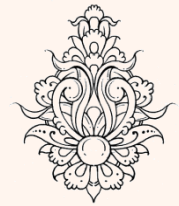
فشرده‌سازی اطلاعات
۴۰-۱۰-۷۰۲-۰۱
بخش هشتم
قسمت اول

مقدمه‌ای بر MPEG2

دانشگاه شهید بهشتی
پژوهشکده‌ی فضای مجازی
بهار ۱۳۹۸
احمد محمودی ازناوه

فهرست مطالب

- پیش‌گفتار
- ادراک حرکت در سیستم بینایی بشر
- فشرده‌سازی ویدئو
  - انواع نمونه‌برداری
    - پیش‌رونده
    - درهم تنیده
- افزونگی زمانی و مکانی
- انواع فریم
- تخمین بردار حرکت



پیش‌گفتار

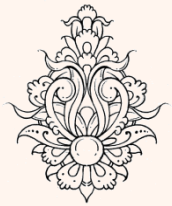
- محیط زندگی ما، پر از دوربین‌های مداربسته است.
- افزون بر این، روزانه تعداد قابل توجهی ویدئو که با دوربین‌های موبایل و تبلت گرفته می‌شوند، در سایت‌های مختلف بارگذاری می‌شوند.
- این دنباله‌های ویدئویی که با اهداف مختلف اخذ شده‌اند، حاوی اطلاعات مفیدی است که باید ذخیره، ارسال و پردازش شوند.
- از این رو، «پردازش سیگنال‌های ویدئویی» اهمیت روزافزونی یافته است.

برخی کاربردها

- ویدئوکنفرانس
- ارسال ویدئو از طریق شبکه‌های ارتباطی
  - فشرده‌سازی، یکی از مهم‌ترین نیازمندی‌های این حوزه است.
- افزایش کیفیت ویدئو
- شناسایی هویت با استفاده از حرکات
- تجزیه و تحلیل حرکات
- تشخیص اعمال انسان(Human Action Recognition)
- نهان‌نگاری و واترمارکینگ ویدئو
- تشخیص جعل در دنباله‌های ویدئو

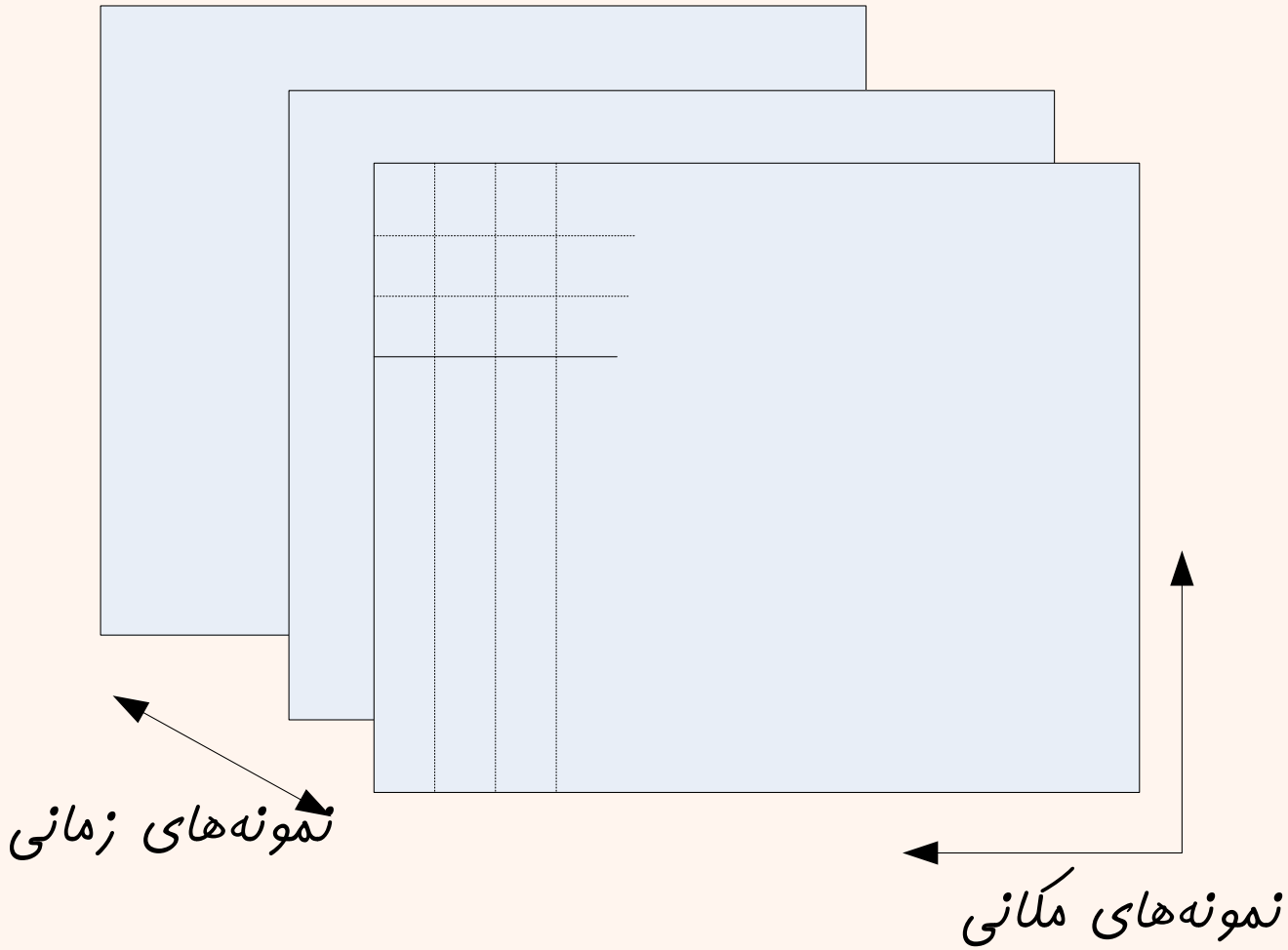
معرفی

- یک صحنه‌ی طبیعی ویدئویی متشکل از تعداد بی‌شماری شئ همراه با خصوصیات ظاهری مانند عمق، طرح و قالب، بافت و روشنایی می‌باشد.
- خصوصیات یک صحنه ویدئویی:
  - شامل خواص زمانی
  - خصیصه‌های وابسته به مکان



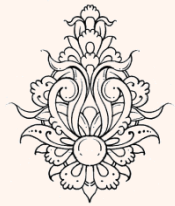
شمای کلی

نمونه‌های مکانی و زمانی یک جریان ویدئویی



نمونه‌برداری زمانی

- نمونه‌برداری زمانی همان ذخیره نمودن تصاویر لحظه‌ای در فواصل زمانی مشخص می‌باشد.
  - نمایش این تصاویر به صورت پی‌درپی حرکت اشیا را نشان می‌دهد.
  - نرخ نمونه‌برداری بسیار اهمیت دارد. هر چه نرخ مذکور بالاتر باشد، حرکت نرم‌تر و به حالت طبیعی نزدیک‌تر است.
  - برای سیستم تلویزیونی ۲۵ تا ۳۰ فریم در ثانیه می‌باشد.



Motion perception

- هر تصویر حدود 25ms بر روی شبکیه ثابت می‌ماند.

Persistence of vision

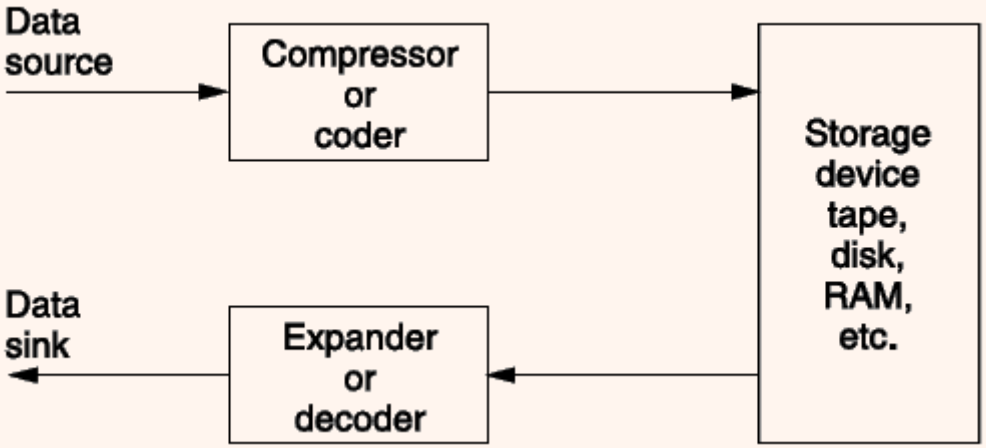
- دو شی که دارای تغییر سریع هستند، در سیستم بینایی بشر به صورت یک حرکت پیوسته ادراک می‌شود.

تصاویر از ویکی‌پدیا گرفته شده است

اطلاعات حجیم

- برخلاف تصویر حجم اطلاعات در رسانه‌ی ویدئو بسیار بالاست.
- از این رو معمولاً کدینگ بدون‌اتلاف در این مقوله کاربرد ندارد.

Too Much Information !!



انواع نمونه‌برداری

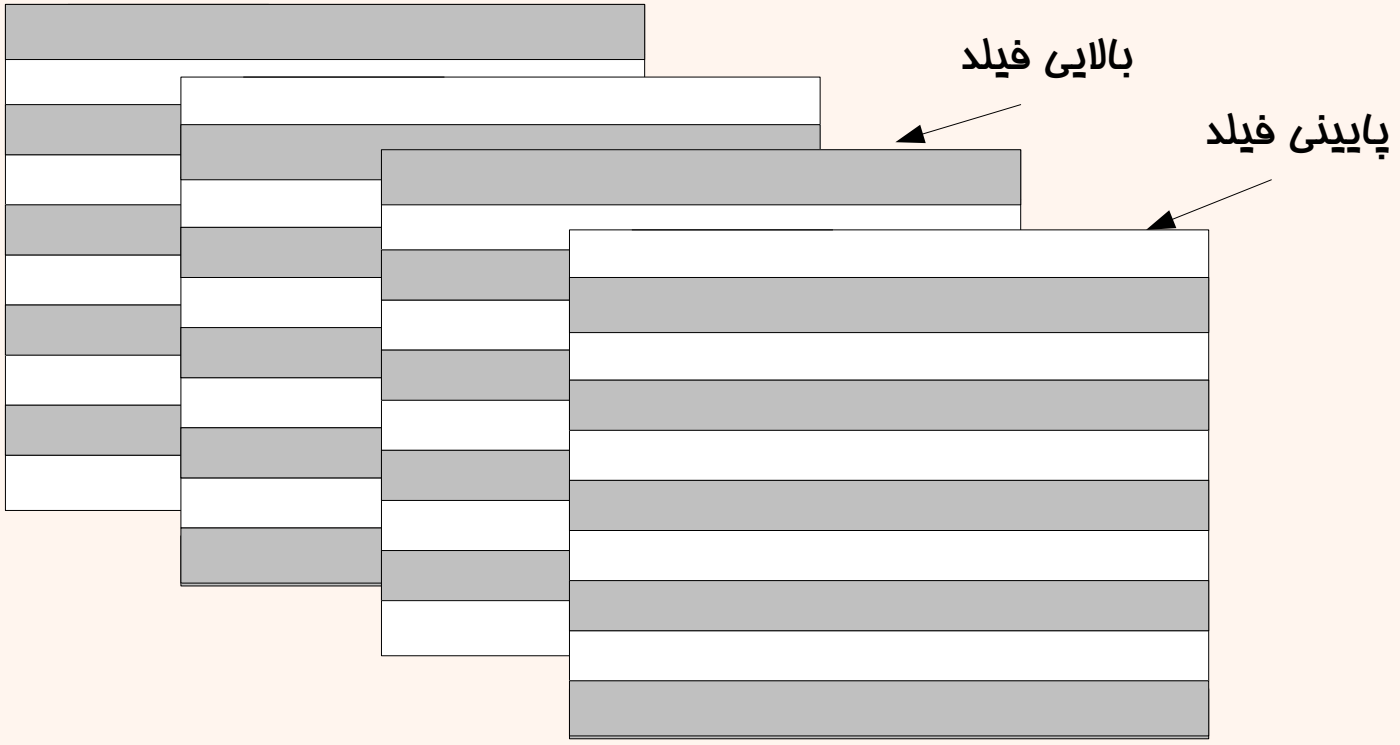
Progressive Sampling

- پیش‌رونده
  - هنگامی‌که یک سیگنال ویدئویی به صورت سری کاملی از فریم‌های کامل نمونه‌برداری گردد، نمونه‌برداری را پیش‌رونده می‌نامند.

Interlaced Sampling

- نمونه‌برداری درهم‌تنیده
  - در هر فریم نیمی از داده‌ها(یک ناحیه) نمونه‌برداری صورت می‌پذیرد. یک ناحیه متشکل از تعداد زوج و یا فرد از خطوط داده‌های هر فریم است. در روش نمونه‌برداری درهم‌تنیده، در هر بازه زمانی، نیمی از داده‌ها (ناحیه بالایی و یا ناحیه پایینی) جهت نمونه‌برداری در نظر گرفته می‌شوند.
    - مزیت این روش به گونه‌ای است که در این شیوه با نرخ بیت یکسان نسبت به روش پیش‌رونده، قابلیت ارسال داده دوبرابر از هر ناحیه وجود دارد.

