EBU5303

Multimedia Fundamentals

JPEG

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

Reading

http://burg.cs.wfu.edu/TheScienceOfDigitalMedia/Chapter1/Ch1ScienceOfDigitalMedia.pdf

1.5.5 Transform Encoding

1.5.6 Compression Standards and Codecs

1.6 Standards and Standardisation Organisations for Digital Media

http://burg.cs.wfu.edu/TheScienceOfDigitalMedia/Chapter2/ch2scienceofdigitalmedia.pdf

2.3 Frequency in Digital Images

2.4 The Discrete Cosine Transform

http://burg.cs.wfu.edu/TheScienceOfDigitalMedia/Chapter3/Ch3ScienceOfDigitalMedia.pdf

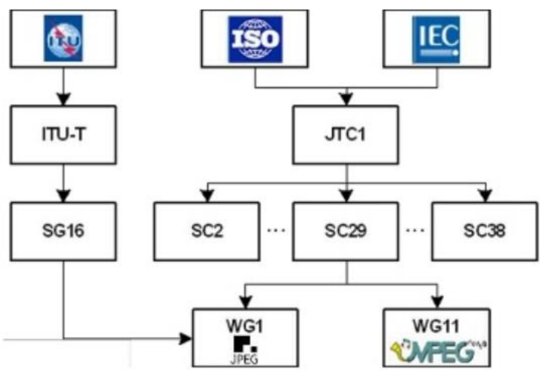
3.10.3 JPEG Compression

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

What is JPEG?

- JPEG = Joint Photographic Experts Group
- Standardised set of algorithms so that the various implementations all produce files in the same format.



- Joint Photographic Experts Group
  - ISO/IEC
  - ITU-T
- Informally known as JPEG
  - WG1 in official communications

5

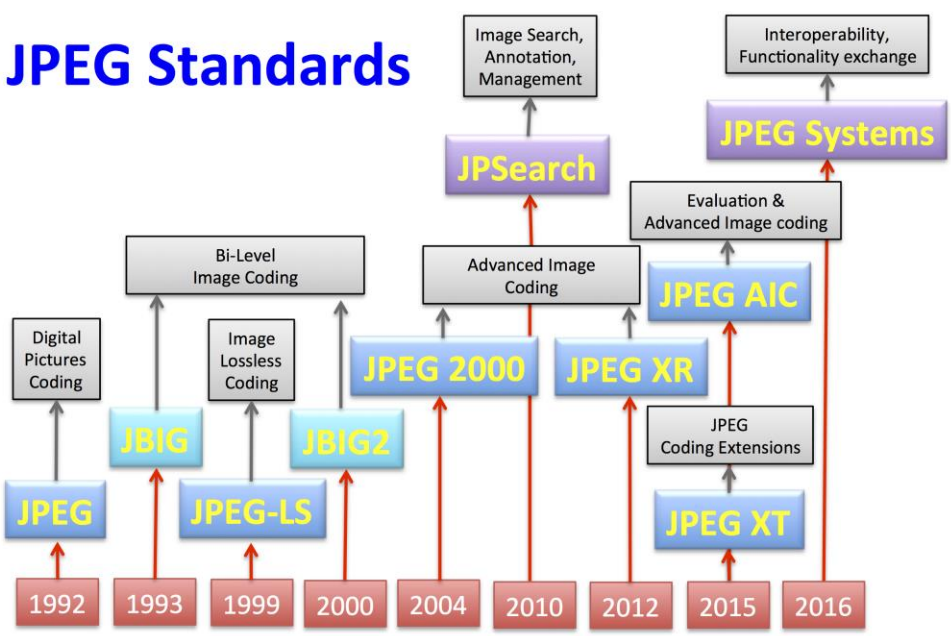
Standardisation Bodies

- *International Telecommunications Union* (*ITU*)
- *International Organization for Standardization* (*ISO*, pronounced “eye-so”),
- *International Electrotechnical Commission* (*IEC*).

