

UNIVERSIDAD DE NAVARRA
ESCUELA SUPERIOR DE INGENIEROS INDUSTRIALES
SAN SEBASTIAN



IMPLANTACION DE ALGORITMOS DE DETECCION DE
BORDES EN ESTRUCTURAS HARDWARE PARA APLICACIONES
DE VISION EN TIEMPO REAL

MEMORIA

que para optar al Grado de Doctor Ingeniero Industrial

presenta

CESAR ALBERTO CHINGUEL ARRESE

San Sebastián, Julio 1994

UNIVERSIDAD DE NAVARRA
ESCUELA SUPERIOR DE INGENIEROS INDUSTRIALES
SAN SEBASTIAN



IMPLANTACION DE ALGORITMOS DE DETECCION DE
BORDES EN ESTRUCTURAS HARDWARE PARA APLICACIONES
DE VISION EN TIEMPO REAL

MEMORIA

que para optar al Grado de Doctor Ingeniero Industrial

presenta

CESAR ALBERTO CHINGUEL ARRESE

San Sebastián, Julio 1994

*A mis padres, Felix y Yolanda,
ya fallecidos.*

*A Cecilia, mi esposa. Doy
gracias a Dios por haber puesto
en mi camino una gran mujer;
no la merezco.*

*A mis queridos hijos Alvaro,
Alejandro, María José y los que
vengan. Estas páginas son
también suyas.*

*A mis hermanos, Enrique,
Jorge, y mi querida Liliana.*

AGRADECIMIENTOS

Quiero expresar mi agradecimiento al Dr. Julián Flórez Esnal, de quien he recibido apoyo constante durante el desarrollo de este trabajo. He tenido la suerte de tener como director de tesis a un profesional excepcional, y deseo hacer público mi reconocimiento por todo lo que me ha enseñado; le agradezco especialmente por haberme hecho merecedor de su amistad.

A la Universidad de Navarra, institución en la que he tenido la suerte de estudiar y trabajar estos años. Trabajar en sus instalaciones es un verdadero placer gracias al ambiente tan humano que se respira. A la Escuela de Ingenieros de San Sebastián, a las autoridades, personal docente y administrativo. Todos ellos me han hecho sentir como en casa durante estos años.

A mis muy queridos amigos del Departamento de Electricidad, Electrónica y Automática del CEIT, a Mario García, Manolo Pérez Toca, Fernando Arizti, Luis Fontán, Joaquín de Nó, Santiago Corera, Jose Luis Pedrejón, Antonio Salterain, David Remondo, etc. He sido muy afortunado al compartir despacho y tiempo con Ana Macarulla, Elisabeth Viles, Iñigo Unanue, Francisco Arregui y Javier Ostolaza, Jose M^a Dorronsoro y deseo mostrarles todo mi agradecimiento por el apoyo recibido. Quiero agradecer especialmente a Igor Armendariz y Ana Macarulla por su generosidad y paciencia para compartir sus conocimientos. Nombrarlos a todos sería interminable, especialmente a la generación de investigadores jóvenes que han ingresado en los últimos años, con la convivencia diaria he aprendido muchas cosas de todos ellos.

Quiero hacer especial mención a mi compañera de trabajo Elisabeth Viles, con quién hemos colaborado estrechamente en el desarrollo de este proyecto. Sus críticas y consejos han contribuido a enriquecer este trabajo.

A mis amigos de los Laboratorios CIM de la Escuela y Robótica del CEIT, les agradezco por invitarme a sus reuniones de trabajo. Gracias a a sus proyectos he podido aproximarme a las necesidades de la industria con respecto a sistemas de visión artificial.

A la Universidad de Piura, a mis amigos de la Facultad de Ingeniería. A todos mis profesores, que se empeñaron en que, pese a mis limitaciones intelectuales, me hiciera ingeniero. Especial mención merecen el Dr. Antonio Mabres y el Ing. Rafael Estartús por su constante aliento y apoyo. Les quiero agradecer a todos en la persona del que fue mi profesor más querido, el Dr. Ramón Múgica Martínez, ya fallecido.

Vaya también mi agradecimiento a la Asociación de Amigos de la Universidad de Navarra, con cuya desinteresada ayuda y constante apoyo he podido contar.

RESUMEN

El objetivo general de la presente tesis es el estudio de estrategias que hagan posible la implantación de algoritmos de detección de bordes en estructuras hardware para aplicaciones de visión en tiempo real. La aplicación inmediata de este estudio consiste en lograr que un sistema de visión pueda ser empleado como un sensor más dentro de una arquitectura hardware encargada de controlar un brazo flexible de robot.

Para ello se estudian los algoritmos de detección de bordes más conocidos. Los criterios tenidos en cuenta para la selección de los mismos son los siguientes,

- que permitan filtrar la señal mientras encuentran los bordes,
- que tengan un costo computacional bajo,
- que puedan ser implantados en arquitecturas hardware,
- que *Detecten y Localicen* con relativa sencillez y precisión.

Con ello se diseña primeramente un nuevo filtro, y se altera después su estructura para hacerlo implantable en hardware. Se establecen los pasos a seguir para convertir sus coeficientes a valores enteros. Asimismo, se proponen criterios para discriminar información de la imagen que no va a ser empleada.

Se proponen y evalúan distintas configuraciones que generan variantes en el hardware final. Se determina una *Configuración Final* que permite su implantación hardware. Por último se propone el hardware, compuesto por componentes disponibles en el mercado, que puede ser empleado en el cálculo de los bordes de una imagen en tiempo real.

INDICE

| | | |
|-----------|---|------------|
| 1. | INTRODUCCIÓN | 1-1 |
| 1.1 | Enfoque Industrial | 1-3 |
| 1.2 | Evolución Histórica | 1-4 |
| 1.3 | Planteamiento del Problema y Objetivos | 1-8 |
| 1.4 | Estructura de la Tesis | 1-11 |
| | | |
| 2. | DETECCIÓN DE BORDES | 2-1 |
| | | |
| 2.1 | Métodos Basados en el Gradiente | 2-2 |
| 2.1.1 | Método de Sobel (1973) | 2-8 |
| 2.1.2 | Método de Robert (1965) | 2-9 |
| 2.1.3 | Otros métodos con Gradiente | 2-10 |
| 2.2 | Métodos Empleando el Laplaciano | 2-11 |
| 2.2.1 | Teoría de Marr y Hildreth (1980) | 2-17 |
| 2.2.2 | Algoritmo de Canny (1986) | 2-22 |
| 2.3 | Sumario de los Métodos Estudiados | 2-24 |
| | | |
| 3. | SEGUNDA DERIVADA DEL GAUSIANO | 3-1 |
| | | |
| 3.1 | Comportamiento de la Segunda del Gausiano | 3-2 |
| 3.1.1 | Alcance del Filtro | 3-10 |
| 3.1.1.1 | Alcance del Filtro en Una Dimensión | 3-10 |
| 3.1.1.2 | Alcance del Filtro en Dos Dimensiones | 3-11 |
| 3.1.2 | Tamaño del Filtro | 3-19 |

| | | |
|------------|---|-------------|
| 3.1.2.1 | Tamaño del Filtro en Una Dimensión | 3-21 |
| 3.1.2.1.1 | Consideraciones en Espacio Continuo | 3-22 |
| 3.1.2.1.2 | Consideraciones en Espacio Discreto | 3-26 |
| 3.1.2.2 | Tamaño del Filtro en Dos Dimensiones | 3-29 |
| 3.1.3 | Localización-Paso por Cero | 3-31 |
| 3.1.3.1 | Empleo de Umbrales | 3-35 |
| 3.1.3.2 | Relaciones con la Desviación | 3-36 |
| 3.1.3.3 | Criterio Considerando Máximos | 3-40 |
| 3.1.4 | Aproximación del Filtro con Valores Enteros | 3-44 |
| 3.1.4.1 | Análisis en Una Dimensión | 3-47 |
| 3.1.4.2 | Análisis en Dos Dimensiones | 3-49 |
| 3.2 | El Umbral y la Primera Derivada del Gausiano | 3-58 |
| 3.2.1 | Consideraciones en Una Dimensión | 3-59 |
| 3.2.2 | Consideraciones del Umbral en Dos Dimensiones | 3-64 |
| 3.3 | Sumario del Tercer Capítulo | 3-70 |
| | | |
| 4. | IMPLANTACIÓN DEL ALGORITMO EN HARDWARE | 4-1 |
| | | |
| 4.1 | Pipeline para Una Línea de Vídeo | 4-3 |
| 4.2 | Tamaño de G1 en Dos Dimensiones | 4-7 |
| 4.3 | Configuraciones Hardware | 4-11 |
| 4.3.1 | Configuración A | 4-12 |
| 4.3.2 | Configuración B | 4-14 |
| 4.3.3 | Configuración C | 4-16 |
| 4.3.4 | Configuración D | 4-18 |
| 4.3.5 | Configuración E | 4-21 |
| 4.4 | Bloques Funcionales | 4-24 |
| 4.4.1 | Conversión Analógica Digital | 4-24 |