جریان داده‌ای ویدئویی همراه با درهم‌تنیدگی



فضا رنگ

$$Y = K_r R + K_g G + K_b B$$

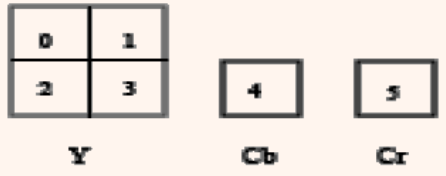
$$C_b = B - Y$$
$$C_r = R - Y$$
$$C_g = G - Y$$

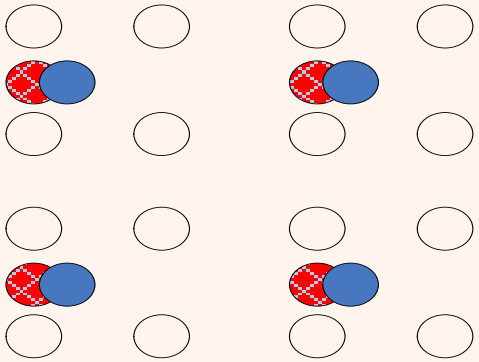
- $Y'C_bC_r$

معادلات تفاضلی رنگ یا مولفه‌های تفاوت کرومینانس

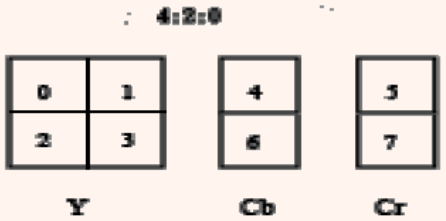
- از آن‌جا که $C_b + C_r + C_g$ مقداری متناسب با Y دارد، ارسال دو مقدار از سه مؤلفه کروما کفایت می‌نماید و مؤلفه سوم با داشتن دومؤلفه دیگر محاسبه می‌گردد.
- در سیستم YC_bC_r مولفه روشنایی و دو مؤلفه رنگی ارسال می‌گردد. مزیت عمده سیستم مذکور قابلیت ارسال دو مؤلفه کروما با دقت کم‌تر از مؤلفه روشنایی است زیرا سیستم بینایی انسان نسبت به روشنایی حساسیت بیشتری دارد.

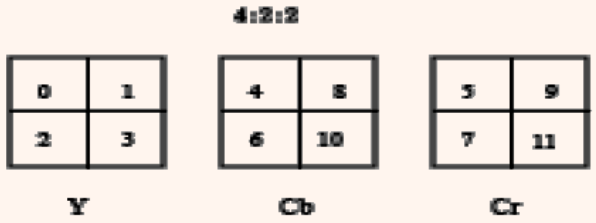
ساختار سیستم نمونه‌برداری



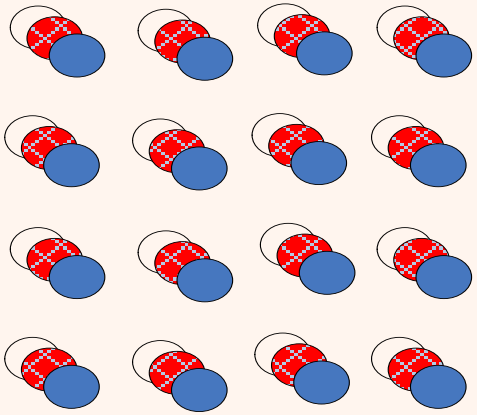


نمونه Y
نمونه Cr
نمونه Cb

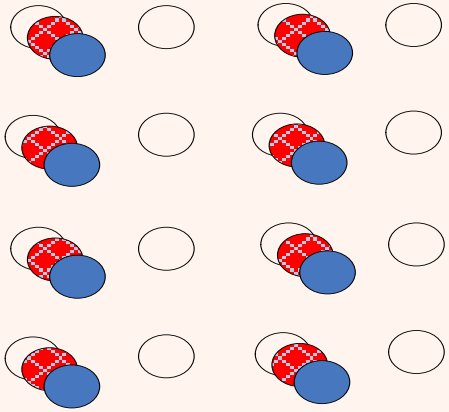




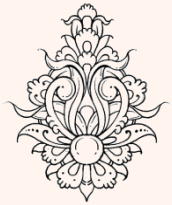
برداری نمونه 4:2:0



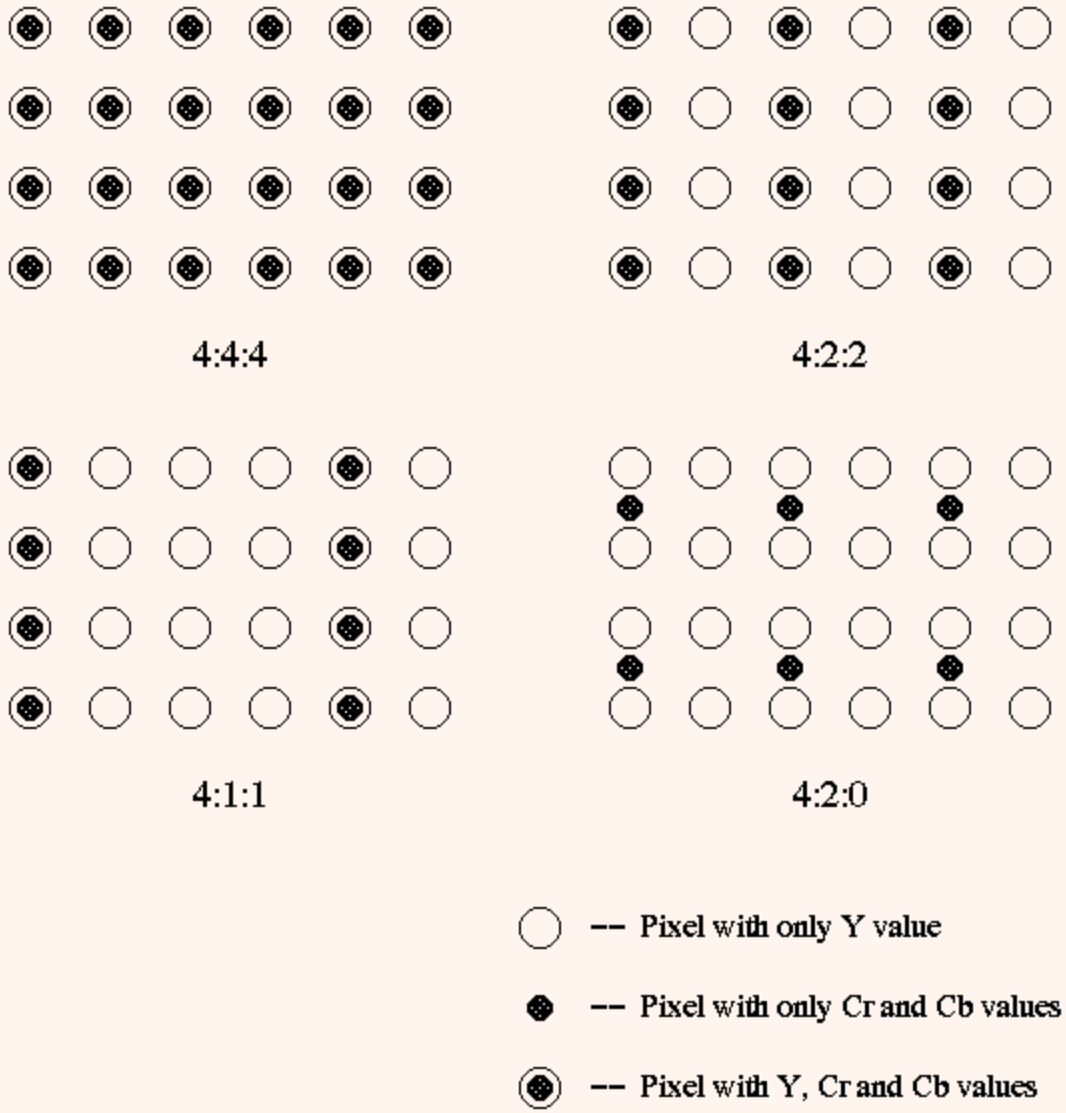
برداری نمونه 4:4:4



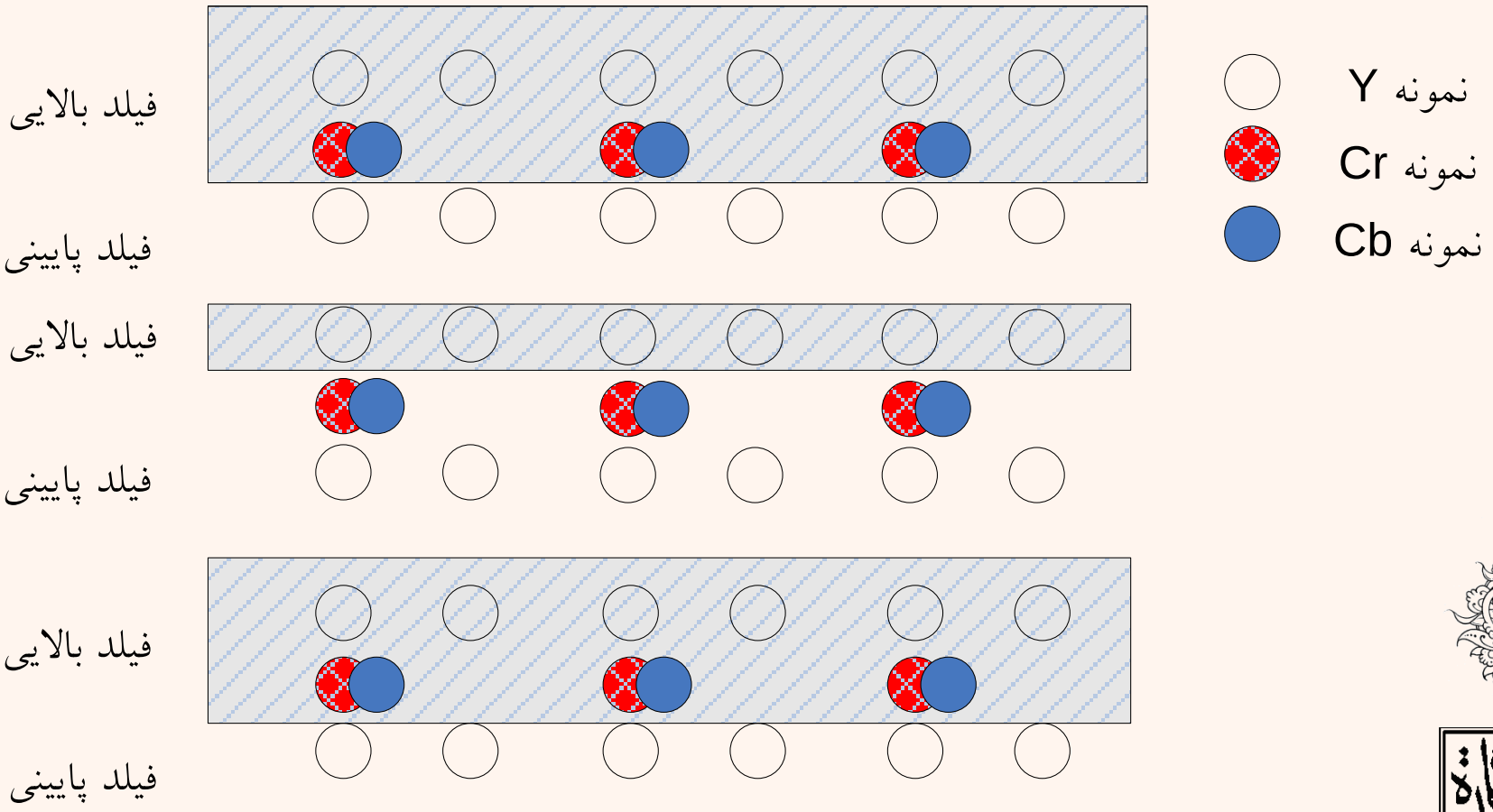
برداری نمونه 4:2:2



ساختار سیستم نمونه‌برداری (ادامه...)

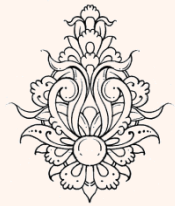


نمونه‌برداری(۴:۲:۰) در ویدئوهای در هم‌تنیده



ساختارهای ویدئویی

- ساختار CIF پایه‌ای برای بسیاری از قالب‌های فایل‌های ویدئویی می‌باشد. بسته به برنامه کاربردی مورد نظر و همچنین ظرفیت کانال مورد استفاده جهت ارسال فایل ویدئویی از ساختارهای متفاوت استفاده می‌شود.
- برای مثال 4CIF برای سیستم استاندارد تلویزیونی و DVD کاربرد دارد و در همین راستا برای کاربردهایی با درجه‌ی دقت پایین‌تر چون ویدئوکنفرانس استاندارد CIF و یا QCIF به کار می‌رود.