6

JPEG Family of Standards

7



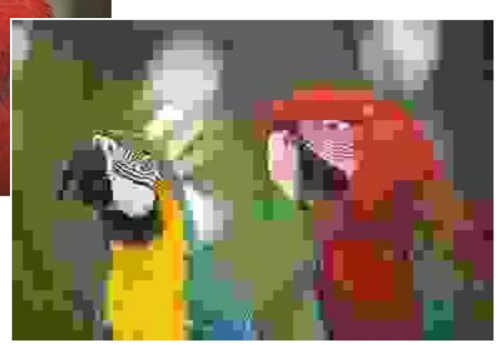
8

JPEG

- Standard for compressing and decompressing still images
- Based on DCT and applies to colour and grey-scaled still images
  – Independent of image sizes
  – Applicable to any aspect ratio (W/H)
  – Colour space and number of colours independent
  – Image content of any complexity, with any spatial and statistical properties
  – Run on many standard processors
  – Compression ratio dictated by user or application

JPEG is lossy!

Compression 10:1



Compression 100:1

Agenda

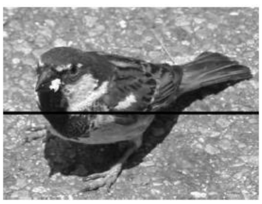
- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

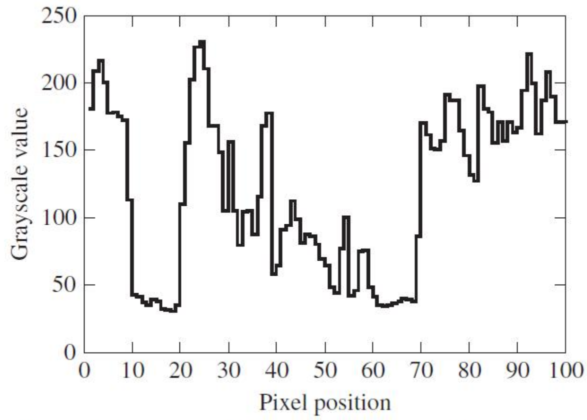
Transform Encoding

- The idea is that changing the representation (domain) of data can sometimes make it possible to **extract details** that won't be missed because they are **beyond the acuity of human perception**.
- Two of the most commonly used transforms in digital media are the ***discrete cosine transform*** (***DCT***) and the ***discrete Fourier transform*** (***DFT***).
- High frequency components correspond to quick fluctuations of colour in a short space—changes that aren't easy for the human eye to see.
- Once you have separated out the high frequency components of an image, you can remove them.

Frequency in Digital Images

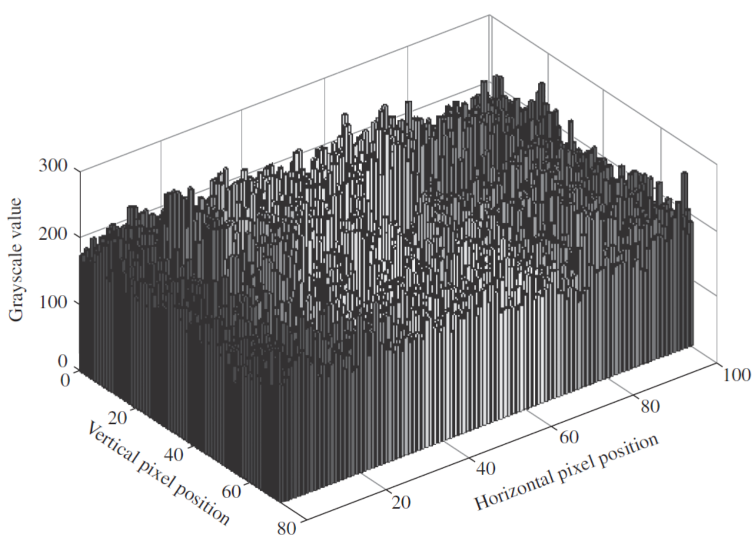
- For digital imaging, *frequency refers to the rate at which color values change*.
- All bitmap images can be viewed as waveforms.





13

Frequency in Digital Images



EBU5303

14

The Discrete Cosine Transform

- DCT tells how to transform an image from the spatial domain (i.e. pixel values), to the frequency domain (coefficients by which the frequency components should be multiplied).
- Rather than being applied to a full MxN image, the DCT is generally applied to 8x8 pixel blocks.

