

Segmentación

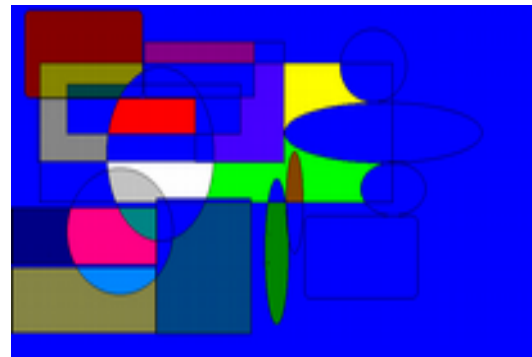
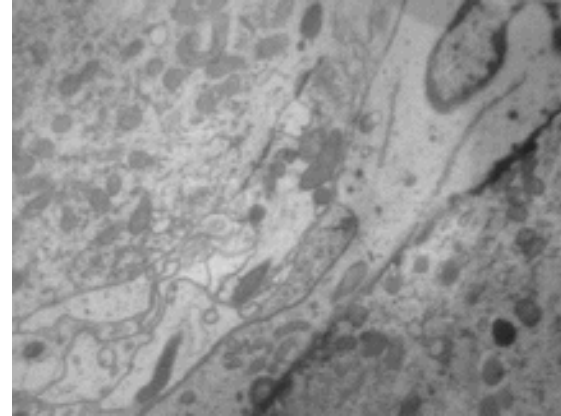
Qué vemos?

- La información visual llega a la retina de manera local y “discreta”: conos y bastones.
- Un buen “modelo” son las imágenes digitales: $f(x) \in \mathbb{R}$ definido en un dominio Ω
- Modelo más general: datos multidimensionales y vectoriales.

Pero “vemos” estructura...

- Los “datos” del ingeniero (o del sistema biológico) son muy distintos de las imágenes en sentido artístico, semántico o perceptual.
- Vemos estructura!

Ejemplos de imágenes



Pregunta

- Cómo pasar de los datos sin estructura a la imagen perceptual estructurada?
- Los psicólogos de la escuela de la Gestalt intentaron responder a esta pregunta en los años 20-30 del siglo pasado.

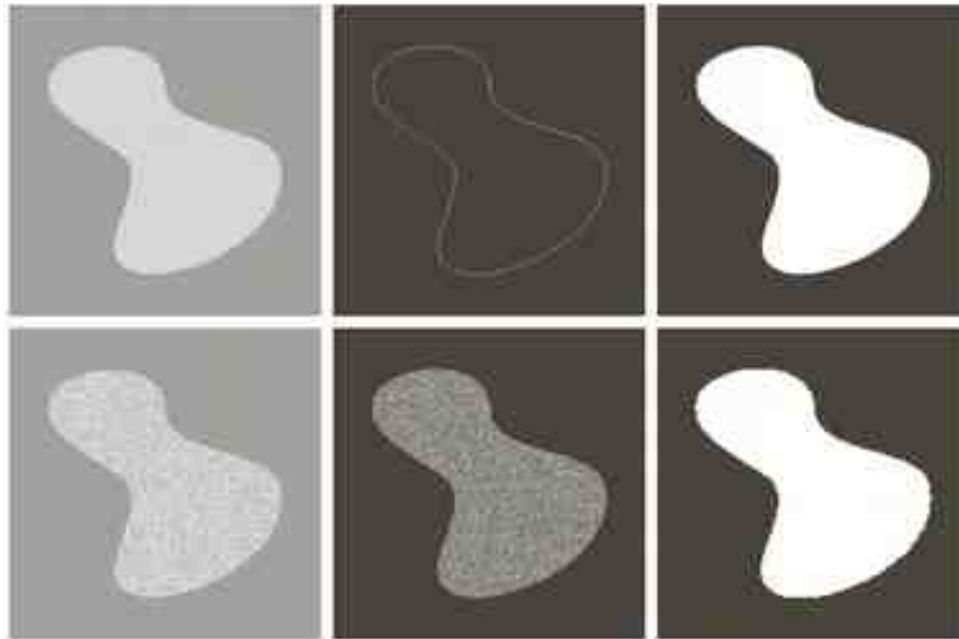
Marr

- Los psicofísicos de los 60 intentaron responder a esta pregunta: por ejemplo midieron cuánto era percibido en los primeros milisegundos luego de llegar la imagen a la retina.
- En los años 60-70 los ingenieros (y otros) intentaron el camino de la “visión artificial”, por ejemplo Marr

El problema de la segmentación

- Dada una imagen digital encontrar una partición razonable en regiones homogéneas y sus fronteras.
 - Partición razonable?
 - Qué propiedades?
 - Regiones homogéneas vs límites.
- Se supone que las regiones homogéneas son objetos perceptualmente significativos y los bordes sus límites.

Segmentación



a b c
d e f

FIGURE 10.1 (a) Image containing a region of constant intensity. (b) Image showing the boundary of the inner region, obtained from intensity discontinuities. (c) Result of segmenting the image into two regions. (d) Image containing a textured region. (e) Result of edge computations. Note the large number of small edges that are connected to the original boundary, making it difficult to find a unique boundary using only edge information. (f) Result of segmentation based on region properties.

Sería una máquina geométrica?

