspección
- Entrada de Datos

Control, en su forma simple, se refiere a la determinación de la posición y generación de comandos apropiados para mecanismos de inicio o transformación de una acción.

Las aplicaciones de inspección en los sistemas de visión efectúan la determinación y/o cuantificación de parámetros tales como dimensiones de partes mecánicas, así como contornos, calidad de superficies, número de cavidades en una parte, presencia o ausencia de rasgos específicos.

La información de los atributos de calidad de un producto, así como el material y el diseño de la producción, pueden formar parte de una base de datos a partir de un sistema de visión. Este método de entrada de datos es altamente preciso y seguro, debido a que el elemento humano se elimina del ciclo. Al mismo tiempo, es extremadamente económico porque el dato ha sido incluido como parte de la inspección del proceso.

2.3 PARAMETROS DE LAS FUNCIONES COMPARATIVAS

El aparato de visión será comparado con la visión humana como un sistema total, en base a parámetros funcionales que son aplicables a la industria o procesos de manufactura. Las funciones de suma importancia en aplicaciones industriales son:

- Adaptabilidad

- Decisiones realizadas
- Calidad de la medida
- Capacidad de dos y tres dimensiones
- Consideraciones económicas

La capacidad de la mejora, como contraste a la reducción del costo conseguido por reemplazo de operadores humanos, es el elemento clave en el uso de sistemas de visión para mejorar la productividad.

2.3.1 ADAPTABILIDAD

La adaptabilidad es la capacidad de un sistema de ajustarse automáticamente o modificar sus operaciones de acuerdo a los parámetros del ambiente para conseguir el resultado deseado.

La capacidad de los aparatos de visión es muy rígida; se determina por el hardware y el software en el sistema. El sistema tendrá que repetir el proceso de visión con un alto grado de certidumbre.

El sistema de visión humano es altamente adaptable, porque imágenes para diferentes ángulos podrán ser utilizadas si hay obstrucción en el campo de vista, o las imágenes podrán ser obtenidas para diferentes rangos si una aplicación es necesaria para determinar detalles específicos. Los sistemas de visión humanos son dinámicos, porque la capacidad se determina por características físicas del ojo, así como del intelecto del individuo. Esto se ilustra en la Figura 2.1.

El intelecto incluye un número de artículos que no pueden ser modificados, tales como el aprendizaje, la asociación, y la percepción.

Los requerimientos de la mayor parte de los procesos de manufactura industrial son muy rígidos y fijos, como lo dictan las especificaciones del producto. De este modo, los costos de manufactura son directamente afectados por lo "ajustado de los requerimientos", esto es deseable para hacer medibles los niveles de precisión necesarios con un alto grado de certidumbre. Las características de no-adaptabilidad de los aparatos de visión, pueden ser una ventaja en aplicaciones de visión industrial que envuelven transformaciones de procesos idénticos en líneas de manufactura paralela, dado que proveen un alto grado de certidumbre medible.

Resumiendo, los sistemas de visión humana son mucho más adaptables que los sistemas de aparatos de visión. La adaptabilidad es el resultado de que el

cerebro humano en un ciclo puede modificar el sistema en la forma "on-line" (en línea), es decir, conforme ocurren los sucesos.

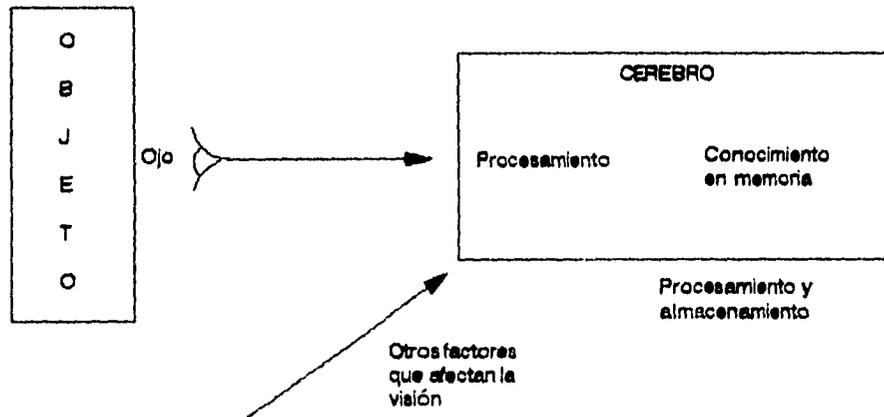


Figura 2.1 Características del Intelecto del Individuo

2.3.2 DECISIONES REALIZADAS

Para una decisión mejor, se requiere transformar alguna parte de la función de un sistema de visión en aplicaciones de automatización industrial.

Los sistemas de visión humana pueden ser efectivos en lugares donde se juzgan atributos para tareas en que se involucra la inspección, como son, el color, contorno, además de las capacidades de percepción e interpretación.

Los sistemas de aparatos de visión, requieren la cuantificación de parámetros medibles utilizados en la decisión de los procesos realizados. Mientras que el sistema de visión humano puede basar su decisión en términos relativos (tono claro u oscuro), por ejemplo el sistema del aparato de visión requiere de valores numéricos; tales como especificar el nivel de gris en un sistema dado, como un número de niveles de gris o especificando el número de píxeles entre dos puntos de referencia en el objeto.

Los sistemas de aparatos de visión podrán ser más consistentes que los sistemas de visión humana, basados en decisiones objetivas. La consistencia de la transformación de los sistemas humanos es grandemente impactada por la fatiga, factores ambientales, y las condiciones físicas del humano.

2.3.3 CALIDAD DE LA MEDIDA

La consistencia de resultados y el nivel de precisión son dos factores principales en la calidad de la medida. Los sistemas de aparatos de visión son claramente superiores sobre los sistemas de visión humana, en el caso de aplicaciones donde la medición se basa en la modificación de los datos de entrada.

Los sistemas de visión humana pueden discernir algunas cosas en el orden de diez a veinte niveles de gris, sin embargo es capaz de distinguir diferencias entre muchos más niveles en una base comparativa. Ayudas mecánicas pueden ser utilizadas para mejorar el número discernible de niveles.

La capacidad de los niveles de gris para los sistemas de aparatos de visión son relativas y limitadas por el número de bits disponibles para el código de representación entera de los niveles de gris. En un sistema de cuatro-bits, pueden existir como máximo 16 niveles de gris. Para un número binario de 8 bits, tenemos que usar la representación de 256 niveles de gris en el sistema de visión. Los procesos de 16 y 32 bits se hacen posibles incrementando el número de niveles de gris, sin embargo comúnmente hay pocos, si acaso algún sistema industrial con más de 256 niveles de gris.

Los sistemas de aparatos de visión no tienen errores aleatorios debidos a la "fatiga" y establecen niveles de funcionamiento que podrán ser constantes para todo propósito práctico, durante la vida de operación del equipo. Una cámara sensitiva es suficiente para suministrar señales como niveles de gris sean necesarios.

2.3.4 DOS Y TRES DIMENSIONES

La capacidad del sistema de visión humano de operar en tres dimensiones proveen la facilidad de generar información a distancia. esto da por seguro la realización de operaciones cotidianas, como guiar un vehículo o determinar cual de dos objetos es más alto en una operación de selección. Los sistemas de aparato de visión pueden tener la capacidad de operar en tres dimensiones (3-D), usando dos o más cámaras y complejas técnicas de procesamiento.

Mientras que la capacidad de dos dimensiones es disponible en ambos sistemas de visión, el humano y el aparato, la diferencia es que los datos en el sistema del aparato de visión están cuantificados y pueden ser utilizados para hacer medible el grado de precisión, limitada por la resolución del arreglo matricial. La visión humana puede hacer una rápida estimación de la distancia entre dos puntos

de una superficie, sin embargo una referencia mecánica, como regla, se requiere para proporcionar la precisión.

2.3.5 CONSIDERACIONES ECONOMICAS

La propuesta de implantación de un sistema de visión debe incluir un análisis del impacto económico de la aplicación. La economía de utilizar un aparato de visión para una aplicación industrial incluye a la vez una mejora del factor capacidad y una directa productividad de la manufactura en el factor costo. Al mismo tiempo se puede obtener un incremento de la productividad reduciendo el costo unitario. El análisis del factor costo relativo a reemplazar una inspección humana por una inspección automática de visión en la estación, podrá proveer una visión dentro de la complejidad del problema donde no siempre es posible cuantificar todos los factores. La determinación del impacto del costo en los obreros, por reemplazar el operador humano con una estación automática, puede no ser fácilmente determinado puesto que se puede involucrar uno o más valores de juicio. El análisis podrá incluir consideraciones de esta controversia.

2.4 VELOCIDAD DE RESPUESTA

La adquisición de una imagen por medio de un sistema de aparato de visión depende del tamaño de la matriz de la imagen, el tiempo de procesamiento del arranque electrónico y del tipo de cámara. La operación de las cámaras tipo tubo, en un modo convencional RS-170, podrá producir 30 imágenes por segundo para monitores estándar comercialmente disponibles; el número de imágenes por segundo puede incrementarse por un factor estimado de cinco o diez usando un modo distinto al RS-170. Las cámaras de estado sólido pueden adquirir la imagen en apenas 10 microsegundos; el tiempo requerido para leer la señal con el sensor dependerá del tamaño de la matriz, la velocidad de procesamiento, y el ancho de banda del sistema. El uso de técnicas de procesamiento en paralelo puede reducir el tiempo, por un factor igual al número de procesos paralelos utilizados.

Los sistemas de aparatos de visión son usados en la industria para inspeccionar etiquetas en botellas a una velocidad de 900 botellas por minuto o alrededor de una botella por 0.07 segundos. Las altas velocidades son posibles incrementando el número de imágenes de botellas en cada trama. El sistema humano opera con un tiempo de respuesta en el orden de un segundo como máximo. La velocidad de respuesta del sistema humano podrá decrementarse por la fatiga y por las condiciones ambientales.

Resumiendo, la adquisición repetitiva de imágenes usando sistemas de aparatos

de visión es alrededor de diez veces mayor que el sistema de visión humano. Esta proporción se incrementa con el tiempo conforme la electrónica mejora, mientras que el sistema humano no cambia. La capacidad de respuesta del sistema de aparato de visión es del orden de quince veces más grande que el sistema humano, en operaciones repetitivas.

2.5 RESPUESTA ESPECTRAL

La respuesta del ojo humano está limitada ya que únicamente puede utilizar luz visible (alrededor de 400 milimicras de espacio en el espectro electromagnético). El rango va del violeta que es como de 380 milimicras, al rojo de 760 milimicras de longitud de onda, como se muestra en la Figura 2.2. La respuesta del sistema de aparato de visión es relativamente amplio en comparación (alrededor de 100,000 milimicras); el rango va de los rayos gamma y X, en la región de microondas cortas, al infrarrojo, en la región de microondas largas del espectro.

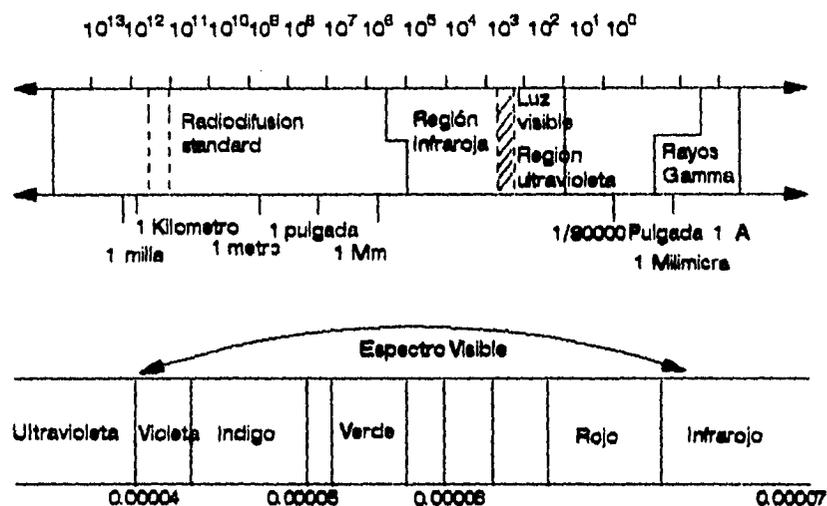


Figura 2.2 Respuesta Espectral

Los sistemas de aparatos de visión tienen la capacidad de crear una nueva imagen basados en la combinación de información para diferentes porciones del espectro. Por ejemplo, la información de una imagen de longitud de onda infrarroja, puede usarse en conjunto con la información de una imagen de longitud de onda visible, para identificar el objeto físico simulado como una fuente de energía calorífica en los procesos de manufactura industrial.

La capacidad de reconocimiento del color del ojo humano es compleja, debido a que no separa la radiación del color observado en sus componentes básicos. En cambio, obtiene el promedio de energía para diferentes longitudes de onda e identifica el color como una longitud de onda intermedia. La capacidad de reconocimiento de color para el aparato de visión requiere de elementos de datos, uno para cada longitud de onda: roja, verde, y azul. La producción de color en un monitor se logra energizando los tres componentes de manera que se produzca el color resultante. La capacidad de color de una imagen, y los componentes R, G, y B del color, aumentan la cantidad de datos para ser procesados con respecto a una imagen monocromática.

Resumiendo, el sistema de aparato de visión tiene una respuesta espectral mucho más amplia que el sistema de visión humano. Este es capaz de combinar y usar información de la imagen para diferentes partes del espectro. La capacidad de reconocimiento de color del sistema de aparato de visión es más consistente y preciso que la del ojo humano.

2.6 COMPONENTES DE UN SISTEMA DE VISION

Un aparato de visión comprende todos los elementos necesarios para obtener la representación digital de una imagen visual, para modificar los datos y presentar una imagen digital de datos al mundo externo. El sistema puede parecer complejo en un ambiente industrial debido a que todos los procesos de manufactura asociados, utilizan equipo en la aplicación. La complejidad se reduce cuando el sistema es visto en términos de tres principales componentes funcionales

- Adquisición de Imágenes
- Procesamiento
- Salida o despliegue

La mayoría de las aplicaciones de los sistemas de visión, comúnmente usados en la industria, son en las áreas de código de barras, publicaciones de sobremesa, preparación de impresiones y automatización de fábricas. El rango de precios para un sistema, varía en alrededor de diez mil dólares para una versión de PC o en varios millones de dólares para un complejo sistema utilizado en la tipografía y en las autoindustrias.

El rango de aplicaciones del código de barras en el mundo comercial, para la colección de datos automáticos de inventarios y control de la producción

incrementan su productividad automatizando los registros exploradores (escaners) en el punto de venta. El segmento del código de barra en la industria esta altamente estandarizado, debido a que el funcionamiento del sistema requerido para propósitos comerciales, es muy similar para almacenes de venta de varios tipos de mercancía.

Las aplicaciones en publicaciones de sobremesa incluyen la capacidad de explorar la tecnología de los aparatos de visión como un mecanismo de entrada de datos digitalizados, impresión de texto y fotografías, la generación de efectos especiales y la preparación de material en formato digital para uso en impresoras laser o dispositivos visuales.

Un simple sistema de visión industrial para automatizar una fábrica puede caracterizarse por una sencilla cámara de monitoreo en una línea de ensamble como se muestra en la figura 2.3. El sistema de visión observa el objeto, determina si esta dentro de las especificaciones y genera una señal de comando de acuerdo a los resultados establecidos. El equipo para la adquisición de datos de imágenes incluye la iluminación, la cámara y posiblemente el marco de recolección. El equipo de procesamiento incluye "hardware" y "software" en la unidad de visión, y el equipo de salida es un sistema de interface electrónica de varias partes del mundo de la manufactura.

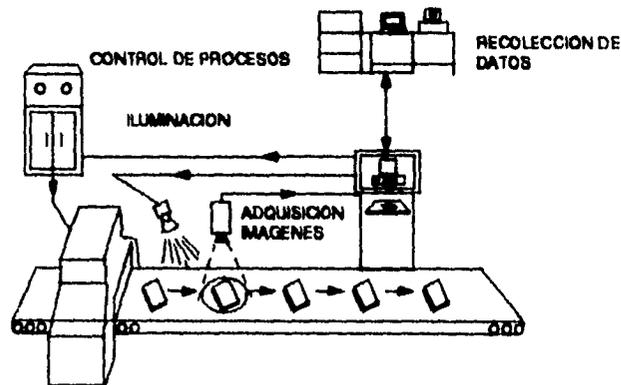


Figura 2.3 Equipo de Adquisición de imágenes

Las señales de instrucción eléctrica controlan la unidad, toman objetos de la línea de ensamble y aceptan o rechazan cajas, de acuerdo a la calidad.

Los aspectos de un sistema de visión de manufactura en un sistema de

ensamble se indican en la figura 2.4. Todas las funciones o partes están altamente interrelacionadas, y el entendimiento de cada una de sus partes es importante para conseguir el objetivo de la aplicación o tarea.

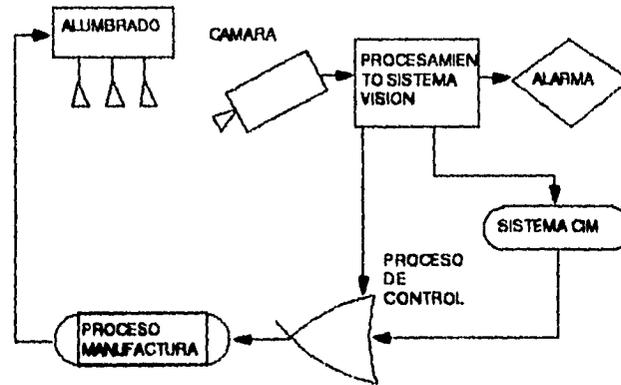


Figura 2.4 Sistema de Visión de la Manufactura

2.7 COMPARACION VISION HUMANA CONTRA APARATO DE VISION

En general todas las capacidades de la visión humana, pueden ser desempeñadas por los sistemas de aparatos de visión, y el nivel de desempeño es superior al de la visión humana, desde un punto de vista de confianza y precisión en operaciones repetitivas. En la tabla 2.1 se observa la comparación entre la visión humana y el aparato de visión [RB1].

TABLA 2.1 COMPARACION VISION HUMANA VS. APARATO DE VISION		
	VISION HUMANA	APARATO DE VISION
Flexibilidad	Muy adaptable y flexible como la tarea y el tipo de entrada.	Rígido como la tarea, requiere de datos cuantizados.
Habilidad	Puede hacer estimaciones correctas en materias subjetivas. Ejemplo: Detectar frutas en mal estado basados en el color, la textura, el contorno.	Puede hacer medidas dimensionales basados en datos de entrada predeterminados. Longitud de partes o dimensión de una cavidad en el cálculo de piezas.
Color	Subjetivo en color.	Medidas de magnitud de parámetros monocromáticos.
Sensibilidad	Adaptivo a las condiciones de luminosidad, características físicas, así como superficies y distancia de los objetos. Limitado en la habilidad de distinguir entre sombras de gris. Varían sus funciones individuales día con día. Puede identificar de 10 a 15 niveles de gris.	Sensitivo a nivel y frecuencia de iluminación, así como de la naturaleza física de la superficie y distancia del objeto. La habilidad de cuantizar es relativamente alta y fija por sensor, medio ambiente y sistema característicos. Puede identificar fácilmente 256 niveles de gris.
Respuesta	La velocidad de respuesta es lenta y del orden de 1/10 de segundo como máximo.	La velocidad de respuesta es muy alta y dependiente de un sistema de cómputo y del ancho de banda. Velocidad de respuesta del orden de 1/100 de segundo.
2-D y 3-D	Puede manipular tareas en 3-D y mult-longitudes de onda (colores de longitudes de onda) fácilmente.	Puede manipular tareas en 2-D fácilmente, sin embargo son limitados en tareas de 3-D. Requieren de 2 cámaras y es muy lento en 3-D. Puede desarrollar operaciones matemáticas de datos, mejorar datos y usar placas para calcular áreas, controles y lo que se desee.
Entrada de datos	Puede suministrar manualmente una entrada de datos discretos secundarios. El costo es alto y tiende a una tasa alta de error en los datos.	Puede suministrar automáticamente datos de entrada discretos precisos a la base de datos de la manufactura continua por un costo relativamente bajo incrementando el costo base.
Percepción	Percebe brillo en una escala logarítmica. Es afectado por el área circundante.	Puede percibir brillo en la escala lineal o logarítmica.
Espectro	Limitado por un espectro visual de 300 a 600 milímetros de longitud de onda.	Puede analizar mediciones en espectro completo desde rayos X a IR. Pero con sensores diferentes.

3 ADQUISICION Y REPRESENTACION DE UNA IMAGEN

3.1 ADQUISICION DE IMAGENES

La adquisición de imágenes transforma la imagen visual de un objeto físico y de sus características intrínsecas propias, las cuales establecen datos digitalizados que pueden ser usados dentro de la unidad de proceso del sistema. La función de adquisición puede ser dividida en cuatro fases:

- Iluminación
- Formación de imágenes o enfocamiento
- Detección de imágenes o sensibilidad
- Formateo en la cámara de la señal de salida

3.1.1 ILUMINACION

La iluminación es un parámetro clave que afecta al sistema de aparato de visión, debido a que de él depende directamente la calidad de entrada de datos. Es necesario acostumbrarse al diseño de iluminación para cada aplicación, debido a que es esencial no estandarizar propósitos generales en los equipos de iluminación de los aparatos de visión. El método y la fuente específica de energía luminosa afecta la cantidad de procesamiento y logro de resultados.

Gran número de sistemas industriales de aparatos de visión en el pasado utilizaron luz visible, debido a que las fuentes eran fácilmente disponibles y la aplicación frecuentemente fué la automatización de tareas de inspección manual. Las tareas de inspección se basaron en las capacidades del operador humano quienes trabajan dentro de la porción visible del espectro de luz. Los cuatro tipos de lámparas visibles que muy frecuentemente se utilizaron en el ambiente industrial son: la incandescente, fluorescente, de vapor de mercurio y de vapor de sodio. Sin embargo, el uso de iluminación fuera del espectro visible, como son los rayos X, ultravioleta, e infrarrojos se incremento debido a la necesidad de conseguir inspecciones especiales no posibles con la luz visible.

Los métodos para aplicaciones industriales pueden ser subdivididos dentro de cuatro categorías:

- Iluminación posterior
- Iluminación frontal

- Iluminación estructurada
- Iluminación estroboscópica

La iluminación ambiental afecta todos los métodos de iluminación porque altera el nivel total de iluminación en el objeto, que se manifiesta como ruido en los datos. El efecto de la iluminación ambiental puede minimizarse por el uso de escudos de luz o barreras, que prevengan o reduzcan el número de radiación desviada entrante al lente.

Las fuentes de luz convencional no siempre son suficientemente estables para asegurar una calidad de imagen específica.

Iluminación Posterior. La iluminación posterior es cuando un objeto se localiza entre la fuente de luz y la cámara como se ilustra en la Figura 3.1. Esto proporciona como resultado la creación de una silueta del objeto por la luz no interceptada por el objeto opaco.

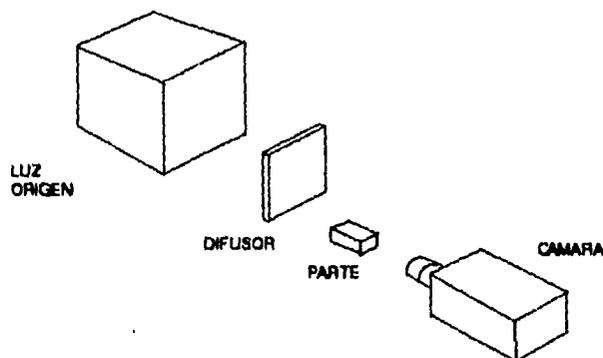


Figura 3.1 Iluminación Posterior

La iluminación posterior tiene la ventaja de producir imágenes de alto contraste en el perímetro de objetos opacos. El alto contraste minimiza la tarea de procesar la imagen y reduce la sensibilidad del sistema a las variaciones de la fuente de iluminación.

La iluminación posterior puede ser combinada con iluminación estructurada proyectando un patrón en la superficie de fondo y midiendo la diferencia entre el objeto y el patrón. La información de las características de la superficie, en superficies no visibles en la silueta, asemeja la presencia repentina de huecos ocultos y los objetos localizados en el límite de cada uno no pueden obtenerse mediante este método de iluminación.

La imagen, en el método de iluminación posterior es básicamente monocromática, porque se obtiene la silueta del objeto. Sin embargo los bordes del objeto pueden no corresponder a los píxeles límites en el sensor, los píxeles del borde pueden tomar valores entre un mínimo y un máximo nivel de gris del sistema de visión. Por ejemplo, el objeto de la Figura 3.2, parcialmente cubre el área de algunos píxeles y los valores de niveles de gris obtenidos con un sistema de 16 niveles se indican en el arreglo de la imagen.

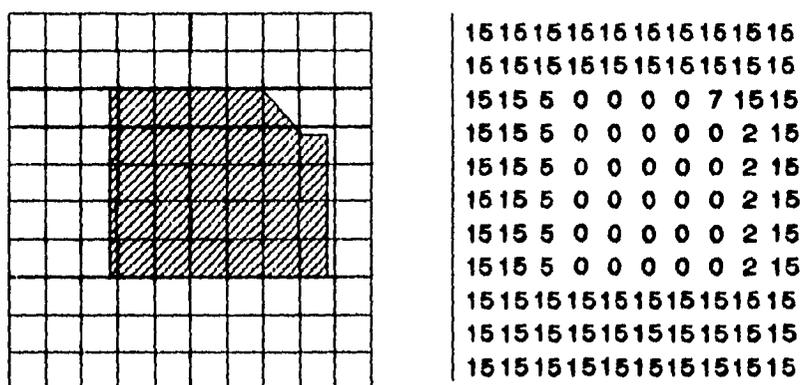


Figura 3.2 Iluminación

Iluminación Frontal: La iluminación frontal emplea luz reflejada por el objeto. La fuente de iluminación y la cámara, están del mismo lado del objeto, como se ilustra en la Figura 3.3. Este método de iluminación se utiliza para obtener información de la textura de la superficie o características, así como sus dimensiones.

Iluminación estroboscópica: La iluminación estroboscópica es la iluminación del objeto por un pulso corto de luz de alta intensidad (5 a 500 microsegundos de duración). El pulso corto de luz puede utilizarse para congelar el movimiento de un objeto mientras que la imagen es adquirida, o para reducir el efecto adverso de la iluminación ambiental.

La fuente de luz y la cámara tienen que estar sincronizados, debido a que el pulso es de corta duración. La duración del pulso es de importancia para permitir que el objeto aparezca esencialmente estacionario durante el pulso.

Iluminación estructurada: La iluminación estructurada es el uso de iluminación con un patrón especial o rejilla. La intersección del objeto y la

Iluminación proyectada resulta en un patrón único dependiendo del contorno y las dimensiones del objeto.

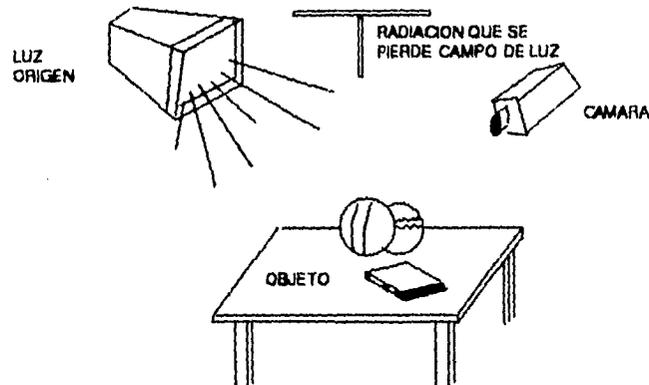


Figura 3.3 Iluminacion Frontal

3.1.2 FORMACION DE LA IMAGEN Y ENFOQUE

La imagen del objeto es enfocada a través de un elemento sensitivo con una lente, de manera similar al utilizado por una cámara fotográfica. La diferencia entre la cámara fotográfica y el sistema de aparato de visión, es que la cámara utiliza una cinta, en contraste con el sensor utilizado por el sistema de aparato de visión. El sensor convierte la imagen visual en una señal eléctrica.

La cámara del aparato de visión es usualmente especificada por separado para sistemas de visión, según la capacidad y características, y dependiendo de las aplicaciones. Resumiendo, los parámetros del lente de la cámara pueden ser especificados como un elemento que se adapta a la cámara para una aplicación específica.

3.1.3 DETECCION DE LA IMAGEN

Los aparatos de visión tienen un dispositivo optoelectrónico que convierte la radiación electromagnética, para la imagen de un objeto físico, en una señal eléctrica utilizada en el procesamiento por la unidad de visión. La imagen es enfocada en el sensor de la cámara por un lente. El elemento sensor localizado en el plano sensor de la cámara produce una señal eléctrica, representando la imagen visible. Las cámaras pueden tener elementos sensores, ya sean tubos al vacío o en estado sólido.

Las cámaras del tipo "tubo al vacío" fueron originalmente desarrolladas para la televisión comercial alrededor de los años 30's, antes de que hubiera cualquier

material de estado sólido. Ellos utilizaron un tubo al vacío conteniendo un elemento sensible de luz para captar imágenes; la imagen internamente capturada por un elemento sensible de luz es explorada para producir una señal proporcional de voltaje analógico [RB1].

Las cámaras de estado sólido tienen divididos fotodiodos, asignados para cada área de pixel como detectores de la iluminación del objeto. Los elementos son usualmente ordenados en un arreglo lineal o en un arreglo rectangular. Los detectores de estado sólido fueron desarrollados a fines de los años 60's como consecuencia de los dispositivos de carga acoplada (CCD) inventados por los laboratorios Bell de Telefonía [RB1]. La señal de salida de la cámara se obtiene sensando cada diodo en orden de secuencia, obteniendo así una serie de pulsos de voltaje que representan el valor del pixel con respecto a su ubicación. Los pulsos de voltaje se deben convertir al formato RS-170 para observar la imagen en un monitor estándar de televisión, o digitalizada si se desean introducir los datos en la memoria de una computadora.

3.2 PROCESAMIENTO

Las limitaciones físicas del hardware utilizado para adquirir la imagen, introducen ruido aleatorio y sistemático dentro de los datos de la imagen. Entonces, la imagen puede contener datos o características que no son de interés. La función primaria del procesamiento de imágenes, es crear una nueva imagen modificando los datos de tal modo que las características de interés sean realzadas y los efectos del ruido sean reducidos o eliminados.

El procesamiento contiene elementos electrónicos y programas (software) para llevar a cabo:

- Captura de la imagen
- Realce de la imagen
- Extracción de las características
- Formato de salida

El procesamiento de datos en un sistema de aparato de visión específico puede llevarse a cabo por hardware o software. El hardware incrementa el costo inicial, reduce la habilidad de hacer cambios en el procesamiento, e incrementa la probabilidad de que todo sistema similar, una planta de manufactura sean ejecutadas por el mismo procesamiento.

Los algoritmos básicos de procesamiento, para técnicas comunes, tales

como eliminación de ruido, realce de bordes, filtrado y modificación de escala de gris son incluidos en la mayor parte de los aparatos de visión comercialmente disponibles [RB1].

3.3 CONCEPTOS FUNDAMENTALES DE PROCESAMIENTO DE IMAGENES

3.3.1 PIXEL:

El área más pequeña que es posible presentar en cualquier proceso de reproducción se le conoce como elemento de imagen, para las imágenes digitales se le conoce como pixel. Una imagen esta compuesta de un conjunto de pxeles.

La imagen será descrita por una matriz $N \times M$ de valores de pixel (los elementos $P(i,j)$ son escalares no negativos) que indican el flujo de la intensidad de luz en el elemento de cuadro en (x,y) representado por el pixel.

Esto es ilustrado en la figura 3.4, la cual proporciona la relación entre el elemento de cuadro y la matriz de pxeles.

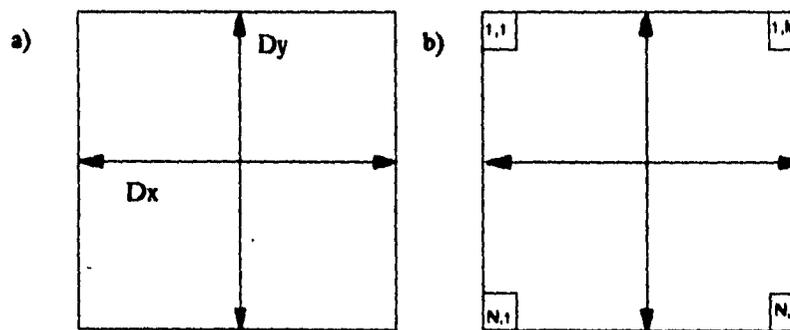


Figura 3.4 a) Elemento de cuadro b) Matriz pixel

Donde N = número máximo de pxeles en una columna
 M = número máximo de pxeles en un renglon

3.3.2 LOCALIZACION DE PIXELES:

Un pixel en su forma más elemental es identificado por sus coordenadas en un arreglo $N \times M$. El pixel en la localidad (n,m) tiene un valor numérico, el cual representa el valor común de iluminación en el área de la figura representada por el pixel.

En imágenes binarias un sistema binario será usado para representar la intensidad de iluminación; una región sin luz será indicada por un cero y la más alta iluminación con un uno.

La figura 3.5 contendrá una imagen que será indicada por una matriz 4x5 (4 renglones y 5 columnas). Cada elemento de cuadro se le asignará un número, dependiendo de la cantidad de luz en el área. Por ejemplo, en la esquina superior izquierda de la figura 3.5, representada por la localidad (1,1) en la matriz 4x5 contendrá un 0, indicando la carencia de luz y en la esquina inferior derecha de la figura que es la quinta columna y el cuarto renglon, representada por la localidad (4,5), contendrá un 1 indicando la máxima iluminación. Note que el valor del pixel (4,5) sería 15, si se usa un sistema de 16 niveles de grises de un sistema binario.

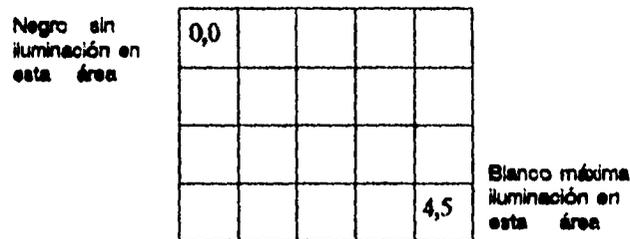


Figura 3.5 Localización de píxeles

Para este ejemplo, no hay información en el valor del píxel para valores intermedios y el sistema debe indicar un umbral de iluminación donde la representación varía de 0 a 1.

En el ejemplo los elementos de cuadro son rectangulares. El área en la figura puede ser rectangular o circular, dependiendo del sensor. En el caso de un tubo de cámara con sensor de área circular, puede ser traslapada como muestra la figura 3.6.