KEY EQUATION

Let $f(r, s)$ be the pixel value at row r and column s of a bitmap. $F(u, v)$ is the coefficient of the frequency component at (u, v), where $0 \leq r, u \leq M - 1$ and $0 \leq s, v \leq N - 1$.

$$F(u, v) = \sum_{r=0}^{M-1} \sum_{s=0}^{N-1} \frac{2C(u)C(v)}{\sqrt{MN}} f(r, s) \cos\left(\frac{(2r + 1)u\pi}{2M}\right) \cos\left(\frac{(2s + 1)v\pi}{2N}\right)$$

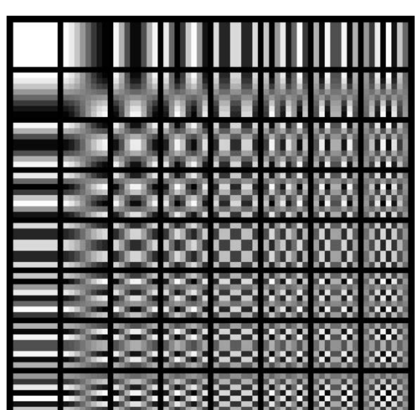
$$\text{where } C(\delta) = \frac{\sqrt{2}}{2} \text{ if } \delta = 0 \text{ otherwise } C(\delta) = 1$$

15

Base Frequencies

Each graph at position (u,v) represents the function:

$$\cos\left(\frac{(2r + 1)u\pi}{16}\right) \cos\left(\frac{(2s + 1)v\pi}{16}\right)$$



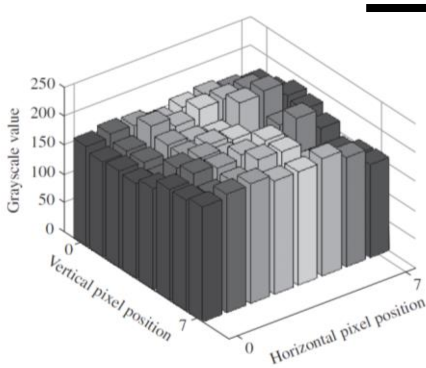
16

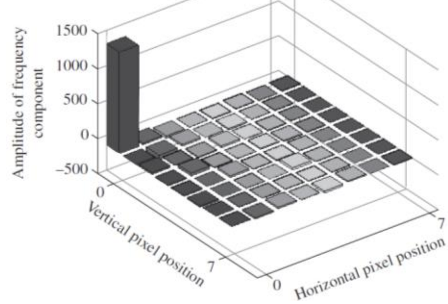
The Discrete Cosine Transform

- DCT takes a bitmap image in the form of a matrix of color values – ***f*** (*r*, *s*) – and returns the frequency components of the bitmap in the form of a matrix of coefficients – ***F***(*u*,*v*).
- The coefficients give the amplitudes of the frequency components.
- The first element ***F***(0,0)—is called the ***DC component***.
- All the other components are called ***AC components***
- A negative coefficient amounts to adding the inverted waveform.
- The discrete cosine transform is invertible.

The Discrete Cosine Transform

Spatial Domain **DCT** → **Frequency Domain**





← **Inverse DCT**

Spatial Domain

255	255	255	255	255	255	159	159
255	0	0	0	255	255	159	159
255	0	0	0	255	255	255	255
255	0	0	0	255	255	255	255
255	255	255	255	255	255	100	255
255	255	255	255	255	255	100	255
255	255	255	255	255	255	100	255
255	255	255	255	255	255	100	255

DCT ↓ ↑ **Inverse DCT**

1628	−61	39	234	173	−128	171	22
−205	−163	74	222	1	74	−30	111
81	150	−95	−82	−42	−11	−6	−53
188	231	−135	−188	−53	−36	−2	−103
96	71	−42	−78	−32	2	−17	−32
25	−42	25	3	−14	27	−26	19
70	15	−6	−51	−17	7	−16	−13
94	72	−38	−87	−18	−14	−4	−40

Frequency Domain

19

Steps of JPEG compression

algorithm jpeg
/*Input: A bitmap image in RGB mode.
Output: The same image, compressed.*/

{
(Convert image to Y'CbCr and do chroma sub-sampling)
Divide image into 8 x 8 pixel blocks

Use Discrete Cosine Transform (DCT) to transform the pixel data from the spatial domain to the frequency domain

Quantise frequency values

DPCM and zigzag order

Do run-length encoding

Do entropy encoding (e.g., Huffman)
}

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

Steps of JPEG compression

algorithm jpeg
/*Input: A bitmap image in RGB mode.
Output: The same image, compressed.*/

{

(Convert image to Y'CbCr and do chroma sub-sampling)
Divide image into 8 x 8 pixel blocks

Use Discrete Cosine Transform (DCT) to transform the pixel data from the spatial domain to the frequency domain

