

## CC5213-1 Recuperación de Información Multimedia

Prof. de Cátedra: Juan Manuel Barrios

Estudiante: Andrés Calderón Guardia



## Mini-Control 1

September 16, 2023

## Pregunta 1

## a) Solución:

Manteniendo el borde con los valores que ya posee, si aplicamos este filtro resulta la siguiente imagen:

100	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	100	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	100	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	100

## b) Solución:

Para este caso al aplicar la convolución descartaremos el borde de la imagen original en el resultado:

300	-300	300	-300	300
-300	500	-300	400	-300
300	-300	300	-300	300
-300	400	-300	500	-300
300	-300	300	-300	300

## c) Solución:

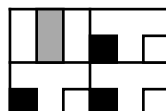
Aplicando el umbral  $t$  sobre el resultado anterior nos queda:

0	0	0	0	0
0	255	0	0	0
0	0	0	0	0
0	0	0	255	0
0	0	0	0	0

## Pregunta 2

## a) Solución:

Primero por como es la imagen utilizaremos tan solo 3 bins, pues solo hay 3 tonos en la escala de grises (negro, gris y blanco), resultando en el siguiente histograma de grises:



La suma de frecuencias para cada una de las 4 zonas es de 64, por lo que dividimos la altura de cada bin por 64, de modo que la altura para cada bin del histograma de grises es como sigue: para la primera

zona es de 1 para el gris y 0 para los demás tonos, mientras que en el resto de zonas se cumple que la frecuencia del negro y blanco es de  $\frac{1}{2}$  mientras que la del gris es 0.

### b) Solución:

Ahora para identificar los píxeles adecuadamente los enumeraremos de la siguiente manera:

(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)	(1,8)	(1,9)	(1,10)	(1,11)	(1,12)	(1,13)	(1,14)	(1,15)	(1,16)
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)	(2,7)	(2,8)	(2,9)	(2,10)	(2,11)	(2,12)	(2,13)	(2,14)	(2,15)	(2,16)
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)	(3,7)	(3,8)	(3,9)	(3,10)	(3,11)	(3,12)	(3,13)	(3,14)	(3,15)	(3,16)
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)	(4,7)	(4,8)	(4,9)	(4,10)	(4,11)	(4,12)	(4,13)	(4,14)	(4,15)	(4,16)
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)	(5,7)	(5,8)	(5,9)	(5,10)	(5,11)	(5,12)	(5,13)	(5,14)	(5,15)	(5,16)
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)	(6,7)	(6,8)	(6,9)	(6,10)	(6,11)	(6,12)	(6,13)	(6,14)	(6,15)	(6,16)
(7,1)	(7,2)	(7,3)	(7,4)	(7,5)	(7,6)	(7,7)	(7,8)	(7,9)	(7,10)	(7,11)	(7,12)	(7,13)	(7,14)	(7,15)	(7,16)
(8,1)	(8,2)	(8,3)	(8,4)	(8,5)	(8,6)	(8,7)	(8,8)	(8,9)	(8,10)	(8,11)	(8,12)	(8,13)	(8,14)	(8,15)	(8,16)
(9,1)	(9,2)	(9,3)	(9,4)	(9,5)	(9,6)	(9,7)	(9,8)	(9,9)	(9,10)	(9,11)	(9,12)	(9,13)	(9,14)	(9,15)	(9,16)
(10,1)	(10,2)	(10,3)	(10,4)	(10,5)	(10,6)	(10,7)	(10,8)	(10,9)	(10,10)	(10,11)	(10,12)	(10,13)	(10,14)	(10,15)	(10,16)
(11,1)	(11,2)	(11,3)	(11,4)	(11,5)	(11,6)	(11,7)	(11,8)	(11,9)	(11,10)	(11,11)	(11,12)	(11,13)	(11,14)	(11,15)	(11,16)
(12,1)	(12,2)	(12,3)	(12,4)	(12,5)	(12,6)	(12,7)	(12,8)	(12,9)	(12,10)	(12,11)	(12,12)	(12,13)	(12,14)	(12,15)	(12,16)
(13,1)	(13,2)	(13,3)	(13,4)	(13,5)	(13,6)	(13,7)	(13,8)	(13,9)	(13,10)	(13,11)	(13,12)	(13,13)	(13,14)	(13,15)	(13,16)
(14,1)	(14,2)	(14,3)	(14,4)	(14,5)	(14,6)	(14,7)	(14,8)	(14,9)	(14,10)	(14,11)	(14,12)	(14,13)	(14,14)	(14,15)	(14,16)
(15,1)	(15,2)	(15,3)	(15,4)	(15,5)	(15,6)	(15,7)	(15,8)	(15,9)	(15,10)	(15,11)	(15,12)	(15,13)	(15,14)	(15,15)	(15,16)
(16,1)	(16,2)	(16,3)	(16,4)	(16,5)	(16,6)	(16,7)	(16,8)	(16,9)	(16,10)	(16,11)	(16,12)	(16,13)	(16,14)	(16,15)	(16,16)