- Los gestaltistas y psicofísicos coinciden en que:
 - El proceso de segmentación funciona en las primeras etapas de la visión.
 - El proceso es independiente de modelos a priori o de procesos de aprendizaje: sería una máquina geométrica.

Un modelo general de segmentación

- Morel y Solimini en “Variational Methods in Image Segmentation” proponen que todos los algoritmos de segmentación corresponden a un mismo modelo general.
- Minimizar una “energía de segmentación”, esencialmente el funcional de Mumford y Shah

Energía de segmentación

$$E(u, K) = \int_{\Omega \setminus K} |(\nabla u)|^2 dx + \int_{\Omega \setminus K} (I - u)^2 dx + \text{long}(K)$$

Ω es el dominio y K es el conjunto de puntos de las fronteras

- En el dominio sin los bordes:
 - Regularidad: u es la aproximación por “trozos” de la imagen I .
 - Similitud: entre la imagen aproximada u y la imagen original I .
- En los bordes:
 - Bordes regulares: El conjunto de discontinuidad de “longitud” mínima.
- Minimizar la energía.

Dos aproximaciones...

- **Discontinuidad** de características:
 - líneas, bordes.
- **Regularidad** de características:
 - regiones.
- Son dos aspectos del mismo problema

Parámetros

- Buscamos que el algoritmo tenga “pocos” parámetros y que sean “intuitivos”.
- Dos parámetros “aceptables”:
 - Qué es borde y qué es ruido? -> un umbral
 - Escala: qué detalles significativos en cada nivel de resolución?

Análisis multiescala

- Dada una imagen $u_0(x)$ definida en un dominio Ω .
- Generar una secuencia de imágenes simplificadas $u_\lambda(x)$ cada vez más groseras conforme aumenta λ .
- Los detalles y bordes permanecen si su “escala” es mayor que λ

Propiedades que debe cumplir

- Fidelidad: $u_\lambda(x) \rightarrow u_0(x)$ cuando $\lambda \rightarrow 0$
- Causalidad: $s_\lambda(u_0)$ solo depende de $s_{\lambda^1}(u_0)$ si $\lambda > \lambda^1$
- Invariancia euclidiana: Si A es una isometría: $s_\lambda(u_0 \circ A) = s_\lambda(u_0) \circ A$
- Causalidad fuerte (para detección de bordes): $K_\lambda \subset K_{\lambda^1}$ si $\lambda > \lambda^1$

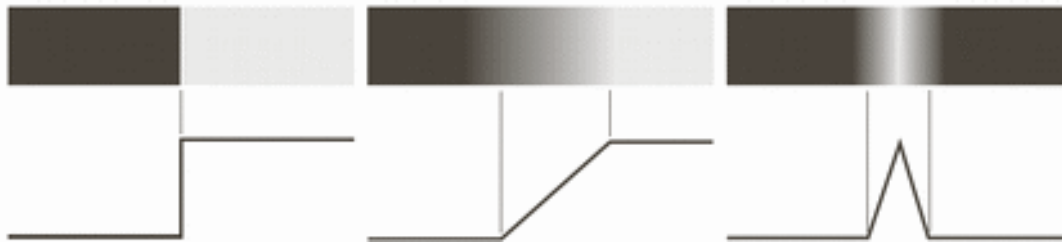
Aproximación lineal

- Hildreth, Marr, Witkin, Koenderink...
- Marr hablaba del primal sketch.
- Convolucionar la imagen con “detectores locales de borde”, detectar, filtrar y agrupar las respuestas.
 - Cambio de signo del laplaciano.
 - Extremas del gradiente.

Detección de bordes.

- Un borde definido como el límite entre dos regiones con *propiedades* diferentes.
- A veces se puede modelar como una discontinuidad en el nivel de gris (en este caso la propiedad de cada región está representada por el nivel de gris).
- La derivada local da cuenta de esas variaciones.

Modelos de borde



a b c

FIGURE 10.8
From left to right, models (ideal representations) of a step, a ramp, and a roof edge, and their corresponding intensity profiles.



Digital image processing, Gonzalez y Woods

Bordes en otras dimensiones...

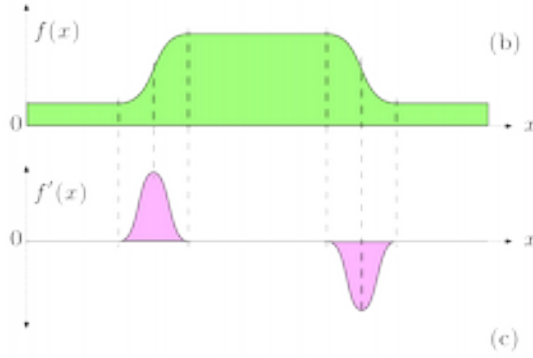
- En 3D: discontinuidades en todas las direcciones: superficies en vez de bordes.
- 2D+t: discontinuidades entre 2 imágenes sucesivas en el eje del tiempo por ejemplo.
- Etc...

bordes...