4CIF
704x57
CIF
352x28
QCIF
176x14
SQCIF
128x96

(Source Input Format)

is practically identical to CIF, but taken from MPEG-1 rather than ITU standards

Format	Video Resolution
SQCIF	128 × 96
QCIF	176 × 144
SCIF	256 x 192
SIF(525)	352 x 240
CIF/SIF(625)	352 × 288
4SIF(525)	704 x 480
4CIF/4SIF(625)	704 × 576
16CIF	1408 × 1152
DCIF	528 × 384

Sub Quarter CIF, sometimes subQCIF

Quarter CIF

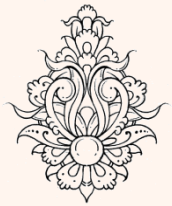
("NTSC") based systems

("PAL") based systems

video conferencing systems

Double CIF

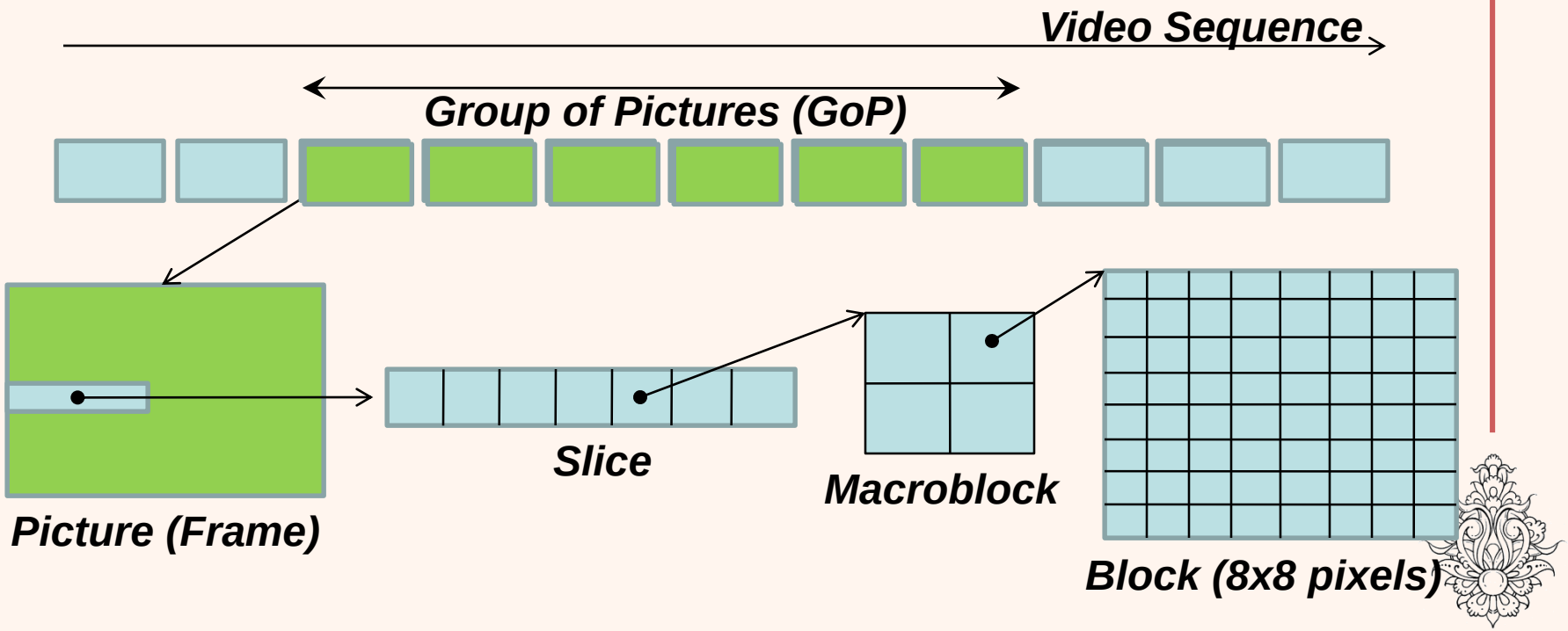
common CCTV equipment

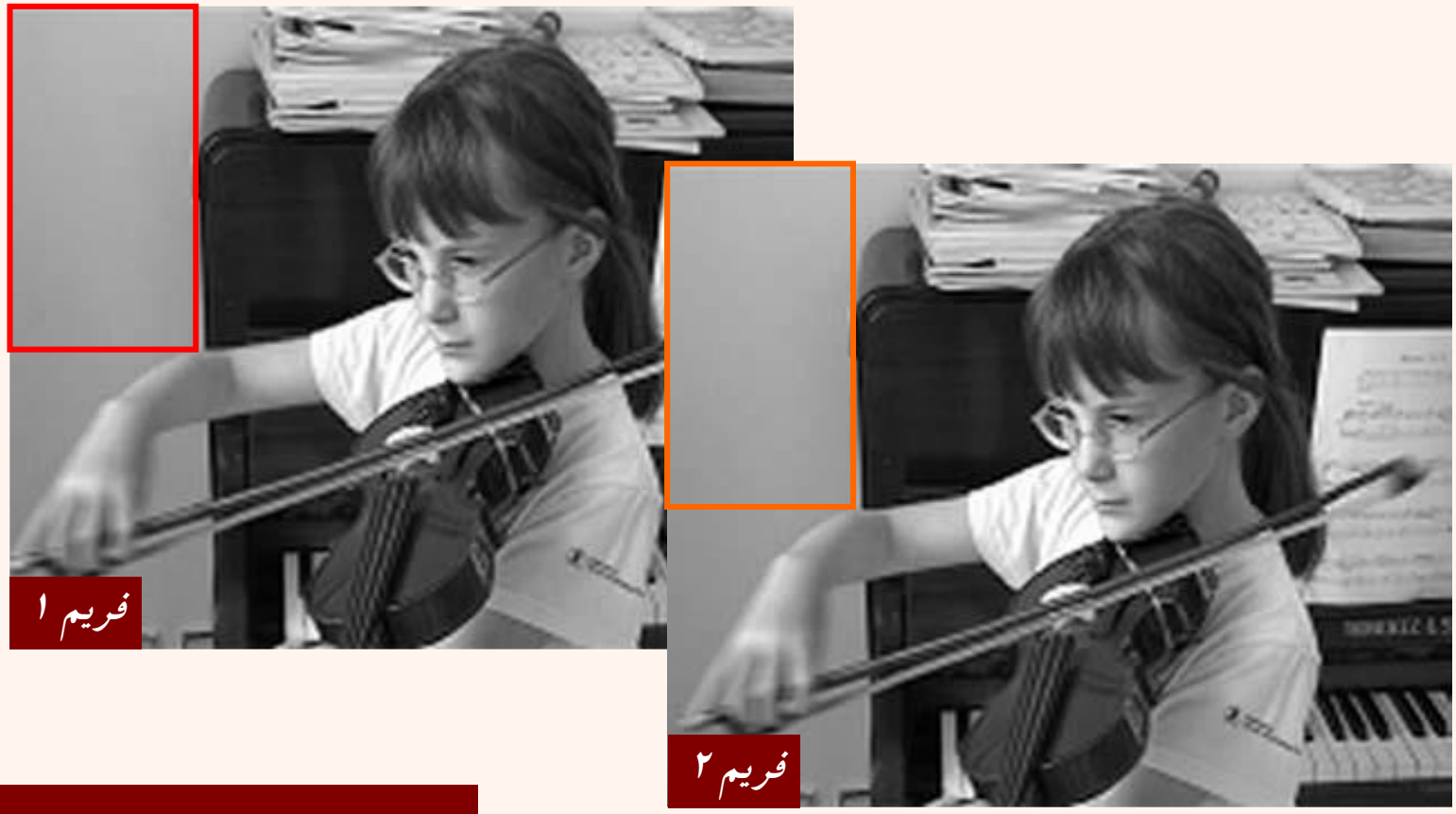


مفهوم کدینگ در ویدئو

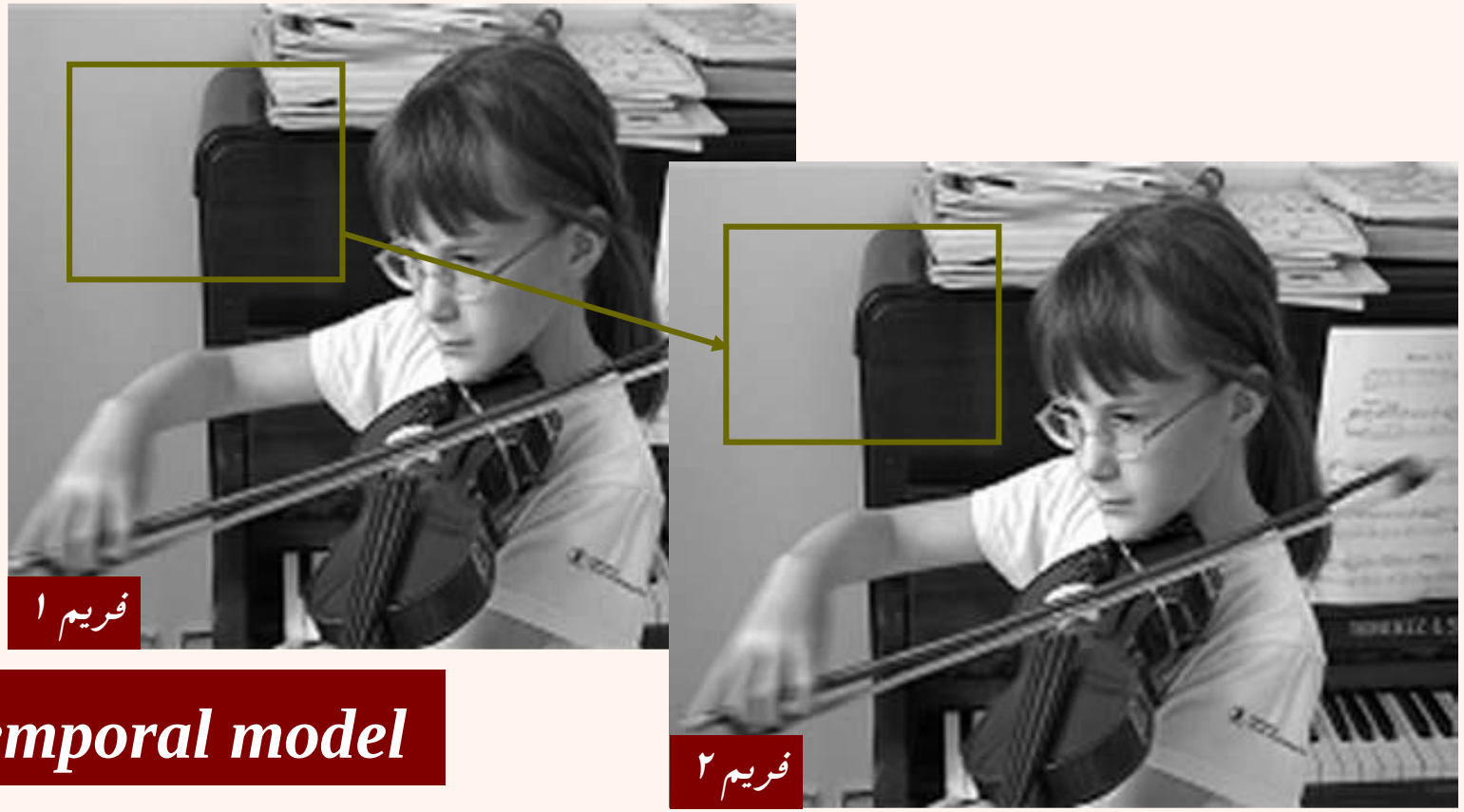
- فشرده‌سازی فرایندی را مطرح می‌نماید که در آن داده به تعداد بیت‌های کم‌ترفشرده می‌گردد.
- بسیاری از روش‌های کدینگ ویدئو از افزونگی زمانی و مکانی داده‌ها در یک ساختار ویدئویی سود می‌جویند. در حوزه زمان در بیشتر مواقع نرخ این شباهت میان فریم‌های متوالی بسیار بالاست.
- فریم‌های متوالی که به لحاظ زمانی‌های متوالی که به لحاظ زمانی مرتب شده‌اند، به ویژه در مواردی که نرخ نمونه‌برداری بالا باشد، بسیار شبیه به هم می‌باشند.
- در حوزه مکان، هنگامی که یک فریم در نظر گرفته می‌شود، این شباهت در پیکسل‌های مجاور هم به چشم می‌خورد.