Quantise frequency values

DPCM and zigzag order

Do run-length encoding

Do entropy encoding (e.g., Huffman)
}

Colour Conversion: RGB to Y'CbCr

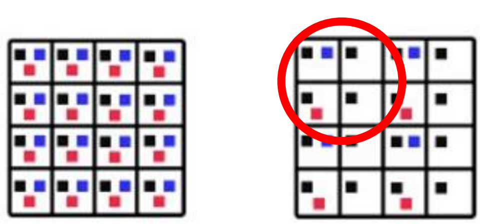
$$\begin{aligned}
Y' &= 0 + (0.299 \cdot R'_D) + (0.587 \cdot G'_D) + (0.114 \cdot B'_D) \\
C_B &= 128 - (0.168736 \cdot R'_D) - (0.331264 \cdot G'_D) + (0.5 \cdot B'_D) \\
C_R &= 128 + (0.5 \cdot R'_D) - (0.418688 \cdot G'_D) - (0.081312 \cdot B'_D)
\end{aligned}$$

And back:

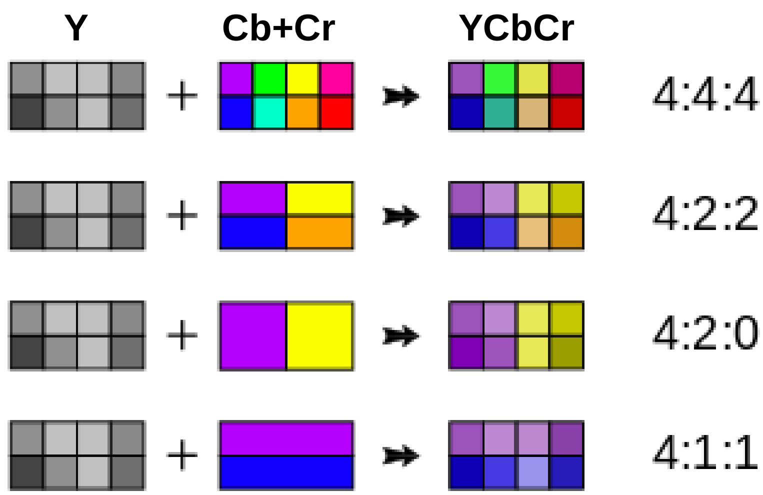
$$\begin{aligned}
R'_D &= Y' + 1.402 \cdot (C_R - 128) \\
G'_D &= Y' - 0.344136 \cdot (C_B - 128) - 0.714136 \cdot (C_R - 128) \\
B'_D &= Y' + 1.772 \cdot (C_B - 128)
\end{aligned}$$

Chroma Sub-Sampling

- The human eye is more sensitive to changes in light (*i.e.*, luminance) than in colour (*i.e.*, chrominance).
- ***Chroma sub-sampling*** (also called ***chrominance downsampling***) is a process of throwing away some of the bits used to represent colour information.
- For example, we might choose to save only one Cb value and one Cr value but four Y values for every four pixel values.



Chroma Sub-Sampling



Chroma Sub-Sampling

Y Cb,Cr	Y	Y	Y
Y Cb,Cr	Y	Y	Y
Y Cb,Cr	Y	Y	Y
Y Cb,Cr	Y	Y	Y

4:1:1

Y Cb,Cr	Y	Y Cb,Cr	Y
Y	Y	Y	Y
Y Cb,Cr	Y	Y Cb,Cr	Y
Y	Y	Y	Y

4:2:0

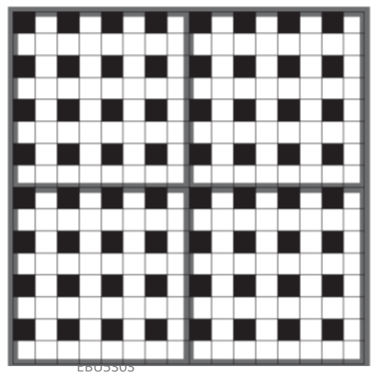
Y Cb,Cr	Y	Y Cb,Cr	Y
Y Cb,Cr	Y	Y Cb,Cr	Y
Y Cb,Cr	Y	Y Cb,Cr	Y
Y Cb,Cr	Y	Y Cb,Cr	Y

4:2:2

Blocks

- With YCbCr colour mode, we begin by dividing the image into 16x16 pixel macroblocks.
- Then, with 4 : 2 : 0 chroma downsampling, we get four blocks of 8x8 Y data for every one block of 8x8 Cb and one block of 8x8 Cr data.