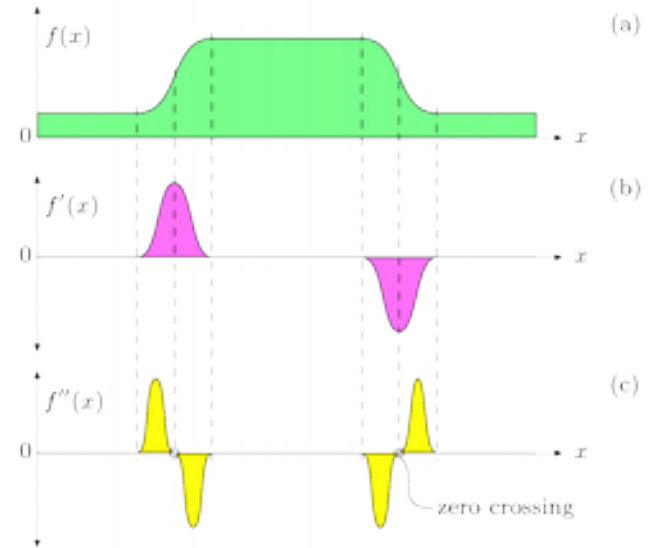
- La primera derivada indica la presencia de un borde.
- El signo de la 2da.derivada indica si un pixel pertenece a la parte “oscura” o “clara” del borde.
- ¿Donde está el borde?
- Ver <http://www.ipol.im/pub/art/2015/35/>



(a)



(c)



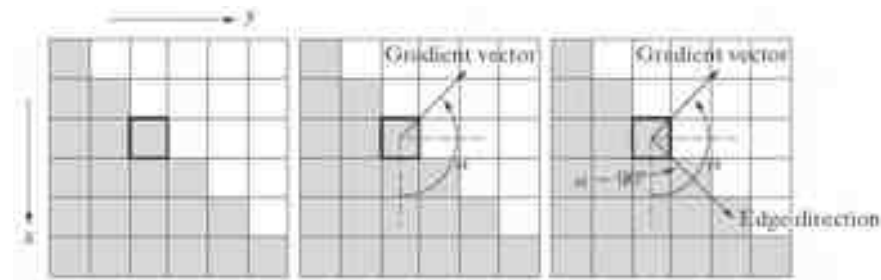
(a)

(b)

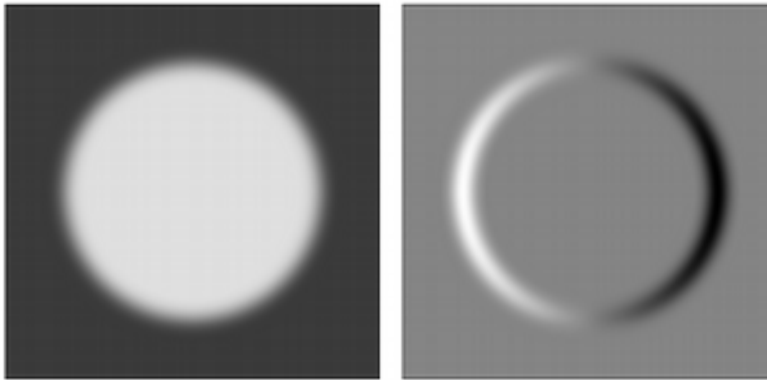
(c)

Gradiente

- Operador asociado a la **1ra. derivada**.
- El gradiente de una imagen está formado por un campo de vectores asociados a cada pixel. El módulo de cada vector da cuenta de la variación de intensidad y la dirección da cuenta de la dirección principal de esa variación.



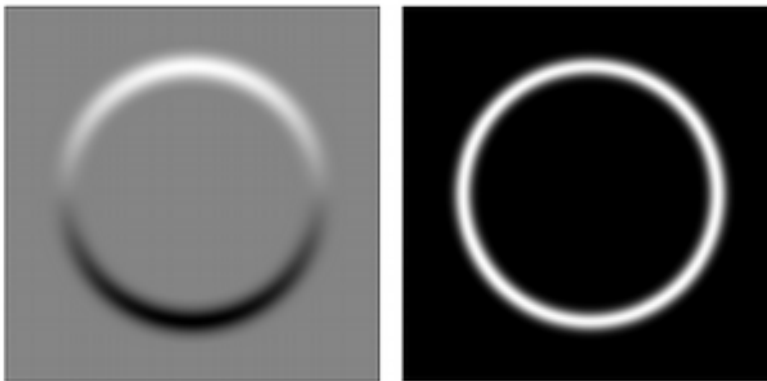
Derivadas parciales según la horizontal, según la vertical y módulo del gradiente.



(a)

(b)

$$\nabla I(x,y) = \begin{pmatrix} \frac{\partial I}{\partial x}(x,y) \\ \frac{\partial I}{\partial y}(x,y) \end{pmatrix}$$



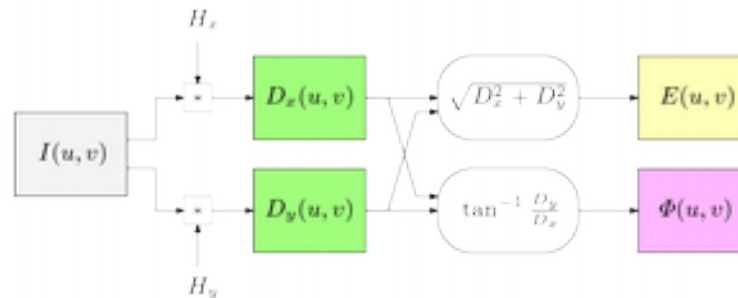
(c)

(d)

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

$$\Phi I(x,y) = \tan^{-1} \frac{\frac{\partial I}{\partial y}(x,y)}{\frac{\partial I}{\partial x}(x,y)}$$

Método general



Aproximar el operador

- Aproximación discreta: Sobel, Prewitt, etc.
- Filtro a respuesta impulsional finita (FIR):
Canny
- Filtro a respuesta impulsional infinita (IIR):
Deriche

Suavizado previo

- Las derivadas amplifican el ruido.
- La idea es filtrar previamente con un pasabajos.
- Filtros separables: se suaviza en una dirección y se deriva en la dirección perpendicular.