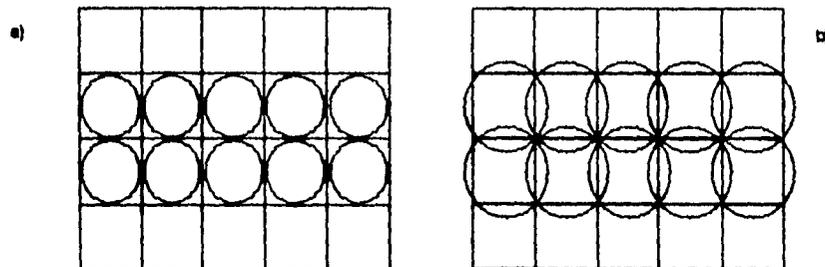


Figura 3.6 a) Área de píxel circular inscrita en la parrilla rectangular.
b) Área de píxel circular inscritos en los cuadros de la parrilla rectangular.

3.3.3 ESCALA O NIVELES DE GRIS:

Para proveer información en valores intermedios de iluminación, es necesario conocer el número de bits que representan el valor de un píxel. Por ejemplo, si se necesitan cuatro niveles de iluminación, se requieren dos bits. Cuatro bits son necesarios para 16 niveles y 8 bits para 256 niveles; El total de niveles de grises es usualmente una potencia de 2.

El valor más bajo 0 es asignado para el negro y el valor 1 para el blanco. Por ejemplo, el 15 para el sistema de 16 niveles es el valor para el blanco. El valor asignado a un píxel es siempre entero. [RB1]

<u>Escala de gris</u>	<u>Niveles de gris</u>
2 valores	0,1
8 valores	0 a 7
16 valores	0 a 15
256 valores	0 a 256

La representación binaria (2 valores o niveles) fué usada al principio de los sistemas de visión. Hoy en día se utilizan sistemas de 8, 16, 64 y 256 niveles. Los niveles de 64 y 256 proveen más distinciones que el ojo humano puede ver.

Un individuo puede diferenciar entre intensidades del orden de 40 diferentes niveles pero puede identificar solamente 10 a 15 niveles. La escala de 16 niveles es inferior al ojo humano, mientras que 64 y 256 niveles de gris provee una discriminación más alta que la de ojo humano.

4 SEGMENTACION DETERMINISTICA DE IMAGENES

4.1 INTRODUCCION

El análisis de imágenes básicamente involucra el estudio de extracción de características, segmentación y técnicas de clasificación

En los sistemas de visión como el mostrado en la figura 4.1. La imagen de entrada es primero preprocesada, la cual puede comprender: restauración, realce o la conveniente representación de los datos, después ciertas características son extraídas por segmentación de la imagen. La segmentación subdivide una imagen en sus partes o componentes, la segmentación debe detenerse cuando los objetos de interés han sido aislados.

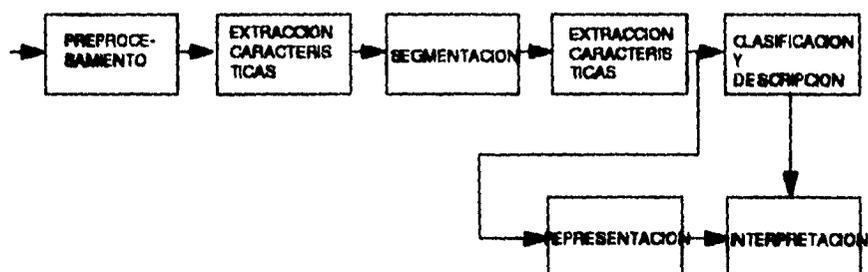


Figura 4.1 Sistema de Visión

Los algoritmos de segmentación para imágenes generalmente están basados en una de dos propiedades básicas de valores de niveles de gris: similitud y discontinuidad. En la primera propiedad, la proposición es partir de una imagen basada en cambios abruptos de niveles de gris. Las principales áreas de interés dentro de esta propiedad es la detección de puntos aislados y detección de líneas y bordes en una imagen. Las principales proposiciones dentro de la segunda propiedad son: umbral, región de crecimiento y región de partición y fusiónamiento.

Los métodos de segmentación de imágenes asumen que los objetos tienen una superficie de suavidad homogénea que responde a regiones de intensidad constante. En una imagen la intensidad cambia abruptamente en los límites, si bien esta suposición se usa muy frecuentemente, no siempre es válida, dado que la textura de la superficie y el ruido causan problemas adicionales.

En el método basado en discontinuidad y bordes, las discontinuidades locales son detectadas y conectadas a lo largo de la forma, encontrando así el límite del objeto. En el método de regiones, las áreas de la imagen con propiedades homogéneas son aisladas y así se identifican los diferentes objetos [R85]

4.2 DETECCION DE DISCONTINUIDADES

Los tres tipos básicos de discontinuidades en una imagen digital son: puntos, líneas y bordes. En la práctica, una forma para buscar las discontinuidades es barrer una máscara a través de toda la imagen. Para la máscara 3x3 mostrada en la figura 4.2, un procedimiento consiste en la suma de los coeficientes con los niveles de gris contenidos en la región abarcados por la máscara. Esto es, la respuesta de la máscara en cualquier punto de la imagen es.

$$R = W_1 Z_1 + W_2 Z_2 + \dots + W_9 Z_9 \quad (4.1)$$

donde Z_i Es el nivel de gris del píxel asociado con el coeficiente W_i

W_1	W_2	W_3
W_4	W_5	W_6
W_7	W_8	W_9

Figura 4.2 Máscara general 3x3

La respuesta de la máscara es definida con respecto a la localidad de su centro. Cuando la máscara es centrada en un píxel límite, la respuesta es calculada usando la apropiada vecindad parcial. Usando la máscara mostrada en la figura 4.3, se dice que un punto ha sido detectado si:

$$|R| > T$$

donde T es un umbral positivo y R esta dado por la ecuación (4.1). Basicamente la fórmula mide la diferencia entre el punto centro y sus vecinos. La idea es que el nivel de gris de un punto aislado será totalmente diferente del nivel de gris de sus vecinos [R88].

4.3 DETECCION DE LINEAS

Para la detección de líneas considere la máscara mostrada en la figura 4.4.

Si la primera máscara fuera movida alrededor de una imagen, esta respondería

-1	-1	-1
-1	8	-1
-1	-1	-1

Figura 4.3 Máscara usada para detectar puntos aislados diferentes de una constante

mas fuertemente a líneas orientadas horizontalmente, la máxima respuesta resultaría cuando la línea pasara por la mitad de la máscara, la segunda máscara respondería mejor a líneas orientadas a 45° , la tercera máscara respondería a líneas verticales y la cuarta máscara a líneas en la dirección de -45° .

	-1	-1	-1		-1	-1	2
HORIZONTAL	2	2	2	45 GRADOS	-1	2	-1
	-1	-1	-1		-1	2	-1
	-1	2	-1		2	-1	-1
VERTICAL	-1	2	-1	-45 GRADOS	-1	2	-1
	-1	2	-1		-1	-1	2

Figura 4.4 Máscaras para detectar líneas

Si R_1, R_2, R_3 Y R_4 son las respuestas de las máscaras de la figura 4.4 de izquierda a derecha, donde las R's estan dadas por la ecuación (4.1). Si se supone que todas las máscaras pasan a través de una imagen, si en un cierto punto de la imagen $R_i > R_j$, para toda j diferente de i, ese punto estará asociado en la dirección de la máscara i. Por ejemplo, si para un punto en la imagen, $R_1 > R_j$ para $j=2,3,4$, ese punto particular será asociado con una línea horizontal [RBB].

4.4 DETECCION DE BORDES

Un problema de importancia fundamental en el análisis de imágenes es la detección de bordes. Los bordes caracterizan los límites de los objetos y son útiles

para la segmentación e identificación de objetos en la escena.

La suposición para la detección de bordes, es que la intensidad es más o menos uniforme a lo largo del borde.

4.4.1 METODO DE UMBRAL PARA LA DETECCION DE BORDES

Uno de los métodos más útiles para extraer una figura o una característica de interés particular de una imagen es el método de umbral. El operador de umbral produce una imagen negra y blanca en la cual el objeto de interés es negro y el fondo blanco. Esto proporciona una imagen binaria donde hay solamente 2 valores 1 (blanco) y 0 (negro).

Considere la imagen f dada por la matriz de la figura 4.5. En la figura, el valor de cuantización de niveles de grises va de 0 a 8, donde 8 representa al blanco y el 0 representa al negro.

```
* 0 0 0 0 * 0 0 0 0 1 2 3 * 3 3
0 0 0 0 0 1 1 8 0 0 2 3 3 3 3 4
0 0 0 2 1 1 0 0 * 4 4 6 5 5 5 4
0 0 2 7 7 7 5 0 2 5 8 8 8 8 5 5
0 0 2 6 7 7 2 0 4 6 8 8 8 8 5 3
0 0 2 6 7 8 2 0 3 5 8 8 8 8 7 6 5
0 1 2 7 8 7 7 7 8 8 8 8 8 6 6 4
0 0 4 7 7 7 7 7 7 8 8 8 8 7 5 6
0 2 2 7 7 7 3 4 4 7 8 8 8 6 4 3
1 1 4 7 8 8 2 0 1 * 8 8 8 5 6 7
1 0 2 7 7 8 2 2 1 5 8 8 8 6 6 5
0 1 3 3 4 2 2 0 1 4 6 8 7 5 4 3
0 0 2 1 1 2 0 0 0 3 4 4 6 6 5 3
* 0 0 0 0 0 0 0 2 2 3 2 3 5 3 3
```

Figura 4.5 Imagen f

Para el umbral de la imagen f , se considera 1 a los valores iguales o superiores a 7 y 0 a los valores menores a 7. Lo anterior se representa en la figura 4.6, la cual tiene solamente dos niveles de grises 0 y 1. La letra H es distinguible en la imagen de umbral, es importante notar que, mientras los valores pequeños son parte de la imagen y denotan el nivel de gris 0, el asterisco (*) no está en el dominio de la imagen y los pxeles asteriscos no están envueltos en la operación

umbral, ellos son invariantes.

La operación fundamental "umbral" es una operación binaria. Las entradas son una imagen f y un número real t . La salida es una imagen blanca y negra pura, es decir sin niveles de gris intermedios. La imagen umbral es denotada por $\text{THRESH}(f;t)$. Esto es una imagen binaria y esta definida por:

$$[\text{THRESH}(f;t)](i,j) \begin{cases} 1 & \text{si } f(i,j) \geq t \\ 0 & \text{si } f(i,j) < t \\ * & \text{si } f(i,j) = * \end{cases}$$

```
* 0000 * 00000000 * 00
000000000000000000
0001110000111000
0000110000111000
0000110000111100
0001111111111000
0001111111111100
0001110001111000
0001110001111000
000111000*111001
0001110000111000
0000000000011000
0000000000000000
0000000000000000
```

Figura 4.6 THRESH (t;7)

El diagrama de bloques de THRESH se ilustra a continuación, en la figura 4.7:

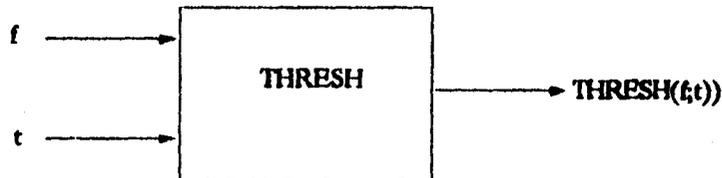


Figura 4.7 Diagrama de Bloques de THRESH.

Note que el asterisco ocurre en el píxel donde el sensor no tiene información. Los píxeles de asteriscos no están en el dominio de la imagen y esos píxeles están fuera de la representación de la matriz de la imagen.

Análogamente a la pérdida de dos píxeles en la letra H a la izquierda (figura 4.6), existen siete píxeles negros extras en el lado derecho. Mientras que la pérdida

de píxeles a la izquierda podría ser recuperada con un umbral $t=8$, todos podrían ser removidos excepto uno con un umbral $t=8$. Sin embargo la entrada $t=7$ parece ser la mejor. Las figuras 4.8 y 4.9 muestran el resultado de aplicar los umbrales $t=6$ y $t=8$ respectivamente. Una figura crea exceso de distorsión por la inclusión de muchos píxeles a la derecha, mientras que la otra figura resulta con pocos píxeles a la izquierda.

```
*0000*0000000*00
0000000100000000
00000000*0010000
0001110000111100
0001110001111110
0001110000111110
0001111111111110
0001111111111101
0001110001111100
000111000*111011
0001110000111110
0000000000111000
0000000000011000
0000000000000000
```

Figura 4.8 THRESH (t:6)

```
*0000*0000000*00
0000000100000000
00000000*0010000
0001110000111100
0001110001111100
0001110000111110
0001111111111110
0001111111111101
0001110001111100
000111000*111011
0001110000111110
0000000000111000
0000000000011000
0000000000000000
```

Figura 4.9 THRESH (t:8)

4.4.2 GRADIENTE Y NORMAL

Para una imagen continua $f(x,y)$ su derivada asume un máximo local en la dirección del borde. Por lo tanto la técnica para detectar un borde es medir el gradiente $f_{[RS]}$ a lo largo de r en la dirección θ (Figura 4.10), esto es:

$$\frac{df}{dr} = \frac{df}{dx} \frac{dx}{dr} + \frac{df}{dy} \frac{dy}{dr} = f_{\cos\theta} + f_{\sin\theta} \quad (4.2)$$

El máximo valor de $f(x,y)$ se obtiene cuando df/dr es igual a cero, esto es:

$$-f_{\sin\theta} + f_{\cos\theta} = 0 \quad \text{y} \quad \theta_g = \tan^{-1}(f_y/f_x) \quad (4.3a)$$

$$Df = (df/dr)_{\max} = (G_x^2 + G_y^2)^{1/2} \quad (4.3b)$$

donde θ es la dirección del borde.

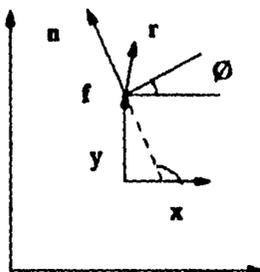


Figura 4.10 Gradiente de $f(x,y)$.

En la práctica es común aproximar el gradiente con el valor absoluto:

$$Df \approx |G_x| + |G_y| \quad (4.4)$$

La dirección del vector gradiente es una cantidad muy importante. $\theta(x,y)$ representa el ángulo del vector Df en (x,y) . Entonces del análisis vectorial

$$\theta(x,y) = \tan^{-1}(G_y/G_x) \quad (4.5)$$

donde el ángulo es medido con respecto al eje x . El gradiente de una imagen se basa en obtener la derivada parcial de f para cada $pxel$. La magnitud del vector Df es:

$$Df = \{ (a f/a x)^2 + (a f/a y)^2 \}^{1/2} \quad (4.6)$$

Considere la región de la imagen mostrada en la figura 4.11, donde las Z 's denotan el valor de los niveles de grises. La ecuación (4.6) puede ser expresada en el punto Z_5 en un número de formas distintas. El uso de la diferencia $(Z_5 - Z_6)$ en

la dirección x y (Z_5-Z_3) en la dirección y , que combinados:

$$Df = [(Z_5-Z_3)^2 + (Z_6-Z_4)^2]^{1/2} \quad (4.7a)$$

Se pueden obtener resultados similares usando el valor absoluto

$$Df = |Z_5-Z_3| + |Z_6-Z_4| \quad (4.7b)$$

Otra proposición para aproximar la ecuación (4.6) es:

$$Df = [(Z_5-Z_3)^2 + (Z_8-Z_7)^2]^{1/2} \quad (4.8a)$$

usando el valor absoluto:

$$Df = |Z_5-Z_3| + |Z_8-Z_7| \quad (4.8b)$$

Z ₁	Z ₂	Z ₃
Z ₄	Z ₅	Z ₆
Z ₇	Z ₈	Z ₉

Figura 4.11 Región 3x3 de una imagen.

Basado en estos conceptos, dos tipos de operadores para la detección de bordes han sido introducidos: el operador gradiente y el operador compass (generalización de ocho máscaras del operador gradiente). El operador gradiente está representado por un par de máscaras G_1 y G_2 , las cuales miden el gradiente de la imagen $u(m,n)$ en dos direcciones ortogonales. Definiendo el gradiente bidireccional.

$$g_1(m,n) = \langle U, G_1 \rangle_{m,n}$$

$$g_2(m,n) = \langle U, G_2 \rangle_{m,n}$$

La magnitud del vector gradiente y la dirección están dados por:

$$g(m,n) = (g_1^2(m,n) + g_2^2(m,n))^{1/2} \quad (4.9)$$

$$\theta g(m,n) = \tan^{-1}(g_2(m,n)/g_1(m,n)) \quad (4.10)$$

A menudo la magnitud del gradiente es calculada como:

$$g(m,n) = |g_1(m,n)| + |g_2(m,n)| \quad (4.11)$$

El operador compass mide el gradiente en un número seleccionado de direcciones (8 direcciones por lo que se necesitan 8 máscaras).

Dado el vector $V=(x,y)$ hay tres formas comunes de definir la normal o magnitud del vector $|V_{max}|$. La primera es llamada l_{max} y esta dada por:

$$|V_{max}| = | \max(x,y) | \quad (4.12)$$

La segunda es la normal h y esta dada por:

$$|V_1| = |X| + |Y| \quad (4.13)$$

Y finalmente la normal I_2 :

$$|V_2| = (X^2 + Y^2)^{1/2} \quad (4.14)$$

Se pueden observar las tres normales en los diagramas de bloques de la figura 4.12. En los puntos 4.4.3 y 4.4.4 se analizarán los diferentes métodos correspondientes al operador gradiente y al operador compass respectivamente (los cuales fueron implantados y usados durante esta tesis)

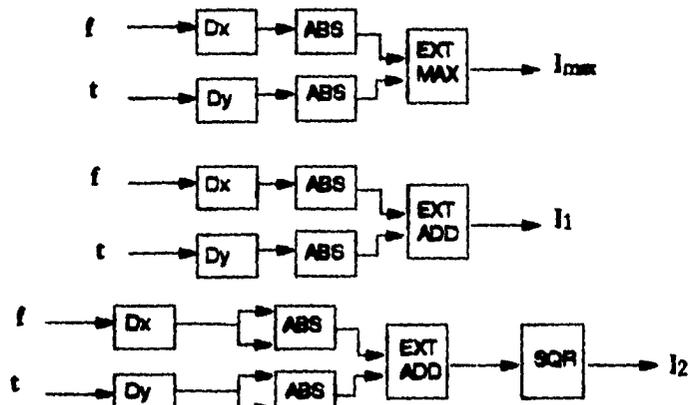


Figura 4.12 Normales para el operador Gradiente.

4.4.3 DIFERENTES TIPOS DE OPERADORES GRADIENTE

Para calcular los bordes de una imagen, el método más simple es calcular el gradiente de la imagen, un borde es dado cuando la magnitud del gradiente excede un cierto umbral. Muchos métodos han sido desarrollados utilizando el gradiente digital, los más comunes son: el gradiente Roberts, Laplaciano, diferencia de pixeles, Prewitt, Sobel, Frei-chen y Kirsh.

Si dos regiones adyacentes existen dentro de una imagen que difiere substancialmente con respecto a los niveles de gris en los bordes, entonces el cálculo de la velocidad de cambio con respecto a los niveles de gris puede ayudar a determinar el borde.

El operador gradiente que es requerido mide la velocidad de cambio de los niveles de gris. El valor de éste operador será cero en dominios constantes dentro de la imagen y variará según la velocidad de las fluctuaciones de niveles de gris en otra parte.

Es necesario que el operador mida la velocidad de cambio de niveles de

gris en la dirección vertical y horizontal, por lo que existirán dos máscaras G_1 en la dirección horizontal (D_x) y G_2 en la dirección vertical (D_y). La máscara G_1 medirá la velocidad de cambio de niveles de gris en la dirección X y G_2 medirá la velocidad de cambio de los niveles de gris en la dirección Y .

La ocurrencia de un borde no depende de la dirección de la intensidad del contraste. D_x y D_y deben ser usados junto con el operador valor absoluto ABS y con el operador umbral $THRESH$, que producirán la figura deseada. El borde es entonces representado en una imagen blanco y negro.

Las imágenes obtenidas $VERTEDGE$ y $HOREDGE$ tendrán 2 entradas, cada una respectivamente: la imagen f y el umbral t . El diagrama de bloques para la imagen $VERTEDGE$ se muestra en la figura 4.13

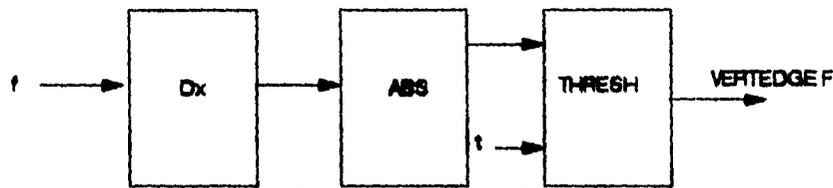


Figura 4.13 Diagrama de bloques $VERTEDGE$

D_y será denotada $HOREDGE$ y representada por un diagrama de bloques similar al de la figura 4.13.

4.4.3.1 OPERADOR GRADIENTE ROBERTS

Este operador [R82] tiene las siguientes máscaras:

$$G_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad G_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Que se obtiene para cada píxel:

$$\begin{aligned} [FILTER(f;G_1)](i,j) &= [f(i,j) - f(i+1,j+1)] \\ [FILTER(f;G_2)](i,j) &= [f(i,j+1) - f(i+1,j)] \end{aligned}$$

Donde el $_i$ indica el píxel actual donde se están efectuando los cálculos, es decir corresponde al píxel (i,j) .

Después de obtener los filtros se recomienda calcular la normal t como se muestra a continuación.

$$f_{\text{ABS}}(i,j) = |[\text{FILTER}(f;G1)](i,j)| + |[\text{FILTER}(f;G2)](i,j)|$$

Después se le aplica el operador umbral y se obtiene la imagen de bordes.

4.4.3.2 OPERADOR DIFERENCIA DE PÍXELES

El operador de diferencia de píxeles [RB3] tiene las siguientes máscaras:

$$G1 = \begin{vmatrix} 0 & 0 \\ 1 & -1 \end{vmatrix} \quad G2 = \begin{vmatrix} -1 & 0 \\ 1 & 0 \end{vmatrix}$$

Obteniendo para cada píxel la siguiente salida:

$$\begin{aligned} Dx &= [\text{FILTER}(f;G1)](i,j) = [f(i,j) - f(i,j+1)] \\ Dy &= [\text{FILTER}(f;G2)](i,j) = [f(i,j) - f(i-1,j)] \end{aligned}$$

Aplicando la normal:

$$f_{\text{ABS}} = |Dx| + |Dy|$$

Por último se aplica el operador umbral y se obtiene la imagen de bordes.

4.4.3.3 OPERADOR DE DIFERENCIA DE PÍXELES SEPARADA

El operador de diferencia de píxeles separada [RB3] tiene las siguientes máscaras:

$$G1 = \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{vmatrix} \quad G2 = \begin{vmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix}$$

Se obtiene para cada píxel la salida:

$$\begin{aligned} Dx &= [\text{FILTER}(f;G1)](i,j) = [f(i,j) - f(i,j+2)] \\ Dy &= [\text{FILTER}(f;G2)](i,j) = [f(i+1,j+1) - f(i+1,j-1)] \\ f_{\text{ABS}} &= |Dx| + |Dy| \end{aligned}$$

4.4.3.4 OPERADOR LAPLACIANO

El operador Laplaciano [RB4] tiene la siguiente máscara:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\text{FILTER}(i,j) = f(i,j-1) + f(i,j+1) + f(i-1,j) + f(i+1,j) - 4f(i,j)$$

Aplicando el operador valor absoluto y el operador umbral se obtiene la imagen de bordes.

4.4.3.5 OPERADOR PREWITT 2 MASCARAS ,3 NIVELES

El operador Prewitt [PB2] tiene las siguientes dos máscaras:

$$G1 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad G2 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

$$[\text{FILTER}(f;G1)](i,j) = 1/3\{f(i-1,j+1) - f(i-1,j-1) + f(i,j+1) - f(i,j-1) + f(i+1,j+1) - f(i+1,j-1)\}$$

$$[\text{FILTER}(f;G2)](i,j) = 1/3\{f(i-1,j-1) + f(i-1,j) + f(i-1,j+1) - f(i+1,j) - f(i+1,j+1)\}$$

$$f_{\text{ABS}} = |[\text{FILTER}(f;G1)](i,j)| + |[\text{FILTER}(f;G2)](i,j)|$$

4.4.3.6 OPERADOR DIFERENCIA DE PÍXELES 8 NIVELES

El operador de diferencia de píxeles de 8 niveles [PB4] tiene la siguiente máscara:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\text{FILTER}(i,j) = f(i-1,j-1) + f(i-1,j) + f(i-1,j+1) + f(i,j-1) + f(i,j+1) + f(i+1,j-1) + f(i+1,j) + f(i+1,j+1) - 8f(i,j)$$

Aplicando el operador valor absoluto y el operador umbral se obtiene la

imagen de bordes.

4.4.3.7 OPERADOR SOBEL

El operador Sobel [RB3] tiene las siguientes máscaras:

$$G1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\begin{aligned} [\text{FILTER}(f;G1)(i,j)] &= 1/4\{f(i-1,j+1) - f(i-1,j-1) + 2f(i,j+1) \\ &\quad - 2f(i,j-1) + f(i+1,j+1) - f(i+1,j-1)\} \\ [\text{FILER}(f;G2)(i,j)] &= 1/4\{f(i-1,j-1) + 2f(i-1,j) + f(i-1,j+1) \\ &\quad - f(i+1,j-1) - 2f(i+1,j) - f(i+1,j+1)\} \end{aligned}$$

Aplicando el operador valor absoluto y el operador umbral se obtiene la imagen de bordes.

4.4.3.8 OPERADOR FREI - CHEN

El operador Frei-chen [RB3] tiene las siguientes máscaras.

$$G1 = \begin{bmatrix} -1 & 0 & 1 \\ \sqrt{2} & 0 & \sqrt{2} \\ -1 & 0 & 1 \end{bmatrix} \quad G2 = \begin{bmatrix} 1 & \sqrt{2} & 1 \\ 0 & 0 & 0 \\ -1 & \sqrt{2} & -1 \end{bmatrix}$$

$$\begin{aligned} [\text{FILTER}(f;G1)(i,j)] &= 1/(2^{1/2} + 2^{1/2})\{f(i-1,j+1) - f(i-1,j-1) + \sqrt{2}f(i,j+1) \\ &\quad - \sqrt{2}f(i,j-1) + f(i+1,j+1) - f(i+1,j-1)\} \\ [\text{FILER}(f;G2)(i,j)] &= 1/(2^{1/2} + 2^{1/2})\{f(i-1,j-1) + \sqrt{2}f(i-1,j) + f(i-1,j+1) \\ &\quad - f(i+1,j-1) - \sqrt{2}f(i+1,j) - f(i+1,j+1)\} \end{aligned}$$

Aplicando el operador valor absoluto y el operador umbral se obtiene la imagen de bordes.

4.4.3.9 OPERADOR KIRSH

El operador Kirsh [RB4] tiene las siguientes máscaras:

$$G1 = \begin{vmatrix} -5 & 3 & 3 \\ -5 & 0 & 3 \\ -5 & 3 & 3 \end{vmatrix} \quad G2 = \begin{vmatrix} 3 & 3 & 3 \\ 3 & 0 & 3 \\ -5 & -5 & -5 \end{vmatrix}$$

$$\begin{aligned} [\text{FILTER}(f;G1)(i,j)] &= 1/15\{-f(i-1,j-1) + f(i-1,j) + f(i-1,j+1) \\ &\quad - f(i,j-1) + f(i,j+1) - f(i+1,j-1) + \\ &\quad f(i+1,j) + f(i+1,j+1)\} \\ [\text{FILTER}(f;G2)(i,j)] &= 1/15\{f(i-1,j-1) + f(i-1,j) + f(i-1,j+1) \\ &\quad f(i,j-1) + f(i,j+1) - f(i+1,j-1) \\ &\quad - f(i+1,j) - f(i+1,j+1)\} \end{aligned}$$

Aplicando el operador valor absoluto y el operador umbral se obtiene la imagen de bordes.

4.4.4 GENERALIZACION DE OCHO MASCARAS DEL OPERADOR GRADIENTE (OPERADOR COMPASS)

Cada uno de los gradientes detectores de bordes envuelven el uso de dos máscaras. Para los operadores o gradientes Sobel, Kirsch, Prewitt y Frei - chen, existe una máscara que responde al cambio en la dirección X y otra que responde al cambio en la dirección Y. Con el gradiente Roberts una máscara responde al cambio en la dirección de -45 grados y la otra responde al cambio en la dirección de 45 grados.

El operador compass surge de la necesidad de observar máscaras que respondan a cambios en todas las direcciones que son múltiplos enteros de 45 grados. En esencia, la metodología es aplicar filtros por medio de máscaras que respondan también al cambio en una dirección en particular. Hay ocho direcciones de 45 grados, cada una produce filtros envolventes para 8 máscaras diferentes. Cada máscara empieza en un ciclo de 45 grados de la máscara previa. Una normal máxima se aplica a las ocho salidas para obtener una imagen que indica la velocidad del cambio de nivel de gris en cada pixel. Después a ésta imagen se le aplica el operador umbral y produce una imagen de bordes. El siguiente diagrama de bloques ilustra el método a ser aplicado (figura 4.14) [RB2].

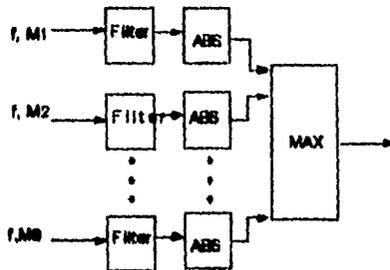


Figura 4.14 Diagrama de bloques del operador Compass.

4.4.4.1 OPERADOR PREWITT, 8 MASCARAS, 3 NIVELES

Las máscaras para este operador Prewitt [RB2] son las siguientes:

-1 0 1	0 1 1	1 1 1
-1 0 1	-1 0 1	0 0 0
-1 0 1	-1 -1 0	-1 -1 -1
1 1 0	1 0 -1	0 -1 -1
1 0 -1	1 0 -1	1 0 -1
0 -1 -1	1 0 -1	1 1 0
-1 -1 -1	-1 -1 0	
0 0 0	-1 0 1	
1 1 1	0 1 1	

4.4.4.2 OPERADOR PREWITT, 8 MASCARAS, 5 NIVELES

Las máscaras para este operador Prewitt [RB2] son las siguientes:

-1 1 1	1 1 1	1 1 1
-1 2 1	-1 2 1	1 2 1
-1 1 1	-1 -1 1	-1 -1 -1
1 1 1	1 1 -1	1 -1 -1
1 2 -1	1 2 -1	1 2 -1
1 -1 -1	1 1 -1	1 1 1
-1 -1 -1	-1 -1 1	
1 2 1	-1 2 1	
1 1 1	1 1 1	

4.4.4.3 OPERADOR SOBEL, 6 MASCARAS

Las máscaras para este operador Sobel [RB3] son las siguientes:

-1 0 1	0 1 2	1 2 1
-2 0 2	-1 0 1	0 0 0
-1 0 1	-2 -1 0	-1 -2 -1

2 1 0	1 0 -1	0 -1 -2
1 0 -1	2 0 -2	1 0 -1
0 -1 -2	1 0 -1	2 1 0

-1 -2 -1	-2 -1 0
0 0 0	-1 0 1
1 2 1	0 1 2

4.4.4.4 OPERADOR KIRSH, 6 MASCARAS

Las máscaras para este operador de kirsh [RB4] son las siguientes:

-5 3 3	3 3 3	3 3 3
-5 0 3	-5 0 3	3 0 3
-5 3 3	-5 -5 3	-5 -5 -5

3 3 3	3 3 -5	3 -5 -5
3 0 -5	3 0 -5	3 0 -5
3 -5 -5	3 3 -5	3 3 3

-5 -5 -5	-5 -5 3
3 0 3	-5 0 3
3 3 3	3 3 3

4.4.4.5 OPERADOR DE DIFERENCIA DE PIXEL SEPARADA, 6 MASCARAS

Las máscaras para este operador de diferencia de píxeles separada [RB3] son las siguientes:

0 0 0	0 0 1	0 1 0
-1 0 1	0 0 0	0 0 0
0 0 0	-1 0 0	0 -1 0

$$\begin{array}{ccc}
 1 & 0 & 0 \\
 0 & 0 & 0 \\
 0 & 0 & -1
 \end{array}
 \quad
 \begin{array}{ccc}
 0 & 0 & 0 \\
 1 & 0 & -1 \\
 0 & 0 & 0
 \end{array}
 \quad
 \begin{array}{ccc}
 0 & 0 & -1 \\
 0 & 0 & 0 \\
 1 & 0 & 0
 \end{array}$$

$$\begin{array}{ccc}
 0 & -1 & 0 \\
 0 & 0 & 0 \\
 0 & 1 & 0
 \end{array}
 \quad
 \begin{array}{ccc}
 -1 & 0 & 0 \\
 0 & 0 & 0 \\
 0 & 0 & 1
 \end{array}$$

4.5 TRANSFORMADA HOUGH

Los bordes son transformados a otro espacio llamado el espacio Hough. La ecuación general de la línea recta puede ser escrita como:

$$x \cos \varnothing + y \sin \varnothing = r \quad (4.15)$$

donde \varnothing es el ángulo de una normal de la recta al eje X y r es la distancia de esta normal (Figura 4.15a). Cualquier punto del borde (X_i, Y_i) en esta línea debe satisfacer:

$$X_i \cos \varnothing + Y_i \sin \varnothing = r \quad (4.16)$$

La ecuación (4.16) puede ser interpretada como una curva sinusoidal en el espacio (r, \varnothing) y vista como la transformada Hough [RHS] de el punto (X_i, Y_i) en el espacio (r, \varnothing) . Cualquier punto en la curva corresponde a una línea contenida en (X_i, Y_i) . Si un número de puntos son colineales en el espacio (x, y) , su curva de transformada Hough debe intersectar en el mismo punto en el espacio (r, \varnothing) . Para una implementación digital se debe cuantizar el espacio (r, \varnothing) . En la figura 4.15b se observa la transformada Hough de la recta de la figura 4.15a.

