

بینایی کامپیوتر

فصل سوم

فشرده سازی ۳- ویدئو

VIDEO COMPRESSION

محمد جواد فدائی اسلام

بهمن ۱۴۰۳

فشرده سازی ویدئو

- حجم ویدئوی فشرده نشده بسیار زیاد است.
- راه حل: کدگذاری پیشگو (Predictive Coding) مبتنی بر فریم قبلی است.
- چون قسمت زیادی از تصاویر ویدئویی ثابت می ماند. در این صورت تفاضل باعث می شود هیستوگرام سیگنال خطا در نقطه صفر تیزی بیشتری داشته باشد که باعث کاهش آنتروپی می شود.

فشرده سازی ویدئو (ادامه)

JPEG = حذف افزونگی فضایی

ویدئو = حذف افزونگی زمانی و فضایی

- افزونگی زمانی: اغلب محتوای فریم های پشت سرهم یکسان است.
- در ویدئو یک فریم ذخیره نمی شود بلکه اختلاف آن ذخیره می شود.

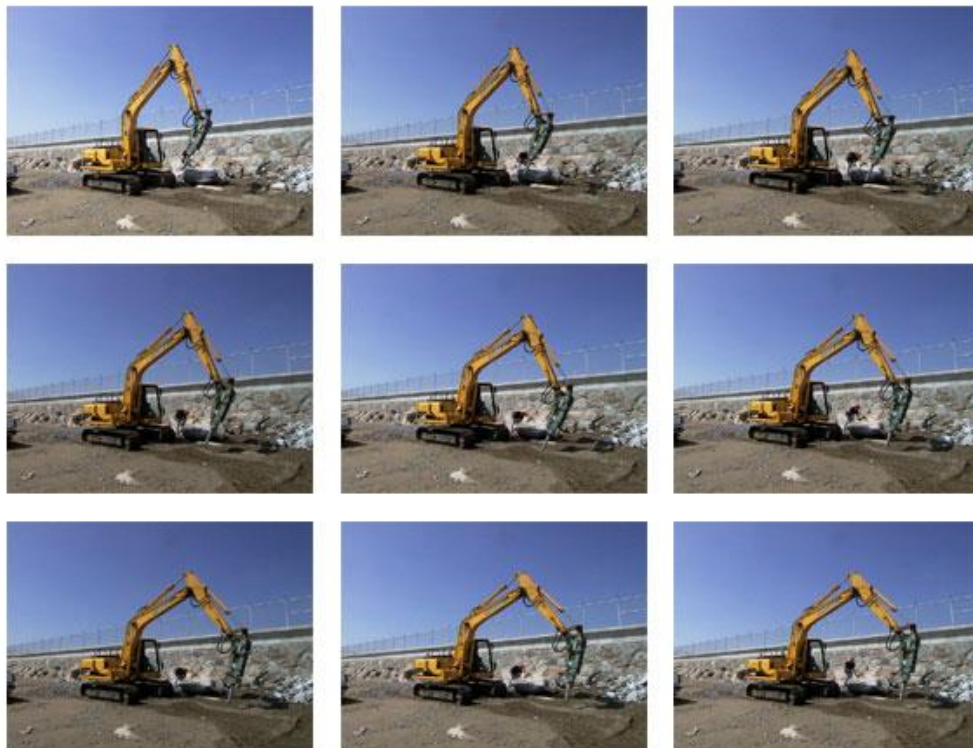
فشرده‌سازی ویدئویی (افزونگی فضایی)

• Spatial Redundancy (افزونگی فضایی) پیکسل‌های کنار هم

شبیه هم هستند



- فشرده‌سازی ویدئویی (افزونگی زمانی)
- Temporal Redundancy (افزونگی زمانی)
 - فریم‌های متوالی در یک ویدئو، مشابه هستند.



فشرده سازی ویدئو (ادامه)

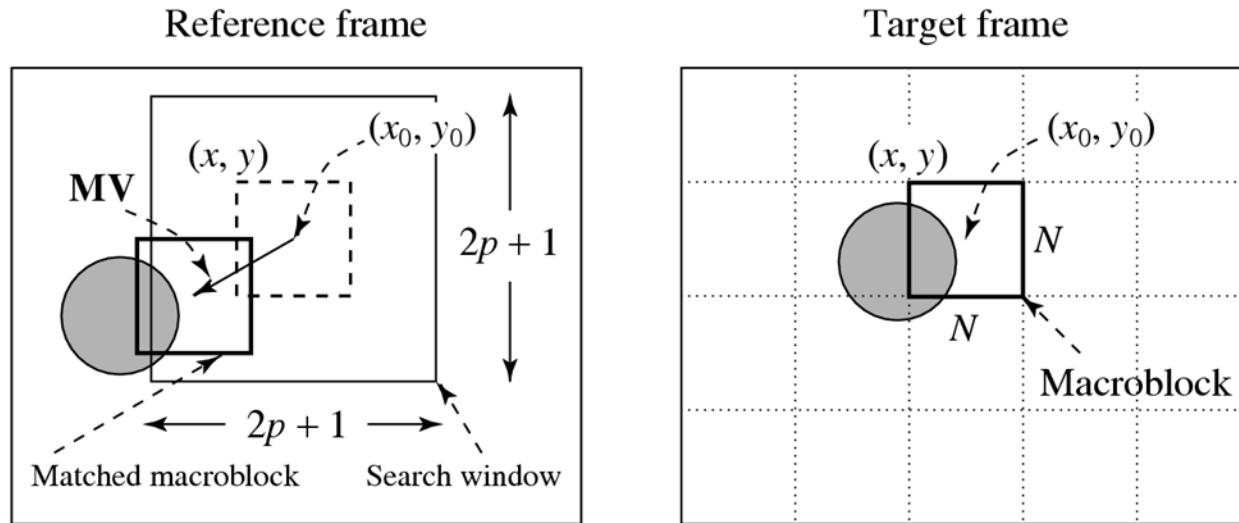
راه حل بهتر جست و جوی قسمتی از تصویر در فریم بعد به منظور تفریق بهتر است.