ساختار اصلی


Video Sequence
Group of Pictures (GoP)
Picture (Frame)
Slice
Macroblock
Block (8x8 pixels)

مفهوم فشرده‌سازی در ویدئو

Spatial model

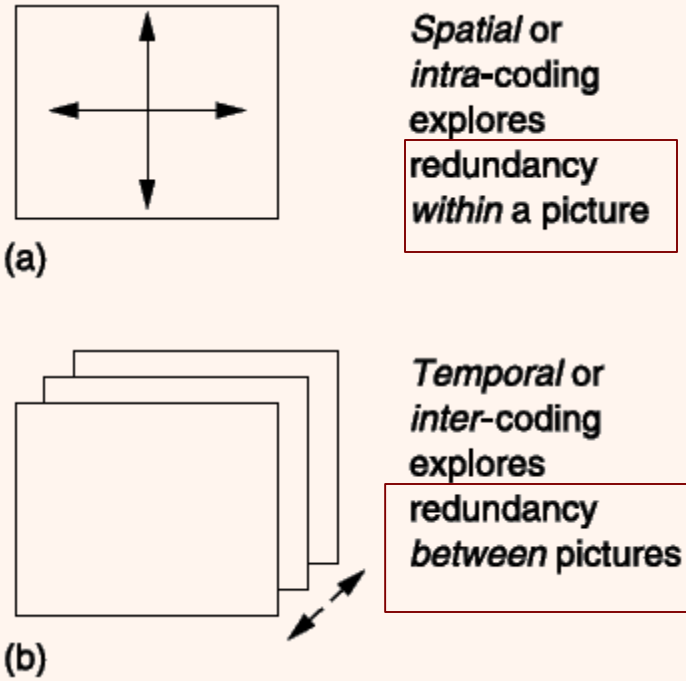
مفهوم فشرده‌سازی در ویدئو

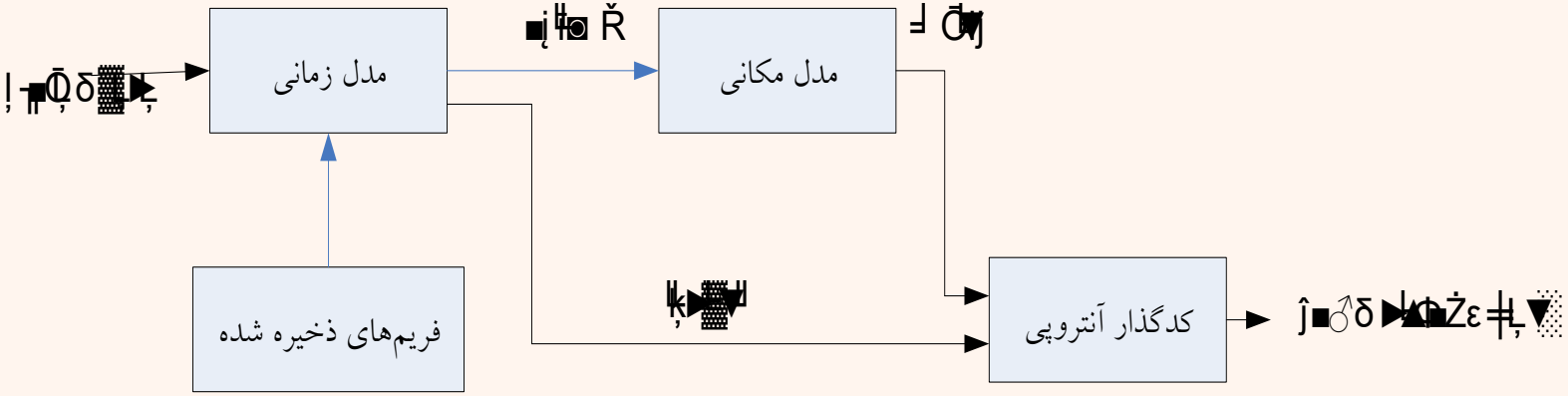


Temporal model

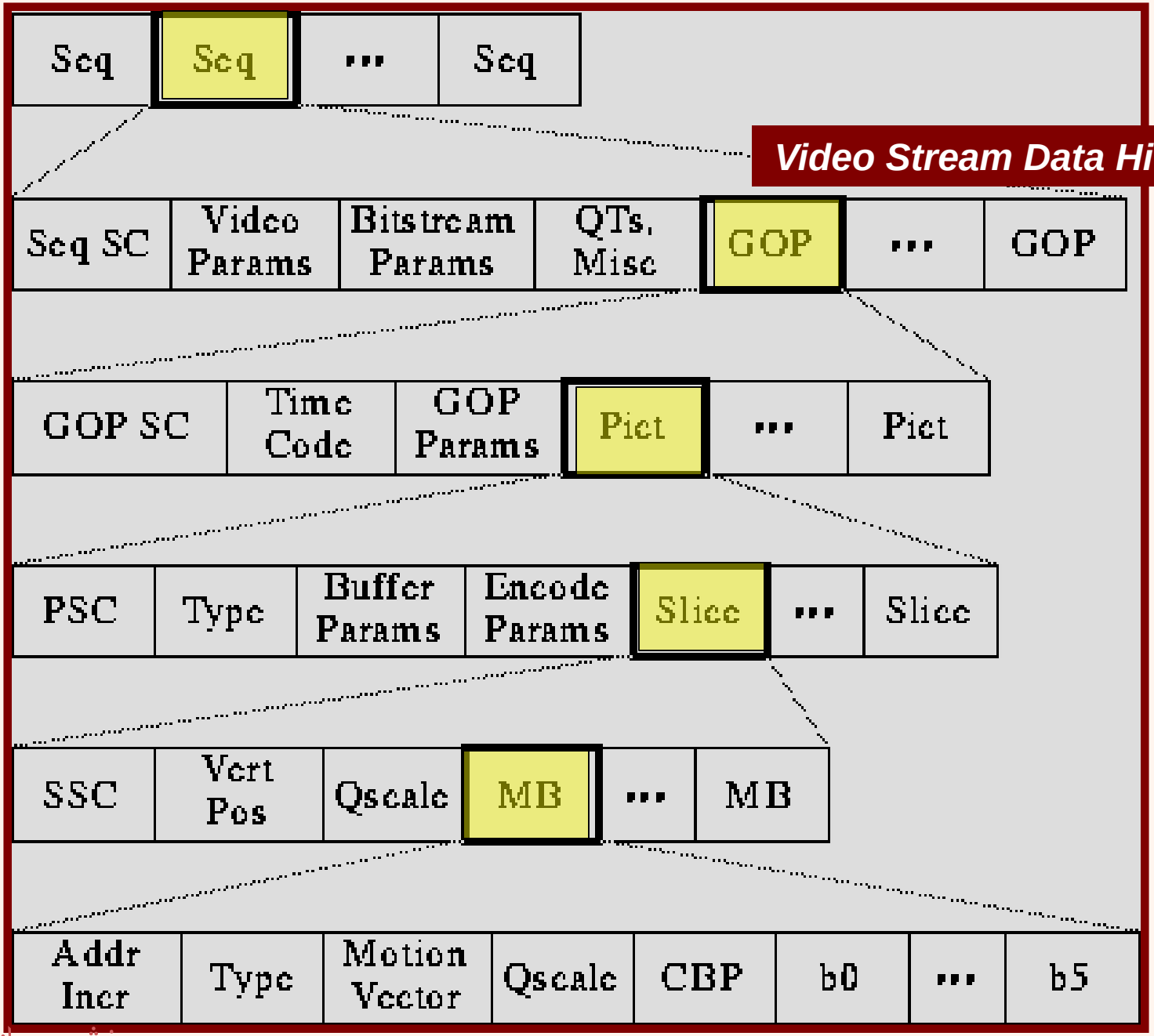
کدگذار ویدئو

- مدل زمانی
- مدل مکانی
- کدگذار آنتروپی





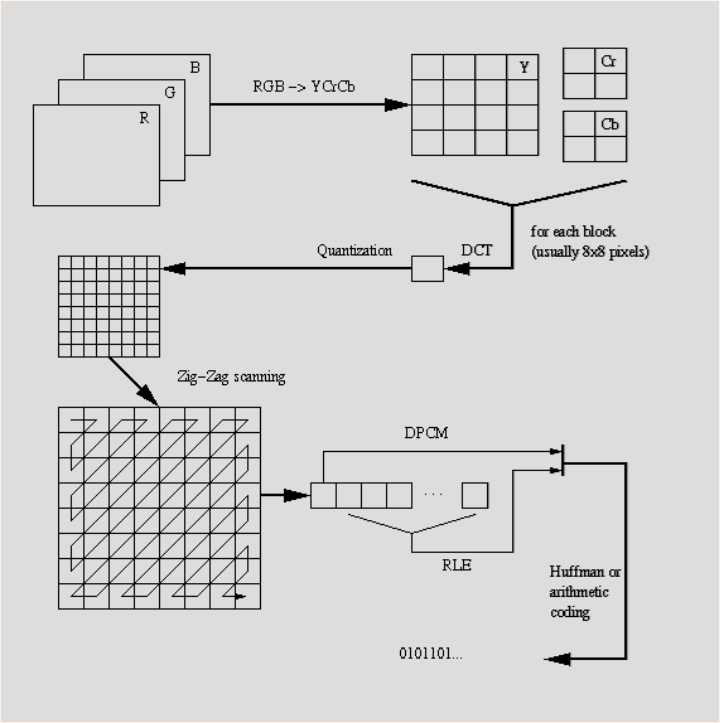
Video Stream Data Hierarchy



فشرده‌سازی

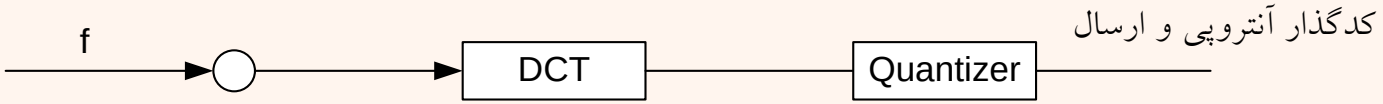
ارسال اطلاعات یک فریم

Intra Frame



- یک فریم

$$f(n_1, n_2)$$



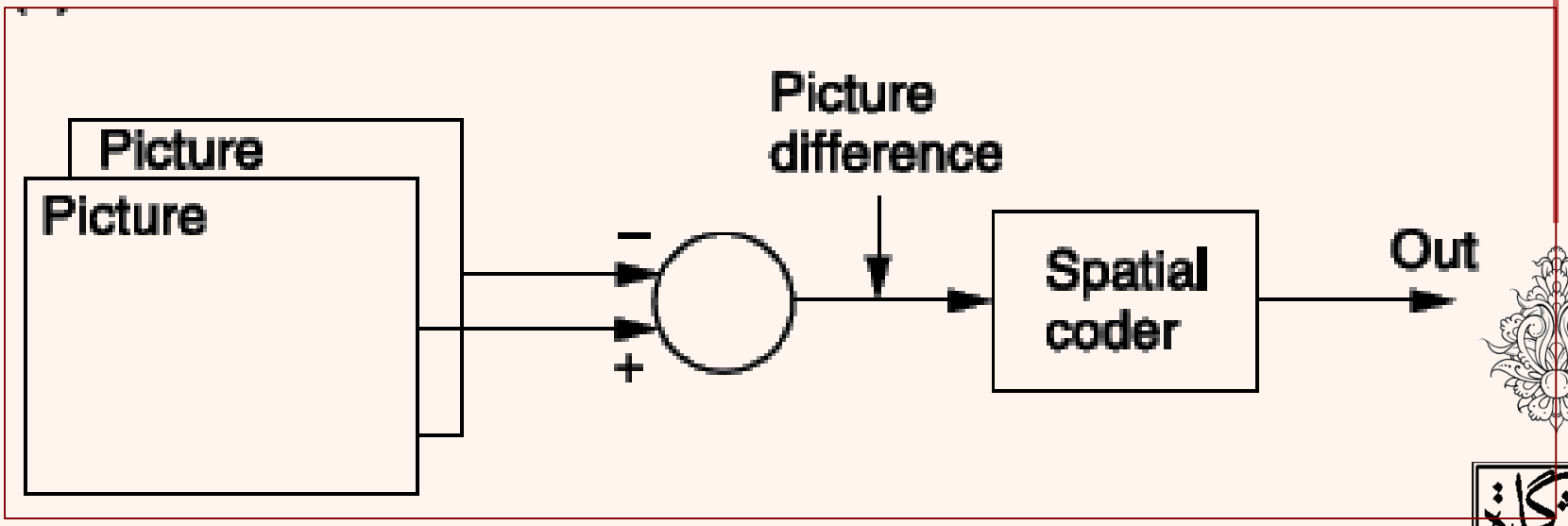
- چند فریم متوالی

Inter Frame

$$f(n_1, n_2, n_3)$$

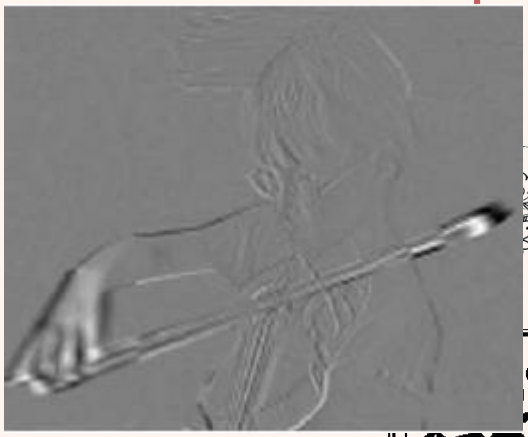
ارسال اطلاعات دو فریم متوالی

- ساده ترین روش این است که دو فریم متوالی را از یکدیگر کم کرده اختلاف را برای کد کردن ارسال کنیم.