Máscaras.

- Se utilizan máscaras que recorren la imagen ponderando a los vecinos. El valor del pixel en la imagen de salida es la suma de los productos de los coeficientes de la máscara por el valor del pixel homólogo en la imagen de entrada.

máscara

w1	w2	w3
w4	w5	w6
w7	w8	w9

vecinaje 8

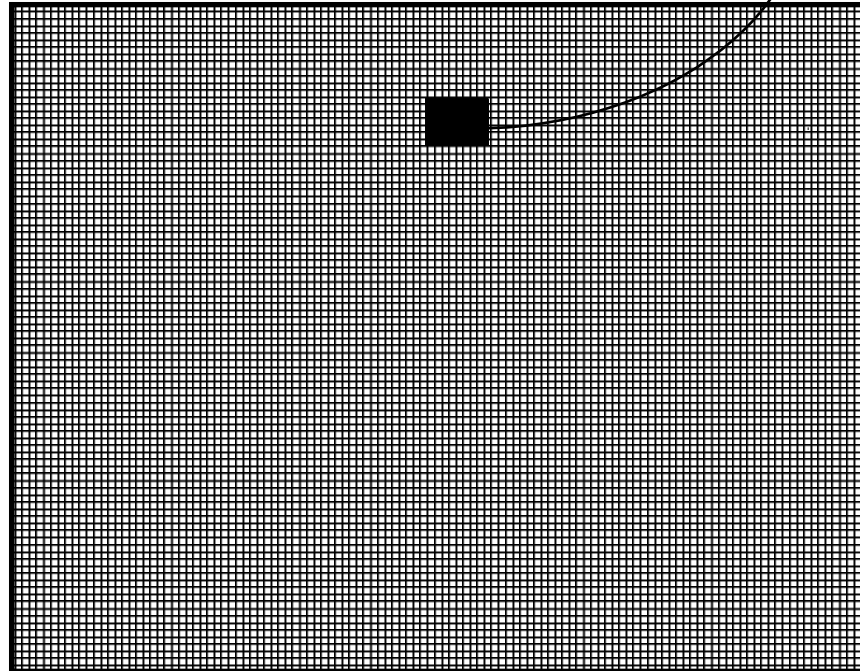
x1	x2	x3
x4	x5	x6
x7	x8	x9

$P(i,j) =$

\times

j

i



Expresado vectorialmente.

- Para una ventana de orden m , se trabaja con vectores de dimensión $n=2m+1$.
- Para cada pixel de la imagen de entrada se obtiene un valor en la salida según:

$$\sum_{i=0}^n \sum_{i=0}^n w_i x_i$$

Detección de puntos aislados.

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

- En zona homogénea = 0.
- Centrado en un punto aislado > 0 .
- Definición de un umbral.

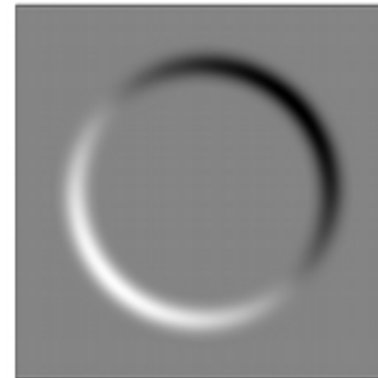
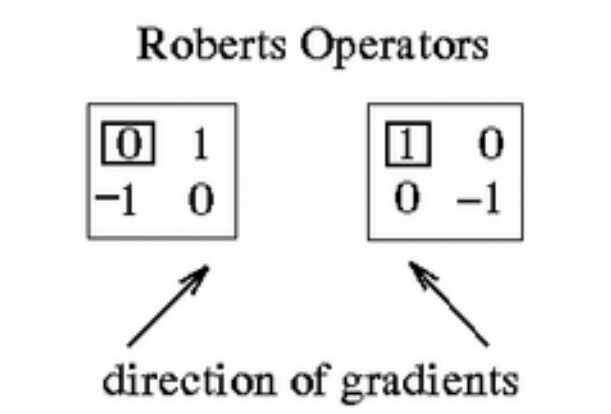
Gradiente

$$\nabla I(x, y) = \begin{pmatrix} \frac{\partial I}{\partial x}(x, y) \\ \frac{\partial I}{\partial y}(x, y) \end{pmatrix}$$

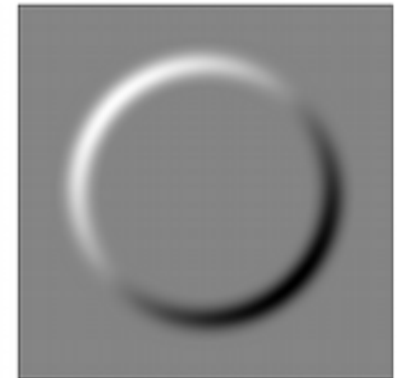
$$|\nabla I(x, y)| = \sqrt{\left(\frac{\partial I}{\partial x}(x, y)\right)^2 + \left(\frac{\partial I}{\partial y}(x, y)\right)^2}$$

$$\Phi I(x, y) = \tan^{-1} \left(\frac{\frac{\partial I}{\partial y}(x, y)}{\frac{\partial I}{\partial x}(x, y)} \right)$$