Black cell: YCbCr
White cell: Y only

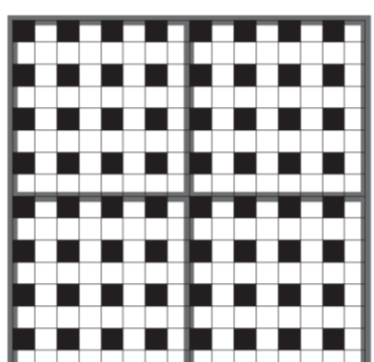


27

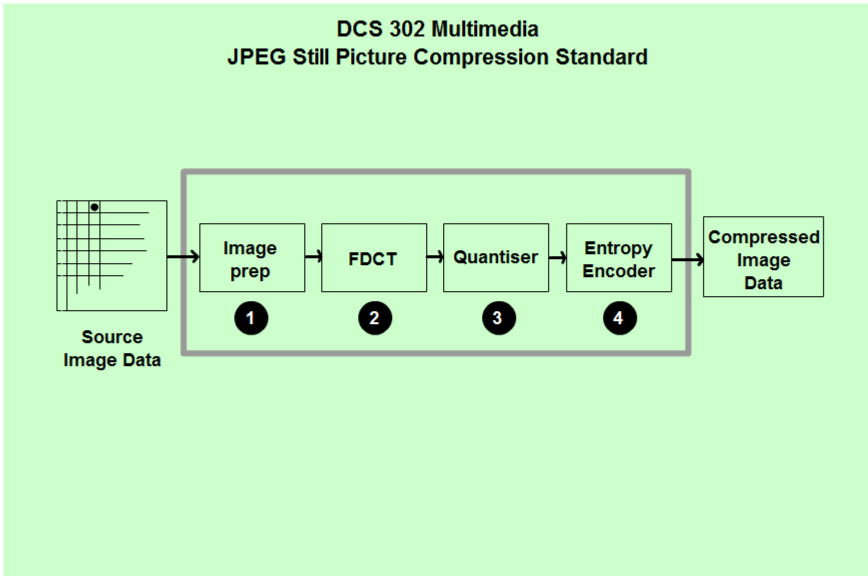
Exercise

With 4:2:0 YCbCr chrominance subsampling, a 16 x 16 macroblock yields: four 8 x 8 blocks of Y values; one 8 x 8 block of Cb values; and one 8 x 8 block of Cr values.

What is the compression rate achieved?



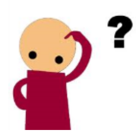
28



EBU5303

29

Question



Which of the following statements describes best what is chroma sub-sampling?

- It is a process where colours are converted to a different model.
- It is a process where some colour information is thrown away.
- It is a process to separate colour from brightness information.
- It is a process to eliminate some of the pixels.

EBU5303

30

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

Steps of JPEG compression

algorithm jpeg
/*Input: A bitmap image in RGB mode.
Output: The same image, compressed.*/

{
(Convert image to Y'CbCr and do chroma sub-sampling)
Divide image into 8 x 8 pixel blocks

Use Discrete Cosine Transform (DCT) to transform the pixel data from the spatial domain to the frequency domain

Quantise frequency values

DPCM and zigzag order

Do run-length encoding

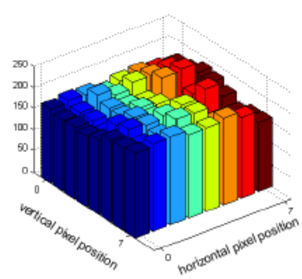
Do entropy encoding (e.g., Huffman)
}

Shifting Pixel Values

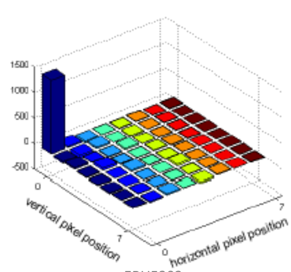
- On an intuitive level, shifting the values by –128 is like looking at the image function as a waveform that cycles through positive and negative values.
- This step is a preparation for representing the function in terms of its frequency components.

