TÉCNICO LISBOA

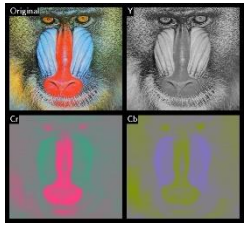
JPEG CODING STANDARD

Laboratory session

Fernando Pereira

Instituto Superior Técnico

DCT Based Image Coding: Let's Create Bits !



Spatial Redundancy

Quantization tables

Coding tables

Statistical Redundancy

PCM component samples

Block splitting

DCT

Quantization

Entropy coder

bits

8×8 samples block

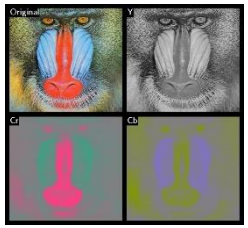
8×8 DCT coeffs

8×8 quantized DCT coeffs

Irrelevancy

Transmission or storage

≠



Quantization tables

Coding tables

Block assembling

IDCT

Inverse quantization

Entropy decoder

TÉCNICO LISBOA

Exploiting the Spatial Redundancy

Discrete Cosine Transform (DCT)

The DCT is one of the several sinusoidal transforms available; its basis functions correspond to discretized sinusoisal functions.

DCT Transform coefficients

Image block

$$F(u,v) = \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

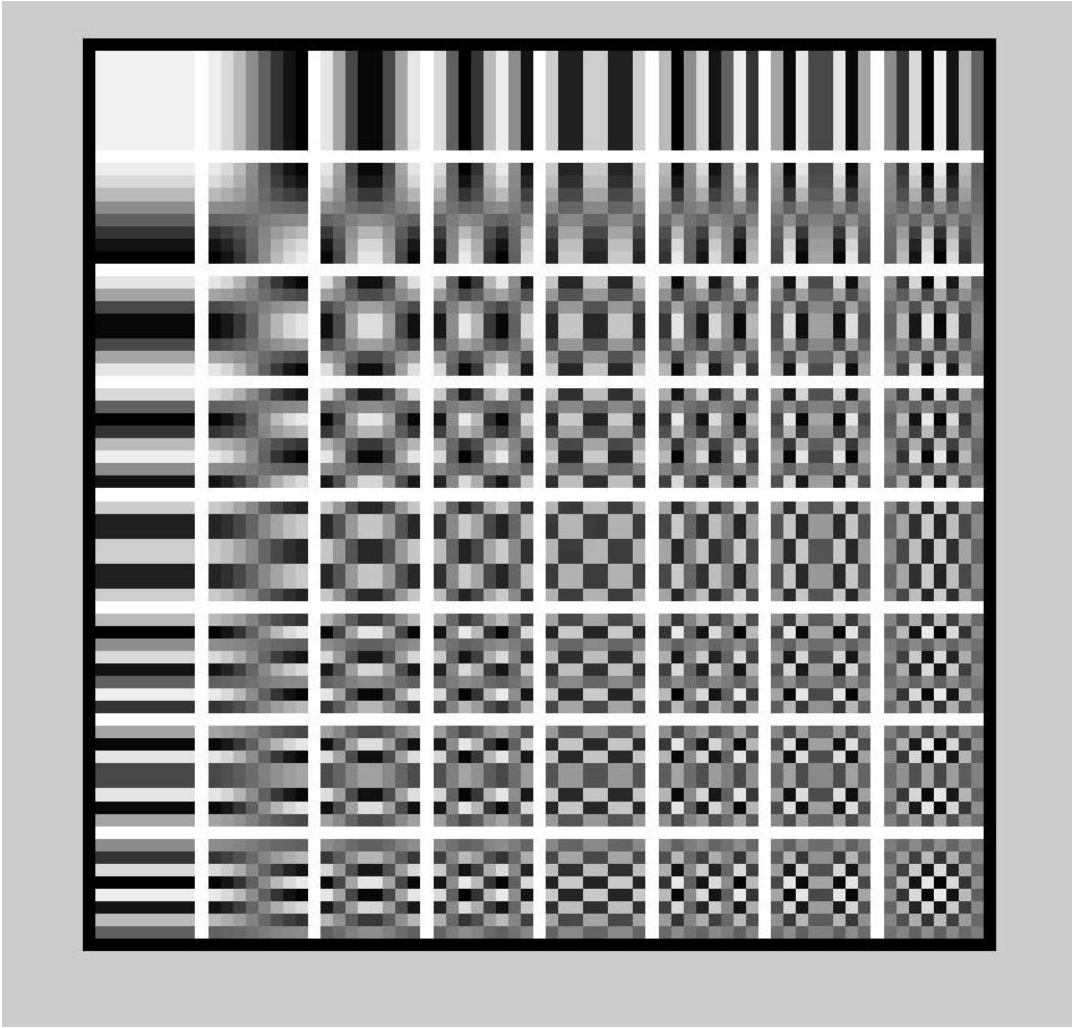
$$f(i,j) = \frac{2}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u,v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

