

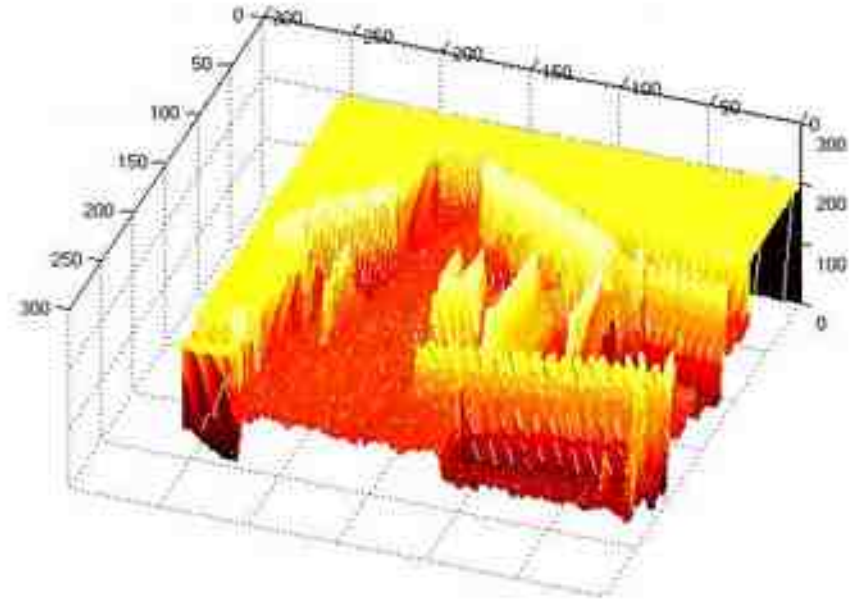
Tratamiento de Imágenes por Computadora

# Representaciones de imágenes digitales y algo más

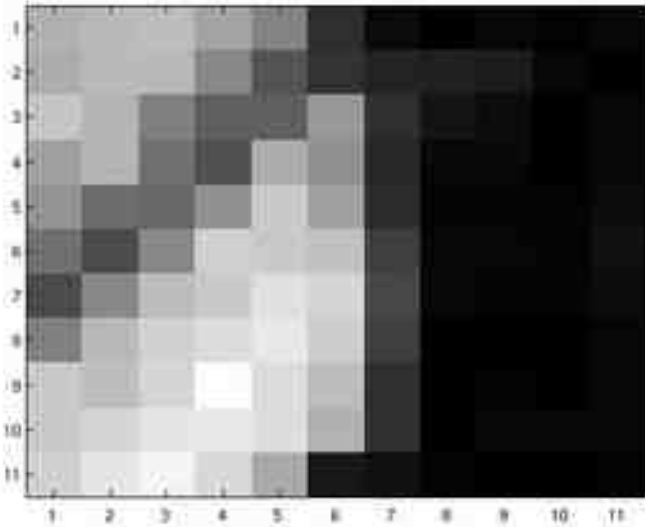
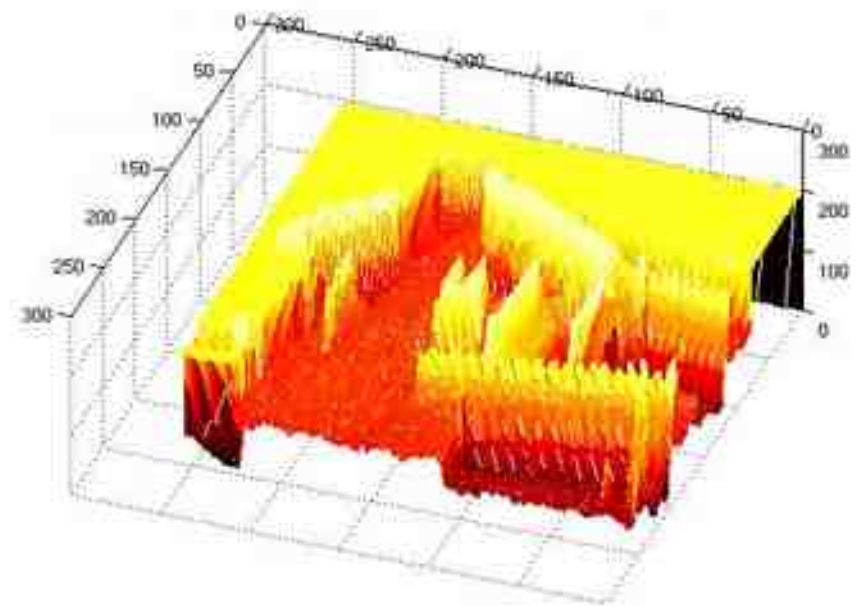
Facultad de Ingeniería, Universidad de la República

# Modelo de imagen

- Una imagen monocromática o en niveles de gris es una función  $u : \Omega \subset \mathbb{R}^2 \rightarrow S \subset \mathbb{R}$
- Imagen digital: discretización del soporte espacial y de los niveles de gris.  
Típicamente:
  - $\Omega$  grilla rectangular,  $\{1, \dots, M\} \times \{1, \dots, N\}$   
*pixels*
  - $S = \{0, \dots, 255\}$  (8 bits)



