EE421/521 Image Processing

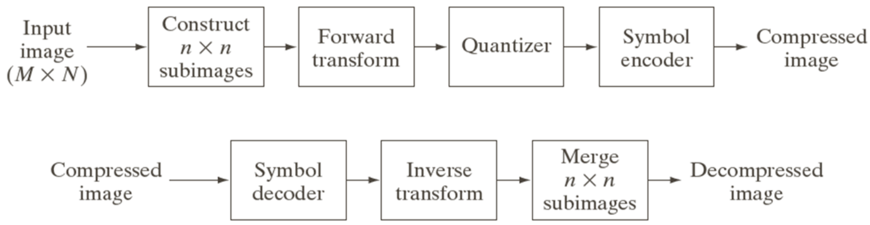
Lecture 12b
LOSSY IMAGE COMPRESSION

1

Transform Coding

2

Transform Coding



3

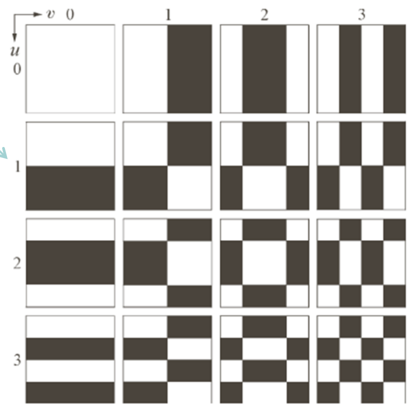
Walsh-Hadamard Transform

Basis functions

$$f(x,y) = \sum_u \sum_v F(u,v) r_{u,v}(x,y)$$

4x4 image block

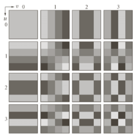
Transform coefficients



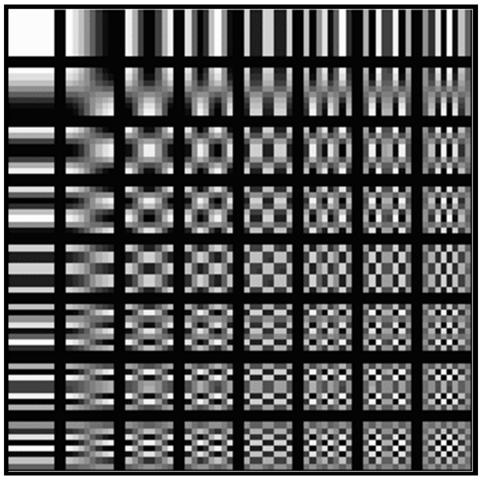
16 basis functions for n=4

4

Discrete Cosine Transform



16 basis functions for n=4
(image divided into 4x4 blocks)



64 basis functions for n=8
(image divided into 8x8 blocks)

8x8 Discrete Cosine Transform

Adopted by the international standards (type-II)

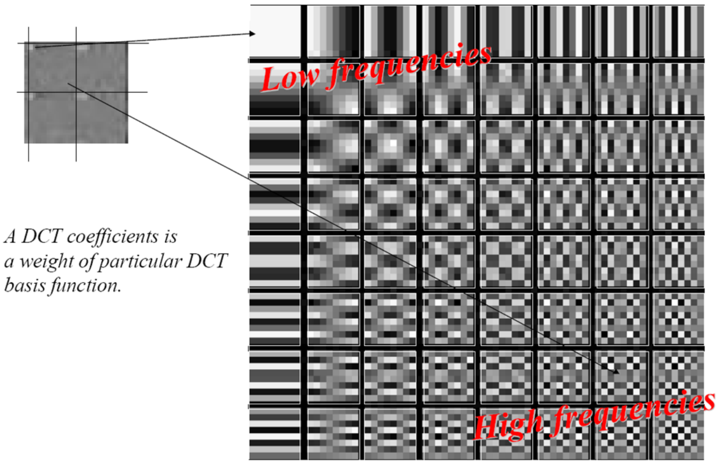
$$F(u,v)=\frac{1}{4}C(u)C(v)\sum_{m=0}^{7}\sum_{n=0}^{7}f(m,n)\cdot\cos\left(\frac{(2m+1)u\pi}{16}\right)\cos\left(\frac{(2n+1)v\pi}{16}\right)$$

$$f(m,n)=\frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7}C(u)C(v)F(u,v)\cdot\cos\left(\frac{(2m+1)u\pi}{16}\right)\cos\left(\frac{(2n+1)v\pi}{16}\right)$$

where $C(u), C(v)=1/\sqrt{2}$, for $u,v=0$;
$C(u), C(v)=1$, otherwise.