Image block

DCT Transform coefficients

The DCT is the most used transform for image and video coding since its performance is close to the KLT performance for highly correlated signals; moreover, there are fast implementation algorithms available.

DCT Bidimensional Basis Functions (N=8)



You see here 64 8x8 sample blocks !

All existing and future image blocks can be rather efficienctly represented with these 64 (8×8) basic images !!!

TÉCNICO LISBOA

Luminance Samples, Y =

87	89	101	106	118	130	142	155
85	91	101	105	116	129	135	149
86	92	96	105	112	128	131	144
92	88	102	101	116	129	135	147
88	94	94	98	113	122	130	139
88	95	98	97	113	119	133	141
92	99	98	106	107	118	135	145
89	95	98	107	104	112	130	144

64 PCM samples are transformed into 64 DCT coefficients !

But more perceptual compression friendly !

DCT

DCT Coefficients =

898.0000	-149.5418	26.6464	-14.0897	0.7500	-5.7540	3.5750	0.0330
12.1982	-16.5235	-7.6122	5.2187	-0.2867	-1.9909	8.4265	1.2591
5.3355	-2.6557	2.3410	-9.9277	2.4614	4.4558	-3.1945	-3.1640
1.9463	-2.7271	1.5106	2.8421	-2.1336	-2.7203	-2.7510	5.4051
0.7500	-2.0745	0.8610	0.2085	2.5000	1.8446	2.0787	2.4750
7.9536	-2.6624	2.6308	0.4010	0.4772	3.3000	1.7394	0.3942
-4.1042	-0.1650	-0.6945	0.0601	0.0628	-0.7874	-0.8410	0.3496
-3.4688	2.3804	0.1559	0.8696	0.1142	-0.5240	-3.9974	-5.6187

How Does the DCT Work ?