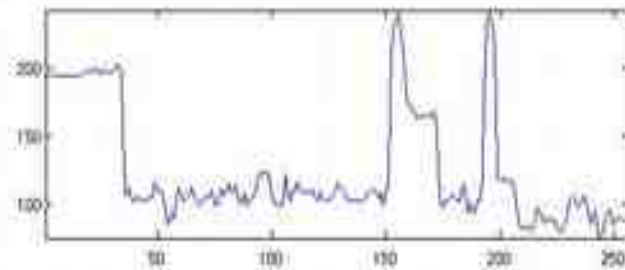
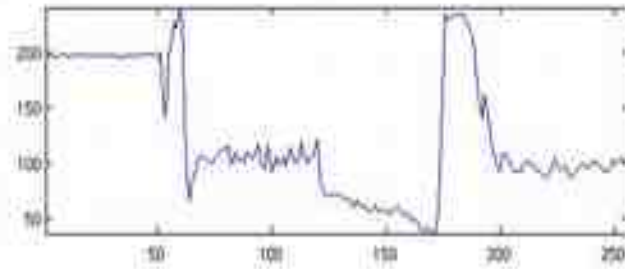
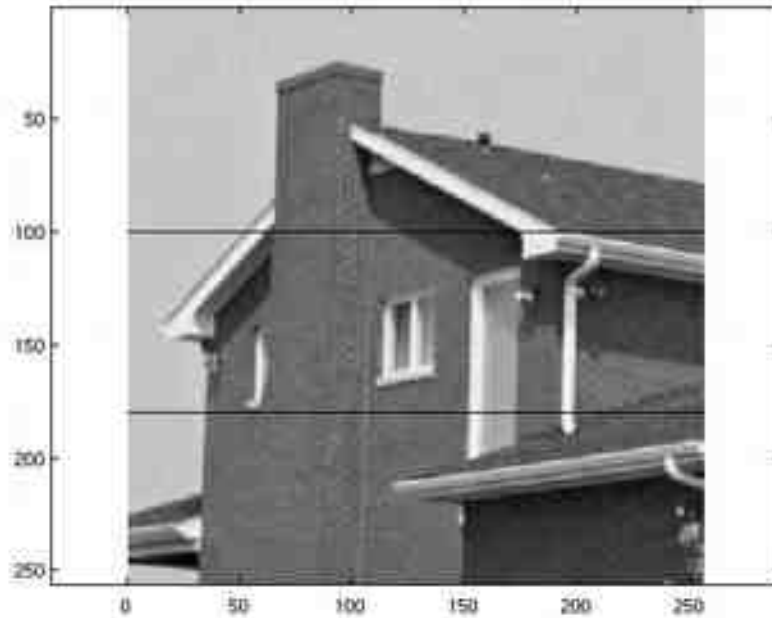
Sacado de "Digital Image Processing" Rafael Gonzalez & Richard Woods



199	204	207	193	173	124	103	96	100	98	100
198	205	204	177	145	128	117	115	112	100	95
212	202	172	153	153	186	125	107	104	97	100
191	202	162	143	197	182	120	101	100	97	102
183	161	157	182	215	190	122	99	99	98	104
163	140	177	219	215	210	133	102	100	99	105
140	176	206	215	229	222	139	102	98	98	103
172	204	219	227	234	216	133	97	96	96	101
215	208	222	246	225	206	125	95	98	97	101
215	221	231	232	225	201	124	97	102	101	101
220	231	240	224	196	111	105	98	95	96	99

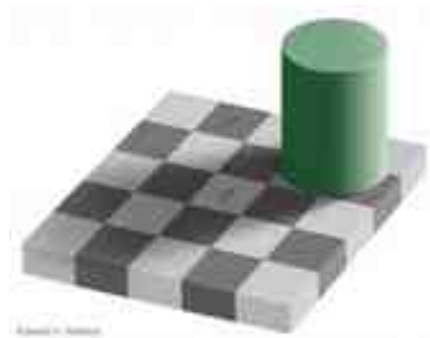
Sacado de "Digital Image Processing" Rafael Gonzalez & Richard Woods

# Perfiles



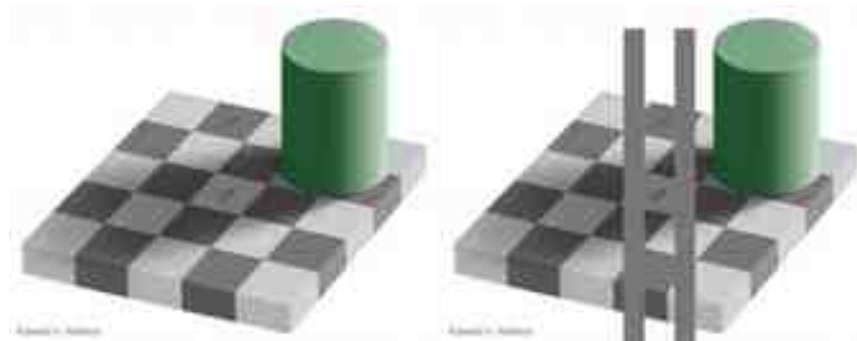
Sacado de "Digital Image Processing" Rafael Gonzalez & Richard Woods

# El modelo morfológico



[https://en.wikipedia.org/wiki/Checker\\_shadow\\_illusion](https://en.wikipedia.org/wiki/Checker_shadow_illusion)

# El modelo morfológico



[https://en.wikipedia.org/wiki/Checker\\_shadow\\_illusion](https://en.wikipedia.org/wiki/Checker_shadow_illusion)

- **Observaciones:**

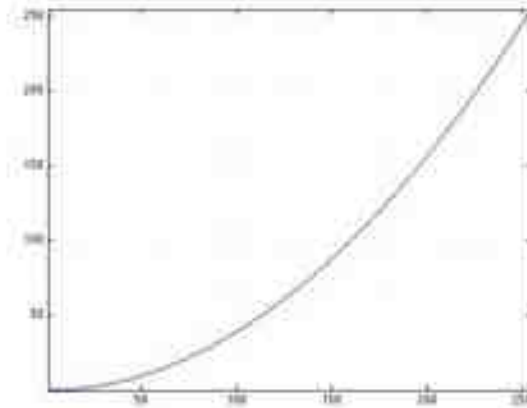
- La información visual que percibimos no está basada en el valor absoluto de la intensidad (nivel de gris en imágenes)
- La percepción de intensidades está dominada por el contraste local (escuela de la Gestalt, Kanisza)



