

## 5. CODAREA VIDEO MPEG

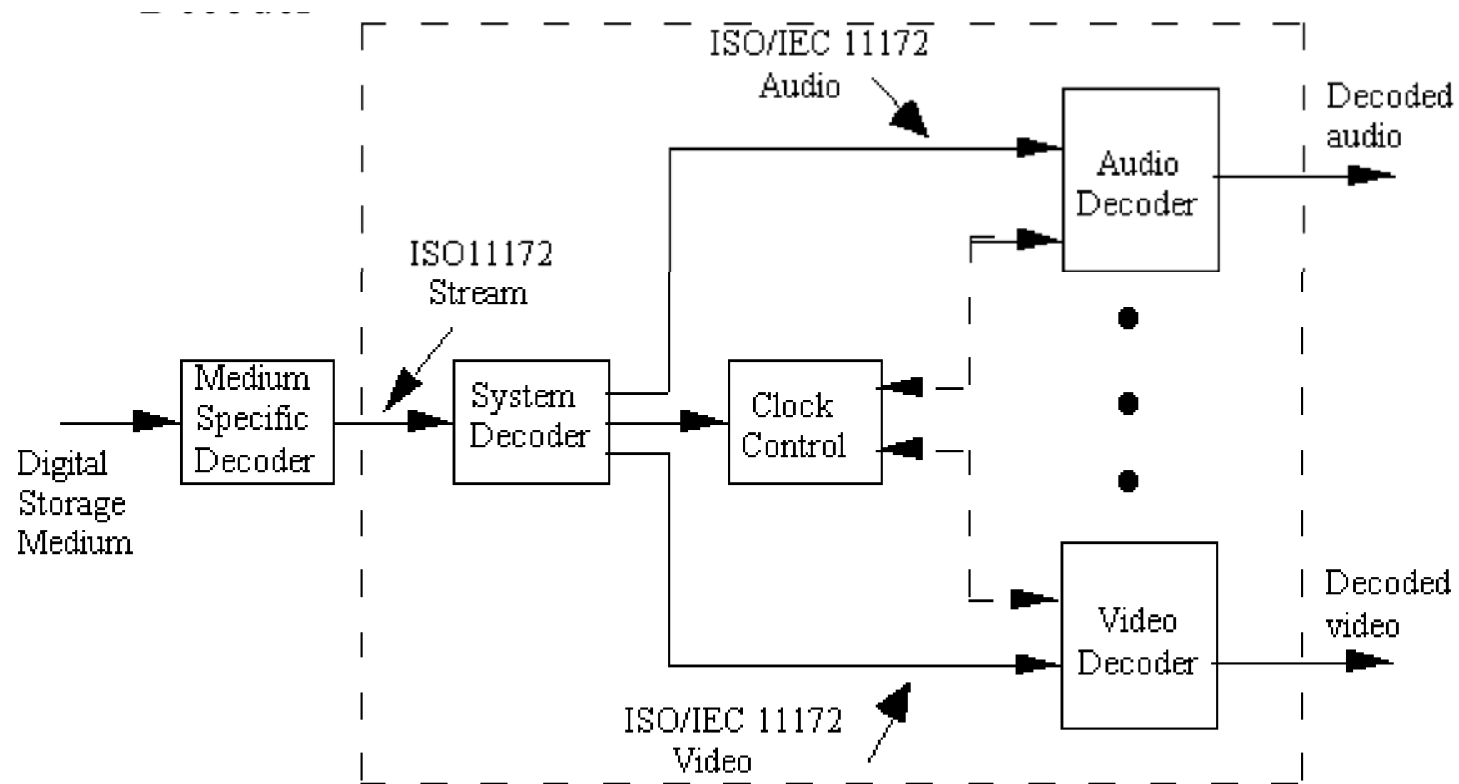
- MPEG = Moving Picture Expert Group
  - MPEG-1 (ISO/IEC 11172, Nov 1992)
  - MPEG-2 (ISO/IEC 13818, Nov 1994)
  - MPEG-4 (ISO/IEC 14496, Oct 1998)
  - MPEG-7 (ISO/IEC N6828, Oct 2004)
  - MPEG-21 în curs de elaborare (definitivare 2012)
- Standardul MPEG1 a fost dezvoltat pentru **compresia video și a sunetului pentru stocarea pe CD**.
  - CD-ROM 1x: 150kB/s = 1,2 Mbiți/s.
  - MPEG-1 tinde la un flux de date video comprimate de aprox. **1,2Mbiți/s**.
  - Datele audio sunt codate cu **32 - 448kbiți/s**.
- MPEG-2 este utilizat pentru stocarea pe **DVD** și în **televiziunea digitală**: SDTV, HDTV.
  - Rata de bit variabilă între **4 și 80Mbiți/s**. (optim: 4Mbiți/s).
  - Suportă codare video **întrepătrunsă și scalabilă**.

## 5.1. MPEG-1

- Cerințe de bază:
  - Codarea video de **calitate bună** (aproape calitatea VHS) la rate între **1 și 1,5Mbiți/s**.
  - **Acces aleator** la un cadru într-un timp limitat (puncte de acces în fluxul video).
  - **Derulare înainte/înapoi** (căutare și afișare folosind punctele de acces).
  - Sistemul să suporte **sincronizarea audio-video**.
- Cerințe suplimentare:
  - **Redare inversă** (necesită decodarea și stocarea cadrelor între punctele de acces).
  - Compromis între **calitatea codării** și **întârzierea la decodare**.
  - Schema de codare și organizarea biților să fie **robustă la erori**.
  - **Formatele imaginilor** să suporte o varietate de aplicații.
  - Codorul în timp real să aibă un **cost rezonabil**.
- Standardul MPEG ia în considerare și implementările din alte standarde:
  - **JPEG**. Deoarece o secvență video poate fi privită ca o **secvență de cadre (imagini fixe)** și deoarece JPEG a fost dezvoltat înaintea MPEG, acesta utilizează ideile din JPEG.
  - **H.261**. Întrucât H.261 era deja disponibil în timpul elaborării MPEG, s-a încercat **compatibilitatea** cu acest standard.