| | | |
|-----------|--|-------------|
| 4.4.2 | Primera Derivada del Gausiano (PDG) | 4-24 |
| 4.4.3 | Bloque MaxAbs | 4-27 |
| 4.4.4 | Bloque MaxAbsSign | 4-28 |
| 4.4.5 | Bloque Ceros | 4-28 |
| 4.5 | Sumario del Cuarto Capítulo | 4-36 |
| 5. | DISEÑO DEL HARDWARE | 5-1 |
| 5.1 | Dispositivo de Carga Acoplada (BDCA) | 5-3 |
| 5.2 | Convertidor de Video Analógico Digital (BVAD) | 5-4 |
| 5.3 | Convolución con G1 (BCG1) | 5-7 |
| 5.3.1 | Procesador PDSP16401 | 5-8 |
| 5.3.2 | Procesador ZR33771-45 | 5-9 |
| 5.3.3 | Procesador PDSP16488 | 5-11 |
| 5.4 | Bloque Mayor en Valor Absoluto (BMVA) | 5-17 |
| 5.5 | Bloque memoria Buffer (BMBF) | 5-20 |
| 5.6 | Bloque Mayor en Valor Absoluto más Signo (BMAS) | 5-21 |
| 5.7 | Bloque Ceros (BDLC) | 5-22 |
| 5.8 | Retardo Acumulado | 5-25 |
| 6. | CONCLUSIONES Y TRABAJOS FUTUROS | 6-1 |
| 6.1 | En Cuanto al Algoritmo Detector de Bordes | 6-2 |
| 6.2 | En Cuanto al Hardware | 6-5 |
| 6.3 | En Cuanto a la Captura de Información | 6-6 |
| 6.4 | Líneas de Trabajo | 6-8 |
| 7. | REFERENCIAS BIBLIOGRÁFICAS | 7-1 |

| | | |
|-------------------|--|-------------|
| APÉNDICE A | EL GAUSIANO Y SUS DERIVADAS | A-1 |
| A.1 | Relaciones en Una Dimensión | A-2 |
| A.1.1 | El Gausiano | A-2 |
| A.1.2 | Primera Derivada | A-3 |
| A.1.3 | Segunda Derivada | A-4 |
| A.2 | Relaciones en Dos Dimensiones | A-5 |
| A.2.1 | El Gausiano | A-5 |
| A.2.2 | Primera Derivada | A-6 |
| A.2.3 | Segunda Derivada | A-9 |
| | | |
| APÉNDICE B | DESCRIPCIÓN DE LA SEÑAL DE VÍDEO | B-1 |
| B.1 | Modo Entrelazado | B-11 |
| B.2 | Modo No-entrelazado | B13 |
| | | |
| APÉNDICE C | PROCESADOR DIGITAL DE SEÑAL PDSP14188 | B-1 |
| C.1 | Operación Básica | B-1 |
| C.2 | Precisión en la Salida | C-2 |
| C.3 | Saturación de Salida | C-2 |
| C.4 | Salida Binaria | C-3 |
| C.5 | Multiplicador | C-3 |
| C.6 | Operación con Líneas de Retardo | C-4 |
| C.7 | Vídeo Entrelazado | C-5 |
| | | |
| APÉNDICE D | PRUEBAS Y EJEMPLOS | D-1 |

1. INTRODUCCION

Desde un punto de vista general, la visión artificial es considerada como el conjunto de técnicas que permite a una máquina ser capaz de reproducir, parcialmente, el sistema de visión humano y sus características asociadas. El objetivo final de este planteamiento es el de llegar a dar una interpretación de la imagen similar a la que podría dar un ser vivo. Este punto de vista distingue dos grandes campos de estudio, el de la captación y el de la interpretación e incluye el procesamiento dentro del estudio de la interpretación.

El impulso para desarrollar la visión artificial proviene esencialmente de dos motores, el del mercado industrial actual, que consiste en promover nuevas técnicas para mejorar las prestaciones de los sistemas orientados a proporcionar un Control de Calidad Total en menos tiempo y con mayor precisión, y segundo, el de los profesionales dedicados a la Investigación y

Desarrollo, para los que la visión artificial consiste en un conjunto de técnicas de captación y procesamiento de imágenes orientadas a obtener de ellas la información necesaria para aplicaciones muy concretas.

Un tercer punto de vista no tocado en esta tesis consiste en desarrollar aplicaciones en las que la visión artificial no se limita a un campo específico, sino que parte de una información bi o tridimensional, a la que se añade la obtenida por medio de sensores táctiles y de otro tipo para generar imágenes virtuales. Este enfoque acentúa la diferenciación entre sistema de visión artificial y sistema humano de visión, ya que cada uno dispone de técnicas diferentes.

Se puede afirmar que hay tareas que son muy fáciles de realizar para un sistema de visión artificial y muy difíciles para el sistema de visión humano, por el contrario, es impensable que se realicen otras con visión artificial siendo muy fáciles para el sistema humano de visión. Así, la evolución de la visión artificial no tiene por qué estar ligada a imitar estrictamente el sistema humano de visión.

Sin embargo, la referencia al sistema de visión humano es válida en el sentido que es el único sistema de visión ordenado a su fin que funciona correctamente, esto lo convierte en el primer punto de referencia para cualquier desarrollo de visión, sin que esto quiera decir que sea susceptible de ser imitado con exactitud. Una cosa es comprender cómo se resuelve un problema fisiológicamente, y otra muy diferente es implantar una solución física del mismo.

1.1 Enfoque Industrial

En la industria de los noventa, la visión artificial se está caracterizando por tener una demanda latente, existiendo un desfase entre el grado de automatización que tienen actualmente las empresas y el número de sistemas de visión instalados, ya sea en España o en el mercado europeo. Se puede afirmar que este retraso no es tecnológico, sino de mercado. Se posee un "knowhow" suficiente para afrontar varias aplicaciones, pero la tecnología no está introducida en el mercado o lo hace con lentitud. Esto se debe al desconocimiento de la tecnología por parte de los usuarios, a la desconfianza debida en parte a la irregularidad de la oferta, a algunos fracasos iniciales, y al precio relativamente elevado de estas tecnologías emergentes.

Estos factores negativos están desapareciendo debido a que el coste del hardware y de las instalaciones está descendiendo notablemente. Sin embargo, existe mayor oferta que demanda, en buena parte porque no se conoce el grado de información que se puede obtener por esta técnica. Industrialmente casi no se utiliza visión artificial en aplicaciones tridimensionales en tiempo real, y lo mismo pasa con las aplicaciones en color. Las aplicaciones bidimensionales en grises, casi todas orientadas al control de calidad y a la monitorización del tráfico, están bastante extendidas dentro de lo que cabe, siempre con limitaciones de resolución y velocidad de proceso.

Los sistemas de visión artificial que comenzaron a ser utilizados de forma muy limitada hace veinte años están todavía en estado incipiente, considerando la cantidad de tiempo que se viene trabajando en el tema.

Actualmente los investigadores dedicados al campo de la visión artificial tienen la posibilidad de desarrollar sistemas de visión que, bien utilizados, pueden mejorar mucho la calidad de vida del ser humano.

1.2 Evolución Histórica

Los problemas de percepción visual han atraído a científicos de todas las épocas, desde Newton (1709) que sentó las bases para los posteriores trabajos de visión en color; Helmholtz (1910) aportó su tratado de óptica fisiológica; Wertheimer (1912,1923) trabajó las imágenes representadas secuencialmente; Miles (1931), Wallach (1953), y O'Connell (1953) trabajaron experimentalmente la visión en color; Julesz (1960) trabajó en visión binocular; Brinbley (1970) estableció su teoría tricromática del color a partir de los trabajos de los anteriores. Los primeros trabajos sobre detectores de rasgos los realizaron Barlow (1953), Hubel y Weisel (1965).

La motivación que impulsó los estudios sobre procesamiento de imágenes se sitúa en torno a 1977 y tiene sus orígenes en el deseo de analizar la gran cantidad de información sobre satélites con una herramienta como el ordenador. Así, Hendlich (1971) estudió una técnica de reconocimiento para determinar el movimiento de nubes a partir de secuencias de fotografías de satélite. Leese, (1971) publica una técnica automática para la obtención del movimiento de nubes a partir de datos obtenidos con satélites geoestacionarios usando correlación cruzada. Un buen resumen fue publicado por Aggarwal (1975) estudiando las técnicas para la determinación del movimiento a partir de imágenes poligonales.

Por la cantidad de trabajos publicados por esta época, se puede apreciar que el mayor esfuerzo investigador se centró en analizar secuencias de imágenes de escenas dinámicas, dándole menor importancia al procesamiento de la información a nivel más bajo. El optimismo llevó a trabajar antes en la interpretación que en la captación correcta de información.

Al disponer de ordenadores cada vez más potentes, el análisis de imágenes

comienza a interesar, como tal, a partir de 1977; esto puede observarse con los trabajos de Roberts(1979) encaminados a la percepción de sólidos en tres dimensiones. Este trabajo presenta un estudio sobre la segmentación y el reconocimiento de objetos, y sobre el análisis requerido para determinar la posición de un objeto en tres dimensiones. Desde entonces el análisis de secuencias de imágenes de objetos en movimiento recibe una atención relativa a las dificultades que se iban encontrando. Estas dificultades obligaron a reorientar casi por entero la atención al estudio de escenas en dos dimensiones.

A partir de aquí se suceden muchos artículos que estudian geoméricamente el problema, entre los más destacados se puede nombrar a Chow (1977) que estudia el movimiento planar curvilíneo de imágenes, Martin (1979) estudia también una secuencia dinámica de imágenes curvilíneas, Jain (1979) estudió y analizó una secuencia de imágenes por diferencias acumulativas de las mismas.

Si bien hasta entonces los estudios basados en imágenes dinámicas muestran a peatones y coches en movimiento, no se nota ningún interés en utilizar esta información en algún tipo de control, entre otras razones por la tecnología disponible.

Antes de 1980 la investigación está referida principalmente al movimiento en dos dimensiones porque la interpretación de la rotación y profundidad es más compleja en tres que en dos dimensiones. En tres dimensiones algunas partes del objeto pueden desaparecer de la vista como resultado de la rotación del objeto en el espacio tridimensional. Se puede nombrar los trabajos de Roach (1980) sobre la determinación del movimiento de objetos de una secuencia de imágenes partiendo siempre de dos dimensiones.