Spatial Domain, samples

X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X

8×8×8=512 bits

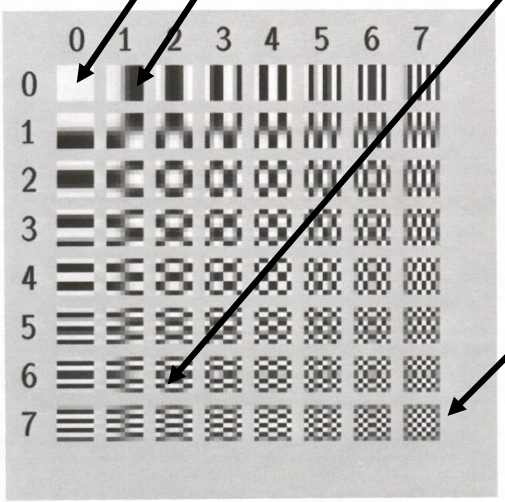
DCT

Frequency Domain, DCT coefficients

x	y		a				
C	f	d	B	c			
	H			k			
Y			i	p			
w	q	d					
	n	m					
							z

Low frequency, high energy

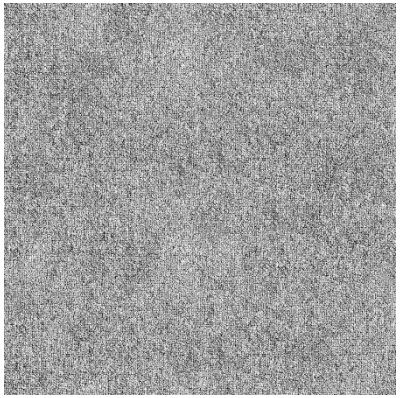
High frequency, low energy



Easy/Cheap versus Though/Expensive Blocks

8×8 samples

8×8 samples



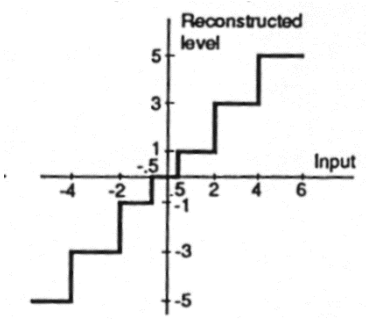
8×8 samples

All blocks above have the same price (8×8×8)=512 bits in the PCM/spatial domain because redundancy is not exploited !

In the DCT/frequency domain, simpler blocks will be cheaper and vice-versa because 'information' is bought with more DCT coefficients and associated rate.

Exploiting the Perceptual Irrelevance

How Does DCT Coding Work ?



Samples (spatial domain) s_{ij}

DCT

DCT Coefficients Sij

Quantization

Round (S/Q)

Level for Quantized coefficients Sqij

Transmission or storage

≠ ≠

Quantization tables Qij

Dec. samples (spatial domain) r_{ij}

IDCT

Reconstructed DCT coefficients Rij

e.g. R = Sq*Q

Level for Quantized coefficients Sqij

Inverse quantization

Quantization Matrices

For transparent quality, JPEG suggests to quantize the DCT coefficients using the values for the 'minimum perceptual difference' (for each coefficient) multiplied by 2; for more compression, a multiple of them may be used.

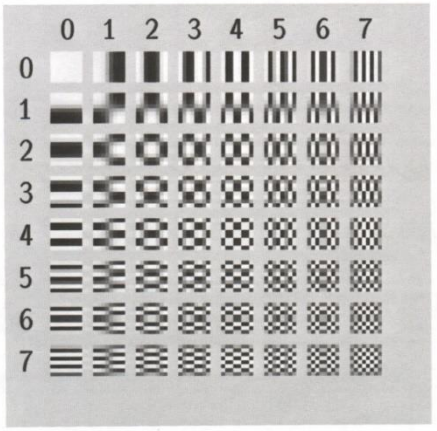
The quantization matrixes have to be always transmitted or at least signalled.

Luminance

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Chrominances

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99



Situation: Luminance and crominance with 2:1 horizontal subsampling; samples with 8 bits (*Lohscheller*)

From DCT Coeffs to Quantized DCT Coeffs

898.0000	-149.5418	26.6464	-14.0897	0.7500	-5.7540	3.5750	0.0330
12.1982	-16.5235	-7.6122	5.2187	-0.2867	-1.9909	8.4265	1.2591
5.3355	-2.6557	2.3410	-9.9277	2.4614	4.4558	-3.1945	-3.1640
1.9463	-2.7271	1.5106	2.8421	-2.1336	-2.7203	-2.7510	5.4051
0.7500	-2.0745	0.8610	0.2085	2.5000	1.8446	2.0787	2.4750
7.9536	-2.6624	2.6308	0.4010	0.4772	3.3000	1.7394	0.3942
-4.1042	-0.1650	-0.6945	0.0601	0.0628	-0.7874	-0.8410	0.3496
-3.4688	2.3804	0.1559	0.8696	0.1142	-0.5240	-3.9974	-5.6187

Quantizing with selected quantization matrix …

56	-14	3	-1	0	0	0	0
1	-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	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