Frame 1

Frame 2



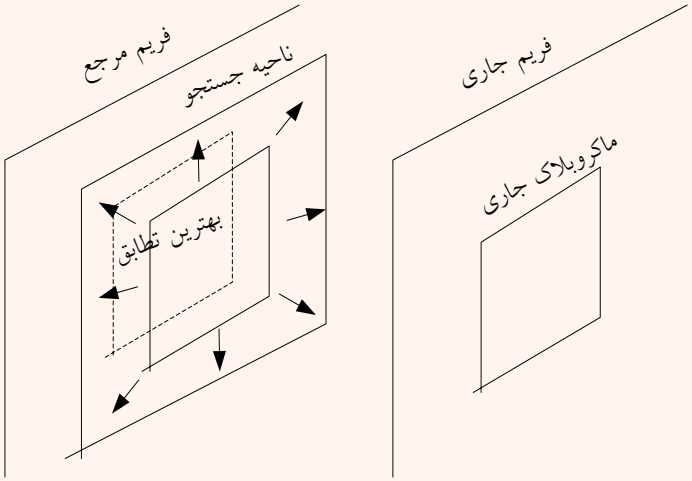
Difference

تخمین حرکت (برپایه‌ی بلوک)

Motion Estimation

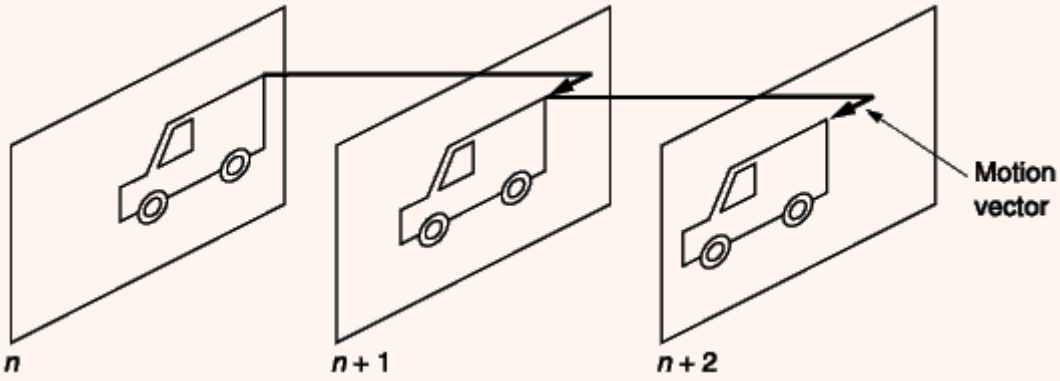
- «تخمین حرکت» یک ماکروبلاک یافتن ناحیه‌ای ۱۶×۱۶ در فریم مرجع است که بیشترین میزان شباهت را با فریم جاری دارا باشد.
- فریم مرجع، فریم قبلی کدگذاری شده است که می‌تواند به لحاظ ترتیب نمایش یکی از فریم های پسین یا پیشین را شامل گردد. به بهترین گزینه انتخابی «بهترین انطباق» گفته می شود.

Best Match

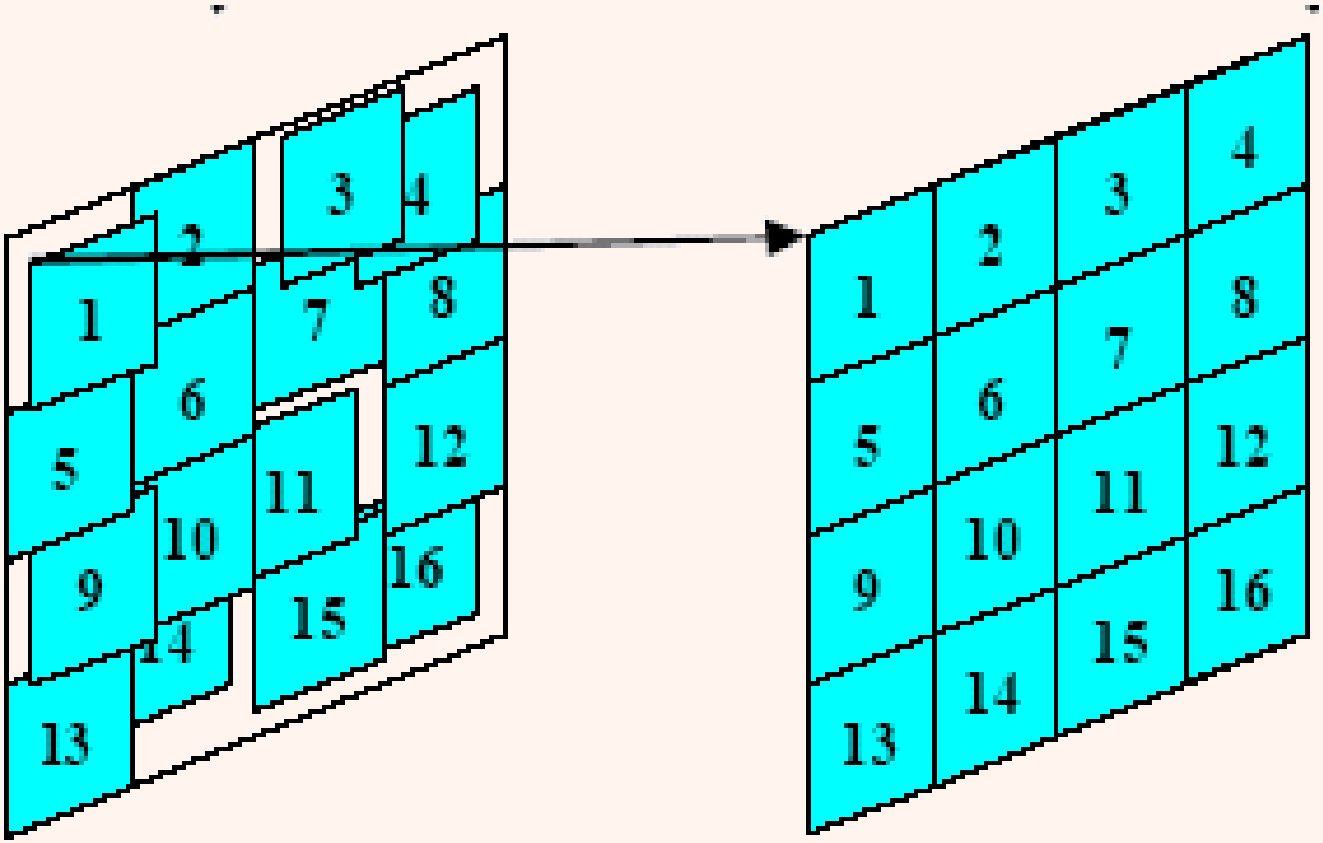


جبران حرکت (ادامه...)

- روشی است که به واسطه‌ی آن میزان افزونگی داده در فریم‌های متوالی را به کمینه‌ی آن برسانیم.
- به واسطه‌ی این فرآیند داده‌هایی از نوع بردار حرکت و خطای ناشی از تخمین خواهیم داشت که می‌باید ارسال گردد.



جبران حرکت (ادامه...)



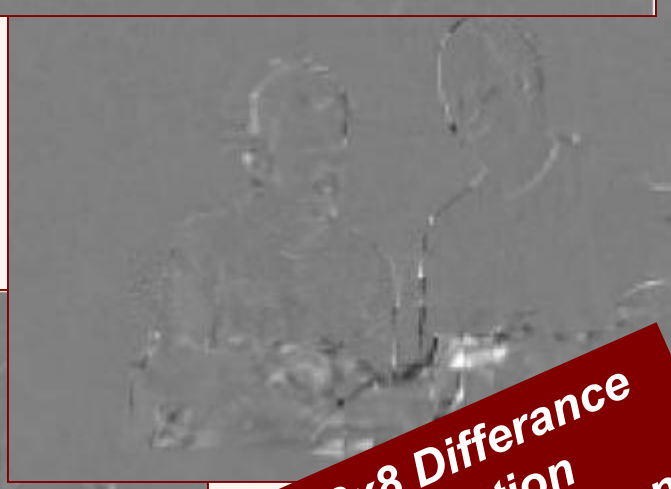
http://www.cs.ucf.edu/courses/cap5015/lectures.html

اندازه‌ی بلوک‌ها

Frame 1

Frame 2

4x4 Differance Motion Compensatio



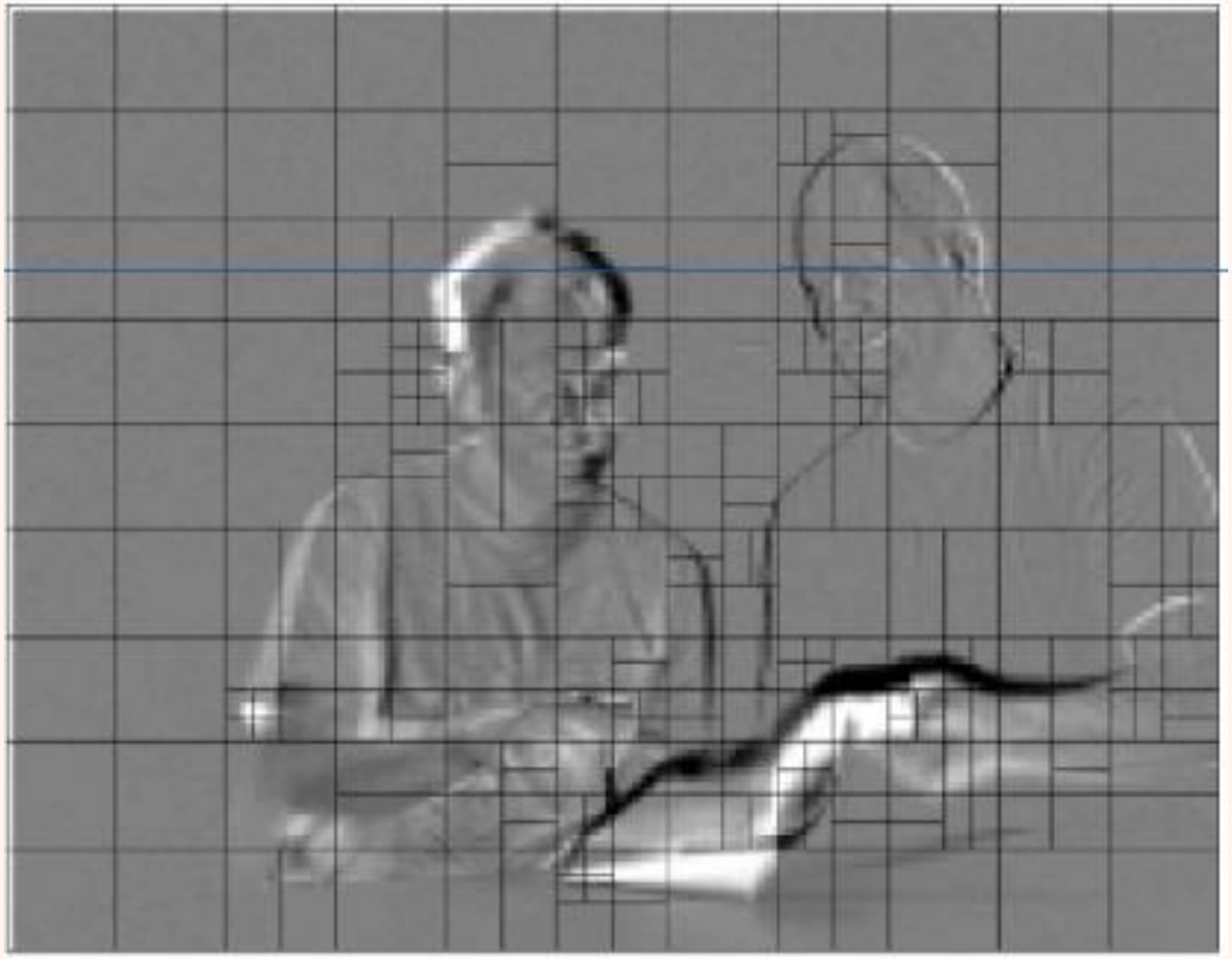
8x8 Differance Motion Compensation

Differance No Motion Compensation

16x16 Differance Motion Compensation

فشرده‌سازی

اندازه‌ی بلوک‌ها



H.264

Picture Types

- Intra

کم‌ترین میزان فشرده‌سازی

همان فریم نوع I است که به صورت مستقل کد می‌گردد.

- Predicted

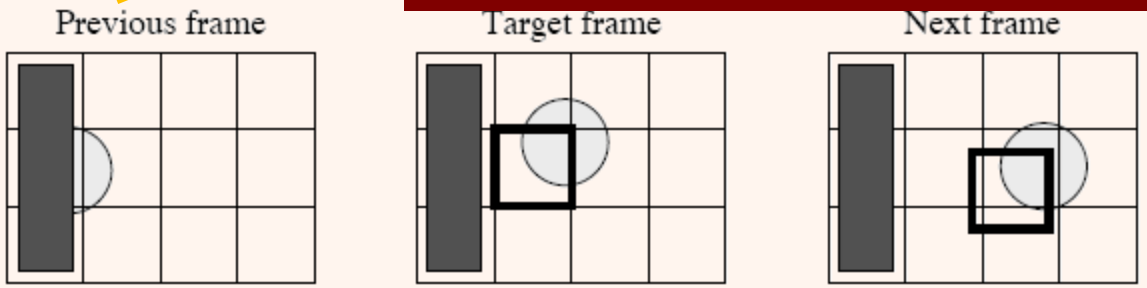
فشرده‌سازی میانه

همان فریم نوع P است که با توجه به نزدیک‌ترین فریم نوع I یا P کد می‌شود.
فریم‌های نوع P می‌توانند خطا را در کل جریان داده در یک GOP انتقال دهند.