Roach (1980) realiza una formalización matemática haciendo un paralelo con la manera como el humano percibe la profundidad a partir de dos imágenes

determinando los gradientes de textura. Formula las coordenadas espaciales que relacionan los puntos en el plano detectados por la cámara, con las distancias relativas de las cámaras a los puntos y la distancia de la cámara a unos ejes espaciales de referencia. En consecuencia, los cambios entre dos puntos en una secuencia de imágenes, llamado disparidad, es causado por el movimiento relativo entre el objeto y la cámara. Teniendo la información de otra cámara, por triangulación, se puede determinar la profundidad de estos puntos. Este trabajo asumía una segmentación de objetos óptima.

Por las dificultades encontradas, a principios de los ochenta se dividió el problema en tres campos de estudio claramente diferenciados, la detección, la segmentación, y la interpretación, con sus respectivas implantaciones en un ordenador. Para continuar con la secuencia del párrafo anterior, se comentará brevemente algunos trabajos importantes sobre detección y análisis de movimiento que aparecen posteriores a esta fecha, haciendo hincapié en que todas ellas asumen que se ha realizado una buena detección de bordes, premisa que simplifica mucho el problema, y que no es necesariamente cierta. Lo cierto es que en 1980 aparecen publicados algunos estudios referidos a la detección y localización de los bordes de un objeto dentro de una escena estática.

A partir de 1980 se publican diversos métodos más o menos eficaces para reconocer y segmentar objetos en movimiento. Huang (1989) presenta un trabajo basado en las conclusiones de Roach (1980) y estudia el problema de la posición de objetos en movimiento formulando una solución algebraica desde ecuaciones polinómicas simultáneas. Después de 1990 aparecen otros trabajos orientados a reconocer objetos en movimiento que no se comentarán dado que el interés se centrará en la detección de bordes, objetivo de esta tesis. A continuación se mencionan algunos trabajos encaminados a mejorar la detección y ubicación de los bordes dentro de una imagen.

Rosenfeld y Thurston (1971) aplicaron operadores de diferentes tamaños

para aumentar la sensibilidad y reducir el ruido, pero sin evaluar seriamente los algoritmos aplicados; Hueckel (1973) trabajó un operador que descubría la dirección del borde que mejor se ajustaba a un cambio de intensidad. Los trabajos desarrollados hasta aquí no restringían las condiciones de la imagen, sino que trataban de encontrar soluciones teóricas generales para cualquier imagen, sin importar su tamaño, intensidad luminosa, número de bytes por pixels, etc., constituyendo éste el primer enfoque del problema.

El segundo enfoque restringía el mundo real a imágenes de escenas constituidas por objetos opacos, blancos, sobre un fondo negro, y con ángulos siempre rectos. Así, y bajo estas restricciones, Binford y Horn (1976) y Shirai (1973) publicaron sendos "descubridores de líneas".

Los trabajos de Roberts (1965), Guzmán (1968), Walts(1975), y Mackworth (1973) enfocaron el problema de visión suponiendo resuelta la búsqueda de bordes. Walts probó de modo explícito que un análisis exhaustivo de todas las posibles disposiciones locales de superficies, bordes, y sombras podía llevar a un algoritmo eficiente a interpretar una imagen real; siempre suponiendo que el problema de la búsqueda de bordes había sido superado.

Un tercer enfoque, con Marr y Poggio (1977), consiste en plantear el problema como un complejo sistema jerarquizado de procesamiento de información teniendo como base el trabajo de Horn (1975) que publicó su Análisis de Formas a Partir de lo Sombreado interrelacionando la geometría, intensidad, y reflectancia en una ecuación diferencial.

Marr y Poggio (1979), Poggio y Ullman (1979), junto a Marr y Hildreth (1980), enfocaron un primer esbozo de la imagen haciendo un estudio sobre la detección de ceros. La formación del esbozo primario en bruto lo enfocó Marr (1976). Marr y Hildreth (1980) publican su Teoría de Detección de Bordes, trabajo que ha servido de base para importantes publicaciones posteriores

como la de Canny (1986) donde propone un método computacional para detección de bordes, o la de Torre y Poggio (1986) que hace una buena crítica a la publicación de Canny y otros. Se considera que los trabajos publicados a partir de esta fecha influyen directamente, en menor o mayor medida, en el desarrollo de la presente tesis, por lo que se irán discutiendo a lo largo de la misma.

1.3 Planteamiento del Problema y Objetivos

Con costo computacional elevado, el procesamiento de imágenes está en sus etapas elementales de segmentación, y la impaciencia investigadora ha hecho que se de un salto hacia las etapas de interpretación, lo cual no es necesariamente malo. Sin embargo, esto ha llevado a que los sistemas de visión no sean todo lo robustos que debieran, lo que limita su empleo en la industria. Todo esto convierte a la visión artificial en un problema atacable desde distintos sectores de la ingeniería.

Las nuevas técnicas de multiprocesamiento de información en paralelo, y la aparición en el mercado de procesadores electrónicos más rápidos, ha hecho que los investigadores se vuelquen en el diseño de sistemas que permitan procesar información en tiempo real.

En este marco se encuadra la presente tesis que tiene como finalidad estudiar y fijar las pautas que permitan el desarrollo sistemas de visión que sean capaces de obtener información válida para controlar en tiempo real un brazo de robot flexible. Esto implica reducir notablemente el costo computacional y la velocidad de proceso empleando dispositivos electrónicos, como Transputers y DSPs, optimizados para el procesamiento en paralelo.

- ⊕ El objetivo general es el estudio de estrategias que hagan posible la implantación de algoritmos de detección de bordes en estructuras hardware para aplicaciones de visión en tiempo real. La aplicación inmediata de este estudio consiste en lograr que un sistema de visión pueda ser empleado como un sensor más dentro de una arquitectura hardware encargada de controlar un brazo flexible de robot.

- ⊕ Esta tesis analizará la problemática de los filtros detectores de bordes, buscando soluciones que puedan implantarse directamente en hardware. Esencialmente se trata de poder procesar tanta información como haya en una escena captada por una cámara de vídeo, y hacerlo en tiempo real.

- ⊕ Si la imagen es captada con una cámara convencional, el objetivo primario del sistema a desarrollar será el de poder digitalizar y procesar una imagen de 512x512 pixels ¹ a una cadencia de 25 imágenes por segundo, y una frecuencia de muestreo de 15 MHz. para el sistema actual español. Se puede deducir rápidamente que la electrónica asociada deberá ser bastante rápida, y que la misma juega un papel muy importante en la solución del problema.

- ⊕ Este trabajo forma parte de un proyecto de control más general, siendo los límites establecidos para esta parte los de: digitalizar una señal de vídeo blanco y negro, filtrar esta señal, obtener los bordes de los objetos representados en la escena, y transmitir esta información a otro procesador que se encargará

¹ Pixel es la denominación utilizada para referenciar cada elemento de la imagen.

de realizar el "tracking" de los objetos que se muevan en la imagen. Todas estas tareas en tiempo real.

- ⊕ Los bordes de un objeto contienen información de las características físicas de la escena que los contiene. Estas características pueden consistir en cambios de intensidad originados por los bordes reales de un objeto, límites en una región de un área captada a través de un microscopio, una escena, etc. De esta manera, los resultados que se obtengan de este trabajo serán aplicables tanto a desarrollos simples de control de calidad industrial, como a sistemas de visión más complejos.

- ⊕ Además, se busca establecer criterios que permitan la eliminación de detalles no deseados de la imagen. Estos detalles no son necesariamente ruido, que se supone filtrado, sino que pueden consistir en una parte de la información contenida en la imagen que no interesa para una aplicación en particular.

- ⊕ También, se pretende trazar los perfiles del hardware que permita aplicar algoritmos de visión. Como las características del hardware son cambiantes, se intenta determinar las características de los dispositivos que podrían ser empleados en la aplicación antes mencionada, de tal manera que un sistema de visión artificial pueda en el futuro implantarse en una aplicación de control como un simple sensor más.

1.4 Estructura de la Tesis

El presente trabajo ha sido dividido en siete capítulos cuyo contenido se describe brevemente a continuación.

De este modo, en el Capítulo Segundo se estudian los algoritmos de detección de bordes más conocidos, y se selecciona el filtro Segunda Derivada del Gaussiano para obtener los bordes. En el Capítulo Tercero se analiza la capacidad de este filtro para su implantación en hardware y se establecen sus limitaciones.

En el Capítulo Cuarto se modifica el filtro y se propone una configuración adecuada para la implantación propuesta. En el Capítulo Quinto se propone el hardware que puede ser empleado en el cálculo de los bordes de una imagen en tiempo real. En el Capítulo Sexto se describen las conclusiones y líneas de trabajo futuras. Por último, en el Capítulo Séptimo se nombran las referencias bibliográficas más importantes empleadas en esta Tesis.

Se han incluido cuatro apéndices. En el apéndice A se presentan las expresiones empleadas para los gaussianos y sus derivadas, en el Apéndice B se hace una descripción detallada de la señal de vídeo compuesto empleada, en el Apéndice C se resaltan las características técnicas del DSP propuesto, y en el Apéndice D se presentan a manera de ejemplos algunas pruebas realizadas.

2. DETECCION DE BORDES

La mayoría de los sistemas de visión que se ofrecen actualmente en el mercado aplican los algoritmos que se discuten brevemente a continuación. Los sistemas de visión comerciales que ofrecen el procesamiento de la imagen en tiempo real, tienen un costo bastante elevado, y el procesamiento en tiempo real que desarrollan es muy relativo.

A continuación se presenta algunos de los métodos más comentados por la bibliografía técnica. La mayoría de los sistemas de visión comerciales emplean algunos de estos métodos para procesar imágenes. Se puede afirmar con certeza que prácticamente todos estos sistemas emplean métodos basados en gradiente, dando énfasis en la detección de los bordes sobre la localización de los mismos.

2.1 Métodos Basados en el Gradiente

Considerérese una función analítica $f(x)$ la cual representa un borde típico en un espacio de una dimensión, tal como se muestra en la Figura 2.1. Es razonable considerar al punto x_0 como un "punto-borde". Un camino muy empleado para encontrar x_0 es haciendo consideraciones sobre la primera derivada $f'(x)$, o la segunda derivada $f''(x)$, (Nevatia, 1982).