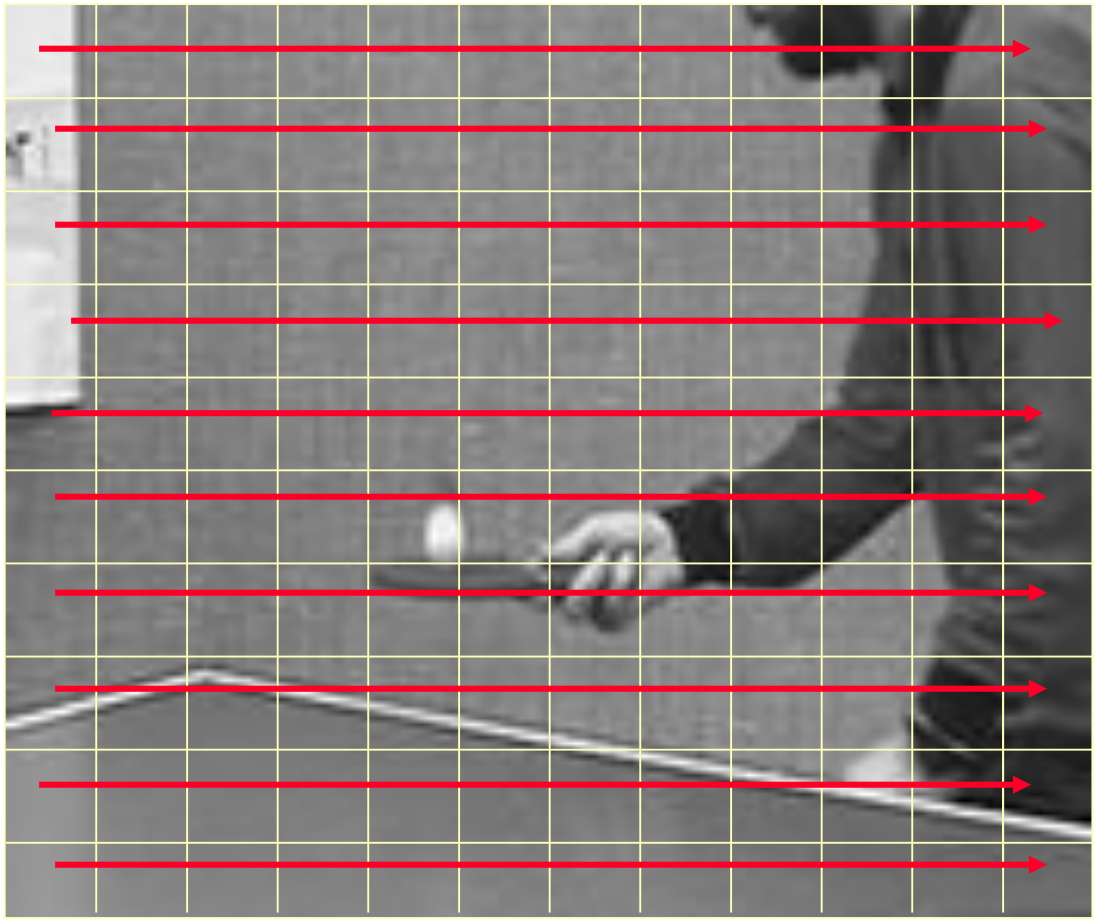
The rate is reduced, eventually at no quality cost !

Finally, the waited miracle !

TÉCNICO LISBOA

The Coding Modes

The JPEG Sequencial Mode: one Single Scan



The image is made available at a single resolution, at a single quality, in a single scan !

No flexibility to serve other 'clients' with different needs in terms of resolution and quality …

JPEG Operational Modes

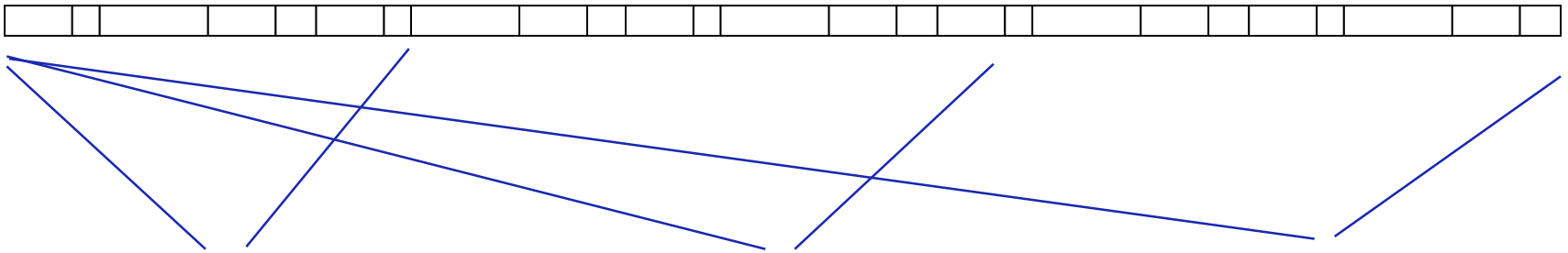
The various JPEG operational modes address the need to provide solutions for a large range of applications with different requirements.