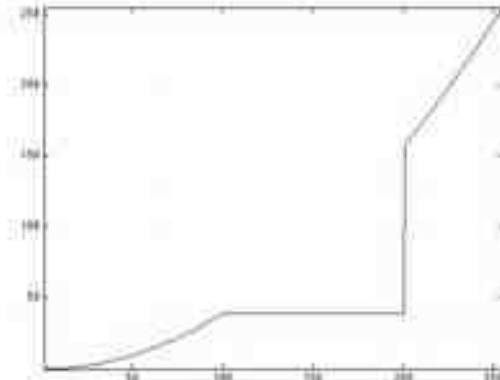
- **Conclusión:**
- El **análisis de imágenes** debe ser independiente de cualquier cambio *razonable* en la escala de grises.
- El principio de **invarianza por cambio de contraste** es un hipótesis central en *Morfología Matemática* (Matheron, Serra)

# Cambio de contraste.

- **Def:** un **cambio de contraste** es una función creciente  $g : IR \rightarrow IR$  que transforma la imagen  $u$  en la imagen  $z \circ u$ .

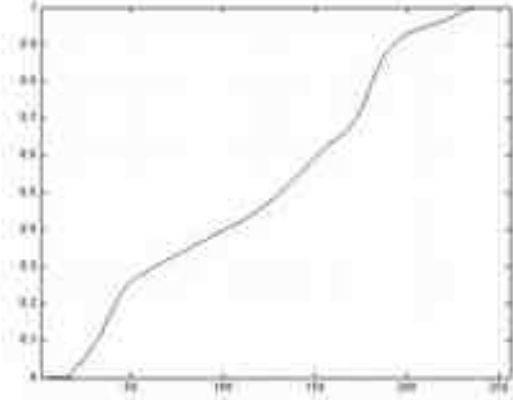
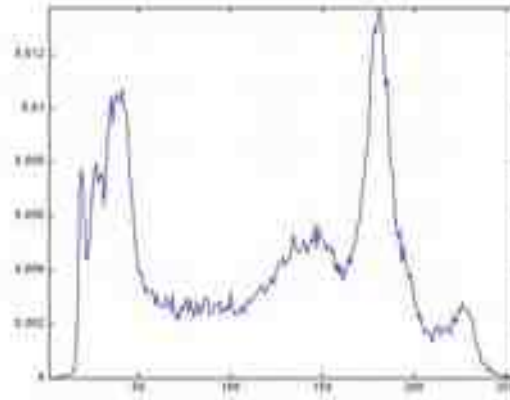


- $g$  estrictamente creciente  $\Rightarrow g$  invertible  
 $\Rightarrow$  Las imágenes  $u$  y  $g(u)$  contienen la “misma información”.
- Las imágenes que se corresponden por cambios de contraste estrictamente crecientes forman un clase de equivalencia.
- Si el crecimiento no es estricto, puede haber pérdida de información:



- **Def:** el histograma  $h_u(t)$  de  $u: \Omega \rightarrow IR$  es la función:

$$h_u(t) = \frac{\left\{ (x,y) \in \Omega : u(x,y) = \lambda_i \right\}}{\Omega}$$



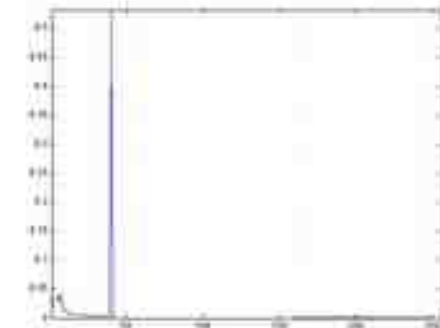
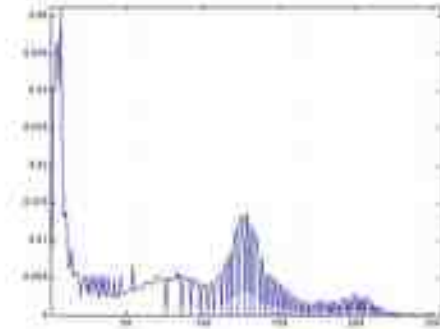
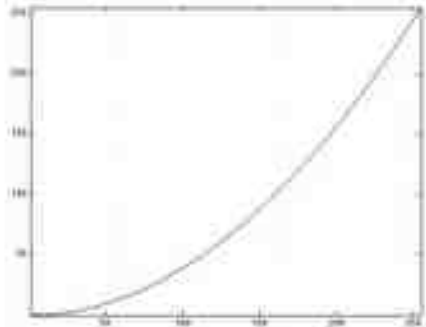
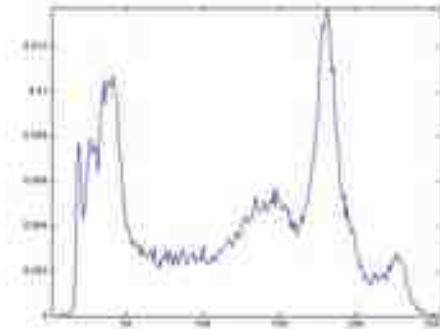
- **Def**: el **histograma acumulativo**  $H_u(t)$  de la imagen  $u : \Omega \rightarrow IR$  es la función creciente

$$H_u(t) = \frac{\#\{(x, y) \in \Omega : u(x, y) \leq t\}}{\#\Omega}$$

- **Obs**: en imágenes digitales  $H_u(t)$  no es continua (salvo si es constante)

- Observaciones:

- El histograma es una *descripción global* de la imagen
  - Qué dice de la imagen?
- Dos imágenes distintas pueden tener el mismo histograma:
  - Ejemplo trivial?
  - Algoritmo para generar imágenes con histograma dado?



# Procesamiento de histogramas

- Es posible definir otras funciones.
- Definición interactiva.
- Look Up Table (LUT):
  - Obtener analíticamente la función deseada y
  - Solución numérica rápida .



# LUT

- Una tabla con valores de una función dada.
- Accedemos al contenido de una celda de la tabla:  $v=f(i)$ .