ایده یافتن یک جسم در فریم بعدی = Motion Estimation

انتقال تکه های تصویر به منظور جبران حرکت به صورتی که اختلاف تفریق دو تصویر

کمینه شود = Motion Compensation

جستجوی ماکرو بلاک ها



جستجوی ماکرو بلاک ها یک مساله تطبیق (Matching) یا مساله یافتن متناظر (Correspondence Problem) است.

فشرده‌سازی ویدئویی توسط جبران حرکت

- به جای ذخیره سازی فریم، تفاوت بین فریم فعلی و سایر فریم ها ذخیره می شود.
- مراحل فشرده‌سازی ویدئویی بر اساس جبران حرکت (Motion Compensation):
 - ۱- برآورد حرکت (جستجوی بردار حرکت)
 - ۲- پیش بینی مبتنی بر جبران حرکت (MC)
 - ۳- به دست آوردن خطای پیش‌بینی (تفاوت)

جستجوی بردارهای حرکتی

- میانگین قدرمطلق تفاضل ها (MAD) :

$$MAD(i, j) = \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} |C(x+k, y+l) - R(x+i+k, y+j+l)|$$

- N - اندازه ماکروبلوک

- k و l - مکان ماکروبلوک در فریم هدف

- i و j - جابجایی‌های عمودی و افقی ماکروبلوک در فریم مرجع برای یافتن بهترین تناظر

- $C(x+k, y+l)$ - پیکسل‌های موجود در ماکروبلوک در فریم هدف

- $R(x+i+k, y+j+l)$ - پیکسل‌های موجود در ماکروبلوک در فریم مرجع

هدف این جستجو پیدا کردن یک بردار (i, j) به عنوان بردار حرکتی $\mathbf{MV} = (\mathbf{u}, \mathbf{v})$ به صورتی است که $MAD(i, j)$ حداقل باشد.

$$(u, v) = [(i, j) | MAD(i, j) \text{ is minimum, } i \in [-p, p], j \in [-p, p]]$$

دو سازمان مهم رایانه دهنده استانداردهای ویدئویی

- Video Coding Expert Group (VCEG)

- گروه بالا متعلق به مجموعه

ITU – International Telecommunication Union

- Moving Pictures Expert Group (MPEG)

- گروه بالا متعلق به مجموعه‌های

ISO/IEC

International Organization for Standard (ISO)

International Electrotechnical Commission (IEC)