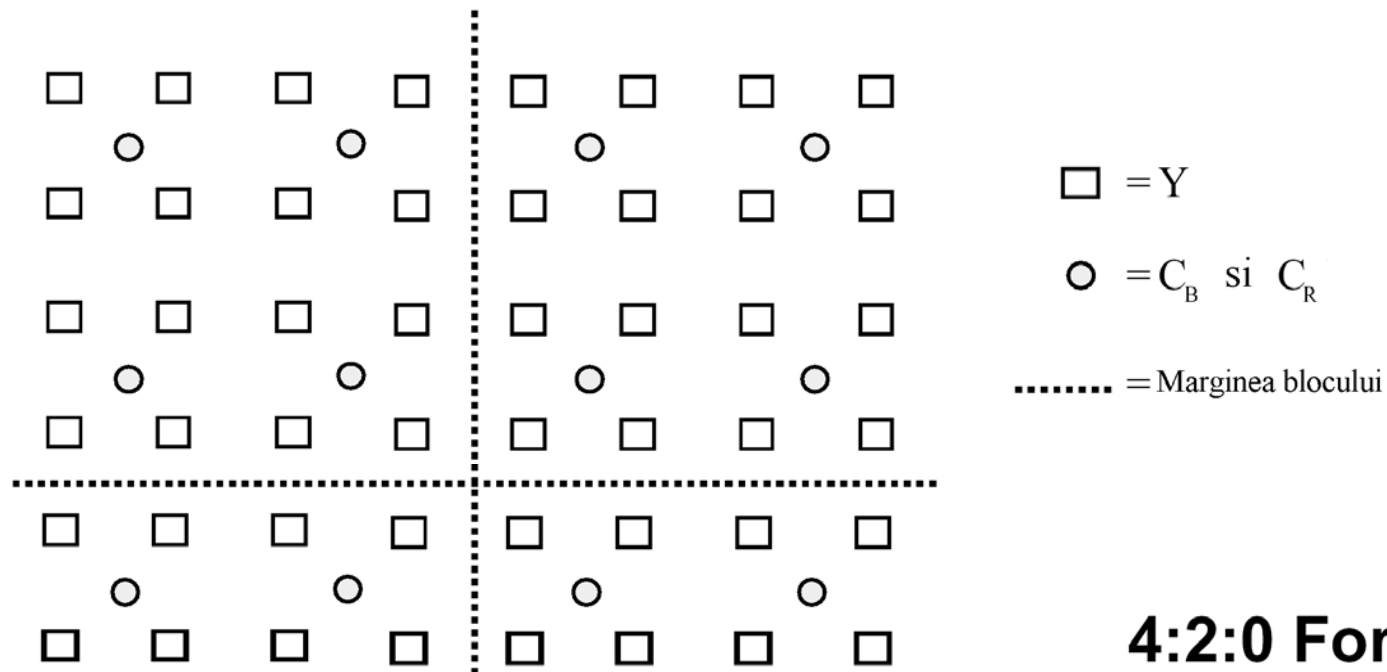
- Facilități noi față de H.261:
  - Dimensiuni diverse ale imaginilor și rate de bit variabile.
  - Structurarea în “felii” (slice) a datelor în loc de GOB.
  - Grupuri de imagini (GOP – Group Of Pictures).
  - Compensare bidirecțională a mișcării.
  - Compensare a mișcării “half pixel”.
  - Tipuri diferite de MacroBlocuri și tabele VLC pentru imaginile I, P și B.
  - Tabele de cuantizare.
  - Codare VLC pentru diferențele vectorilor de mișcare MV.

- Sistemul MPEG include codarea **video**, **audio**.
- Specifică formarea **fluxului de biți** și **decodarea**:



## Codarea video

- Fiecare imagine constă din **trei componente YCbCr** (asemănător formatului YUV).
- Componenta de luminanță are de 2 ori mai multe eșantioane pe orizontală și verticală. Acest procedeu se mai numește și “**color subsampling**”.



