

Codificación Perceptual

Miguel Hernández-Cabronero

<miguel.hernandez@uab.cat>

UAB

Universitat Autònoma
de Barcelona

G|Ci

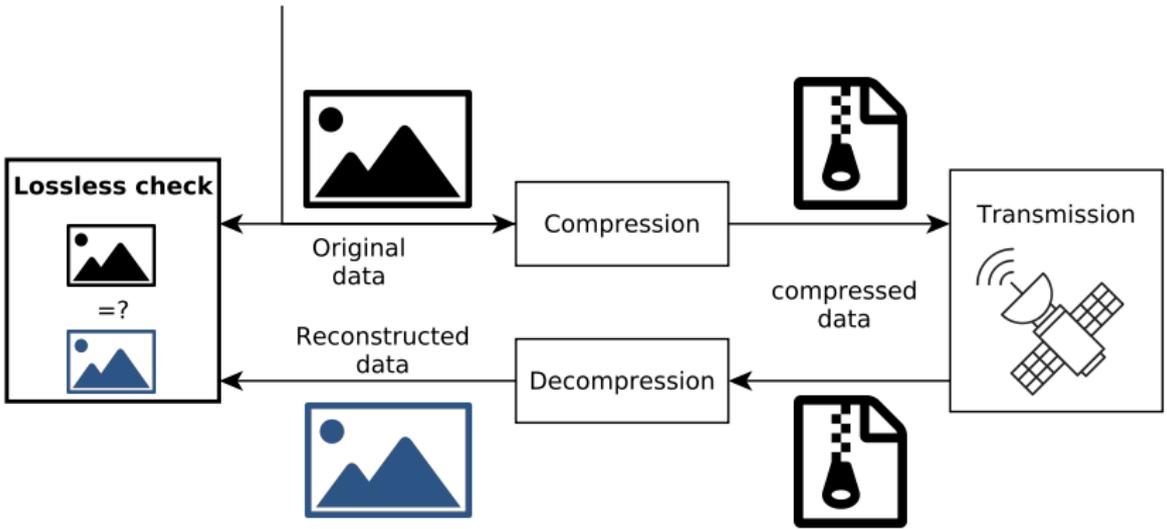
Group on Interactive
Coding of Images

Universidad Miguel Hernández

19 de Noviembre de 2019

- 1 Nos hackean
- 2 Qué tal se recibe?
 - Calidad subjetiva
 - Límites de la percepción
- 3 Cómo te lo dibujo?
 - Básicos de compresión
 - JPEG (classic)
 - JPEG 2000
- 4 Modelos de la Percepción
 - CSF y Thresholds Visuales (VTs)
 - Probabilidad de detección
 - Métricas de calidad

Lossless vs Lossy



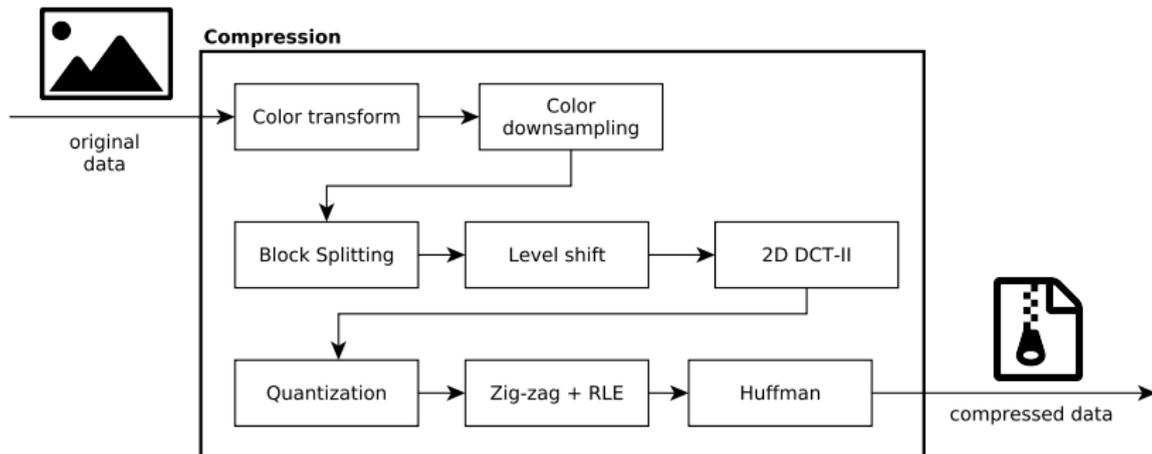
Lossless vs Lossy

- Cuándo es OK tener pérdida?
- Cuánta pérdida?
- Cómo se pierde?
- Cómo controlarla?

Cuantización uniforme

$$x \xrightarrow{\text{quantization}} Q(x) = \left\lfloor \frac{x}{\Delta} \right\rfloor \xrightarrow{\text{reconstruction}} \hat{x} = Q(x) \cdot \Delta + \frac{\Delta}{2}$$

JPEG (classic)



2D DCT

Transformada directa:

Original matrix



Transformed matrix



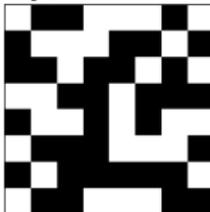
Transformed matrix (no DC)



|Transformed matrix|



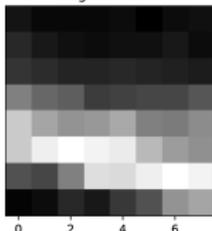
sign(Transformed matrix)



2D DCT

Transformada directa:

Original matrix



Transformed matrix



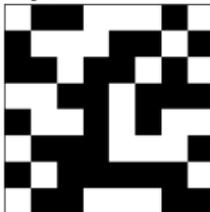
Transformed matrix (no DC)



[Transformed matrix]



sign(Transformed matrix)

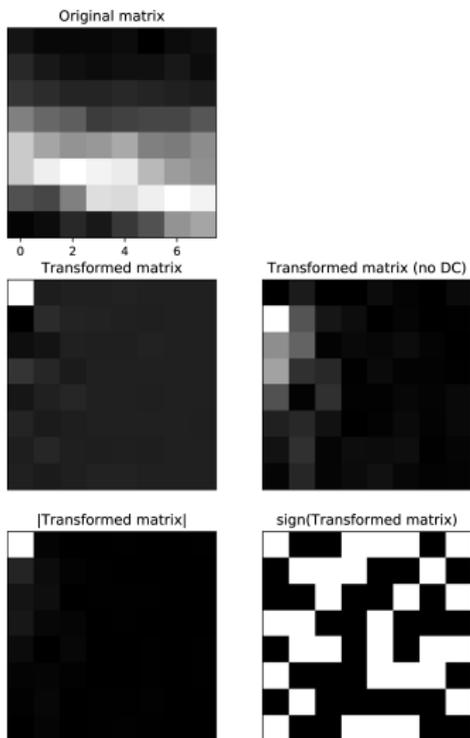


Demo:

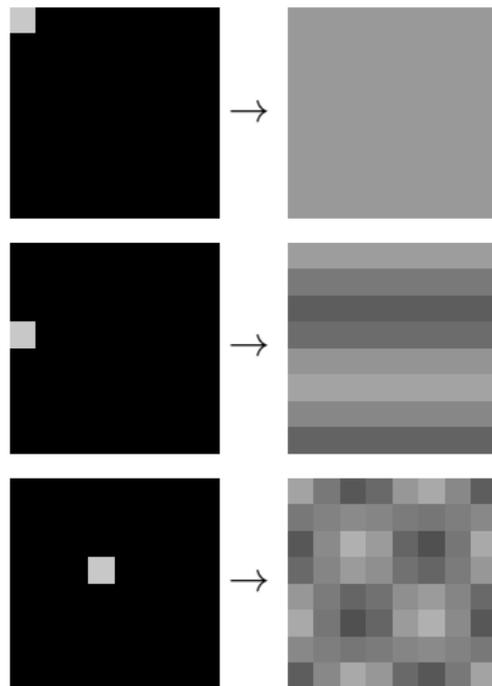


2D DCT

Transformada directa:



Transformada inversa:



2D DCT

Transformada directa:

Animacion (fuelle)

Original matrix



Transformed matrix



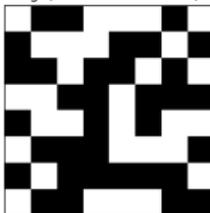
Transformed matrix (no DC)



|Transformed matrix|



sign(Transformed matrix)

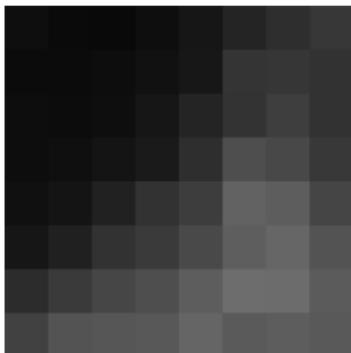


Quantización DCT – Filtrado de frecuencias

- Bandas cuantizadas independientemente
⇒ filtrado selectivo de frecuencias

$$DCT(i,j) \leftarrow Q(i,j) \cdot \lfloor DCT(i,j)/Q(i,j) \rfloor$$

$\Delta =$



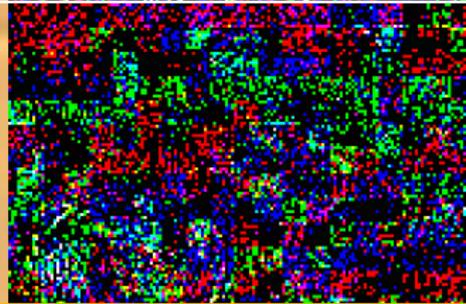
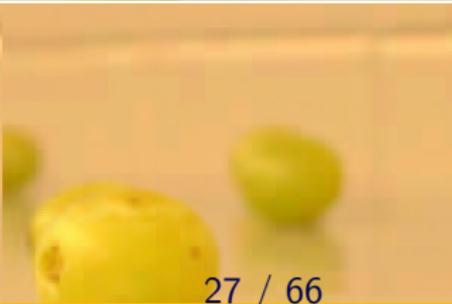
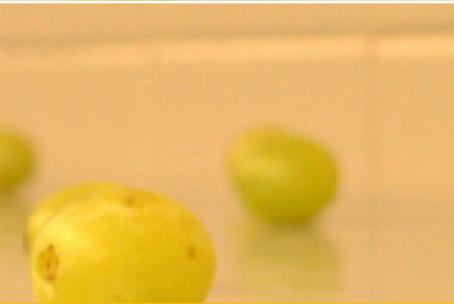
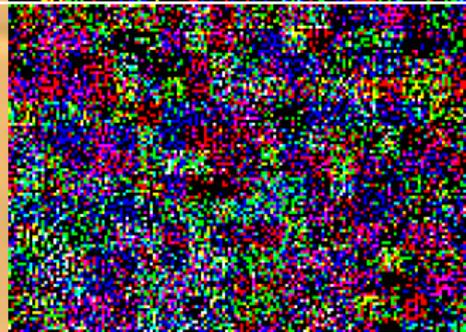
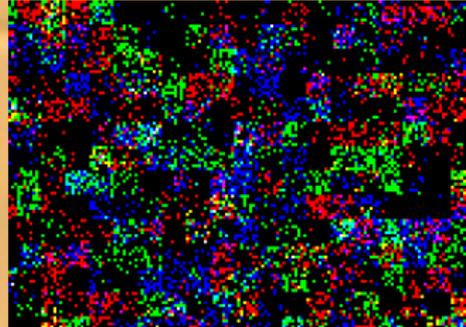
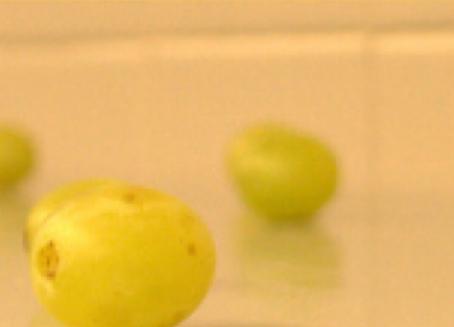
- black: $\Delta = 1$
- white: $\Delta = \infty$

Nos hackean
oooooooo

Qué tal se recibe?
ooooooo

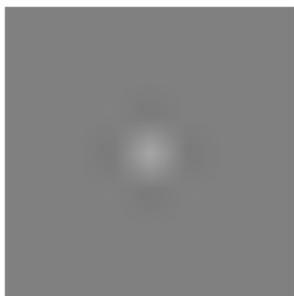
Cómo te lo dibujo?
oooooooooooooooo●oooooooooooo

Modelos de la Percepción
oo



DWT – vectores de síntesis

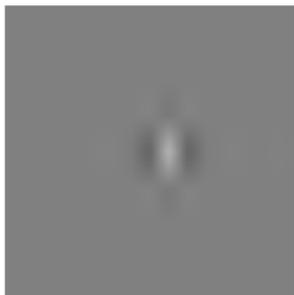
5 niveles DWT inversa de **1 coeficiente**:



LL5



HL5



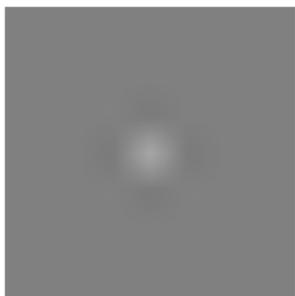
LH5



HH5

DWT – vectores de síntesis

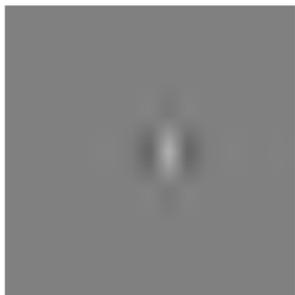
5 niveles DWT inversa de **1 coeficiente**:



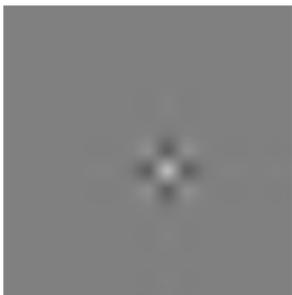
LL5



HL5



LH5



HH5

Cada coeficiente DWT:

- Info de lugar
- Info de frecuencia
⇒ no hay artefactos de bloque

DWT – parámetros visuales

Ajustes visuales en JPEG 2000:

- Color downsampling
- Adaptar paso de cuantización Δ :
 - Bitrate¹
 - Pesos componentes color¹
 - Δ para subbandas HL1, LH1, HH1, ..., HL5, LH5, HH5, LL5¹
 - Δ localmente adaptativo (visual masking^{1,2})