استانداردهای مهم ویدئو

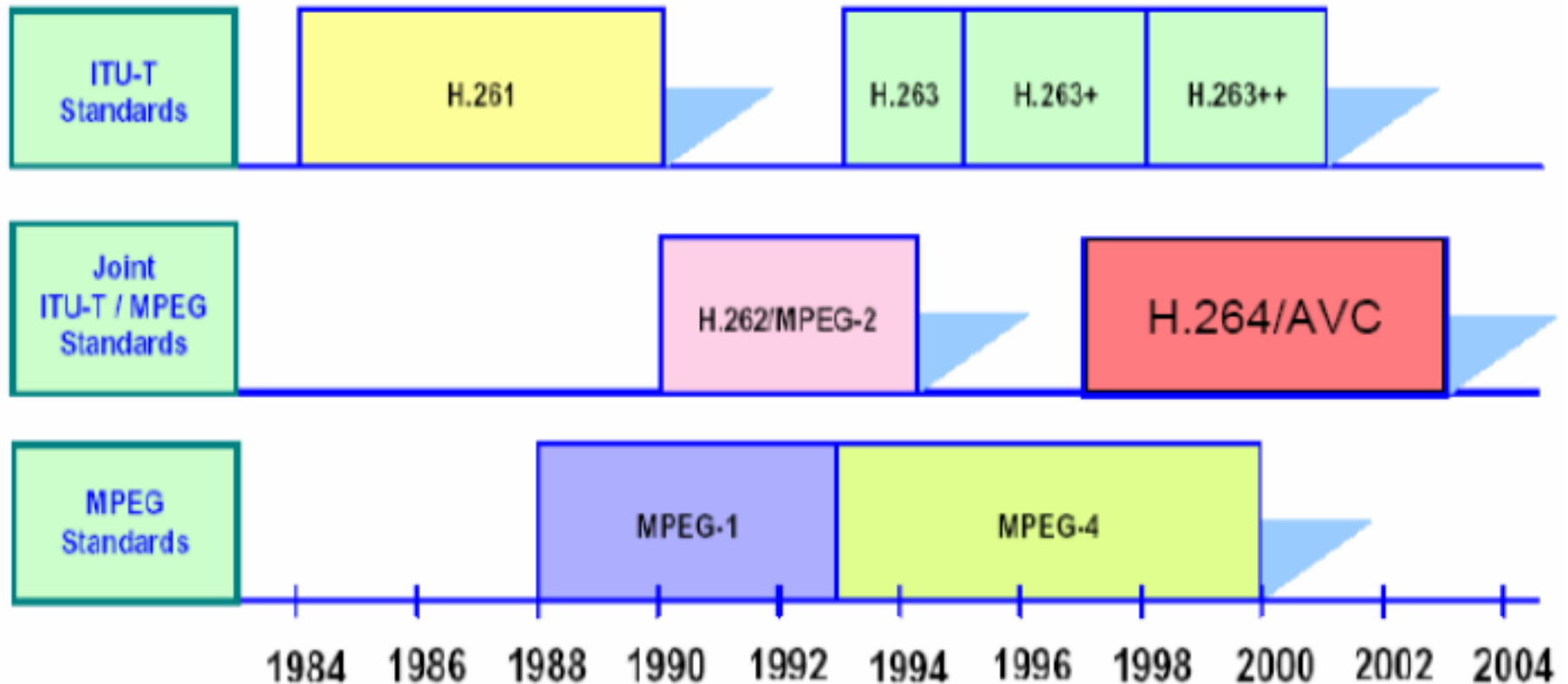


Figure 1. Progression of the ITU-T Recommendations and MPEG standards.

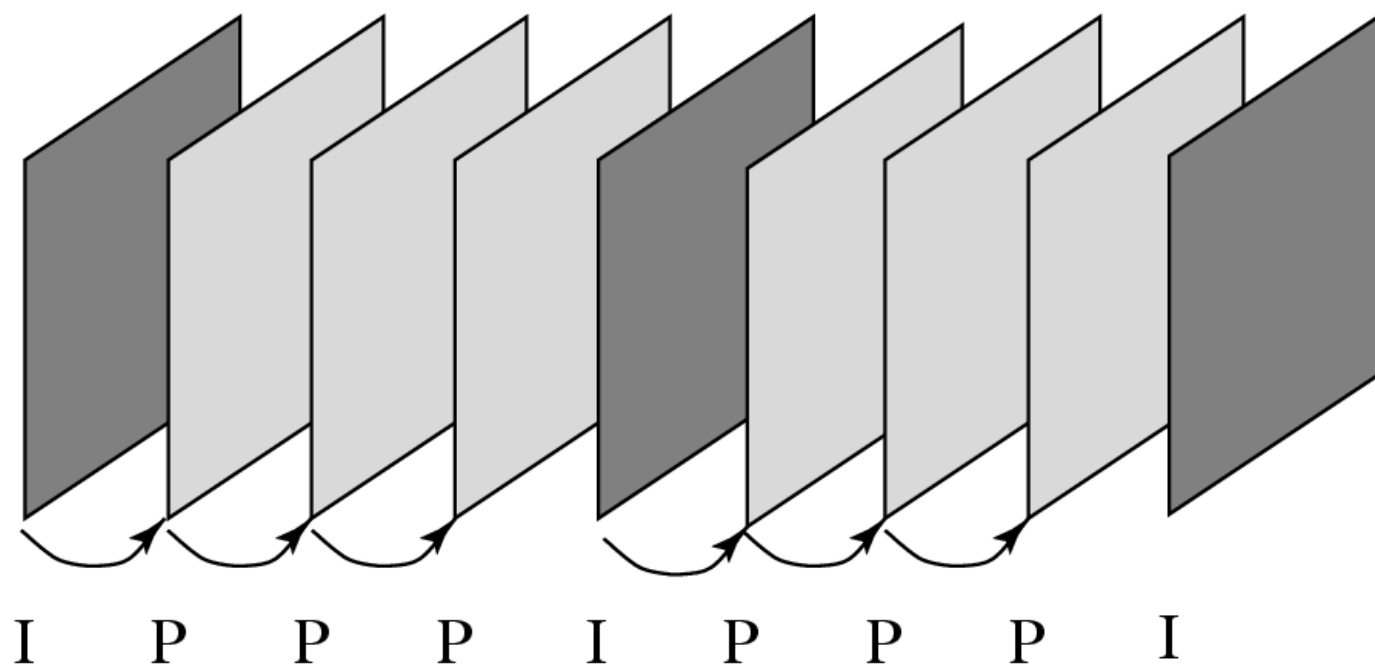
H.264 is also referred to as MPEG-4 Part 10 Advanced Video Coding



استاندارد H.261

H.261

- یک استاندارد اولیه برای فشرده‌سازی دیجیتالی ویدئو است، از جبران حرکت برای فشرده‌سازی استفاده می‌کند. همه استانداردهای فشرده‌سازی ویدئویی از این تکنیک استفاده می‌کنند.
- در سال ۱۹۸۸ این استاندارد عرضه شد.
- این استاندارد برای videophone و video conferencing و سایر خدمات سمعی بصری از طریق ISDN طراحی شده بود.
- این video codec از نرخ بیت $p \times 64$ kbps پشتیبانی می‌کند (p از ۱ تا ۳۰).
- لازم است که تاخیر کدگذار ویدئویی کمتر از ۱۵۰ میلی ثانیه باشد تا اینکه ویدئو بتواند برای video conferencing بلادرنگ دو جهت استفاده شود.
- نمونه برداری رنگ: ۴:۲:۰



H.261 ترتیب فریم

ترتیب فریم H.261

○ دو نوع فریم تصویری تعریف شده است:

- درون فریمی (I-frames)
- میان فریمی (p-frames)

○ (I-frames): تصویر مستقل، کدگذاری مشابه JPEG

○ (p-frames): وابسته، با روش کدگذاری پیشگویانه رو به جلو کدگذاری می‌شود.

