

Práctica 2 Servizos Multimedia

Autores:

Óscar Olveira Miniño

Eloy Calvo Gens

Cuestión 1.1: ¿Al hacer esta transformación de color, se produce alguna pérdida de información? ¿O por el contrario se genera nueva información al pasar de 1 a 3 imágenes?

Ao facer a transformación de RGB a YCbCr non se perde nin se xenera nova información, soamente hai un cambio de modelo de cor.

Cuestión 2.1: Si usamos el modo de subsampling 4:4:4, ¿Estaremos perdiendo algo de información? ¿Qué porcentaje de la imagen original?

Ao usar o subsampling 4:4:4, non estaremos efectuando ningún downsampling polo que non perdemos ningunha información e teremos o 100% da imaxe orixinal.

Cuestión 2.2: Si usamos el modo de subsampling 4:2:0, ¿Estaremos perdiendo algo de información? ¿Qué porcentaje de la imagen original?

En este caso, estamos perdendo información xa que estamos reduciendo a resolución vertical e horizontal. O porcentaxe de perda é do 50% xa que do total ($4+4+4=12$), con este subsampling temos 6 ($4+2+0$), polo que, queda $\frac{1}{2}$ da información orixinal.

Cuestión 3.1: ¿Qué puedes decir acerca de las dos imágenes obtenidas tras el submuestreo? ¿A qué es debido?

A imaxe do submuestreo das crominancias (CrCb) non presenta ningún tipo de cambio mentras que o submuestreo da luminancia (Y) mostrase con menos calidade. Isto débese a que o ollo humano é máis sensible a un cambio na luminancia que nas crominancias.

Cuestión 5.1: ¿Observas alguna diferencia entre los valores del bloque original y los valores recuperados mediante la IDCT? ¿Qué significa eso con respecto a la DCT?

Non hai ningunha diferenza entre eles. Isto significa que si desfecemos a DCT sin cuantificar, vamos recuperar sempre a imaxe orixinal sin perda.

Cuestión 6.1: Anota los valores de salida. A la vista de ellos, ¿Con qué nivel de cuantización se consiguen un mayor número de valores en la matriz ? ¿A qué es debido?

Valores da saída:

Valores no nulos en la DCT original: 64 Valores no nulos en la DCT cuantificada escala 1: 14 Valores no nulos en la DCT cuantificada escala 0.25: 33 Valores no nulos en la DCT cuantificada escala 4: 7

Según os resultados que nos deron, canto maior sea o factor de escalado, máis nulos ten a matriz. A matriz da imaxe orixinal divídese pola matriz de cuantificación, como esta última crece conforme maior sea o número do factor de cuantificación, obtemos una matriz resultante con valores cada vez máis pequenos (máis cercanos a cero).

Cuestión 6.2: ¿Por qué se vuelven nulos algunos valores de las matrices cuantizadas pero no otros?

Os valores nulos son o resultado da división mencionada anteriormente. Estes valores nulos foron escollidos para ser eliminados, xa que ao descartalos afórrase espazo.

Cuestión 7.1: ¿Los valores recuperados diferentes de cero, tienen el mismo valor que en la DCT original? ¿Por qué?

Na maioría dos casos si. Hai algúns píxeles que amosan un valor diferente debido ao redondeo, pero, como a división ten a propiedade conmutativa, na maioría dos casos pódense recuperar os valores orixinais.

Cuestión 8.1: A la vista de las dos imágenes, ¿Ves alguna similitudes entre ambas? ¿A qué es debido?

Tal e como se ve, o fondo da imaxe apenas ten valores parte xa que é a parte máis homoxénea mentras que na zona onde esta o paxaro hai valores, podendo distinguir ben a silueta (xa que é a parte da imaxe onde hai cambios de intensidade)

Cuestión 9.1: Comparando esta imagen con la del pájaro, ¿Qué diferencias observas en cuanto al número de coeficientes de la DCT no nulos? ¿Qué nos dice eso con respecto a la naturaleza (predominio de frecuencias altas o bajas) de cada imagen?

Na imaxe do paxaro, parece que hai unha maior cantidade de valores nulos. Algo similar ocorre en áreas da imaxe do libro, como por exemplo no marxe, onde tamén se observan valores nulos. Isto suxire que a imaxe do libro ten unha frecuencia xeral máis baixa, o que leva a descartar menos coeficientes ao procesala.

Na área do plumaxe do peito do paxaro ou nas ás, onde se deberían mostrar frecuencias altas, atopamos moitos valores nulos. Pola contra, na zona do texto, onde se esperaría unha menor frecuencia debido ao seu menor detalle visual, hai poucos valores nulos. Isto parece contradictorio respecto ao que se esperaría pola natureza das frecuencias nesas áreas.

Cuestión 10.1: ¿Se parecen las versiones cuantizadas entre sí? ¿Cual es la diferencia entre ellas? ¿A qué es debido?

As imáxenes parecense, pero hai certos cambios cando cambiamos de factor de cuantificación. Os bloques teñen valores que se dividen pola matriz de cuantización multiplicada polo factor de escala. Cando o factor de escala aumenta, a matriz de cuantización tamén aumenta, facendo que os valores divididos sexan máis pequenos. Se o resultado da división é moi pequeno, convértese en nulo. Por iso, a medida que o factor de escala crece, xéranse máis valores nulos no bloque entón diferencianse entre elas por este motivo.