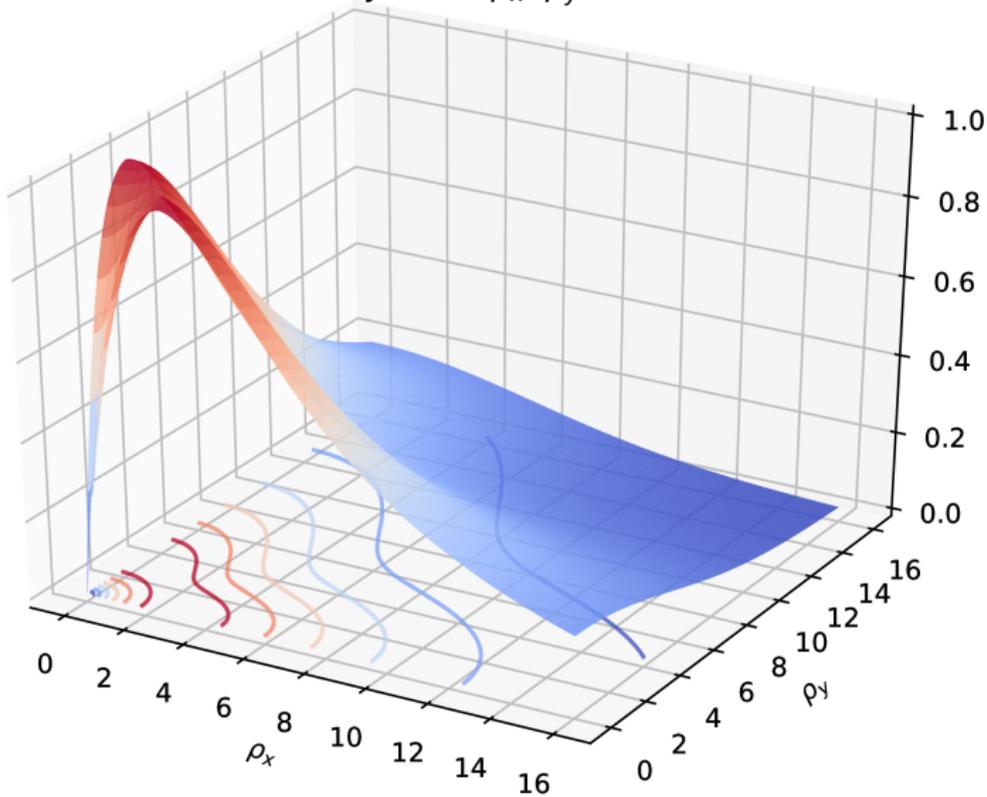
Otros ajustes importantes:

- Niveles de resolución
- Transmisión progresiva
- Region de interés (ROI)

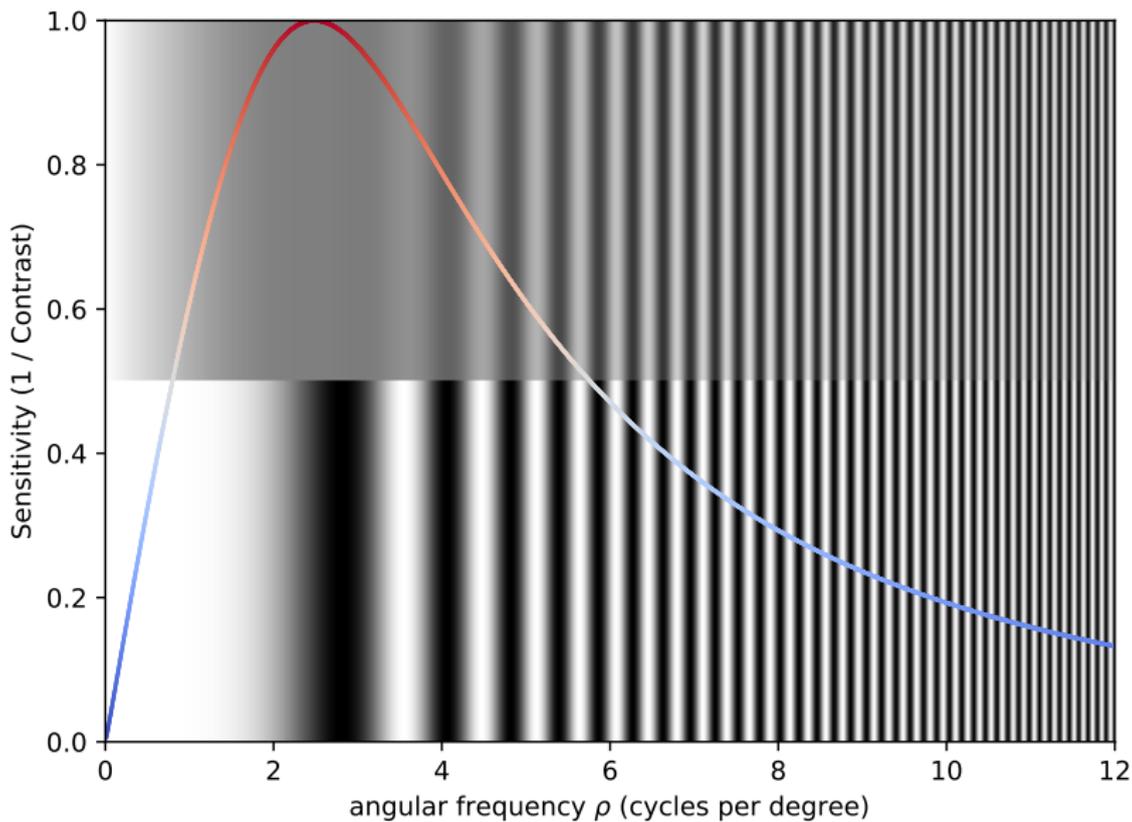
—
¹ Ver [Taubman], esp. Caps. 15.5 y 16.1.

² Ver [Oh]

Daly's CSF(ρ_x, ρ_y)



Mantiuk's Model (2013, HDR)



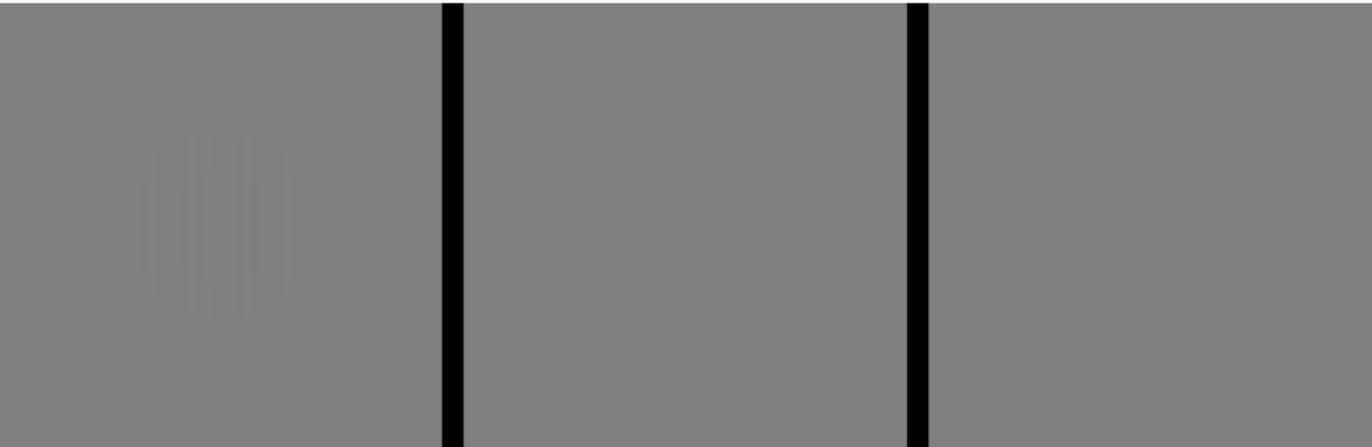
Dónde está el límite?