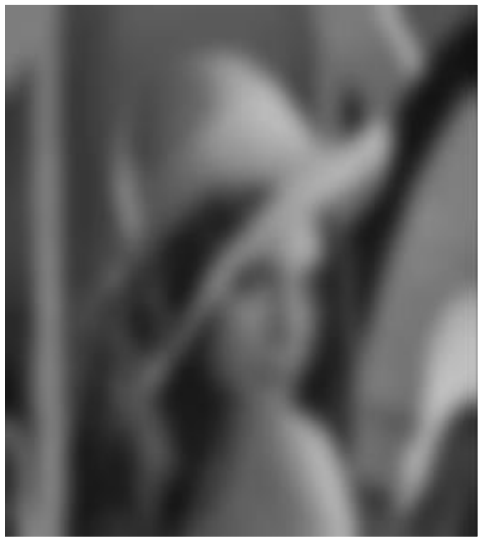
- **SEQUENTIAL MODE – Each image component is coded in a single scan (from top to bottom and left to right).**
- **PROGRESSIVE MODE (scalable) - The image is coded with several scans which offer a successively better quality (but same spatial resolution).**
- **HIERARCHICAL MODE (scalable) - The image is coded in several resolutions exploiting their mutual dependencies, with lower resolution images available without decoding higher resolution images.**
- **LOSSLESS MODE – This mode guarantees the exact reconstruction of each sample in the original image (mathematical equality).**

For each operation mode, one or more codecs are specified; these codecs are different in the sample precision (bit/sample) or the entropy coding method.

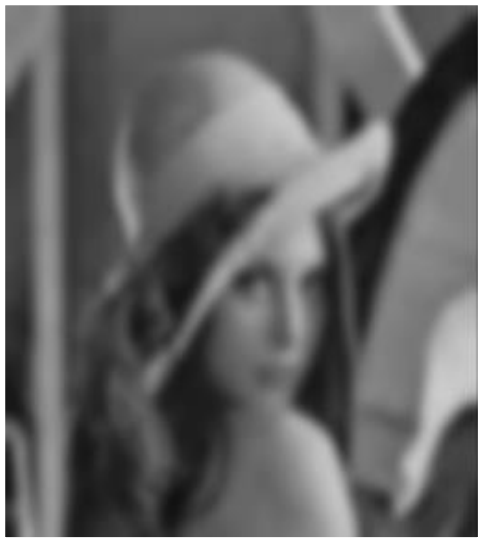
Progressively More Quality: Quality or SNR Scalability

Scalable stream





Decoding 1



Decoding 2

Decoding 3

JPEG Progressive Mode

The image is coded with successive scans. The first scan gives very quickly an idea about the image content; after, the quality of the decoded image is progressively improved with the successive scans, i.e. quality layers.

The implementation of the progressive mode requires a memory with the size of the image to store the quantized DCT coefficients (11 bits for the baseline process) which will be partially coded with each scan.

There are two methods of implementing the progressive mode:

- **SPECTRAL SELECTION – Only a specified 'zone' of the DCT coefficients is coded in each scan (going from lower to higher frequencies)**
- **GROWING PRECISION – DCT coefficients are coded with successively higher precision, bitplane by bitplane**