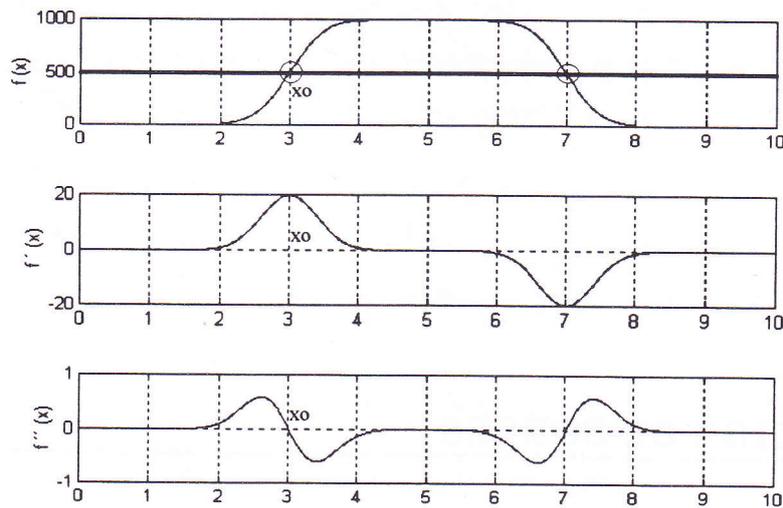


Figura 2.1 Función $f(x)$ y sus dos primeras derivadas.

Como se ve en la Figura 2.1, el punto x_0 puede ser localizado buscando el valor máximo y mínimo de la primera derivada, o encontrando los pasos por cero de la segunda derivada, donde esta última cambia de signo. Además de proporcionar información sobre la localización del punto-borde, la amplitud y el signo de la primera derivada nos indica la fuerza y la dirección del borde; esto es, si el cambio ha sido suave o brusco, y si ese cambio ha sido producido por aumento o disminución de la intensidad de la imagen.

Si el $|f'(x)|$ es muy grande, $f(x)$ está cambiando muy rápidamente lo que indica que ha habido un cambio brusco en la intensidad de la imagen. Si $f''(x)$ es

positiva, la intensidad de la imagen está aumentando, por el contrario si es negativa está disminuyendo.

Teniendo en cuenta las afirmaciones anteriores, se puede aproximar un primer detector de bordes basado en la primera derivada como el mostrado en la Figura 2.2.

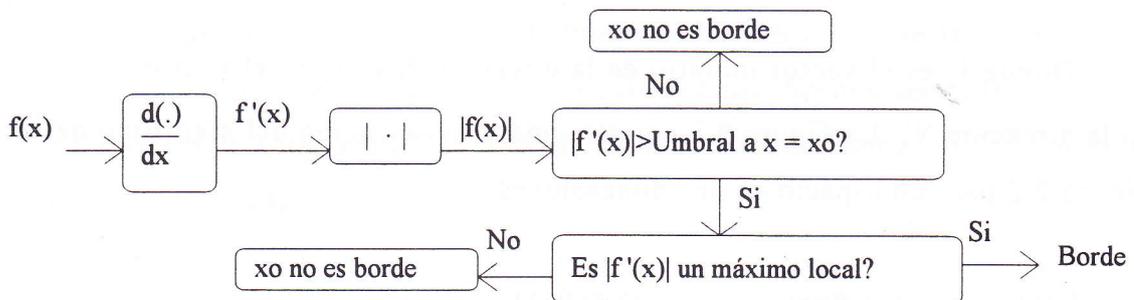


Figura 2.2 Detector de bordes con la primera derivada en una dimensión.

El algoritmo anterior calcula $|f'(x)|$ a partir de $f(x)$. Todos los valores de x cuyo $|f'(x)|$ sea mayor que un cierto umbral son puntos candidatos a borde. Si se representan los puntos que cumplen la condición anterior se obtendrá una línea en vez de un punto como sería deseable. Para evitar este problema se deberá buscar el máximo de entre los puntos candidatos a borde, ese máximo será x_0 . $f'(x)$ proporciona una información adicional, se puede saber si $f(x)$ está creciendo o decreciendo en torno a x_0 .

La elección del umbral depende de la aplicación particular que se esté desarrollando. Si se hace aumentar el umbral, sólo los valores de x cuyo $f(x)$ cambia rápidamente serán registrados como puntos candidatos a punto-borde. Como es difícil elegir un umbral óptimo se emplea a menudo un proceso de prueba-error. Algunos sistemas utilizan un umbral adaptativo, para lo cual se debe conocer con relativa exactitud los objetos que se representan en la escena, (lim, 1990).

La generalización de estos conceptos a un espacio de dos dimensiones $f(x,y)$ es el gradiente $\nabla f(x,y)$ y se puede expresar con la siguiente ecuación matemática:

$$\nabla f(x,y) = \frac{\partial f(x,y)}{\partial x} \hat{i}_x + \frac{\partial f(x,y)}{\partial y} \hat{i}_y \quad (2.1)$$

Donde \hat{i}_x es el vector unitario en la dirección X e \hat{i}_y es el vector unitario en la dirección Y. La Figura 2.3 muestra una generalización del algoritmo de la Figura 2.2 para un espacio en dos dimensiones.

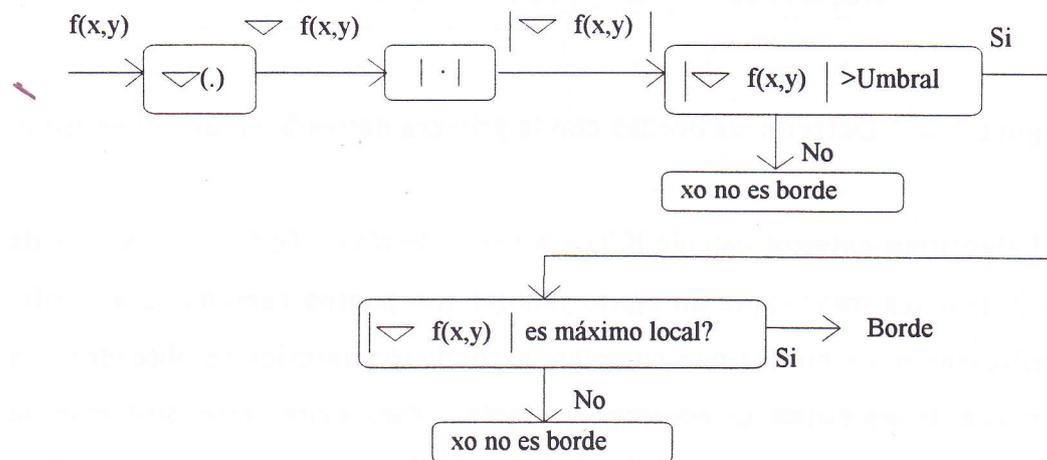


Figura 2.3 Detector de bordes con la primera derivada en dos dimensiones.

La magnitud $|\nabla f(x,y)|$ es calculada para comparar su valor absoluto con cierto umbral. El procedimiento es análogo al aplicado en el espacio unidimensional. En este caso en lugar de encontrar una línea de puntos candidatos a punto-borde, se encuentra una banda de puntos candidatos a línea-borde. Al proceso que consiste en estrechar esa banda hasta convertirla en una línea-borde, se denomina "adelgazamiento" o "afinamiento". Un algoritmo muy empleado consiste en encontrar el máximo local de $|\nabla f(x,y)|$ en una dirección como mínimo, (Rosenfeld, 1981).

En la mayoría de los casos es suficiente controlar los máximos locales en dos direcciones, la horizontal y la vertical. Si $|\nabla f(x,y)|$ es un máximo local a lo largo de alguna dirección dada, se dice que el punto determinado por (x,y) es un punto borde real.

La mayor dificultad que presenta el algoritmo de afinamiento consiste en la creación de un pequeño número de líneas de bordes falsos en la vecindad de la línea de borde real. Algunos autores denominan bordes secundarios a los bordes falsos porque entienden que de alguna manera contienen información adicional de la imagen, (Lim, 1990).

Segun Lim(1990), un método simple para eliminar parte de los bordes falsos es imponiendo unas restricciones previas:

- I. Si $|\nabla f(x,y)|$ tiene un máximo local en (x_0,y_0) , en la dirección horizontal X pero no en la dirección vertical Y, (x_0,y_0) será un punto-borde cuando:

$$\left| \frac{\partial f(x,y)}{\partial x} \right|_{\substack{x = x_0 \\ y = y_0}} > K \cdot \left| \frac{\partial f(x,y)}{\partial y} \right|_{\substack{x = x_0 \\ y = y_0}}, K \cong 2. \tag{2.2}$$

- II. Si $|\nabla f(x,y)|$ tiene un máximo local en (x_0,y_0) , en la dirección vertical Y pero no en la dirección vertical X, (x_0,y_0) será un punto-borde cuando:

$$\left| \frac{\partial f(x,y)}{\partial y} \right|_{\substack{x = x_0 \\ y = y_0}} > K \cdot \left| \frac{\partial f(x,y)}{\partial x} \right|_{\substack{x = x_0 \\ y = y_0}}, K \cong 2. \tag{2.3}$$

Cuando $|\nabla f(x,y)|$ tiene un máximo local en la dirección horizontal pero no en la vertical, la primera condición requiere que la rapidez de cambio a lo largo de la dirección horizontal sea significativamente mayor que en la dirección vertical. Por lo menos dos veces mayor. La segunda condición de la restricción es análoga a la primera cambiando las direcciones. Un sistema de detección de bordes basado en $|\nabla f(x,y)|$ es llamado "detector no direccional", dado que tales funciones, por trabajar con el valor absoluto, no tienen preferencia por ninguna dirección.