Operadores que implementan el gradiente



$$D_1 = I * H_1^R$$



$$D_2 = I * H_2^R$$

Prewitt

Máscaras para calcular **Gx** y **Gy**:

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

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

Isotrópico

Máscaras para calcular **Gx** y **Gy**:

-1	$-\sqrt{2}$	-1
0	0	0
1	$\sqrt{2}$	1

-1	0	1
$-\sqrt{2}$	0	$\sqrt{2}$
-1	0	1

Sobel

Máscaras para calcular **Gx** y **Gy**:

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

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

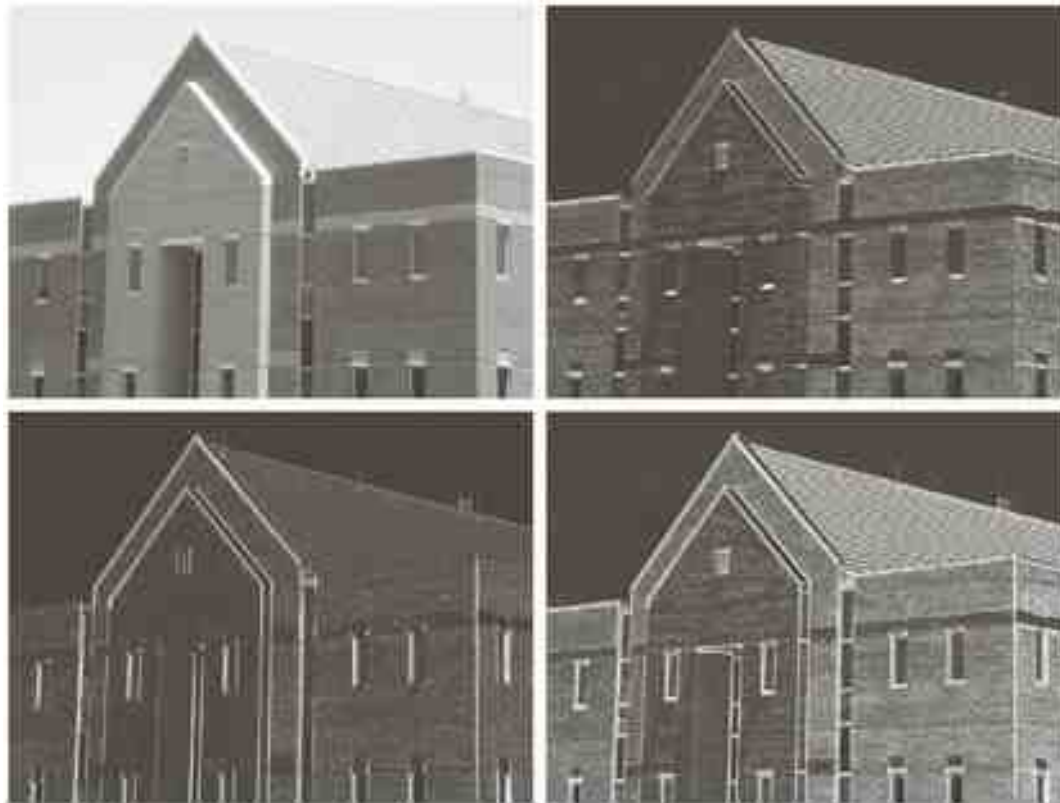
Pasabajos en un sentido y pasaaltos en el otro

- Ejemplo para Prewitt
- Filtro pasabajos promediador: Pasamos una máscara en cada fila $1/3(1 \ 1 \ 1)$
- Derivada por columnas $1/3 (-1 \ 0 \ 1)$
$$G_x = 1/2((x_7+x_8+x_9)/3 - (x_1+x_2+x_3)/3)$$
$$G_y = 1/2((x_3+x_6+x_9)/3 - (x_1+x_4+x_7)/3)$$

Sobel

- Filtro pasabajos más suave: Pasamos una máscara en cada fila $1/4(1 \ 2 \ 1)$
- Derivada por columnas $1/2 (-1 \ 0 \ 1)$
 $G_x = 1/2((x_7 + 2x_8 + x_9)/4 - (x_1 + 2x_2 + x_3)/4)$
 $G_y = 1/2((x_3 + 2x_6 + x_9)/4 - (x_1 + 2x_4 + x_7)/4)$

Sobel, resultado sobre una imagen



a. b.
c. d.

FIGURE 10.16
(a) Original image of size 834×1114 pixels, with intensity values scaled to the range $[0, 1]$.
(b) $|g_x|$, the component of the gradient in the x -direction, obtained using the Sobel mask in Fig. 10.14(f) to filter the image.
(c) $|g_y|$, obtained using the mask in Fig. 10.14(g).
(d) The gradient image, $|g_x| + |g_y|$.

Sobel, resultado sobre una imagen



FIGURE 10.17
Gradient angle
image computed
using
Eq. (10.2-11).
Areas of constant
intensity in this
image indicate
that the direction
of the gradient
vector is the same
at all the pixel
locations in those
regions.