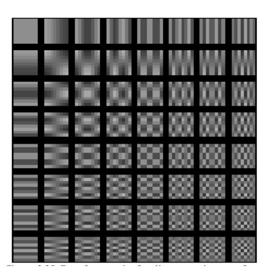
Discrete Cosine Transform (DCT)

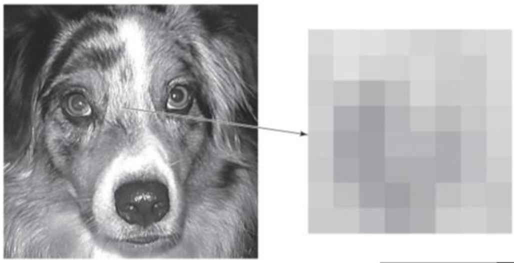
Pixel matrix (Y, Cb or Cr)



DCT coefficient matrix







Pixel values shifted by -128

94	103	101	96	88	85	92	96
88	101	89	87	93	82	81	95
83	74	55	70	90	79	81	93
86	52	36	60	75	65	77	89
81	43	38	62	62	50	71	87
78	49	38	51	52	50	71	82
84	69	45	38	51	70	78	75
80	80	67	46	56	82	86	78

DCT

585.7500	−24.5397	59.5959	21.0853	25.7500	−2.2393	−8.9907	1.8239
78.1982	12.4534	−32.6034	−19.4953	10.7193	−10.5910	−5.1086	−0.5523
57.1373	24.829	−7.5355	−13.3367	−45.0612	−10.0027	4.9142	−2.4993
−11.8655	6.9798	3.8993	−14.4061	8.5967	12.9151	−0.3122	−0.1844
5.2500	−1.7212	−1.0824	−3.2106	1.2500	9.3595	2.6131	1.1199
−5.9658	−4.0865	7.6451	13.0616	−1.1927	1.1782	−1.0733	−0.5631
−1.2074	−5.7729	−2.0858	−1.9347	1.6173	2.6671	−0.4645	0.6144
0.6362	−1.4059	−0.7191	1.6339	−0.1438	0.2755	−0.0268	−0.2255

35

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

EBU5303

36

Steps of JPEG compression

algorithm jpeg
/*Input: A bitmap image in RGB mode.
Output: The same image, compressed.*/

{

(Convert image to Y'CbCr and do chroma sub-sampling)
Divide image into 8 x 8 pixel blocks

Use Discrete Cosine Transform (DCT) to transform the pixel data from the spatial domain to the frequency domain

Quantise frequency values

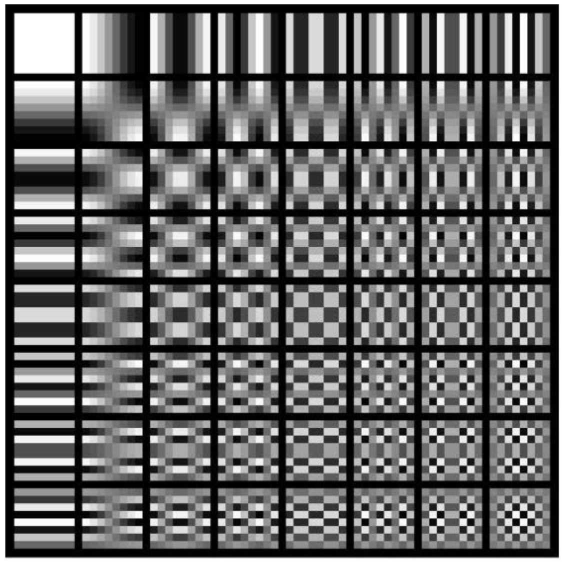
DPCM and zigzag order

Do run-length encoding

Do entropy encoding (e.g., Huffman)
}

DCT coefficients matrix

- Describes each 8x8 block in terms of how much the colour detail changes.
- Values are arranged from lowest frequencies to highest frequencies
  - Lowest frequencies represent average value for the block
  - Highest frequency represent fine detail (they can be dropped)
    - Eye unable to perceive brightness levels above or below thresholds
    - Gentle gradation of brightness of colour are more important to the eye than abrupt changes



