



Università degli Studi di Bergamo



**DIP. DI INGEGNERIA DELL'INFORMAZIONE E
METODI MATEMATICI**

RETI INTERNET MULTIMEDIALI

JPEG, TIFF, Facsimile

JPEG

Joint Photographic Experts Group

Introduzione

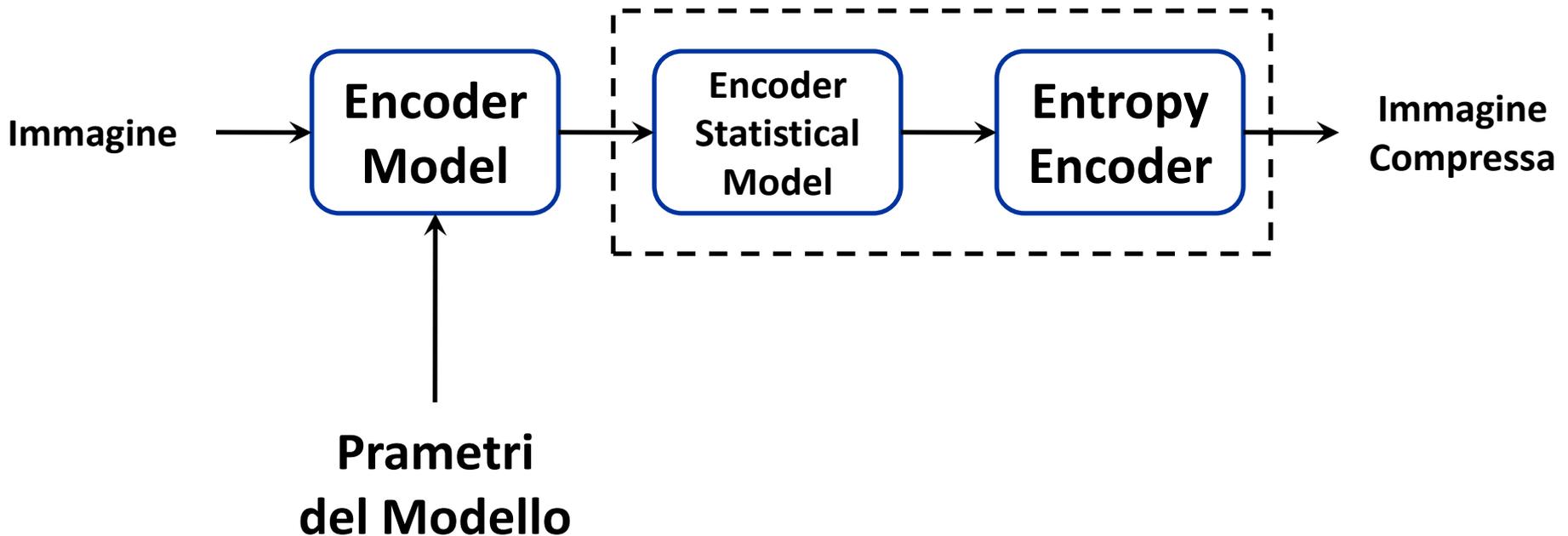
- JPEG: Joint Photographic Experts Group
 - Gruppo di lavoro della ISO e ITU-T
- Standard di compressione con perdita (compressione lossy) delle immagini con profondità 6-24 bits per pixel
- La compressione e decompressione sono eseguite in modo efficiente e rapido (risorse richieste)
- Non definisce un unico formato per il file che contiene i dati che rappresentano l'immagine
 - Esistono diversi formati che supportano JPEG

Introduzione

- JPEG sfrutta le caratteristiche di percezione visiva dell'occhio umano
 - Sensibilità maggiore all'intensità del colore rispetto alla sua velocità di variazione
- Può raggiungere un rapporto di compressione elevato, oltre 20:1
- Garantisce un buona qualità dell'immagine anche ad alti rapporti di compressione
- La codifica JPEG è applicabile a qualsiasi immagine
 - Color e Grayscale (variazione continua)
 - Composta da un massimo di 65535 linee e 65535 pixel per linea

