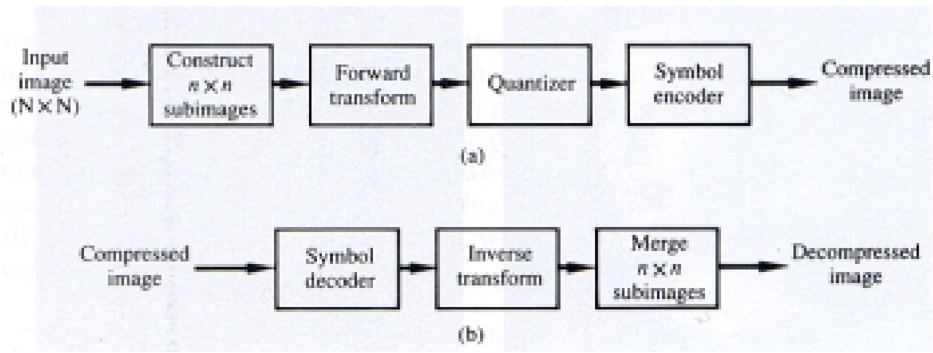
6.5.2 Transform Coding

- for most natural image
  - a significant number of coefficients. : small magnitude
    - → coarsely quantized or discarded

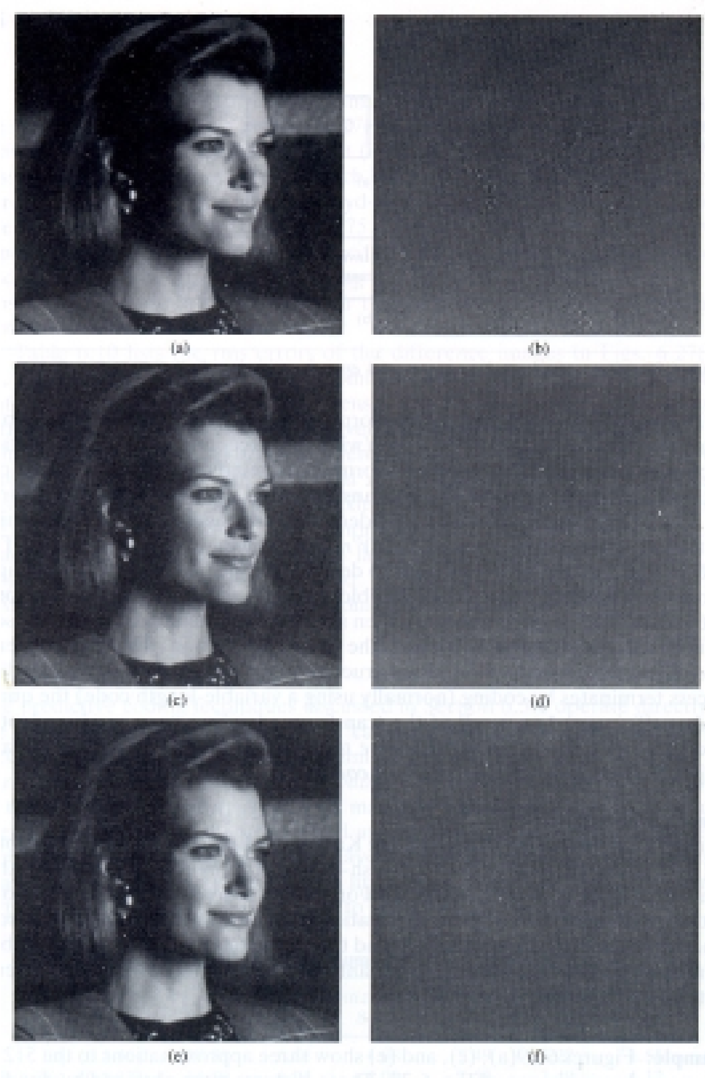


- operations
  - subimage decomposition,   transformation,   quantization, c oding

1) Transform selection

- Karhunen-Loève (KLT)
  - DFT
  - DCT
  - Walsh-Hadamard (WHT)
- Choice depends on
  - Error tolerence
  - Computational resources available
- Compression : by quantization not during the transformation step

Ex.)