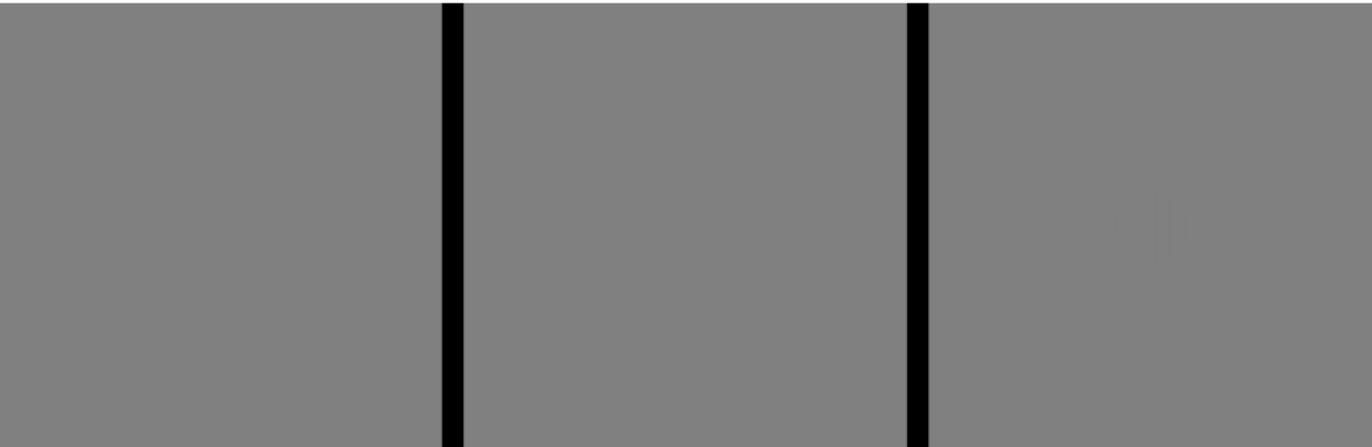
Dónde está el límite?



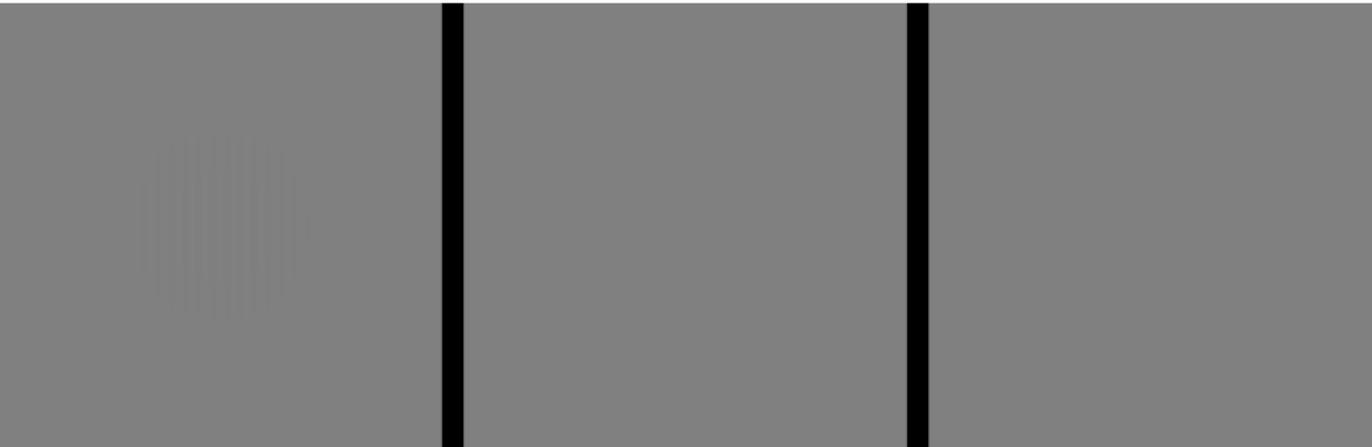
3AFC – 3-alternative forced choice



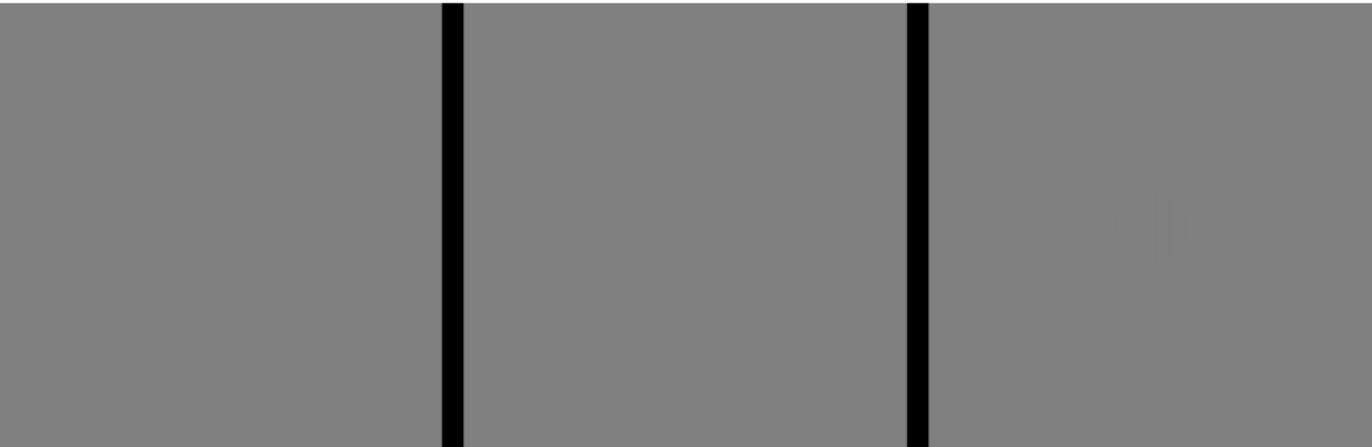
3AFC – 3-alternative forced choice



3AFC – 3-alternative forced choice



3AFC – 3-alternative forced choice



Cómo medir

Repetir:

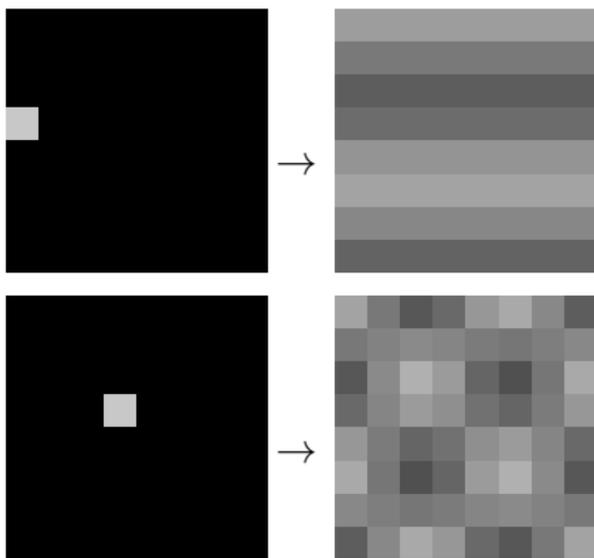
- × cada frecuencia
- × cada orientación
- × cada color
- × cada brillo de fondo



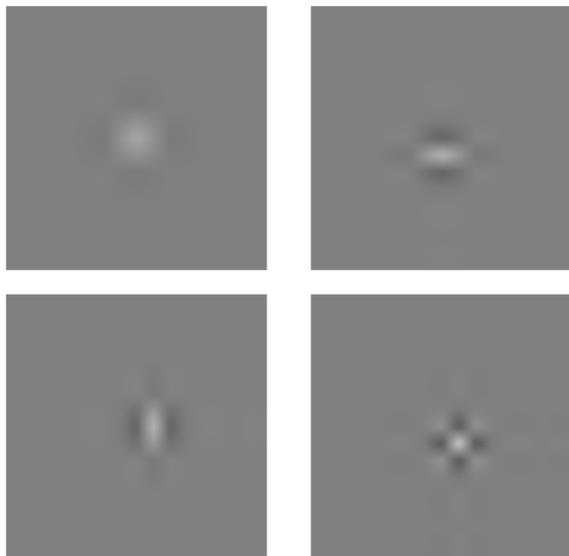
Otras CSF

Podemos medir Visibility Thresholds (VT) para otros estímulos:

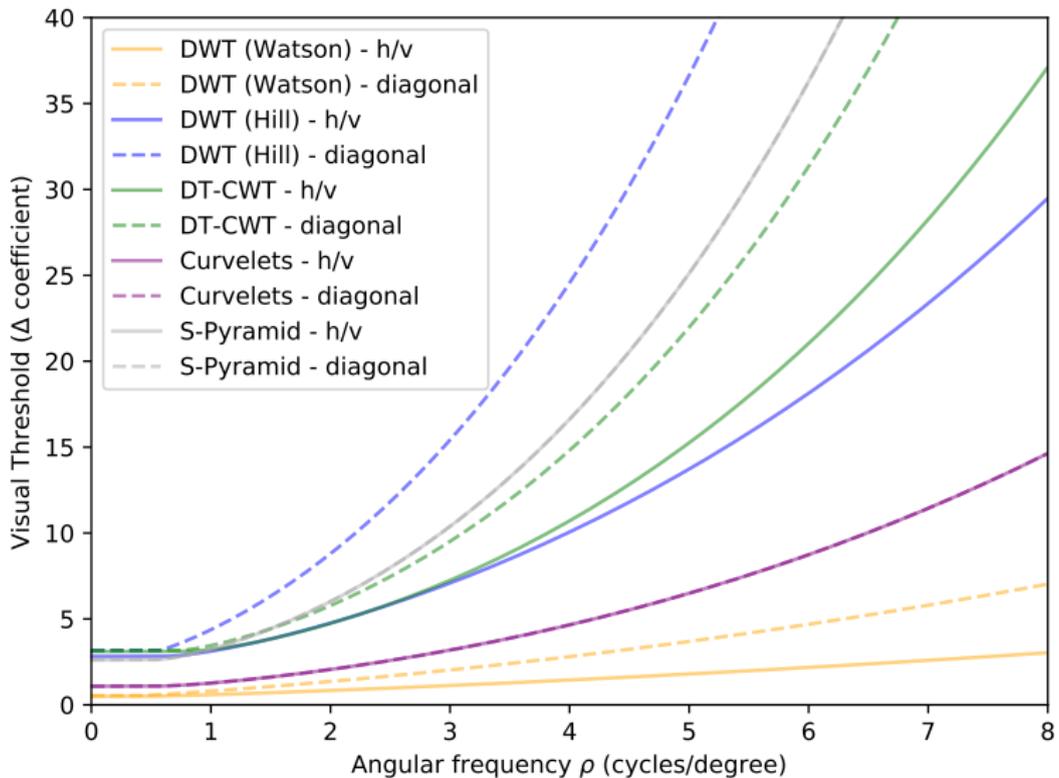
DCT:



DWT:



[Hill] 2016

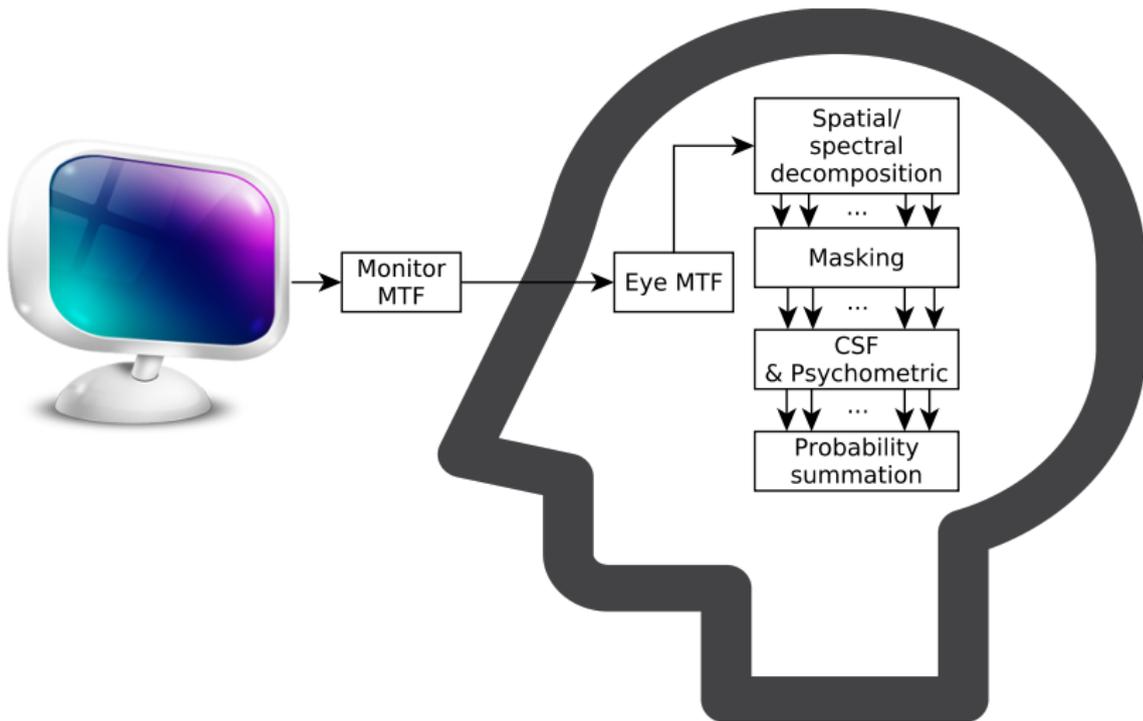


Problemas prácticos de las CSF:

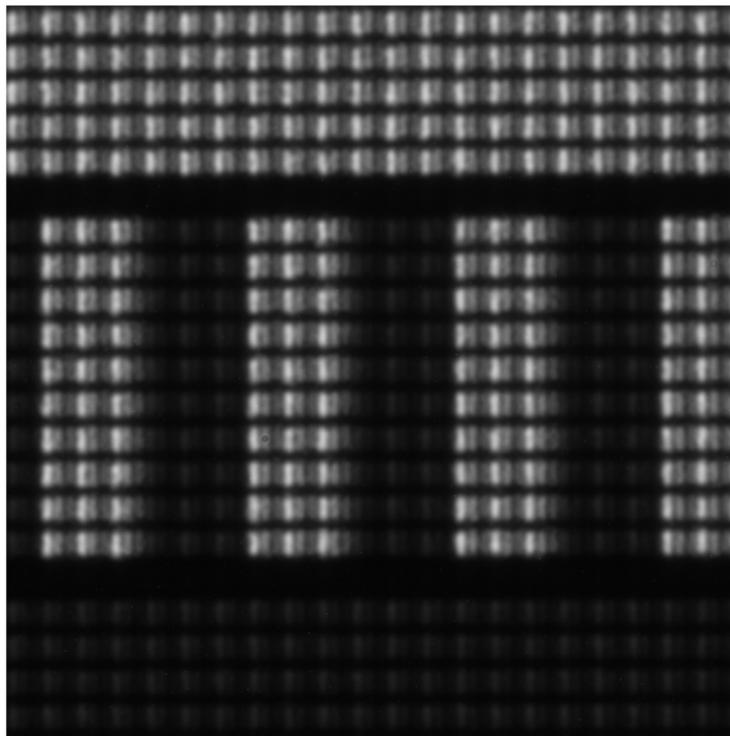
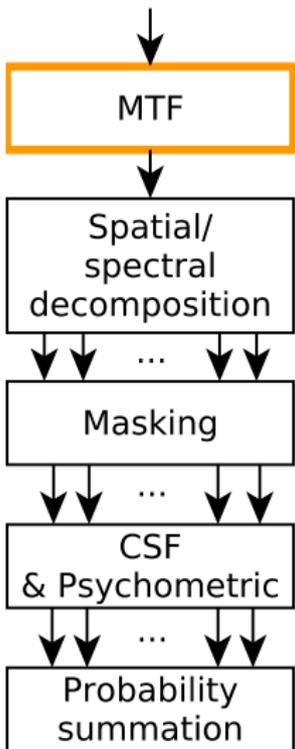
- Cómo aplicarla al codificar?
- Qué ocurre al combinar estímulos?
- Qué vemos sobre otro fondo?
- En otras condiciones?