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

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

# Ecuación de histograma

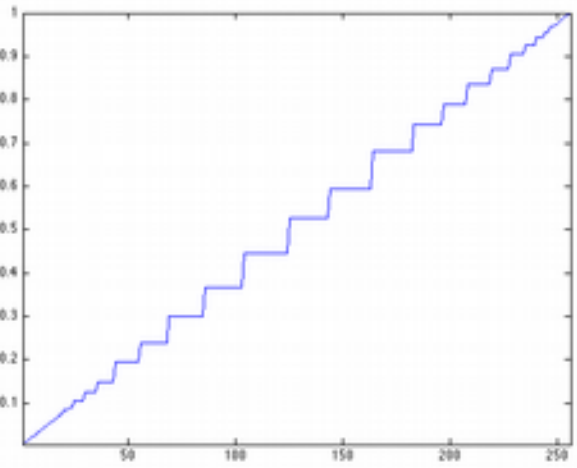
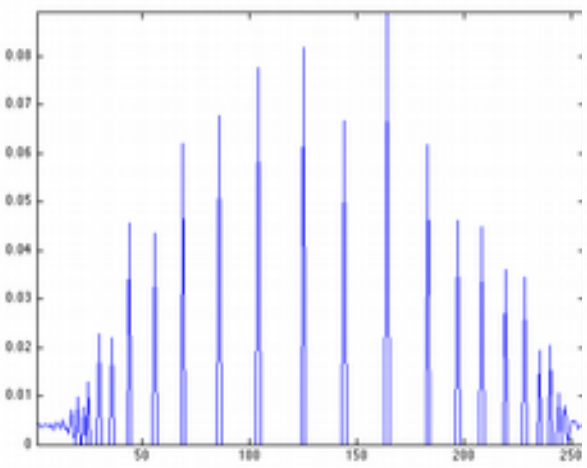
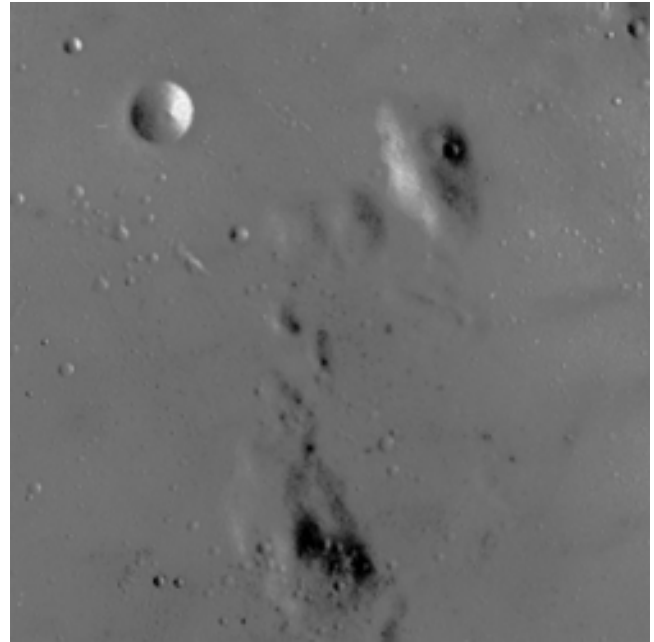
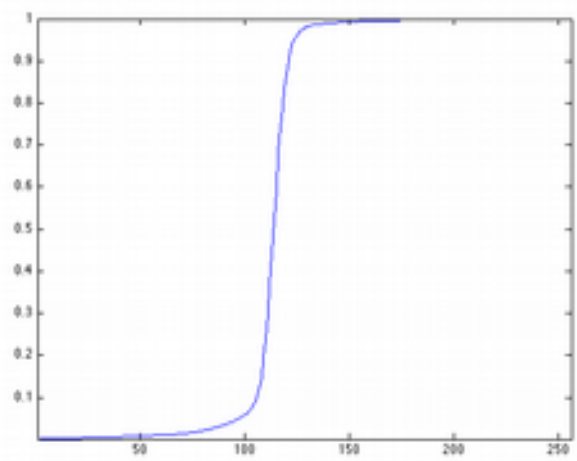
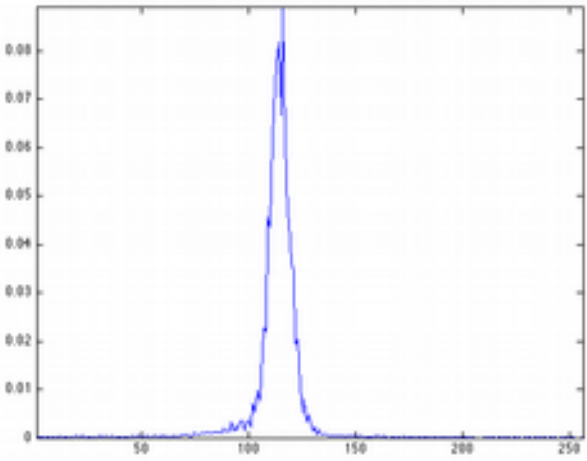
- **Objetivo:** expandir el rango dinámico de la imagen (corregir una foto sobre o sub-expuesta, ganar en detalle en imágenes radiológicas, etc)
- **Objetivo concreto:** aplicar un cambio de contraste a la imagen para que el histograma sea lo más constante posible.

- Caso general:

-  $\lambda_1 < \lambda_2 < \dots < \lambda_n$  niveles de grises de  $u$  en  $\Omega$

$$\Rightarrow g_k = [H_u(\lambda_k) + H_u(\lambda_{k-1})] / 2$$

- Utilizar el histograma acumulado y aproximar.
- Es automático.



Sacado de "Digital Image Processing" Rafael Gonzalez & Richard Woods

# Ecuación de histograma: desventajas

- **Es totalmente global.** La presencia de regiones muy claras o muy oscuras, no necesariamente muy grandes, influye en fuertemente en toda la imagen.
- **Puede hacer explotar el ruido** en zonas poco contrastadas de la imagen (problemas de cuantificación)
- Conservación de la **información geométrica?**

# Modelo morfológico: conjuntos de nivel

- **Afirmación informal:**

*Las imágenes que se corresponden por un cambio de contraste estrictamente creciente g contienen la “misma información”*

- Preguntas:

- De qué información hablamos?