Introduzione

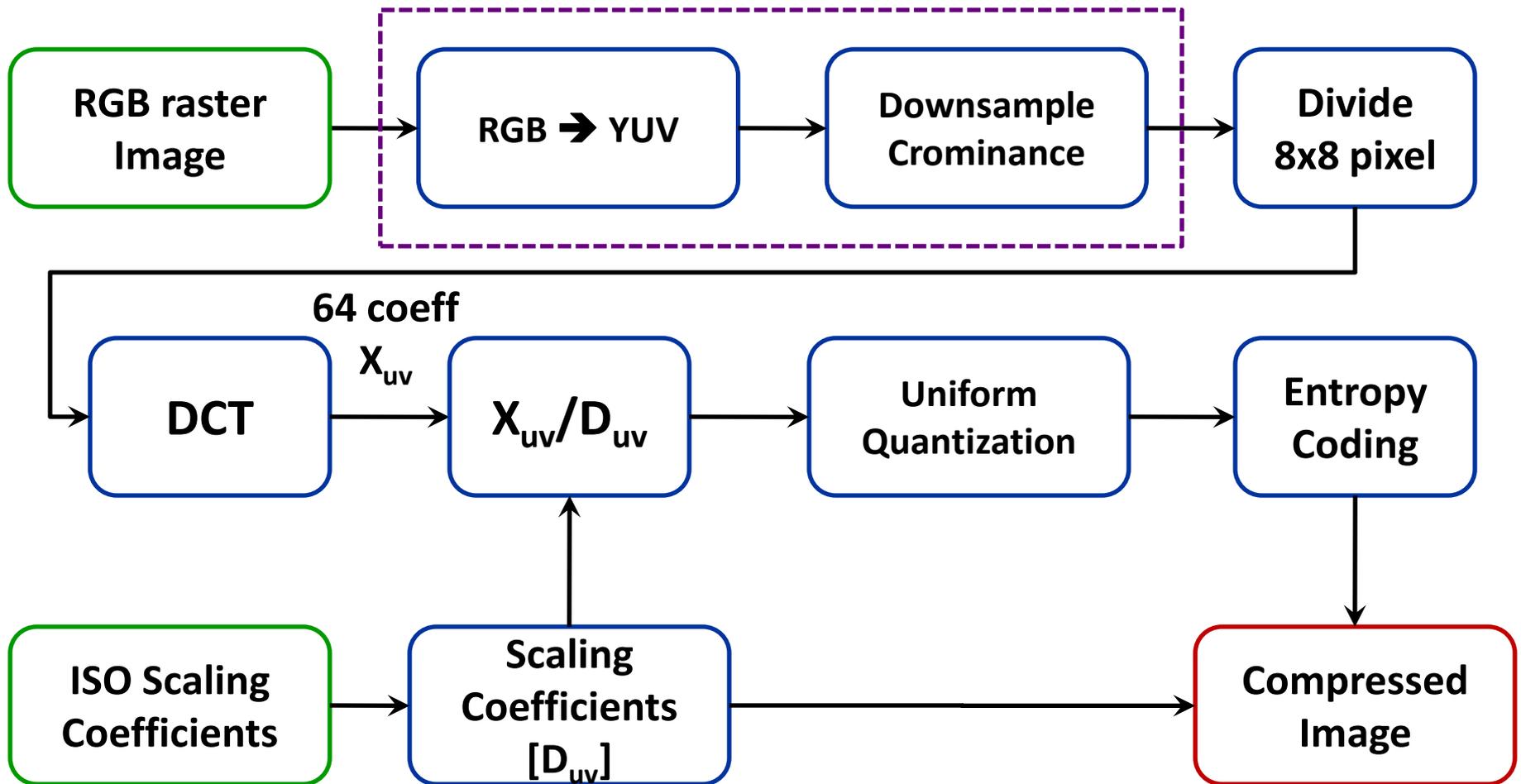
- La codifica JPEG è composta da 3 componenti principali



Introduzione

- Lo standard JPEG definisce 4 diversi tipi di modalità operative (tipi di codifica):
 - Sequential DCT-based (standard)
 - Ogni componente dell'immagine è codificata in una sola scansione
 - Progressive DCT-based
 - Sequential Lossless
 - Hierarchical
- Inoltre sono supportate 2 diverse tipi di codifica lossless (entropy coding):
 - Codifica di Huffman
 - Codifica Aritmetica

Compressione JPEG



Compressione JPEG

- Algoritmo di compressione composto da 5 fasi
 1. Trasformare l'immagine dalla rappresentazione RGB alla rappresentazione YUV [Opzionale]
 2. Dividere il file in blocchi di 8 pixel x 8 pixel
 3. Applicare la Discrete Cosine Transform (DCT) a ogni blocco 8x8 (spazio → frequenza)
 4. Quantizzare i coefficienti della DCT, dividendo ognuno di essi per un valore intero e approssimando all'intero più vicino
 5. Eseguire una codifica di Huffman (o codifica aritmetica) dei coefficienti ottenuti al passo precedente (scansione zig-zag).

Fase 1a: Conversione RGB → YUV

- L'immagine sorgente è composta da una matrice raster RGB (Red, Green, Blue), che utilizza tre valori numerici (8 bit) per rappresentare ogni pixel che dell'immagine
- La rappresentazione RGB viene trasformata in rappresentazione YUV, in cui ogni i colori che compongono ogni pixel sono rappresentati in termini di luminanza (luminosità) e cromaticanza (colore)
- Infatti l'occhio umano è più sensibile alla componente di luminosità Y
- La trasformazione RGB → YUV non ai fini della compressione JPEG, ma permette di ottenere un maggiore rapporto di compressione

Fase 1a: Conversione RGB → YUV

- La trasformazione è eseguita semplicemente attraverso un'operazione aritmetica
 - La formula utilizzata specifica il contributo relativo di ognuno dei tre colori alle componenti di luminosità e cromaticità
- Esistono diverse versioni della formula di conversione, a seconda del dispositivo utilizzato per visualizzare l'immagine