Cuestión 12.1: Viendo la imagen en la que solo queda el coeficiente de DC tras la cuantificación, ¿Qué es lo que observas? ¿A qué es debido?

Este componente representa o valor medio de intensidade dun bloque de píxeles. Se eliminamos todos os valores e deixamos ese píxel específico, a imaxe resultante parecería ter un só píxel grande que contén o cor medio de todos os píxeles eliminados.

Cuestión 12.2: Observando las otras dos gráficas, ¿por qué se parece más al original la imagen con solo 3 coeficientes que la que usa 61? ¿Qué es lo que se estaría viendo realmente en esta última con respecto a las diferentes frecuencias que componen la imagen?

A clave encontrase no manexo das compoñentes de baixa frecuencia e DC durante a cuantización e reconstrución. Ao manter os coeficientes preservase as compoñentes de baixa frecuencia (DC incluído) polo que mantense información sobre a intensidade; aspecto máis perceptible ao ollo xa que captamos mellor cambios en zonas de baixas frecuencias, por eso neste caso apreciamos mellor a imaxe.

Na imaxe que usa 61 coeficientes haberíamos perdido información de DC e compoñentes de baixas frecuencias; polo que, a información que aparece representa as zonas con cambios bruscos de intensidade, é dicir, compoñentes de alta frecuencia. Por eso observamos o “contorno” do paxáro e a rama, ao ser o contorno (e partes da plumaxe) as zonas donde se procude o salto máis brusco de intensidade co fondo da imaxe.

Cuestión 13.1: ¿Cual es la versión con mayor calidad? ¿A qué se debe?

Según esta táboa (sacada dos arquivo Jupyter) establécese:

Ruido	Calidade
$20 < \text{pSNR(dB)}$	mala calidade
$20 \leq \text{pSNR(dB)} < 30$	calidade aceptable
$30 \leq \text{pSNR(dB)} < 40$	calidade buena
$\text{pSNR(dB)} \geq 40$	calidade muy buena

Resultado do exercicio:

pSNR(Y), escalado 1: 33.37 dB

pSNR(Y), escalado 0.25: 40.77 dB

pSNR(Y), escalado 4: 29.24 dB

Si comparamos a táboa de arriba cos resultados, vemos con escalado 0.25 obtemos a mellor calidade mentras que con escalado 4 obtemos a peor calidade. Esto débese a que con factor de cuantificación 0.25, comprimimos a imaxe con menos diferenzas que a imaxe orixinal polo que MSE tamén é menor.

Cuestión 14.1: ¿Qué versión consigue el menor tamaño? ¿A qué se debe esto?

Con escala 4 obtemos un tamaño menor xa que ten poucos valores non nulos, polo que ao comprimir, comprímense máis nulos que no resto de escalados.

Cuestión 15.1: Haz una tabla en el fichero de respuestas, y apunta para cada factor de escalado y modo de subsampling cual es el factor de compresión obtenido y el valor de pSNR(RGB).

Factor de escalado	Modo de subsampling	Factor obtenido	Valor de pSNR(RGB) en dB
1	4:4:4	4.53	35.49
0.25	4:4:4	2.94	38.84
4	4:4:4	6.11	32.63
1	4:2:2	6.47	35.03
0.25	4:2:2	3.94	37.32
4	4:2:2	8.90	32.44
1	4:2:0	8.07	34.70
0.25	4:2:0	4.77	36.32
4	4:2:0	11.56	32.29
1	4:1:1	8.18	34.38

Factor de escalado	Modo de subsampling	Factor obtenido	Valor de pSNR(RGB) en dB
0.25	4:1:1	4.86	36.09
4	4:1:1	11.62	32.08

Cuestión 15.2: A la vista de los resultados obtenidos, ¿Qué combinación consigue el mayor factor de compresión? ¿Y el menor? ¿Por qué crees que es así?

A combinación con maior factor de compresión é a que utiliza un subsampling de 4:1:1 e un factor de escalado de 4. A combinación co menor é que utiliza un subsampling de 4:4:4 e un factor de escalado de 0,25.

Débese a que o subsampling 4:1:1 reduce **á un cuarto da orixinal a resolución horizontal e vertical das crominancias** ademais utiliza o factor de escalado máis alto (4) ; polo que, ao realizar a cuantificación, será na que máis valores sean establecidos nulos (redondeados a 0). Por outro lado, o subsampling 4:4:4 supón que se **manteña a resolución orixinal** das crominancias e o factor de escalado 0,25 ao ser o menor dos utilizados é o que ao momento de realizar a cuantificación redondee menos valores a nulo.

Cuestión 15.3: De la misma forma, detalla en que combinación de parámetros se obtiene el mejor y el peor valor de pSNR(RGB), y cuales crees que son los motivos de ello.

O mellor valor de pSNR(RGB) prodúcese ca combinación dun downsampling de 4:4:4 e un factor de escalado de 0.25 (un valor de 38.84dB) e o peor valor de pSNR(RGB) prodúcese cun downsampling de 4:1:1 e un factor de escalado de 4.

Esto débese a cantidade de información con respecto a orixinal que se perde durante o proceso de codificación JPEG: canto menos se vexan alteradas as resolucións das crominancias (menor downsampling) e menor sea o factor de escalado (para producir menos 0 ao redondear na cuantificación) máis información se conservará da imaxe orixinal e por conseguinte mellor pSNR(RGB) (a cambio de reducir o factor de compresión).