Primero notemos que los píxeles que están tanto en la fila 16 como en la columna 16 diremos que no tendrán borde por simplicidad, pues por la fórmula que se nos da los bordes dependen de los píxeles que estén directamente abajo y a la derecha del que estamos revisando.

Ahora definiremos los valores para cada color, que en este caso el blanco correspondería al 255, el gris (usando un cuentagotas) valdría 150, y el negro vale 0, después necesitamos un umbral para considerar un píxel como uno de borde, el cual simplemente diremos que debe ser mayor a 0. Luego, dada la fórmula con la que calcularemos los gradientes tenemos que los píxeles grises entre (1,1) y (7,7) no poseen un borde, pues al calcular su gradiente obtendríamos 0 en ambas componentes, y un detalle es que si obtenemos 0 para el cálculo del gradiente en el eje x asumiremos un ángulo de  $\frac{\pi}{2}$ . Para el resto de casos se tiene que los gradientes son:

- Desde (1,8) hasta (7,8):

$$\left( \frac{\partial I}{\partial x} = 0 - 150 = -150, \frac{\partial I}{\partial y} = 150 - 150 = 0 \right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{0}{-150}\right) = 0$$

- Desde (8,1) hasta (8,7) los píxeles con casillas negras debajo:

$$\left( \frac{\partial I}{\partial x} = 0 - 0 = 0, \frac{\partial I}{\partial y} = 0 - 150 = -150 \right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{-150}{0}\right) = \frac{\pi}{2}$$

- Desde (8,2) hasta (8,6) los píxeles con casillas negras debajo:

$$\left( \frac{\partial I}{\partial x} = 0 - 0 = 0, \frac{\partial I}{\partial y} = 255 - 150 = 105 \right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{105}{0}\right) = \frac{\pi}{2}$$

- Píxel en (8,8):

$$\left(\frac{\partial I}{\partial x} = 0 - 150 = -150, \frac{\partial I}{\partial y} = 255 - 150 = 105\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{105}{-150}\right) = \frac{7\pi}{36}$$

- Columnas negras que van desde las filas 1 a la 7 (de (1,9) a (7,9) y así las demás de la derecha):

$$\left(\frac{\partial I}{\partial x} = 255 - 0 = 255, \frac{\partial I}{\partial y} = 0 - 0 = 0\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{0}{255}\right) = 0$$

- Columnas blancas que van desde las filas 1 a la 7 (de (1,10) a (7,10) y así las demás de la derecha exceptuando la última):

$$\left(\frac{\partial I}{\partial x} = 0 - 255 = -255, \frac{\partial I}{\partial y} = 255 - 255 = 0\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{0}{-255}\right) = 0$$

- Desde (8,9) hasta (8,15) los píxeles negros:

$$\left(\frac{\partial I}{\partial x} = 255 - 0 = 255, \frac{\partial I}{\partial y} = 0 - 0 = 0\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{0}{255}\right) = 0$$

- Desde (8,10) hasta (8,14) los píxeles blancos:

$$\left(\frac{\partial I}{\partial x} = 0 - 255 = -255, \frac{\partial I}{\partial y} = 0 - 255 = -255\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{-255}{-255}\right) = \frac{\pi}{4}$$

- Píxeles negros en el rectángulo con esquinas en (9,1) y (15,8):

$$\left(\frac{\partial I}{\partial x} = 255 - 0 = 255, \frac{\partial I}{\partial y} = 255 - 0 = 255\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{255}{255}\right) = \frac{\pi}{4}$$

- Píxeles blancos en el rectángulo con esquinas en (9,1) y (15,8):

$$\left(\frac{\partial I}{\partial x} = 0 - 255 = -255, \frac{\partial I}{\partial y} = 0 - 255 = -255\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{-255}{-255}\right) = \frac{\pi}{4}$$

- Píxeles negros en el rectángulo con esquinas en (9,9) y (15,15):

$$\left(\frac{\partial I}{\partial x} = 0 - 0 = 0, \frac{\partial I}{\partial y} = 255 - 0 = 255\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{255}{0}\right) = \frac{\pi}{2}$$