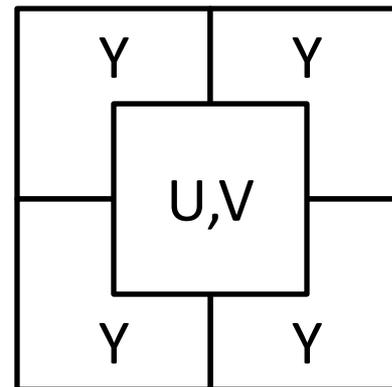
$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$

$$U = -0.1687 \cdot R - 0.3313 \cdot G + 0.5 \cdot B + 128$$

$$V = 0.5 \cdot R - 0.4187 \cdot G - 0.813 \cdot B + 128$$

Fase 1b: Sottocampionamento

- Poichè l'occhio umano è più sensibile alla luminosità, le componenti di cromaticità possono essere sottocampionate
- Esistono diverse alternative di sotto-campionamento
- La notazione **4:1:1** indica che per ogni blocco di 4 pixel si hanno 4 campioni di luminosanza (Y), e 1 campione per ognuna delle due componenti di cromaticità (U e V)
- JPEG usa un sottocampionamento 4:2:0
- MCU: Minimum Coded Unit



Fase 2: Divisione in blocchi 8x8

- L'immagine viene scomposta in blocchi di dimensione 8 pixel x 8 pixel
- Si noti che con un sottocampionamento 4:1:1 si hanno 16 pixel in ognuna delle due direzioni di luminanza (Y), ma solo 8 per ognuna delle componenti di cromaticità (U e V)
- Se l'immagine non può essere scomposta perfettamente in blocchi di dimensione 8x8, vengono aggiunti ulteriori pixel (padding)

Fase 3: DCT

- La DCT trasforma i dati dell'immagine dal dominio dello spazio a quello delle frequenze
 - Il dominio dello spazio mostra la variazione di intensità del colore nello spazio
 - Il dominio delle frequenze mostra quanto rapidamente l'intensità del colore varia spostandosi da un pixel a quello adiacente

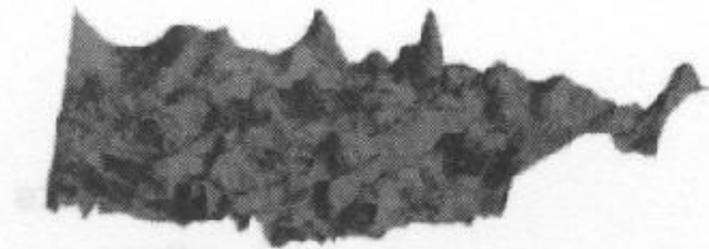


Figure 5.4

Fase 3: DCT

- La rappresentazione nel dominio delle frequenze è più semplice da analizzare
 - Si distingue più facilmente l'informazione che è essenziale ai fini della percezione umana da quella che risulta invece superflua
- L'occhio umano non è molto sensibile a repentini cambiamenti del colore nelle fotografie
 - Componenti in alta frequenza possono essere filtrate

Fase 3: DCT

- Input: blocco (matrice) composto da 8x8 pixel

$$[p_{xy}] \quad 0 \leq x < 8, 0 \leq y < 8$$

- Output: matrice composta dai coefficienti in frequenza della DCT

$$[X_{uv}] \quad 0 \leq u < 8, 0 \leq v < 8$$

$$X_{uv} = \frac{1}{\sqrt{2N}} C_u C_v \sum_{x=0}^7 \sum_{y=0}^7 p_{xy} \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$C_u C_v = \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & u, v \neq 0 \end{cases}$$

Fase 4: Quantizzazione

- La trasformata DCT non comprime il contenuto informativo, lo trasforma soltanto
- E' possibile ricostruire l'immagine originaria anti-trasformando se si hanno a disposizione tutti i coefficienti della trasformata
- I coefficienti della trasformata sono valori reali (floating-point)
 - E' necessario comprimerli
- I coefficienti della trasformata vengono scalati dividendoli per opportuni coefficienti D_{uv} (tabella ISO)

$$X_{uv} \rightarrow X_{uv} / D_{uv}$$

Fase 4: Quantizzazione

- La fase di quantizzazione rappresenta ogni coefficiente modificato della DCT con un valore intero tra 0 e 255
- La quantizzazione viene eseguita in modo uniforme rispetto a tutti i coefficienti della trasformata

$$F_{uv} = \text{round}\left(\frac{X_{uv}}{D_{uv}}\right)$$

- I coefficienti utilizzati per scalare i coefficienti della DCT sono trasmessi insieme all'immagine compressa

Fase 4: Quantizzazione

- Lo scopo della quantizzazione è di ridurre il numero di bit necessari a rappresentare l'immagine compressa
- Alti valori dei coefficienti D_{uv} permettono di ridurre il contenuto informativo
 - Alti valori dei coefficienti D_{uv} vengono utilizzati per le alte frequenze
- Esistono tabelle standard dei coefficienti D_{uv}
 - Sono basate su studi di percezione visiva
 - Massimizzare il rapporto di compressione con una minima distorsione della percezione visiva

Fase 4: Quantizzazione

- Possono essere utilizzate diverse tabelle/matrici di quantizzazione (variando i valori di default)
 - Trade off tra qualità e rapporto di compressione
 - Alti valori dei coefficienti D_{uv} implicano maggiore compressione e minore qualità
- I valori utilizzati per variare i coefficienti D_{uv} standard devono essere inviati al decodificatore