**4:2:0 Format**

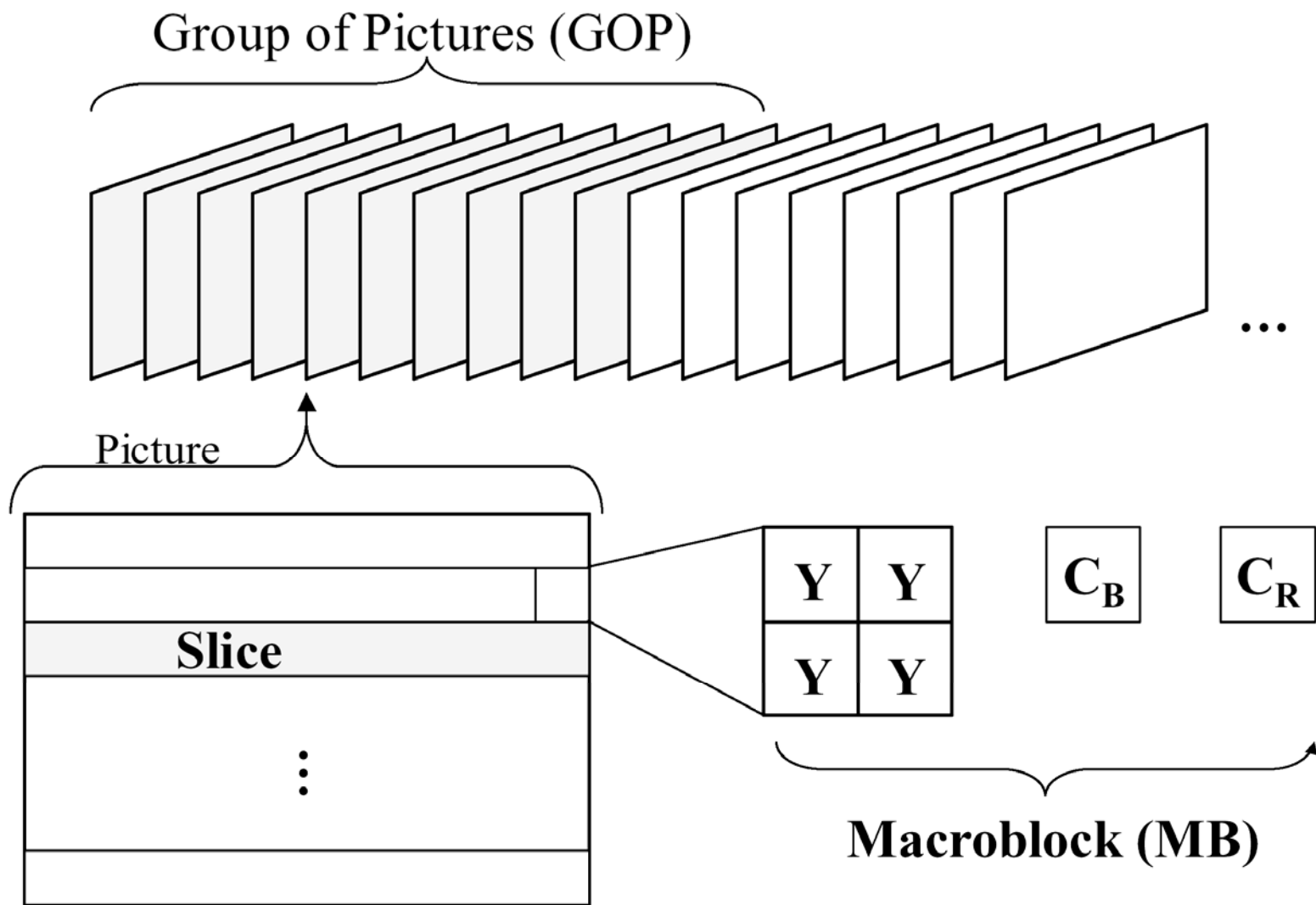
- Rezoluția componentei de luminanță < **976 x 768 pixeli**
- fiecare pixel este codat cu **8 biți**.
- Fluxul de date MPEG conține mai multă informație decât JPEG.
- **Raportul de aspect** pentru imagine:
  - pentru cadre cu 702 x 575 pixeli sau 711 x 487 se definește un **raport 4:3**
  - pentru cadre cu **625 linii** (sistemele HDTV europene) sau **525 linii** (sistemele HDTV din SUA) se definește un **raport 16:9**.
- **Frecvența cadrelor**:
  - 23,976Hz; 24Hz (cinema); 25Hz (PAL); 29,97Hz (NTSC); 30Hz; 50Hz; 59,94Hz; 60Hz.

Principiul compresiei:

- predicția temporală a carelor
- Imaginile animate conțin adesea **mișcări netransiționale** ca de exemplu rotația unui obiect. Asemenea mișcări pot fi reduse printr-o **codare intra-cadru**.
- În cele mai multe cazuri o **codare predictivă** are sens pentru o **parte a unui cadru** și nu pentru întreaga cadru.

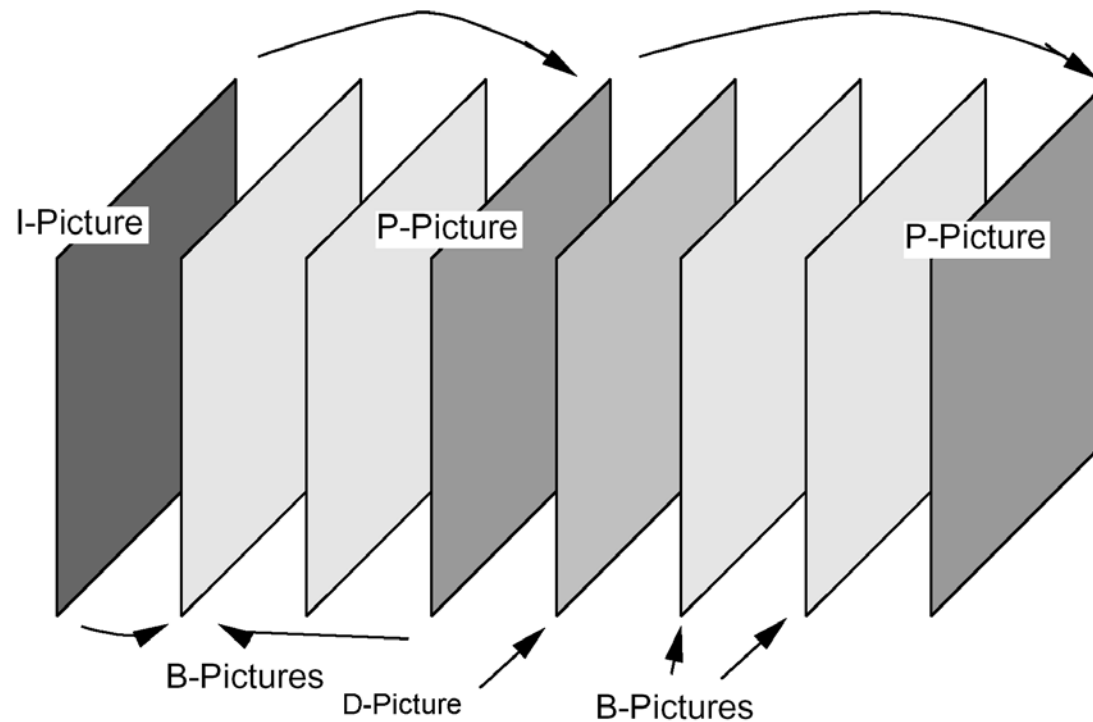
## Structura datelor:

- Fiecare imagine este împărțită în suprafețe numite **macro-blocuri (MB)**.
- Fiecare **macro-bloc** este format din:
  - 16 x 16 pixeli pentru luminanță.
  - 8 x 8 pixeli pentru componentele de crominanță.
- Macro-blocurile sunt ordonate de la stânga la dreapta și de sus în jos în **slice-uri**.
- Dacă fluxul de biți conține o eroare, decodorul poate sări până la începutul următorului **slice**.
- Mai multe **slice-uri** - comportament mai bun la erori dar mai mulți biți de codare.
- **Imaginea** constă din trei matrici dreptunghiulare reprezentând luminanța Y și cele două valori ale crominanței  $C_B$  și  $C_R$ .
- Matricea Y - număr par de linii și coloane.
- Matricile de crominanță - jumătate din dimensiunea matricei Y pe ambele direcții.
- **Grup de imagini** - unitatea care permite accesul aleator la fluxul video.
  - header + unul sau mai multe cadre I, P și/sau B.