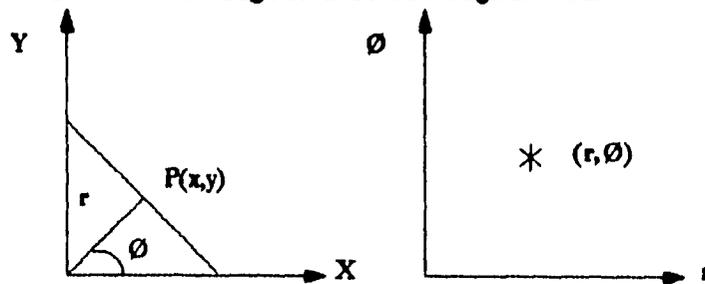


Figura 4.15 a) línea recta b) Transformada Hough

4.6 TÉCNICAS DE SEGMENTACION

El diagrama de la figura 4.16 lista varias técnicas de segmentación de imágenes [R55]:

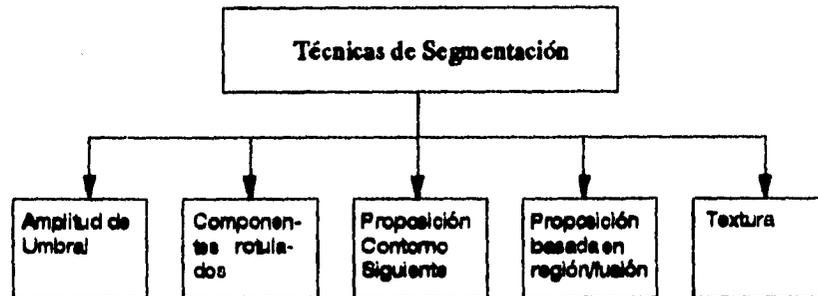


Figura 4.16 Técnicas de segmentación

4.6.1 AMPLITUD DE UMBRAL

Primero necesitan ser identificadas las regiones de una imagen que tienen el potencial para ser alguna parte del objeto, por simplicidad, aunque no necesariamente la mejor manera de identificar estas regiones de objetos es mejorar la etiqueta de umbral de operación en la cual cada píxel que tiene un valor bastante alto se le da el valor binario 1. El valor binario 1 es aquí asignado al píxel que tiene alguna posibilidad de ser parte de un objeto de interés. Cada píxel que no tiene un valor bastante alto se le asigna el valor binario 0. Esta designación significa que tiene poca posibilidad de ser parte de un objeto de interés. Esta técnica agrupará o segmentará la imagen y permite aislar los objetos de interés.

El valor del Umbral separa el valor oscuro del objeto de brillo, la separación puede ser idealmente realizada si la distribución de los píxeles oscuros y de brillo son conocidos, el valor umbral puede entonces ser determinado.

Es decir, uno de los métodos más útiles para extraer una figura o una característica de interés particular de una imagen es aplicar el método de umbral. El operador umbral produce una imagen negra y blanca, en el cual el objeto de interés es negro y el fondo blanco (o viceversa).

Otra manera de aplicar el valor de umbral para segmentar una imagen es, por ejemplo, si en una determinada imagen un objeto tiene un valor de nivel de gris T , puede ser segmentado o aislado de la imagen aplicando un umbral T a la imagen,

y entonces se tendrá segmentado el objeto.

4.6.2 COMPONENTES ROTULADOS O ETIQUETADOS

Un simple y efectivo método de segmentación de imágenes binarias es examinar la conectividad de píxeles con sus vecinos y entonces se etiqueta esta conexión. Dos algoritmos son pixel rotulado y análisis de conectividad.

PIXEL ROTULADO.- Una imagen binaria es explorada de izquierda a derecha y de arriba a abajo, el píxel corriente X (Figura 4.17) es etiquetado 1 para un objeto y 0 para un agujero, en seguida se examina la conectividad de los vecinos A;B;C y D. Por ejemplo, si X=1, entonces este es asignado como un objeto y entonces es conectado, si existen dos o más objetos calificados, entonces estos objetos son declarados equivalentes y entonces son fusionados. Un nuevo objeto etiquetado es asignado cuando ocurre una transición de 0 a 1. Una vez etiquetado el píxel, las características de ese objeto son almacenadas, al final de la exploración, características tales como centroide, área y perímetro son almacenadas para cada región de 1's conectados.

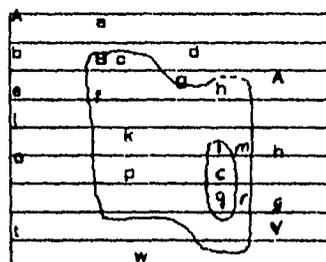
C	A	D
B	X	

Figura 4.17 Vecindad del píxel X en el algoritmo de píxel rotulado.

ANÁLISIS DE CONECTIVIDAD: Un método alternativo de segmentación de imágenes binarias es analizar la conectividad del muestreo de líneas sucesivas. Para ilustrar esta idea, considere la figura 4.18, donde el cambio de negro a blanco es denotada por a,b,c,...A, la tabla de segmentación es creada, donde el cambio a de la primera línea explorada es puesta en la primera columna. El objeto del primer cambio a es nombrado A. El primer cambio de la siguiente línea explorada, b, es del mismo color que a, el traslape b pertenece al objeto A y es colocado abajo de a en la primera columna, c es de diferente color y es colocado en una nueva columna, para un objeto etiquetado B, el cambio d es del mismo color que a, b y d ambos se traslapan, por lo tanto ocurre una divergencia y una nueva columna del objeto A es creada, donde d es colocada. Una bandera divergente ID1 es puesta en esta columna para indicar que el objeto B ha causado esta divergencia. También la bandera ID2 de B (Columna2) es puesta en A para indicar que B ha causado divergencia en A. Similarmente, la convergencia ocurre cuando dos o más cambios de 0's y 1's dan una línea de traslape, con un cambio del mismo

color en la línea previa. La convergencia ocurre en el cambio u, la bandera de convergencia IC1 corresponde a C en la columna 4 y IC2 a B en la columna 6. similarmente w pone la bandera de convergencia IC2 a A en la columna 2 y la columna 5 es etiquetada como perteneciente al objeto A. La tabla de segmentación da los datos relevantes de cada objeto

Columna	1	2	3	4	5	6
Nivel	1	2	1	2	1	3
Objeto		A	B	A	B	A
IC1				B	C	
ID1				B	C	
IC2		A				B
ID2		A				B



columna	1	2	3	4	5	6
		a				
		b	c	d		
		e	f	g	h	i
DATOS	j	k	m	n	l	
		o	p	r	s	q
		t	u		v	
						w

Figura 4.18 Algoritmo de conectividad

4.6.3 PROPOSICION BASADA EN CONTORNO SIGUIENTE

Contorno siguiente es una simple proposición, la cual sigue el límite o contorno de la figura. La base de la técnica es empezar en un punto que es un borde y se examinan los vecinos a este punto en la dirección de los bordes o del contorno. La base para seleccionar un candidato varia de tarea a tarea y de algoritmo a algoritmo pero normalmente es dependiente de la dirección del gradiente y la magnitud del gradiente. La proposición basada en contorno siguiente es a menudo basada en operadores gradientes detectores de bordes, estas proposiciones son normalmente más exitosas en imágenes con poco ruido.

Para empezar con el algoritmo no se puede asumir que el límite constituye una curva cerrada. Así el límite es seguido en dos direcciones. Primero en la

dirección hacia adelante y después en la dirección de reversa. la dirección hacia adelante es arbitraria pero comunmente asignada a la dirección normal, dirección $+90^\circ$ y la dirección de reversa corresponde a la dirección -90° . La proposición de este algoritmo es seguir una secuencia de pixeles límites y generar un modelo de los límites, y compararlo con una base de datos ya establecida.

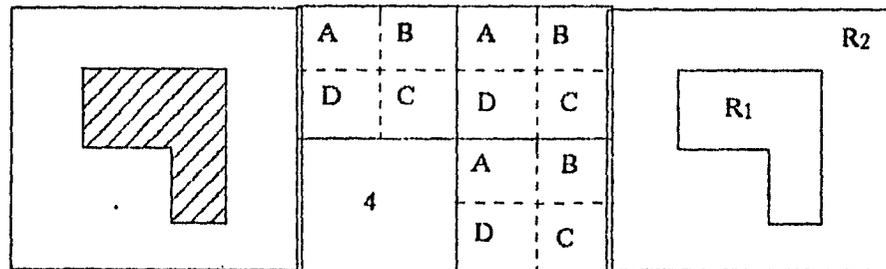
4.6.4 PROPOSICION BASADA EN REGION Y AGRUPAMIENTO

Las técnicas de segmentación basadas en región, es identificar regiones en una imagen que tienen características similares:

1.- Unir dos regiones R_i y R_j si $W/P_m > Q_1$ donde $P_m = \min(P_i, P_j)$, P_i y P_j son los perímetros de R_i y R_j , y W es el número de límites (Un pixel en otro lado tiene su magnitud diferente menor que un umbral T). El parámetro Q_1 controla el tamaño de la región a ser unida. por ejemplo $Q_1 = 1$ implica que dos regiones serán unidas si solamente una de las regiones casi rodea a la otra.

2.- Unir R_i y R_j si $W/l > Q_2$, donde l es el largo del límite común entre dos regiones. Tipicamente $Q_2 = 0.75$. Este paso se aplicará después del paso 1.

3.- Unir R_i y R_j si una distancia similar es menor que un umbral T



Un ejemplo de esta técnica se observa en la siguiente figura 4.19.

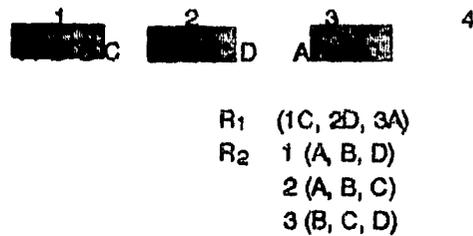
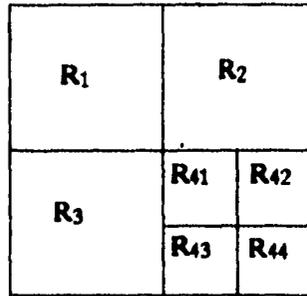
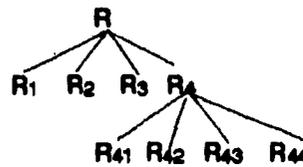


Figura 4.19 Proposición basada en Región y agrupamiento

Si R representa la imagen entera. Para una imagen cuadrada, para segmentar R es necesario subdividir esta sucesivamente hasta una región cuadrada más pequeña. Esta particular técnica de partición tiene una conveniente representación en la forma llamada árbol cuadrado (Esto es un árbol en el cual cada nodo tiene exactamente cuatro descendientes) ilustrado en la figura 4.20 (En este caso solamente R_4 fue subdividido).



a)



b)

Figura 4.20 a) Imagen Particionada b) Cuadratura correspondiente

En el siguiente capítulo se realizará la implantación de los métodos de identificación de figuras geométricas y de piezas mecánicas. Para realizar esta identificación se utilizará la teoría proporcionada en este capítulo.

Para la identificación de figuras y piezas mecánicas se utilizarán los siguientes métodos: Método de Umbral para la detección de bordes, Gradiente y normal, Operadores gradiente y compas para la detección de bordes, la transformada Hough, y de las técnicas de segmentación de imágenes digitales se utilizarán todas excepto una (componente rotulado "análisis de conectividad"; pero si se utilizará componente pixel rotulado).

La técnica de análisis de conectividad es útil para segmentar una imagen que contenga varios niveles de gris en su interior o que tenga una o más regiones

de un determinado valor de gris en el interior de la figura, o también puede ser utilizada en objetos en los que existan cambios abruptos en su contorno. Las imágenes que se utilizaran para realizar la identificación de figuras y de piezas mecánicas tendrán un solo valor de gris y no tendrán orificios en la figura, por lo que no es necesario utilizar la técnica de segmentación análisis de conectividad; pero todos los métodos descritos en este capítulo se utilizarán en los capítulos siguientes.

5 IMPLEMENTACION DE LOS METODOS DE IDENTIFICACION DE FORMAS

5.1 INTRODUCCION

El objetivo de la tesis es la identificación de objetos, para esta identificación se analizará el contorno de las figuras por regiones, las cuales pueden tener formas geométricas diversas.

El primer paso para la identificación es conocer el borde o contorno de las figuras, para obtener una imagen de bordes, para ello se programaron todos los métodos de detección de bordes analizados en el capítulo 4, el segundo paso es aplicar el método del "operador Umbral" para obtener una imagen de bordes binaria, y por último, el tercer paso es aplicar las técnicas de segmentación de imágenes para realizar la identificación de las figuras geométricas.

En el punto 5.2 se muestra el análisis de todos los métodos de detección de bordes para realizar una selección de algunos de ellos y poder utilizarlos posteriormente en la identificación de las figuras geométricas. En el punto 5.3 se muestra la implementación de los métodos de identificación de figuras. En cada uno de los puntos anteriores se muestra el algoritmo de programación. El punto 5.6 muestra una aplicación a objetos industriales más complejos (pinzas, candados, desarmadores y martillos).

5.2 ANALISIS Y SELECCION DE LOS METODOS OPERADORES DETECTORES DE BORDES

Para obtener una imagen de bordes se programaron todos los métodos de detección de bordes descritos en el capítulo anterior. Esos 14 métodos de detección de bordes se probaron para seis imágenes diferentes: esferas encontradas, figuras geométricas, una imagen de objetos, un estadio en forma de pentágono y una imagen de una mujer.

Los 14 métodos de detección de bordes que se muestran, no son todos los métodos de detección de bordes que se pueden encontrar, pero son los suficientes para permitir hacer una selección adecuada de algunos de ellos para utilizarlos posteriormente en la identificación de figuras.

A continuación se resume el análisis de los diferentes métodos de detección de bordes, con la finalidad de realizar una selección de los más convenientes para

utilizarlos en la identificación.

OPERADOR DE DIFERENCIA DE PÍXELES.- Este operador tiene dos máscaras de 2×2 y es muy sencillo para programarse y consiste en obtener la diferencia entre el píxel de cálculo y el píxel adyacente superior y la diferencia entre el píxel de cálculo y el píxel adyacente a la derecha, este operador tiene algunos problemas para la identificación de esquinas, las imágenes resultantes al aplicar este operador tienen una excelente resolución para las 6 imágenes de prueba.

OPERADOR DE DIFERENCIA SEPARADA DE PÍXELES.- Este operador tiene dos máscaras e involucra el valor de muy pocos píxeles para su cálculo (izquierda, derecha, inferior y superior del píxel donde se estén efectuando los cálculos $\langle i, j \rangle$), las máscaras de este operador son de 3×3 pero cada una envuelve el valor de únicamente dos píxeles para el cálculo, por lo que no permite tener una comparación entre más píxeles vecinos al píxel de cálculo. El borde que se obtiene con este operador es un borde doble, las imágenes resultantes con este operador tienen una buena resolución para las 6 imágenes.

OPERADOR GRADIENTE ROBERTS.- Este operador tiene dos máscaras de 2×2 y es muy sencillo de programarse sus máscaras tienen orientación de 45° y -45° por lo que permite localizar más fácilmente las esquinas de las figuras geométricas. El borde que se obtiene con este operador es un borde sencillo. Las imágenes resultantes al aplicar este operador tienen una excelente resolución para las 6 imágenes.

OPERADOR LAPLACIANO.- Este operador tiene una sola máscara y utiliza la expresión matemática de la segunda derivada parcial o laplaciano. Observando la máscara de este operador consiste en obtener la diferencia entre cada uno de los píxeles adyacentes y el píxel de cálculo (superior, inferior, izquierda y derecha), el borde producido por este operador es un borde doble. Las 3 primeras imágenes resultantes al aplicar este operador tienen una buena resolución pero las siguientes tres imágenes no la tienen, esto se debe a que este operador produce un borde doble y también por el tipo de máscara que maneja.

OPERADOR KIRSH 2 MASCARAS.- Este operador tiene 2 máscaras y no involucra el valor del píxel en el cálculo, es decir, que no considera el valor del píxel que se está calculando, pero involucra el valor de los 8 píxeles adyacentes, este operador maneja el valor de altos coeficientes (5 y 3) por lo que es más fácil la identificación al aplicar el operador umbral. Las imágenes resultantes al aplicar este operador tienen una excelente resolución.

OPERADOR SOBEL DE 2 MÁSCARAS.- Este operador tiene 2 máscaras de 3x3 y no involucra el valor del pixel en el cálculo. La resolución de las imágenes resultantes no es buena debido al tipo de máscara que tiene, esto se debe a que involucra el valor de 6 pixeles para su cálculo y también porque el coeficiente 2 no es el adecuado, esto se comprueba observando el resultado al aplicar un coeficiente $2^{1/2}$ (Operador Frei-Chen) o un coeficiente 1 (Operador Prewitt).

OPERADOR PREWITT DE 2 MÁSCARAS.- Este operador produce un borde doble, las imágenes resultantes al aplicar este operador tienen una excelente resolución.

OPERADOR FREI-CHEN.- Este operador produce un borde doble. Las imágenes resultantes al aplicar este operador tienen una buena resolución.

OPERADOR DE DIFERENCIA DE PÍXELES 8 NIVELES.- Este operador tiene una máscara y consiste en obtener la diferencia de todos los pixeles adyacentes al pixel de cálculo y este último multiplicado por 8, la resolución de las imágenes resultantes al aplicar este operador se observa que para las 3 primeras es muy buena, pero para las tres últimas es mala debido a su tipo de máscara, ya que involucra el valor de muchos pixeles para su cálculo y las 3 últimas imágenes contienen muchos niveles de grises, por lo que para aplicar esta máscara no deben existir muchos valores diferentes de gris en la imagen, esta conclusión se comprueba observando las imágenes de los objetos y de las figuras geométricas originales (debido a que estas no contienen muchos niveles de gris).

Las imágenes resultantes al aplicar el operador compass, es decir la generalización a 8 máscaras del operador gradiente, tienen un borde más grueso, esto puede mejorar o afectar la resolución de las imágenes. Las mejores imágenes resultantes fueron las de los operadores Kirsh y Prewitt de 8 máscaras, los operadores de diferencia de pixeles y el operador Sobel de 8 máscaras producen imágenes con una resolución inferior que las de los operadores anteriores.

Los métodos de 8 máscaras son más lentos que los de 2 máscaras (en tiempo de procesamiento). Con el análisis de cada uno de los métodos detectores de bordes, se concluye que los mejores métodos de detección de bordes son:

- Operador Gradiente Roberts
- Operador de diferencia de pixeles de 2 máscaras
- Operador Kirsh
- Operador de diferencia de pixeles separada

Cualquiera de los operadores anteriores pueden ser utilizados para la identificación de figuras, pero el Gradiente Roberts resulta ser el más adecuado por su simplicidad y efectividad.

Para obtener la imagen de bordes es necesario aplicar el "operador Umbral" a la imagen resultante después de aplicar el operador de detección de bordes

En las siguientes páginas se muestran las imágenes resultantes al aplicar cada uno de los operadores detectores de bordes a las 6 imágenes descritas en este análisis. La primera página corresponde a las imágenes originales.

5.2.1 ALGORITMOS DE PROGRAMACION

A continuación se muestra el algoritmo para los operadores detectores de bordes:

- 1.- Se leen los primeros 2 ó 3 renglones de la imagen y se guardan en 2 ó 3 arreglos vectoriales.
- 2.- Se aplican las máscaras respectivas para cada tipo de operador, y se almacena el resultado en otro archivo
- 3.- Se lee el siguiente renglón y se repiten los puntos 2 y 3 hasta terminar toda la imagen.

El algoritmo para el programa que muestra en pantalla las imágenes es el siguiente:

- 1.- Se realiza una función para inicializar en modo gráfico.
- 2.- Se leen hasta 4 archivos diferentes (imágenes) y se muestra cada uno en pantalla .