6

DCT Coefficients



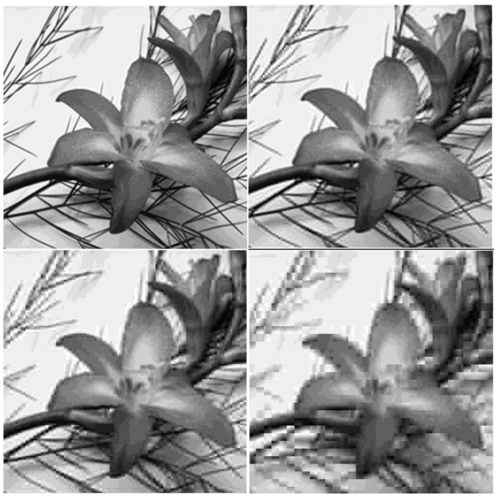
A DCT coefficients is a weight of particular DCT basis function.

7

Approximating an Image with a Subset of DCT Coefficients

Original

With 16/64 Coefficients

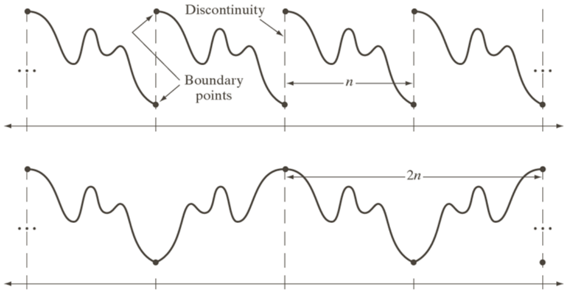


With 8/64 Coefficients

With 4/64 Coefficients

8

DFT vs. DCT



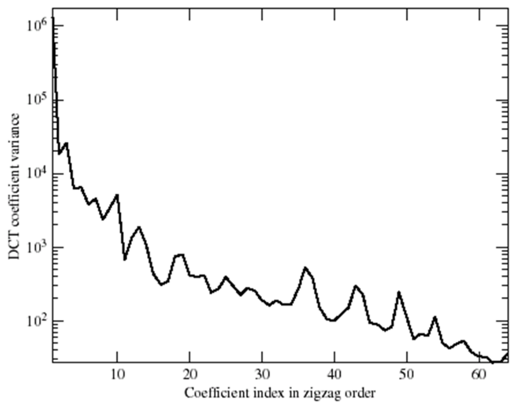
a
b

FIGURE 8.25 The periodicity implicit in the 1-D (a) DFT and (b) DCT.

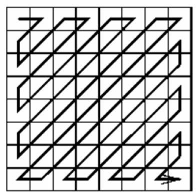
- DCT provides near optimal energy compaction (true also for DFT)
- Fast implementation available for DCT (true also for DFT)
- DCT does not introduce spurious frequencies (not true for DFT)
- DCT coefficients are real (DCT coefficients are complex)