- MPEG definește 4 tipuri de cadre:



- Ideea celor 4 tipuri de imagini provine de la faptul că cerințele de **codare eficientă** și un **acces aleator rapid** sunt **contradictorii**.
- rate mari de compresie - exploatarea redundanțelor temporale ale cadrelor (codare inter-cadre).

## Cadre I (Intra-coded frames)

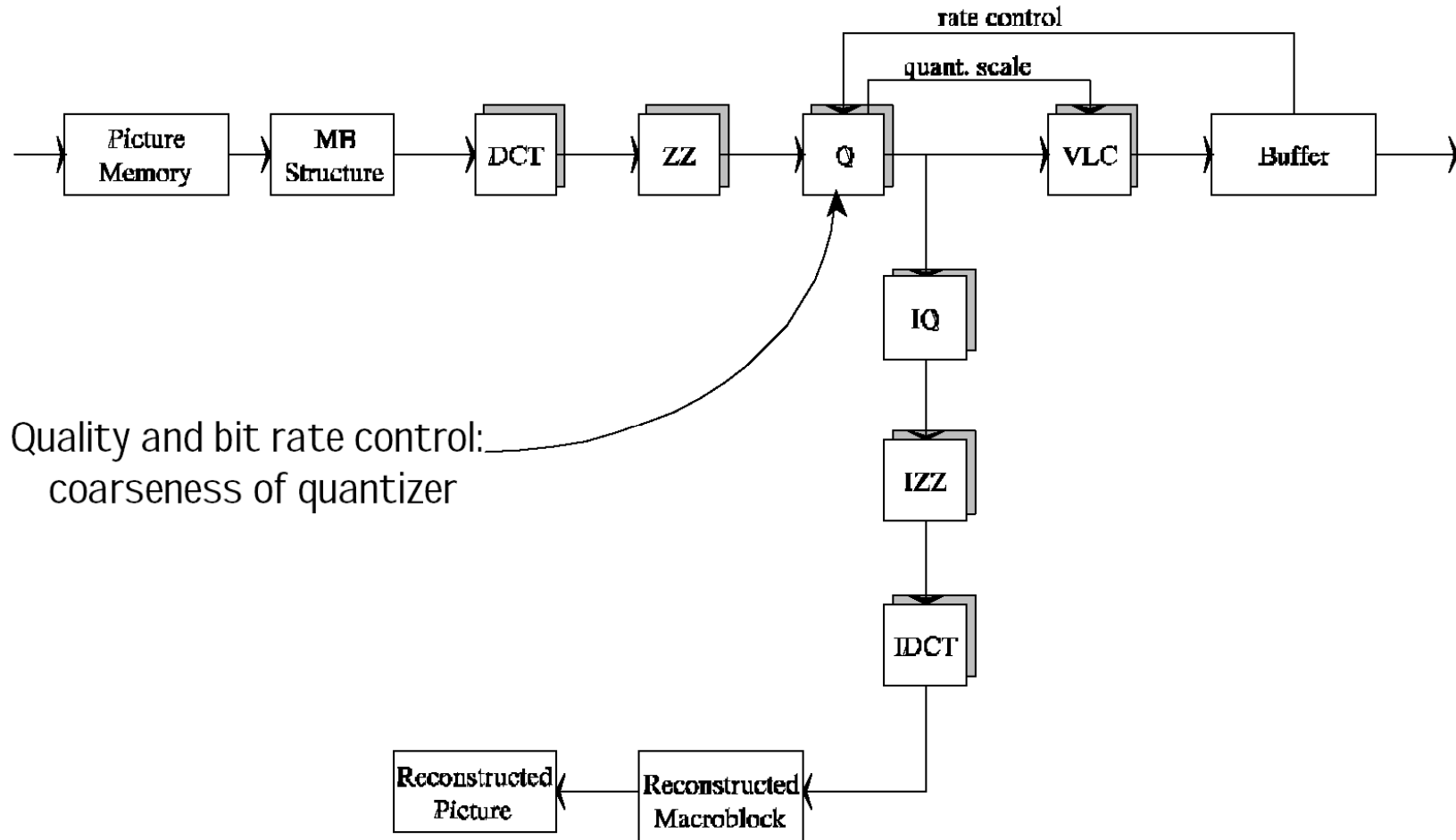
- codate **fără referințe la alte cadre** (imagine fixă).
- MPEG folosește codarea **JPEG** pentru aceste cadre.
- compresia trebuie făcută în **timp real** - Rata de compresie a cadrelor I < **imagini JPEG**.
- Cadrele I sunt **punctele de acces aleator** în fluxul MPEG.
- DCT se calculează pe fiecare bloc de 8x8 pixeli.
- Coeficienții DC sunt codați DPCM.
- Coeficienții AC sunt scanați în zigzag și codați PCM.
- Fluxul de coeficienți este codat zero run – length folosind tabele VLC standard
- Pasul de cuantizare poate fi controlat și adaptat prin:
  - parametrul Q (factor de calitate) – se poate schimba la slice sau la macro-bloc.
  - matrice de cuantizare – se poate schimba la grupuri de imagini (GOP).