Si existe preferencia por alguna dirección, el sistema en cuestión se llamará "detector direccional" (Torre, 1986). Así, si en el sistema de la Figura 2.3 se reemplaza el gradiente $\nabla f(x,y)$ por la derivada direccional $\partial f(x,y)/\partial x$, el sistema detectará los bordes en esa dirección y no responderá a los cambios de intensidad en la otra dirección.

Para una secuencia de imágenes en dos dimensiones, las derivadas parciales $\partial f(x,y)/\partial x$ y $\partial f(x,y)/\partial y$ se suelen aproximar por alguna forma de diferenciación numérica. Por ejemplo, $\partial f(x,y)/\partial x$ se suele aproximar por una de las siguientes expresiones:

$$\begin{aligned} & [(f(n_1, n_2) - f(n_1 - 1, n_2)) / T \\ & [(f(n_1 + 1, n_2) - f(n_1, n_2)) / T \\ \text{o} & [(f(n_1 + 1, n_2) - f(n_1 - 1, n_2)) / 2T \end{aligned} \tag{2.4}$$

Siendo la distribución espacial:

$$\begin{bmatrix} f(n_1 - 1, n_2 - 1) & f(n_1, n_2 - 1) & f(n_1 + 1, n_2 - 1) \\ f(n_1 - 1, n_2) & f(n_1, n_2) & f(n_1 + 1, n_2) \\ f(n_1 - 1, n_2 + 1) & f(n_1, n_2 + 1) & f(n_1 + 1, n_2 + 1) \end{bmatrix}$$

Donde (n_1, n_2) denotan el pixel central en dos dimensiones. Como las derivadas así calculadas se comparan posteriormente con un umbral ajustable, los factores de escala $1/T$ y $1/2T$ pueden ser omitidos.

Es común encontrar librerías de software, como IP-LIB de Matrox Electronic System USA e IPLIB del National Engineering Laboratory UK, que combinan las expresiones de (2.4) para obtener mejor rendimiento y continuidad en la estimación de $\partial f(x, y) / \partial x$. Entre las combinaciones más comunes se encuentran las siguientes:

$$\begin{aligned} \frac{\partial f(x, y)}{\partial x} &\cong \\ &[(f(n_1 - 1, n_2 - 1) - f(n_1 + 1, n_2 - 1))] + \\ &[(f(n_1 + 1, n_2) - f(n_1 - 1, n_2))] + \\ &[(f(n_1 - 1, n_2 + 1) - f(n_1 + 1, n_2 + 1)]. \\ \frac{\partial f(x, y)}{\partial x} &\cong \\ &[(f(n_1 - 1, n_2 - 1) - f(n_1 + 1, n_2 - 1))] + \\ &2 \cdot [(f(n_1 + 1, n_2) - f(n_1 - 1, n_2))] + \\ &[(f(n_1 - 1, n_2 + 1) - f(n_1 + 1, n_2 + 1)]. \end{aligned} \tag{2.5}$$

Los factores de escala han sido omitidos. La operación de diferenciación en las ecuaciones (2.4) y (2.5) puede interpretarse como una convolución de $f(n_1, n_2)$ con la máscara $h(n_1, n_2)$ como filtro. Las máscaras más empleadas para estimar la primera derivada direccional se comentan brevemente a continuación.

La máscara (2.6.a) es un estimador de la derivada direccional en la dirección horizontal; (b), (c) y (d), son estimadores en las direcciones: vertical, 45°, y 135° grados sexagesimales.

$$\begin{matrix} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} & \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix} \\ \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)} \end{matrix}$$

(2.6)

Haciendo combinaciones no lineales de los términos empleados en el desarrollo de detectores de bordes direccionales se puede desarrollar detectores de bordes no direccionales, como el que se muestra en la ecuación (2.7).

$$|\nabla f(x,y)| = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}$$

(2.7)

Para una variable discreta en dos dimensiones la ecuación (2.7) se puede escribir:

$$|\nabla f(x,y)| = \sqrt{(f_x(n_1, n_2))^2 + (f_y(n_1, n_2))^2},$$

(2.8)

siendo, $f_x(n_1, n_2) = f(n_1, n_2) * h_x(n_1, n_2),$
y $f_y(n_1, n_2) = f(n_1, n_2) * h_y(n_1, n_2).$

2.1.1 Método de Sobel (1973)

Por aparecer en la mayoría de las librerías de software sobre tratamiento de imágenes, en (2.9) se muestra una aproximación conocida como método de Sobel (1973), la cual fue publicada por Duda y Hart en 1973. Es un "buen" detector de bordes, sin que ello quiera decir que localice éstos de manera eficiente.

Detección y localización son dos conceptos que se repetirán frecuentemente a lo largo de la presente memoria. Normalmente las funciones de visión que proporcionan las librerías dan como resultado un mapa de bordes cuyos valores van de 0 a 255, y dejan al usuario la posibilidad de establecer el umbral que decida cuales son bordes y cuales no.

$$|\nabla f(x,y)| = \sqrt{\left(\underset{\uparrow}{\underset{\uparrow}{\overbrace{f_x(n_1, n_2)}{f(n_1, n_2) * h_x(n_1, n_2)}}}} \right)^2 + \left(\underset{\uparrow}{\underset{\uparrow}{\overbrace{f_y(n_1, n_2)}{f(n_1, n_2) * h_y(n_1, n_2)}}}} \right)^2}$$

$$\begin{matrix} \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \end{matrix}$$

(2.9)

2.1.2 Método de Robert (1965)

Otro método muy empleado es el conocido como filtro de Robert (1965), el cual, al igual que Sobel, se basa en las expresiones (2.7) y (2.8), y utiliza para $h_x(n_1, n_2)$ y $h_y(n_1, n_2)$ máscaras cuya dimensión es dos por dos, a diferencia de Sobel que emplea máscaras de tres por tres. A continuación se muestran los filtros de Robert.

$$h_x(n_1, n_2) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad h_y(n_1, n_2) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

(2.10)

2.1.3 Otros Métodos con Gradiente

De la bibliografía técnica se conocen numerosas variantes de las expresiones presentadas en (2.8) y (2.9) que emplean diferentes combinaciones de gradientes, como por ejemplo los trabajos de Prewitt (1970), Arggle y Maclead (1972), Abdou(1973), Frei y Chen (1977), Navia (1980) y otros. Todas ellas son especialmente sensitivas al ruido con imágenes reales. La probabilidad de que aparezcan puntos-borde aislados es proporcional al filtro aplicado, al ruido de fondo, y a la cantidad de detalles muy finos de la imagen.

En la mayoría de aplicaciones es necesario someter a la imagen a un suavizado como procedimiento previo a la detección de bordes. También se aplican algoritmos para eliminar los puntos-borde aislados; solamente si el tiempo no es determinante en el procesamiento de la imagen, estos procedimientos son aplicables (Bovik, 1987).

A manera de ejemplo, en la Figura 2.4 se muestra una imagen real de poca definición y el resultado obtenido de aplicar el gradiente en dos direcciones ortogonales. Los bordes encontrados son bastante gruesos debido a la aplicación de un umbral demasiado bajo evitar discontinuidades en las líneas.

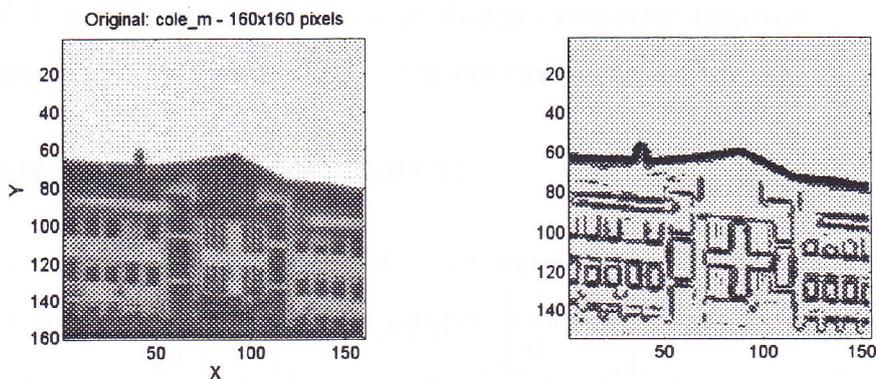


Figura 2.4 Imagen y su primera derivada

Normalmente los métodos basados en el gradiente generan discontinuidad en el contorno de los bordes reales obtenidos. Se pueden aplicar algoritmos que solucionen este problema elevando el tiempo de procesamiento de la imagen, con el consiguiente costo computacional.

2.2 Métodos Empleando el Laplaciano

El objetivo final de un algoritmo de detección de bordes es la localización correcta de las regiones de la imagen donde ocurren cambios bruscos de intensidad. En el caso de una función unidireccional $f(x)$, la búsqueda de los cambios rápidos se corresponden con la búsqueda de los máximos y mínimos en $f'(x)$.

Para los métodos basados en gradientes, $f'(x)$ es considerada suficientemente grande cuando la magnitud de su valor absoluto es mayor que cierto valor umbral. Otro camino para determinar ese máximo es considerando que $f'(x)$ es máxima-mínima cuando $f''(x)$ es cero (ver Figura 2.1). Si todos los pasos por cero son considerados como puntos-borde, el mapa de bordes estará formado por muchos puntos; esto puede verse con más claridad en la Figura 2.5.

Si no se guarda la información de $f'(x)$, cualquier pequeña variación en $f(x)$ es suficiente para generar un punto-borde. Debido a esta sensibilidad al ruido, la aplicación de un sistema reductor de ruido previo a la detección de bordes es muy aconsejable al procesar imágenes con ruido de fondo.

Por otro lado, la propia toma de información por medio de la cámara genera pequeños gradientes en la intensidad de imagen aunque la escena enfocada esté compuesta de un mismo color de intensidad uniforme.