- 1 Nos hackean
- 2 Qué tal se recibe?
 - Calidad subjetiva
 - Límites de la percepción
- 3 Cómo te lo dibujo?
 - Básicos de compresión
 - JPEG (classic)
 - JPEG 2000
- 4 Modelos de la Percepción
 - CSF y Thresholds Visuales (VTs)
 - Probabilidad de detección
 - Métricas de calidad

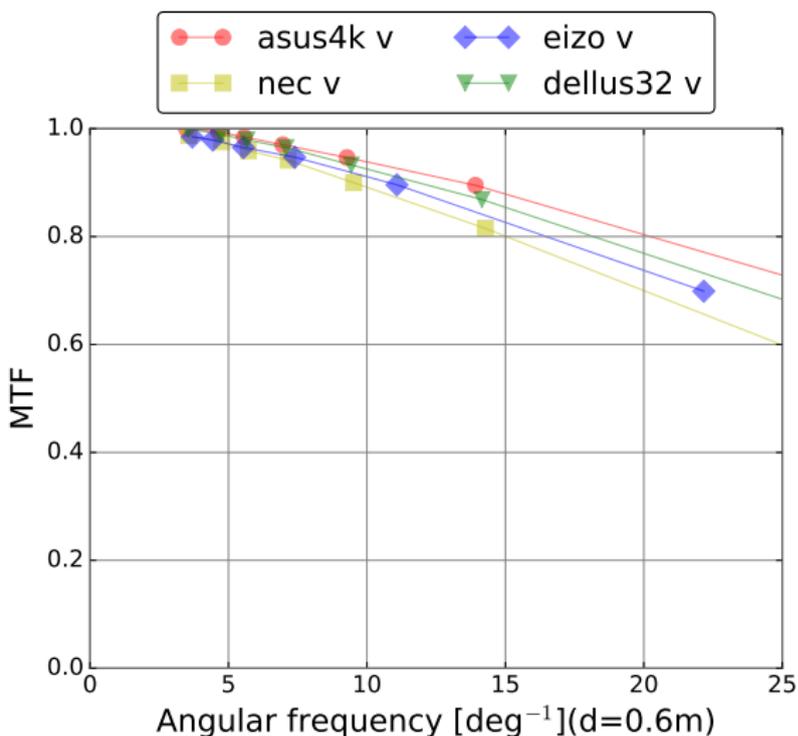
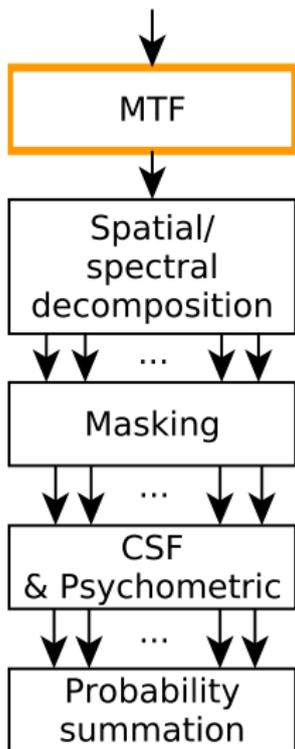
Modelo completo



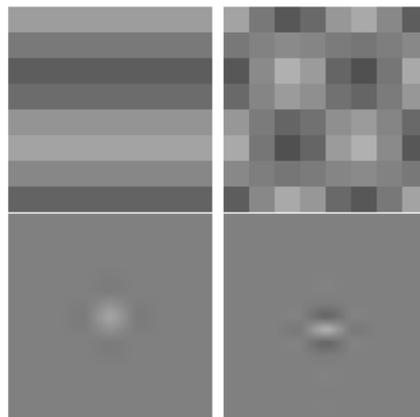
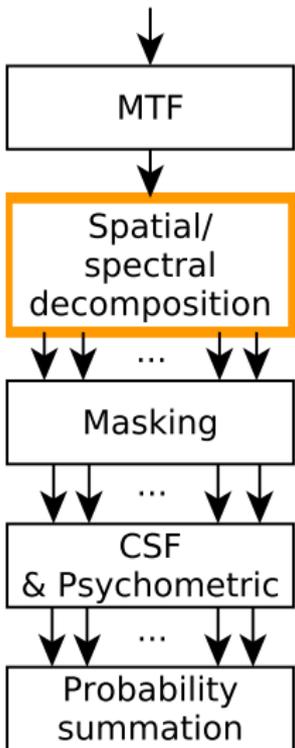
MTF



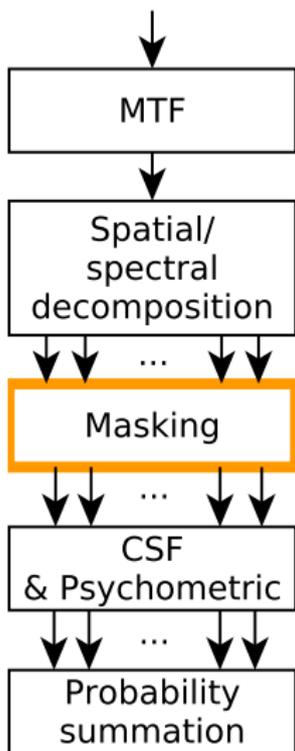
MTF



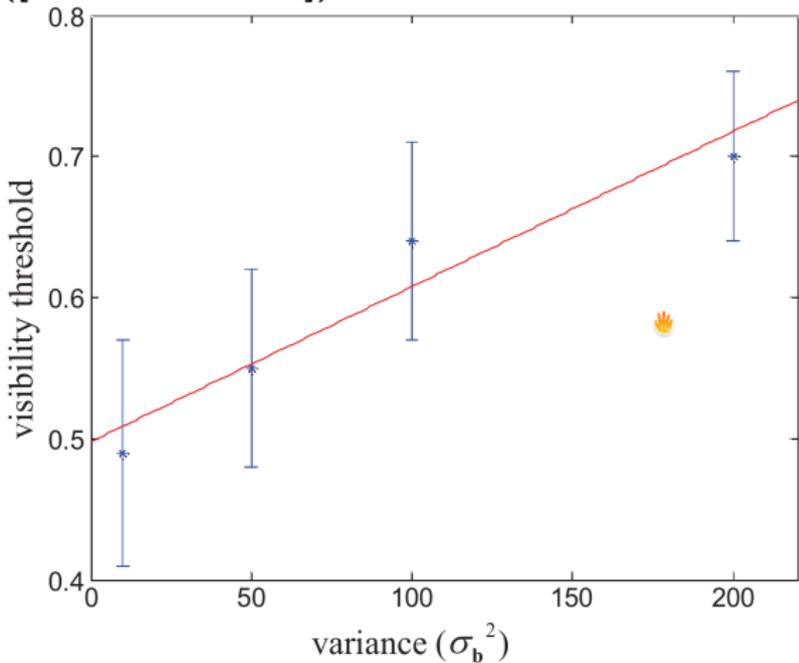
Band decomposition



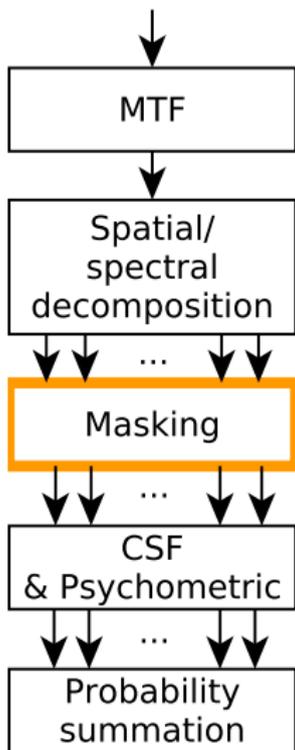
Masking



Señales activas ocultan estímulos ([Oh13, Watson06]):