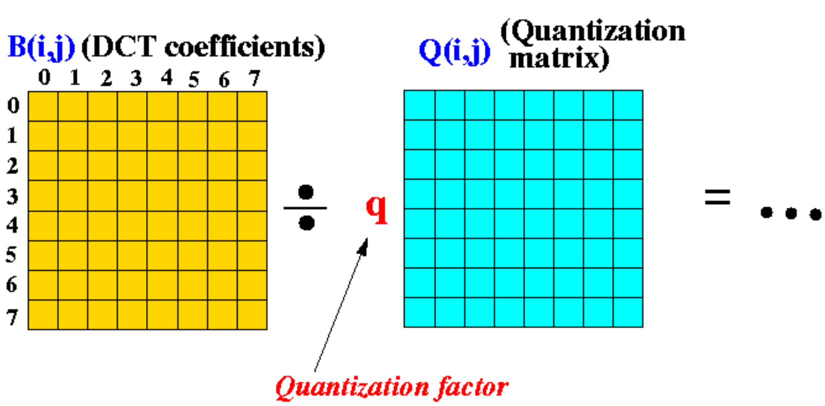
Quantisation = Division

- Quantisation involves dividing each frequency coefficient by an integer and rounding off.
- Rounding during quantisation makes JPEG lossy.
- The coefficients for high-frequency components are typically small, so they often round down to 0—which means, in effect, that they are thrown away.
- Not every value in the matrix needs to be divided by the same integer.
- A quantisation table is be stored with the compressed image

Quantisation = Division

$$F^{Q}(u,v) = Integer\ Round\left(\frac{F(u,v)}{Q(u,v)}\right)$$

Quantisation = Division



Quantisation table examples

a. Low compression

1	1	1	1	1	2	2	4
1	1	1	1	1	2	2	4
1	1	1	1	2	2	2	4
1	1	1	1	2	2	4	8
1	1	2	2	2	2	4	8
2	2	2	2	2	4	8	8
2	2	2	4	4	8	8	16
4	4	4	4	8	8	16	16

b. High compression

1	2	4	8	16	32	64	128
2	4	4	8	16	32	64	128
4	4	8	16	32	64	128	128
8	8	16	32	64	128	128	256
16	16	32	64	128	128	256	256
32	32	64	128	128	256	256	256
64	64	128	128	256	256	256	256
128	128	128	256	256	256	256	256

EBU5303

43

Matrix of coefficients

585.7500	−24.5397	59.5959	21.0853	25.7500	[illegible]93	−8.9907	1.8239
78.1982	12.4534	−32.6034	−19.4953	[illegible]	[illegible]	−5.1086	−0.5523
57.1373	24.829	−7.5355	−13.3367	[illegible]	10.0027	4.9142	−2.4993
−11.8655	6.9798	3.8993	[illegible]	[illegible]	12.9151	−0.3122	−0.1844
5.2500	−1.7212	[illegible]	[illegible]	1.2500	9.3595	2.6131	1.1199
−5.9658	−4.0865	[illegible]	[illegible]	−1.1927	1.1782	−1.0733	−0.5631
−1.2074	−5.7729	[illegible]	−1.9347	1.6173	2.6671	−0.4645	0.6144
0.6362	−1.4059	[illegible]	1.6339	−0.1438	0.2755	−0.0268	−0.2255

Divided by Quantisation Table

8	6	6	7	6	5	8	7
7	7	9	9	8	10	12	20
13	12	11	11	12	25	18	19
15	20	29	26	31	30	29	26
28	28	32	36	46	39	32	34
44	35	28	28	40	55	41	44
48	49	52	52	52	31	39	57
61	56	50	60	46	51	52	50

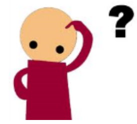
Equal

Compressed Matrix of Integer Coefficients

73	−4	10	3	4	0	−1	0
11	2	−4	−2	1	−1	0	0
4	2	−1	−1	−4	0	0	0
−1	0	0	−1	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

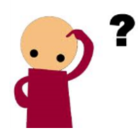
44

Questions



- What known limitations of the human visual system are used in JPEG to achieve compression?
- Is JPEG always lossy?
- In general, does JPEG achieve greater compression rates on colour images or on gray scale images?

Question



In JPEG, what is the effect of using quantisation tables that contain small values? Choose one answer.

- It achieves great compression and loses much quality.
- It achieves little compression and preserves much quality.
- It achieves great compression and preserves much quality.
- It achieves little compression and loses much quality.

Agenda

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values

Steps of JPEG compression

algorithm jpeg
/*Input: A bitmap image in RGB mode.
Output: The same image, compressed.*/

{
(Convert image to Y'CbCr and do chroma sub-sampling)
Divide image into 8 x 8 pixel blocks

Use Discrete Cosine Transform (DCT) to transform the pixel data from the spatial domain to the frequency domain

Quantise frequency values

DPCM and zigzag order

Do run-length encoding

Do entropy encoding (e.g., Huffman)
}

DPCM

- ***DPCM*** = ***differential pulse code modulation***
- DPCM is a compression technique that works by recording the difference between consecutive values rather than the actual values.
- It is effective in cases where consecutive values don’t change very much because fewer bits are needed to record the change than the value itself.
- In JPEG, the upper leftmost value in an block (the DC component) is stored as the difference from the DC component in the previous block.