9

Typical Energy Distribution of DCT Coefficients

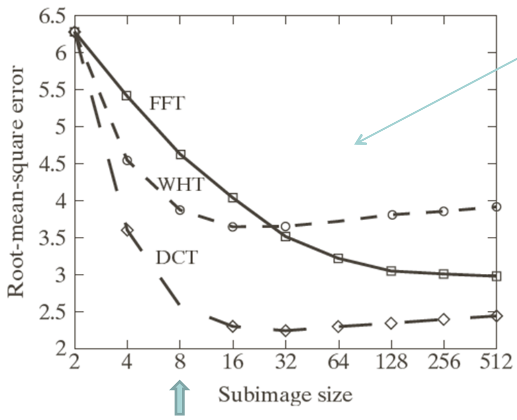


Zig-zag scan



10

Selection of Block Size



A good choice for block size

Only the first 25% of the transform coefficients are retained

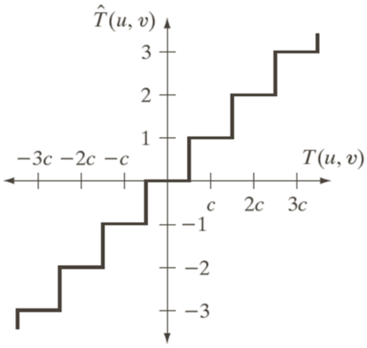
−76	−73	−67	−62	−58	−67	−64	−55
−65	−69	−62	−38	−19	−43	−59	−56
−66	−69	−60	−15	16	−24	−62	−55
−65	−70	−57	−6	26	−22	−58	−59
−61	−67	−60	−24	−2	−40	−60	−58
−49	−63	−68	−58	−51	−65	−70	−53
−43	−57	−64	−69	−73	−67	−63	−45
−41	−49	−59	−60	−63	−52	−50	−34

−415	−29	−62	25	55	−20	−1	3
7	−21	−62	9	11	−7	−6	6
−46	8	77	−25	−30	10	7	−5
−50	13	35	−15	−9	6	0	3
11	−8	−13	−2	−1	1	−4	1
−10	1	3	−3	−1	0	2	−1
−4	−1	2	−1	2	−3	1	−2
−1	−1	−1	−2	−1	−1	0	−1

Quantization

Quantization

- Quantization of transform coefficients is needed to achieve high rates of compression



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

13

Quantization Error

- For a memoriless Gaussian scalar source, the distortion D (mean squared error) caused by quantization of its symbols having variance σ^2 into r bits is given by:

$$D = \left(\frac{\sigma}{N}\right)^2$$

$$r = \log_2 N$$ ← Total number of quantization levels

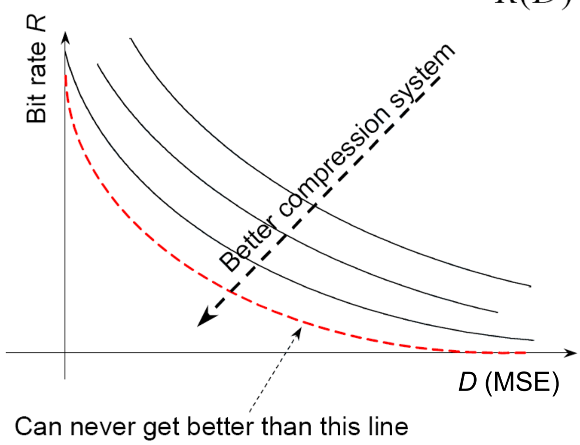
$$D = \sigma^2 2^{-2r}$$

14

Bit Rate vs. Distortion

$$R(D) = \frac{1}{2}\log_2 \frac{\sigma^2}{D}$$

$R(0) = H$



Bit Allocation Problem

Suppose $\mathbf{X} = (X_1, \cdots, X_M)$ is a random vector with zero mean and variances $\sigma_1^2, \cdots, \sigma_M^2$. Let D_m denote the distortion incurred by quantizing X_m using a uniform quantizer that has resolution r_m.

The problem of bit allocation is to find r_m, $m = 1, \cdots, M$ to minimize the overall distortion $D = \sum_{m=1}^{M} D_m$ subject to the constraint that $\sum_{m=1}^{M} r_m \leq R$.

139	144	149	153	155	155	155	155
144	151	153	156	159	156	156	156
150	155	160	163	158	156	156	156
159	161	162	160	159	159	158	159
159	160	161	162	162	155	155	155
161	161	161	161	160	157	157	157
162	162	161	163	162	157	157	157
162	162	161	161	163	158	158	158

Image Block (Lena)

1260	-1	-12	-5	2	-2	-3	-1
-23	-17	-6	-3	-3	0	0	-1
-11	-9	-2	2	0	-1	-1	0
-7	-2	0	1	1	0	0	0
-1	-1	1	2	0	-1	1	1
2	0	2	0	-1	1	1	-1
-1	0	0	-1	0	2	1	-1
-3	2	-4	-2	2	1	-1	0

NINT[DCT Block]

Space domain or DCT domain quantization?