Masking

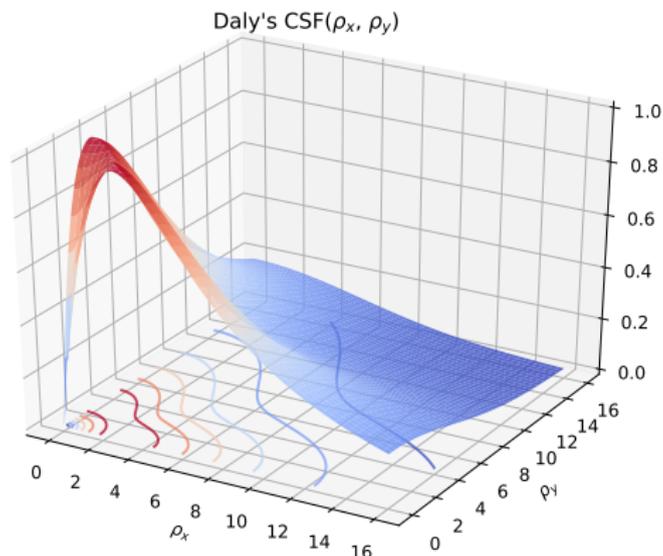
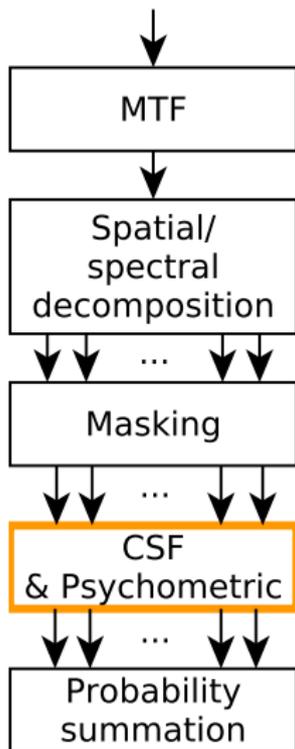


Más brillo: menos sensibilidad:

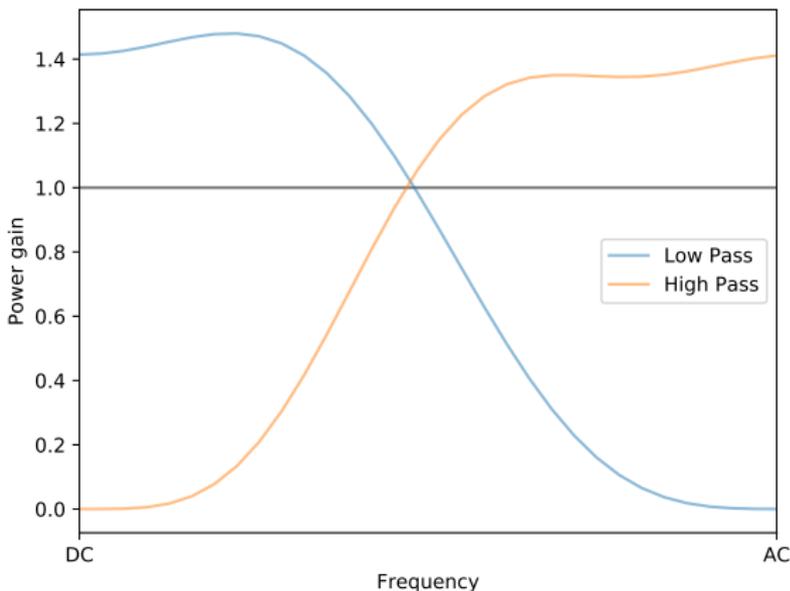
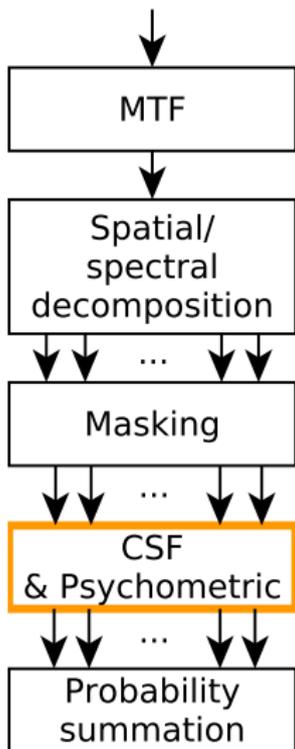
$$a_I(\lambda, \theta) = \left(\frac{v_{\lambda_{\max}, LL}}{v_{bg}} \right)^{a_T}$$

threshold $\Delta'_{\lambda, \theta} = \Delta_{\lambda, \theta} / a_I(\lambda, \theta)$

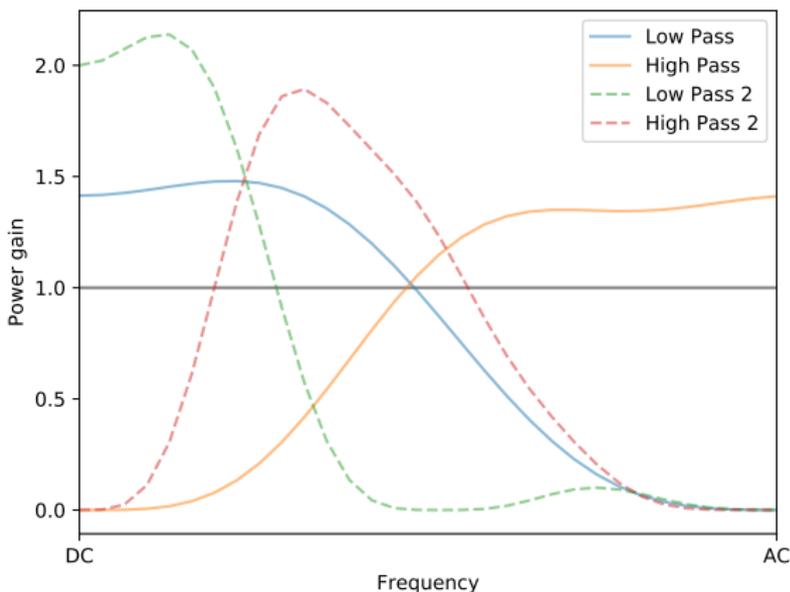
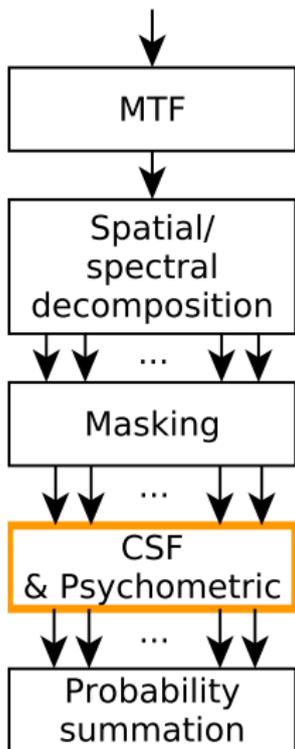
Aplicación de la CSF