- n× n image f(x,y)

$$f(x,y)=\sum_{u=0}^{n-1}\sum_{v=0}^{n-1}T(u,v)h(x,y,u,v)$$

$T(u,v)$: 2-D transform

$h(x,y,u,v)$: Inverse kernel

matrix form

$$\mathbf{F}=\sum_{u=0}^{n-1}\sum_{v=0}^{n-1}T(u,v)\mathbf{H}_{uv}$$

$$\mathbf{H}_{uv} = \begin{bmatrix} h(0,0,u,v) & h(0,1,u,v) & \cdots & h(0,n-1,u,v) \\ h(1,0,u,v) & & \cdots & \\ \vdots & \vdots & \cdots & \vdots \\ h(n-1,0,u,v) & h(n-1,1,u,v) & \cdots & h(n-1,n-1,u,v) \end{bmatrix}$$

subimage **F** : a linear combination of n^2 $n \times n$ matrices

- masking function
  - $m(u,v) = \begin{cases} 0 & \textit{if } Y(u,v) \textit{ satisfies a specified truncation criterion} \\ 1 & \textit{otherwise} \end{cases}$

  u,v = 0, 1, 2, …, n-1

- approximation of F

  $$\hat{\mathbf{F}} = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} T(u,v) m(u,v) \mathbf{H}_{uv}$$

  - mean square error
  - $e_{ms} = E\{\|\mathbf{F} - \hat{\mathbf{F}}\|^2\}$

    $= E\{\|\sum\sum T(u,v)\mathbf{H}_{uv}(1 - m(u,v))\|^2\}$

    $= \sum\sum \sigma^2_{T(u,v)}[1 - m(u,v)]$ ($\because$ **F** is zero mean and **H** is orthonormal)

  - total mean-square approximation Error

    = sum of variance of the discarded transform coefficients.

- DCT : superior to DFT, WHT

  KLT
  - the optimal transform
  - minimum MSE
  - data dependent → basis image (kernel) : nontrivial computational task
  - seldom used

  DCT : normally selected

  WHT : simple to implement (non sinusoidal transform)

  DCT : minimize blocking artifact

  DFT : blocking artifact

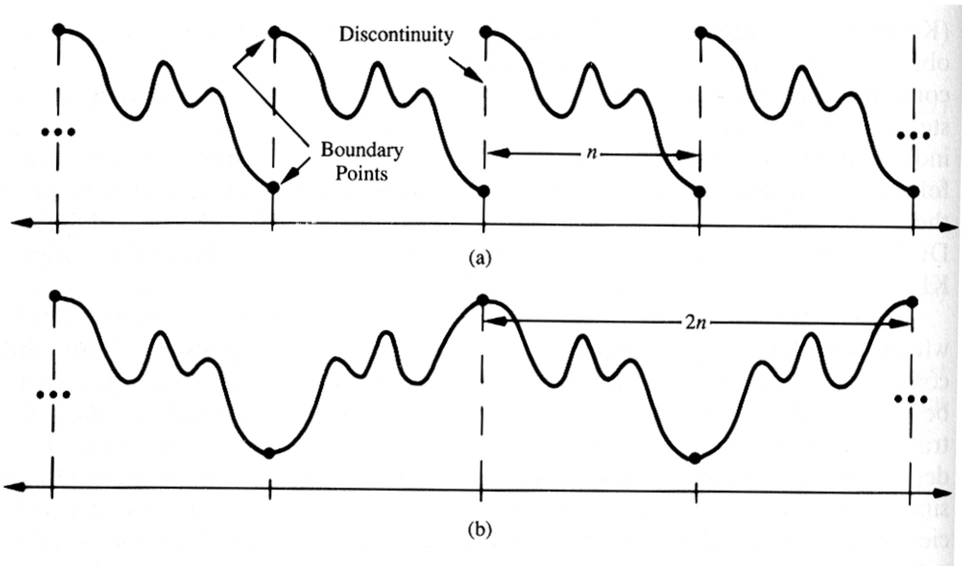


Figure 6.30 The periodicity implicit in the 1-D (a) DFT and (b) DCT.