IMAGENES DE PRUEBA

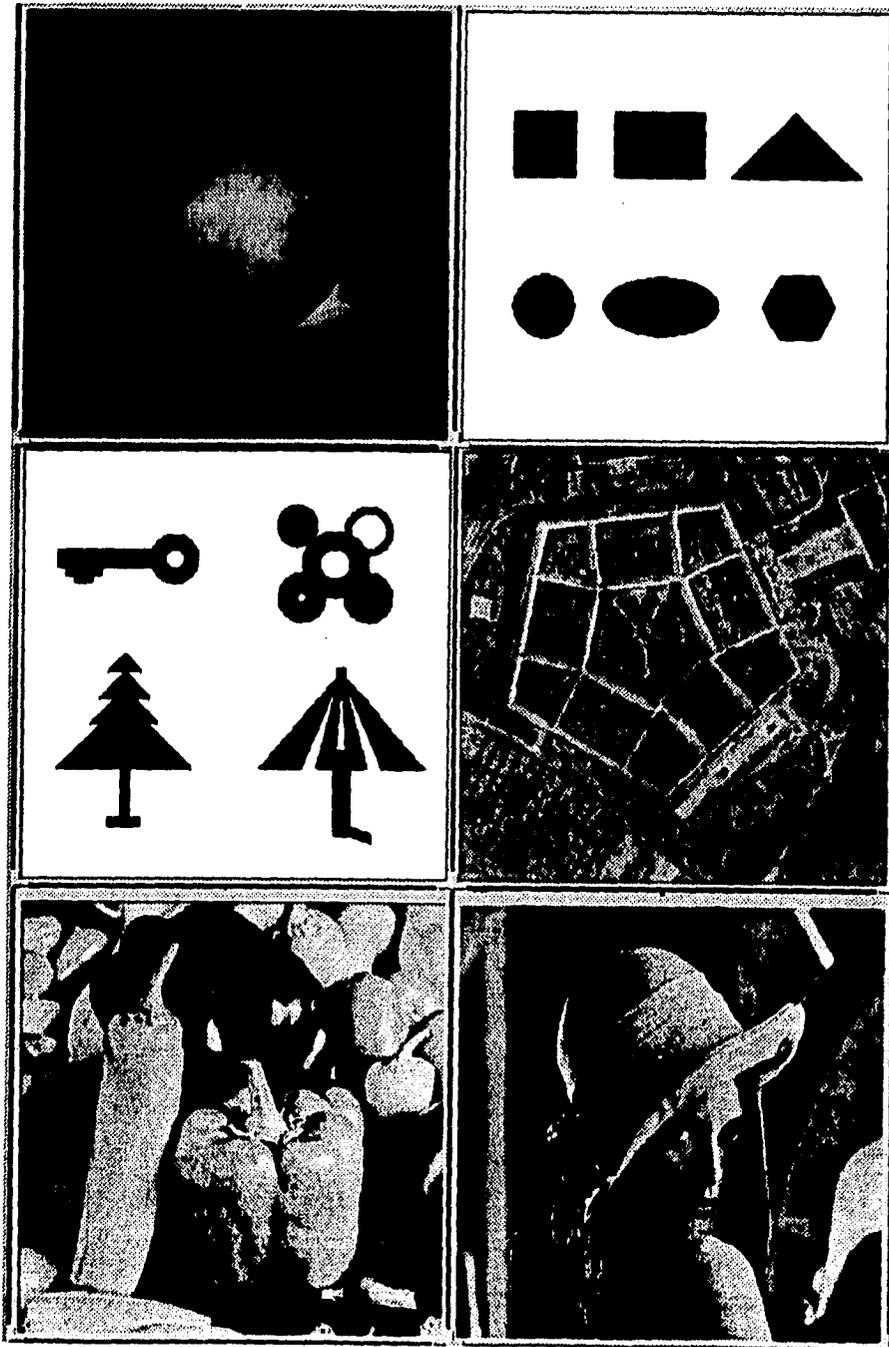


Figura 5.1 Imágenes Originales

METODO DETECTOR DE BORDES

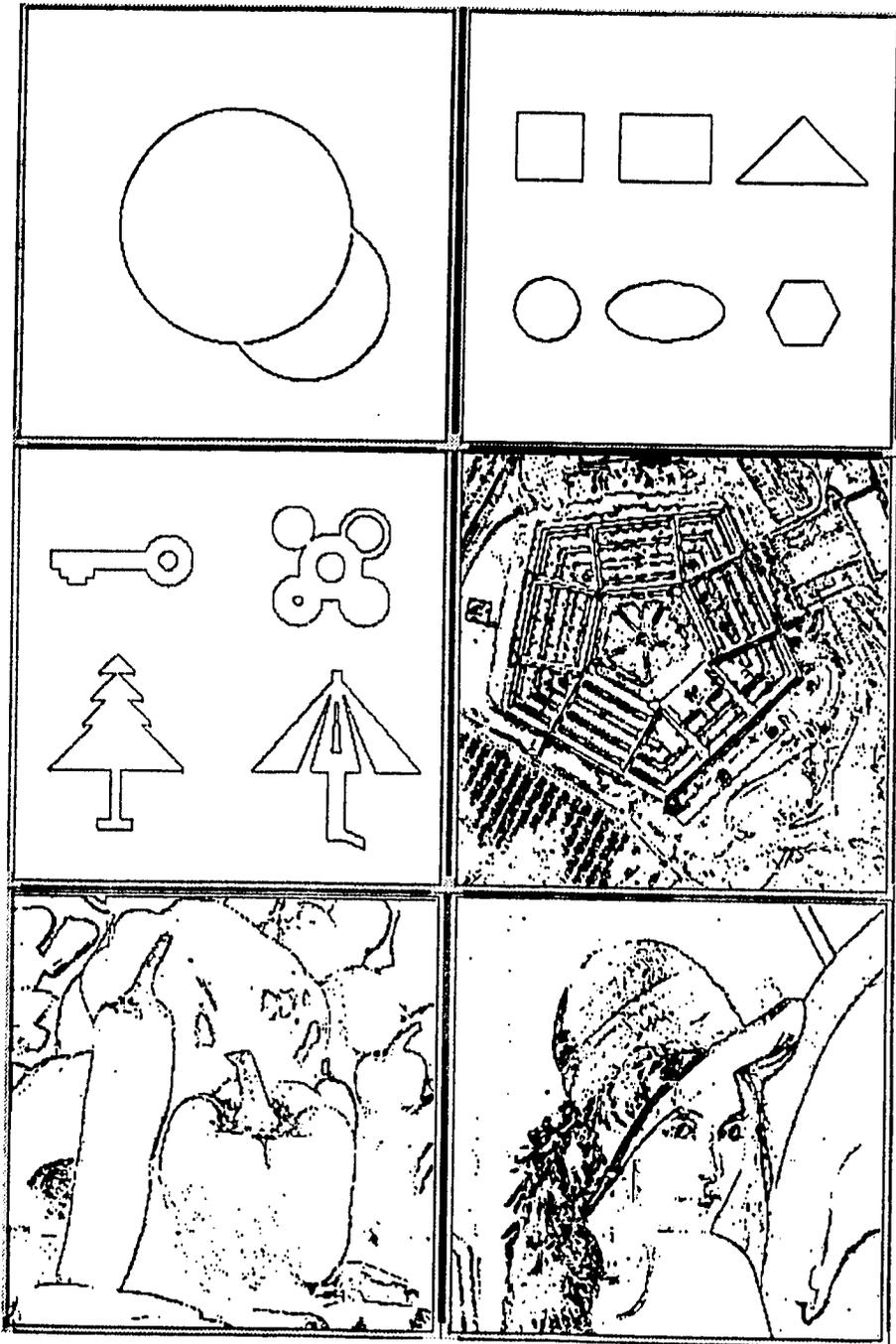


Figura 6.2 Operador de diferencia de pixeles 53

METODO DETECTOR DE BORDES

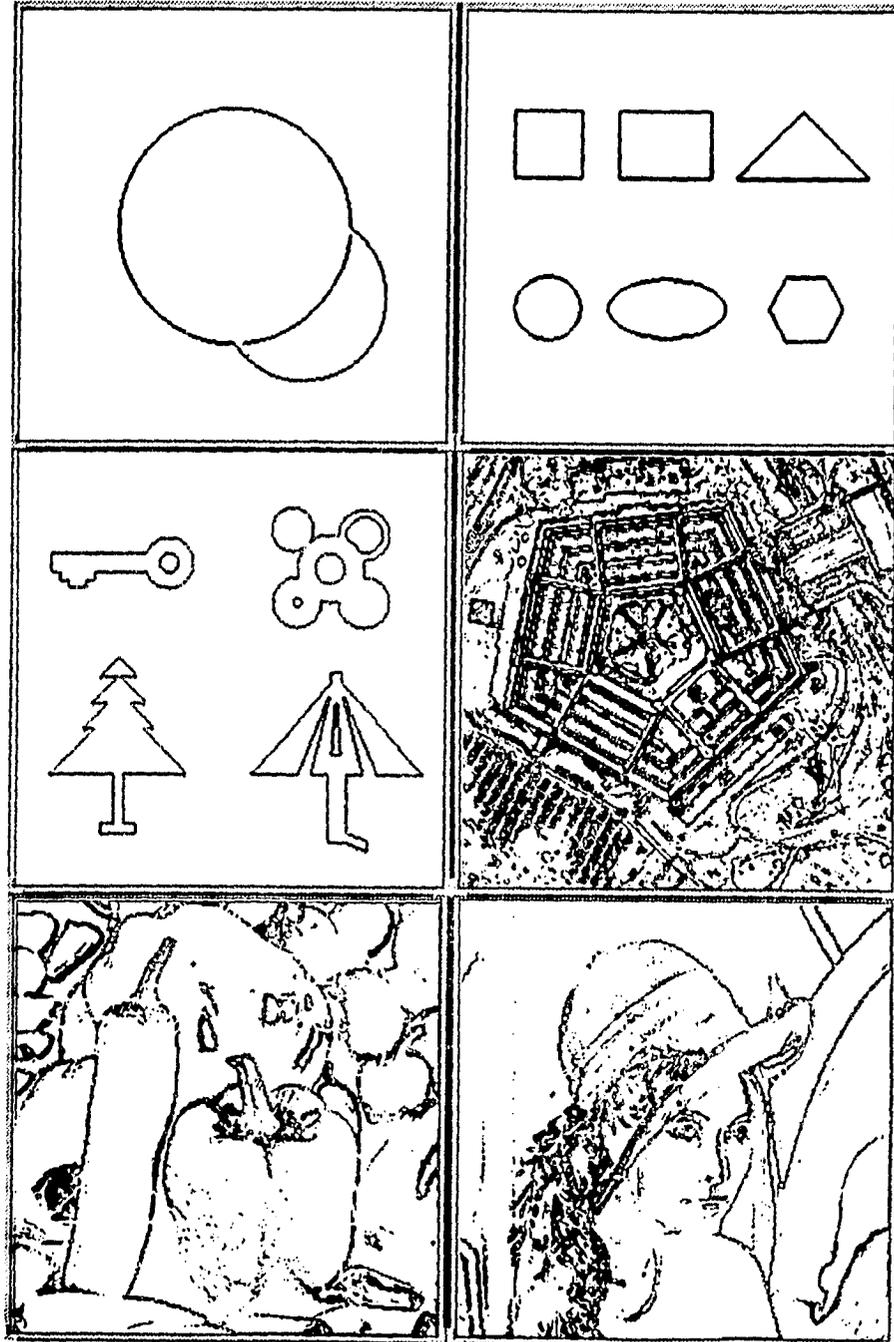


Figura 53 Diferencia separada de pixeles

METODO DETECTOR DE BORDES

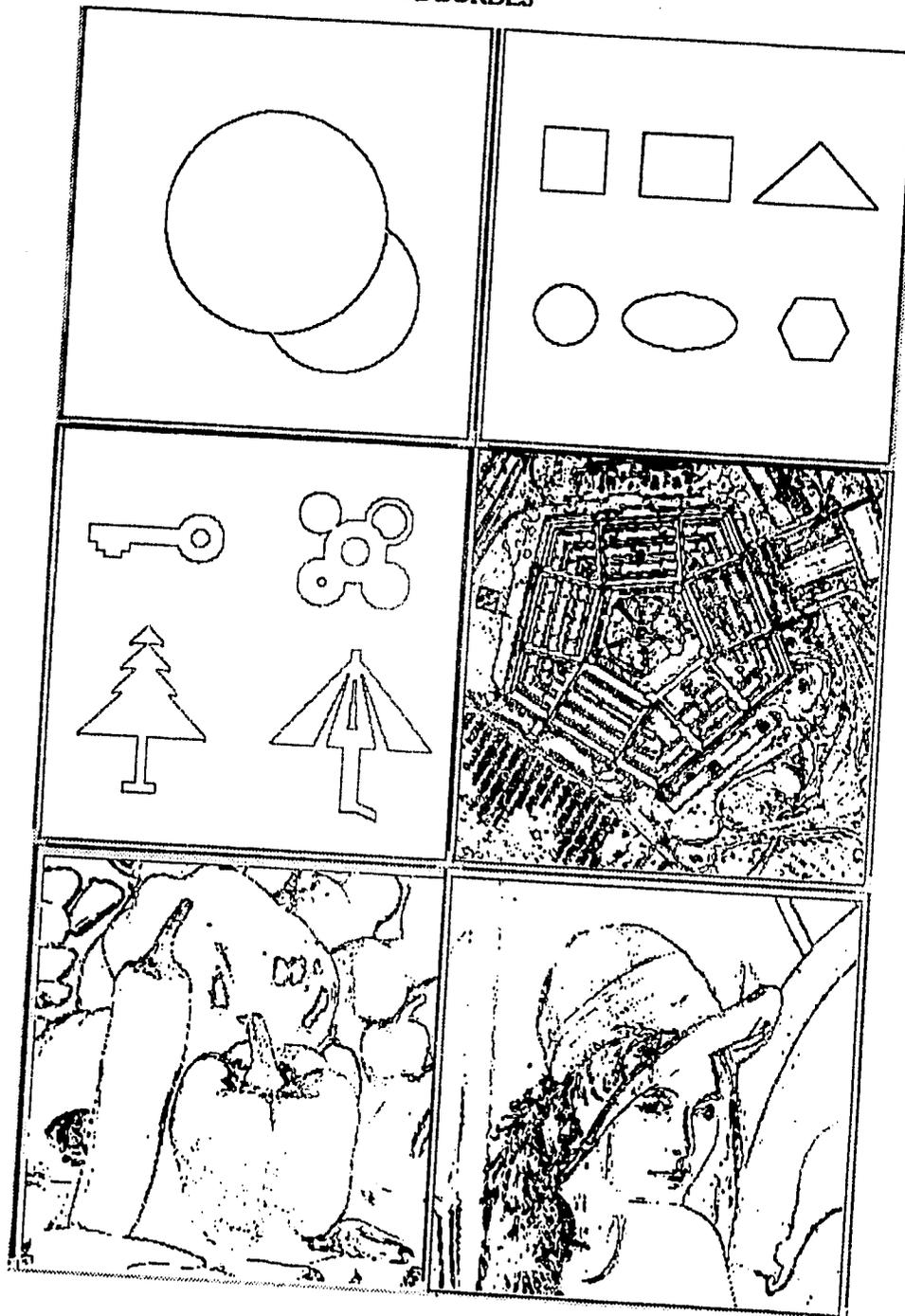


Figura 5.4 Operador Gradiente Roberts

METODO DETECTOR DE BORDES

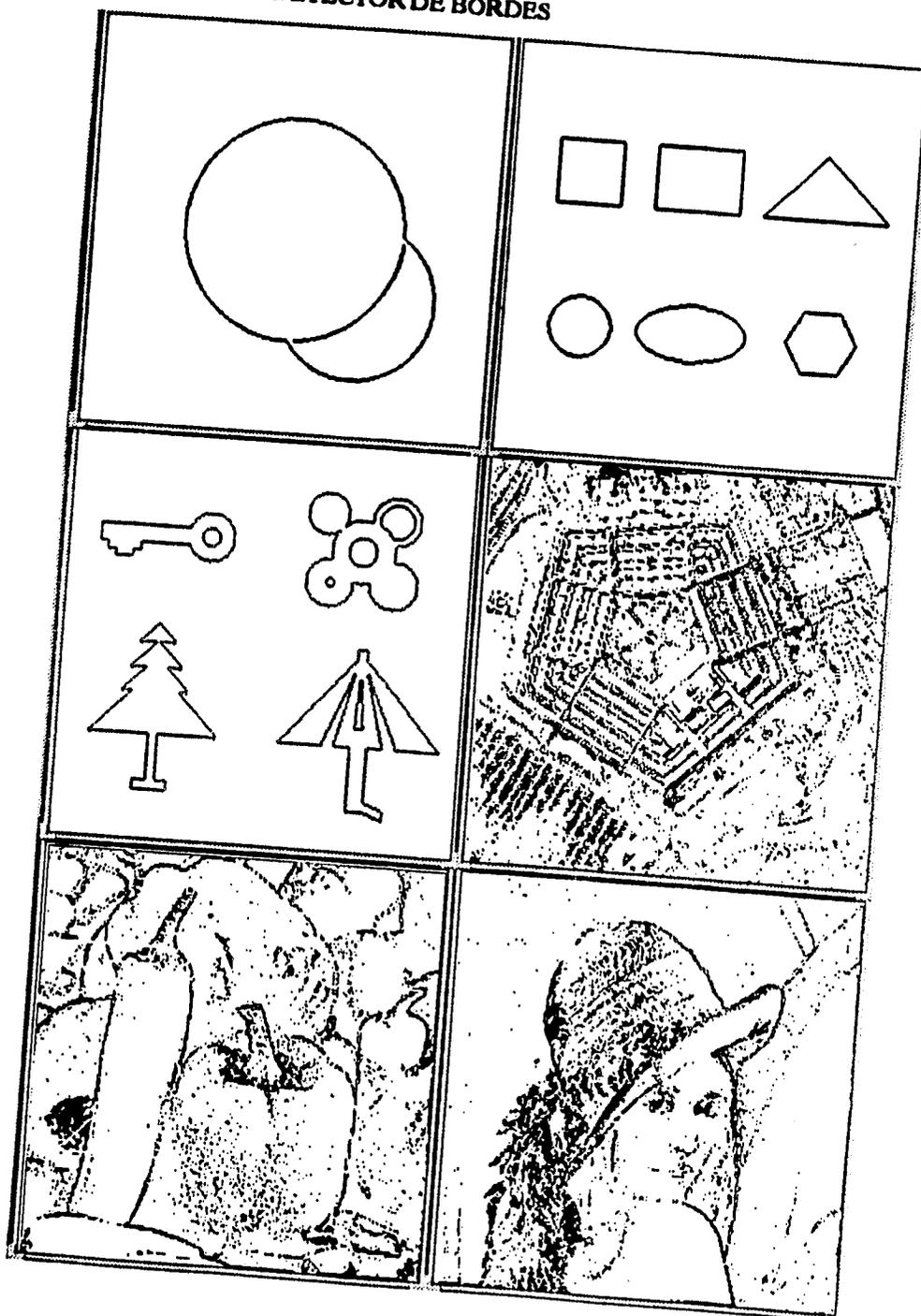


Figura 5.5 Operador Laplaciano

METODO DETECTOR DE BORDES

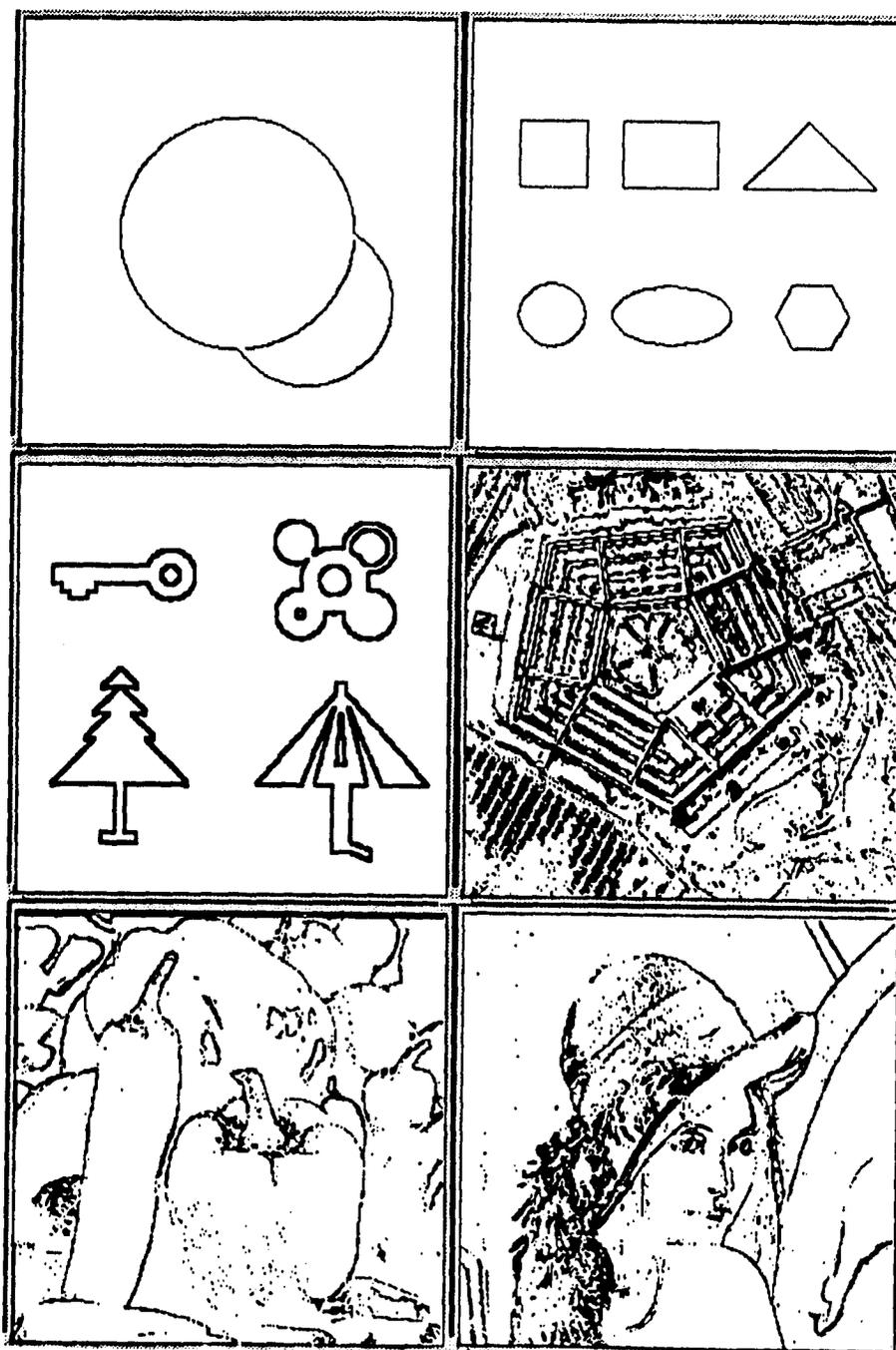


Figura 66 Operador kirsh, 2 Máscaras

METODO DETECTOR DE BORDES

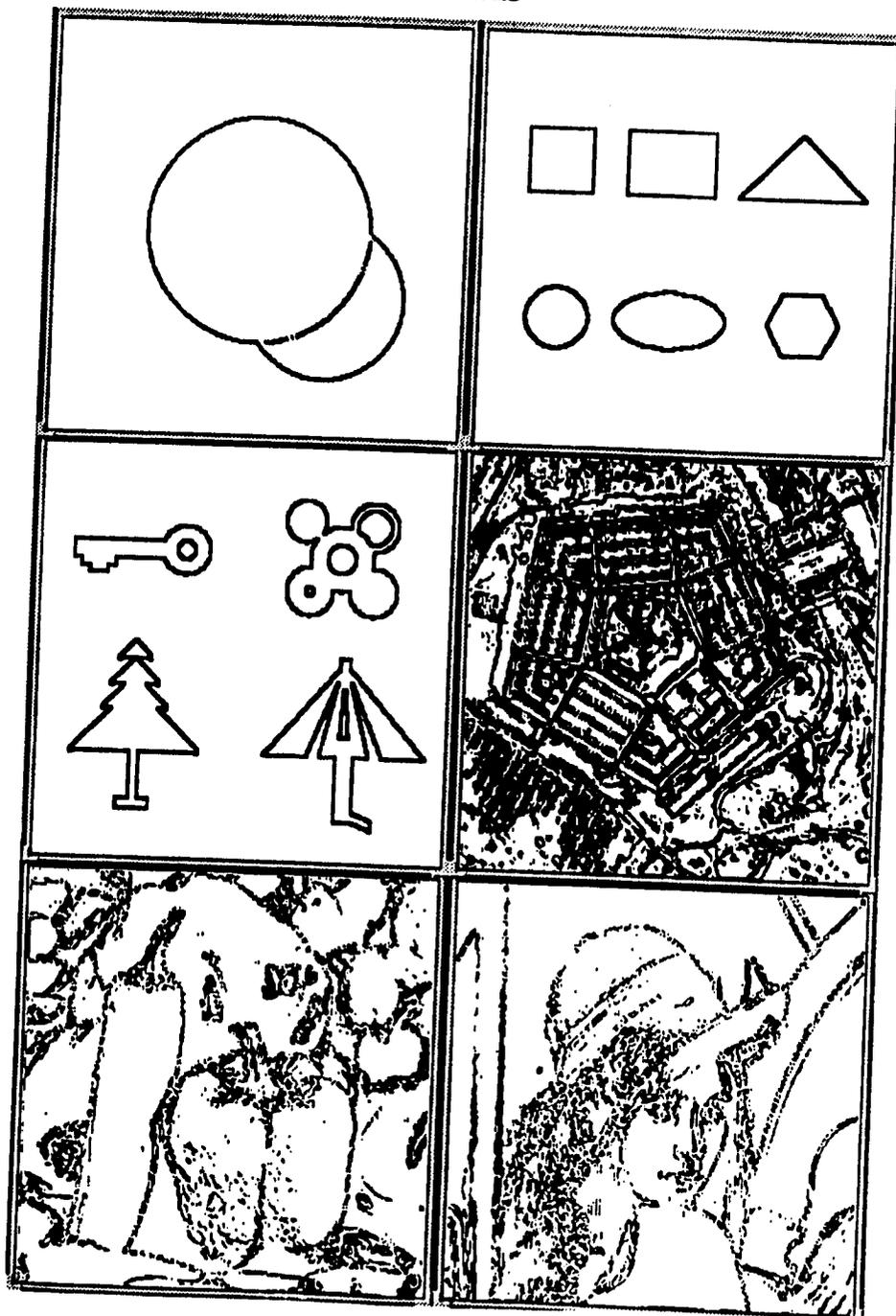


Figura 5.7 Operador Sobel, 2 Máscaras

METODO DETECTOR DE BORDES

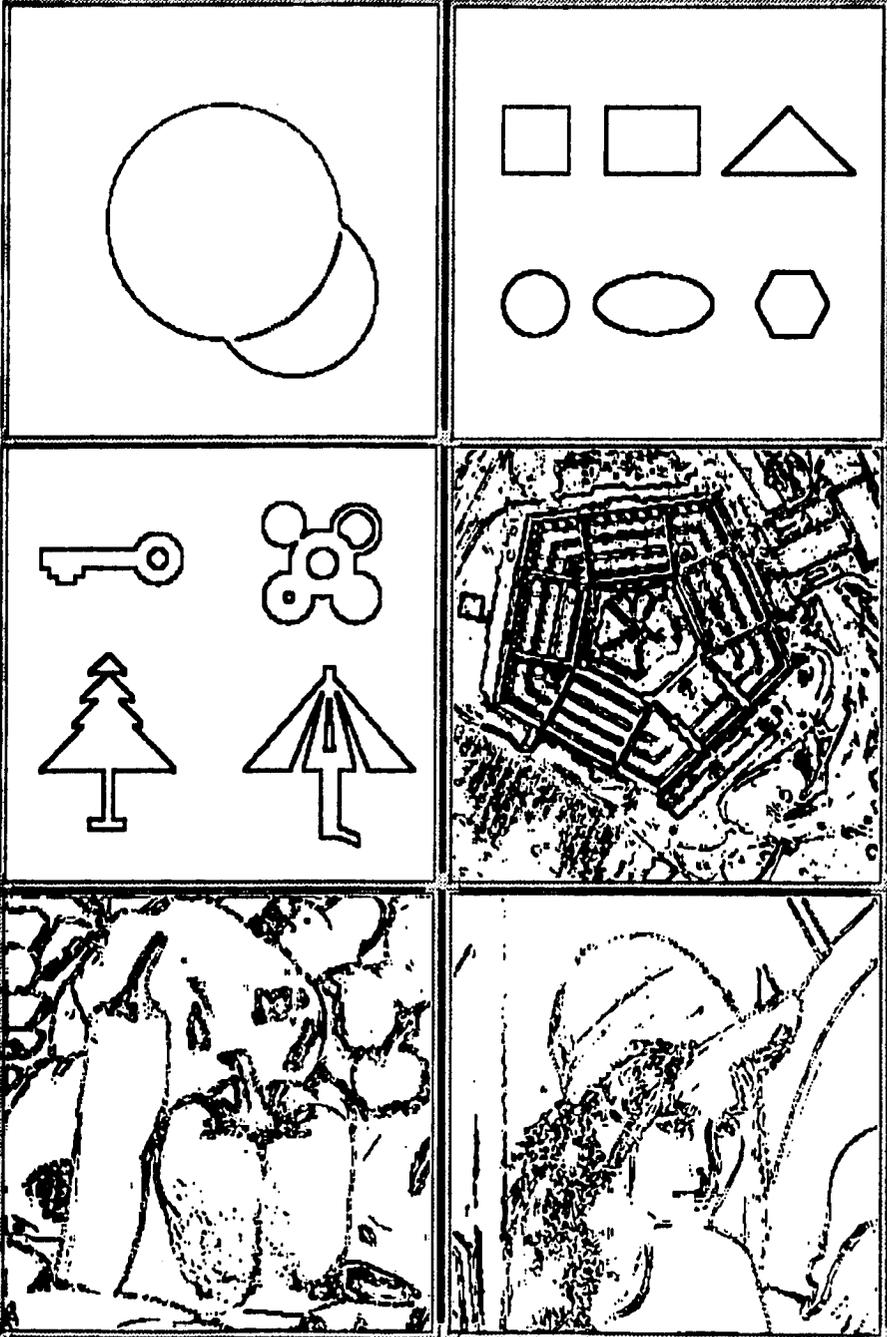


Figura 5.8 Operador Prewitt, 2 Máscaras

METODO DETECTOR DE BORDES

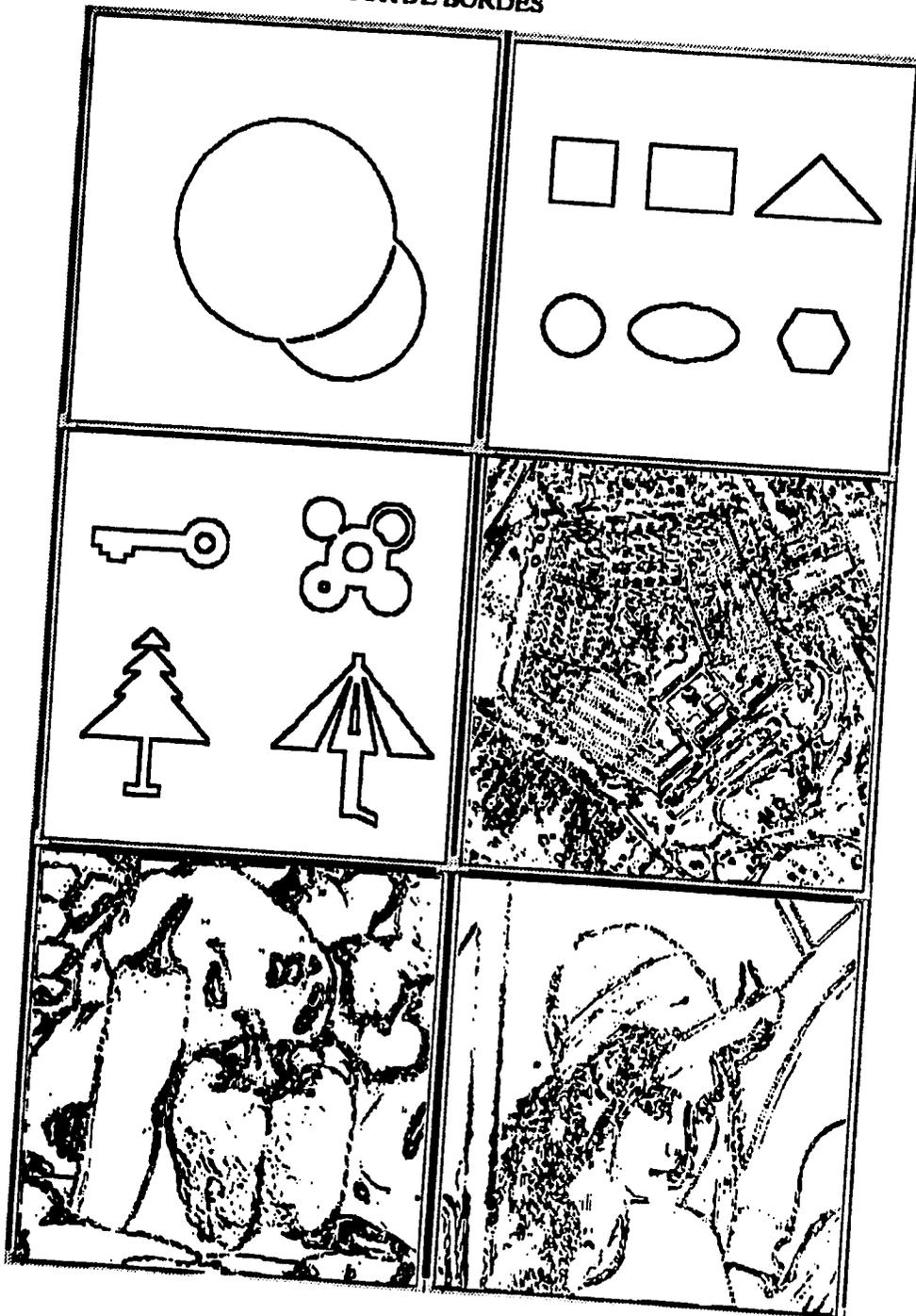


Figura 5.9 Operador Frei-Chen, 2 Máscaras. 60

METODO DETECTOR DE BORDES

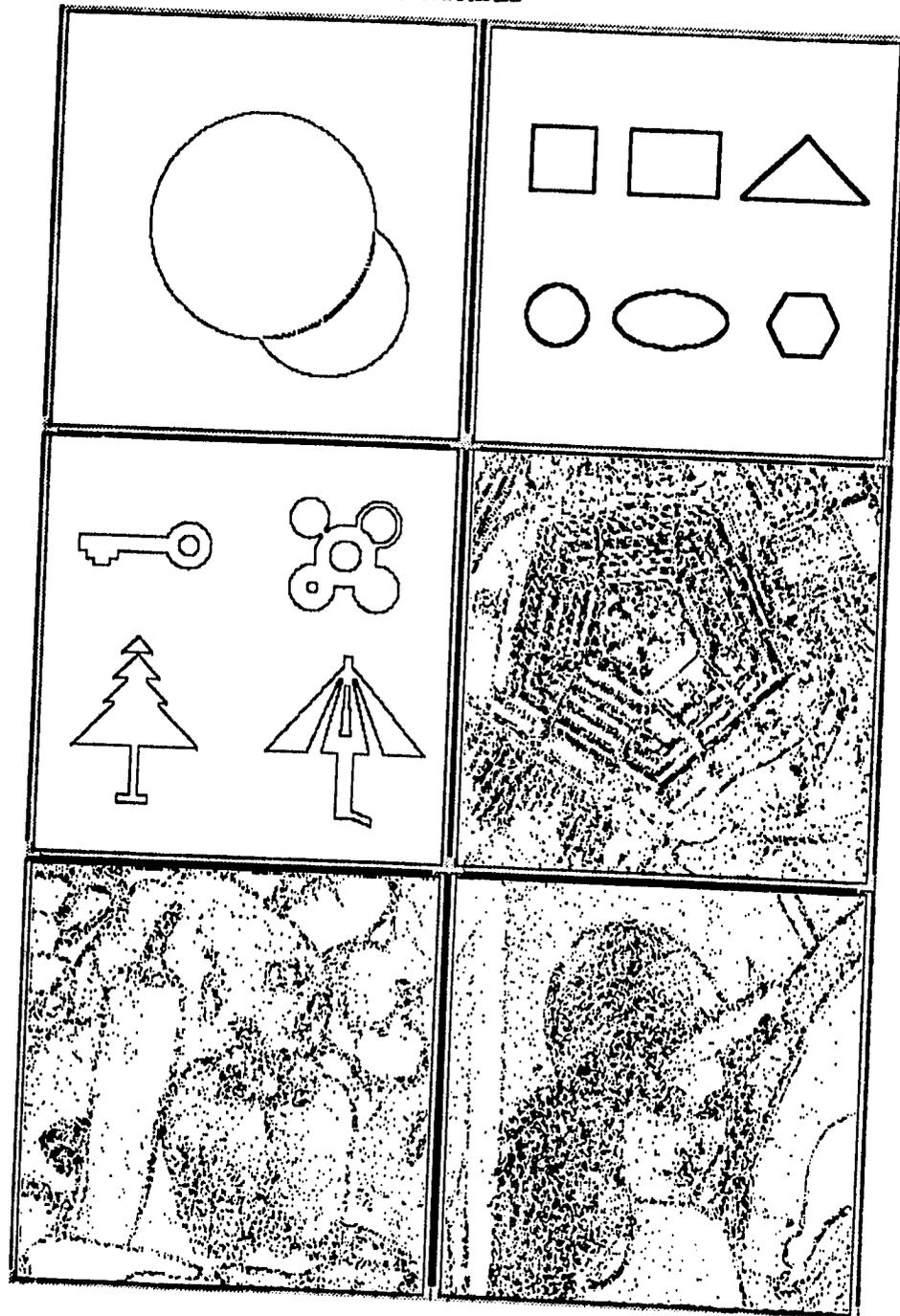


Fig 5.10 Diferencia separada de pixeles, 8 Nvls. 61

METODO DETECTOR DE BORDES

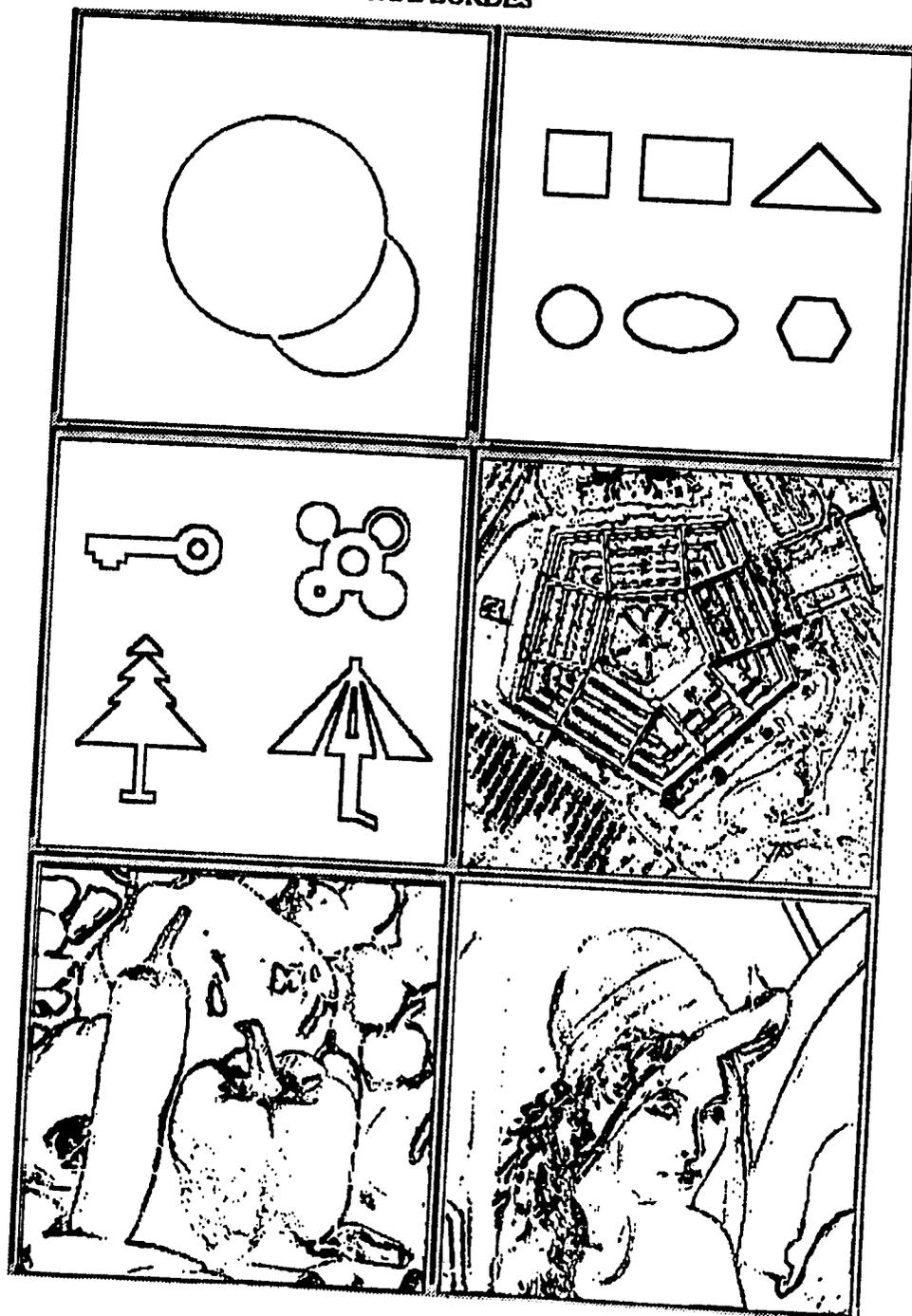


Figura 5.11 Operador Kirsh, 8 Máscaras

METODO DETECTOR DE BORDES

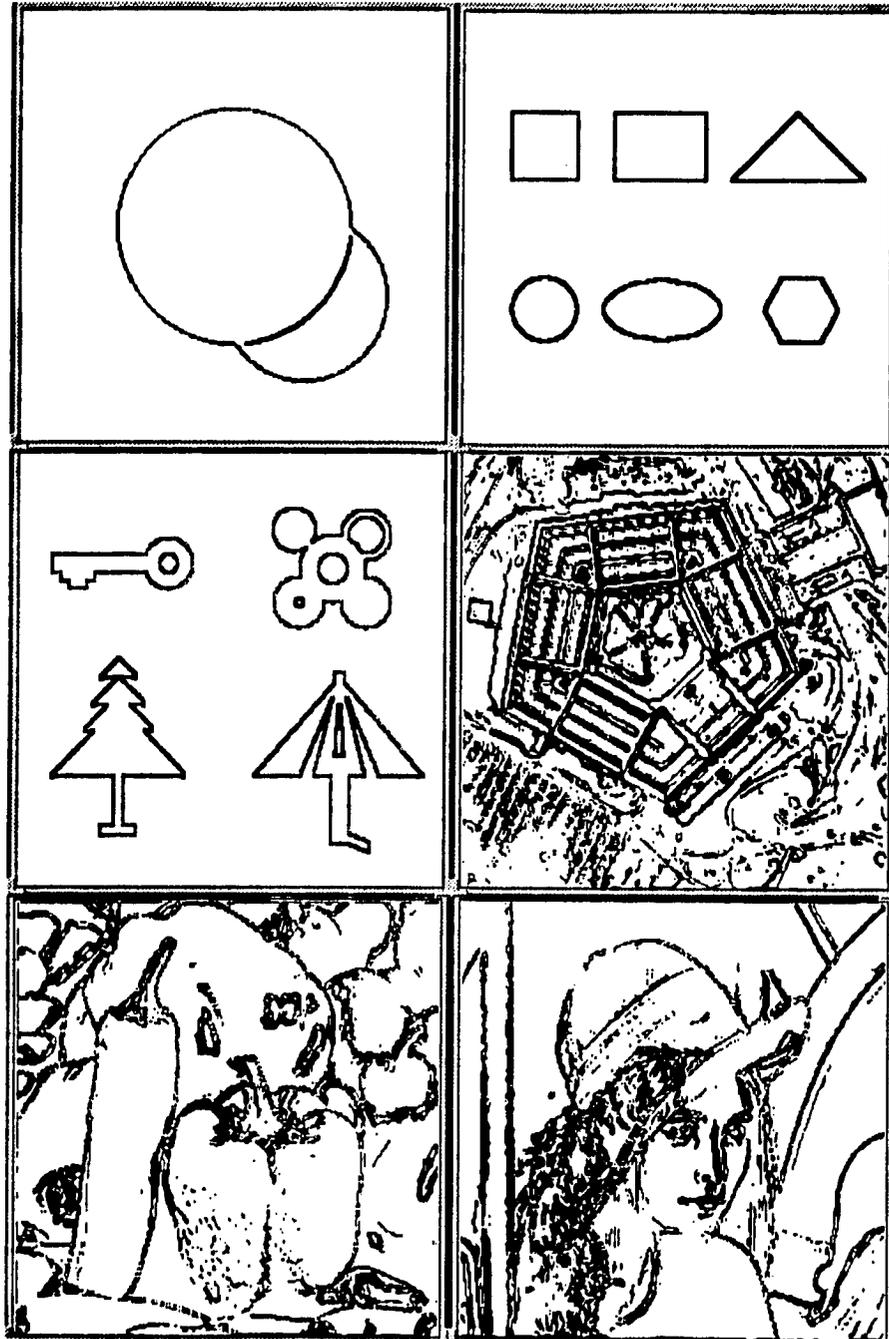


Figura 5.12 Diferencia de pixeles, 8 Máscaras 63

METODO DETECTOR DE BORDES

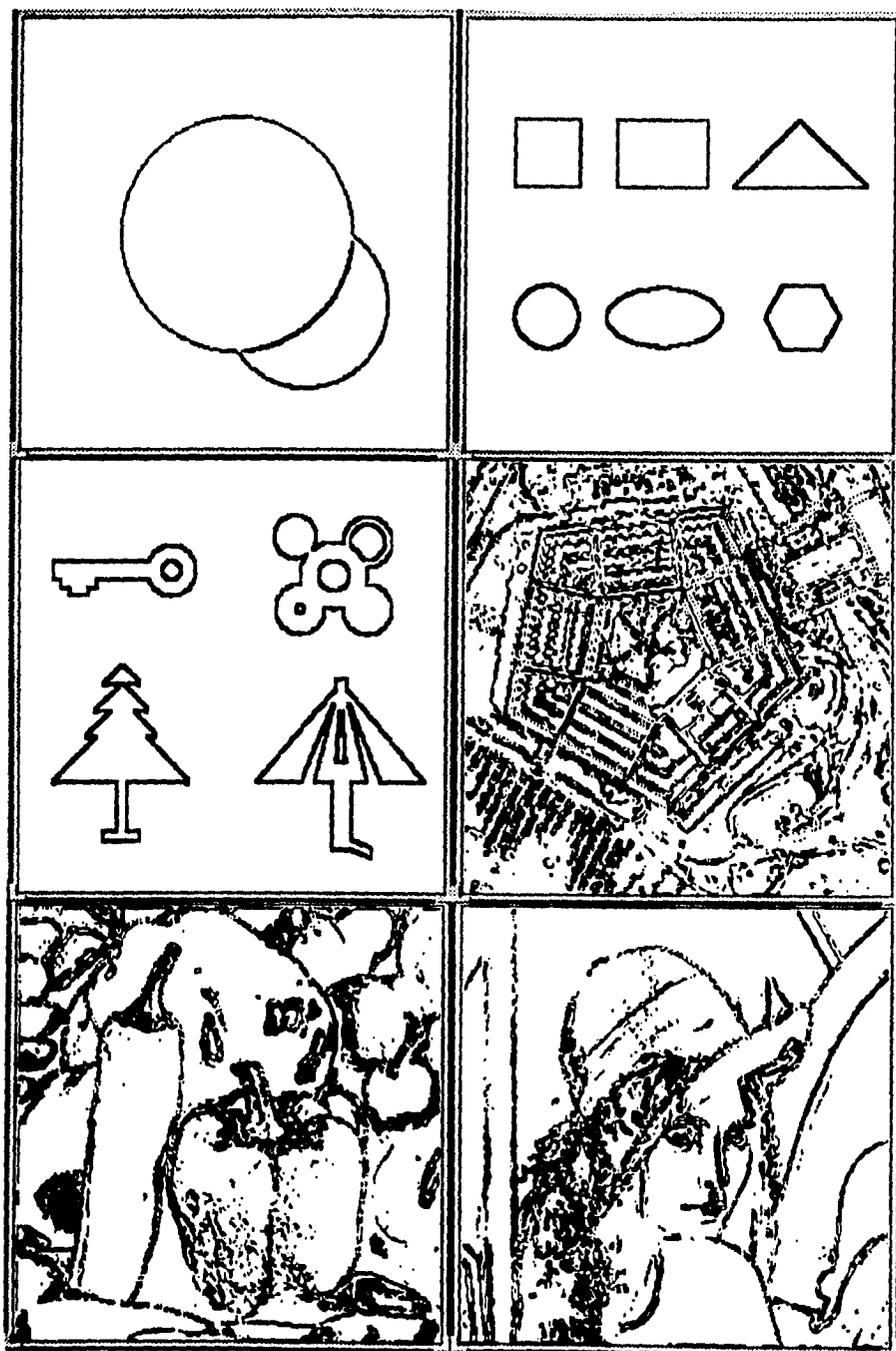


Figura 5.13 Operador Sobel, 8 Máscaras

METODO DETECTOR DE BORDES

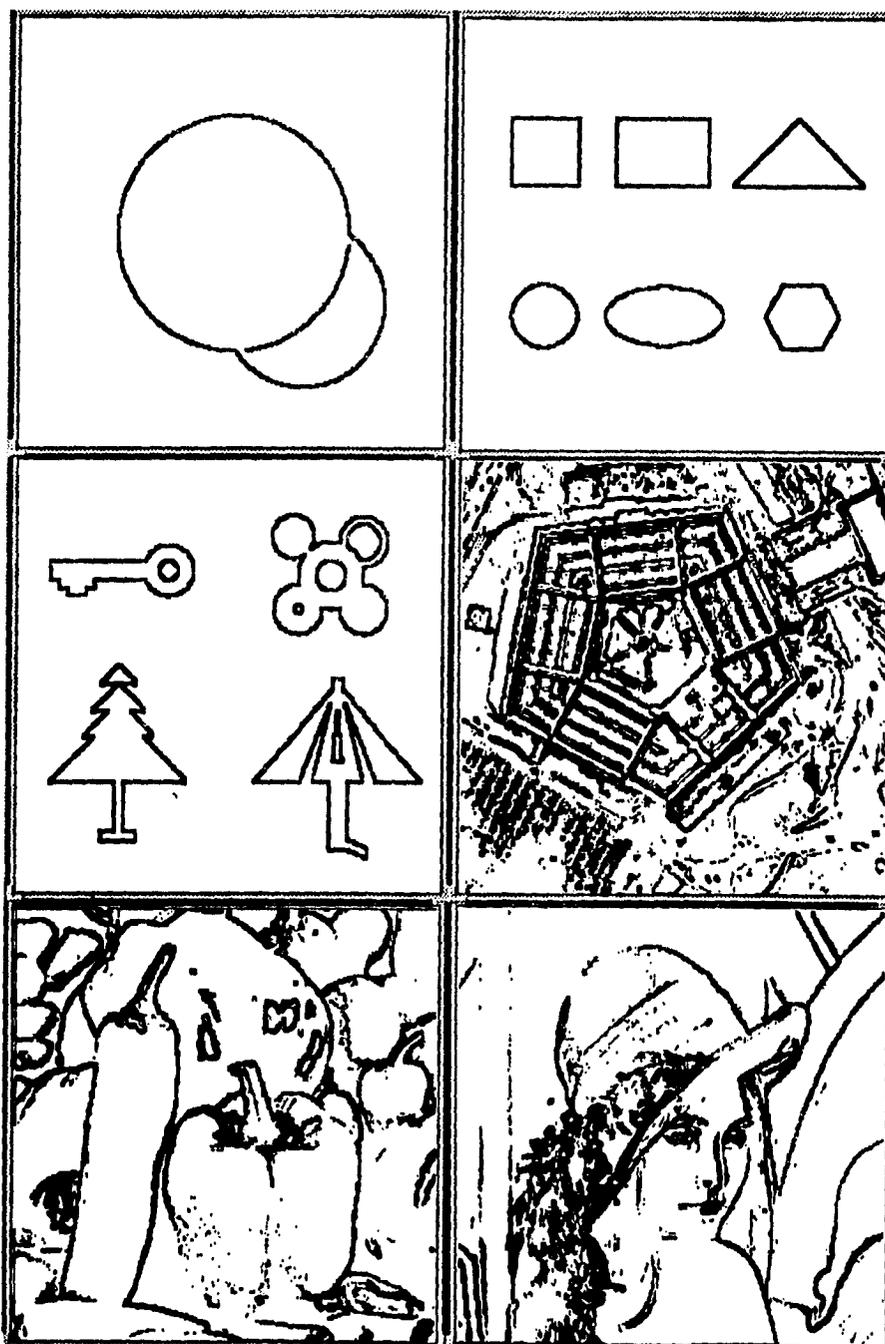


Fig. 5.14 Operador Prewitt, 8 Máscaras, 3 Nvlis. 65

METODO DETECTOR DE BORDES

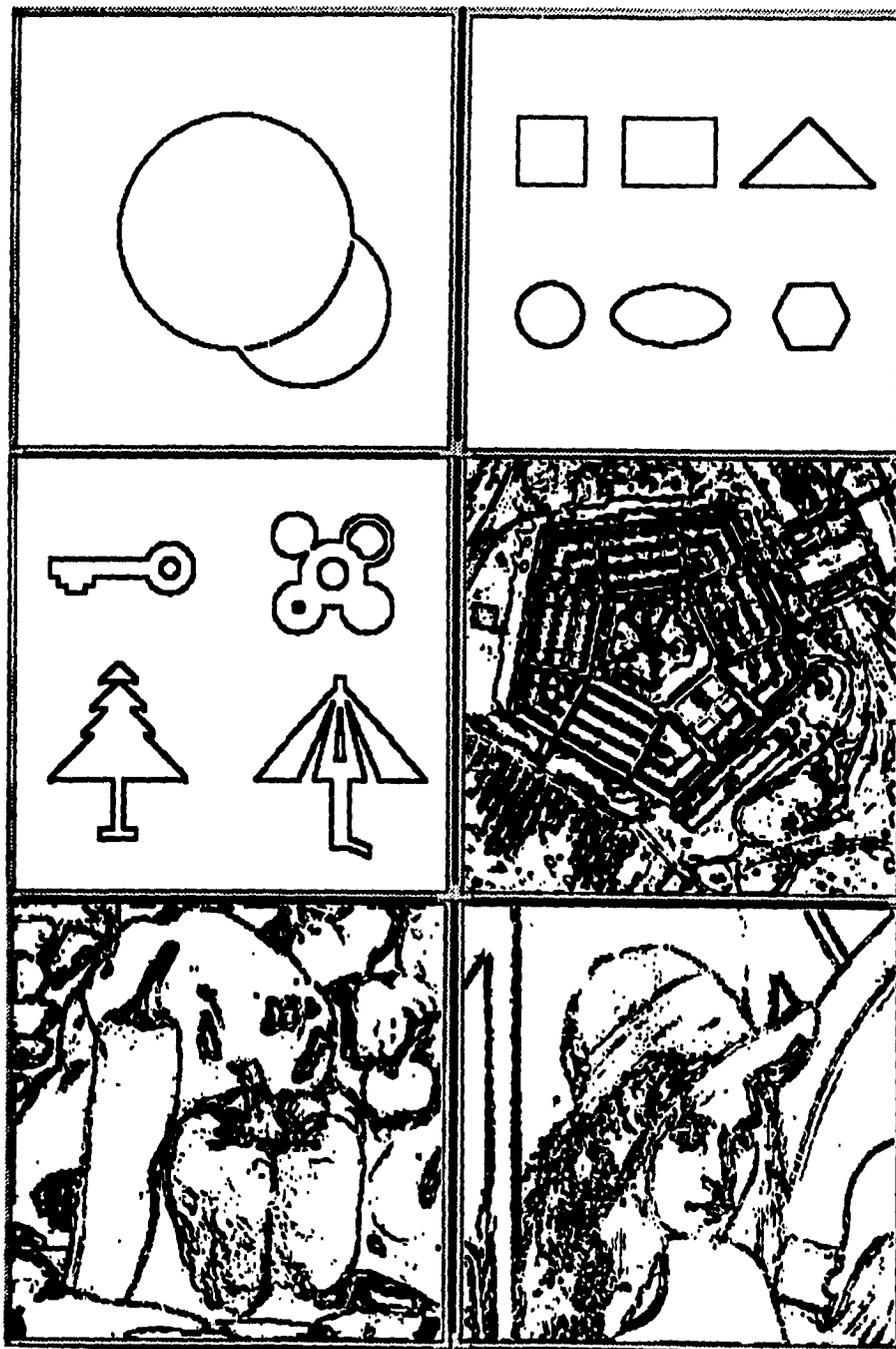


Fig. 5.15 Operador Prewitt, 8 Máscaras, 5 Nvs. 86

5.3 IMPLEMENTACION DE LOS METODOS PARA LA IDENTIFICACION DE FORMAS

De las técnicas de segmentación descritas en el capítulo 4 (Amplitud de umbral, componentes etiquetados <pixel rotulado y análisis de conectividad>, contorno siguiente y proposición región y agrupamiento), se programaron todas excepto una (análisis de conectividad), porque esta técnica realiza la partición a figuras u objetos que tengan agujeros, como es el caso de una tuerca.

Para realizar la identificación de las figuras se elaboraron 4 programas de cómputo:

El primer programa identifica la figura geométrica que se le proporciona (cuadrado, rectángulo, círculo, triángulo, hexágono o una elipse) y muestra sus características más importantes dependiendo del tipo de figura que se trate (diámetro, área, pixel centro, base, lado mayor, etc.). Para este programa no se necesita tener la imagen original binaria y la imagen de bordes binaria, cualquiera de las dos imágenes anteriores que se tenga será suficiente para identificar el tipo de figura. Esto se debe a la técnica de segmentación que se utilizó (proposición componentes etiquetados "pixel rotulado"), la cual identifica el borde de la figura explorando toda la imagen de izquierda a derecha y de arriba a abajo, el pixel de cálculo es comparado con sus vecinos (pixeles adyacentes), si se cumple la propiedad de borde se etiqueta y se identifican todos los bordes, posteriormente se realiza una comparación con la base de datos ya establecida en el programa, y finalmente se proporciona la información correspondiente a este tipo de figura.

El segundo programa que se elaboró utiliza dos técnicas de segmentación: contorno siguiente y proposición de región y agrupamiento. Este programa primero identifica todas las figuras que existen, proporciona la información de las figuras existentes y por último pregunta que figura se quiere buscar o localizar. Para este programa se necesitan las dos imágenes (la imagen original y la imagen de bordes), con la imagen original se divide la imagen por regiones, esta partición encierra inicialmente cada figura lo que permite conocer más fácilmente los pixeles bordes de cada figura, a continuación se aplica la técnica de contorno siguiente para identificar cada figura.

En el segundo programa después de identificar las figuras existentes muestra, por ejemplo, el siguiente menú y respuestas:

SE TIENEN LAS SIGUIENTES FIGURAS

ELIPSE
CUADRADO
RECTANGULO
HEXAGONO

QUE QUIERE DETECTAR

- 1.- CUADRADO
- 2.- RECTANGULO
- 3.- TRIANGULO -- NO EXISTE
- 4.- ELIPSE
- 5.- CIRCULO -- NO EXISTE
- 6.- HEXAGONO

El tercer programa identifica la figura geométrica que se proporcione, pero esta puede tener una rotación o no. A este programa se le proporciona la imagen de bordes. Este programa utiliza la técnica de contorno siguiente, pixel rotulado, transformada Hough y un cálculo de pendientes para cada tramo de la imagen. Este programa identifica todas las figuras que se le proporcionen al programa 1.

El cuarto programa que se realizó localiza la figura que se pida buscar de entre 6 figuras que contenga la imagen. Las figuras pueden tener una rotación o no, (esta rotación puede ser de 0° a 360° para cualquier figura geométrica). Este programa es parecido al segundo, pero el segundo no identifica figuras con rotación. Este programa es muy extenso por lo que no identifica primero las figuras que existen (como en el programa 2), sino que, localiza la figura que se pida buscar. Este programa identifica todas las figuras que puede analizar el programa 2. A este programa se puede proporcionar ya sea la imagen original binaria, o bien la imagen de bordes binaria en monoresolución como en multiresolución (capítulo 6). Este programa utiliza las técnicas de segmentación: contorno siguiente, porposición de región y agrupamiento, pixel rotulado, amplitud de umbral, transformada Hough y un cálculo de pendientes para cada tramo de la imagen.

5.3.1 ALGORITMOS DE PROGRAMACION

El algoritmo para el programa que identifica el tipo de figura (programa 1) es:

- 1.- Se aplica un método de detección de bordes (operador Gradiente).
- 2.- Se aplica el "operador Umbral" a la imagen resultante del punto anterior, obteniendo la imagen de bordes.
- 3.- Se lee la imagen resultante renglón por renglón y se almacena en un

arreglo vectorial.