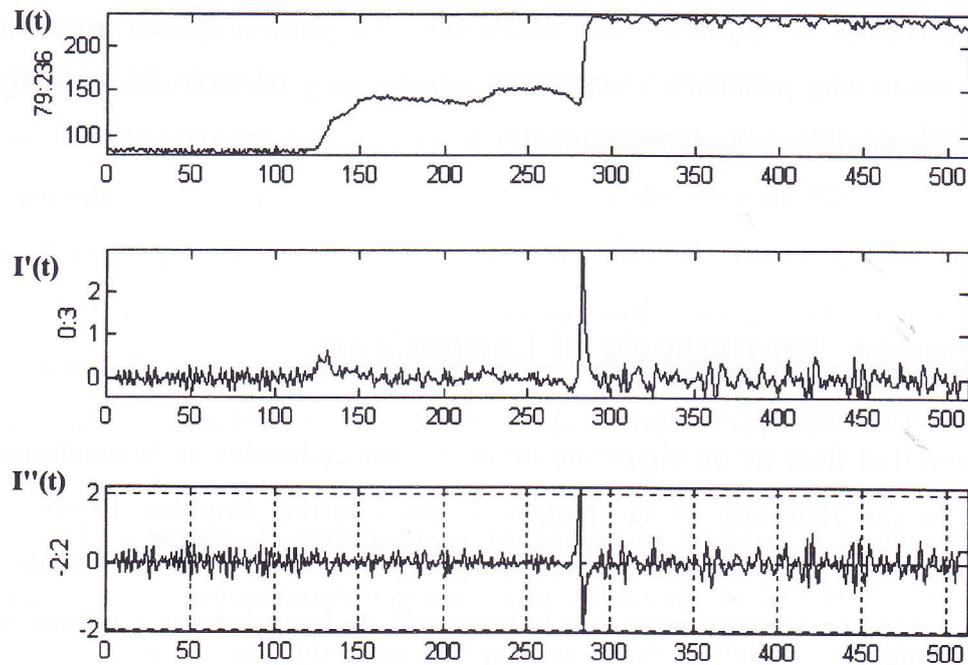


Figura 2.5 Una línea de vídeo de 512 pixels, y sus dos primeras derivadas.

Una generalización de $\partial^2 f(x)/\partial x^2$ a una función en dos dimensiones $f(x,y)$ es el Laplaciano $\nabla^2 f(x,y)$.

$$\nabla^2 f(x,y) = \nabla(\nabla f(x,y)) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \quad (2.11)$$

Para una secuencia $f(n_1, n_2)$, la segunda derivada parcial $\partial^2 f(x)/\partial x^2$ y la $\partial^2 f(x)/\partial y^2$ pueden reemplazarse por alguna forma de diferenciación discreta de segundo orden. Esta diferenciación puede ser visualizada como una operación de convolución de $[f(n_1, n_2) * h(n_1, n_2)]$.

A continuación se muestra que $[f(n_1, n_2) * h(n_1, n_2)]$ es una buena aproximación discreta de Laplaciano $\nabla^2 f(x,y)$.

Sean n_1 y n_2 las coordenadas en el plano de la máscara h , según la siguiente distribución discreta:

$$\begin{bmatrix} (n_1 - 1, n_2 - 1) & (n_1, n_2 - 1) & (n_1 + 1, n_2 - 1) \\ (n_1 - 1, n_2) & (n_1, n_2) & (n_1 + 1, n_2) \\ (n_1 - 1, n_2 + 1) & (n_1, n_2 + 1) & (n_1 + 1, n_2 + 1) \end{bmatrix}$$

(2.12)

Las referencias de dirección y sentido se indican en la Figura 2.6. Se ha tomado como punto de referencia el representado por la esquina superior izquierda de una imagen. La mayoría de los sistemas de representación gráfica, y por tanto, de imágenes, ubica el punto de referencia para sus coordenadas de la misma forma.

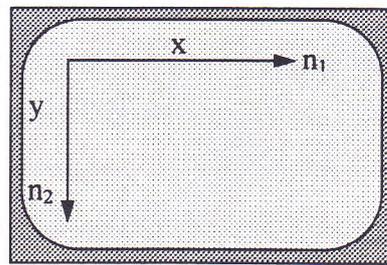


Figura 2.6 Referencias de los ejes.

Se puede escribir f_x empleando la diferencia hacia adelante y la diferencia hacia atrás según (2.13) y (2.14), respectivamente:

$$\frac{\partial f(x,y)}{\partial x} \Leftrightarrow f_x(n_1, n_2) = f(n_1 + 1, n_2) - f(n_1, n_2),$$

(2.13)

$$\frac{\partial f(x,y)}{\partial x} \Leftrightarrow f_x(n_1 - 1, n_2) = f(n_1, n_2) - f(n_1 - 1, n_2),$$

(2.14)

empleando (2.13) y (2.14) se puede aproximar la segunda derivada mostrada en la ecuación (2.15). De esta manera solamente se utiliza una máscara de tres por tres elementos:

$$\frac{\partial^2 f(x,y)}{\partial x^2} \Leftrightarrow f_{xx}(n_1, n_2) = f_x(n_1, n_2) - f_x(n_1 - 1, n_2), \quad (2.15)$$

reemplazando (2.13) y (2.14) en (2.15) y simplificando se obtiene:

$$\frac{\partial^2 f(x,y)}{\partial x^2} \Leftrightarrow f_{xx}(n_1, n_2) = f(n_1 + 1, n_2) - 2 \cdot f(n_1, n_2) + f(n_1 - 1, n_2), \quad (2.16)$$

de la misma forma, y por analogía en el razonamiento, se puede obtener las segunda derivada en el eje y:

$$\frac{\partial^2 f(x,y)}{\partial y^2} \Leftrightarrow f_{yy}(n_1, n_2) = f(n_1, n_2 + 1) - 2 \cdot f(n_1, n_2) + f(n_1, n_2 - 1), \quad (2.17)$$

sumando (2.16) y (2.17) se puede encontrar una expresión matemática para el Laplaciano de (2.11):

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} = f_{xx}(n_1, n_2) + f_{yy}(n_1, n_2), \quad (2.18)$$

reemplazando y reduciendo,

$$\nabla^2 f(x,y) = f(n_1 + 1, n_2) - 4 \cdot f(n_1, n_2) + f(n_1 - 1, n_2) + f(n_1, n_2 + 1) + f(n_1, n_2 - 1). \quad (2.19)$$

La expresión (2.19) puede escribirse en forma de un filtro bidimensional:

$$h(n_1, n_2) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (2.20)$$

El Laplaciano $\nabla^2 f(x,y)$ puede obtenerse fácilmente haciendo la convolución $[f(n_1, n_2) * h(n_1, n_2)]$. Dependiendo de como sea la aproximación discreta de las derivadas parciales, se puede obtener variaciones de (2.20).

Los puntos de paso por cero tienden a aparecer por el contorno de regiones relativamente homogéneas en intensidad. Por lo tanto, al convolucionar con el Laplaciano la tendencia será a formar líneas continuas, lo que hará innecesaria la operación de "afinamiento", imprescindible con el método del gradiente.

Sin embargo, como se puede apreciar en la Figura 2.7 este método genera un gran número de puntos-borde que constituyen bordes falsos, y que aparecen donde la varianza local de la intensidad de la imagen es pequeña.

Aún cuando el fondo de imagen sea uniforme, y por lo tanto el valor del Laplaciano cero dentro de esta región, cualquier perturbación debida al ruido probablemente generará un punto-borde falso.

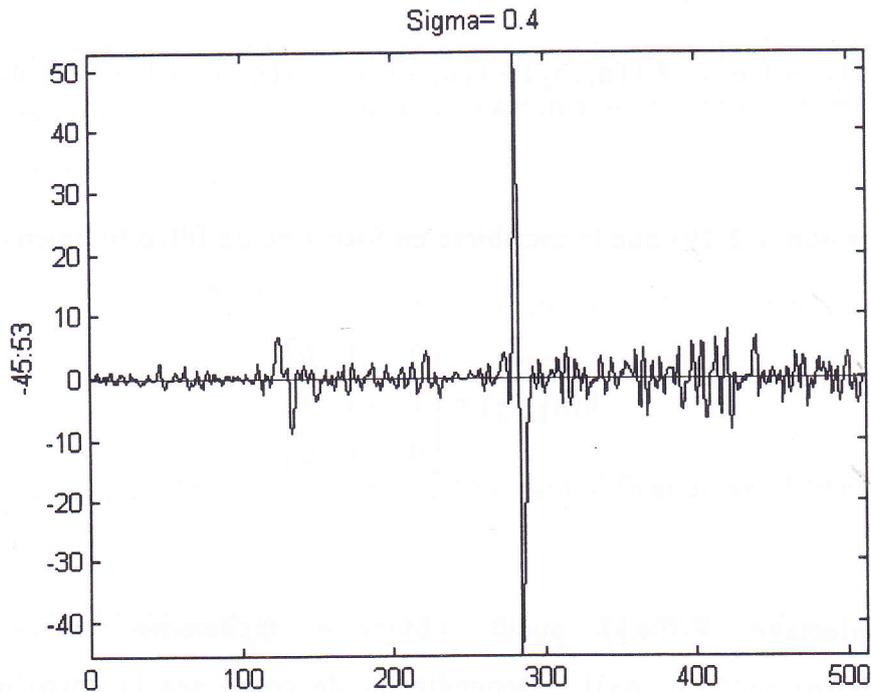


Figura 2.7 Detalle de la Figura 2.5 correspondiente a $I''(t)$.