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

- Codarea coeficienților DC intra-cadre:

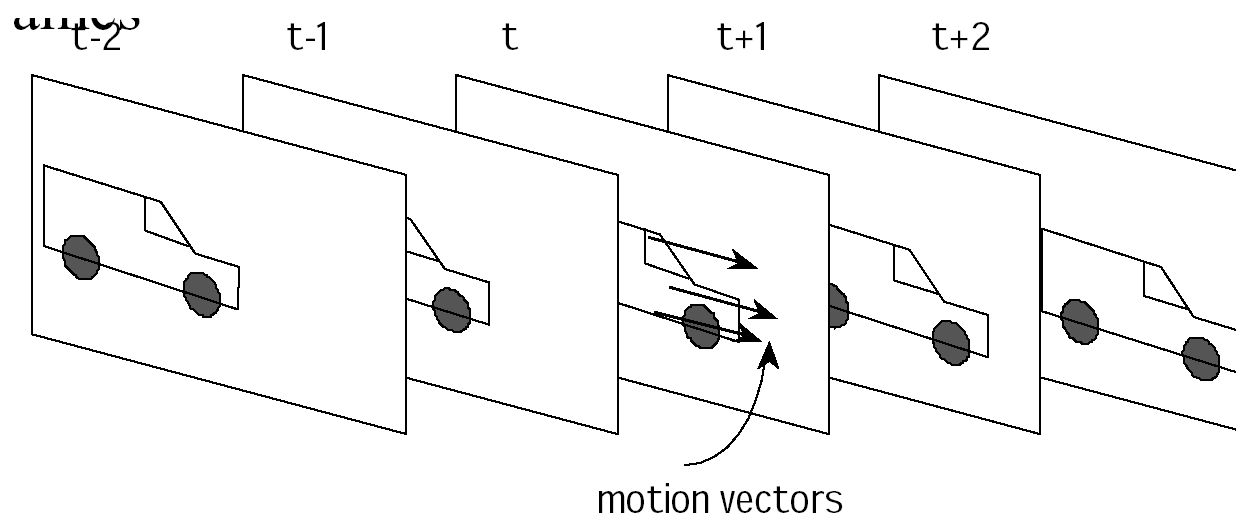
Differential DC	SIZE	VLC Luminance	VLC Chrominance
-255 to -128	8	1111 110	1111 1110
-127 to -64	7	1111 10	1111 110
-63 to -32	6	1111 0	1111 10
-31 to -16	5	1110	1111 0
-15 to -8	4	110	1110
-7 to -4	3	101	110
-3 to -2	2	01	10
-1	1	00	01
0	0	100	00
1	1	00	01
2 to 3	2	01	10
4 to 7	3	101	110
8 to 15	4	110	1110
16 to 31	5	1110	1111 0
32 to 63	6	1111 0	1111 10
64 to 127	7	1111 10	1111 110
128 to 255	8	1111 110	1111 1110