○ حذف افزونگی زمانی در کدگذاری p-frames وجود دارد.

○ کدگذاری I-frame فقط حذف افزونگی فضایی را انجام می‌دهد.

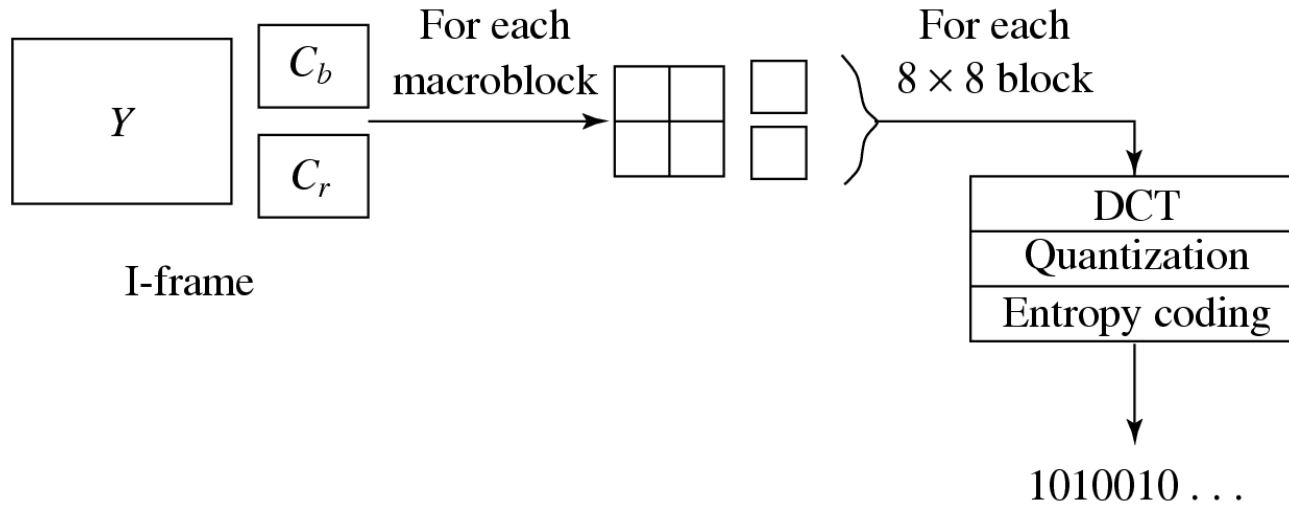
○ جهت اجتناب از انتشار خطاهای کدگذاری معمولاً یک I-frame چندین بار در هر ثانیه از ویدئو فرستاده می‌شود .

○ بردارهای حرکتی در H.261 همیشه در واحد از پیکسل اندازه‌گیری می‌شوند و $p=15$ است.



کدگذاری و رمزگشایی H.261

کدگذاری I-FRAME:

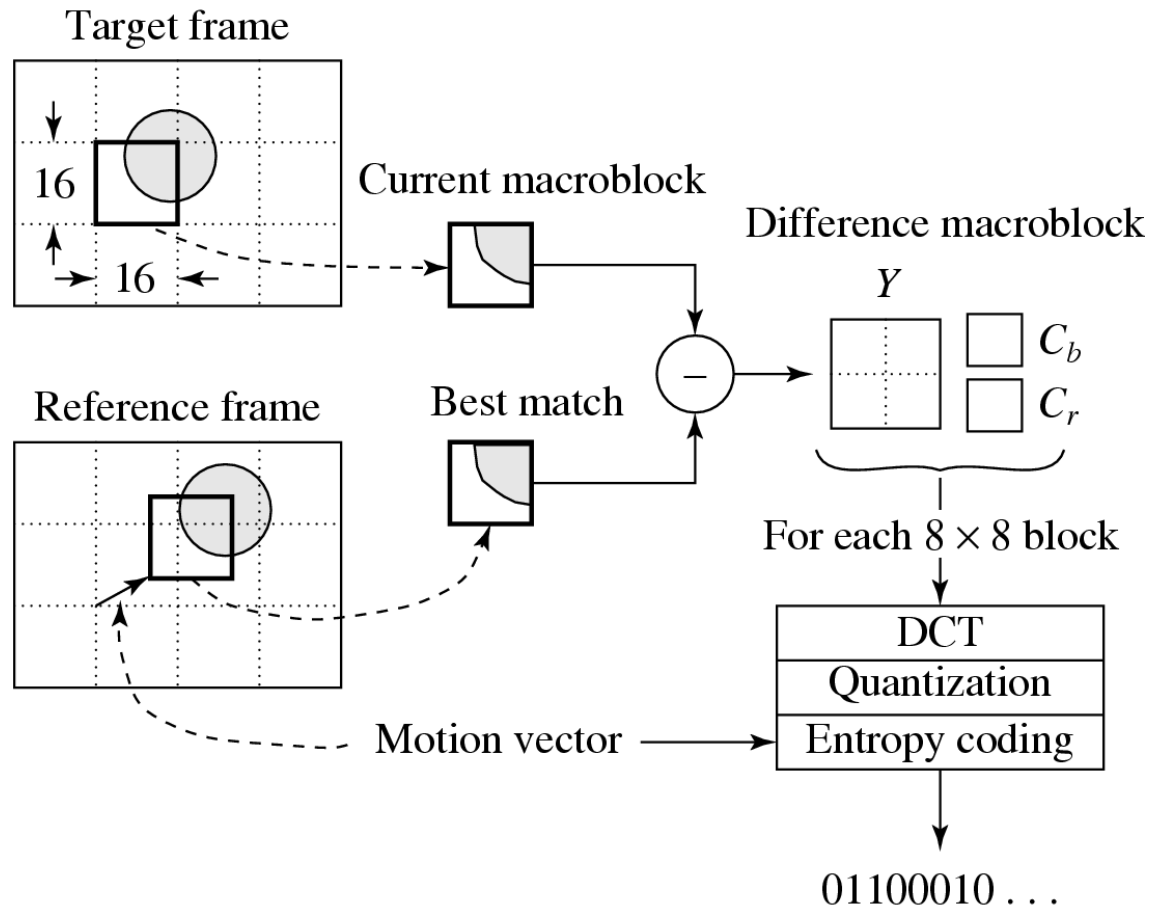


- ماکروبلوک ها به اندازه پیکسل های 16×16 برای فریم Y و 8×8 برای فریم های C_b و C_r هستند چون از نمونه برداری رنگی $4:2:0$ استفاده می شود.
- برای هر بلاک 8×8 یک انتقال DCT استفاده می شود. سپس ضریب های DCT وارد اسکن زیگزاگی تدریجی و کدگذاری آنتروپی می شوند.

کدگذاری پیش‌بینی کننده (P-FRAME)، مبتنی بر جبران حرکت

- برای هر ماکروبلاک در فریم هدف یک بردار حرکتی توسط یکی از روش‌های جستجو که قبلاً توضیح داده شد اختصاص داده می‌شود.
- پس از پیش‌بینی تفریق صورت می‌گیرد تا خطا به دست آید.
- p-frame ماکروبلاک تفاضل را (نه خود ماکروبلاک هدف) کدگذاری می‌کند.
- هر یک از این بلاک‌های 8×8 وارد رویه‌های DCT، Quantization، اسکن زیگزاگی و رمزگشایی آنتروپی می‌شوند.

کدگذاری P-FRAME



QUANTIZATION در H.261

○ Quantization در H.261، برای همه ضریب‌های DCT در داخل یک ماکروبلوک از یک گام ثابت استفاده می‌کند.

○ اگر از DCT و QDCT برای نشان دادن ضریب‌های DCT پیش و بعد از تدریج استفاده کنیم برای ضریب‌های DC در فریم‌های p فرمول زیر را داریم:

$$QDCT = \text{round} \left(\frac{DCT}{\text{step_size}} \right) = \text{round} \left(\frac{DCT}{8} \right)$$