Un método para eliminar muchos de esos falsos contornos es exigir que la varianza local sea suficientemente grande para que determinado paso por cero sea considerado punto-borde. La varianza local, para $f(x,y)$ cuadrada, se puede estimar según Lim (1990) de la siguiente manera :

$$\sigma_f^2(n_1, n_2) = \frac{1}{(2.M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} [f(k_1, k_2) - m_f(k_1, k_2)]^2 \quad (2.21)$$

siendo la media:

$$m_f(n_1, n_2) = \frac{1}{(2.M+1)^2} \sum_{k_1=n_1-M}^{n_1+M} \sum_{k_2=n_2-M}^{n_2+M} [f(k_1, k_2)] \quad (2.22)$$

Típicamente se emplea $M=2$.

Dado que $\sigma_f^2(n_1, n_2)$ debe ser comparada con cierto umbral, el factor de escala $1/(2.M+1)^2$ en (2.21) puede omitirse. La varianza local σ_f^2 solo se debe calcular para todo (n_1, n_2) en los que $\nabla^2 f(n_1, n_2) = 0$, o en los que existe un cambio de signo para dos puntos consecutivos en una misma dirección. Como se ve en la Figura 2.7, pueden ser muchos los puntos que cumplen esta condición. El algoritmo se puede expresar con el siguiente diagrama de flujo:

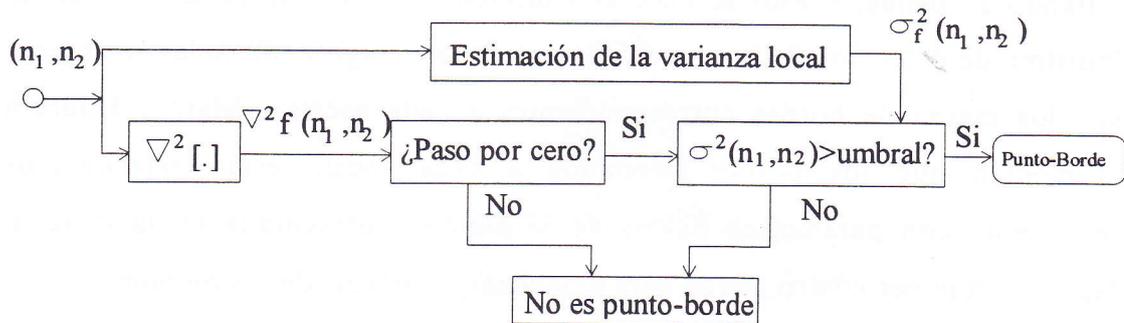


Figura 2.8 Algoritmo Laplaciano-Varianza para detectar bordes.

La varianza local $\sigma_f^2(n_1, n_2)$ está íntimamente relacionada con la magnitud del gradiente. El comparar la varianza $\sigma_f^2(n_1, n_2)$ con un cierto umbral es prácticamente lo mismo que hacer la comparación con el gradiente.

El requerimiento que el $\nabla^2 f(n_1, n_2)$ sea igual a cero puede interpretarse como la condición para realizar el proceso de "afinamiento" a partir del método del gradiente.

2.2.1 Método de Marr y Hildreth (1980)

En las secciones previas se ha discutido brevemente algunos algoritmos de detección de bordes que generan sendos mapas de bordes a partir de una imagen digitalizada. Marr y Hildreth (1980) observaron que ocurren cambios

significativos de intensidad a diferentes escalas dentro de una imagen, a esta característica se le puede equiparar con la resolución de una imagen.

Dentro de una imagen se presentan indistintamente, a diferentes escalas, zonas borrosas y zonas de variaciones finas. Por lo tanto, una detección de bordes "óptima" deberá detectar los cambios significativos de intensidad con operadores que respondan a diferentes escalas. Si la imagen original es filtrada limitando su banda, y esto se hace con diferentes filtros, se puede aplicar un algoritmo de detección de bordes a cada una de las imágenes filtradas, y obtener así, los mapas de bordes correspondientes a cada escala. Marr y Hildreth argumentan que los bordes obtenidos a cada escala están relacionados fuertemente con parámetros físicos de la escena representada en la imagen. Llegan a proponer cuatro filtros para modelizar el sistema de visión humano.

El mundo visual está formado por elementos tales como contornos, rayas aisladas y sombras, que deberían ser localizados a sus propias escalas. También influyen de manera importante los cambios físicos de reflectancia e iluminación. Si el mismo borde está presente en varios mapas de borde, obtenidos a distintas escalas, este borde es debido a un mismo fenómeno físico. Si por el contrario, el borde aparece en un mapa y no en los demás, puede deberse a que en la región donde aparece dicho borde inciden más de un fenómeno físico.

Para limitar el ancho de banda de la imagen, el filtro propuesto por Marr y Hildreth es el Gausiano (Figura 2.9). Al aumentar el valor de la desviación σ disminuye la frecuencia de corte. El Gausiano es expresado en el tiempo y en la frecuencia de la siguiente manera:

$$\text{tiempo: } G(x,y) = e^{\left(-\frac{x^2+y^2}{2\pi\sigma^2} \right)}, \quad (2.23)$$

frecuencia:

$$H(\Omega_x, \Omega_y) = 2\pi^2\sigma^2 e^{-\frac{\pi\sigma^2(\Omega_x^2 + \Omega_y^2)}{2}}$$

(2.24)

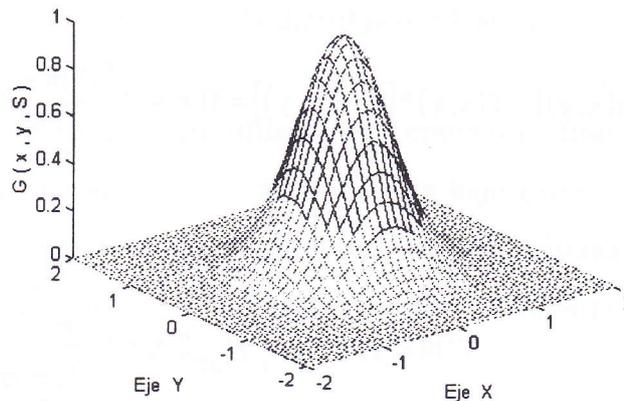


Figura 2.9 Representación de un gaussiano en 2D.

La razón por la que se elige un filtro gaussiano es porque éste se comporta suavemente tanto en el dominio del tiempo como en el de la frecuencia, lo que disminuye la probabilidad de que introduzca algún cambio que no haya estado presente en la imagen original. Asimismo, el Gaussiano es la única función que se localiza de modo óptimo simultáneamente en ambos dominios; característica que evita el desplazamiento de los bordes, o por lo menos, disminuye notablemente esa posibilidad. $\nabla^2 G(x, y)$, es un operador con simetría circular cuya distribución en dos dimensiones puede expresarse en términos de la distancia radial a partir del origen mediante la fórmula siguiente:

$$\nabla^2 G(r) = \left(\frac{-1}{\pi\sigma^4}\right) \left(\frac{1-r^2}{2\sigma^2}\right) e^{\left(\frac{-r^2}{2\sigma^2}\right)}$$

(2.25)

Marr y Hildreth, considerando que el cálculo del Laplaciano y la convolución son operaciones lineales, combinaron ambos operadores para encontrar un operador resultante que derive y filtre al mismo tiempo. Para la notación matemática $(G*I)$, su Laplaciano se representa mediante $\nabla^2(G*I)$ y la identidad matemática permite trasladar el operador ∇^2 dentro de la convolución, así, la expresión queda finalmente: $\nabla^2(G*I) = (\nabla^2 G)*I$, (Figura 2.10). Este razonamiento se puede expresar como sigue:

$$\nabla^2(f(x,y)*h(x,y)) = f(x,y)*[\nabla^2 h(x,y)] = f(x,y)*\left[\frac{\partial^2 h(x,y)}{\partial x^2} + \frac{\partial^2 h(x,y)}{\partial y^2}\right], \quad (2.26)$$

para el tiempo:
$$\nabla^2 h(x,y) = (x^2 + y^2 - 2 \cdot \pi \cdot \sigma^2) \frac{e^{-\frac{(x^2+y^2)}{2 \cdot \pi \cdot \sigma^2}}}{(2 \cdot \pi \cdot \sigma^2)^2} \quad (2.27)$$

para la frecuencia:
$$F[\nabla^2 h(x,y)] = -2 \cdot \pi^2 \cdot \sigma^2 \cdot e^{-\frac{\pi \sigma^2 (\Omega_x^2 + \Omega_y^2)}{2}} \cdot (\Omega_x^2 + \Omega_y^2) \quad (2.28)$$

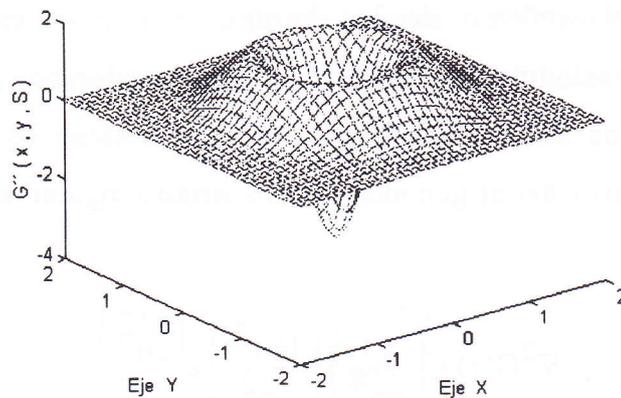


Figura 2.10 Segunda derivada del Gausiano en 2D.

Limitando el ancho de banda se puede reducir la influencia del ruido, y por lo tanto, disminuir el número de los bordes indeseados que aparecen en la Figura 2.7 . En la Figura 2.10 se puede ver la representación gráfica de la segunda derivada de un Gausiano en dos dimensiones con desviación $\sigma=0.5$.

Para elegir un filtro adecuado ∇^2G , se tendrá en cuenta que su parte Gausiana difumina la imagen suprimiendo toda estructura en escalas muy inferiores a la constante espacial de la distribución Gausiana.