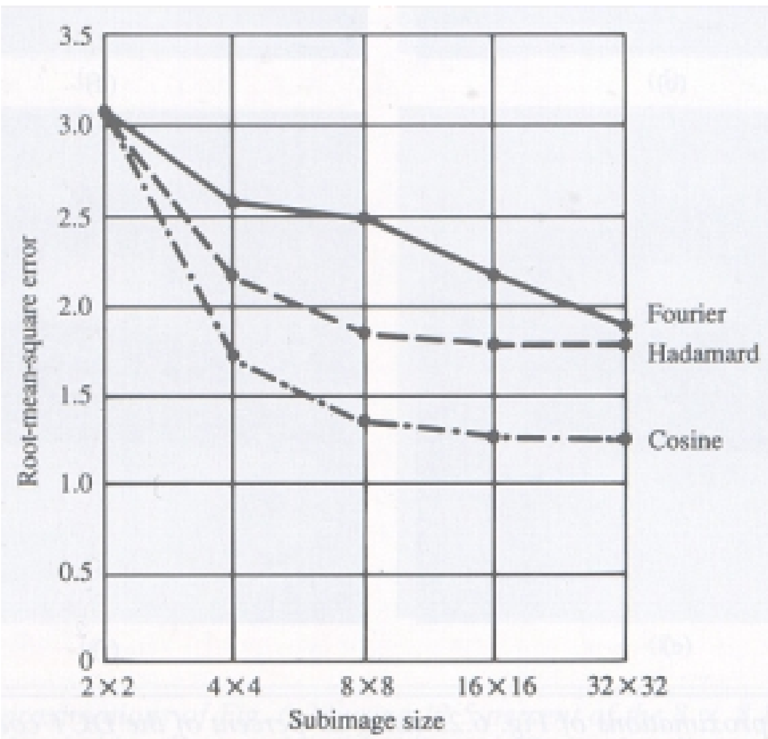
3) Subimage size selection

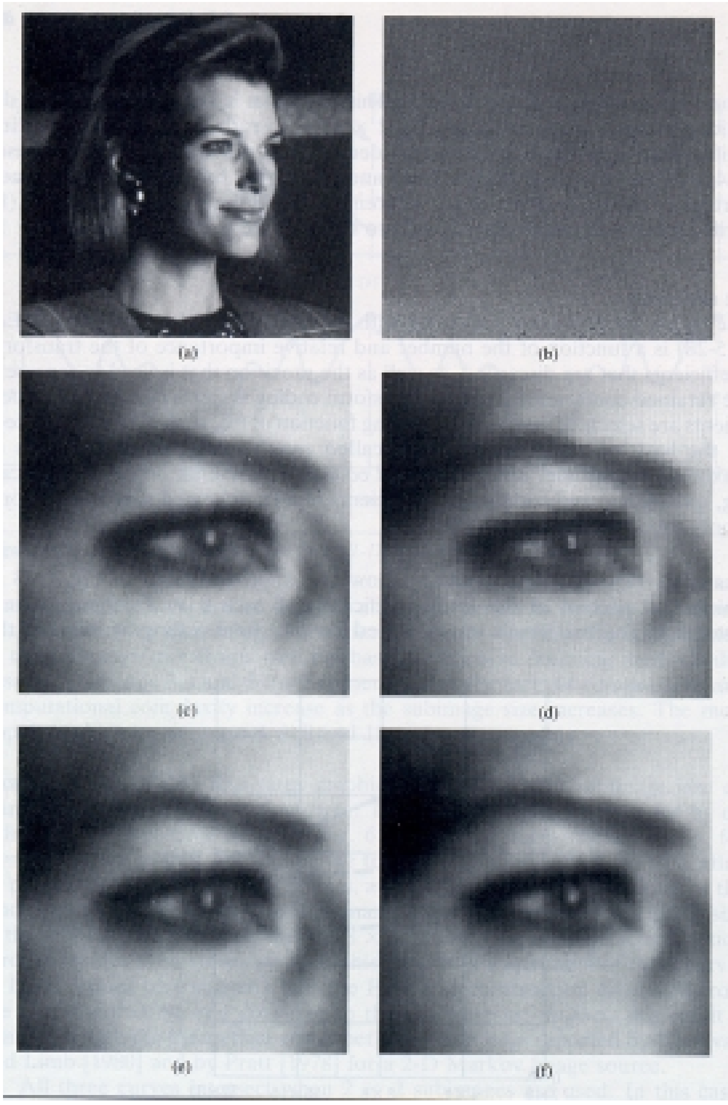
- If subimage size ↑

  → compression ratio ↑ , computational complexity ↑

- Mostly 8× 8, 16× 16

Ex.)





4)Bit allocation

→ bit allocation : the overall process of truncating, quantizing, and coding the coefficients of a transformed subimage

- Zonal coding : max. variance
- Threshold coding : max. magnitude

ex.)

5) Zonal Coding

- Information = uncertainty
- Transform coeffs. of max. variance
  - carry the most picture information
  - should be retained in the coding process
- Variance
  - calculated directly from the ensemble of $(N/n)^2$ transformed subimage arrays
  - or based on an assumed image model (ex. Markov autocorrelation function)

- Zonal mask
  - (a) location of max. variance : 1

    all other location : 1

    (b) zonal mask including bit allocation information
  - in most cases, allocated the same no. of bits or unequally allocated
- Equal(uniform) bit alloc.
- Unequal(nonuniform) bit alloc.
  - Lloyd-Max quantizer : designed for each coeffs.
  - DC coeffs : modeled Rayleigh density function

    The remaining coeffs. : modeled by a Laplacian or Gaussian density
  - no. of quantization levels allotted to each quantize
    - ✓ proportional to $\log_2 \sigma^2_{T(u,v)}$
    - ✓ consistent with rate distortion theory
- Zonal coding : use of a single fixed mask for all subimages