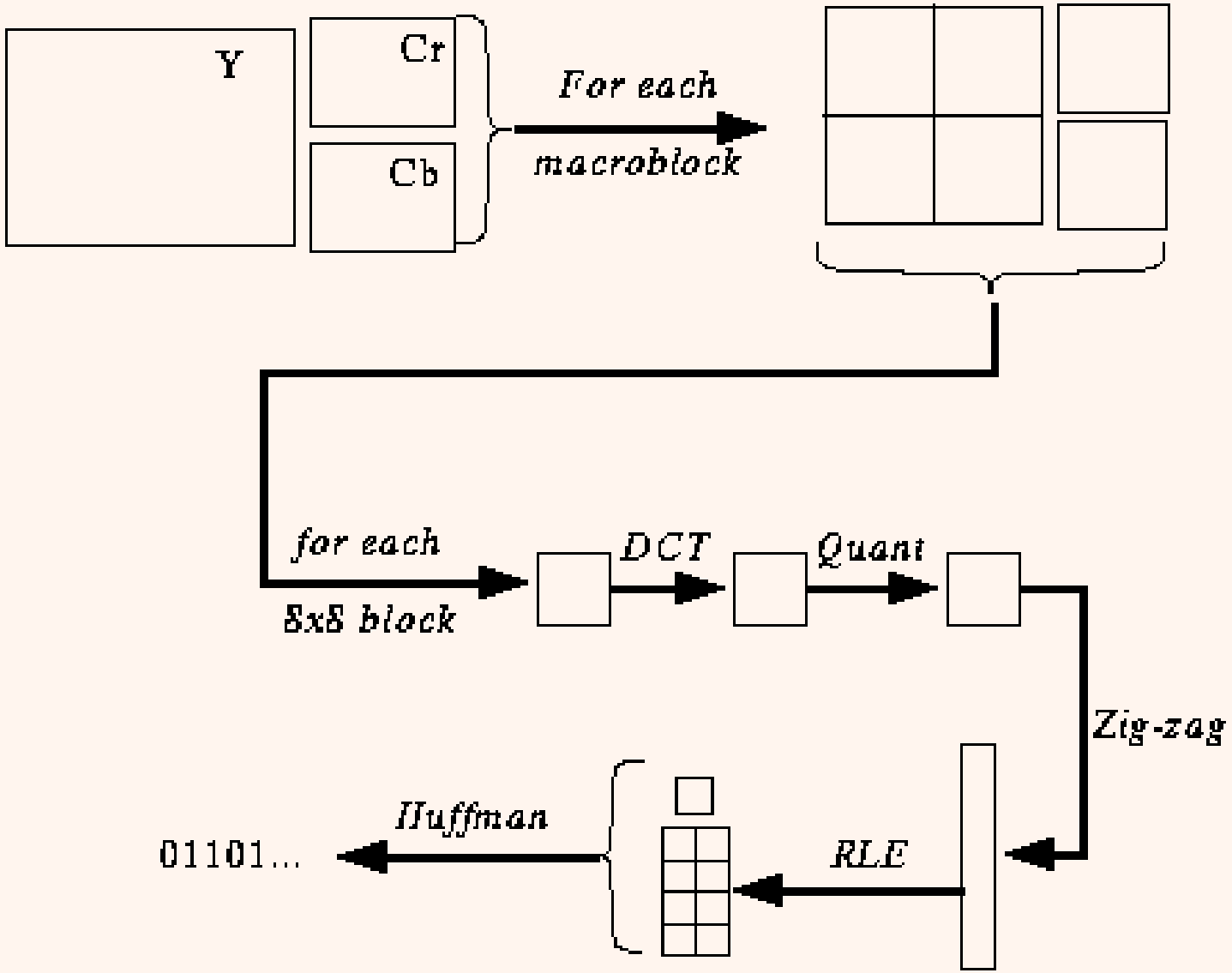
- Bidirectional

بیشترین میزان فشرده‌سازی

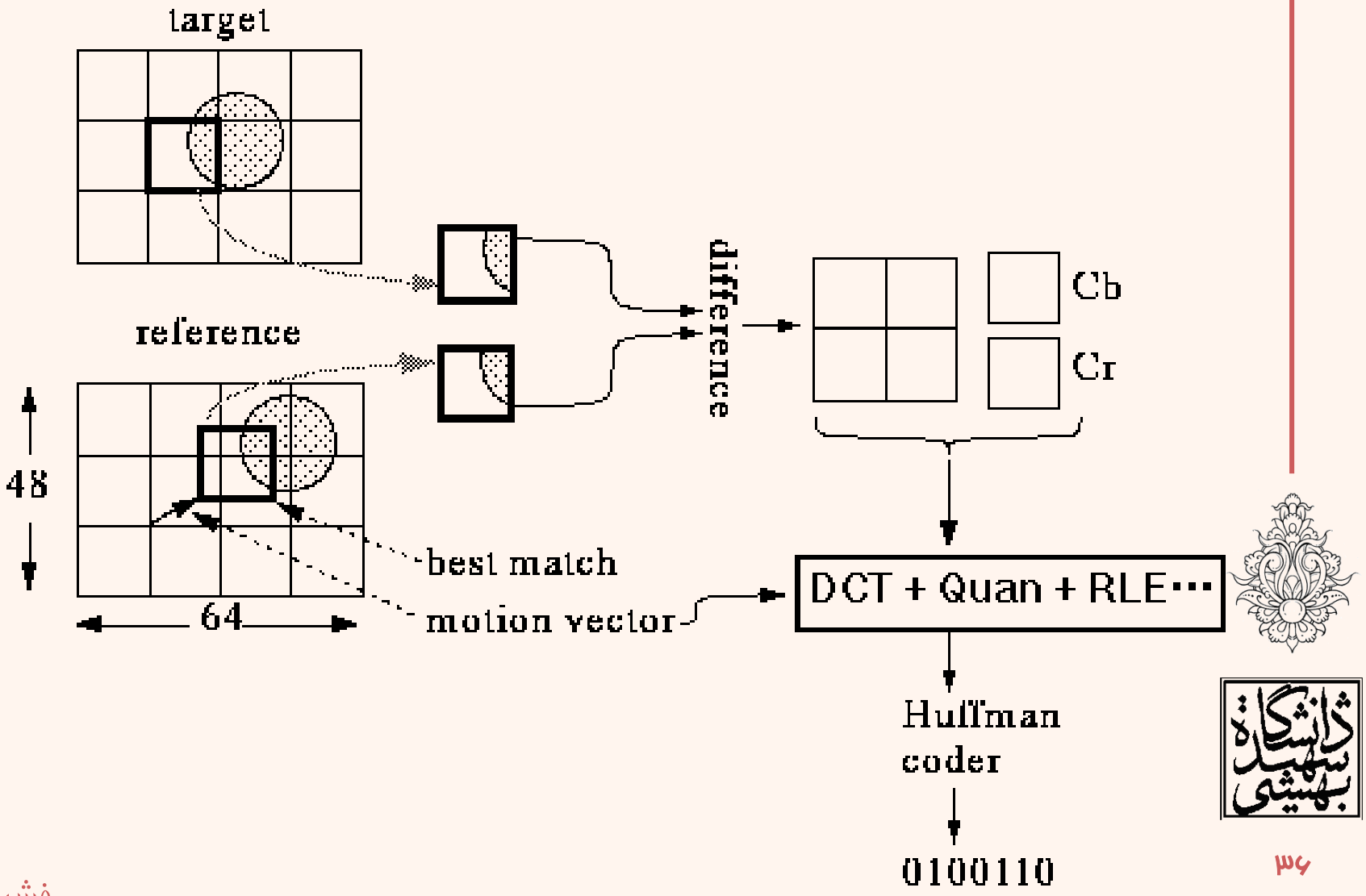
همان فریم نوع B است که با توجه به دو فریم دیگر کد می‌شود. (قبلی و بعدی)
چون هیچ گاه مبدا واقع نمی‌شود خطا را منتقل نمی کند.
به دلیل استفاده از میانگین دو فریم عموماً اثر نویز را هم کاهش می‌دهد.