Marr y Hildreth proponen utilizar un operador diferencial isotrópico independiente de la orientación, siendo el de más bajo orden el Laplaciano ∇^2 . En general, las imágenes satisfacen localmente las condiciones planteadas, y en la práctica es posible hacer uso de este operador. Además, el operador ∇^2 proporciona gran economía de cálculo.

Marr, Hildreth y Poggio (1980), basándose en los trabajos de Campell y Ropson (1968), Wilson y Giese (1977), Wilson y Bergen (1979), sugieren utilizar cuatro diferentes tamaños de máscaras para representar la información. Así, proponen un modelo de la observación humana, modelizándola como un sistema que analiza un campo visual después de haber sido tratado previamente por cuatro filtros ajustados a campos receptivos espaciales diferentes.

Los trabajos de detección de bordes de Marr y Hildreth no presentaron una argumentación matemática rigurosamente probada, esto dio pie a múltiples trabajos posteriores entre los que destaca notoriamente el de Canny (1986). Casi toda la bibliografía técnica de la época hace referencia a los trabajos publicados por Marr y Hildreth en 1980 con el nombre de "Teoría de Detección de Bordes". Los libros y artículos publicados en los últimos años parten casi siempre de esta teoría al comenzar sus referencias bibliográficas.

2.2.2 Algoritmo de Canny (1986)

Canny desarrolló una aproximación analítica para los operadores propuestos por Marr y Hildreth. El desarrollo de Canny está basado en un modelo unidimensional de un paso por borde, de amplitud h_e , más la suma de un ruido blanco gaussiano con una desviación estándar σ_n .

Asume que la detección del borde es realizada por convolución en el dominio continuo de una señal ruidosa $f(x)$ con una función antisimétrica $h(x)$, la cual es de amplitud cero fuera del rango $(-w, w)$.

Un punto es marcado como borde en el máximo local del gradiente obtenido al convolucionar $f(x)*h(x)$. La función $h(x)$ es elegida atendiendo a los siguientes criterios:

1. Buena Detección. La amplitud de la relación señal-ruido (snr) correspondiente al gradiente es maximizado para obtener una probabilidad de fallos baja tanto para la omisión de bordes como para la inclusión de bordes falsos. El snr para el modelo es:

$$\text{snr} = \frac{h_e}{\sigma_n} S(h), \quad S(h) = \frac{\int_0^w h(w) dx}{\int_{-w}^w [h(x)] dx} \quad (2.29)$$

2. Buena Localización. Los puntos-borde encontrados deben ser localizados tan cerca a los bordes reales como sea posible. Se propone el siguiente factor de localización:

$$\text{Loc} = \frac{h_e}{\sigma_n} L(h), \quad L(h) = \frac{h'(0)}{\int_{-w}^w [h'(x)]^2 dx} \quad (2.30)$$

3. Respuesta Única. Debería haber una sola respuesta para un determinado borde. La distancia entre picos del gradiente cuando esta presente sólo el ruido, es una fracción K del ancho del operador w , y se denomina x_m .

$$x_m = K \cdot w \quad (2.31)$$

Canny combinó estos tres criterios maximizando el producto $S(h) \cdot L(h)$, ajustándose a la constante x_m . Dada la complejidad de la formulación, Canny no encontró una solución analítica, pero llegó a la conclusión de que la solución se acerca a la primera derivada del Gausiano según crece x_m .

El estudio desarrollado por Canny desemboca en un resultado muy interesante desde el punto de vista de implementación y de propiedades. El filtro óptimo obtenido numéricamente puede aproximarse por la primera derivada de un Gausiano. La extensión al caso bidimensional es muy eficiente (Torre, 1986).

Considerando que los puntos que son borde son los máximos locales en la respuesta del operador a la imagen dada, los puntos que son borde para el caso n -dimensional mediante la aplicación de este filtro, son los máximos en la magnitud del gradiente de una imagen suavizada por un Gausiano. Para mejorar la detección de estos bordes, Canny (1986) propone la búsqueda de pasos por cero en la segunda derivada direccional.

Este método, sin embargo, presenta el inconveniente de necesitar una gran carga computacional para valores de σ mayores que uno, lo cual no facilita su

empleo en técnicas de visión artificial en tiempo real (Viles, 1994).

Por otro lado, las versiones de los operadores de tamaño grande definidos en el dominio continuo pueden ser obtenidas por muestreo simple de sus funciones impulso respuesta sobre una ventana de tamaño $W \times W$. El tamaño de la ventana se debe elegir lo suficientemente grande, de tal manera que el truncamiento de la función empleada no introduzca componentes artificiales de frecuencias altas. Evidentemente lo ideal sería no hacer ningún truncamiento, pero el costo computacional que resulta de convolucionar con ventanas grandes es elevado para una aplicación en tiempo real. Este es el principal problema a resolver.

2.3 Sumario de los Métodos Estudiados

Con las técnicas vistas hasta ahora, los gradientes son calculados en dos direcciones ortogonales, normalmente coincidiendo con filas y columnas, y la dirección del borde es deducida calculando el vector suma de los dos gradientes (Danielson, 1990). Otra aproximación es calcular los gradientes en un gran número de direcciones, convolucionando con los filtros correspondientes a esas direcciones (Lim, 1990).

También existen trabajos que enfocan el problema de detección de bordes con modelización de Señal, patrones de reconocimiento, y posprocesamientos morfológicos, a partir del conocimiento previo que se tenga de los objetos que aparecerán en la imagen. Un claro ejemplo es el trabajo de Abrams (1983) en el que propone un método para detectar los bordes de las arterias coronarias. Se considera que por el tiempo relativamente elevado que supone la implementación de estos métodos, su empleo en aplicaciones en tiempo real es inviable por ahora.

Se puede concluir, que para encontrar los bordes, o cambios de intensidad, se debe considerar que éstos ocurren a diferentes escalas, y que para detectarlos de modo óptimo, es preciso emplear operadores de diferentes tamaños. Se debe tener en cuenta que un cambio brusco de intensidad dará lugar a una punta o un valle en la primera derivada y un cero en la segunda derivada de la imagen. Estas dos ideas sugieren que se debe buscar un filtro que extraiga la primera o segunda derivada de la imagen, y que se pueda sintonizar para actuar a la escala deseada: con filtros grandes para detectar bordes difuminados, y pequeños para detectar bordes finos y bien enfocados. El filtro más eficiente encontrado es la primera derivada del Gausiano (Marr, 1986).

En la Figura 2.11 a se muestra un impulso de intensidad, en 2.11.b un filtro Gausiano, en 2.11.c la intensidad filtrada con dicho Gausiano, en 2.11.d la primera derivada, y en 2.11.e la segunda. Se puede observar que el máximo de la primera derivada coincide con un cambio en el valor de la intensidad, lo mismo que para los pasos por cero.

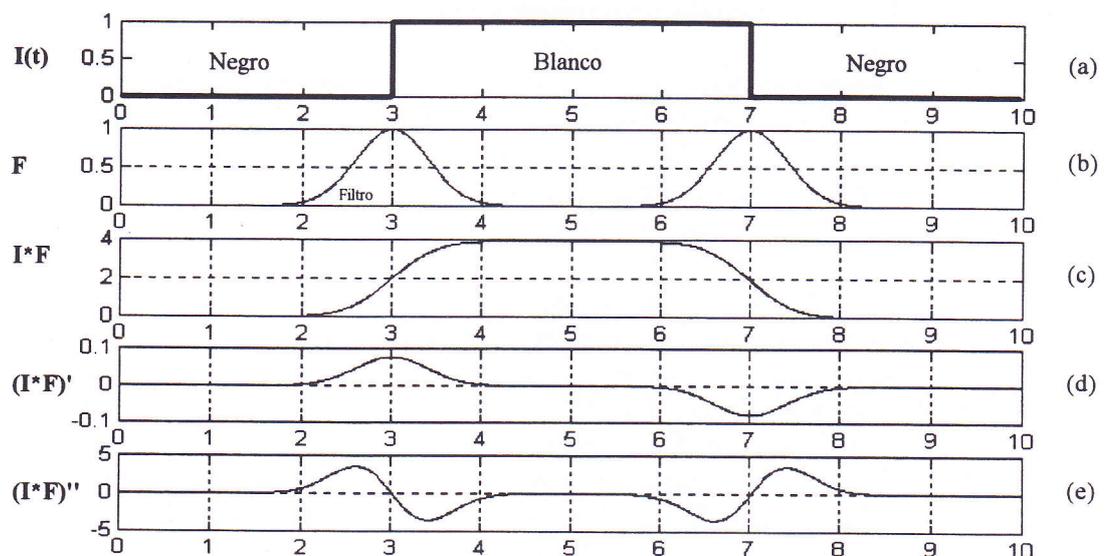


Figura 2.11 Detección y localización de dos bordes en una dimensión

En secciones posteriores se verá que cuando se trabaja con información digitalizada de imágenes reales, la limpieza de los bordes obtenidos no tiene casi nada que ver con la simulación presentada en la Figura 2.11. Esto es debido a la influencia del ruido y a que los niveles de luz constante en una imagen nunca generan niveles constantes de intensidad en la señal de la imagen de vídeo compuesto generada.

Por esta razón, en el siguiente capítulo se someterá al filtro Segunda Derivada del Gaussiano a un análisis para saber si es posible su implantación en hardware; y si lo es, averiguar que modificaciones hay que hacer para alterar lo menos posible su estructura. Se estudiará qué influencia tiene el tamaño del filtro, y cómo influye éste en su rendimiento. Se intentará aproximar sus coeficientes a valores enteros, y se analizarán los resultados obtenidos. Asimismo, se establecerán criterios para determinar umbrales de intensidad tanto para la primera como para la segunda derivada del gaussiano. Por último, se intentará estimar la segunda derivada del gaussiano con procedimientos alternativos para reducir el costo computacional ocasionado por operaciones de convolución con máscaras excesivamente grandes.