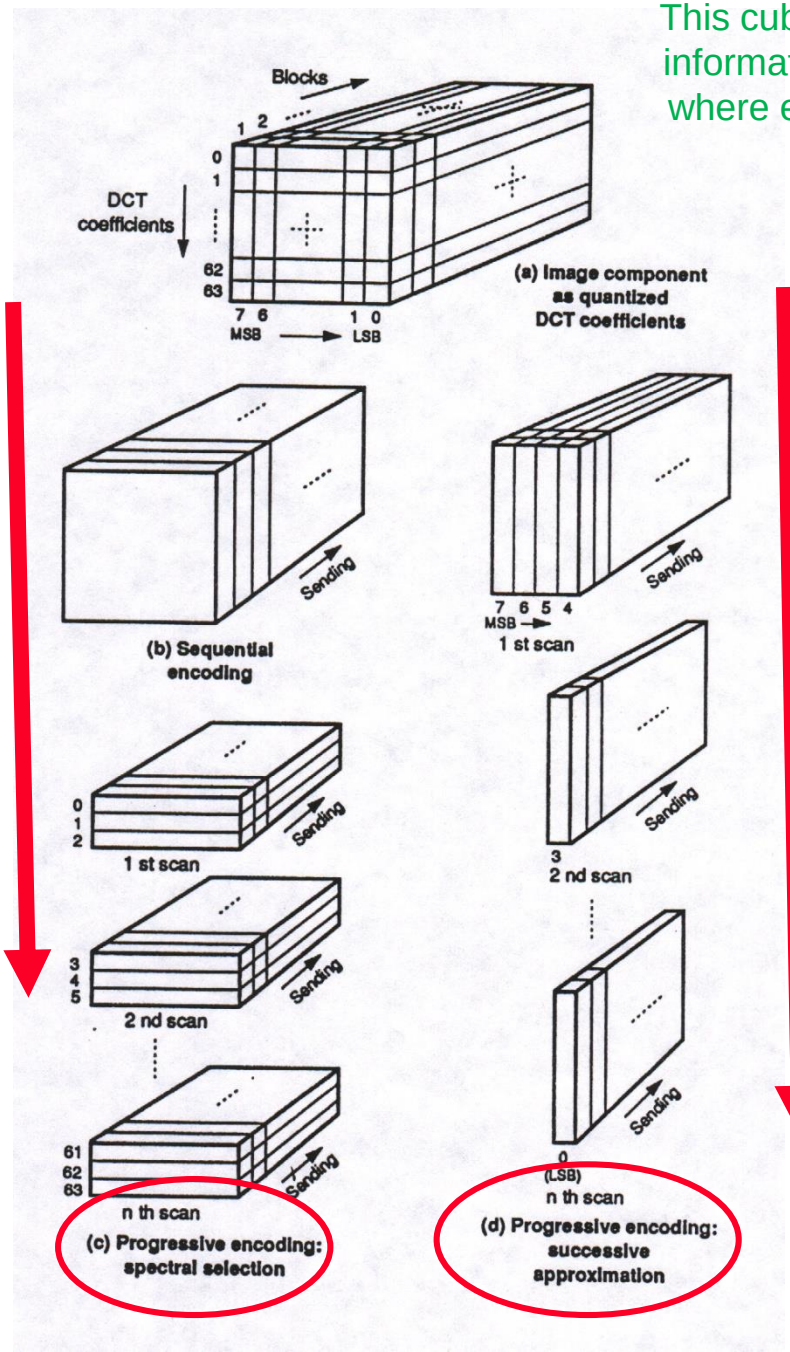
The spectral selection and successive approximations methods may be applied separately or combined.

TÉCNICO LISBOA

This cuboid includes all (quantized) information representing the image where each slice corresponds to a 8×8 block.

Progressive Modes:

Spectral Selection and Growing Precision



Spectral selection: Each layer brings an increasing number of DCT coefficients, and thus frequencies

Very 'natural' quality improvement process

Successive approximation: Each layer brings an increasing (mathematical) precision for all coefficients