Optimal Bit Allocation

$$r = \frac{1}{2}\log_2 \frac{\sigma^2}{D} \Rightarrow D = \sigma^2 2^{-2r}$$

Less distortion if smaller geometric mean for variances

We have

$$D = \sum_{m=1}^{M} D_m = \alpha \sum_{m=1}^{M} \sigma_m^2 2^{-2r_m} \geq \alpha M \left(\prod_{m=1}^{M} \sigma_m^2 2^{-2r_m} \right)^{\frac{1}{M}} = \alpha M 2^{-2\bar{r}} \left(\prod_{m=1}^{M} \sigma_m^2 \right)^{\frac{1}{M}}$$

with equality if and only if (iff)

$$\sigma_m^2 2^{-2r_m} = \text{constant} = C \quad \text{for all } m = 1, \cdots, M.$$

$$r_m = \log \sigma_m - \frac{1}{2} \log C.$$

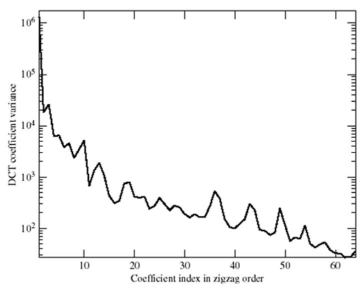
$$\frac{1}{M} \sum_{m=1}^{M} r_m \leq \frac{1}{M} R = \bar{r}$$

More bits allocated if higher variance

17

Better to Quantize in the DCT Domain

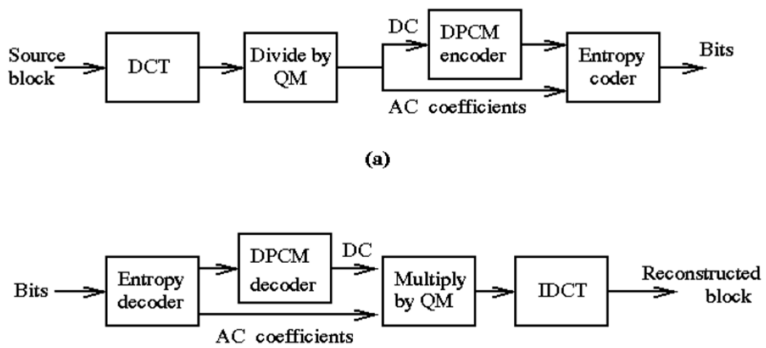
- The **arithmetic mean** of space domain and DCT domain variances are the same due to **Parseval's relation**
- Space domain variances are nearly the same for all pixels
- **DCT provides energy compaction** which implies smaller geometric mean for the variances of DCT coefficients



18

JPEG Algorithm

JPEG Coding/Decoding of DCT Coefficients

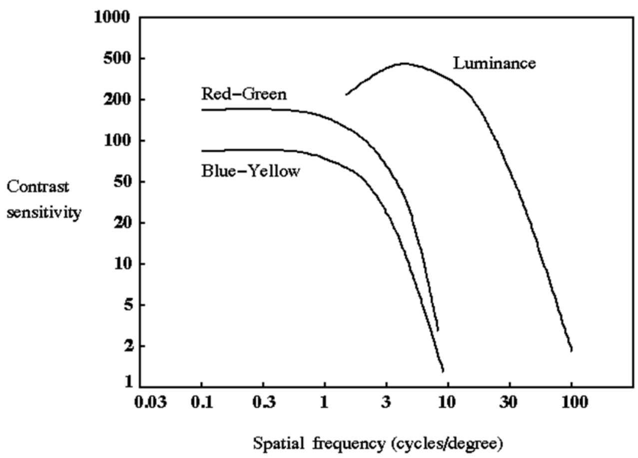


Quantization Matrix

- **Flat Quantization:** A single step size (threshold) may be used for all DCT coefficients.
- **HVS Weighted Quantization:** The thresholds vary by frequency according to human visual system response. A single threshold matrix is used for all blocks.
- **Adaptive Quantization:** The quantization matrix may be allowed to change from block to block by simple scaling. The scale parameter is called "mquant."
  - In all cases, the location of the retained coefficients vary from block to block.