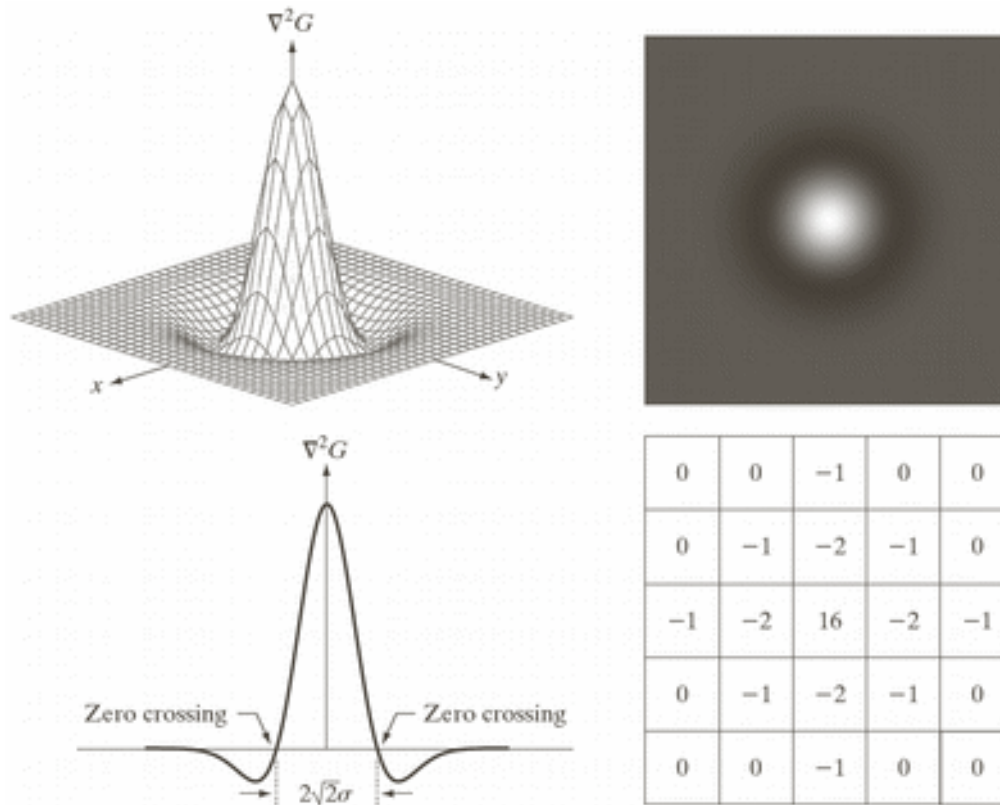
Marr-Hildreth

- Busca que el filtro sea sensible a la escala. El parámetro de la gaussiana permite aproximar la escala del detalle. Un operador que estima la 1ra. o 2da. derivada en cada punto de la imagen:
- Laplacian of Gaussian (LoG)

$$G(x, y) = \exp\left(\frac{-x^2 - y^2}{2\sigma^2}\right)$$

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

LoG



a b
c d

FIGURE 10.21
 (a) Three-dimensional plot of the *negative* of the LoG. (b) Negative of the LoG displayed as an image. (c) Cross section of (a) showing zero crossings. (d) 5×5 mask approximation to the shape in (a). The negative of this mask would be used in practice.

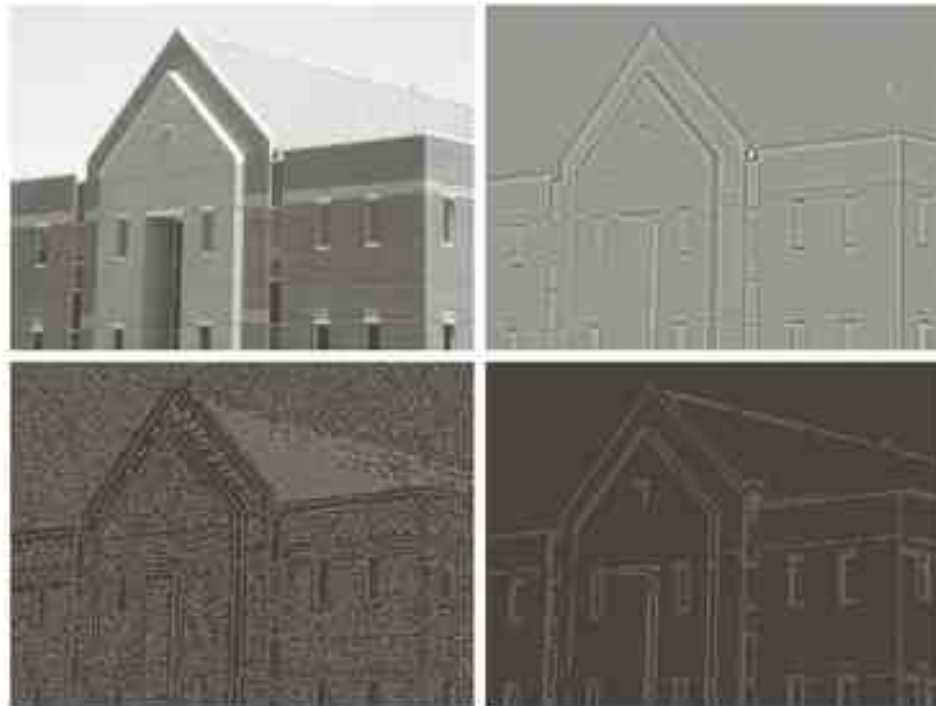
LoG

- La gaussiana filtra pasabajo para detalles menores que σ
- Los cruces por cero de la 2da derivada indican la presencia de cambios.
- Ambos son isotrópicos.
- Son operaciones lineales. Convolucionar la imagen con el LoG es como aplicar el laplaciano a la imagen I filtrada por una gaussiana G .