- 4.- Se obtienen los puntos extremos (superior izquierda, superior derecha inferior izquierda, inferior derecha, extremo izquierda, extremo derecha, extremo izquierda superior, extremo izquierda inferior, extremo derecha inferior, extremo derecha superior).
- 5.- Con la teoría de píxel rotulado y con los datos del punto 4, se compara la figura con cada tipo ya establecido en la base de datos (cuadrado, rectángulo, triángulo, círculo, hexágono y elipse).
- 6.- Dependiendo del tipo de figura de que se trate se calculan diferentes parámetros como son: píxel centro, área, diámetro, base, etc.
- 7.- Se muestra en pantalla la imagen de bordes pedida y los datos correspondientes a esta imagen.

Algoritmo del programa que localiza e identifica la figura solicitada (Programa 2)

- 1.- Se aplica un método de detección de bordes (operador Gradiente)
- 2.- Se aplica el operador umbral a la imagen resultante del punto anterior obteniendo la imagen de bordes.
- 3.- Se proporciona la imagen original y la imagen de bordes.
- 4.- Con la imagen original se particiona la imagen localizando cada figura, con la teoría de región y agrupamiento.
- 5.- Se localizan las coordenadas de la ubicación de cada figura que se localice en la imagen.
- 6.- Se identifica el tipo de figura con la teoría de contorno siguiente.
- 7.- Se proporciona la información de las figuras existentes (círculo, cuadrado, rectángulo, triángulo, hexágono y elipse), y el programa pide que figura se quiere localizar.
- 8.- El programa la identifica y la muestra en pantalla.
- 9.- El programa consta de 6 funciones (detección de círculos, detección de cuadrados, detección de rectángulos, detección de triángulos, detección de hexágonos y detección de elipses)

Algoritmo del programa que identifica el tipo de figura, (con rotación o sin rotación). (programa 3):

- 1.- Se obtiene el borde de la imagen con la técnica de píxel rotulado.
- 2.- Se obtienen todos los puntos extremos (punto superior izquierdo, punto superior derecho, punto izquierdo, punto derecho, punto inferior izquierdo y el punto inferior derecho).
- 3.- Con la teoría de contorno siguiente, transformada Hough y un cálculo

de pendientes, se obtiene si un tramo del borde es una recta, curva o un vertice

- 4.- Se compara el resultado del punto anterior con la base de datos y se identifica el tipo de figura.
- 5.- Se proporciona el tipo de figura de que se trata.
- 6.- Se muestra en pantalla la imagen.

Algoritmo del programa que localiza la figura que se pida buscar de entre 6 figuras (programa 4).

- 1.- Se divide por regiones toda la imagen, encerrando cada figura en una región en base a la técnica de región y agrupamiento.
- 2.- Se obtiene el borde de la imagen para cada región con el método de píxel rotulado.
3. Se obtienen todos los puntos extremos para cada figura.
- 4.-Se pide que figura se quiere localizar.
- 5.- El programa consta de 6 funciones (detecta círculos, detecta rectángulos, detecta triángulos, detecta cuadrados, detecta elipses y detecta hexágonos).
- 6.- Con la teoría de contorno siguiente, transformada Hough y un cálculo de pendientes se identifica si existe el tipo de figura en cada función.
- 7.- Se muestran en pantalla unicamente las figuras existentes que se pidieron buscar.

5.4 RESULTADOS DE LA APLICACION DE LOS METODOS DE IDENTIFICACION DE FIGURAS

En este punto se mostrará el resultado de aplicar los cuatro programas de cómputo. Se mostrará el resultado de una sola imagen para cada programa, esto con la finalidad de ilustrar el resultado de cada programa (excepto para el primero y tercer programa, en los cuales se mostraran imágenes de prueba). Cada programa se probó con varias imágenes diferentes, algunas de ellas se pueden observar en el punto 5.5 (Robustez de los métodos de identificación de figuras)

Las siguientes páginas muestran las imágenes resultantes al aplicar cada uno de los tres programas de cómputo. La página 72 muestra algunas de las imágenes que se utilizaron para probar el primer programa, el cual identifica el tipo de figura de que se trata.

En las páginas 73 y 74 se pueden ver las imágenes resultantes al aplicar el segundo programa, el cual identifica y localiza la figura que se pida buscar. Se puede observar que el programa identifica correctamente una elipse, un círculo, un cuadrado, un rectángulo, un triángulo y un hexágono. En el punto 5.5 se observan algunas imágenes que se utilizaron para probar la robustez de este programa.

En las páginas 75 y 76 se muestran algunas de las imágenes que se utilizaron para probar el tercer programa, el cual identifica el tipo de figura de que se trata.

En las páginas 77 y 78 se puede ver el resultado de aplicar el cuarto programa de cómputo a una imagen. Este programa localiza la figura que se pida buscar, de entre 6 figuras.

5.5 PRUEBAS DE ROBUSTEZ DE LOS METODOS DE IDENTIFICACION DE FIGURAS

La robustez del primer programa se probó con imágenes originales e imágenes de bordes, este programa se analizó para diferentes imágenes de un cuadrado, de un rectángulo, de un círculo, de un triángulo, de una elipse o de un hexágono, y se observó que el programa identifica correctamente cualquier tipo de figura.

La robustez del segundo programa (el cual identifica o localiza la figura que se solicite buscar) se probó con diferentes imágenes como son:

- Que existan las seis imágenes geométricas (cuadrado, rectángulo, triángulo, elipse, círculo y hexágono).
- Que falte una o hasta las seis figuras .
- Que existan figuras diferentes a las geométricas antes mencionadas.
- Que existan varias figuras iguales (2 cuadrados, 2 rectángulos, etc.).
- Que las figuras no se encuentren alineadas

Este programa identifica correctamente todas las diferentes imágenes.

La robustez del tercer programa (el cual identifica la figura que se le proporcione, esta figura puede tener una rotación) se probó con diferentes imágenes de un cuadrado, de un rectángulo, de un triángulo, de una elipse y de

IDENTIFICA UNA SOLA FIGURA GEOMETRICA

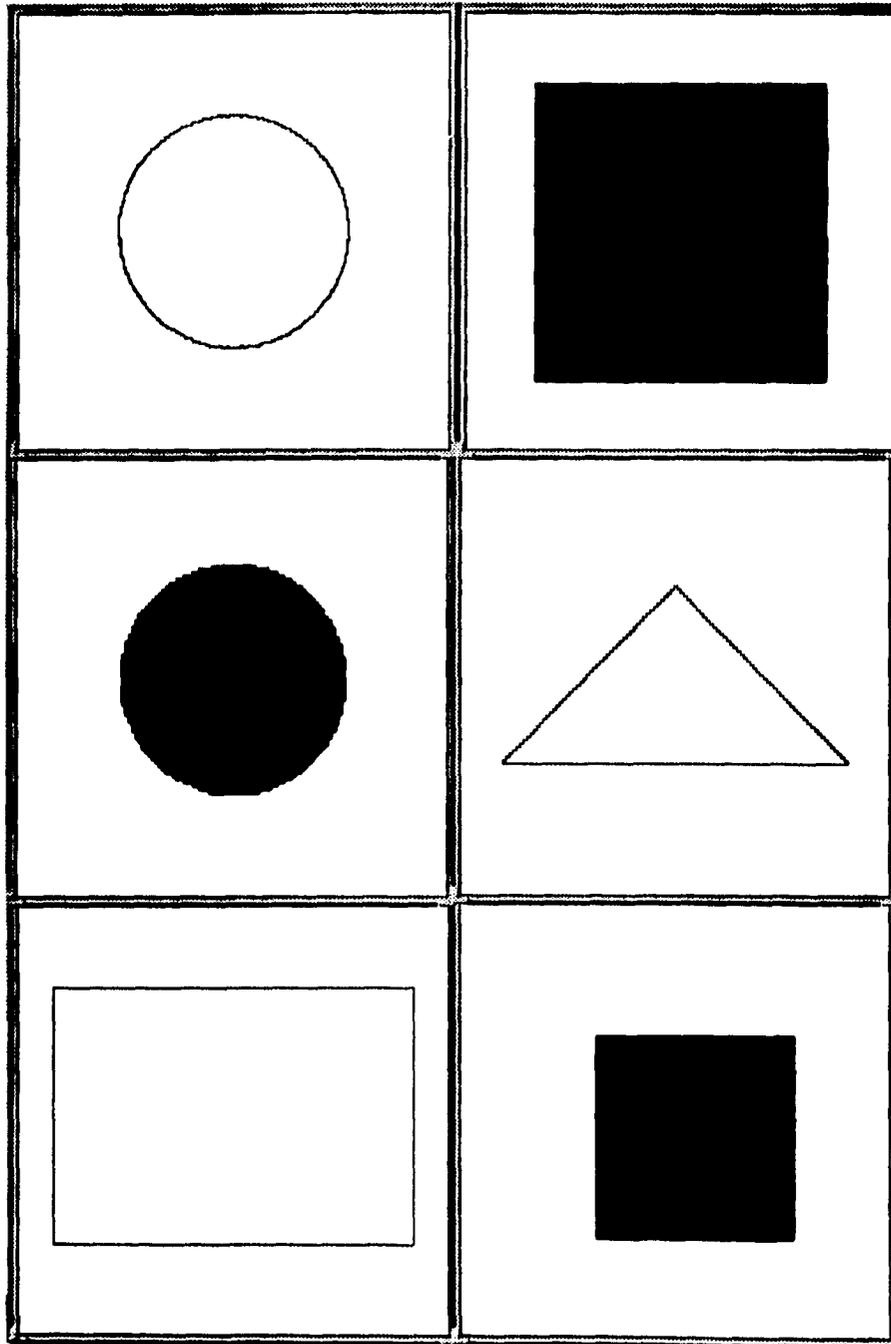


Fig. 5.16 Imágenes de prueba: Programa 1 72

IDENTIFICA Y LOCALIZA LA FIGURA A BUSCAR

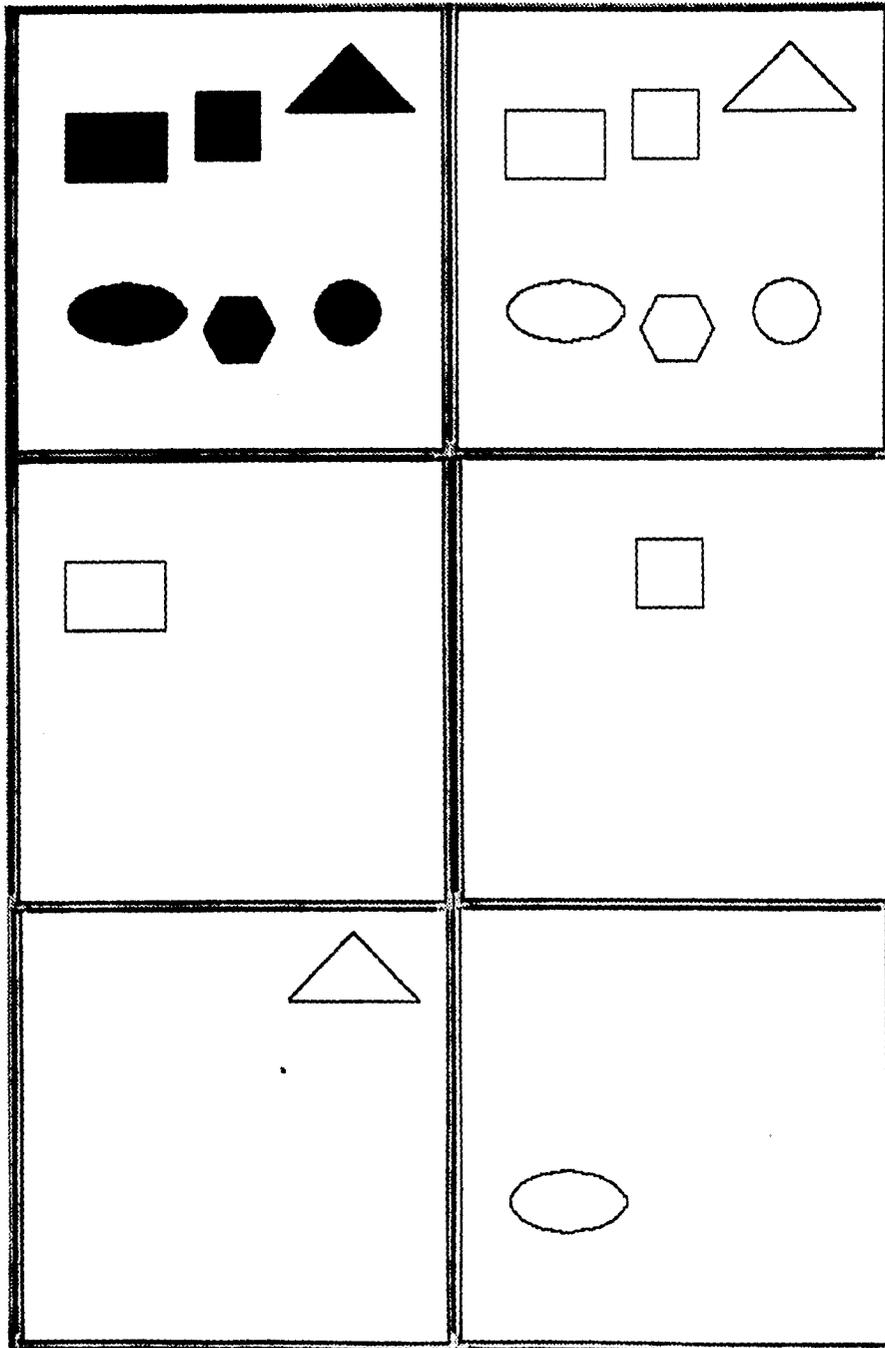
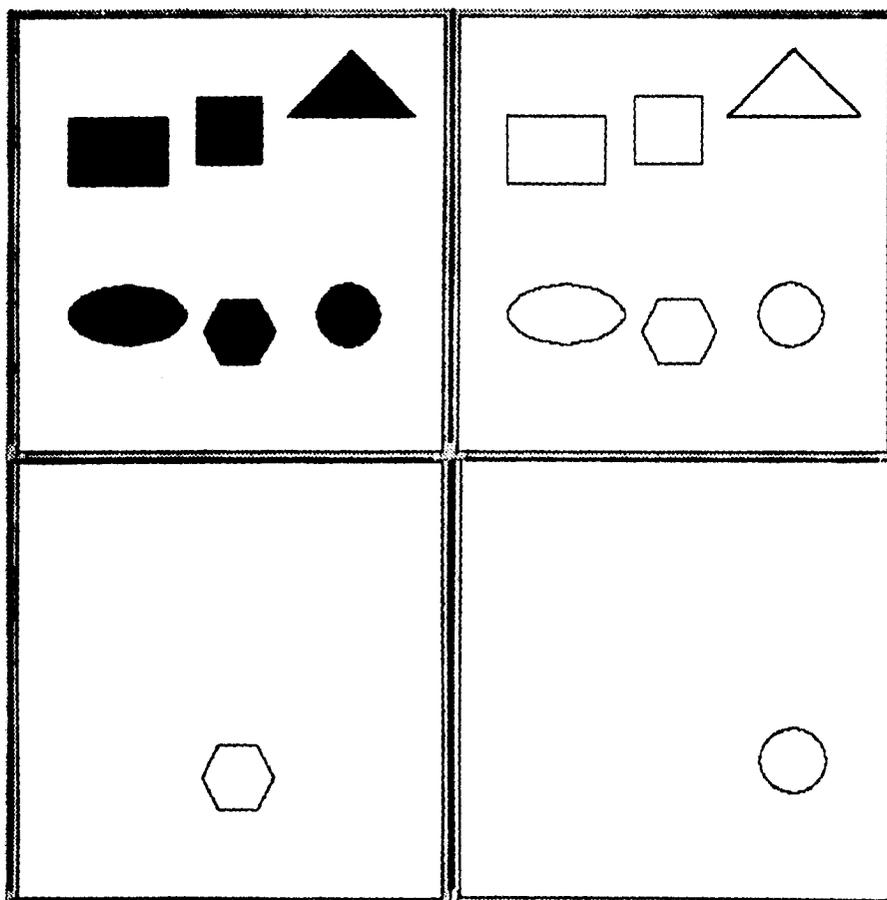