- Qué se conserva al aplicar  $g$ ?

- Qué tienen en común esas imágenes?

# La descomposición en conjuntos de nivel

- **Def:** los conjuntos de nivel superiores de una imagen  $u : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  son los conjuntos

$$\mathcal{X}_\lambda(u) = \{(x, y) \in \mathbb{R}^2 : u(x, y) \geq \lambda\}$$

donde  $\lambda$  es un real cualquiera.

- **Independencia del nivel de gris absoluto:** para cualquier cambio de contraste creciente estricto  $g$ ,

$$\mathcal{X}_\lambda(u) = \mathcal{X}_{g(\lambda)}(g \circ u)$$



- Los conjuntos de nivel contienen toda la información geométrica de la imagen;  
**propiedad de reconstrucción:**

$$u(x, y) = \sup\{\lambda \in IR : (x, y) \in \chi_\lambda(u)\}$$

- Los conjuntos de nivel verifican **la propiedad de inclusión**

$$\forall \lambda, \mu t. q. \lambda \geq \mu, \chi_\lambda(u) \subseteq \chi_\mu(u)$$

# Líneas de nivel, Mapa Topográfico

- **Líneas de nivel**: fronteras topológicas de los componentes conexos de los conjuntos de nivel
- Las líneas de nivel son curvas cerradas (si no tocan el borde). Si la imagen  $u$  es  $C^1$ , una línea de nivel  $L$  cuyo gradiente no se anula es una curva de Jordan de clase  $C^1$
- **Mapa topográfico**: el conjunto de líneas de nivel

# Propiedades

- **Invariancia al cambio de contraste:** Dos imágenes a las que se les cambió el contraste mediante una función creciente estricta, tienen el mismo mapa topográfico.
- **Representación.** Estructura que cumple la propiedad de inclusión lo que habilita una representación mediante un árbol.

- Con unos pocos conjuntos de nivel, se puede lograr una comprensión suficiente de la imagen

original(100,000)

reconstrucción con 400 líneas



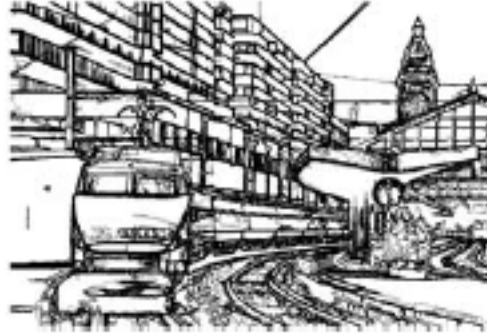
Original (84000)

reconstrucción con 850 líneas

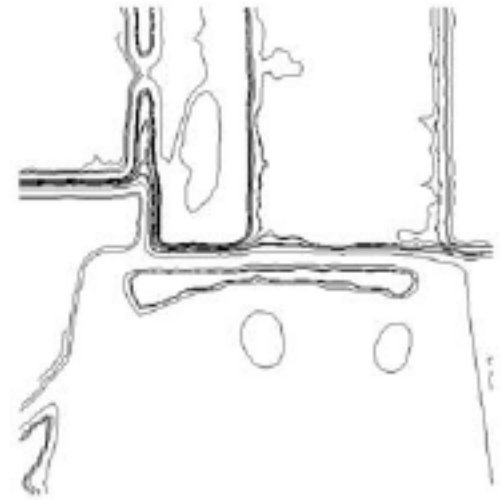


original 84000 11500

900 líneas



- Ventaja: juntas T



## En suma: el mapa topográfico tiene buenas propiedades

- Invariante con respecto a cambios de contraste (estrictamente crecientes)
- Las líneas de nivel son curvas (y no pixels disjuntos) más fáciles de manejar
- Estructura de árbol de inclusión
- En general los contornos de los objetos en las imágenes siguen curvas o pedazos de curvas de nivel.

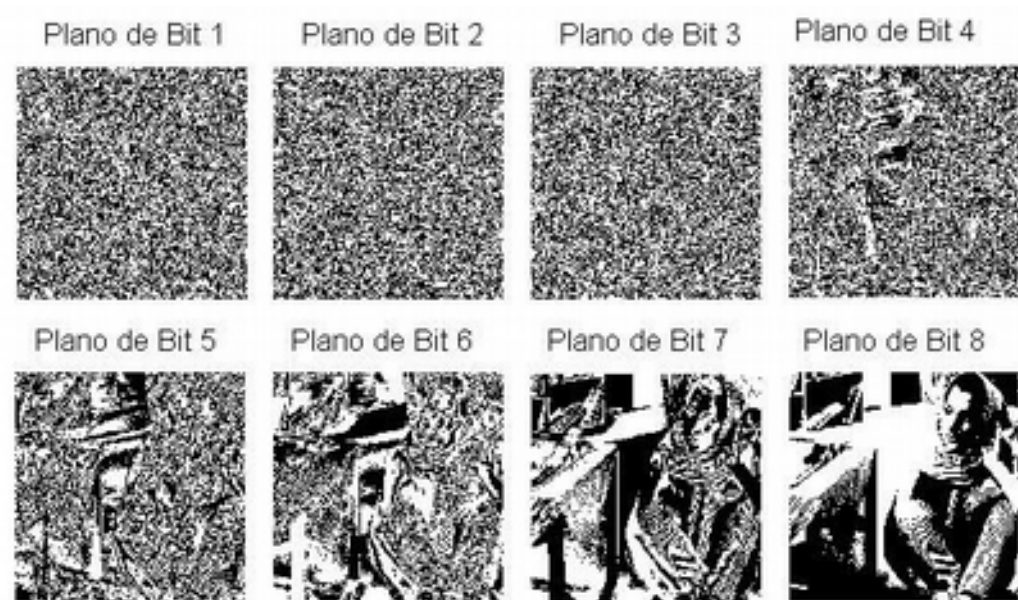
# Descomposición en planos de bit

- Descomposición de una imagen de niveles de gris cuantificada en  $m$  bits, en un conjunto de  $m$  imágenes binarias