Blocchi Originali e Codificati DCT

■ Blocco altamente omogeneo

Luminanza (Y) del blocco 8x8



p_{xy}

200	202	189	188	189	175	175	175
200	203	198	188	189	182	178	175
203	200	200	195	200	187	185	175
200	200	200	200	197	187	187	187
200	205	200	200	195	188	187	175
200	200	200	200	200	190	187	175
205	200	199	200	191	187	187	175
210	200	200	200	188	185	187	186

$$X_{uv} = DCT(p_{xy})$$



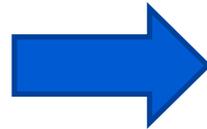
515	65	-12	4	1	2	-8	5
-16	3	2	0	0	-11	-2	3
-12	6	11	-1	3	0	1	-2
-8	3	-4	2	-2	-3	-5	-2
0	-2	7	-5	4	0	-1	-4
0	-3	-1	0	4	1	-1	0
3	-2	-3	3	3	-1	-1	3
-2	5	-2	4	-2	2	-3	0

Blocchi Originali e Codificati DCT

■ Blocchi quantizzati e ricostruiti

$$F_{uv} = \text{round}\left(\frac{X_{uv}}{D_{uv}}\right)$$

32	6	-1	0	0	0	0	0
-1	0	0	0	0	0	0	0
-1	0	1	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
0	0	0	0	0	0	0	0



$$\tilde{F}_{uv} = F_{uv} \cdot D_{uv}$$

512	66	-10	0	0	0	0	0
-12	0	0	0	0	0	0	0
-14	0	16	0	0	0	0	0
-14	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

IDCT e Differenze dall'Originale

- Errore commesso a causa della quantizzazione nel processo di ricostruzione

$$\tilde{p}_{xy} = IDCT(\tilde{F}_{uv})$$

$$\varepsilon_{xy} = p_{xy} - \tilde{p}_{xy}$$

199	196	191	186	182	178	177	176
201	199	196	192	188	183	180	178
203	203	202	200	195	189	183	180
202	203	204	203	198	191	183	179
200	201	202	201	196	189	182	177
200	200	199	197	192	186	181	177
204	202	199	195	190	186	183	181
207	204	200	194	190	187	185	184



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

Blocchi Originali e Codificati DCT

■ Blocco poco omogeneo

Luminanza (Y) del blocco 8x8



p_{xy}

70	70	100	70	87	87	150	187
85	100	96	79	87	154	87	113
100	85	116	79	70	87	86	196
136	69	87	200	79	71	117	96
161	70	87	200	103	71	96	113
161	123	147	133	113	113	85	161
146	147	175	100	103	103	163	187
156	146	189	70	113	161	163	197

$X_{uv} = DCT(p_{xy})$



-80	-40	89	-73	44	32	53	-3
-135	-59	-26	6	14	-3	-13	-28
47	-76	66	-3	-108	-78	33	59
-2	10	-18	0	33	11	-21	1
-1	-9	-22	8	32	65	-36	-1
5	-20	28	-46	3	24	-30	24
6	-20	37	-28	12	-35	33	17
-5	-23	33	-30	17	-5	-4	20

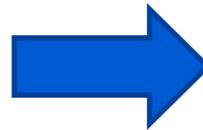
Blocchi Originali e Codificati DCT

■ Blocchi quantizzati e ricostruiti

$$F_{uv} = \text{round}\left(\frac{X_{uv}}{D_{uv}}\right)$$

$$\tilde{F}_{uv} = F_{uv} \cdot D_{uv}$$

-5	-4	9	-5	2	1	1	0
-11	-5	-2	0	1	0	0	-1
3	-6	4	0	-3	-1	0	1
0	1	-1	0	1	0	0	0
0	0	-1	0	0	1	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



-80	-44	90	-80	48	40	51	0
-132	-60	-28	0	26	0	0	-55
42	-78	64	0	-120	-57	0	56
0	17	-22	0	51	0	0	0
0	0	-37	0	0	109	0	0
0	-35	55	-64	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

IDCT e Differenze dall'Originale

- Errore commesso a causa della quantizzazione nel processo di ricostruzione

$$\tilde{p}_{xy} = IDCT(\tilde{F}_{uv})$$

$$\varepsilon_{xy} = p_{xy} - \tilde{p}_{xy}$$

70	60	106	94	62	103	146	176
85	101	85	75	102	127	93	144
98	99	92	102	74	98	89	167
132	53	111	180	55	70	106	145
173	57	114	207	111	89	84	90
164	123	131	135	133	92	85	162
141	159	169	73	106	101	149	224
150	141	195	79	107	147	210	153