Fig. 5.17 Resultados: Programa 2

IDENTIFICA Y LOCALIZA LA FIGURA A BUSCAR



IDENTIFICA EL TIPO DE FIGURA

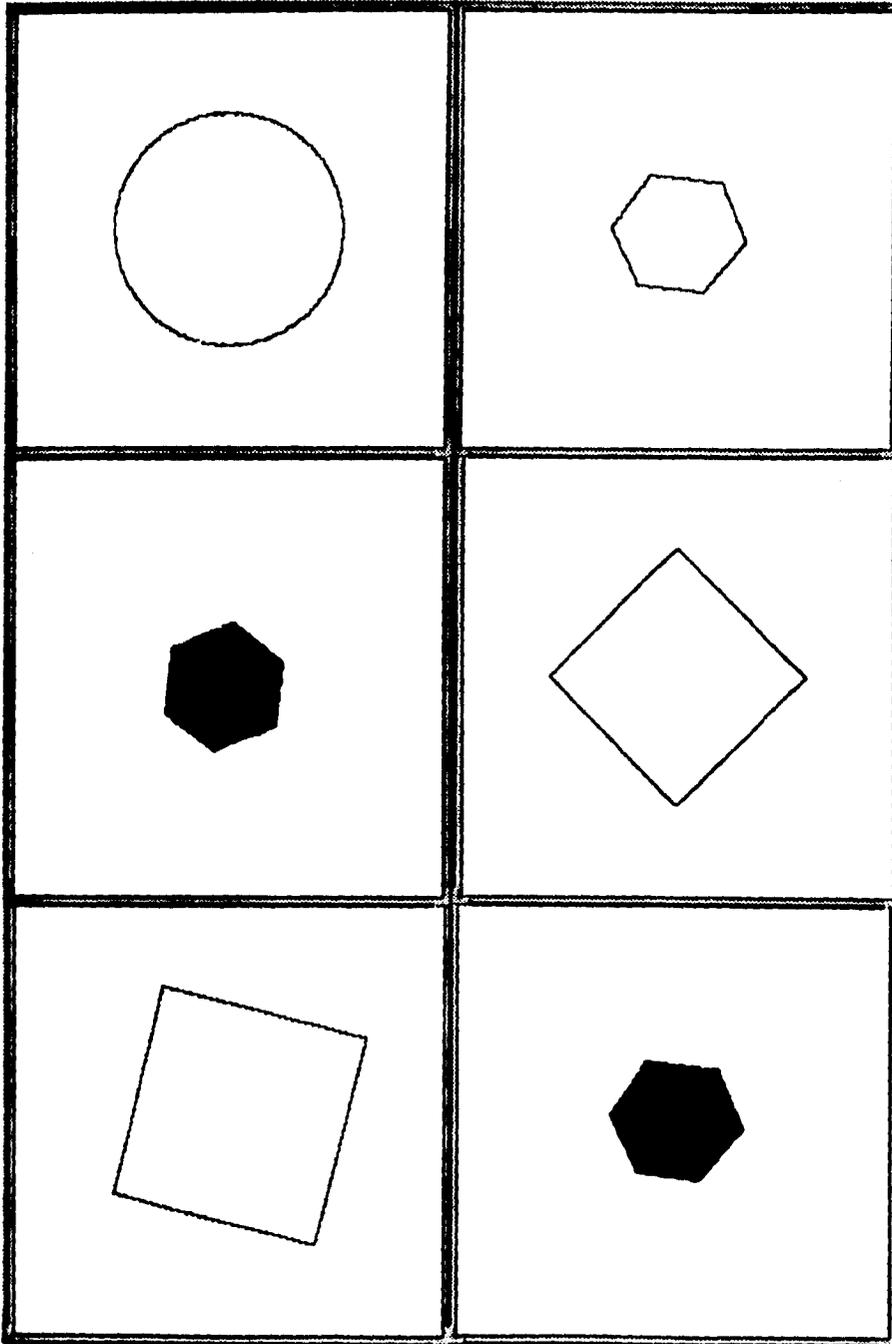


Fig. 5.19 Imágenes de prueba: Programa 3 75

IDENTIFICA EL TIPO DE FIGURA

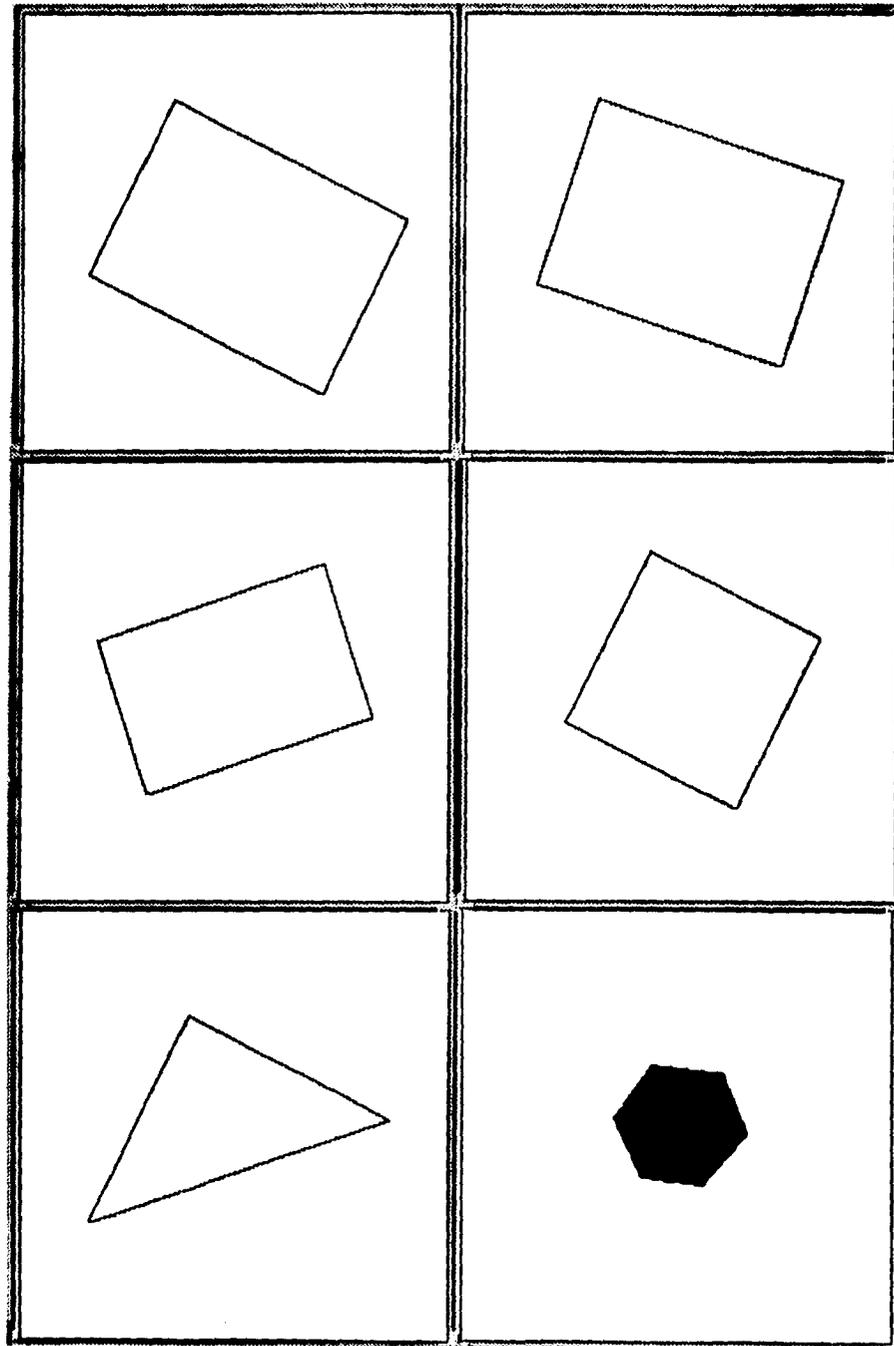


Fig. 5.20 Imágenes de prueba: Programa 3 76

IDENTIFICA LA FIGURA QUE SE PIDA BUSCAR

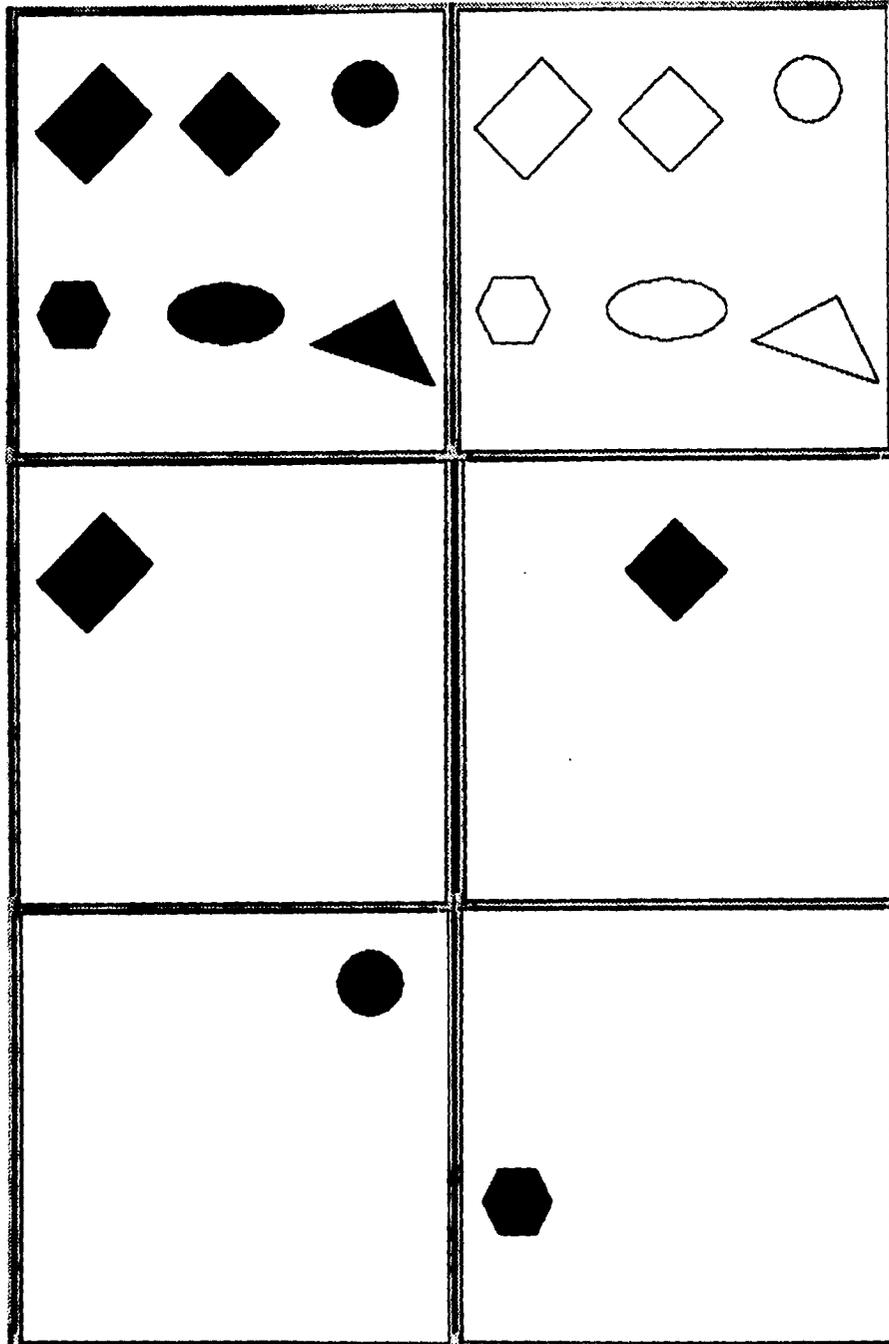
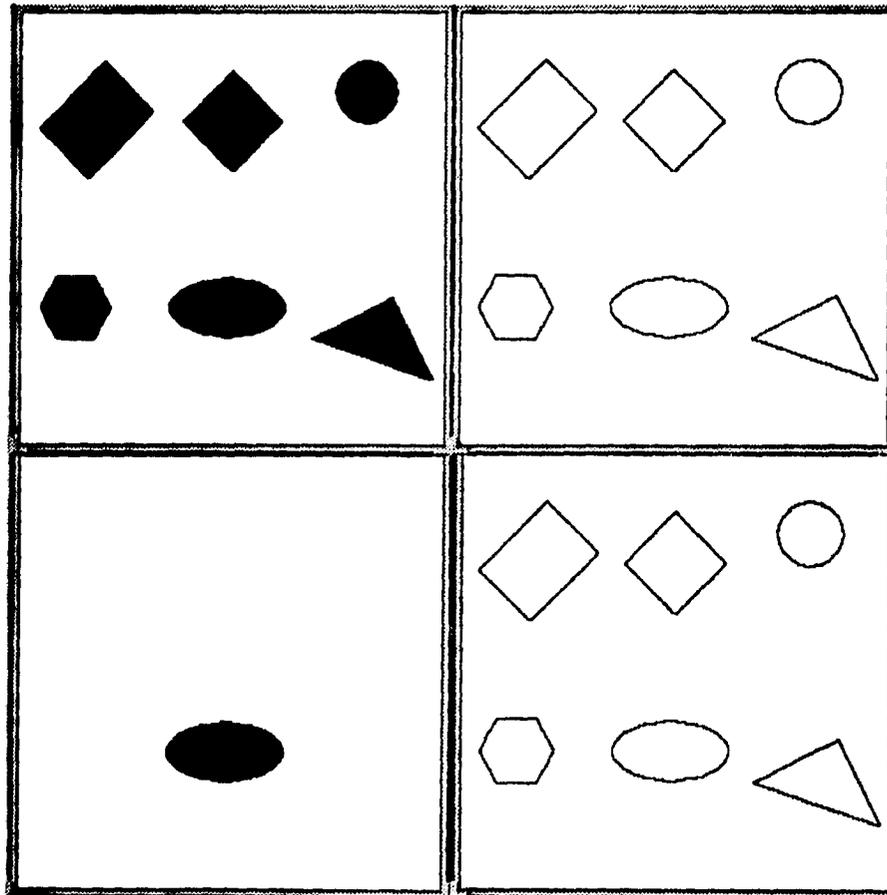


Fig. 6.21 Resultados: Programa 4

IDENTIFICA LA FIGURA QUE SE PIDA BUSCAR



**ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA**

un hexágono, se observó que el programa identifica correctamente cualquier tipo de figura.

La robustez del cuarto programa de cómputo (el cual localiza la figura que se pide buscar) se probó con imágenes diferentes como son:

- Que existan las seis imágenes diferentes.
- Que falte una o hasta 5 figuras.
- Que existan diferentes figuras a las geométricas.
- Que existan varios tipos de figuras iguales (dos cuadrados, tres círculos, etc)
- Que las figuras no se encuentren alineadas.
- Que todas las figuras tengan rotación.

En las páginas 80 y 81 se observan algunas imágenes que se utilizaron para probar la robustez del programa que identifica el tipo de figura de que se trata.

En las páginas 82 y 83 se observan algunas imágenes que se utilizaron para probar la robustez del segundo programa el cual identifica o localiza el tipo de figura que se pide buscar.

En las páginas 84 y 85 se observan algunas imágenes que se utilizaron para probar la robustez del tercer programa, el cual identifica la figura que se le proporcione.

En las páginas 86 y 87 se observan algunas imágenes que se utilizaron para probar la robustez del cuarto programa de cómputo, el cual localiza la figura que se pide buscar.

5.6 APLICACION A OBJETOS INDUSTRIALES

Los puntos anteriores de este capítulo muestran la aplicación de los métodos de detección de bordes y las técnicas de segmentación de imágenes, para la identificación de figuras geométricas, dicha identificación dió muy buenos resultados. Por lo que es importante conocer como se comportarán estos métodos con figuras más complejas, que son las que se encontrarían en un ambiente industrial. Entonces en este punto se implantarán los métodos a dos programas para realizar la identificación de piezas mecánicas, como son: un desarmador, una pinza, un candado, y un martillo.

Para realizar esta identificación se elaboraron dos programas; el primero

PRUEBA DE ROBUSTEZ

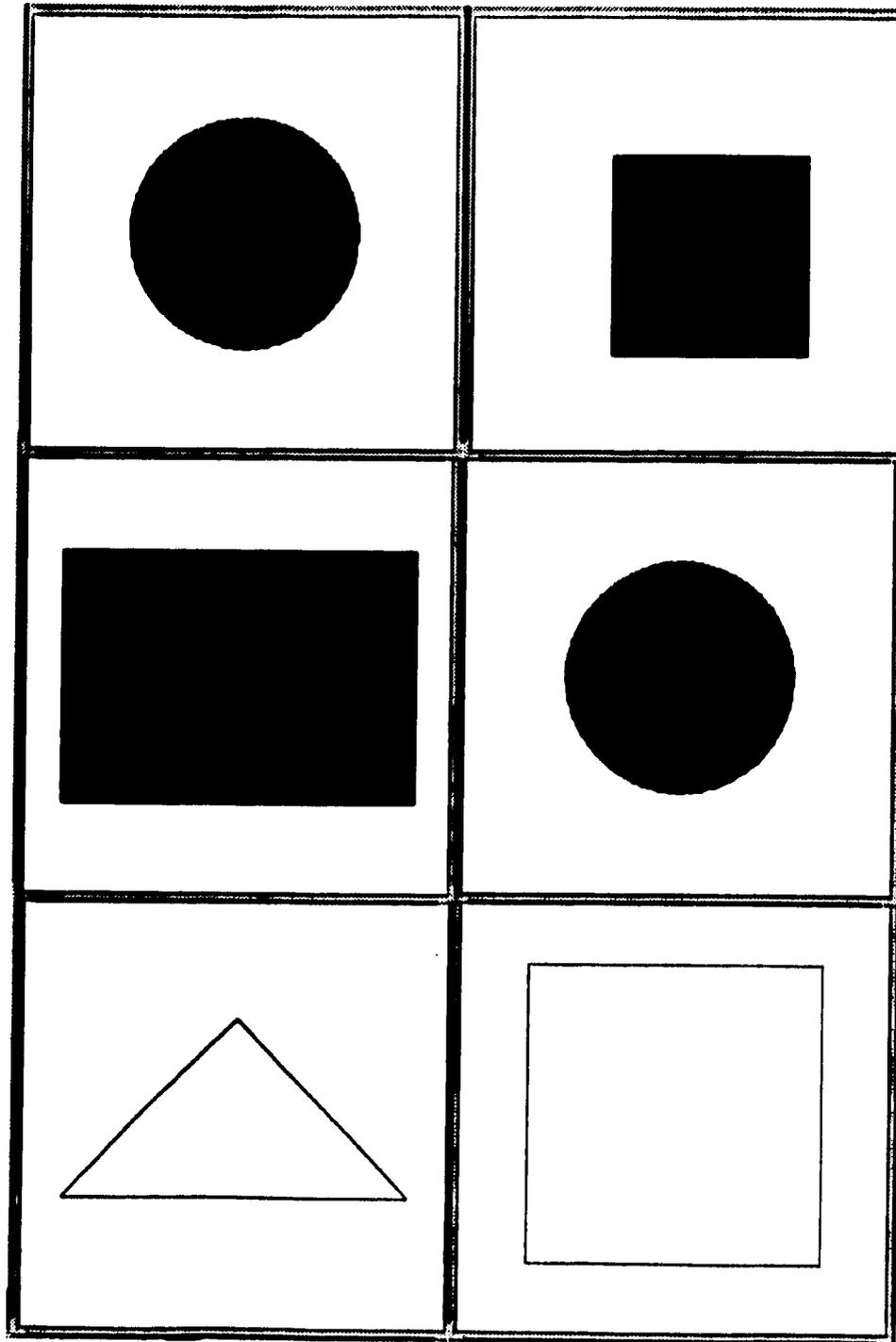


Fig. 5.23 Imágenes de prueba: Programa 1 80

PRUEBA DE ROBUSTEZ

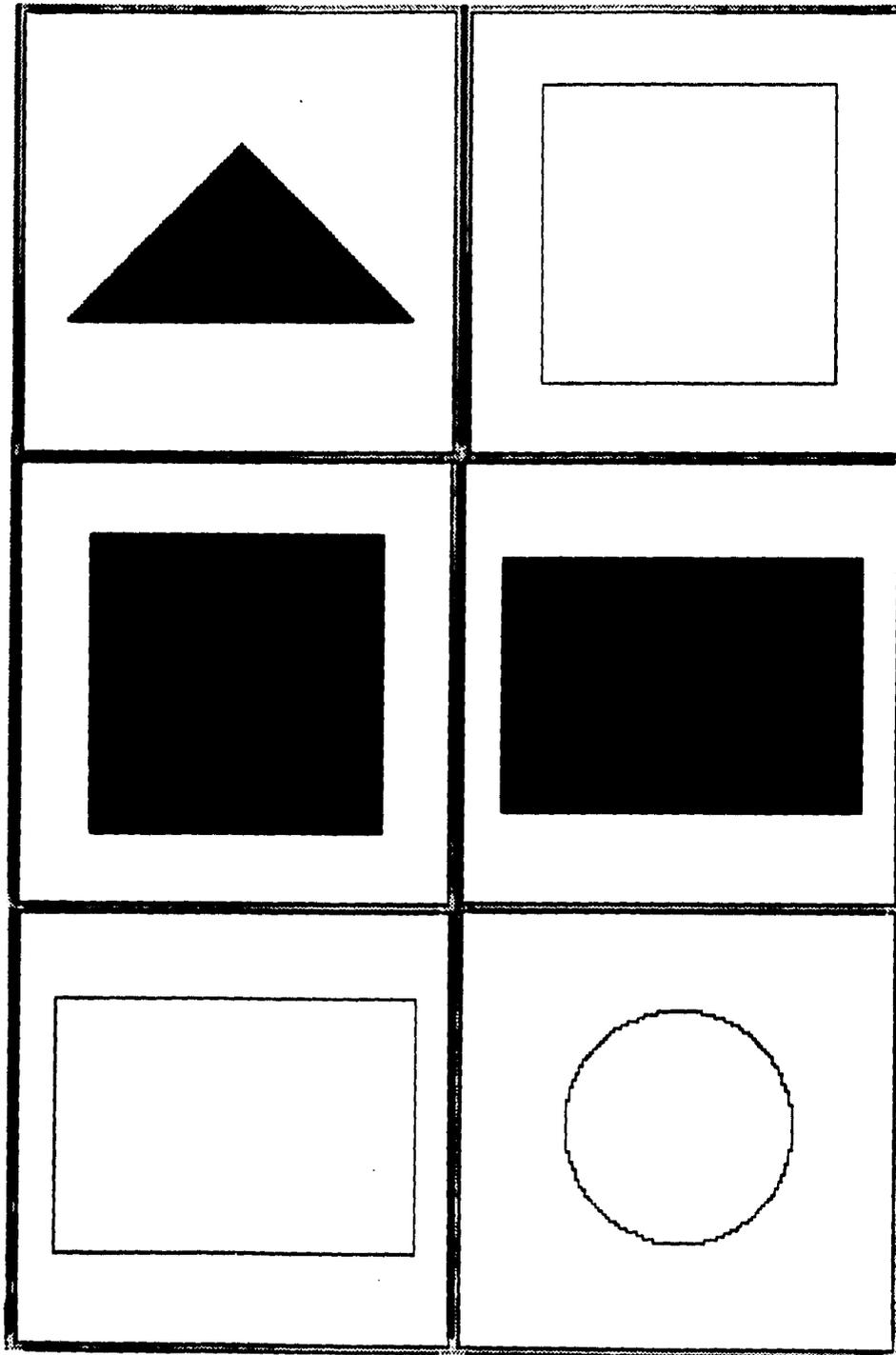


Fig. 5.24 Imágenes de prueba: Programa 1 81

PRUEBA DE ROBUSTEZ

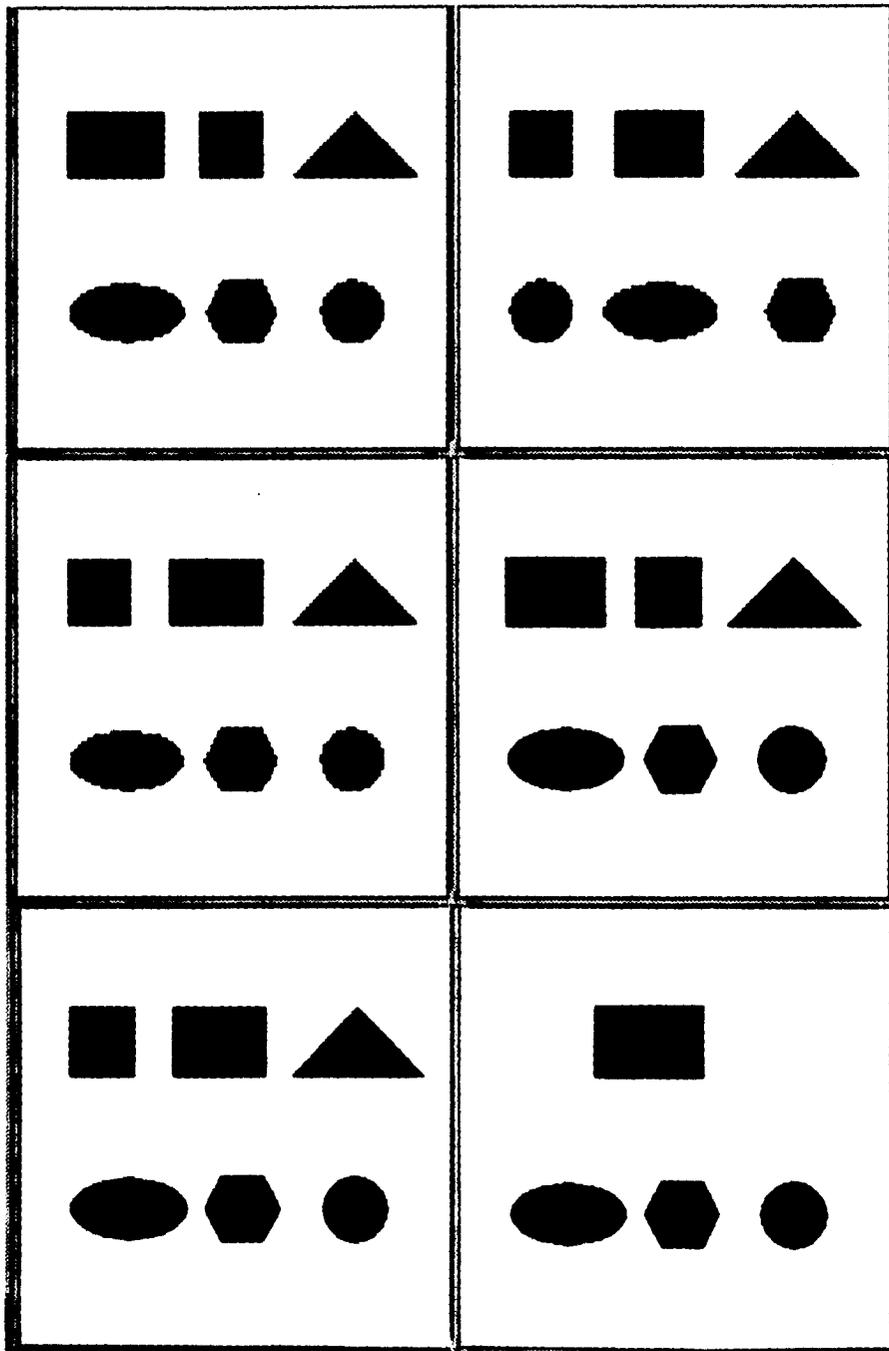


Fig. 6.25 Imágenes de prueba: Programa 2 82

PRUEBA DE ROBUSTEZ

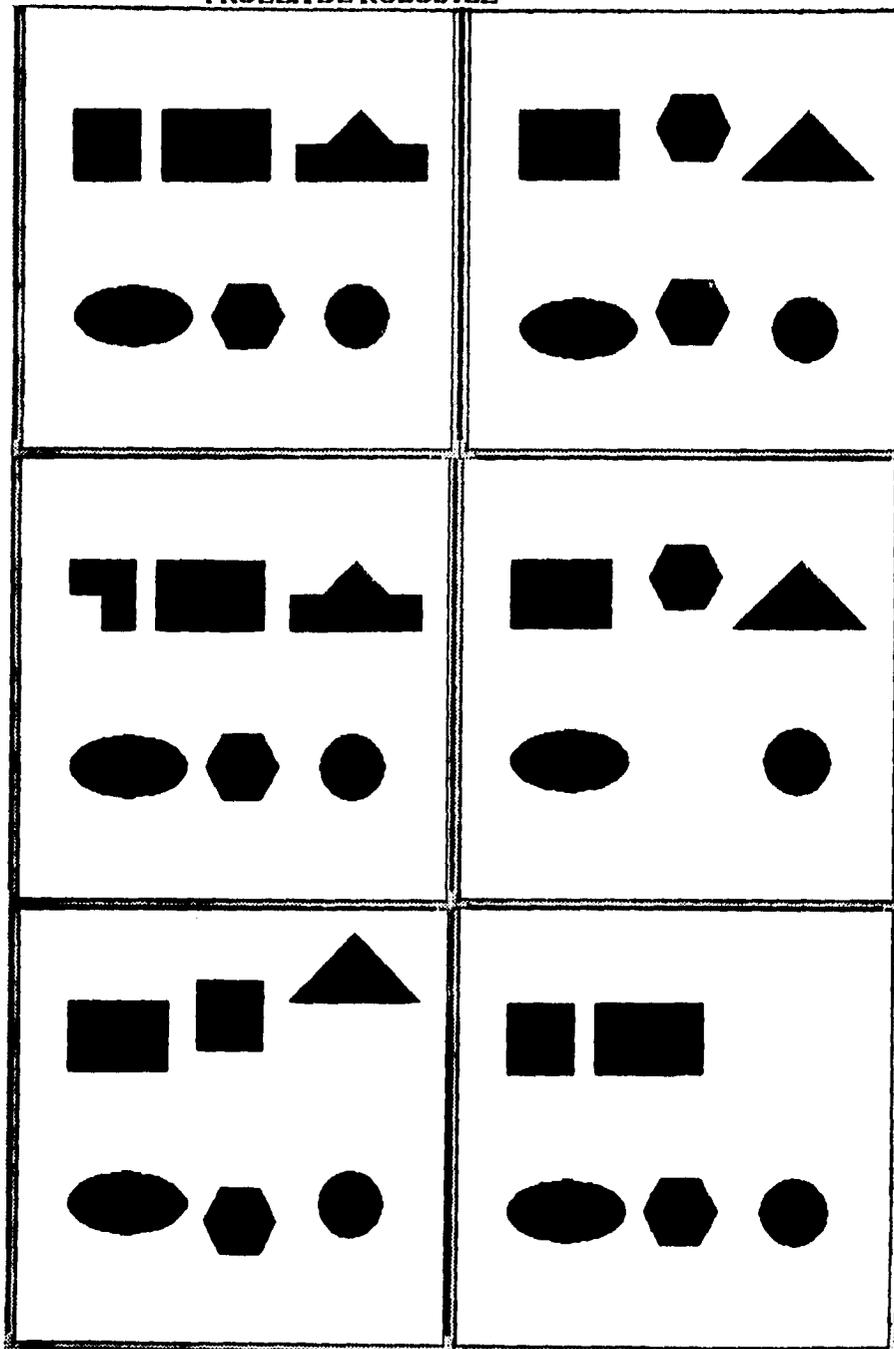


Fig. 5.26 Imágenes de prueba: Programa 2 83

PRUEBA DE ROBUSTEZ

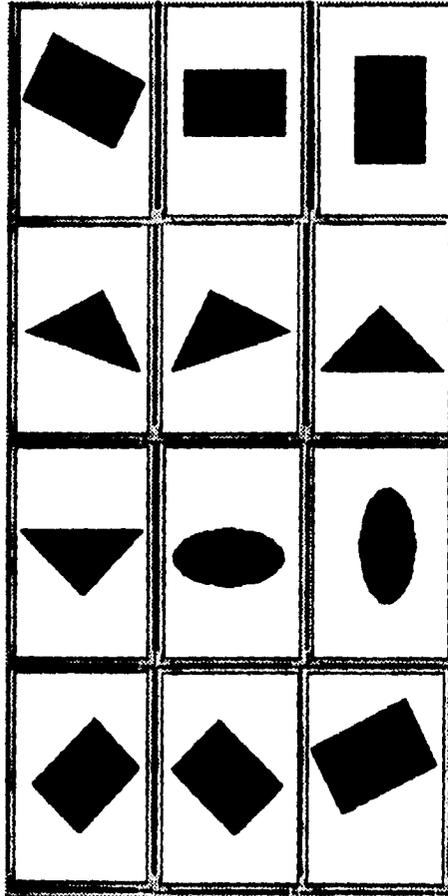


Fig. 5.27 Imágenes de prueba: Programa 3 84

PRUEBA DE ROBUSTEZ

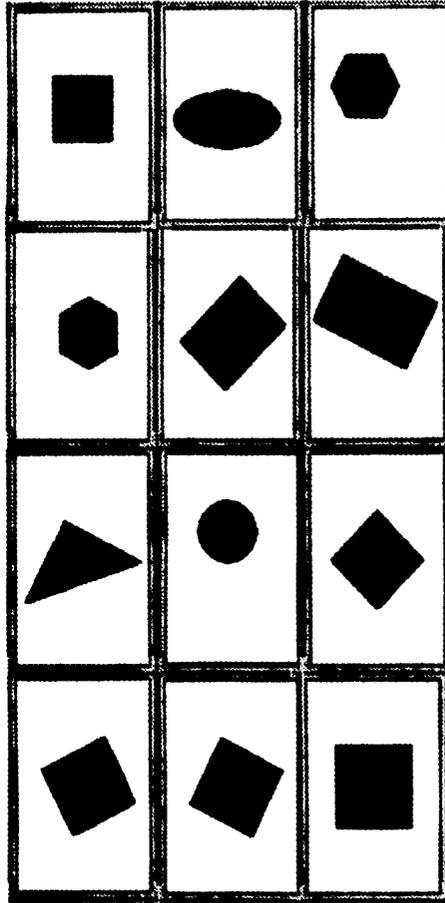


Fig. 6.28 Imágenes de prueba: Programa 3 65

PRUEBA DE ROBUSTEZ

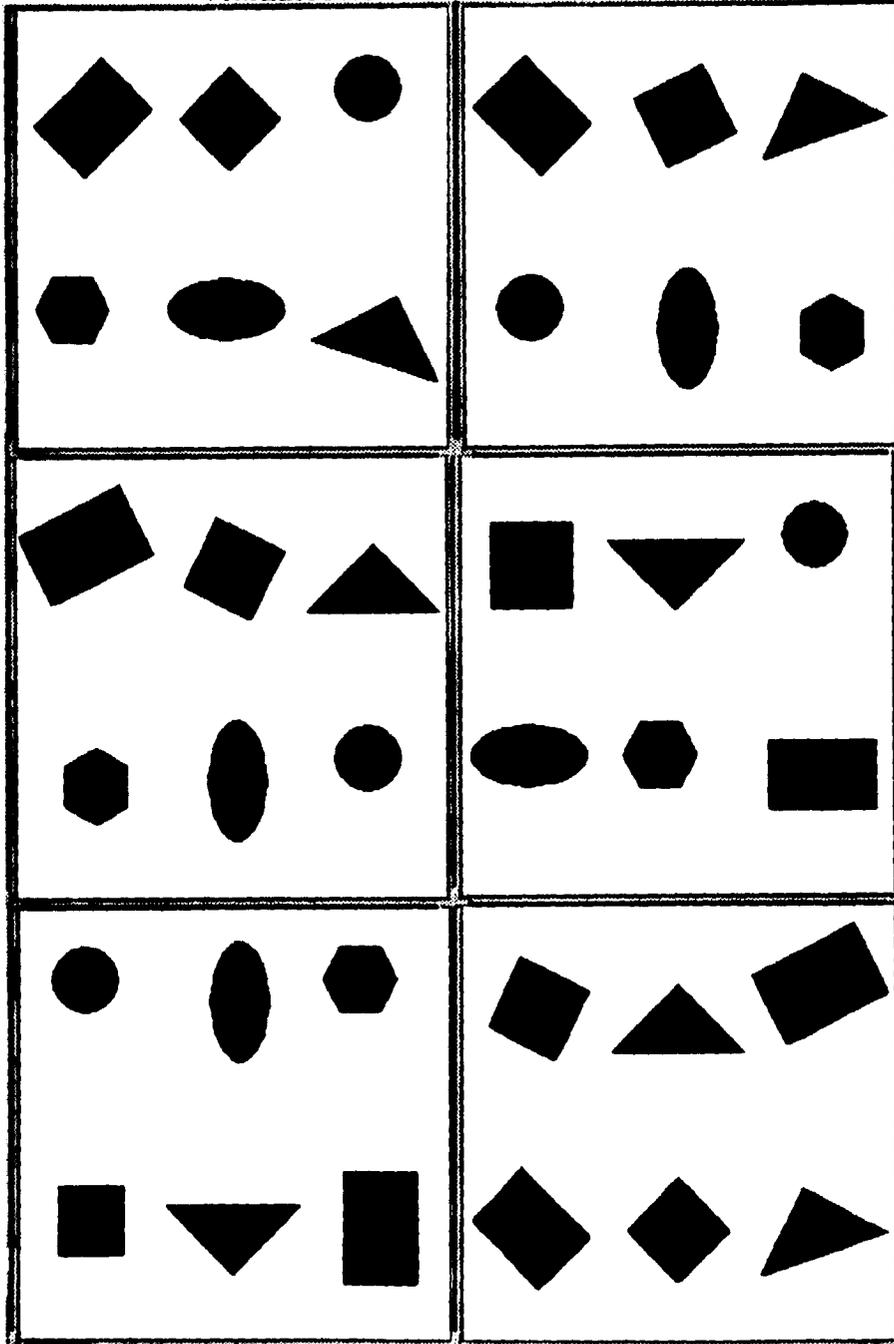


Fig. 5.29 Imágenes de prueba: Programa 4 66

PRUEBA DE ROBUSTEZ

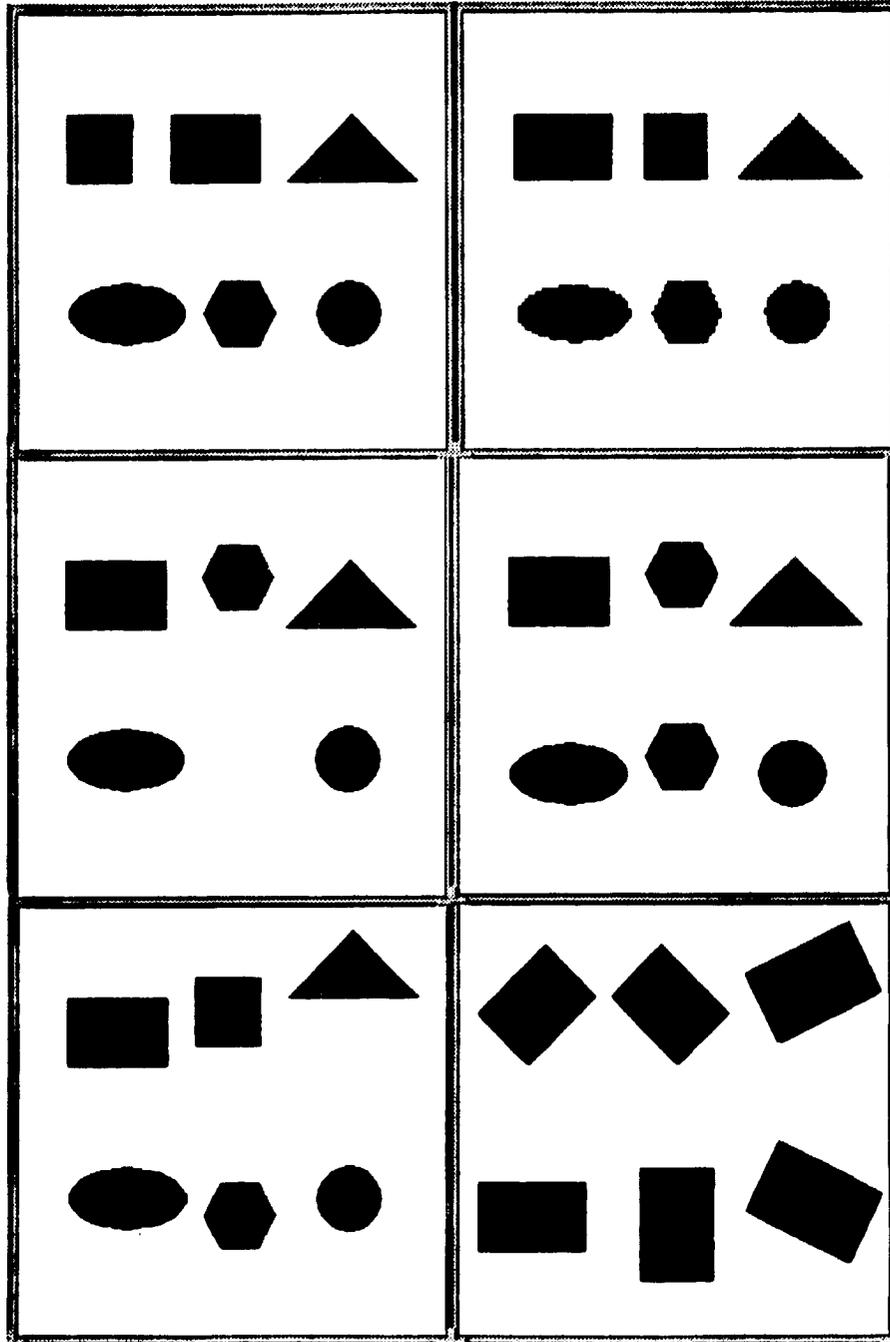


Fig. 5.30 Imágenes de prueba: Programa 4 87

Identifica el tipo de pieza que se le proporciona y el segundo localiza la pieza mecánica que se le pide buscar.