## Codarea cadrelor I

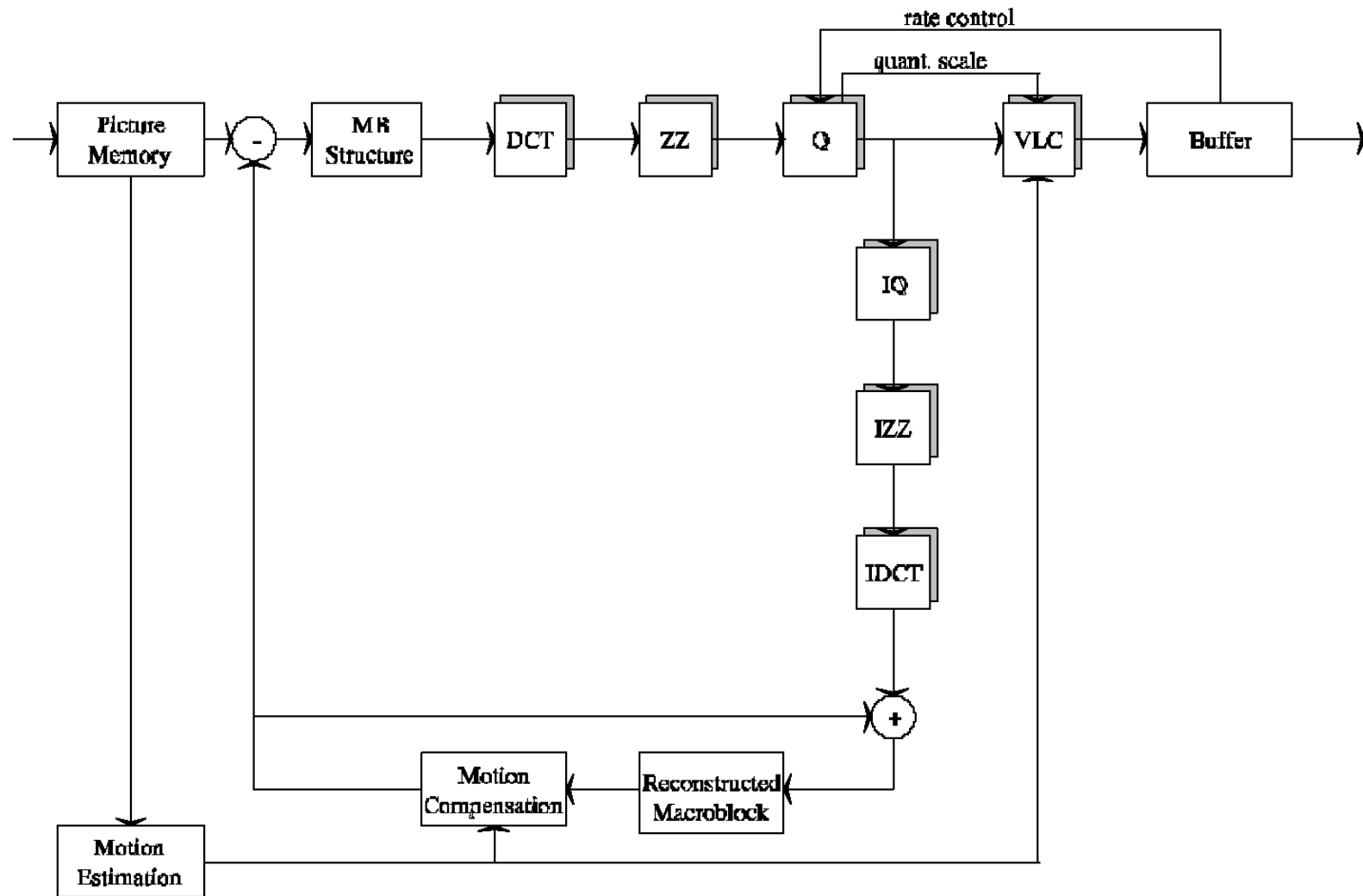


## Cadre P (Predictive-coded frames)

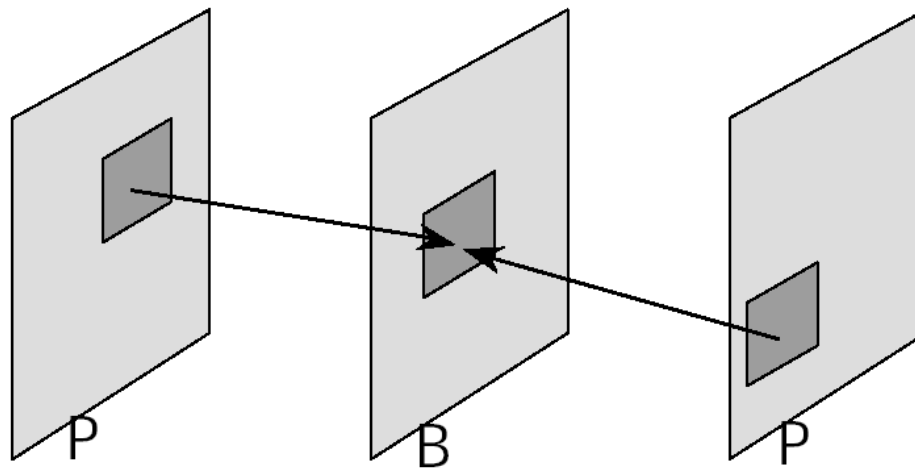
- Necesită informații din cadrele I sau P anterioare pentru codare și decodare.
- Codarea cadrelor P este bazată pe faptul că la cadre succesive, adesea suprafețele nu se schimbă de tot și **întreaga suprafață este deplasată** (redundanță temporală)
- **Mai multe metode** pentru estimarea mișcării
- MPEG nu prevede un algoritm pentru estimarea mișcării dar specifică rezultatele codării.
- Sunt codate numai **vectorul de mișcare** (diferența între localizarea temporală a macroblocului) și **diferența** dintre macroblocul curent și cel obținut prin estimarea mișcării.



- cadrele P conțin macroblocuri **codate intra** (ca și cadrele I) și macroblocuri **codate predictiv**
- un macrobloc este format din 6 blocuri de 8x8.
- Codorul trebuie să determine **dacă** un macrobloc trebuie să fie **codat predictiv** și dacă există un **vector de mișcare** care trebuie codat.
- Codorul specific macrofluxurilor P trebuie să ia în considerare **diferența macroblocurilor** și **vectorul de mișcare**.
- Toate diferențele dintre cele 6 blocuri de 8x8 pixeli din cadrul curent și cele estimate dintr-un cadru anterior sunt transformate utilizând DCT bidimensională.
- Blocurile care au numai **coeficienți DCT 0** nu sunt prelucrate mai departe.
- Coeficienții sunt calculați utilizând **6 biți**
- Coeficienții AC și DC sunt codați utilizând o tehnică **diferită de JPEG** și de cea folosită la cadrele I
- În pasul următor este făcută o **codare run-length**.
- Când **vectorul de mișcare** al macroblocurilor adiacente **diferă puțin**, o codare **DPCM** este utilizată pentru codarea vectorului.
- Modurile de codare a macro-blocurilor din cadrele P:
  - non-intra / intra cadru
  - codarea sau ne-codarea vectorilor de mișcare
  - diferența cuantizată a predicției este zero sau nu.



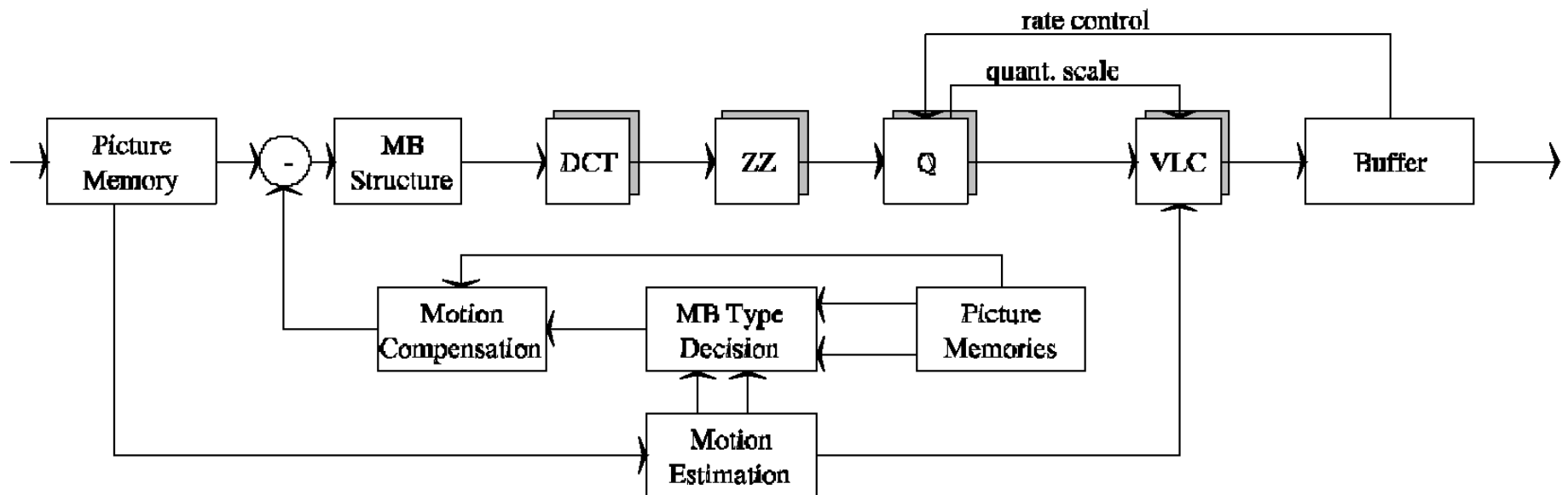
- **Cadre B (Bi-directionally predictive coded frames)**
- Necesită informații din **cadrele I și/sau P anterioare și următoare** pentru codare și decodare.