1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0
1	1	1	0	0	0	0	0
1	1	0	0	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

(a)

8	7	6	4	3	2	1	0
7	6	5	4	3	2	1	0
6	5	4	3	3	1	1	0
4	4	3	3	2	1	0	0
3	3	3	2	1	1	0	0
2	2	1	1	1	0	0	0
1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)

1	1	0	1	0	0	0	0
1	1	1	1	0	0	0	0
1	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	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

(c)

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

(d)

6) Threshold Coding

- Thresholding criteria
  1. A single global threshold : applied to all subimage

- variable code rate

2. A diff. Threshold : used for each subimage
   - N-largest coding constant code rate
3. The threshold : varied as a function of the location of each coeff. within the subimage
   - variable code rate

- The third tech.
  - thresholding and normalization

$$\hat{\mathbf{T}} = round\left[\frac{T(u,v)}{Z(u,v)}\right]$$

where $\hat{\mathbf{T}}[\ ]$: thresholded and normalized appr. of $\mathbf{T}[\ ]$

Z : transform normalization array

Eg. 6.5.31

- denormalization

$$\dot{T}(u,v) = \hat{T}(u,v)Z(u,v)$$

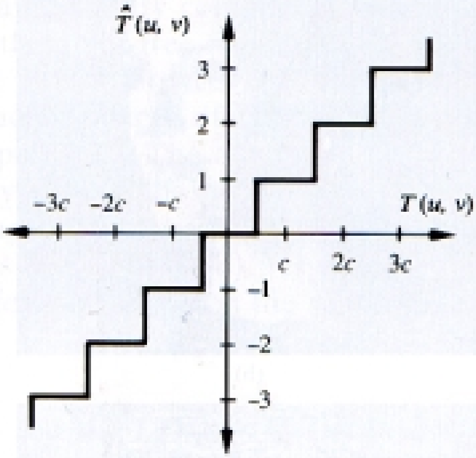
$$\xrightarrow[transform]{inverse}$$ decompressed subimage

- Ex.

If Z[u,v]=c

$$kc - \frac{c}{2} \le T[u,v] \le kc + \frac{c}{2} \rightarrow \hat{T}[u,v] = k$$

fig. 6.35



(a)

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

(b)

(b) typical normalization array (used in JPEG)

ex.)