$$g(x,y)=[\nabla^2 G(x,y)]*I(x,y)$$

$$g(x,y)=\nabla^2[G(x,y)*I(x,y)]$$

LoG



a b
c d

FIGURE 10.22

(a) Original image of size 834×1114 pixels, with intensity values scaled to the range $[0, 1]$. (b) Results of Steps 1 and 2 of the Marr-Hildreth algorithm using $\sigma = 4$ and $n = 25$. (c) Zero crossings of (b) using a threshold of 0 (note the closed-loop edges). (d) Zero crossings found using a threshold equal to 4% of the maximum value of the image in (b). Note the thin edges.

Difference of Gaussians DoG

- Marr y Hildreth propusieron aproximar el filtro LoG mediante la diferencia de dos gaussianas de parámetros σ_1 y σ_2 tal que $\sigma_1 > \sigma_2$

$$\text{DoG}(x, y) = \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}}$$

Laplaciano

- Los bordes están en el cruce por cero del resultado de aplicar el operador.
- Curvas cerradas.
- No hay parámetro de cuan “fuerte” es el borde.
- Muy sensible al ruido.
- No da dirección.

Laplaciano

- La máscara más común:

$$\nabla^2 I(x, y) = 4z_5 - (z_2 + z_4 + z_6 + z_8)$$

0	1	0
1	-4	1
0	1	0

Qué hacer con el gradiente?

- Por el método que sea he aproximado en cada punto la dirección y el módulo del gradiente.
- Qué es ruido y qué información?
 - Relación con el soporte de la estimación del gradiente.
- Bordes cerrados? Bordes continuos?

Detectores de bordes: Canny.

- El modelo de borde es un escalón más ruido: $I(x) = Au_{-1}(x) + \eta(x)$
- Los bordes serían los máximos del resultado de la convolución del borde ruidoso con una función antisimétrica $f(x)$: $\Theta(x_0) = \int_{-\infty}^{\infty} I(x) f(x_0 - x) dx$
- Canny busca esa función $f(x)$ bajo tres premisas.

Criterios de Canny.

- Buena detección
- Buena localización
- Respuesta única

Criterios de Canny.

- **Buena detección:** Baja probabilidad de no detectar un punto de borde y baja probabilidad de marcar como borde un punto que no lo es. Maximizar la relación señal a ruido.

Buena detección.

- Llamemos $f(x)$ a la respuesta impulsional del filtro y llamemos $G(x)$ al borde centrado en $x = 0$. La respuesta del filtro al borde en su centro es:

$$H_G = \int_{-w}^w G(-x)f(x)dx$$

- Para la clase de borde *escalón*: $G(x) = Au_{-1}(x)$

- donde: $u_{-1} = \begin{cases} 0, & x < 0; \\ 1, & x \geq 0; \end{cases}$

Buena detección.

- Lo que resulta en: $H_G = \int_{-W}^0 Af(x)dx$
- La relación señal a ruido puede ser expresada como la relación entre la respuesta del filtro al borde y su respuesta al ruido:

$$SNR = \frac{A}{\eta_0} \frac{\int_{-W}^0 f(x)dx}{\sqrt{\int_{-W}^0 f^2(x)dx}} = \frac{A}{\eta_0} \Sigma$$

Buena localización.

- **Buena localización:** Los puntos marcados por el operador deben estar ubicados lo más cerca posible del centro del borde. El recíproco de la distancia media cuadrática del borde marcado respecto al centro del verdadero borde.

$$L = \frac{A}{\eta_0} \frac{|f'(0)|}{\sqrt{\int_{-w}^w f'^2(x) dx}} = \frac{A}{\eta_0} \lambda$$

Respuesta única.

- Respuesta única a un borde. El detector debe dar una sola respuesta a un borde único.
- Limitar el número de máximos en la ventana $[-W, W]$: la distancia media entre máximas adyacentes es el doble de los cruces por cero en la salida del operador.

Respuesta única.