- $u(i,j) = b_{m-1}(i,j) \times 2^{m-1} + \dots + b_0(i,j) \times 2^0$

$\rightarrow \{b_{m-1}\}_{i,j}, \dots, \{b_0\}_{i,j}$

# Descomposición en planos de bit



Tomado de “Ocultamiento de Datos en Imagenes Digitales”, Alejandro Padrón Godínez



- **Obs:** dos pixeles vecinos de niveles de gris cercano pueden ser representados de forma muy diferente en los planos de bit.  
**Ejemplo:** 127 (01111111) y 128 (10000000)
- **Descomposición alternativa** para reducir la sensibilidad a pequeños cambios en el nivel de gris: representar la imagen como un **código de Gray de  $m$  bits** antes de descomponer en planos de bit
- Código de Gray: EXOR entre la palabra binaria y ella misma desplazada de 1 bit a la derecha. Dos valores sucesivos difieren sólo en 1 bit.

# Código de Gray

Secuencia	Binario	Gray
0	0000	0000
1	0001	0001
2	0010	0011
3	0011	0010
4	0100	0110
5	0101	0111
6	0110	0101
7	0111	0100
8	1000	1100

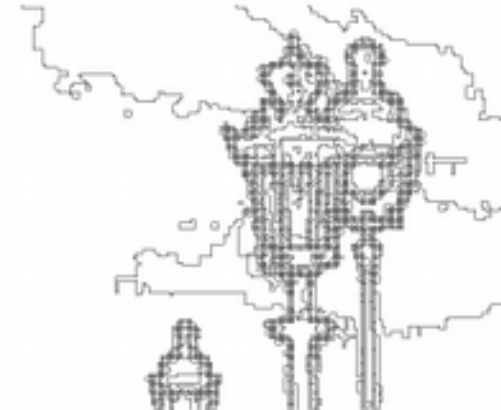
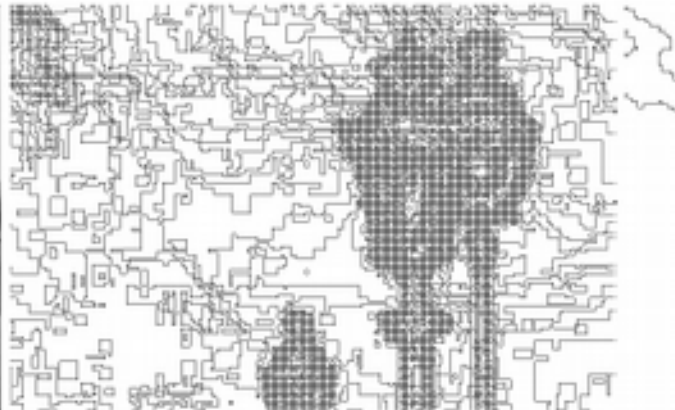
# Interpolación de la imagen y líneas de nivel

- Orden cero:

$$\forall (x, y) \in [i - 1/2, i + 1/2) \times [j - 1/2, j + 1/2), \quad \tilde{u}(x, y) = u(i, j)$$

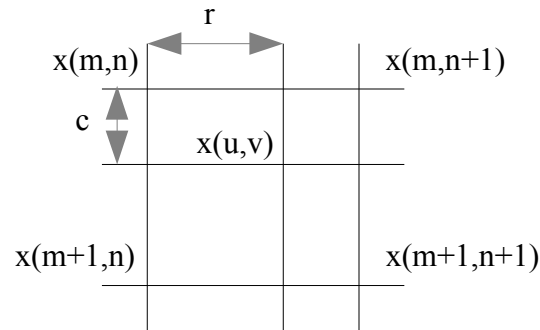
paso de cuantif. 1

paso de cuantif. 10



# Interpolación bilineal (orden 1)

- Preserva el orden entre los niveles de gris.



$$x(u,v) = (1-c)(1-r)x(m,n) + (r)(1-c)x(m,n+1) \\ + (c)(1-r)x(m+1,n) + (r)(c)x(m+1,n+1)$$

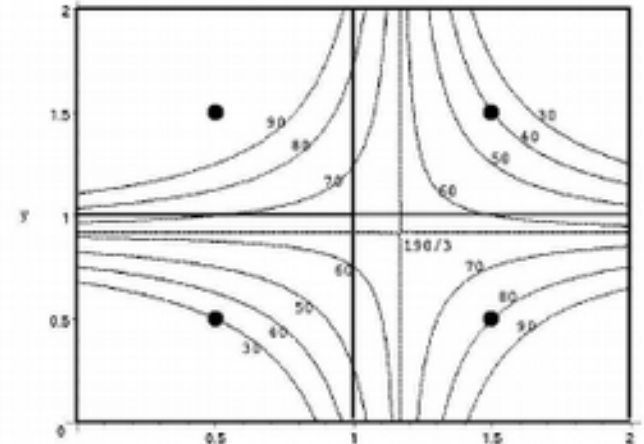
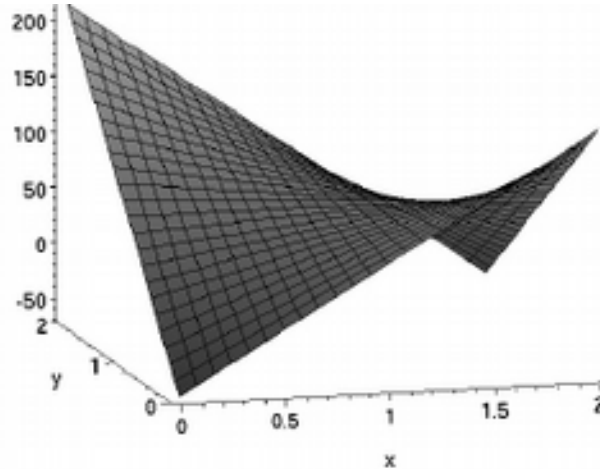
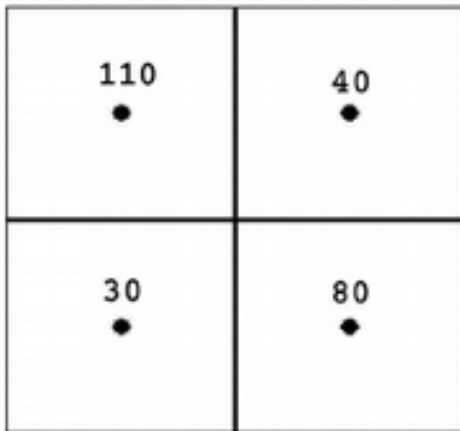
# Interpolación inversa.