Los dos programas funcionan con la imagen original binaria o con la imagen de bordes. Las técnicas de segmentación utilizadas para los dos programas son: pixel rotulado, contorno sigulente, transformada Hough y cálculo de pendientes.

El algoritmo de programación consiste en encontrar el borde de las piezas mecánicas en base a la técnica de pixel rotulado, posteriormente se encuentran todos los puntos extremos, y con la teoría de contorno sigulente, transformada Hough y cálculo de pendientes, se examina el borde de cada pieza por tramos, y en base a las curvas y rectas de estos tramos se compara con la base de datos y se realiza la identificación de cada pieza mecánica.

Las páginas 89 y 90 muestran los resultados al aplicar el segundo programa a una imagen

Las páginas 91 y 92 muestran algunas imágenes que se utilizaron para probar la robustez del primer programa de cómputo.

Las páginas 93 y 94 muestran algunas imágenes que se utilizaron para probar la robustez del segundo programa de cómputo.

APLICACION A PIEZAS MECANICAS

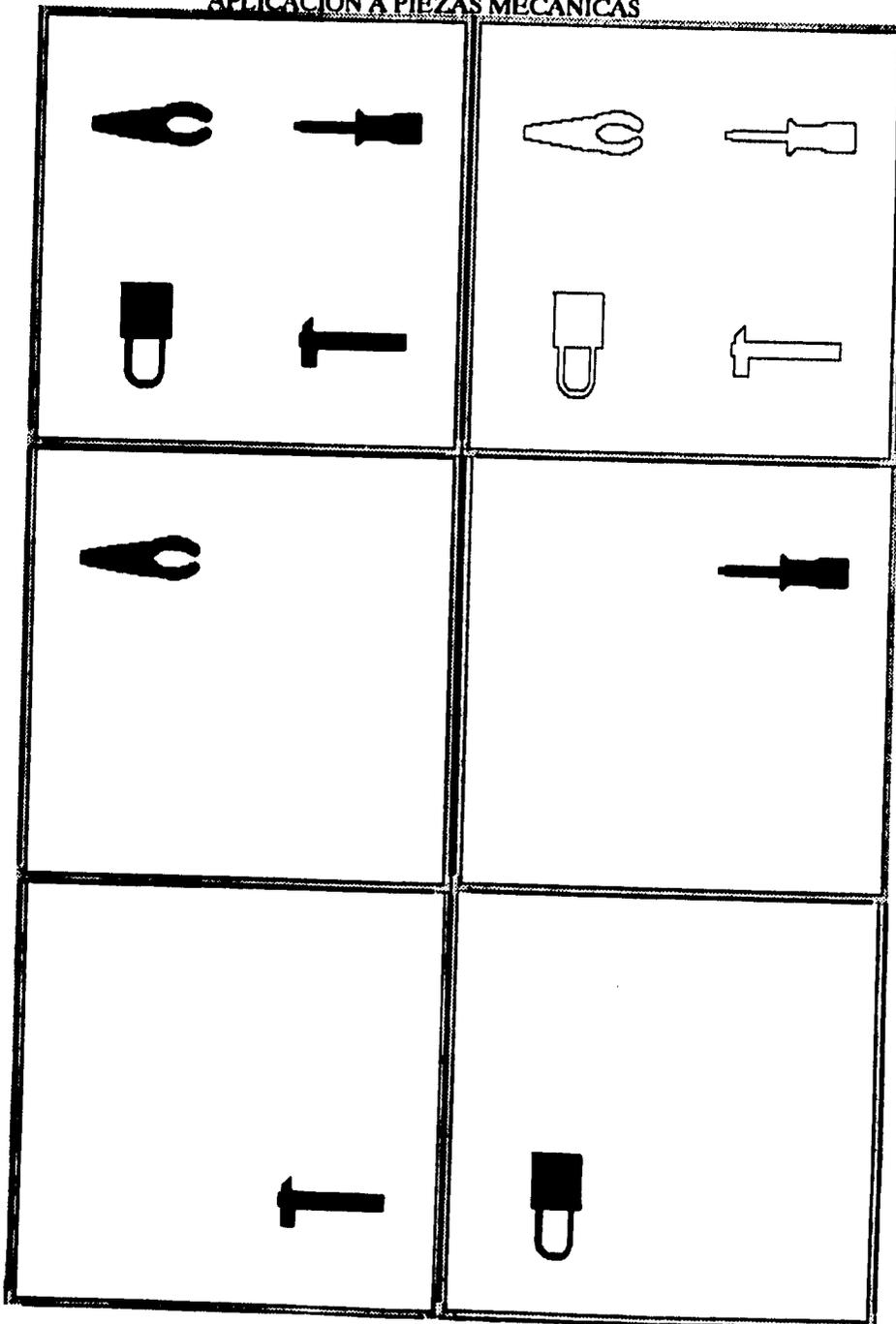


Fig. 6.31 Resultados: Programa 1

APLICACION A PIEZAS MECANICAS

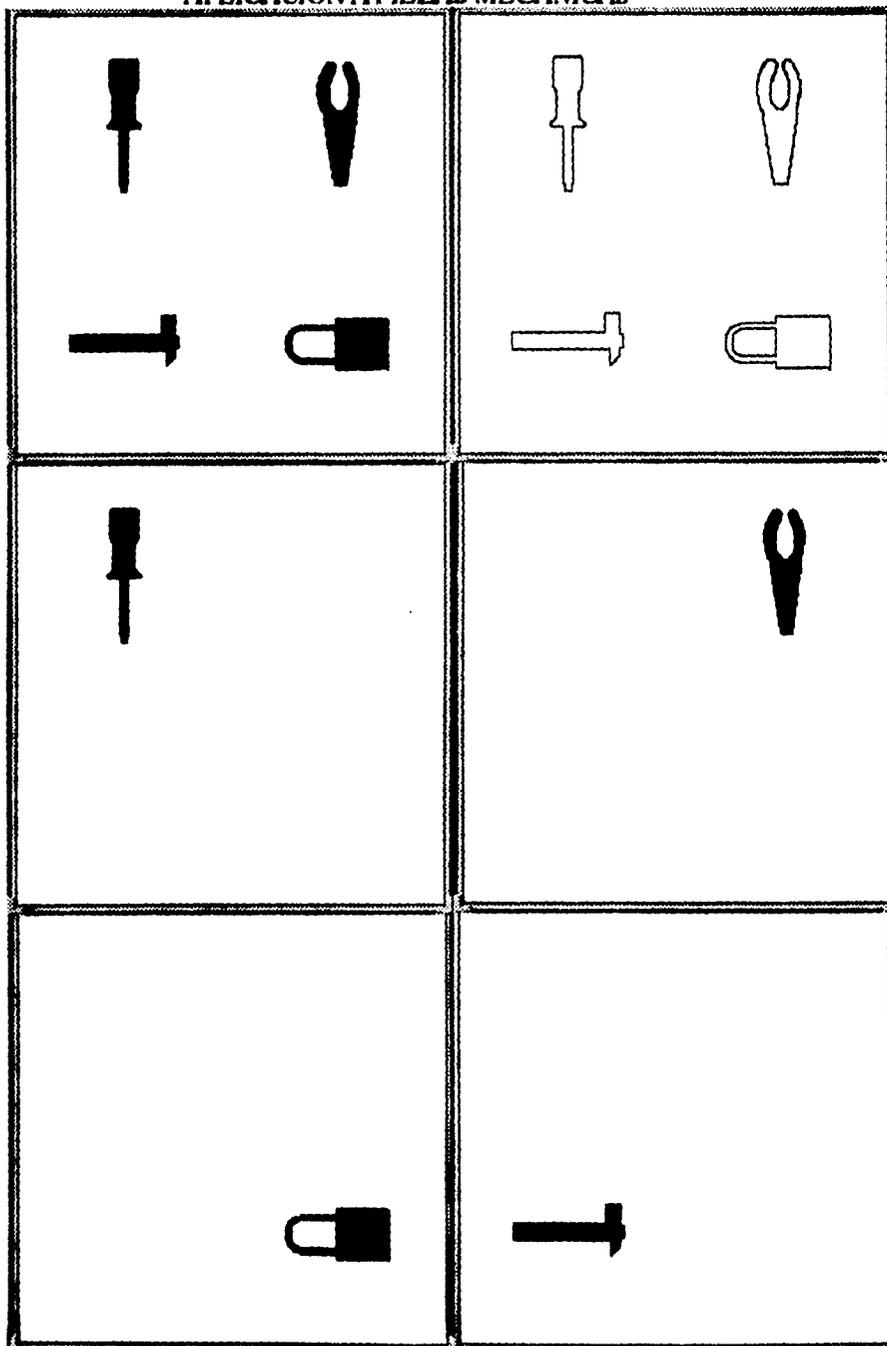


Fig. 6.32 Resultados: Programa 1

APLICACION A PIEZAS MECANICAS

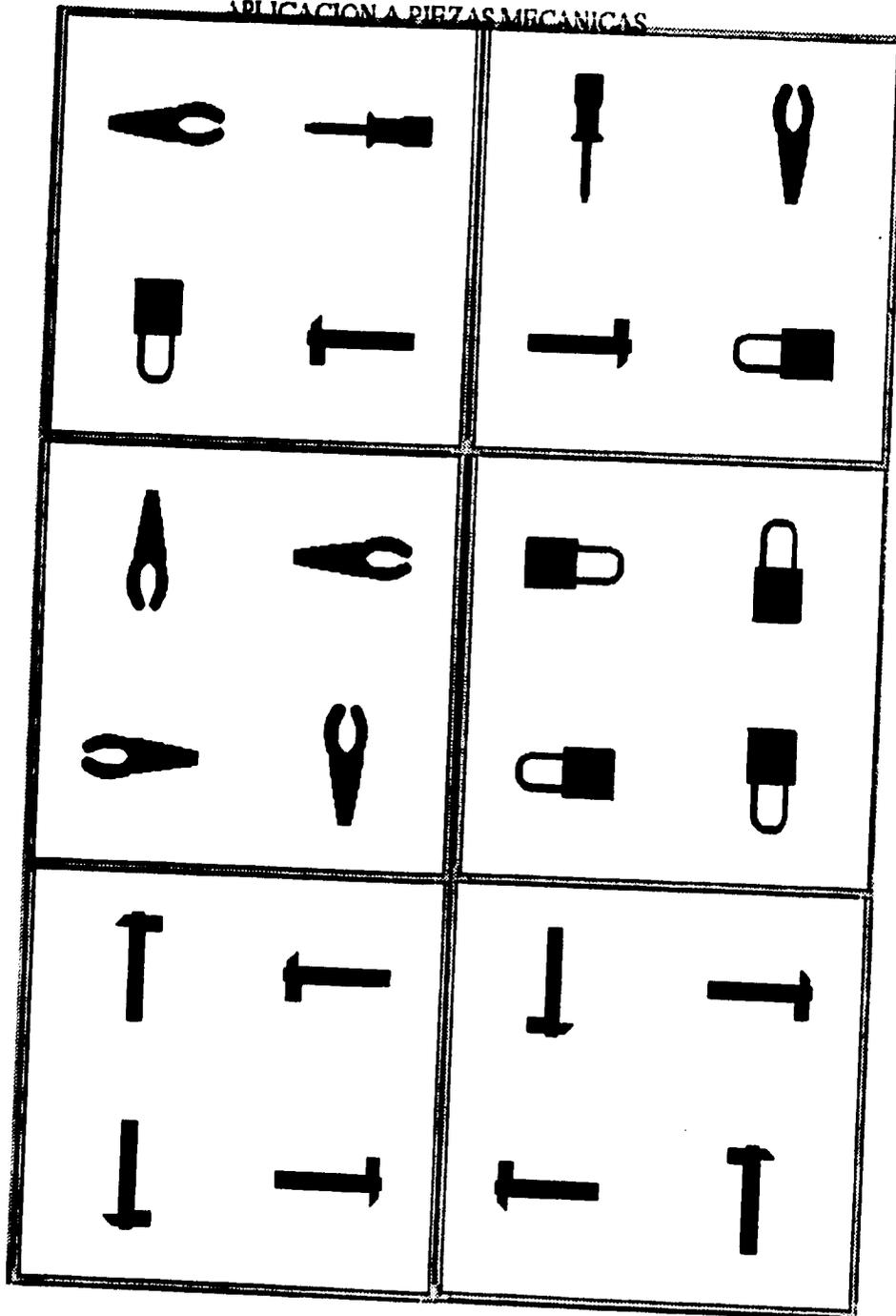


Fig. 5.33 Prueba de Robustez: Programa 1 91

APLICACION A PIEZAS MECANICAS

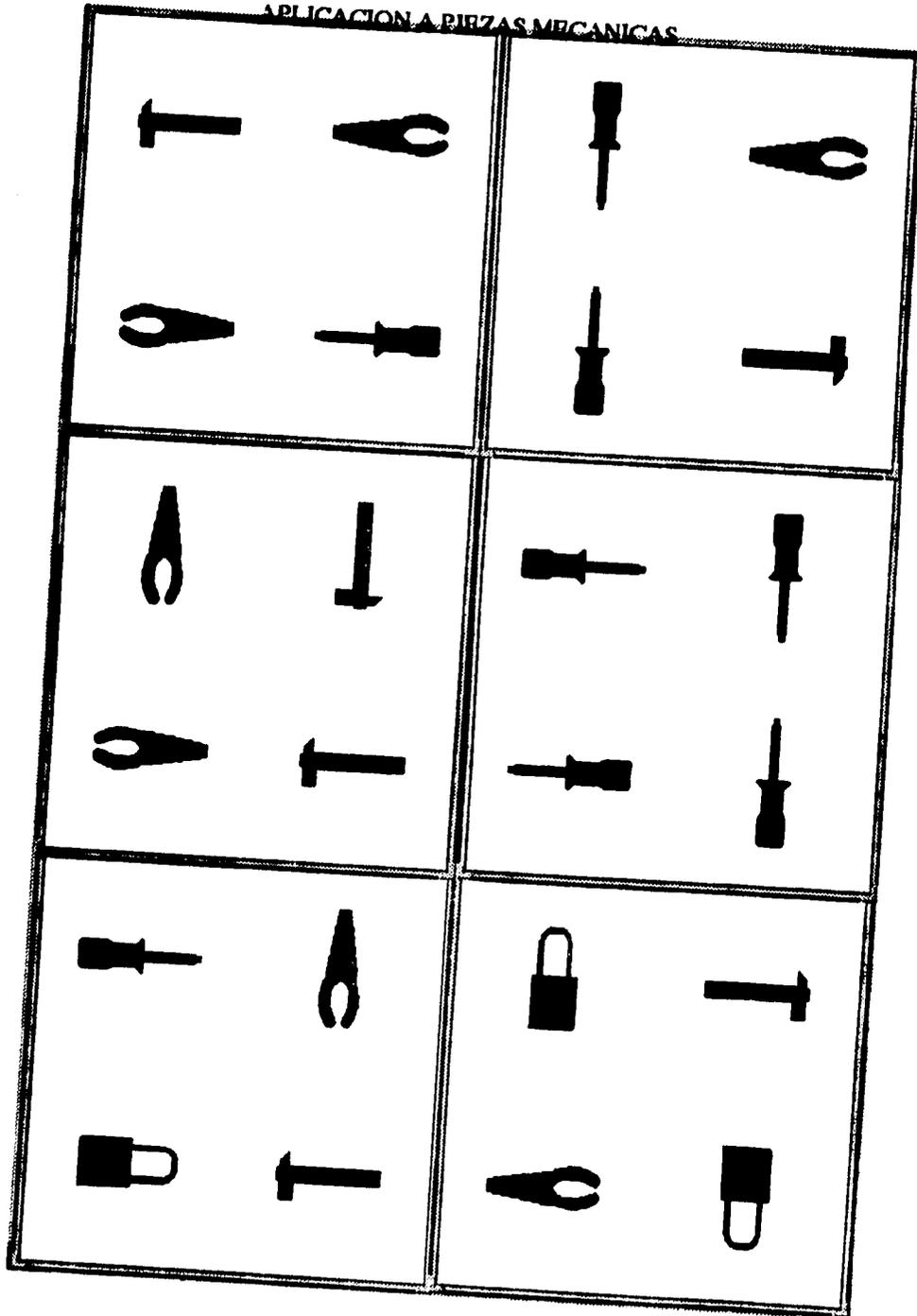


Fig. 6.34 Prueba de Robustez: Programa 1

APLICACION A PIEZAS MECANICAS

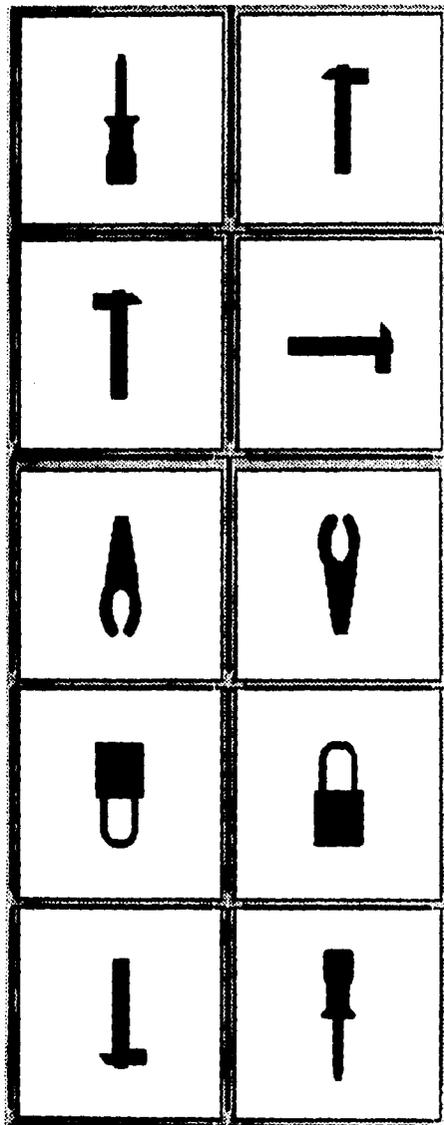


Fig. 5.36 Prueba de Robustez: Programa 2 93

APLICACION A PIEZAS MECANICAS

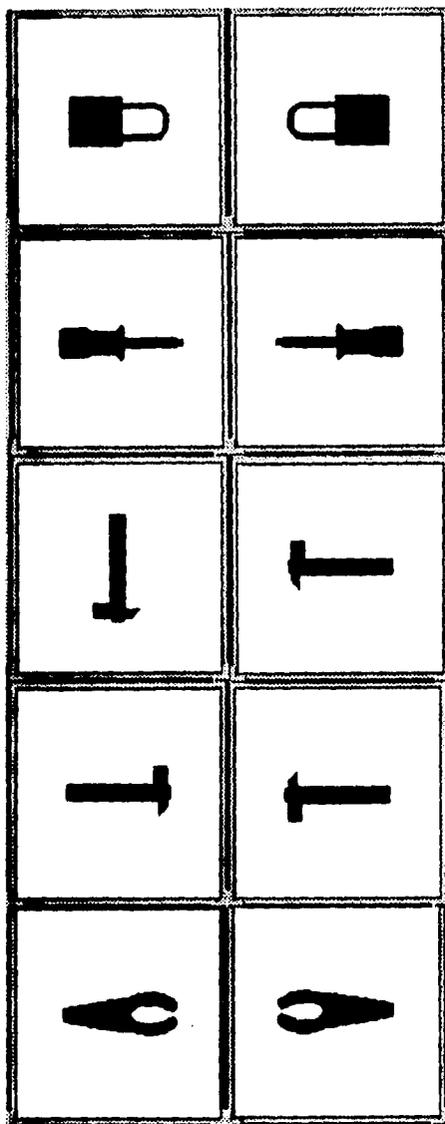


Fig. 5.36 Prueba de Robustez: Programa 2 94

6 ANALISIS DE LA IMPLANTACION DE UN METODO DE IDENTIFICACION DE FORMAS SOBRE UNA ESTRUCTURA DE DATOS MULTIRESOLUCION

6.1 INTRODUCCION

En este capítulo se utilizará una imagen obtenida sobre una estructura de datos multiresolución, esto con la finalidad de observar como responden los programas (del capítulo anterior) con este tipo de imágenes. Para obtener la imagen en multiresolución se utilizarán los programas realizados por anteriores tesis, es decir, en esta tesis no se realizará el algoritmo de descomposición para obtener una imagen en multiresolución.

Las representaciones multiresolución son muy efectivas para analizar el contenido de información de imágenes. La diferencia de información de una señal entre la resolución 2^{j+1} y 2^j puede ser extraída descomponiendo la señal en una base "wavelet" ortonormal de $L^2(\mathbb{R}^n)$. Un wavelet ortonormal es una familia de funciones, las cuales son construidas por dilatación y traslación de una función única. Esta descomposición define una representación multiresolución ortogonal llamada representación "wavelet".

Es difícil analizar el contenido de información directamente de la intensidad de los niveles de gris de una imagen, en general el valor depende de las condiciones de luz. Una descomposición nos habilita a tener una interpretación invariante de la escala de la imagen. La escala de una imagen varía con la distancia entre la escena y el centro óptico de la cámara. Cuando la escala de la imagen es modificada, la interpretación de la escena no debe cambiar. Una representación multiresolución puede ser de escala-invariante si la secuencia de resolución de parámetros $(\eta_j)_{j \in \mathbb{Z}}$ varía exponencialmente. Si se supone que existe una resolución a R tal que para todos los enteros j $\eta_j = a^j$.

6.2 ALGORITMOS DE DESCOMPOSICION Y RECONSTRUCCION EN DOS DIMENSIONES

En base a los trabajos de Mallat [RH1], la transformada "wavelet" bidimensional que se describirá puede ser vista como una transformada "wavelet" unidimensional a lo largo de los ejes X y Y. Se puede mostrar que una transformada wavelet bidimensional puede ser calculada como una extensión separable del algoritmo de descomposición unidimensional. A cada paso se descompone $A_2^{d,j+1} f$ en $A_2^{d,j} f$, $D_1^{d,j} f$, $D_2^{d,j} f$ y $D_3^{d,j} f$. Este algoritmo se muestra en el diagrama de bloques de la figura 6.1 a. Primero se convolucionan las líneas de $A_2^{d,j+1} f$ con un filtro unidimensional, tomando una línea de dos. En seguida se convolucionan las

columnas de las señales resultantes con otro filtro unidimensional, tomando una columna de dos.

La estructura de aplicación de los filtros para calcular $A_{2^j}^{d_j} f$, $D_{2^j}^{1_j} f$, $D_{2^j}^{2_j} f$ y $D_{2^j}^{3_j} f$ esta dada en la figura 6.1a. La transformada wavelet se calcula de una imagen $A_{2^1}^{d_1} f$ repitiendo este proceso para $-1 \geq j \geq -J$.

El algoritmo de reconstrucción unidimensional puede ser extendido a dos dimensiones. En cada paso, la imagen $A_{2^{j+1}}^{d_{j+1}} f$ es reconstruida a partir de $A_{2^j}^{d_j} f$, $D_{2^j}^{1_j} f$, $D_{2^j}^{2_j} f$ y $D_{2^j}^{3_j} f$. Este algoritmo esta ilustrado en el diagrama de bloques de la figura 6.1b.

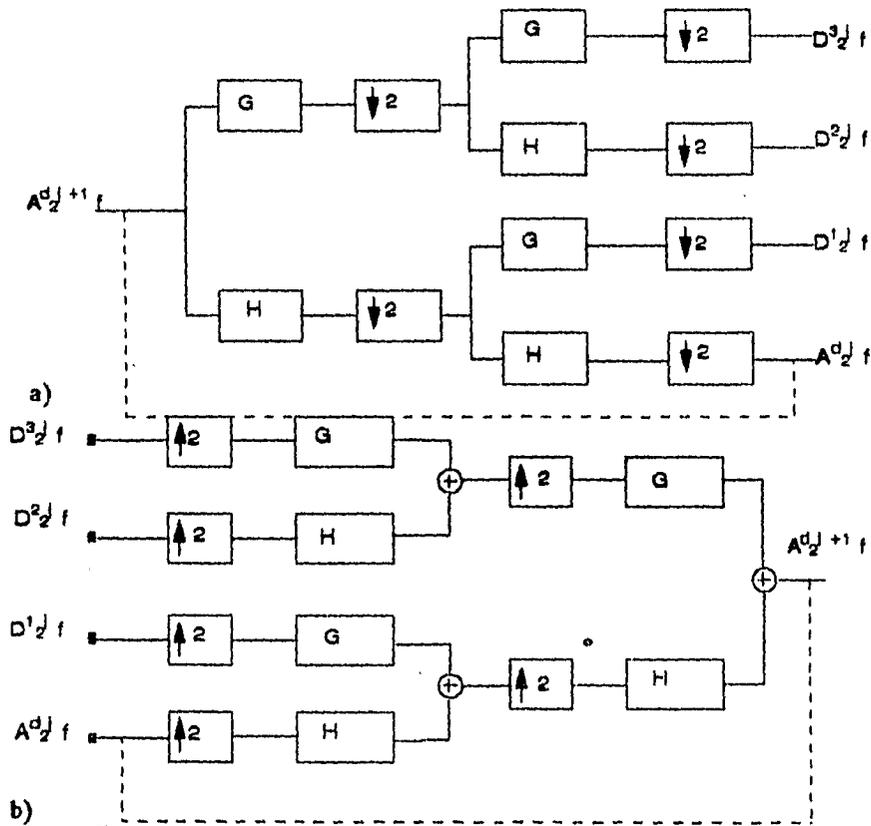


Figura 6.1 Diagrama de bloques (Algoritmos basados en filtros unidimensionales de líneas y columnas)

- a) Descomposición de una aproximación discreta
- b) reconstrucción de una imagen

Entre cada columna de las imágenes $A_{2^j}^d f$, $D_{2^j}^1 f$, $D_{2^j}^2 f$ y $D_{2^j}^3 f$, se añade una columna de ceros, convolucionando las líneas con un filtro unidimensional, se añade una línea de ceros entre cada línea de la imagen resultante y se convolucionan las columnas de la imagen resultante y se convolucionan las columnas con otro filtro unidimensional. La imagen $A_{2^{j+1}}^d f$ es reconstruida a partir de su transformada wavelet repitiendo el proceso para $-j \leq j \leq -1$.

La descomposición wavelet puede ser interpretada como una descomposición de la señal en un conjunto de canales de frecuencias orientadas espacialmente (sub-bandas).

La figura 6.2 muestra en el dominio de la frecuencia como la imagen $A_{2^{j+1}}^d f$ es descompuesta en $A_{2^j}^d f$, $D_{2^j}^1 f$, $D_{2^j}^2 f$ y $D_{2^j}^3 f$. La imagen $A_{2^j}^d f$ corresponde a las más bajas frecuencias, $D_{2^j}^1 f$ conserva las altas frecuencias verticales (bordes horizontales), $D_{2^j}^2 f$ conserva las altas frecuencias horizontales (bordes verticales), $D_{2^j}^3 f$ conserva las altas frecuencias en ambas direcciones (esquinas diagonales).

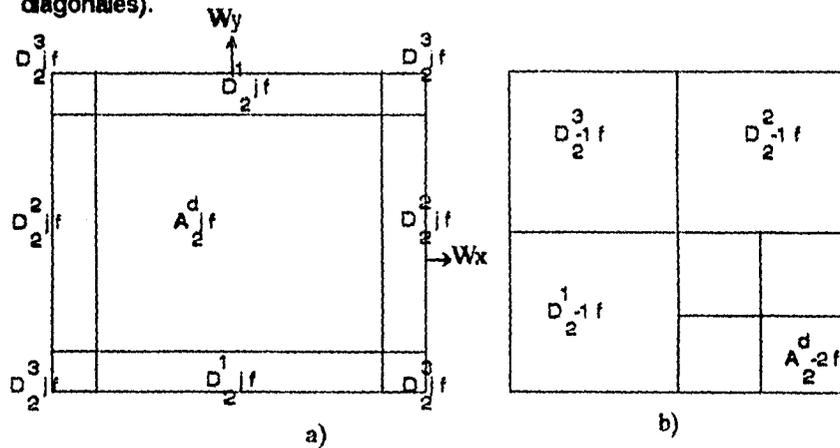


Figura 6.2
a) Descomposición del espectro $A_{2^{j+1}}^d f$ en $A_{2^j}^d f$ y los detalles $D_{2^j}^k f$
b) Disposición de las imágenes $D_{2^j}^k f$ y $A_{2^j}^d f$

6.3 INTRODUCCION A UNA ESTRUCTURA DE DATOS MULTI-RESOLUCION

A partir de [RB11], se tiene una clasificación de las diversas transformaciones piramidales, la cual se ilustra en la figura 6.3

Los criterios de clasificación parten esencialmente de:

- La naturaleza del filtrado

Pasa bajas

Pasa banda

- La propiedad de separación de los filtros utilizados

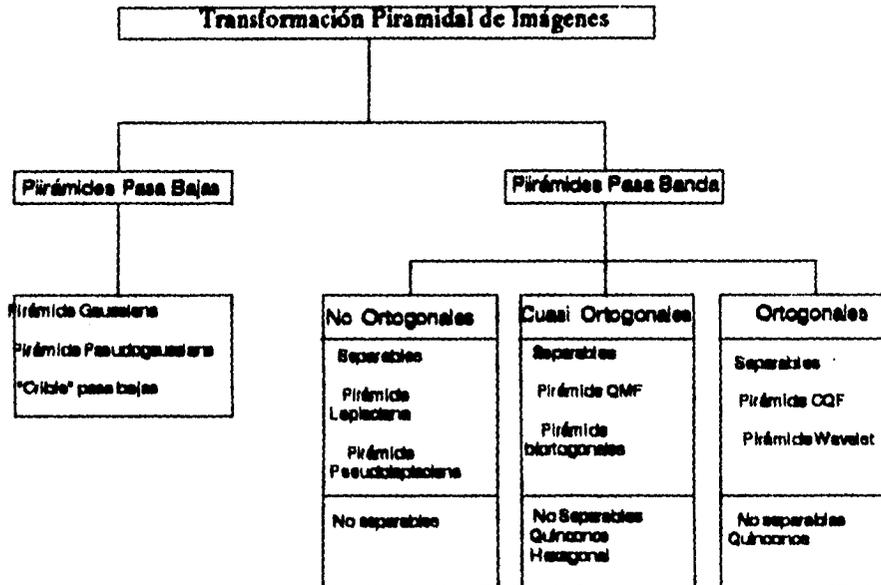


Figura 6.3 Diagrama de métodos piramidales de procesamiento de imágenes

6.3.1 METODO PIRAMIDAL PASA BAJAS

Para una serie de operaciones de filtrado pasa bajas y sus muestreos sucesivos se asocia a la imagen original G_0 una serie de versiones G_j a resoluciones y tamaños decrecientes que se ordenan para formar una estructura piramidal $(G_0, G_1, \dots, G_{N-1})$

Esta forma de representación que permite tomar la información de manera gradual y jerárquica se clasifica también como un método de aproximación multiresolución donde cada versión G_j de la imagen original representa la aproximación discreta a una resolución r_j . De este modo, una imagen piramidal con N niveles, corresponde a la aproximación sobre una secuencia de resoluciones $(r_j)_{j \in [0, N-1]}$ donde los valores condicionan las características de los filtros y las cadencias de muestreos utilizados.

A fin de tener una descomposición uniforme y evitar problemas de la no variación en la escala, la secuencia de resolución debe variar exponencialmente. Esto tiene consecuencias ventajosas sobre la eficiencia puesto que la construcción de una estructura piramidal se simplifica con la utilización de un algoritmo recursivo manipulando un solo filtro discreto y una sola cadencia de muestreos.

La generación recursiva de los niveles sucesivos de una imagen piramidal esta escrita por el esquema de la figura 6.4.

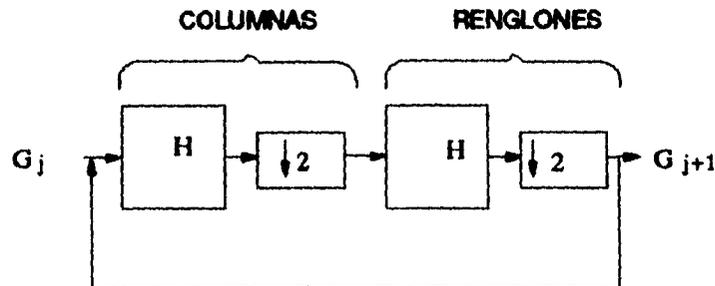


Figura 6.4 Diagrama de construcción de una pirámide pasa bajas

Esta pirámide sobresale por la simplicidad de su construcción.

DECIMADOR DE SEÑALES DISCRETAS

Un decimador de orden p es un sistema que a una señal de entrada discreta $u(x)$ le asocia una señal de salida discreta $v(x)$, tal que:

$$v(x) = u(px)$$

donde p y x son enteros.

Así se puede constatar que la operación de decimación introduce a la vez un submuestreo y una compresión en el tiempo que se traduce por una expansión dentro del dominio de la frecuencia. Esta operación se denota con la flecha hacia abajo (indicativa del sub-muestreo, la cual se ilustra en la figura 6.5).

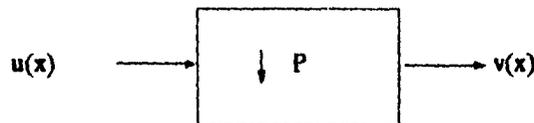


Figura 6.5 Diagrama de Bloques de un Decimador de Señales discretas

El decimador produce un efecto llamado "aliasing". En la práctica antes de

que una señal sea pasada a través de un decimador debe ser limitado en banda usando un filtro pasa banda reduciendo así los efectos del aliasing¹. Tales filtros son llamados filtros de decimación.