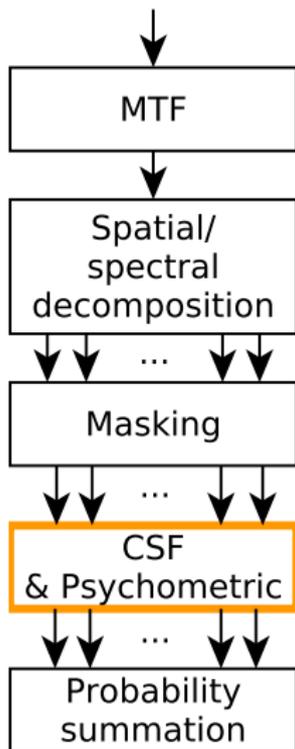
Aplicación de la CSF



Aplicación de la CSF



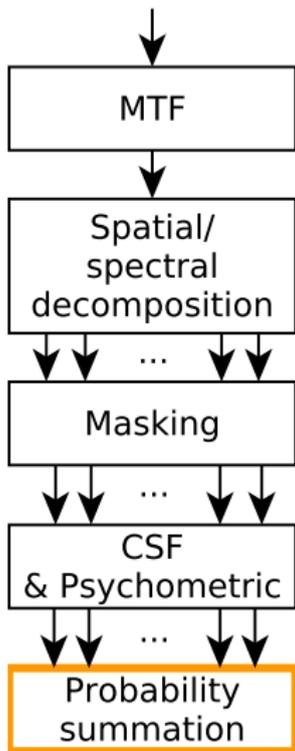
Aplicación de la CSF



Frecuencias representativas:

- Punto medio
- Centro de masa
- Experimentalmente

Integración de probabilidades



P^{det} para cada pixel:

$$P_{map}^{det}(x, y) = 1 - \prod_{\lambda, \theta} (1 - P_{\lambda, \theta}^{det}(\lambda, \theta))$$

Integración espacial:

- Máximo de detección en cada punto [Mantiuk]
- A veces se integran parámetros de masking [Oh13]:

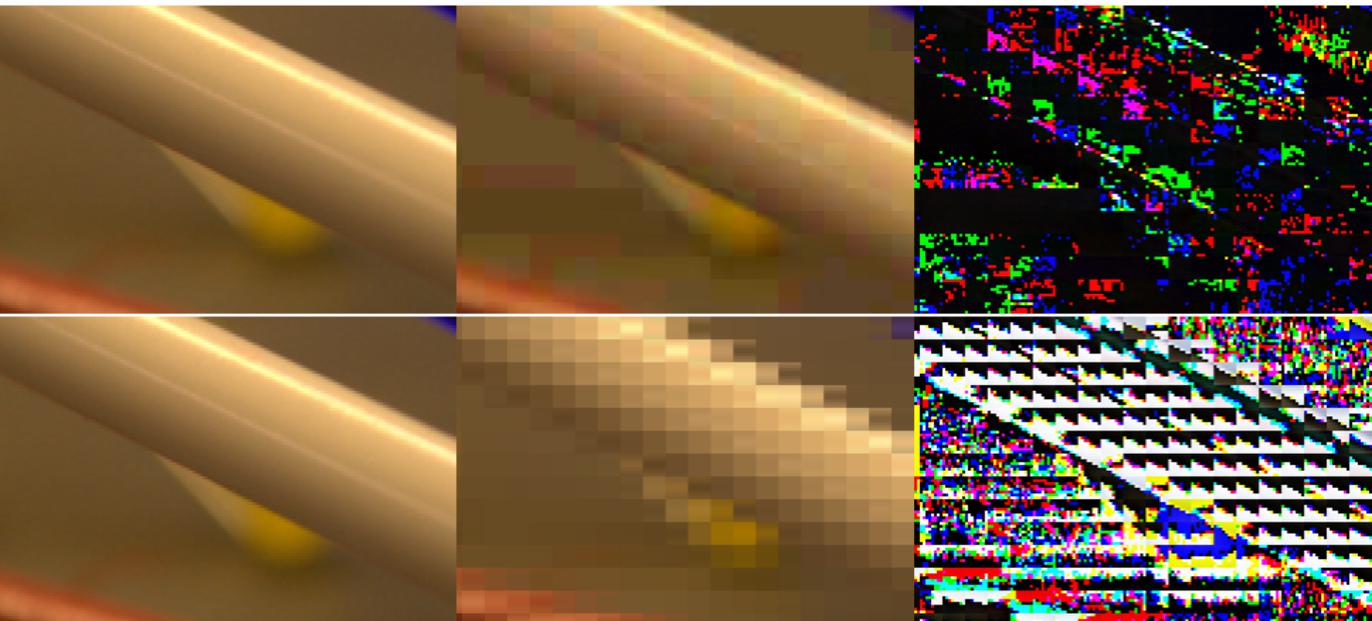
$$\left(\frac{1}{|B|} \sum_{n \in B} mask(n)^\beta \right)^{1/\beta}$$

- 1 Nos hackean
- 2 Qué tal se recibe?
 - Calidad subjetiva
 - Límites de la percepción
- 3 Cómo te lo dibujo?
 - Básicos de compresión
 - JPEG (classic)
 - JPEG 2000
- 4 Modelos de la Percepción
 - CSF y Thresholds Visuales (VTs)
 - Probabilidad de detección
 - Métricas de calidad

MSE

$$MSE(X, \hat{X}) = \frac{1}{|X|} \sum_{(x, \hat{x}) \in (X, \hat{X})} (x - \hat{x})^2$$

Ejemplos cuantización DCT



Top: JPEG Q60

Bottom: H2

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...
- Niveles de ajuste:
 - ① Automático (algoritmo decide) – calidad?

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...
- Niveles de ajuste:
 - 1 Automático (algoritmo decide) – calidad?
 - 2 Basado en MSE/PSNR (típico)

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...
- Niveles de ajuste:
 - 1 Automático (algoritmo decide) – calidad?
 - 2 Basado en MSE/PSNR (típico)
 - 3 Basado en modelo CSF (e.g., Daly)

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...
- Niveles de ajuste:
 - 1 Automático (algoritmo decide) – calidad?
 - 2 Basado en MSE/PSNR (típico)
 - 3 Basado en modelo CSF (e.g., Daly)
 - 4 Basado en modelo completo de la visión (HDR-VDP2)

Conclusiones

- Engañamos al cerebro rutinariamente
- Codificación perceptual \Rightarrow sistemas más eficientes
 - Cuantización \Rightarrow $\left\{ \begin{array}{l} \text{más compresión, y} \\ \text{descarte selectivo de info} \end{array} \right.$
 - Sensitividad B/N vs. color
 - Ajustable en JPEG, JPEG 2000, HEVC, ...
- Niveles de ajuste:
 - 1 Automático (algoritmo decide) – calidad?
 - 2 Basado en MSE/PSNR (típico)
 - 3 Basado en modelo CSF (e.g., Daly)
 - 4 Basado en modelo completo de la visión (HDR-VDP2)
 - 5 Basado en pruebas subjetivas – calidad!

Bibliografía II

- **[Watson93]**: Watson, A.B., 1993, September. "DCT quantization matrices visually optimized for individual images". In Human vision, visual processing, and digital display IV (Vol. 1913, pp. 202-216). International Society for Optics and Photonics.
- **[Watson06]**: Zhen Liu, L. J. Karam and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," in IEEE Transactions on Image Processing, vol. 15, no. 7, pp. 1763-1778, July 2006.
- **[Oh13]**: H. Oh, A. Bilgin and M. W. Marcellin, "Visually Lossless Encoding for JPEG2000," in IEEE Transactions on Image Processing, vol. 22, no. 1, pp. 189-201, Jan. 2013.
- **[Taubman]**: Taubman, D.S. and Marcellin, M.W., 2002. "JPEG2000: Image compression fundamentals. Standards and Practice," pp.295-301, Kluwer.