- Debido al redondeo pueden quedar puntos sin contribución en la imagen de llegada.
- Usar la transformación inversa y en vez de escribir el valor calculado en cada punto de la imagen de llegada, estimar las contribuciones de la imagen de partida.

- Dentro de los rectángulos  $(i,j),(i+1,j),(i,j+1),(i+1,j+1)$ , la función bilinear se escribe

$$f(x,y) = axy + bx + cy + d$$

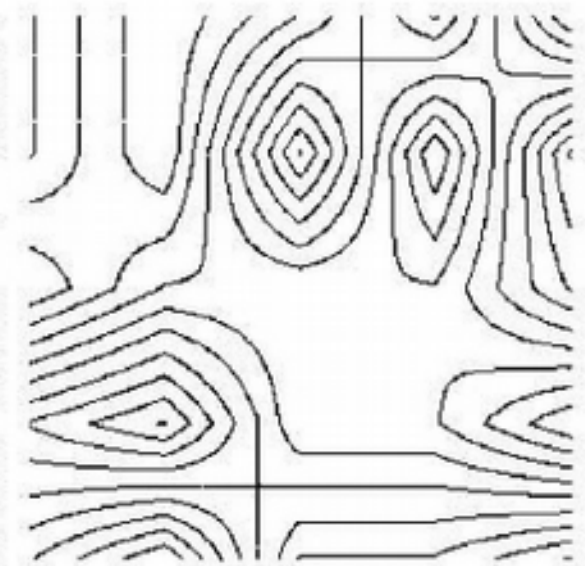
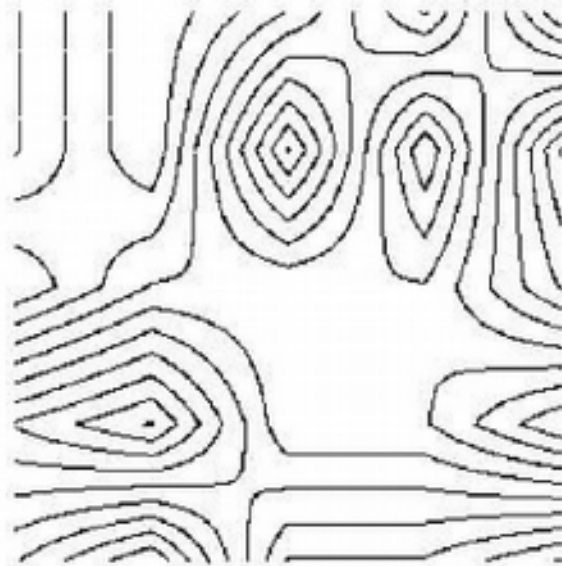
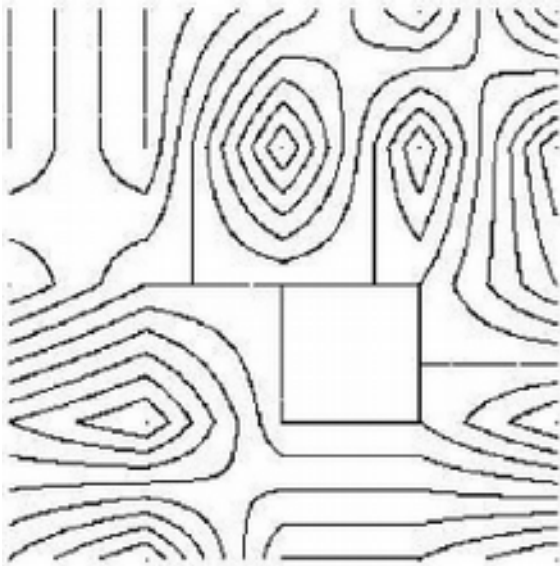
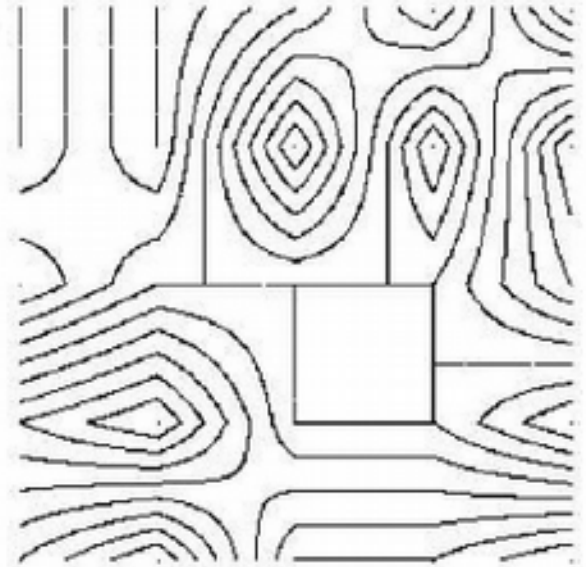
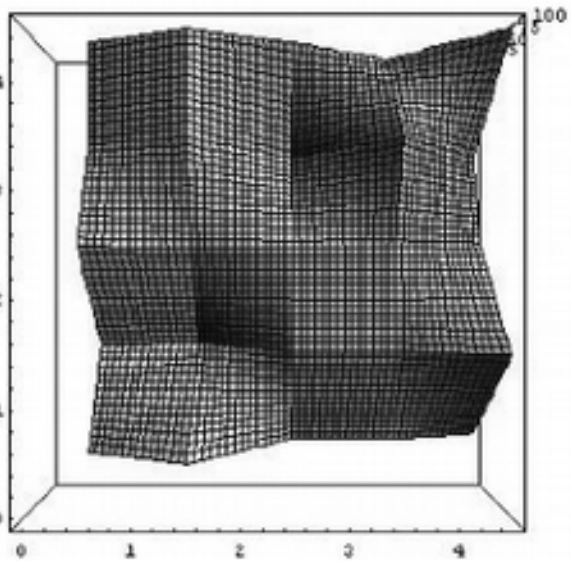
=> Las líneas de nivel de una imagen son concatenaciones de tramos de hipérbolas que tienen un punto silla. El tema es que ese saddle point no caiga dentro del dominio de definición!