- **Cea mai mare rată de compresie** este atinsă la aceste cadre
- Pentru fiecare macrobloc se estimează **2 vectori de mișcare** (unul față de un cadru de referință anterior și unul următor)
- Sunt salvați cei doi vectori de mișcare și eventual diferența dintre macroblocul curent și cel estimat



- Un cadru B conține macroblocuri:
  - Codate predictiv bidirecțional
  - Codate predictiv unidirectional (cu referința anterioară sau următoare)
  - Codate intra
- Cadrul B nu poate fi niciodată accesat direct într-un mod aleator.



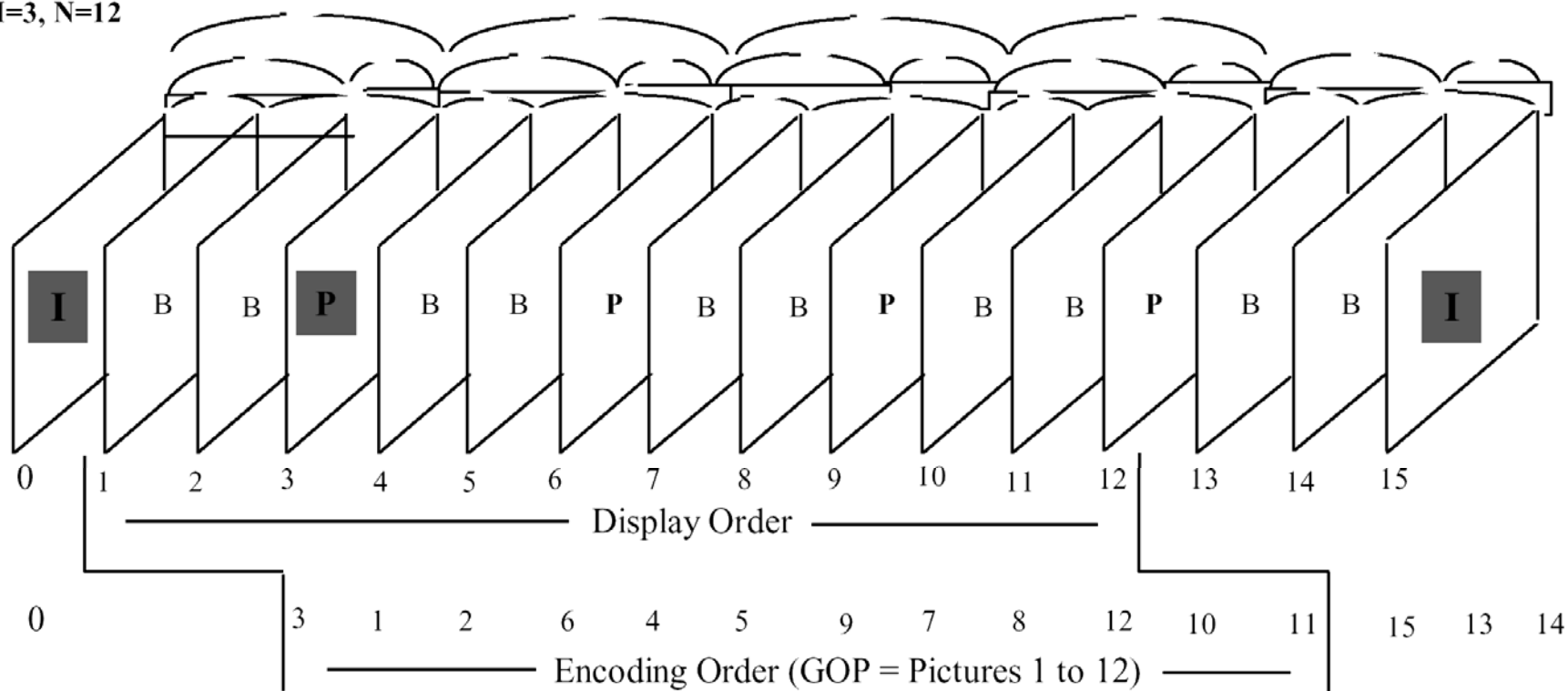
- Următorul exemplu ilustrează avantajele unei predicții bidirecționale.  
Într-o scenă video, o minge se mișcă de la stânga la dreapta în fața unui fundal static. În partea stângă a scenei părți ale imaginii care apar în imagini anterioare nu apăreau fiind acoperite de minge. O predicție a acestor suprafețe poate fi dezvoltată din următoarele imagini dar nu din imaginile anterioare. Un macro-bloc poate fi dezvoltat din macro-blocul anterior sau următor al cadrelor P sau I. În afară de un vector de mișcare de la imaginea anterioară la cea următoare, poate fi utilizat un vector de mișcare în celalaltă direcție.

## Cadre D (DC – coded frames)

- Sunt cadre codate **intra-cadru**
- folosite pentru **derulare rapidă** înainte sau înapoi.
- Numai parametrii **DC** ai DCT sunt codați, coeficienții **AC sunt neglijați**.
- Cadrele D constau numai din componentele de **joasă frecvență** ale cadrelor
- Derularea înainte sau înapoi se poate realiza și cu cadre I
- Pentru aceasta cadrele I trebuie să apară periodic în fluxul de date
- Redarea înapoi însă necesită o **cantitate imensă de stocare** deoarece toate imaginile dintr-un grup trebuie decodate în modul forward și stocate, după care redarea înapoi este posibilă.