INTERPOLACION DE SEÑALES DISCRETAS

Un interpolador de orden p es un sistema que a una entrada discreta $u(x)$ asocia una señal de salida $v(x)$ tal que

$$v(x) = \begin{cases} u(x/p) & \text{x múltiplo de } p \\ 0 & \text{si no es múltiplo} \end{cases}$$

Una flecha hacia arriba indica la operación de interpolación. (Como se muestra en la figura 6.6)

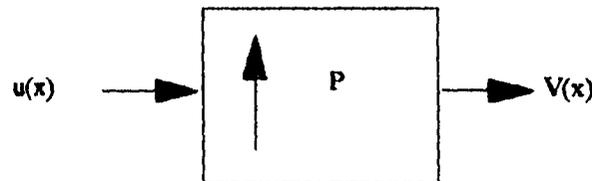


Figura 6.6 Diagrama de bloque de un interpolador de señales discretas

El efecto de la interpolación es un ensanchamiento en el dominio del tiempo y una compresión en el dominio de la frecuencia. El interpolador produce un efecto llamado reflejo.

En la práctica un interpolador es seguido de un filtro de interpolación el cual elimina este efecto.

El decremento en la escala se logra sobre-muestreando con un interpolador y el incremento sub-muestreando con un decimador. El cambio de escala en señales de tiempo continuo no altera la resolución ya que los cambios de escala pueden ser revertidos.

Sin embargo en señales de tiempo discreto el incremento en la escala en el análisis involucra sub-muestreo el cual automáticamente reduce la resolución.

Mientras los parámetros de escala y resolución sean aumentados la imagen digitalizada se parecerá más a la imagen original. Sin embargo, el incremento de

estos parámetros incrementará el tamaño del almacenamiento y consecuentemente el tiempo de procesamiento de la imagen.

6.3.2 METODO PIRAMIDAL PASA BANDA ORTOGONAL

Las técnicas de descomposición no son de aparición reciente. Dos dominios independientes se han desarrollado, uno en el espacio y otro en el tiempo, el primero se basa sobre un formalismo matricial y conceptos de procesamiento de señales, aborda el problema con el estudio de sistemas de análisis y síntesis de señales y la concepción de bancos de filtros de descomposición. El segundo puramente matemático y de más grande envergadura. Se preocupa de los problemas de análisis tiempo-frecuencia de la señal y de la construcción de un modelo matemático más coherente haciendo uso de las herramientas del análisis funcional de la teoría de aproximaciones de funciones [RB11].

ANÁLISIS Y SÍNTESIS

El sistema análisis y síntesis descrito por el diagrama de la figura 6.7 está constituido esencialmente de dos módulos.

MODULO DE ANÁLISIS.- Descompone la señal $i(n)$ en dos canales frecuenciales, con filtrado discreto con la ayuda de los filtros G_0 y H_0 (filtros de análisis pasa bajas) seguidos de un decimador de orden dos.

MODULO DE SÍNTESIS.- Reconstruye la señal a partir de las sub-bandas que salen de un interpolador, seguidos por los filtros G_1 y H_1 (filtros de síntesis pasa altas y pasa bajas respectivamente).

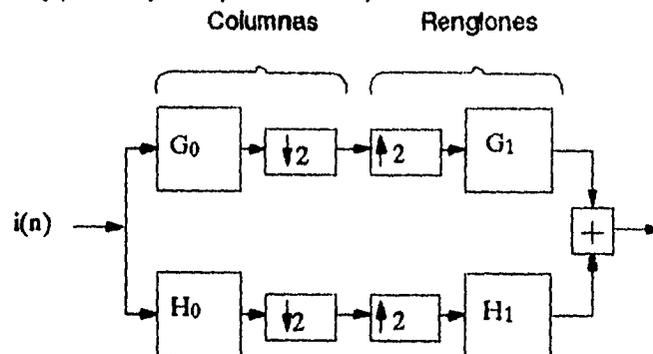


Figura 6.7 Sistema análisis- síntesis de dos canales

Un buen sistema de análisis - síntesis debe garantizar la completa conservación de la información.

Cada filtrado es acompañado de un decimador de orden dos. Las subimágenes obtenidas reducen el tamaño a un cuarto de la figura original como se muestra en la figura 6.8.

ANÁLISIS Y SÍNTESIS CON FILTROS SEPARABLES UNIDIMENSIONALES

Son posibles cuatro combinaciones de los filtros:

Sub-banda Alta-Alta.- El filtro pasa altas G se aplica sucesivamente sobre las líneas y columnas de la imagen. Las altas frecuencias horizontales y verticales son favorecidas dentro de la subimagen HH . La orientación de esta sub-banda favorece y pone en evidencia los contornos diagonales.

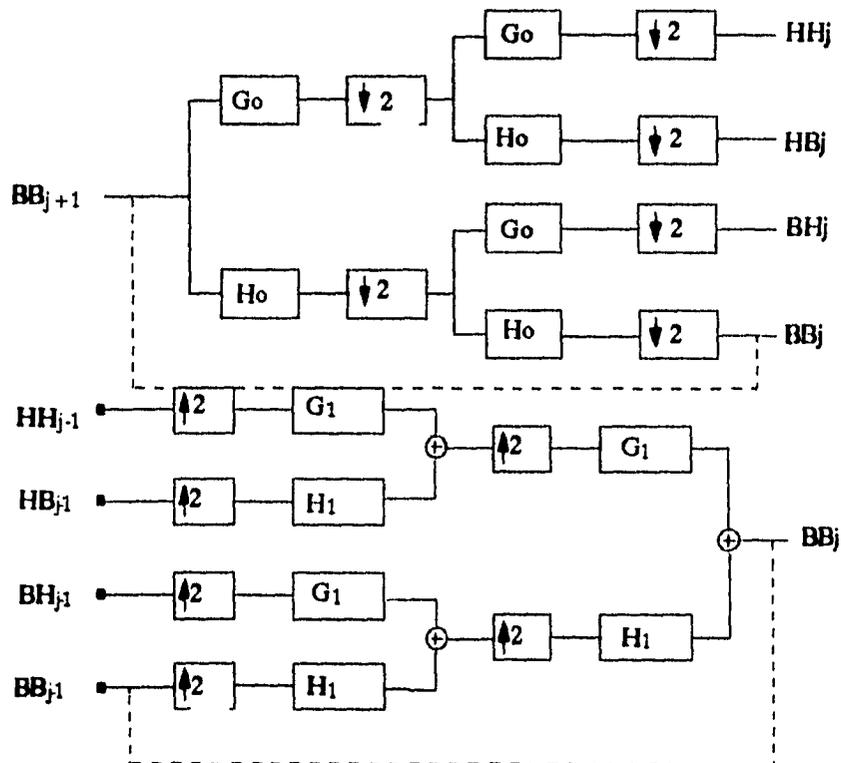


Figura 6.8 Análisis Síntesis de imágenes con filtros separables

Sub-banda Alta-Baja.- El filtro pasa altas G es aplicado a las líneas y el filtro

pasa bajas H a las columnas. Las altas frecuencias horizontales son favorecidas en la sub-imagen HB. Las estructuras de orientación vertical son puestas en evidencia.

Sub-banda Baja-Alta.- La imagen es filtrada pasa bajas en líneas y pasa altas en columnas, las altas frecuencias verticales son favorecidas en la sub-imagen BH. Las estructuras de orientación horizontal son favorecidas.

Sub-banda Baja-Baja.- La imagen es filtrada pasa bajas en líneas y columnas, las frecuencias bajas residuales son recuperadas en la sub-imagen BB. Representando la aproximación de la imagen original a una resolución de 1/2.

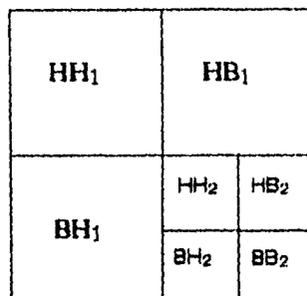


Figura 6.9 Partición frecuencial

6.4 IMPLANTACION DE UN METODO SOBRE UNA ESTRUCTURA DE DATOS MULTIRESOLUCION

Para implantar un método sobre una estructura de datos multiresolución se utilizará la sub-banda baja-baja de la representación piramidal pasa banda ortogonal, que como ya se mencionó, la imagen es filtrada pasa bajas en líneas y columnas las frecuencias bajas residuales son recuperadas en la subimagen BB. Representando la aproximación de la imagen original a una resolución de 1/2. Es decir, que se obtendrá una imagen más pequeña, si la imagen original es de 256x256 pixeles, la imagen en multiresolución será de 128x128 pixeles. A esta imagen se le aplicará el método de detección de bordes Gradiente Roberts, posteriormente se le aplicarán los programas de detección e identificación de figuras y a la imagen resultante se le aplicará una proyección 1:4 para obtener una imagen del mismo tamaño que la imagen original, es decir, no se aplicará el módulo de síntesis para recuperar la imagen. La teoría de los puntos anteriores, se mostró con la finalidad de comprender esta teoría y poder aplicar los programas

elaborados por anteriores tesis, al aplicar estos programas se obtendrá una imagen en multiresolución correspondiente a la sub-banda baja-baja.

6.5 COMPARACION DE LOS RESULTADOS OBTENIDOS EN MONORESOLUCION Y MULTIRESOLUCION

En este punto se realizó una comparación de los tiempos de procesamiento entre los resultados obtenidos en monoresolución y multiresolución (Para ello se utilizó una computadora personal, 386 a 33 Mega Hertz, con 2 Mega de memoria RAM). La imagen original es de 256x256 pixeles y los tiempos de procesamiento están dados en segundos.

El tiempo de procesamiento en monoresolución para obtener la imagen borde de la imagen original fué de 72 segundos.

El tiempo para obtener la imagen en multiresolución de la imagen original fué de 12 segundos (imagen sub-banda baja-baja). El tiempo para obtener la imagen borde de la imagen en multiresolución fué de 18 segundos. El tiempo para realizar la proyección de la imagen borde en multiresolución, para obtener una imagen borde del mismo tamaño que la imagen original, fué de 4 segundos. En conclusión, el tiempo para obtener la imagen borde proyectada fué de 34 segundos.

Analizando los tiempos de procesamiento, se demuestra la hipótesis de que el tiempo es menor aplicando datos en multiresolución.

En las páginas 106 y 107 se muestran las imágenes en monoresolución y en multiresolución al aplicar el primer programa de cómputo, esto con la finalidad de ilustrar las imágenes resultantes.

En las páginas 108 y 109 se muestran las imágenes en monoresolución y multiresolución para aplicar el segundo programa de cómputo, lo mismo se puede observar en las páginas 110 y 111 que corresponden a las imágenes en monoresolución y multiresolución para aplicar el tercer programa de cómputo.

Para las imágenes obtenidas en monoresolución se observan dos imágenes únicamente, una corresponde a la imagen original (imagen superior) y la segunda a la imagen de bordes (imagen inferior).

Para las imágenes obtenidas en multiresolución, se observan cuatro

imágenes. La primera corresponde a la imagen original (imagen superior), la segunda es la imagen obtenida en multiresolución (imagen central izquierda), la tercera es la imagen de bordes en multiresolución (imagen central derecha) y la cuarta imagen corresponde a la proyección de la imagen de bordes anterior (imagen inferior).

Comparando las imágenes de bordes resultantes en monoresolución y multiresolución se observan pequeñas distorsiones en el contorno de las figuras.

Al aplicar los programas de identificación y detección de figuras con las imágenes de bordes obtenidas en monoresolución y multiresolución, se tiene que los programas funcionan correctamente con los dos tipos de imagen de bordes. Esto implica que las pequeñas distorsiones en las figuras de las imágenes obtenidas en multiresolución, no alteran los programas de cómputo.

En conclusión, si lo que se pretende únicamente es la identificación de objetos (como es el propósito de éste trabajo) se pueden aplicar los resultados obtenidos en multiresolución y con esto hacer el sistema más rápido.

En la industria dependería de la calidad o complejidad de los objetos, por ejemplo cierto tipo de engranes o piezas de relojería, muy probablemente no se podrían detectar o identificar correctamente si se aplicaran estructuras en multiresolución, por lo que se tendría que aplicar datos en monoresolución.

IMAGENES EN MONORESOLUCION

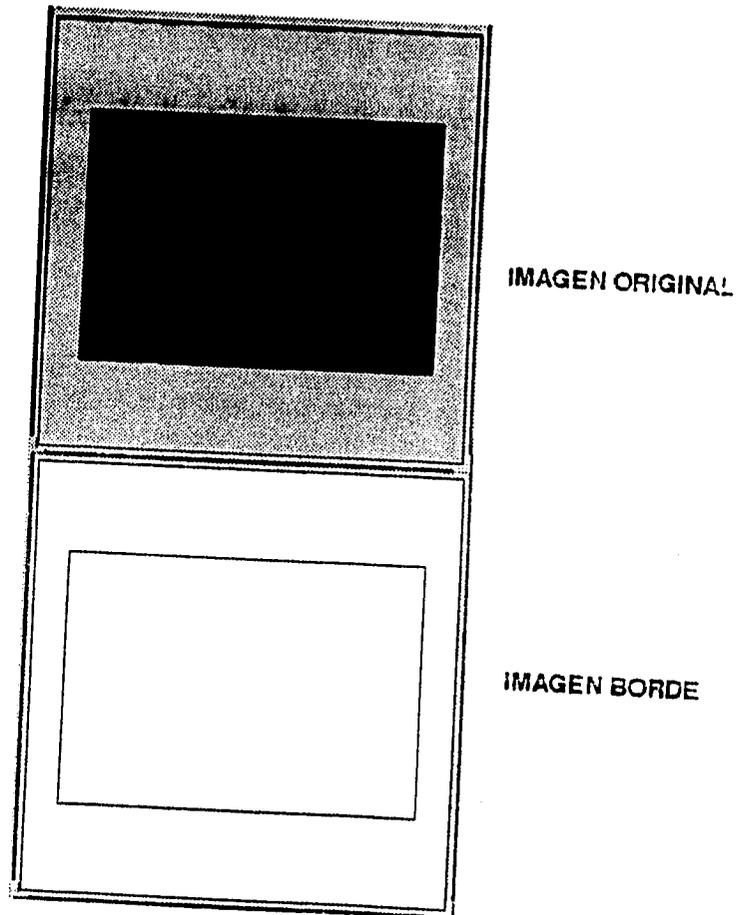


Figura 6 10 Imagen en Monoresolucion:Prog 1 106

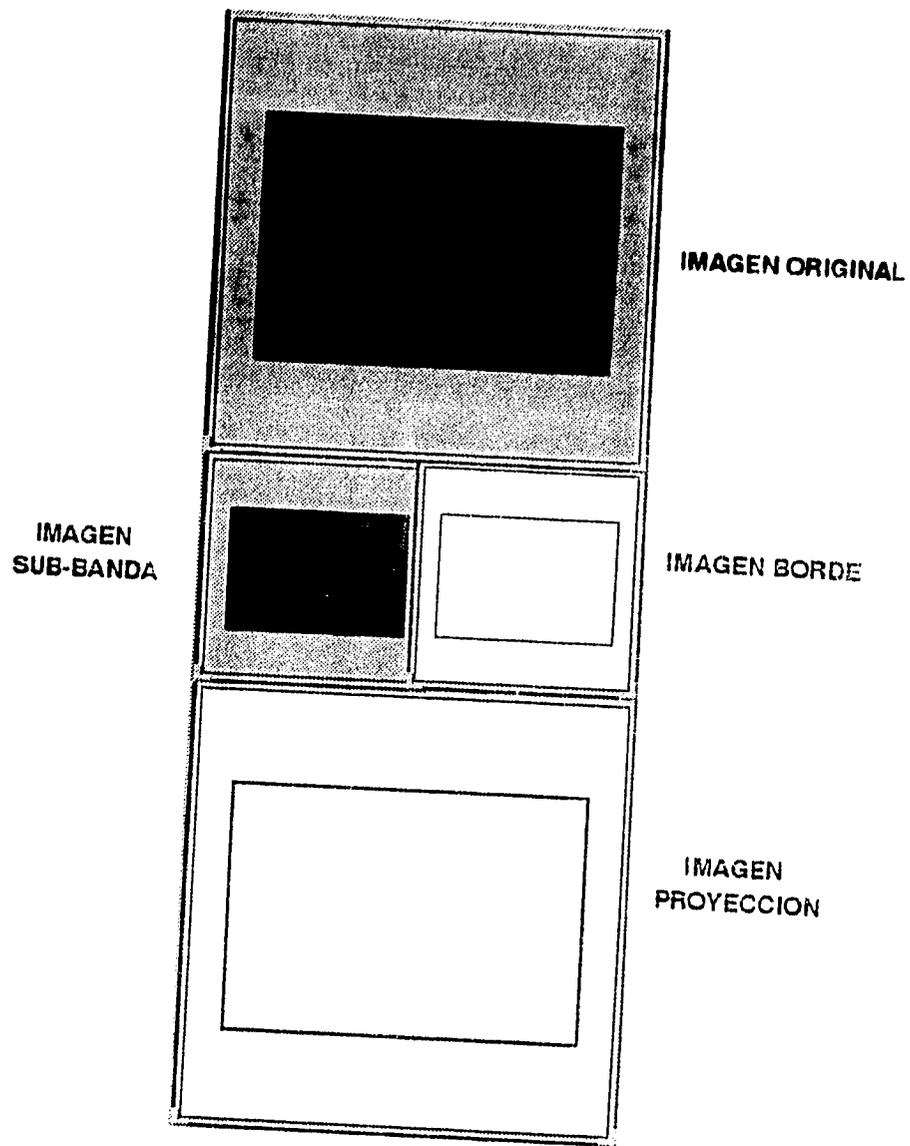


Figura 6.11 Imagen en Multiresolución Prog. 1 107

IMAGENES EN MONORESOLUCION

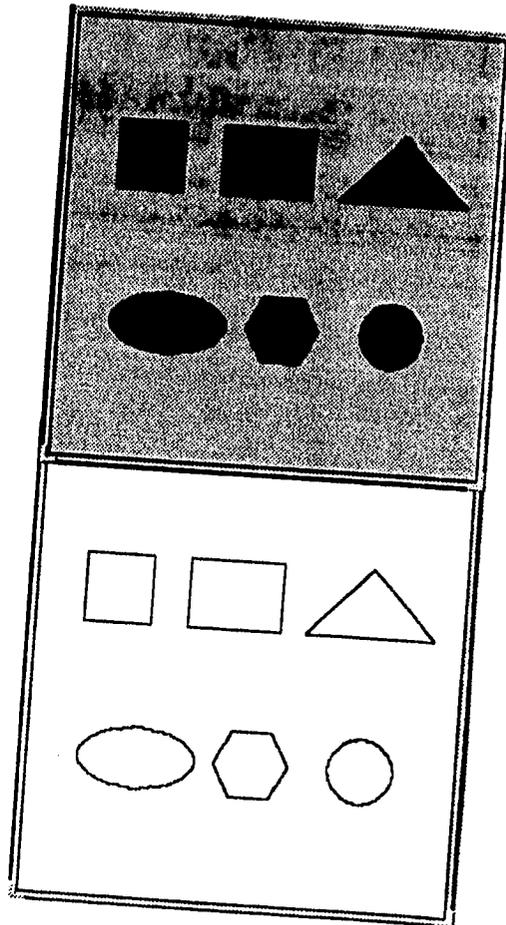


IMAGEN ORIGINAL

IMAGEN BORDE

Figura 6.12 Imagen en Monoresolución: Prog 2 108

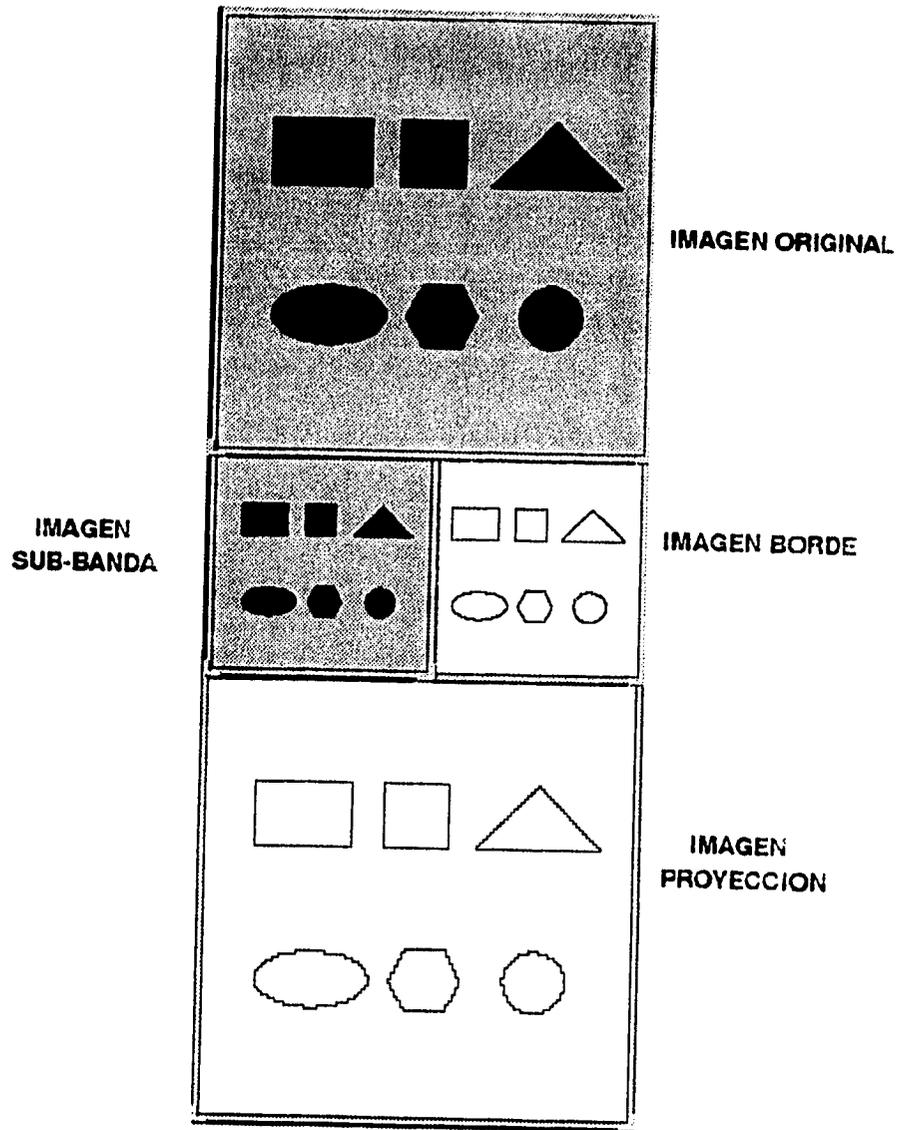


Figura 6.13 imagen en Multiresolucion.Prog 2 109

IMAGENES EN MONORESOLUCION

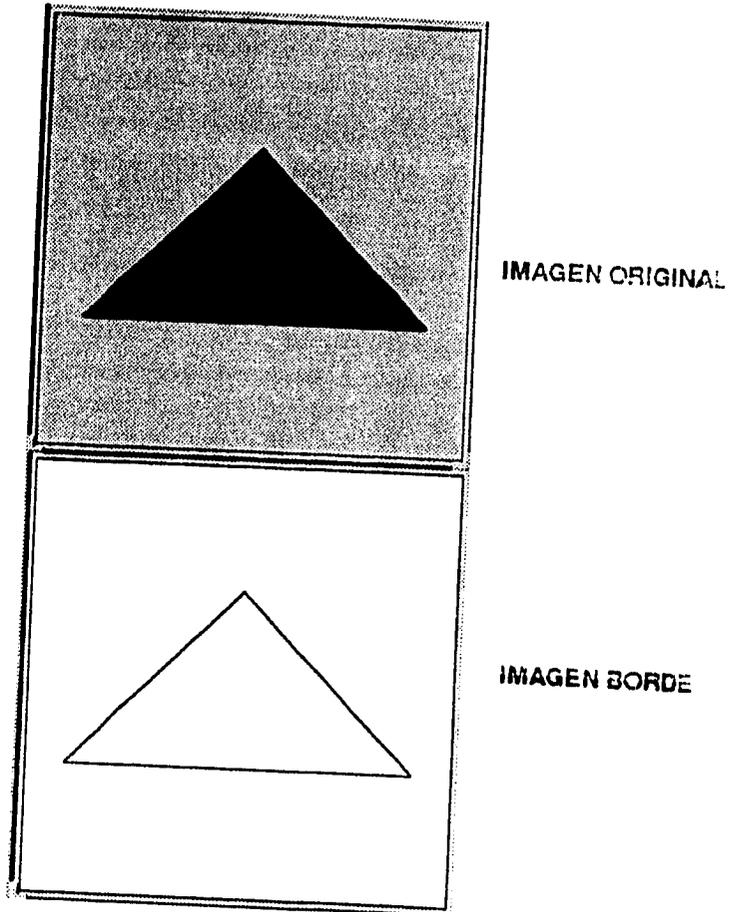


Figura 6 14 imagen en Monoresolucion: Prog 3 110

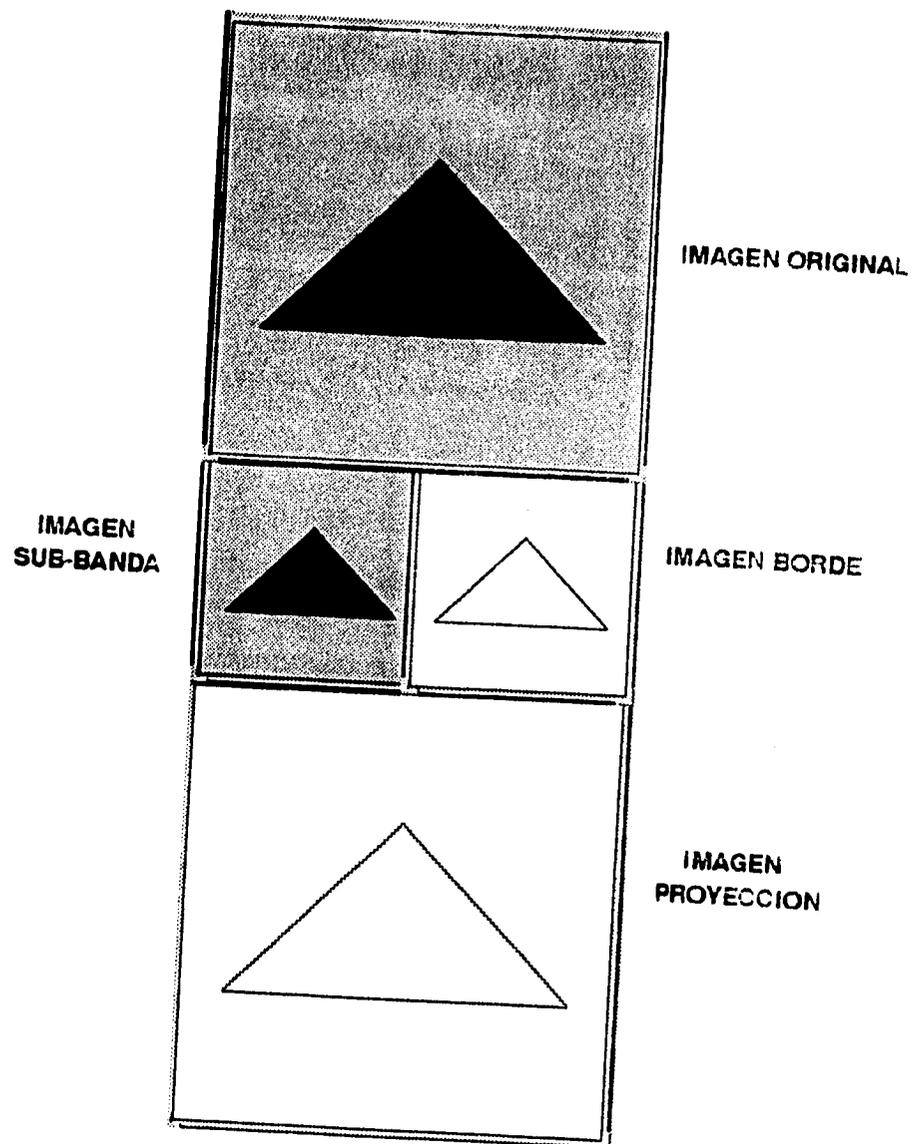


Figura 6.15 imagen en Multiresolución.Prog 3 111

CONCLUSIONES

La identificación de objetos es muy importante en el análisis de imágenes, pero es especialmente importante en la industria, porque los objetos necesitan ser medidos y seleccionados con alta precisión.

Esta tesis proporciona una herramienta para estudiantes e investigadores en el campo del procesamiento digital de imágenes, en el área de reconocimiento de patrones. El objetivo de la tesis es el reconocimiento de objetos industriales, los objetos en la industria pueden tener muy variadas formas; pero estos objetos tienden a tener una forma geométrica, por lo que en este trabajo se analizaron figuras geométricas

En los objetos industriales, al igual que en las figuras geométricas, se analizan sus bordes por regiones, y estas regiones tendrán una forma geométrica (circular, elipsoidal, hexagonal, cuadrada o con forma de recta o curva).

Con la teoría y las pruebas proporcionadas en esta tesis, el reconocimiento de objetos industriales será más sencillo, únicamente los programas serán más elaborados, debido a que los objetos en la industria son más complejos que las figuras geométricas analizadas en esta tesis.

Esta tesis tiene su aplicación en la industria, por ejemplo:

- En una línea de ensamble se pueden tener muchos objetos industriales, en la cual es importante que cada objeto sea medido y/o identificado con gran precisión, si un objeto es detectado que no cumple con ciertas características importantes (forma, tamaño y color), entonces el objeto es aislado de la línea.
- En una línea de ensamble se pueden tener muchos objetos industriales, los cuales pueden ser identificados y localizados con precisión, para que un brazo robot pueda seleccionarlos y ensamblarlos posteriormente.
- Puede ser también aplicada en la industria de la comida, por ejemplo, en una industria seleccionadora de fruta, la cual puede identificar la fruta en base a su forma, tamaño y color.

Observando las tres aplicaciones anteriores, se tiene que dichas aplicaciones se encuentran en base a un sistema de visión. Como se analizó en el

capítulo 2, un sistema de visión comprende 3 componentes principales: Adquisición de imágenes, procesamiento y despliegue. Esta tesis fue enfocada al procesamiento y despliegue de la información. Para aplicarla a la industria, se necesitaría un costo para la adquisición de imágenes, en el que se involucra una cámara de video con salida de video compuesto NTSC, una computadora con una tarjeta digitalizadora de video y un monitor de video, lo anterior implica que el costo para aplicarla a la industria no sería tan grande; pero este costo se vería aumentado si lo que le interesa a cierta industria es el tiempo de respuesta y su eficiencia.

Para la identificación de las piezas mecánicas se observó que estas se identifican correctamente, únicamente los programas de cómputo son más elaborados. Pero al detectar o localizar objetos más complejos, permite comprobar que gracias a los métodos de detección de bordes y a las técnicas de segmentación se puede identificar cualquier tipo de objeto.

La complejidad de los programas depende del tipo de pieza mecánica que se quiera identificar, estos programas se pueden simplificar o complicar más dependiendo de su aplicación en la industria.

Finalmente los objetivos de esta tesis se cumplieron con éxito, al comprobar la confiabilidad de los programas realizados con las figuras de prueba, esos programas se pueden adaptar fácilmente a una situación real en la industria.

REFERENCIAS BIBLIOGRAFICAS

[RB1] MACHINE VISION AND DIGITAL IMAGE PROCESSING FUNDAMENTALS

Galbati Louis J., Jr.
Prentice Hall, Englewood Cliffs
1990 United States of America

[RB2] MATRIX STRUCTURED IMAGE PROCESSING

Dougherty Edward R., Giardina Charles R.
Prentice Hall, Inc, Englewood Cliffs, New Jersey
1987 United States of America

[RB3] DIGITAL IMAGE PROCESSING

Pratt William K.
A Wiley - Interscience publication John Wiley & sons, inc
1991 Second Edition

[RB4] MACHINE VISION: THEORY, ALGORITHMS, PRACTICALITIES

Davies E. R.
Academic Press
1980 London

[RB5] FUNDAMENTALS OF DIGITAL IMAGE PROCESSING

Jain Anil K.
Prentice Hall, Englewood Cliffs
1989 United States of America

[RB6] PATTERN RECOGNITION ENGINEERING

Nodler Norton, Smith Eric P.
John Wiley & Sons Inc
1982 United States of America

[RB7] DIGITAL IMAGE PROCESSING AND COMPUTER VISION

Schalkoff Robert J.
Wiley
1989

[RB8] DIGITAL IMAGE PROCESSING

Gonzalez Rafael C., Wintz Paul
Addison Wesley
1987

[RB9] PATTERN RECOGNITION AND IMAGE PREPROCESSING

Sing- Tze Bow
Marcel Bekker, Inc
1982 New York, USA

[RB10] Computer and Robot Vision Volumen 1

Haralick Robert M, Shapiro Linda G.
Addison Wesley Publishing Company
1992 USA

**[RB11] APPROCHES D'ESTIMATION ET DE COMPENSATION DE
MOUVEMENT MULTIREOLUTIONS POUR LE CODAGE DE
SEQUENCES D'IMAGES**

Nadia Baaziz
Tesis doctoral, L'UNIVERSITE DE RENNES I,
Octubre 1991, Francia

[RB12] TURBO C PROGRAMACION AVANZADA

Schildt Herbert
Borland-Osborne/McGraw-Hill
Segunda Edición 1990

[RB13] EL LENGUAJE DE PROGRAMACION C

Kernighan Brian W, Ritchie Dennis M.
Prentice - Hall Hispanoamérica, S.A.
Segunda Edición 1991

[RB14] MANUAL DE REFERENCIA C

Schildt Herbert
Osborne/McGraw-Hill
Segunda Edición 1990

REFERENCIAS BIBLIOGRAFICAS

- [RH1] A THEORY FOR MULTIREOLUTION SIGNAL DECOMPOSITION:
THE WAVELET REPRESENTATION. Mallat Stephane G.
IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 2
No. 7 Julio 1989
- [RH2] QUADRATURE MIRROR FILTER BANKS M-BAND EXTENSIONS AND
PERFECT RECONSTRUCTION TECHNIQUES. P.P. Vardyanathan
IEEE ASSP Magazine, Julio 1987