Zigzag Ordering

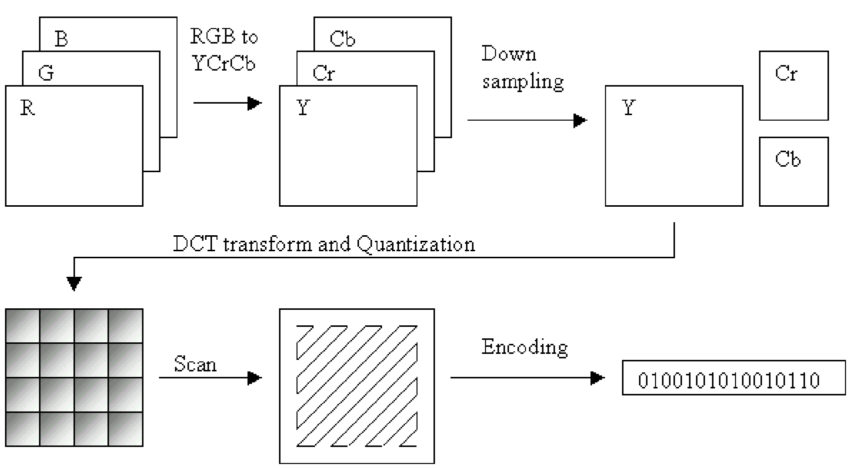
- The zigzag reordering sorts the values from low-frequency to high frequency components.
- The high-frequency coefficients are grouped together at the end.
- If many of them round to zero after quantisation, run-length encoding is even more effective.

Zigzag ordering

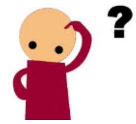
150	80	20	4	1	0	0	0
92	75	18	3	1	0	0	0
26	19	13	2	1	0	0	0
3	2	2	1	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

150	80	92	26	75	20	4	18	19	3	1	2	13	3	1	0	1	2
2	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0								

JPEG overview



Question



In JPEG, which of the following steps contribute the most to compression? Choose one answer.

- Dividing image into blocks.
- Colour conversion.
- Discrete Cosine Transform (DCT).
- DCT coefficients quantisation.

Some comparisons

Original size of color image (Kb)				313.076
Original size of B/W image (Kb)				104.437
	Color JPEG		B/W JPEG	
Quality Factor	File Size (Kb)	Comp. Ratio	File Size (Kb)	Comp. Ratio
75	23.039	13.59	21.02	4.97
20	8.457	37.02	7.599	13.74
5	4.009	78.09	3.257	32.07
3	3.268	95.80	2.522	41.41

Exercise

Consider the 4x4 block of pixels in Table 1, where:

r corresponds to the fully saturated bright red colour; g to the fully saturated bright green colour; b to the fully saturated bright blue colour; w to white; bk to black; and gy to mid-grey.

r	b	g	b
gy	r	gy	gy
gy	w	r	w
gy	bk	bk	r

Table 1

- Extract from the block of pixels shown in Table 1, the 4x4 matrix of values, comprised between 0 and 255, that correspond to the R channel.

Exercise

Consider that the matrix shown in Table 2 is a matrix of DCT coefficients.

- Apply quantisation to this matrix, using the quantisation matrix shown in Table 3.
- Arrange the quantised values in zig zag order and do Run Length encoding.
- To obtain a better quality image after compression, what would you change to the quantisation matrix of Table 3?

1000	800	100	40
600	400	30	10
150	80	10	3
20	10	5	2

Table 2

10	20	50	99
20	50	99	99
50	80	99	99
99	99	99	99

Table 3

JPEG encoding

- After these steps have been performed, the compressed file is put into a standardised format that can be recognised by the decompressor.
- A header contains global information such as: the type of file, the width and height, one or more quantizsation tables, Huffman code tables, and an indication of any pixel-padding.
- An alternative JPEG compression method is called ***JPEG2000***, noted for its high compression rate and good quality, without some of the blocky artifacts of standard JPEG. This method represents digital image data as wavelets as an alternative to the DCT.

Summary

- JPEG is lossy
- JPEG is an example of transform encoding
- Images are pre-processed: colour conversion, chroma-subsampling, 8x8 blocks
- Pixel data is converted to frequency coefficients
- Compression is done by quantisation of the frequency values
- Lossless encoding is also used to compress further the frequency values