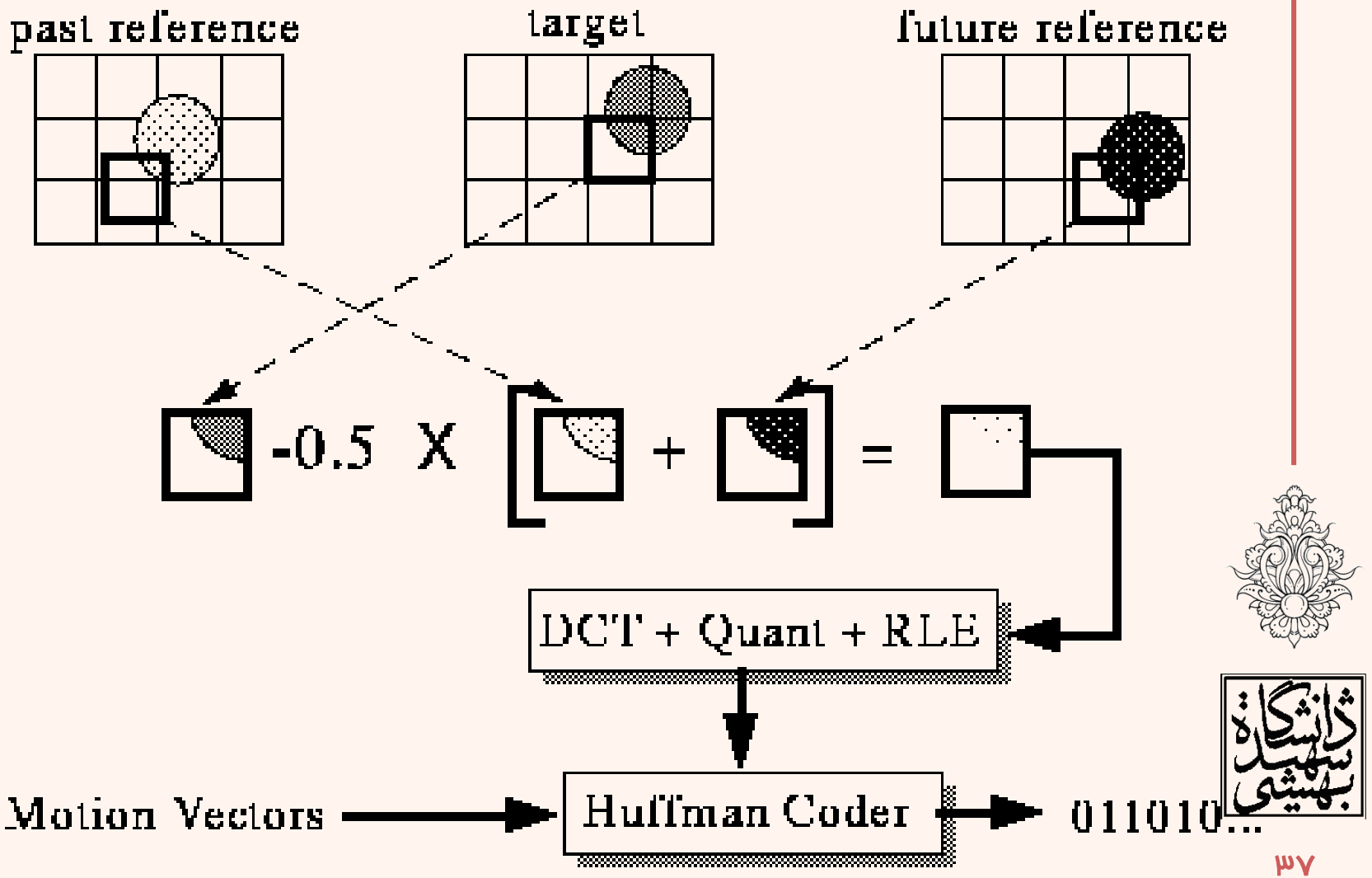
ساختار کلی ارسال فریم ا



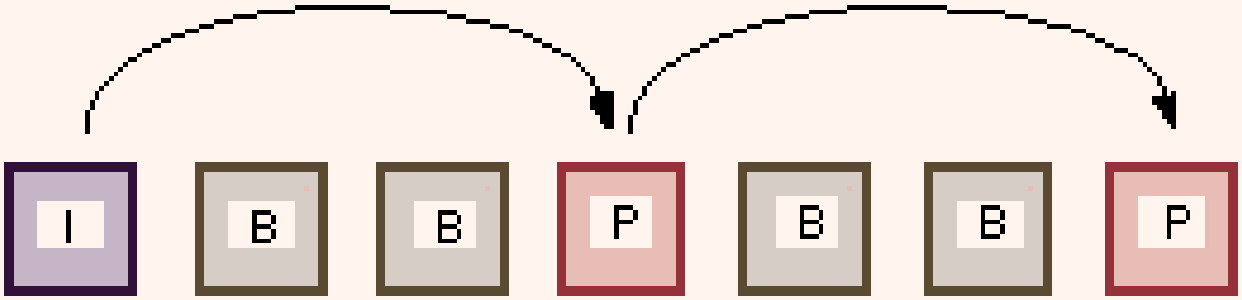
ساختار کلی ارسال فریم P



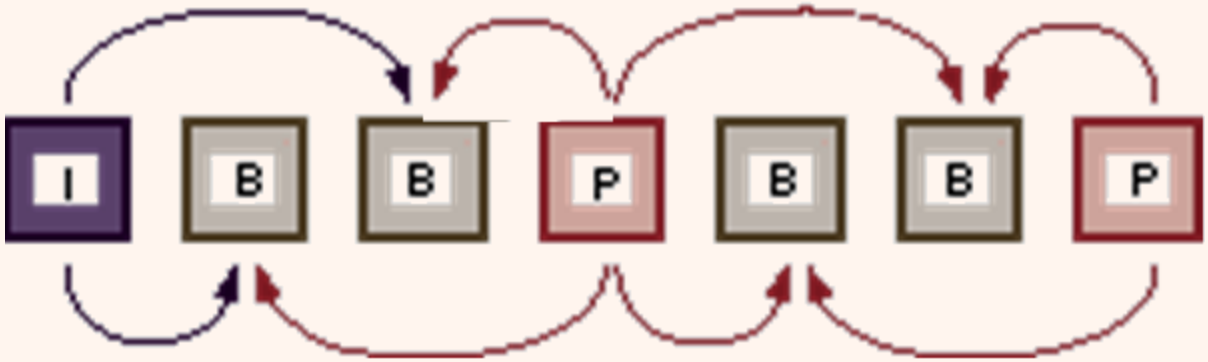
ساختار کلی ارسال B



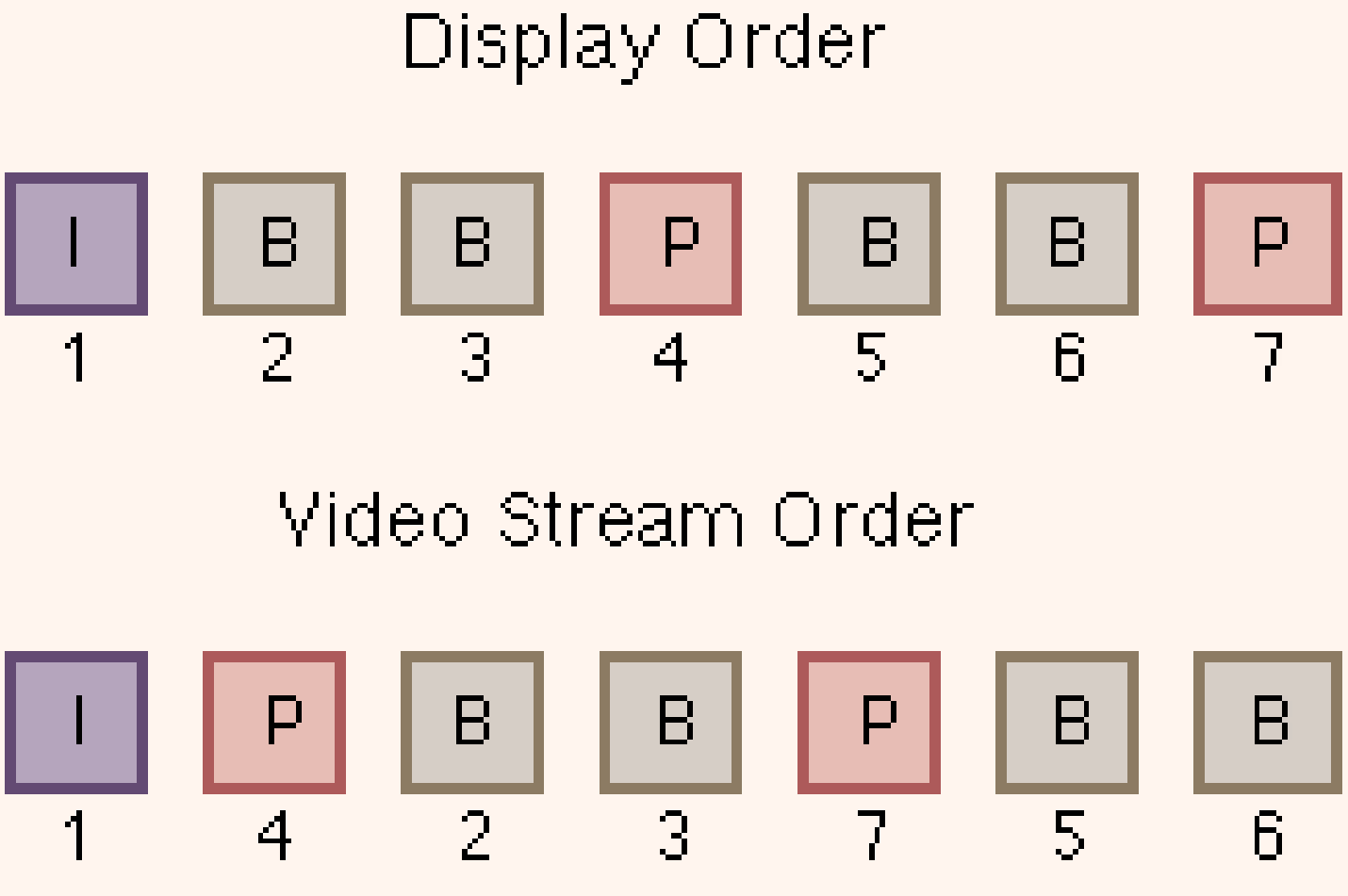
Forward Prediction



Bidirectional Prediction

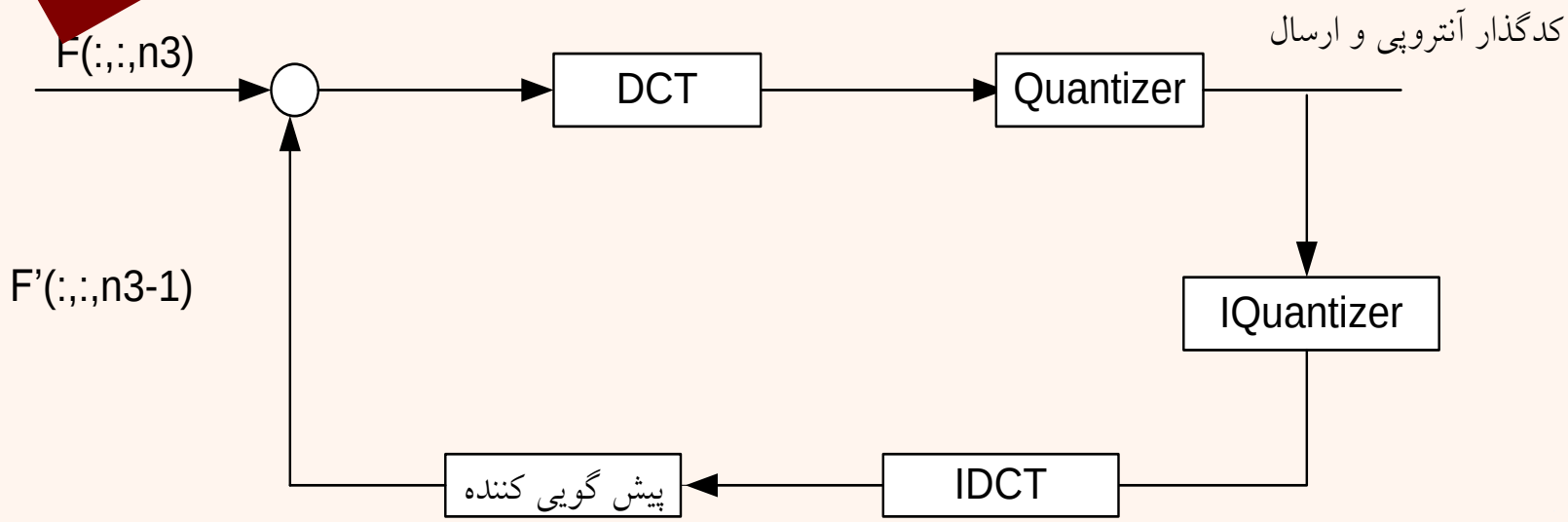


نحوه‌ی ذخیره‌کردن فریم‌ها در کدگذار



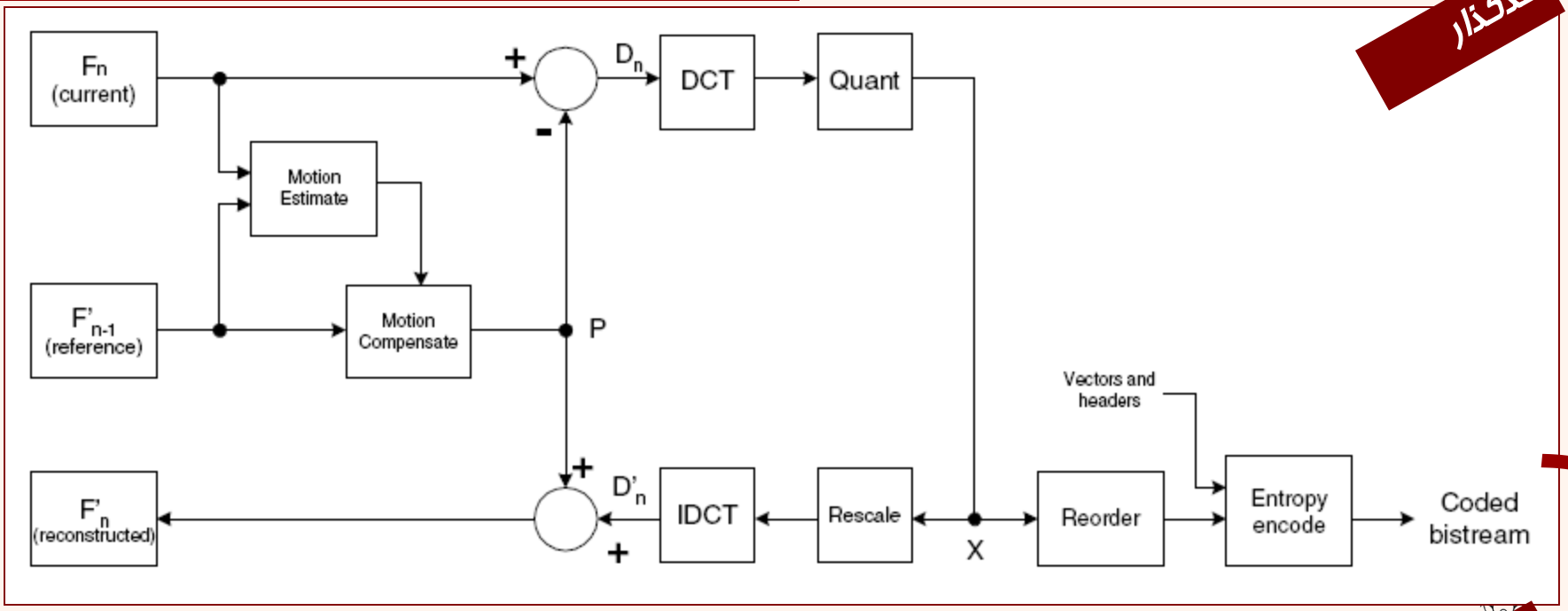
شمای کلی کدگذار

کدگذار

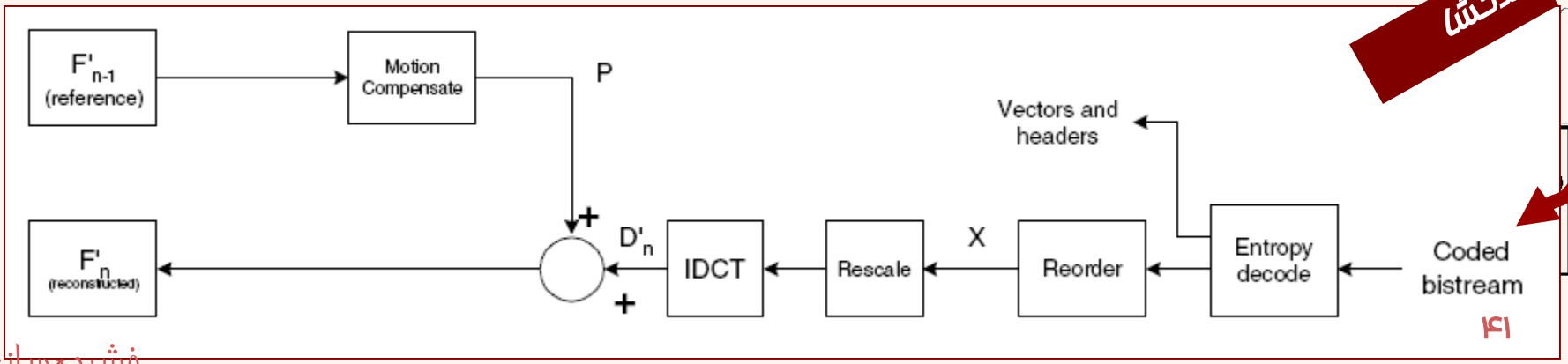


compatible with H.261, H.263, MPEG-1, MPEG-2, MPEG-4 Visual and H.264

کدگذار



کدگشا



فشرده‌سازی

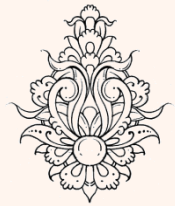
کیفیت زمانی و مکانی

کیفیت مکانی

کیفیت زمانی

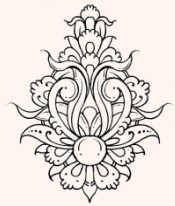
کیفیت در رسانه‌ی ویدئو

- بررسی کیفیت به صورت شکستن ویدئو به سری داده‌های دوبعدی
  - میانگین‌گیری از فریم‌ها
    - PSNR
    - SSIM
- بررسی مستقل رسانه‌ی ویدئو
  - VQM



کیفیت در رسانه‌ی ویدئو(ادامه...)