Very 'non-natural' quality improvement process

Growing Precision, Bit by Bit …

197 -> 1100 0101

197 =

128 (MSB) **= 128**

+ 64 **= 192**

+ 0

+ 0

+ 0

+ 4 **= 196**

+ 0

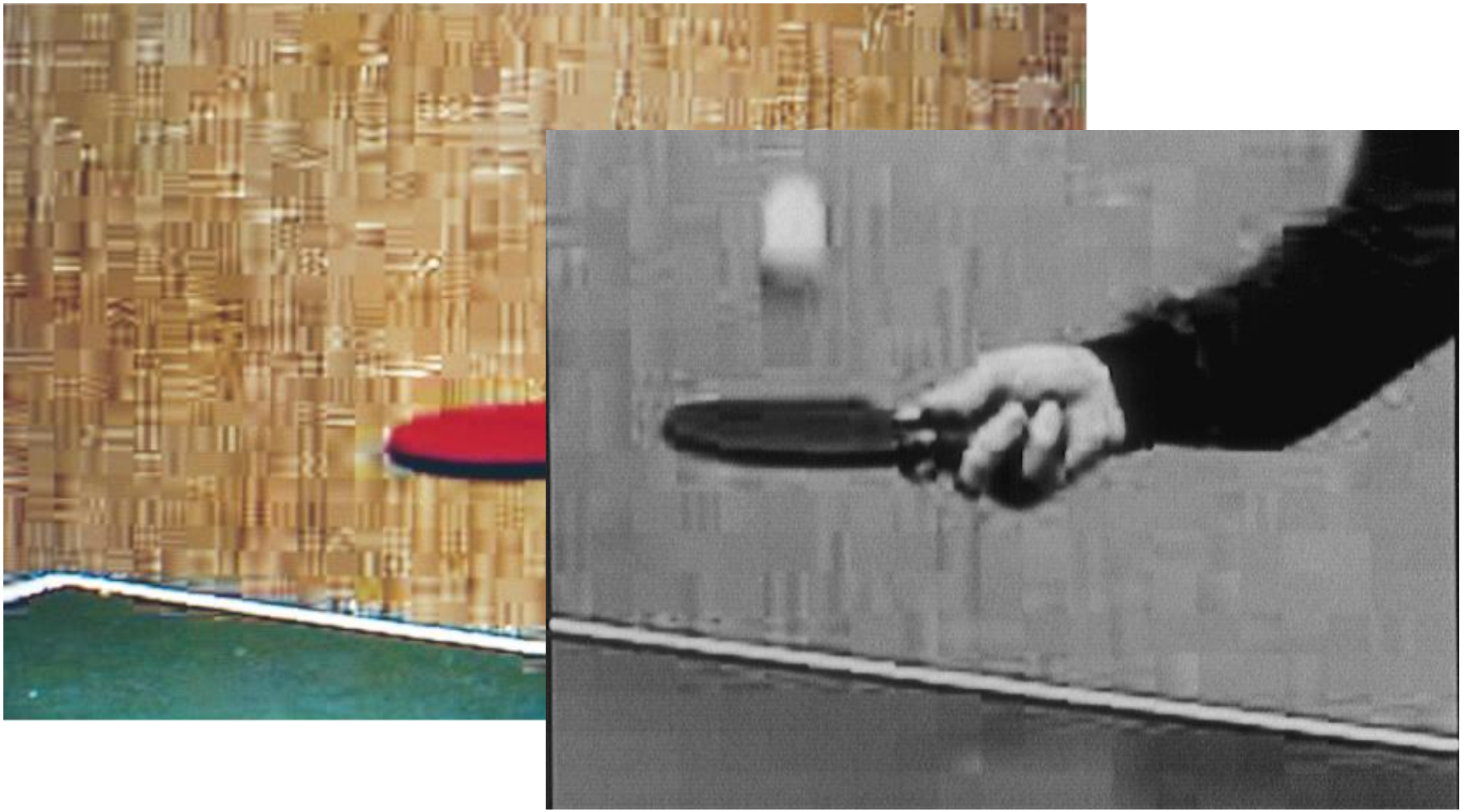
+ 1 (LSB) **= 197**

TÉCNICO LISBOA

Quality Assessment

Visual (Subjective) Inspection

The Block Effect …



Objective Quality Assessment

Compression

X(m,n) Y(m,n)

Original/reference

Decoded

Objective evaluation

x and y are the original and decoded data

$$\mathrm{PSNR(dB)} = 10\log_{10}\frac{255^2}{\mathrm{MSE}}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}}\sum_{i=1}^{M}\sum_{j=1}^{N}(y_{ij} - x_{ij})^2$$

There are other objective quality metrics !

$$SNR = 10\log_{10}\sum_{i=1}^{M}\sum_{j=1}^{N}x^2{}_{ij} \Big/ \sum_{i=1}^{M}\sum_{j=1}^{N}(x_{ij} - y_{ij})^2$$

TÉCNICO LISBOA