21

Sensitivity of the Eye to Luminance Chrominance Variations



22

JPEG Quantization Tables

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

Luminance Quantization Table

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

Chrominance Quantization Table

JPEG Algorithm

- *DCT computation: The image is divided into 8 × 8 blocks.* The 2-D DCT of blocks are computed after level shifting.
- *Quantization of DCT coefficients and zig-zag scanning:* The default quantization chroma and luma matrices can be scaled to provide various compression levels.
- *Variable-length coding (VLC): The AC coefficients are* coded using a VLC code that defines the coefficient's value and the number of preceding zeros. Standard VLC tables are specified. The DC coefficients are DPCM coded.

JPEG

52	55	61	66	70	61	64	73
63	59	66	90	109	85	69	72
62	59	68	113	144	104	66	73
63	58	71	122	154	106	70	69
67	61	68	104	126	88	68	70
79	65	60	70	77	63	58	75
85	71	64	59	55	61	65	83
87	79	69	68	65	76	78	94

−128

−76	−73	−67	−62	−58	−67	−64	−55
−65	−69	−62	−38	−19	−43	−59	−56
−66	−69	−60	−15	16	−24	−62	−55
−65	−70	−57	−6	26	−22	−58	−59
−61	−67	−60	−24	−2	−40	−60	−58
−49	−63	−68	−58	−51	−65	−70	−53
−43	−57	−64	−69	−73	−67	−63	−45
−41	−49	−59	−60	−63	−52	−50	−34

DCT

−415	−29	−62	25	55	−20	−1	3
7	−21	−62	9	11	−7	−6	6
−46	8	77	−25	−30	10	7	−5
−50	13	35	−15	−9	6	0	3
11	−8	−13	−2	−1	1	−4	1
−10	1	3	−3	−1	0	2	−1
−4	−1	2	−1	2	−3	1	−2
−1	−1	−1	−2	−1	−1	0	−1

JPEG

−415	−29	−62	25	55	−20	−1	3
7	−21	−62	9	11	−7	−6	6
−46	8	77	−25	−30	10	7	−5
−50	13	35	−15	−9	6	0	3
11	−8	−13	−2	−1	1	−4	1
−10	1	3	−3	−1	0	2	−1
−4	−1	2	−1	2	−3	1	−2
−1	−1	−1	−2	−1	−1	0	−1

÷

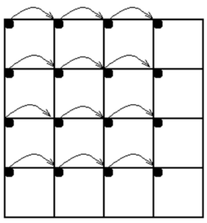
Quantization matrix

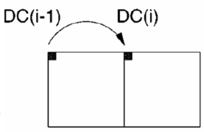
16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

−26	−3	−6	2	2	0	0	0
1	−2	−4	0	0	0	0	0
−3	1	5	−1	−1	0	0	0
−4	1	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

JPEG Coding of DC Coefficients

DPCM coding of DC coefficients (also referred as DC Prediction)



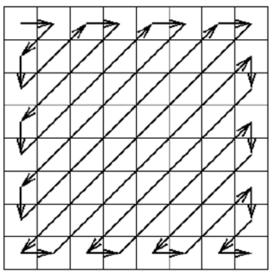


– $DiffDC\ (i) = DC\ (i) - DC\ (i-1),$
where $DC(\cdot)$'s are quantized values.

27

JPEG Coding of AC Coefficients

- Because the location of the retained coefficients vary from block to block, the quantized AC coefficients are zigzag scanned and ordered into (run, level) pairs.
  Level = the value of a nonzero coefficient;
  Run = the number of zero coefficients preceeding it.



- These (run, level) pairs are entropy coded, that is, longer codes for less frequent pairs and vice versa).