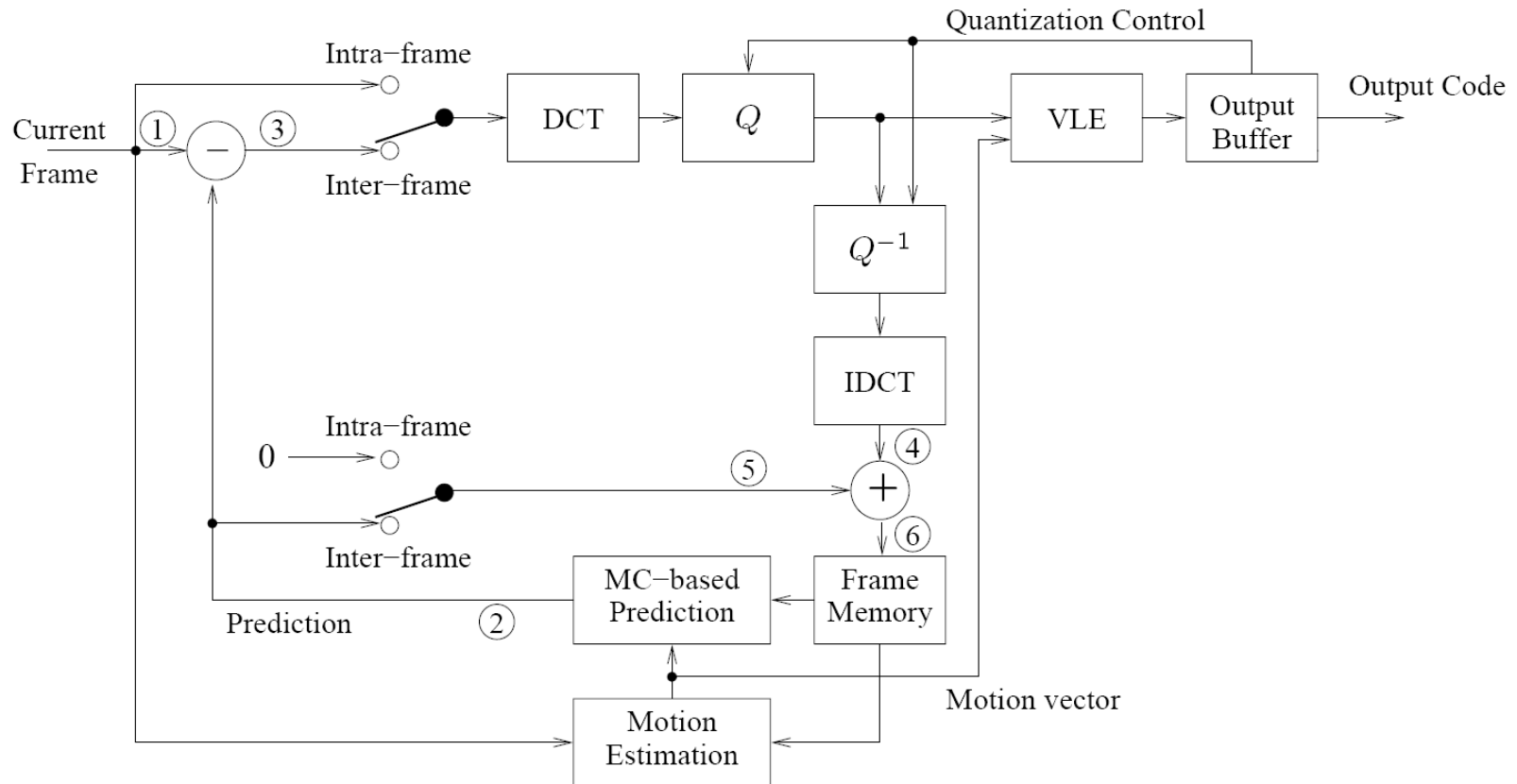
○ برای سایر ضریب‌ها:

$$QDCT = \left\lfloor \frac{DCT}{\text{step_size}} \right\rfloor = \left\lfloor \frac{DCT}{2 * \text{scale}} \right\rfloor$$

عدد ثابت

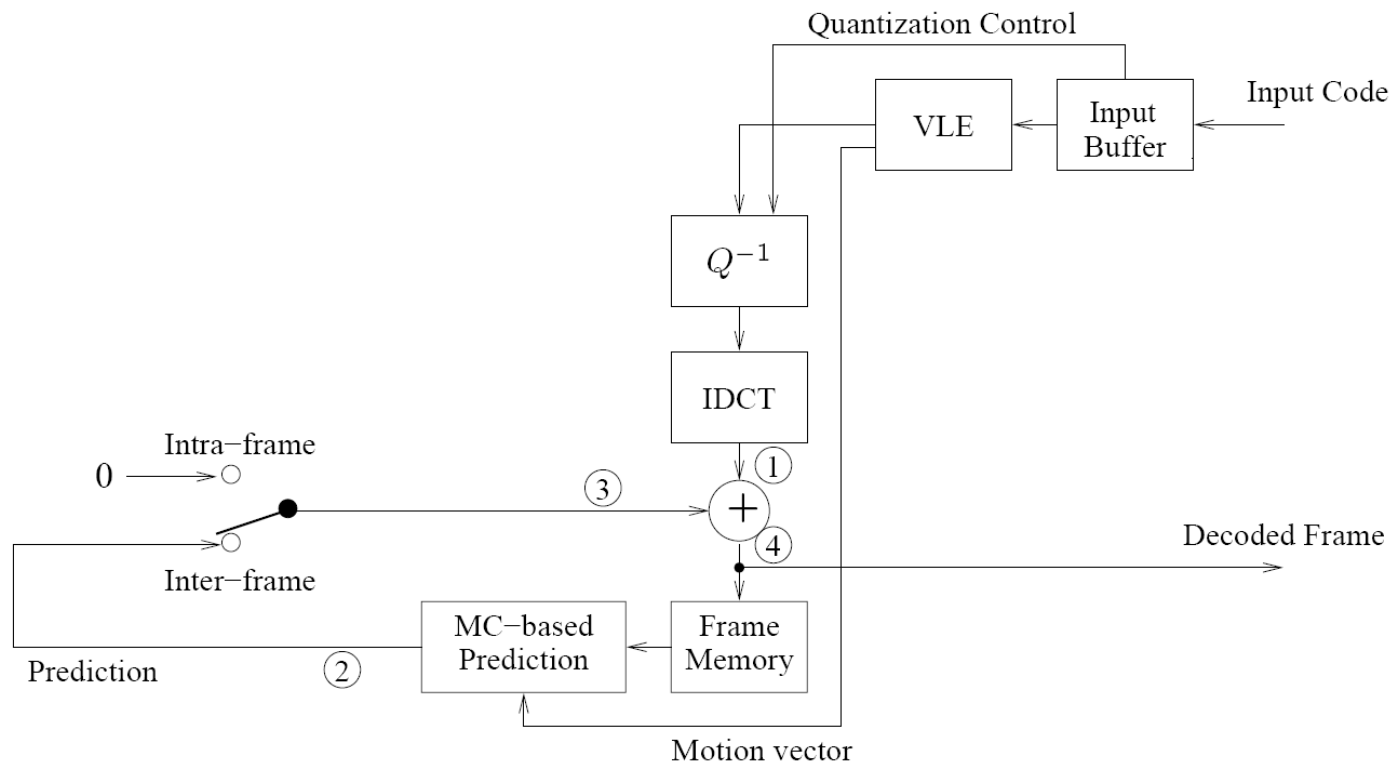
۲۰ مقیاس یک عدد صحیح در محدوده [1,31]

کدگذار H.261



(a) Encoder

کدگشا H.261



(b) Decoder



MPEG-1

○ در سال ۱۹۹۱ این استاندارد تایید شد.