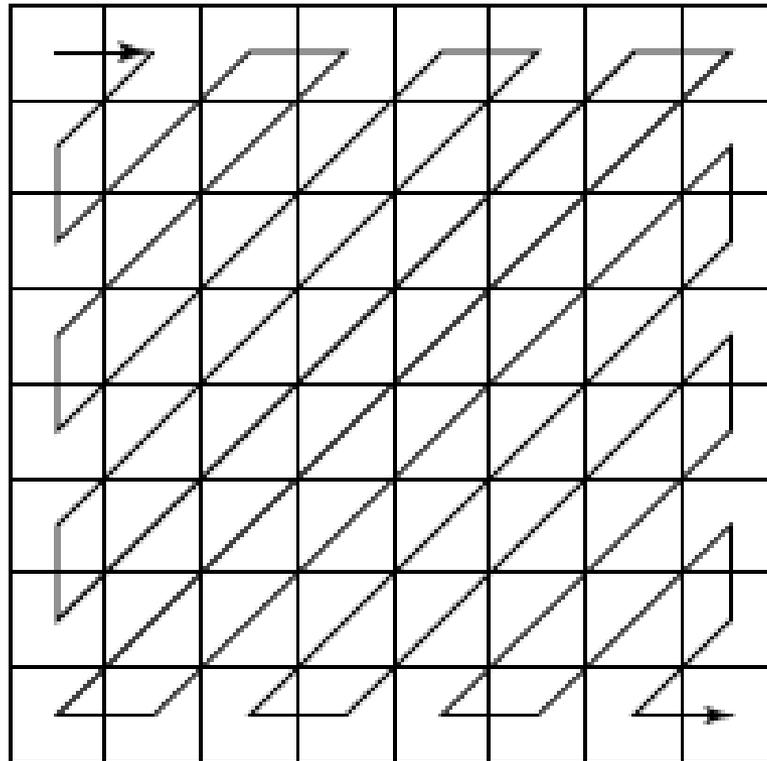
0	10	-6	-24	25	-16	4	11
0	-1	11	4	-15	27	-6	-31
2	-14	24	-23	-4	-11	-3	29
4	16	-24	20	24	1	11	-49
-12	13	-27	-7	-8	-18	12	23
-3	0	16	-2	-20	21	0	-1
5	-12	6	27	-3	2	14	-37
6	5	-6	-9	6	14	-47	44

Fase 5: Entropy Coding

- Ogni elemento della matrice dei coefficienti quantizzati viene compressa utilizzando una codifica lossless (Huffman o Aritmetica)
- La sequenza degli elementi della matrice segue lo schema zig-zag
 - Migliora il rapporto di compressione
- L'informazione codificata e i valori utilizzati per quantizzare i coefficienti della DCT rappresentano l'immagine compressa
- I coefficienti DC vengono generalmente codificati separatamente dai rimanenti coefficienti AC
 - Blocchi adiacenti hanno valori medi simili
 - La codifica della differenza tra i coefficienti DC di blocchi adiacenti produce un maggior guadagno di compressione

Fase 5: Entropy Coding

- Il codificatore segue un percorso zig-zag



Modalità JPEG

- Lo standard JPEG definisce 4 diversi tipi di modalità operative (tipi di codifica):
 - Sequential DCT-based (standard)
 - Progressive DCT-based
 - Sequential Lossless
 - Hierarchical

Modalità JPEG Progressiva

- Progressive mode genera e trasmette diverse versioni dell'immagine a qualità crescente
 - Viene inviata una prima versione a bassa qualità
 - I dettagli vengono aggiunti da scansioni successive
 - Es. Sono inviati i coefficienti DC e pochi AC per primi. Vengono inviati in successione ulteriori coefficienti AC finchè tutti i coefficienti sono stati inviati
- Molto utile per i siti web, in cui l'immagine da scaricare può richiedere molto tempo
 - Il browser può visualizzare la prima versione poco dettagliata e di seguito aggiungere dettagli

Sequenziale vs. Progressiva

- Visualizzazione di un immagine JPEG sequenziale



- Visualizzazione di un immagine JPEG progressiva



Modalità Gerarchica

- La modalità gerarchica codifica l'immagine a risoluzioni diverse
- Le diverse versioni dell'immagine possono essere codificate e trasmesse a risoluzione crescente
- Generalmente ad ogni scansione la risoluzione viene incrementata di un fattore 2
 - Base layer: versione a più bassa risoluzione
 - Enhancement layer: ottenuto dall'interpolazione tra i pixel del base layer e le differenze tra il frame attuale e una versione predetta

Modalità Gerarchica

- Three-level Hierarchical Mode
- Algoritmo di compressione (3-level)
 1. Riduzione della risoluzione dell'immagine ingresso
La risoluzione dell'immagine x viene ridotta progressivamente di un fattore 2
 x (512 x512) \rightarrow x_2 (256x256) \rightarrow x_4 (128x128)
 2. Compressione dell'immagine a più bassa risoluzione
 x_4 viene compressa con la modalità sequenziale $x_4 \rightarrow X_4$

Modalità Gerarchica

■ Algoritmo di compressione (continua)

3. Genera il primo enhancement layer

- Decodifica X_4 per ottenere $\underline{x}_4 \rightarrow \underline{x}_4 = \text{Dec}(X_4)$
- Interpola \underline{x}_4 per ottenere una versione dell'immagine con risoluzione uguale a quella di $x_2 \rightarrow I(\underline{x}_4)$
- Codifica l'immagine ottenuta dalle differenze di x_2 e $I(\underline{x}_4) \rightarrow D_2 = \text{Enc}(x_2 - I(\underline{x}_4))$

4. Genera il secondo enhancement layer

- Decodifica D_2 per ottenere $\underline{d}_2 \rightarrow \underline{d}_2 = \text{Dec}(D_2)$
- Aggiungi $I(\underline{x}_4)$ a \underline{d}_2 per ottenere una versione di $x_2 \rightarrow \underline{x}_2 = \underline{d}_2 + I(\underline{x}_4)$
- Codifica l'immagine ottenuta dalle differenze di x e $I(\underline{x}_2) \rightarrow D_1 = \text{Enc}(x - I(\underline{x}_2))$