- Dacă se utilizează predicția cu cadre B **ordinea imaginilor la codare diferă de cea în timp**
- pentru a fi afișat un cadru B înainte de un cadru P, cadrul P trebuie decodat înaintea lui B deoarece el conține date necesare decompresiei lui B.
- Acest fapt introduce o întârziere

M=3, N=12



N = Numărul de imagini în GOP.

M = Distanța de predicție (M – 1 = numărul de cadre B consecutive)

- Configurația de cadre I, P, B este determinată de MPEG.
- Pentru acces rapid, aleator, cu rezoluție cât mai bună, ar trebui codat numai cu cadre I.
- Pe de altă parte cea mai mare compresie se obține când utilizăm cât mai multe cadre B.
- O secvență de tipul **IBBPBBPBBIBBPBBPBB...** are posibilitatea de acces aleator de 9 imagini (aproximativ 330ms) și are o compresie foarte buna pentru că folosește multe cadre B.
- În ceea ce privește cuantizarea, coeficienții AC ai cadrelor B și P au valori mici, pe când cei ai cadrelor I au valori mai mari.
- Dacă volumul datelor crește peste un anumit prag, cuantizorul o să folosească pași de cuantizare mai mari.

- Tabelele de cuantizare:

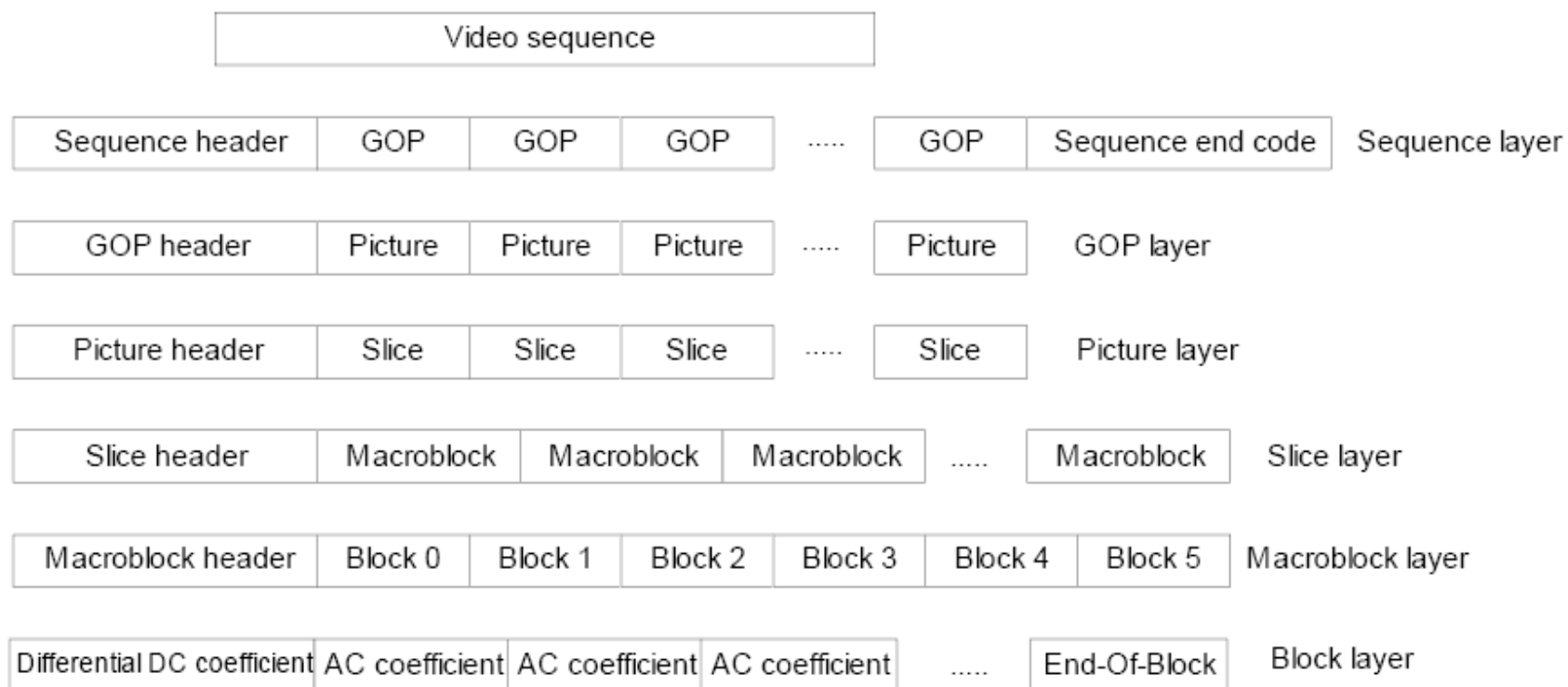
8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

**intra**

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

**inter**

## Structura fluxului de date MPEG-1



## Comparație H.261 – MPEG-1

H.261	MPEG-1
<ul style="list-style-type: none"><li>– acces secvențial</li><li>– o singură rată a cadrelor</li><li>– cadre CIF și QCIF</li><li>– cadre I și P</li><li>– MC folosind 1 cadru</li><li>– MV cu rezoluție de 1 pixel</li><li>– cuantizare uniformă</li><li>– nu există GOP</li><li>– structură GOB</li></ul>	<ul style="list-style-type: none"><li>– acces aleator</li><li>– rată a cadrelor flexibilă</li><li>– dimensiune flexibilă a cadrelor</li><li>– cadre I, P și B</li><li>– MC folosind 1 sau mai multe cadre</li><li>– MV cu rezoluție de 1/2 pixel</li><li>– matrice de cuantizare</li><li>– structură GOP</li><li>– structură „slice”</li></ul>