○ سرعت 1.5mbps

1.2Mbps → video

256kbps → stereo voice

○ کیفیت ضبط در حد VHS و صدا در حد CD .

- 352 x 240 for NTSC video at 30 fps

- 352 x 288 for PAL video at 25 fps

○ MPEG-1 از نمونه برداری رنگی ۴:۲:۰ استفاده می کند.

○ در ذخیره سازی فریم روش interlaced را پشتیبانی نمی کند

MPEG-1

○ در H.261 دو نوع فریم داریم:

I-frame, P-frame

○ در MPEG1 سه نوع فریم داریم

I-frame, P-frame, B-frame

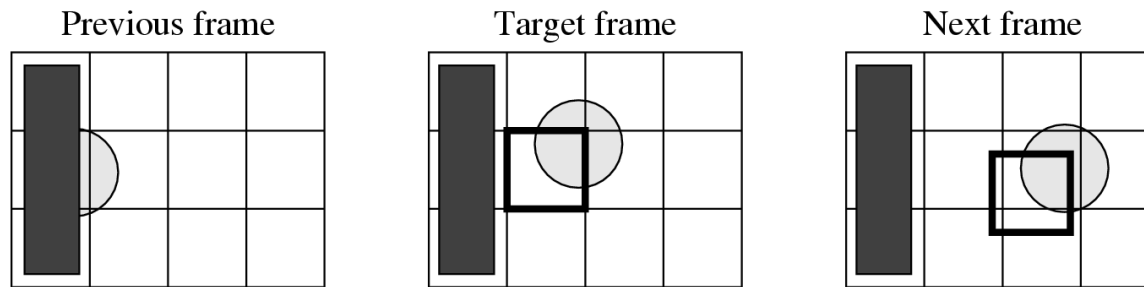
○ در H.261 پیشگویی روبه جلو داریم:

Forward prediction

○ در MPEG1 پیشگویی روبه جلو و عقب داریم

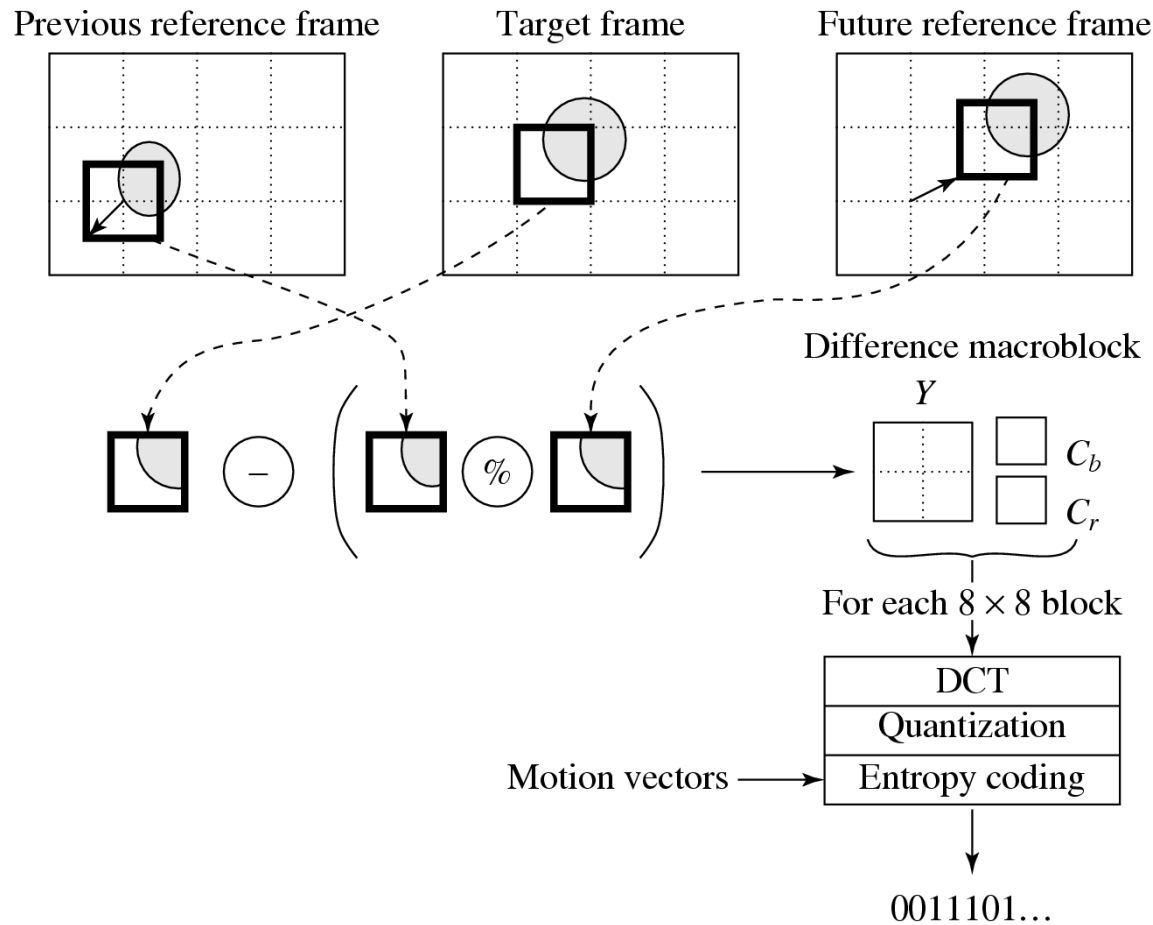
Forward & Backward prediction

جبران حرکت در MPEG-1

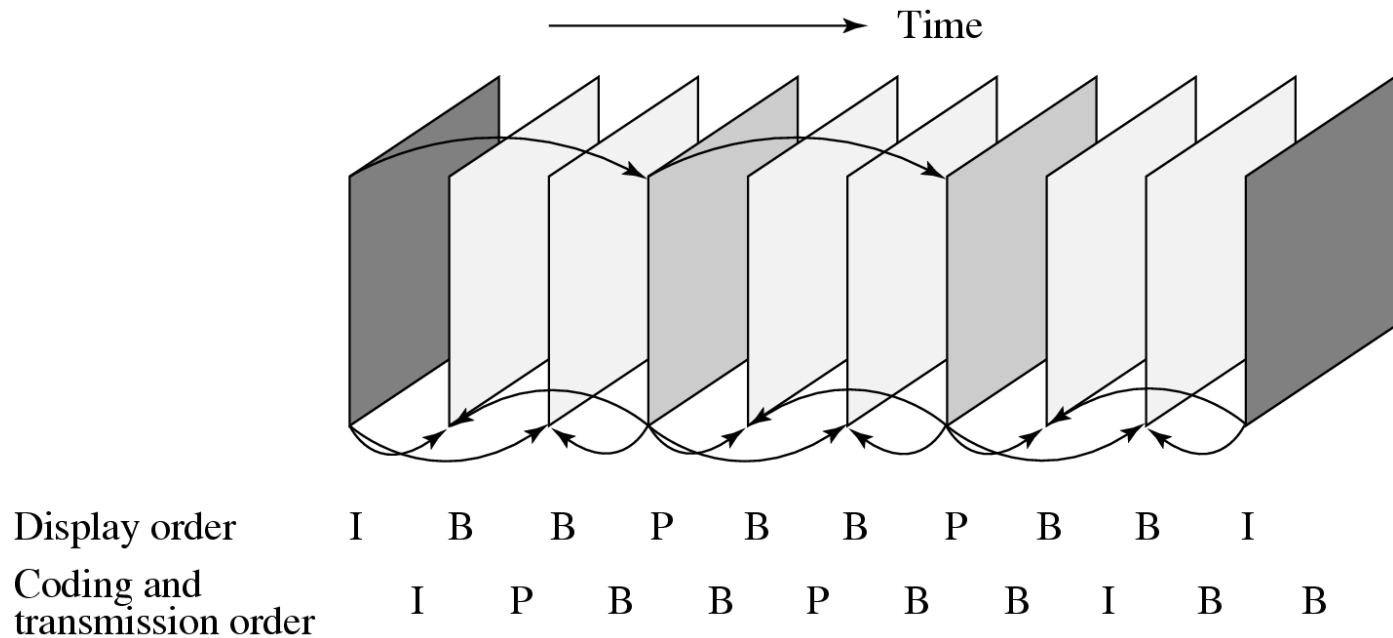


جستجوی دو طرفه

نمایش رمزگذاری B-FRAME بر پایه MC



ترتیب فریم‌ها در MPEG



- میانگین دو پیشگویی به عنوان پیشگویی فریم‌های B در نظر گرفته می‌شود.
- اگر دو پیشگویی بسیار باهم متفاوت باشد تنها یکی مورد استفاده قرار می‌گیرد.

چندی سازی و بردار حرکت

- چندی سازی در MPEG و H.261 متفاوت است.
- در H.261 تنها یک عدد در نظر گرفته می شود. اما در MPEG فریم‌های I با ماتریسی مشابه (و نه مساوی) با ماتریس JPEG چندی می شوند.
- در MPEG بردارهای حرکتی دقتی کمتر از پیکسل دارند که برای استفاده از آنها یک درون یابی نیاز است.
- در MPEG دامنه تغییرات بردار حرکتی بیشتر است.
- $H.261 \rightarrow \pm 15 \text{ pixels}$
- $MPEG1 \rightarrow [-512 \sim 512]$ کمتر از پیکسل
- $MPEG1 \rightarrow [-1024 \sim 1024]$ دقت پیکسل

جدول کارایی MPEG به صورت نوعی

Type	Size	Compression
I	18kb	7:1
P	6kb	20:1
B	2.5kb	50:1
Average	4.8kb	27:1