Modalità Gerarchica

- Algoritmo di decompressione (3-level)
 - Decodifica il base layer
 - Decodifica X_4 utilizzando lo stesso meccanismo JPEG utilizzato per ottenere \underline{x}_4
 - Genera l'immagine a risoluzione intermedia \underline{x}_2 utilizzando il primo enhancement layer
 - Utilizza $D_2 = \text{Enc}(x_2 - I(\underline{x}_4))$ e \underline{x}_4 per ottenere \underline{x}_2
 - Genera l'immagine originaria x utilizzando il secondo enhancement layer
 - Utilizza $D_1 = \text{Enc}(x - I(\underline{x}_2))$ e \underline{x}_2 per ottenere \underline{x}

Modalità Gerarchica

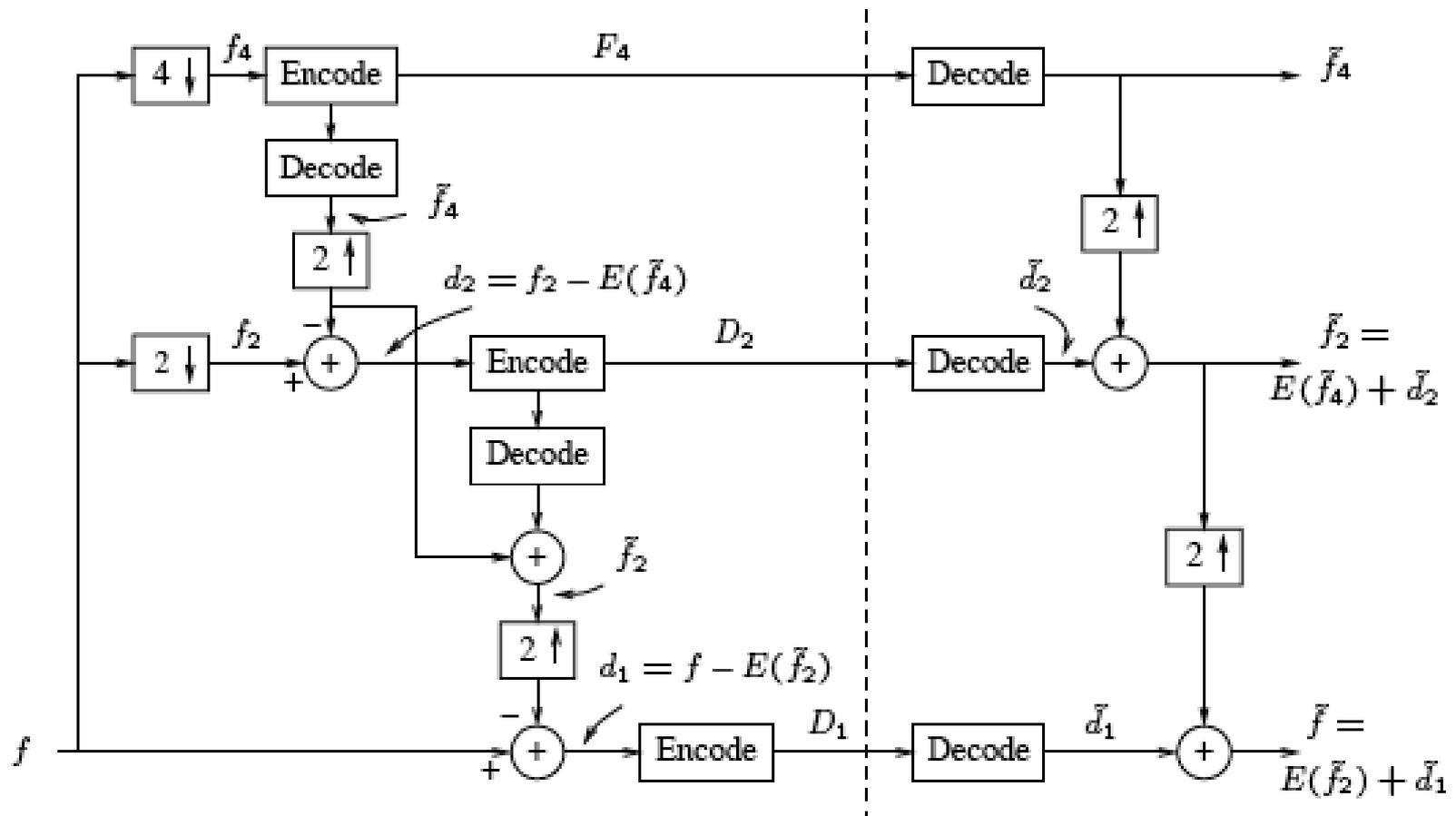
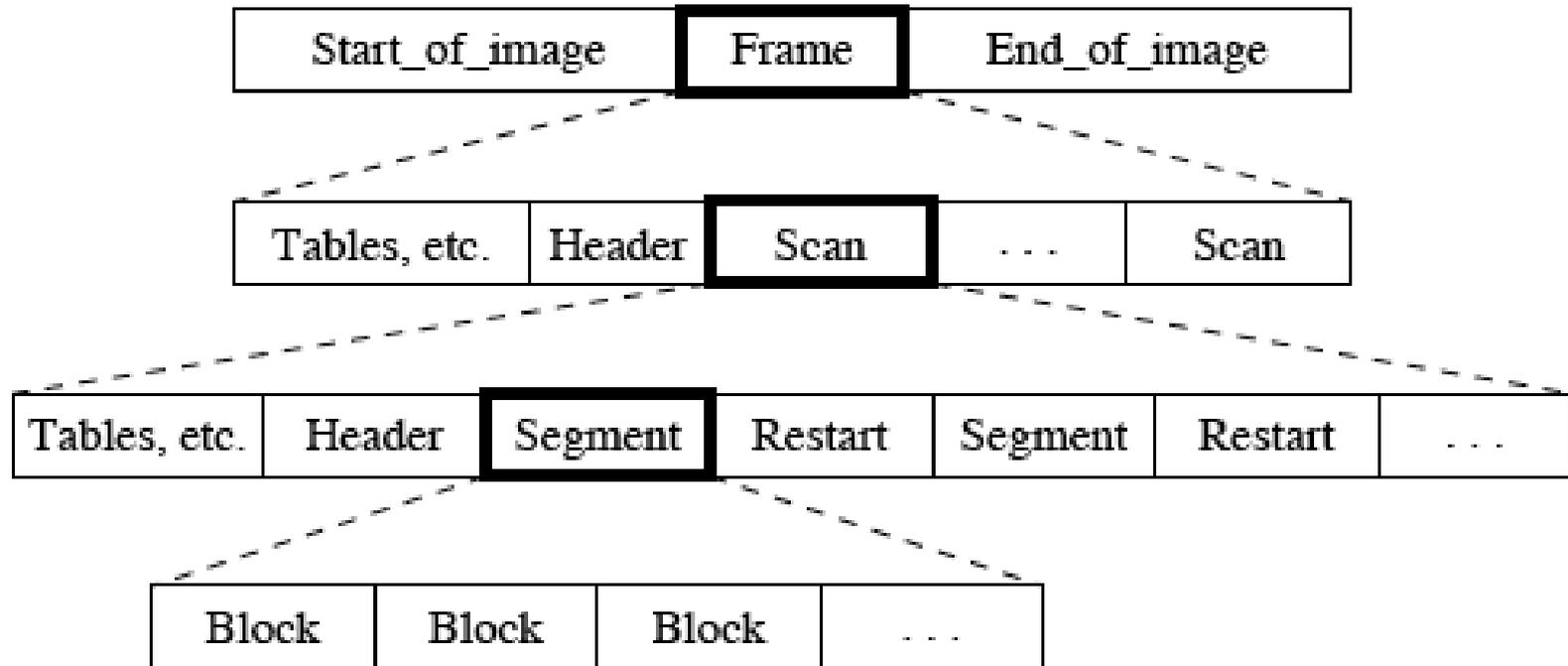