# Evitar el efecto de pixelización

- Cada vez que un punto silla caiga en la zona de definición tendremos efecto de pixelización.
- Ello ocurre cuando los valores interpolados coinciden con el valor original de la imagen.
- La solución es que los valores interpolados sean cercanos pero no iguales.

70	100	70	30	100
70	100	10	100	10
100	70	70	70	30
30	10	70	70	100
70	100	30	30	10



Pablo Musé. On the definition and recognition of planar shapes in digital images. ENS Cachan, 2004.



- Orden 0, cada 10 niveles empezando en 10
- Orden 1, cada 10 niveles empezando en 10
- Orden 1, cada 10 niveles empezando en 0.5
- La solución es interpolar usando valores no enteros.

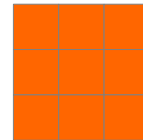
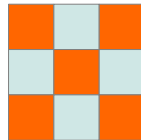
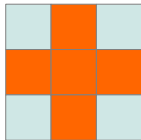


# Ejercicios

- **Extraer las líneas de nivel** de una imagen, con interpolación de orden cero y de orden 1
- **Filtro “extrema killer”**: contaminar una imagen con ruido impulsivo. Filtrarla extrayendo los conjuntos de nivel superiores e inferiores, y eliminando aquellos con área pequeña.
- **Desplegar los planos de bit de una imagen**, usando representación binaria directa y código de Gray

# Vecinos de un píxel

- $V$ : rango de niveles de gris.
- $N_4(p)$  Vecinos horizontales y verticales de  $p$ .
- $N_D(p)$  Vecinos diagonales de  $p$ .
- $N_8(p) = N_4(p) \cup N_D(p)$



# Conectividad

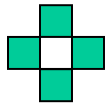
- Los pixels  $p$  y  $q$  son conexos si:
- Pertenecen a un rango  $V$  y son vecinos.
- 4-conectividad:  $q \in N_4(p)$
- 8-conectividad:  $q \in N_8(p)$
- $m$ -conectividad:  $q \in N_4(p)$  o

$$q \in N_D(p) \quad \text{y} \quad N_4(p) \cap N_4(q) = \emptyset$$

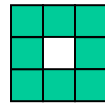
# Conectividad

- Problema: Una región encerrada y que no divida al espacio en 2 zonas diferentes.
- Evitar caminos múltiples.

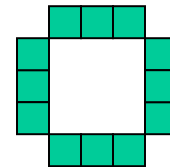
4-conexos



8-conexos



Centro y fondo,  
están conectados?



# Adyacencia

- Dos regiones son adyacentes si algún pixel de una región es conexo a alguno de la otra región.
- Se pueden definir caminos, regiones conexas, etc.
- Esto es muy útil en etiquetado.

# Métricas

- Dados los píxeles  $p, q$  y  $z$  de coordenadas  $(x, y), (s, t)$  y  $(u, v)$ ,  $D$  es una función distancia o una métrica si:

$$D(p, q) \geq 0 \quad (D(p, q) = 0 \text{ si y solo si } p = q)$$

$$D(p, q) = D(q, p)$$

$$D(p, z) \leq D(p, q) + D(q, z)$$

# Métricas

- Distancia euclídea:  $D_e(p, q) = \sqrt{(x - s)^2 + (y - t)^2}$
- Los píxeles que están a una distancia  $r$  de  $(x, y)$  son los contenidos en un círculo de radio  $r$  y centrado en  $(x, y)$ .



# Métricas

- Distancia city-block:  $D_4(p, q) = |x - s| + |y - t|$
- Los píxels que están a una distancia  $r$  de  $(x, y)$  son los contenidos en un rombo de radio  $r$  y centrado en  $(x, y)$ .

# Métricas

- Distancia damero:  $D_8(p, q) = \max(|x - s|, |y - t|)$
- Los píxels que están a una distancia  $r$  de  $(x, y)$  son los contenidos en un cuadrado de radio  $r$  y centrado en  $(x, y)$ .

# Operaciones aritméticas

- Se realizan píxel a píxel en toda la imagen.
- No hay problemas de borde.
- Adición:  $p+q$
- Sustracción:  $p-q$
- Multiplicación:  $pq$
- División:  $p/q$

# Operaciones lógicas

- AND: 1 si los dos pixels valen 1.
- OR: 1 si uno de los dos pixels valen 1.
- NOT: lo contrario del valor del píxel.
- XOR: 1 si uno de los dos pixels vale 1

# Operaciones sobre vecinos

- Imagen de salida: el píxel de salida vale el resultado de la función sobre la vecindad en torno al píxel de la misma posición.
- Vecindad: en general es una ventana cuadrada.

# Estrategias de borde

- Perder 2\*orden líneas y columnas.
- Borde virtual con:
  - valor medio
  - simetría
  - borde opuesto.

# Estrategias de borde