- Utiliza una expresión del valor medio de la distancia entre cruces por cero de la respuesta de una función al ruido gaussiano para llegar a:

$$X = KW = 2\pi \sqrt{\frac{\int_{-\infty}^{\infty} f^2(x) dx}{\int_{-\infty}^{\infty} f'^2(x) dx}}$$

Propuesta de Canny

- Maximizar el producto $\Sigma\lambda$ bajo la exigencia de que el número de máximas locales sea fijo.

$$\Sigma(f) = \frac{|\int_{-w}^0 f(x) dx|}{\sqrt{\int_{-w}^w f^2(x) dx}}$$

$$\lambda(f') = \frac{|f'(0)|}{\sqrt{\int_{-w}^w f'^2(x) dx}}$$

Canny

- Obtuvo una forma teórica del filtro óptimo con valores de performance:

$$k = 0.58 \quad \Sigma\lambda = 1.12$$

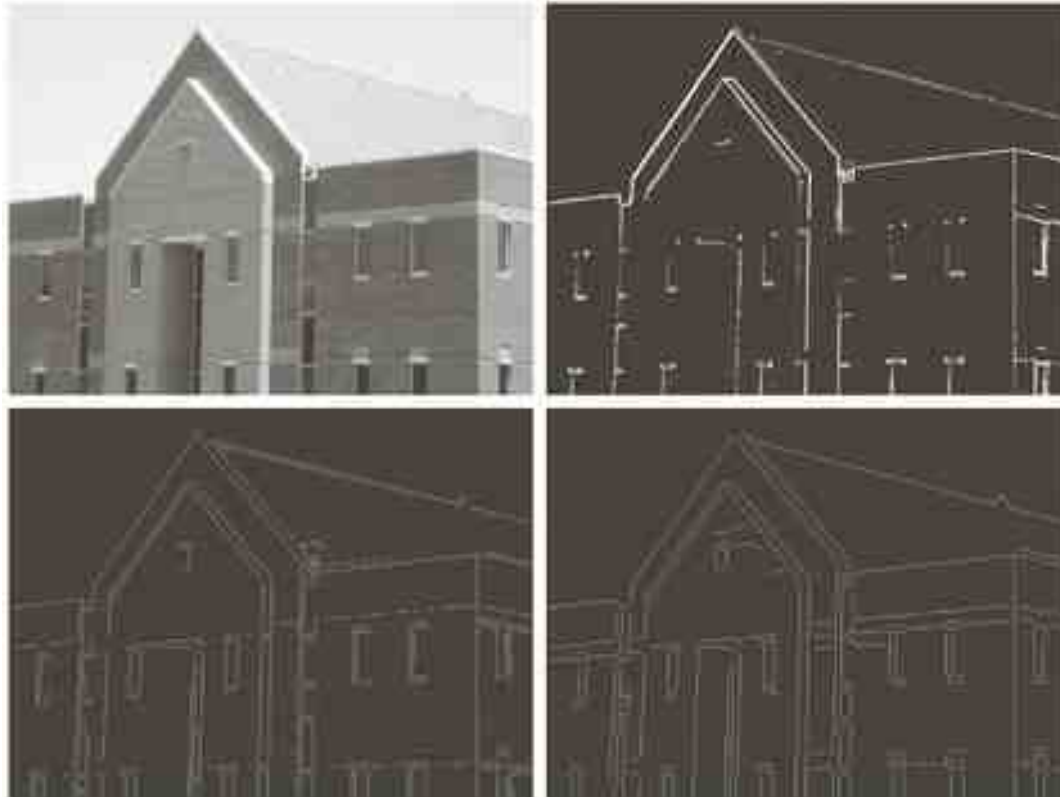
- Observando la forma de ese filtro notó su similitud con la forma de la primera derivada de una gaussiana:

$$f(x) = -\frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \quad k = 0.51 \quad \Sigma\lambda = 0.92$$

NMS: Non maxima suppression

- Marca los máximos la primera derivada de la gaussiana en la dirección perpendicular al borde.
- Umbralización con histéresis sobre el módulo del gradiente.

Canny



a b
c d

FIGURE 10.25

(a) Original image of size 834×1114 pixels, with intensity values scaled to the range $[0, 1]$.

(b) Thresholded gradient of smoothed image.

(c) Image obtained using the Marr-Hildreth algorithm.

(d) Image obtained using the Canny algorithm.

Note the significant improvement of the Canny image compared to the other two.

Deriche

- Canny trabajó con filtros a respuesta impulsional finita. Deriche extendió esa aproximación para filtros a respuesta impulsional infinita y obtuvo un filtro de la forma:
$$f(x) = -cxe^{-\alpha|x|}$$

- Sus índices de performance son:

$$\lambda = \sqrt{2\alpha} \quad \Sigma = \sqrt{\frac{2}{\beta}} \quad k = 0.44 \quad \Sigma\lambda = 2$$

Deriche

- Este operador es mejor que el de Canny y presenta un sólo parámetro α que controla el ancho del filtro. Su ajuste permite jugar con la localización o la SNR. El máximo de $f(x)$ está en $x = 1/\alpha$. Al aumentar α aumenta la localización y baja la SNR.

Deriche. Implementación 2D.

- Se trata de un filtro IIR separable con un número fijo de operaciones por punto independientemente de la talla del filtro.

Detección subpixelimétrica de bordes

- F. Devernay propuso un método para estimarlos:
 - Si $I(x,y)$ es un máximo local del gradiente
 - Interpolar una parábola entre $I(x,y)$ y sus dos vecinos en la dirección del gradiente.
 - El máximo de esa parábola es la posición del borde.
 - A Sub-Pixel Edge Detector: an Implementation of the Canny/Devernay Algorithm.
 - <http://www.ipol.im/pub/art/2017/216/>

Detección subpixelimétrica de bordes



Detección de bordes

- Se filtran las imágenes con filtros cualesquiera para obtener los gradientes G_x y G_y .
- Se calcula el módulo y la dirección del gradiente.
- Segmentación por 2 umbrales con histéresis.
- Supresión de no máximos locales.