Fig. 9.5: Block diagram for Hierarchical JPEG.

Formato file JPEG

- Formato del file JPEG:



Formato file JPEG

- Il formato JPEG è organizzato in una gerarchia di componenti:
 - Frame: immagine
 - Scan: elemento di un'immagine
 - Segment: gruppo di blocchi codificati
- L'intestazione del Frame (Frame Header) contiene
 - Numero di bit per pixel
 - Dimensione dell'immagine
 - Dimensione della tabella di quantizzazione
- L'intestazione dello Scan (Scan Header) contiene:
 - Numero di componenti
 - Tabelle della codifica di Huffman

JPEG 2000

- JPEG2000 (jp2) è la versione più recente dei formati di compressione standardizzati dal gruppo JPEG
- Introduce numerosi miglioramenti rispetto a JPEG:
 - Codifica Wavelet
 - Migliori rapporti di compressione
 - Maggiore compressione senza perdita (lossless)
 - Supporta immagini di maggiore dimensione e con un numero maggiore di componenti rispetto a JPEG
 - Introduce la codifica basata sulla regione di interesse (Region-of-interest coding)

JBIG

- JBIG (Joint Bi-Level Image Processing Group) è uno standard di compressione per le immagini binarie
 - Fax, scansioni di documenti, etc.
 - Hanno caratteristiche diverse dalle immagini a colori o grayscale che richiedono diverse tipologie di codifica
- JBIG è una codifica di tipo senza perdita (lossless)
- JBIG2 è una codifica di tipo con perdita (lossy)
 - Codifica Model-based