- Píxeles blancos en el rectángulo con esquinas en (10,9) y (14,15):

$$\left(\frac{\partial I}{\partial x} = 255 - 255 = 0, \frac{\partial I}{\partial y} = 0 - 255 = -255\right) \Rightarrow \|\vec{\nabla} I\| > 0 \Rightarrow \theta(i, j) = \arctan\left(\frac{-255}{0}\right) = \frac{\pi}{2}$$

Finalmente con esto tenemos los siguientes bins para cada zona con sus respectivas frecuencias:

- Zona superior izquierda:

- 0: 7
- $\frac{\pi}{2}$ : 7
- $\frac{7\pi}{36}$ : 1

- Zona superior derecha:

- 0: 53
- $\frac{\pi}{4}$ : 3

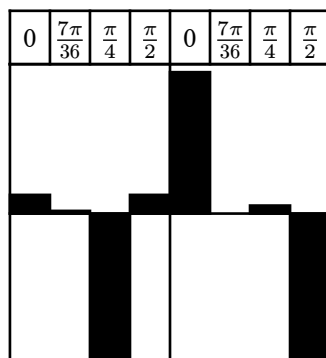
- Zona inferior izquierda:

- $\frac{\pi}{4}$ : 56

- Zona inferior derecha:

- $\frac{\pi}{2}$ : 56

Lo que finalmente nos entrega el siguiente histograma de orientaciones de gradiente (la parte superior es para identificar el bin):



La altura de cada bin en cada zona es como sigue:

- Zona superior derecha: Se normaliza por 15 y entonces la altura para 0 y  $\frac{\pi}{2}$  es  $\frac{7}{15}$  y para  $\frac{7\pi}{36}$  es  $\frac{1}{15}$ .
- Zona superior derecha: La suma de frecuencias es 56 así que se divide cada una por 56 por lo que obtenemos una altura de  $\frac{53}{56}$  para el 0 mientras que para  $\frac{\pi}{4}$  es  $\frac{3}{56}$ .
- Zonas de abajo: Para ambas zonas la altura para  $\frac{\pi}{4}$  es 1.

### c) Solución:

Primero definiremos un bin 0 para el caso en el que al calcular la respuesta a los filtros de orientación todos den el mismo resultado, luego se tendrían los siguientes resultados (cada zona de 2x2 corresponde a la que contiene sus 4 píxeles en la imagen original):

0	0	0	0	1	1	1	1
0	0	0	0	1	1	1	1
0	0	0	0	1	1	1	1
0	0	0	0	1	1	1	1
5	5	5	5	2	2	2	2
5	5	5	5	2	2	2	2
5	5	5	5	2	2	2	2
5	5	5	5	2	2	2	2

De modo que el histograma de bordes es el siguiente (la columna superior indica el valor de cada bin en el histograma original):

0	1	2	3	4	5	0	1	2	3	4	5

Cada cuadrante tiene un solo bin con frecuencia igual a 16, de modo que para normalizar dividimos cada uno por 16, y de esta forma la altura para sus respectivos bins es de 1.

### Pregunta 3

#### a) Solución:

Primero escribamos los descriptores "Signature" de cada imagen:

$$X_1 = \{(0.8, 0.2), (S_1, S_2)\}$$

$$S_1 = \begin{pmatrix} 250 \\ 200 \\ 0 \end{pmatrix} \quad S_2 = \begin{pmatrix} 100 \\ 150 \\ 150 \end{pmatrix}$$

$$X_2 = \{(0.3, 0.6, 0.1), (T_1, T_2, T_3)\}$$

$$T_1 = \begin{pmatrix} 50 \\ 50 \\ 250 \end{pmatrix} \quad T_2 = \begin{pmatrix} 200 \\ 200 \\ 250 \end{pmatrix} \quad T_3 = \begin{pmatrix} 250 \\ 0 \\ 0 \end{pmatrix}$$

Con esto, usando la distancia L1 en el espacio RGB sobre estos descriptores obtenemos la siguiente matriz de costos:

	$T_1$	$T_2$	$T_3$
$S_1$	600	300	200
$S_2$	250	250	450

#### b) Solución:

Ahora con esto obtenemos la matriz de flujos:

	0.3	0.6	0.1
0.8	0.1	0.6	0.1
0.2	0.2	0	0

#### c) Solución:

Finalmente podemos obtener la EMD entre ambos histogramas:

$$EMD(X_1, X_2) = \sum C_{ij} F_{ij} = 310$$