28

Example

- After DCT and quantization

$$\begin{bmatrix} 20 & 5 & -3 & 1 & 3 & -2 & 1 & 0 \\ -3 & -2 & 1 & 2 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & & & & & & & \\ & & & 0 & & & & \end{bmatrix}$$

Zigzag scan
[20,5,-3,-1,-2,-3,1,1,-1,-1,0,0,1,2,3,-2,1,1,0,0,0,0,0,0,1,1,0,1,EOB]

29: previous DC

DPCM   Run level

-9   (0,5)   (0,-3)   (0,-1)   (0,-2)
(0,-3)   (0,1)   (0,1)   (0,1)
(0,-1)   (0,-1)   (2,1)   (0,2)
(0,3)   (0,-2)   (0,1)   (0,1)
(6,1)   (0,1)   (0,1)   (1,1)   EOB

(101 0110/100 101/
01 00/ .../1010)

29

JPEG Examples

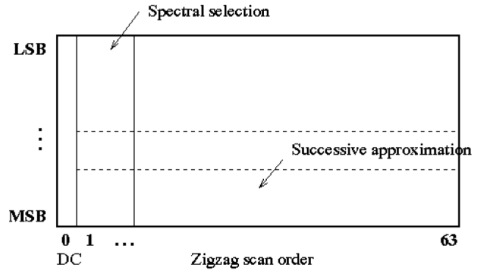
Medium quantization

High quantization

30

JPEG Progressive Coding

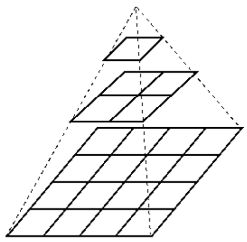
- The progressive mode consists of a sequence of ``scans" each of which codes a part of the quantized DCT coefficients.
- Spectral selection: The DCT coefficients are grouped into spectral bands. The lower frequency bands are usually coded (sent) first.
- Successive approximation: The information is first sent with lower precision, and then refined in later scans.
- Two processes may be combined to provide a graceful progression.



31

JPEG Hierarchical Coding

- The first stage (lowest resolution) is coded using one of the baseline sequential or progressive JPEG modes. The output of each hierarchical stage is then upsampled (interpolated) and used as the prediction for the next stage.
- The image quality at extremely low bit rates surpasses any of the other JPEG modes, but this is achieved at the expense of a higher bit rate at the completion of the progression.

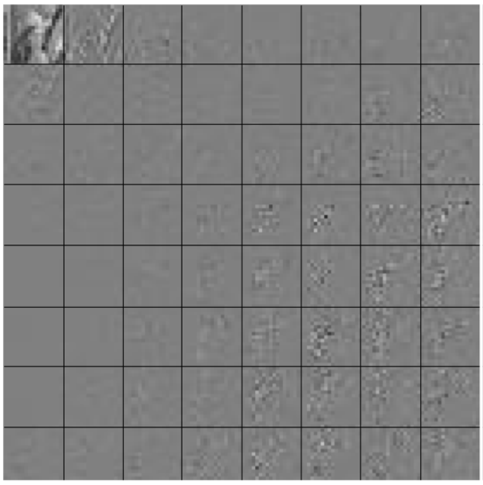


32

DCT vs. Subband Transformation

Collect all DCT coefficients corresponding to same basis image into a band

Take the inverse DCT of each band and tile them

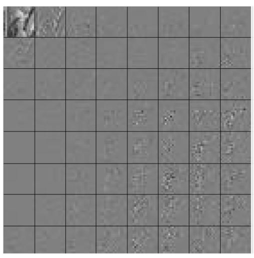


Subband transform of the image

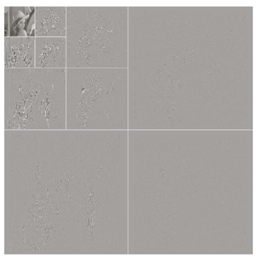
33

Subband Coding

- Decompose an image into critically sampled frequency bands and encode each band separately.



Uniform decomposition

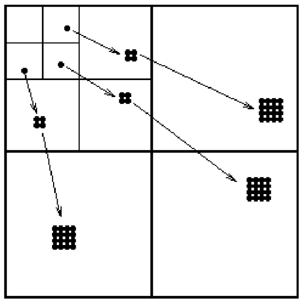


Tree-structured decomposition

34

Coding of the Subbands

- Encode the lowest band using the JPEG algorithm
- For all other bands, apply uniform quantization, followed by run-length coding and entropy coding
- Trace a sequence of zero symbols across bands at the same spatial location



35

Next Lecture

- END OF THE COURSE ☺

36