TIFF

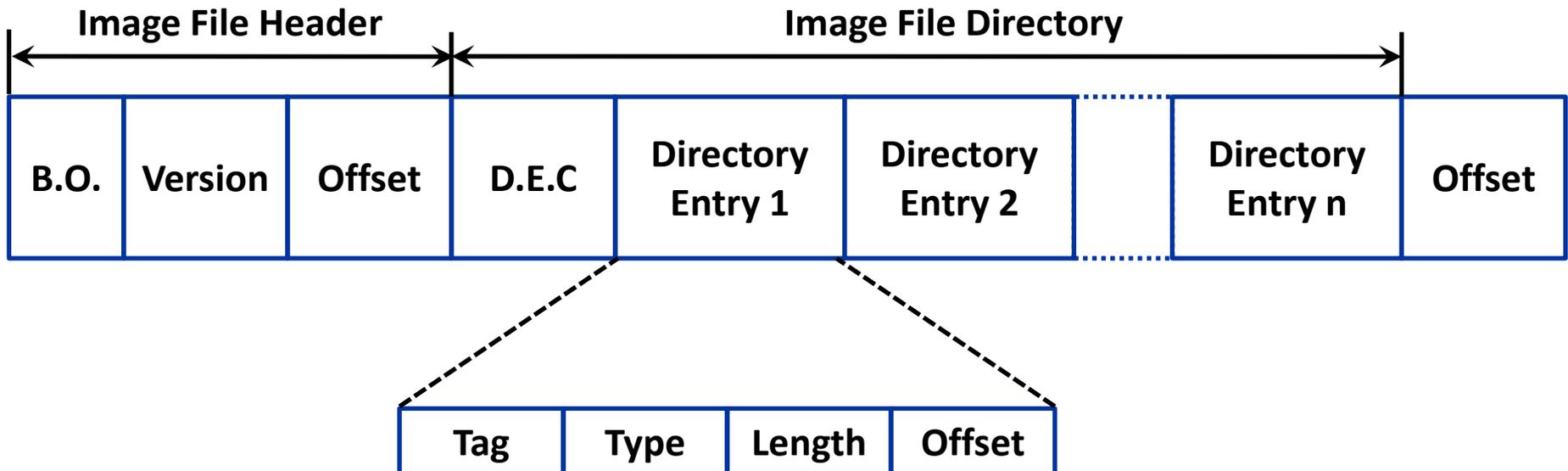
Tag Image File Format

Introduzione

- TIFF (Tag Image File Format)
 - Insieme di formati di immagini grafiche
 - Supporta 6 tipi di compressione lossless
 - E' utilizzata principalmente per la fotografia digitale e altre applicazioni che richiedono un'immagine la cui qualità non ha subito degradazioni (es. editing)
- E' un formato estendibile per cui molto flessibile

TIFF Formato File

- TIFF è composto da:
 - Image File Header (IFH): 8 byte
 - Image File Directory (IFD)
 - Directory Entry (DE): 12 byte



TIFF Formato File

Imege File Header

Campo	Bytes	Descrizione
Byte Order	2	Ordine dei byte utilizzata per creare il file
Version	2	Versione del formato TIFF
Offset	4	Identifica l'inizio del prossimo IFD (in bytes)

Imege File Directory

Campo	Bytes	Descrizione
Directory Entry Count		Numero di Directory Entries appartenenti al IFD corrente
Directory Entry	12	Porzione dell'informazione relativa alla matrice raster che rappresenta l'immagine

TIFF Formato File

- Directory Entry composta da:
 - Tag: deriva da diverse categorie
 - Informazioni di base: colore, dimensioni, compressione, campioni/pixel
 - Informazioni generiche: data, nome, pc
 - Informazioni di documento: nome, posizione
 - Type: formato (ASCII, 32-bit integer)
 - Length: numero di elementi
 - Offset: numero di bytes dall'inizio della DE all'inizio dei dati (se maggiore di 4 bytes indica un puntatore)
- I dati che rappresentano l'immagine sono memorizzati dopo l'IFD raggruppate in linee (corrispondenti alle linee di scansione)

Tipi di Codifica

- Sebbene il formato TIFF possa essere esteso per includere formati che utilizzano compressione lossy, le tecniche di compressione definite dal formato utilizzano solo 6 tipi di codifica lossless
- Le 6 tecniche identificate dal valore assegnato al campo Type della Directory Entry sono:
 - Nessuna compressione , Type = 1
 - CCITT Group 3, Huffman 1-D modificata, Type = 2
 - Facsimile compatible CCITT Group 3 , Type = 3
 - Facsimile compatible CCITT Group 4 , Type = 4
 - Lempel-Ziv and Welch (LZW) , Type = 5
 - PackBits (Macintosh) , Type = 32773

TIFF Conformant Classes

- TIFF Conformant Classes sono sottoinsiemi di tutte le tipologie di codifica supportate da TIFF, che richiedono un numero limitato di Tags per le esigenze dell'applicazione
 - B: immagini con 1 bit per pixel (2 livelli)
 - G: immagini grayscale
 - P: immagini con un numero limitato di colori (palette-color)
 - R: immagini RGB

FACSIMILE COMPRESSIVE CODING

CCITT T.4 (Group 3), T.6 (Group 4)

Introduzione

- Standards definiti dall'ITU (International Telecommunications Union)
- Group 3 definisce le versioni unidimensionali e bidimensionali
- Group 4 definito solo per un sistema bidimensionale (rapporti di compressione tipici 7:1 e 15:1)
- Entrambi gli standards utilizzano parametri costanti per gli algoritmi di compressione
 - Nessun adattamento al messaggio codificato
 - Minore efficacia di codifica

Introduzione

- Facsimile Compression
- ITU standard (A4, 210 x 297 mm)
- 1728 pixels per linea
 - se si utilizza 1 bit per ogni pixel si avrebbero più di 3M bits per pagina
- Una pagina contiene molti pixel consecutivi dello stesso colore
 - Applica Run Length Encoding (RLE)

Group 3

- L'algoritmo di compressione unidimensionale definito da Group 3 utilizza lo schema di Huffman
- La scansione di una linea produce una sequenza di pixel neri (valore '1') e bianchi (valore '0')
- Il codificatore genera una parola di codice che rappresenta la lunghezza della linea scansionata e il colore
- La parola di codice generata risulta generalmente più corta della linea originaria

Approfondimenti

■ Articoli:

- G. K. Wallace, “The JPEG still picture compression standard” IEEE Trans. Consumer Electronics, vol. 38, no. 1, pp. xviii-xxxiv, Feb. 1992.

■ Links:

- ITU-T Recommendation T.81
<http://www.w3.org/Graphics/JPEG/itu-t81.pdf>