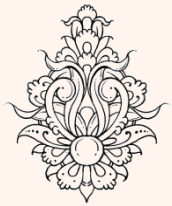
- در صورت بررسی هر فریم ویدئو به صورت یک سیگنال دو بعدی و سپس اعمال میانگین‌گیری
  - خصوصیاتی که مختص سیگنال سه بعدی است ممکن است از دست برود.
  - اثر فرآیندهایی همچون سوسو زدن در این چنین پارامتر کیفیتی لحاظ نخواهد شد.
- بهتر است برای محاسبه‌ی کیفیت در یک سیگنال ویدئویی سیگنال به صورت سه بعدی مورد ارزیابی قرار گیرد.



گام‌های نیم‌پیکسلی

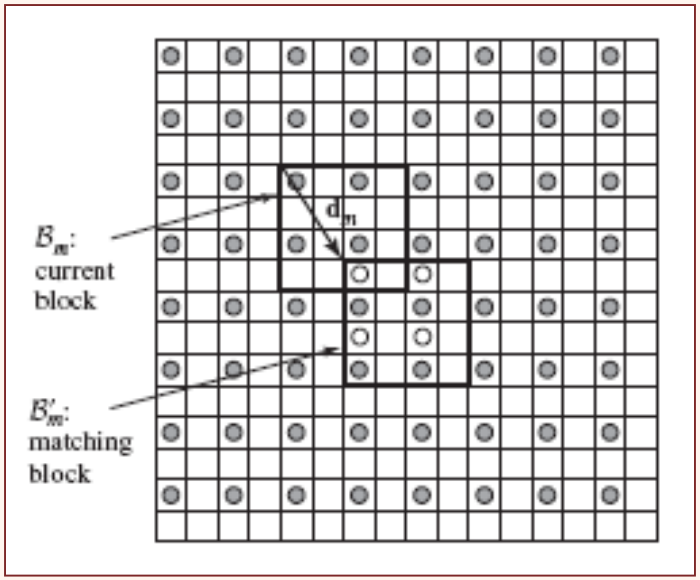
(Exhaustive Block Matching Algorithm)

- ممکن است اندازه‌ی بردارهای ضریب صحیحی از پیکسل‌ها نباشد!
  - اندازه‌ی گام را معمولا ۰.۵ درنظر می‌گیرند تا جستجوی بهتری صورت گیرد.
  - فریم مرجع تنها پیکسل‌های صحیح دارد.
  - راه‌حل:
    - قبل از جستجو میانگین‌گیری صورت گیرد.
    - معمولا Bilinear interpolation صورت می‌گیرد.



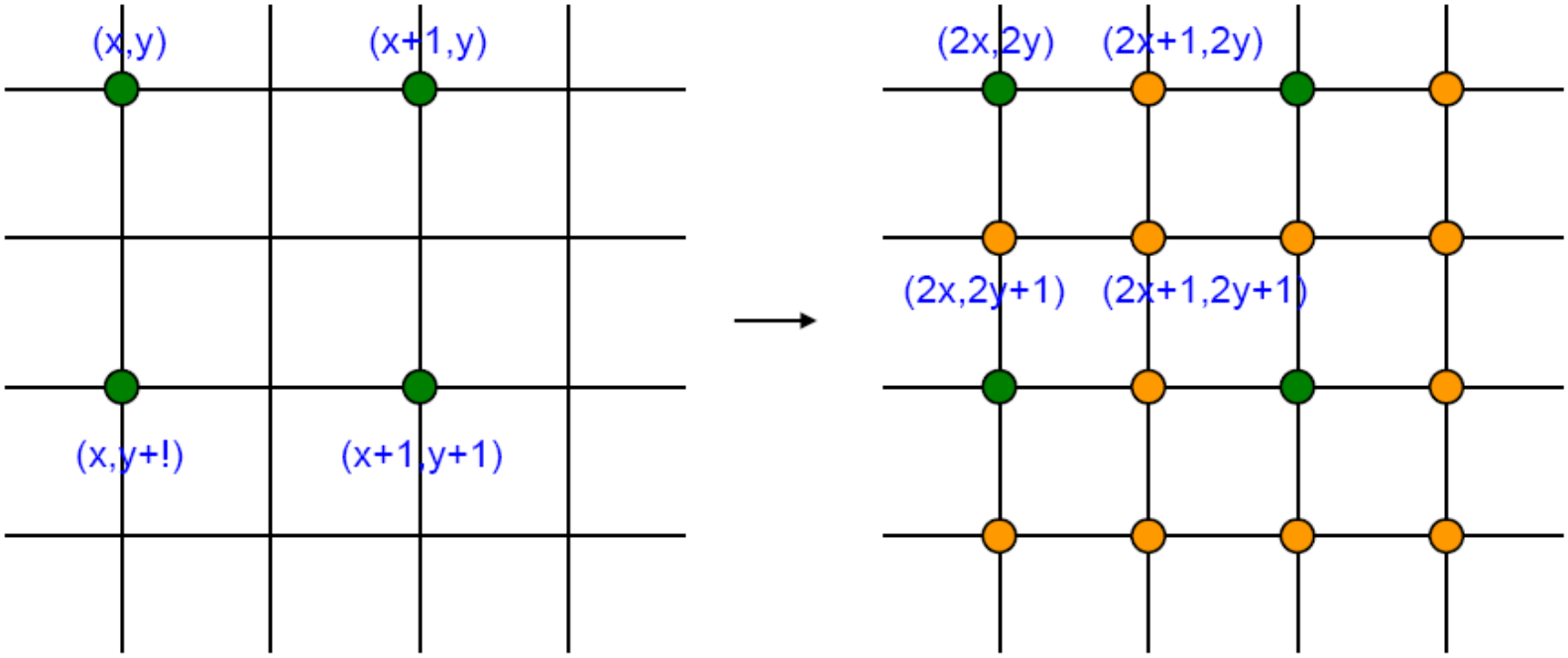
Half-pel EBMA

- **پیچیدگی:** چهار برابر نسخه‌ی صحیح به علاوه‌ی سایر فرآیندهای مورد نیاز
- **الگوریتم سریع:** ابتدا جستجو را برای گام‌های صحیح و سپس برای محدوده‌ی کوچک‌تری گام نیم‌پیکسلی جستجو صورت می‌گیرد.

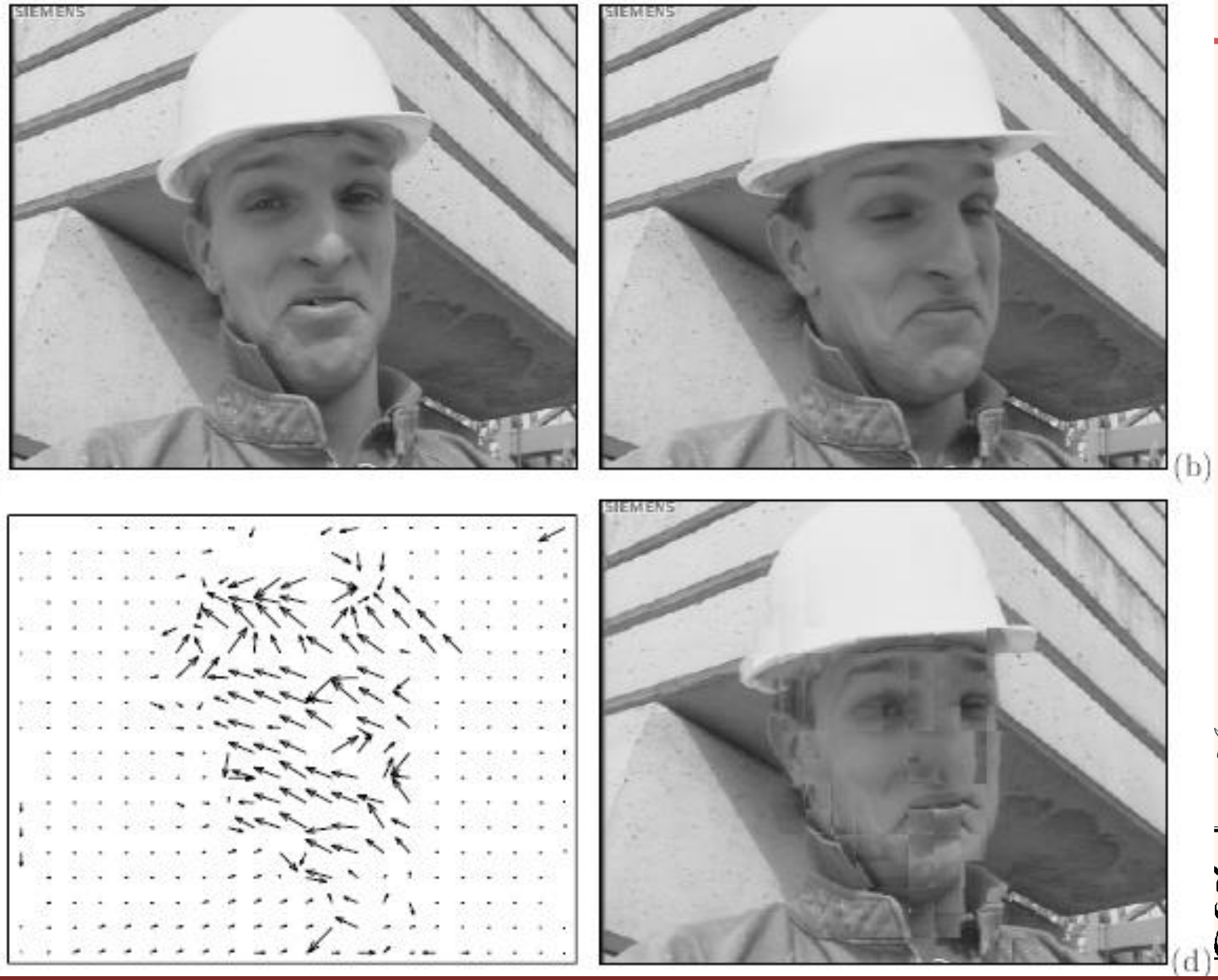


Bilinear Interpolation

درون‌یابی دوخطی



O[2x,2y]=I[x,y]
O[2x+1,2y]=(I[x,y]+I[x+1,y])/2
O[2x,2y+1]=(I[x,y]+I[x+1,y])/2
O[2x+1,2y+1]=(I[x,y]+I[x+1,y]+I[x,y+1]+I[x+1,y+1])/4



(a) the tracked frame; (b) the anchor frame; (c-d) motion field and predicted image the anchor frame (PSNR=29.86 dB) obtained by half-pel accuracy EB

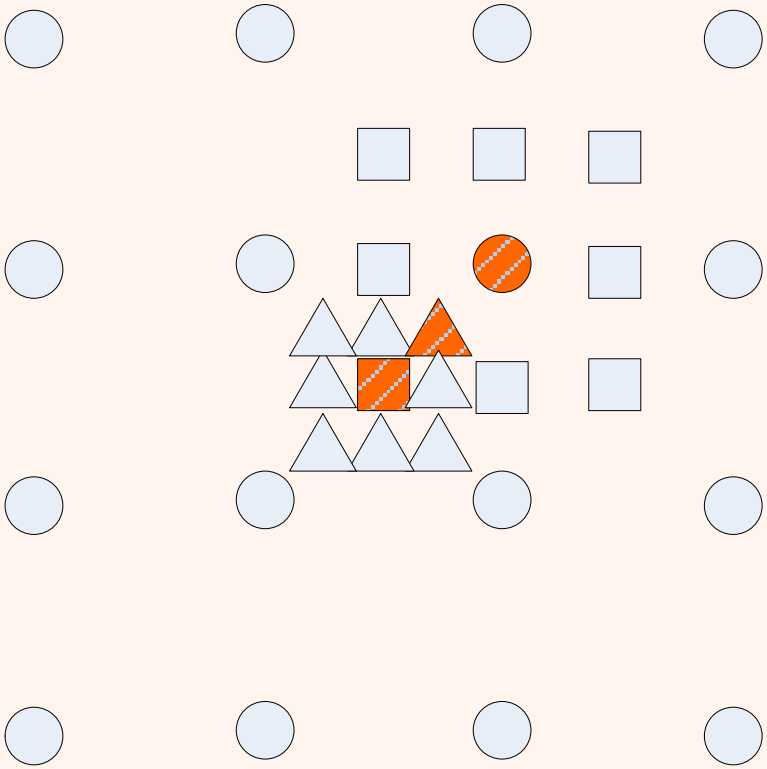
فشرده‌سازی

Quarter-Sample

- مفهوم جستجوی یک چهارم نمونه: در مرحله اول تخمین حرکت با استفاده از نمونه‌های صحیح که با علامت دایره نشان داده شده است صورت می‌پذیرد.
  - کدگذار همزمان در محدوده بهترین انتخاب، مکان‌های نیم‌نمونه را جستجو می‌کند(اشکال مربع)
  - پس از انتخاب بهترین نیم نمونه جستجو به جهت یافتن بهترین ربع‌نمونه صورت می‌پذیرد (اشکال مثلث).
  - در پایان بهترین انتخاب از میان انتخاب‌های صورت گرفته برگزیده خواهد شد.

Quarter-Sample

شمای فرآیند



صحیح جستجوی های مکان

صحیح تطابق بهترین

صحیح نیم جستجوی های مکان

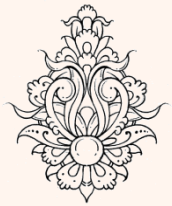
صحیح نیم تطابق بهترین

صحیح ربع جستجوی های مکان

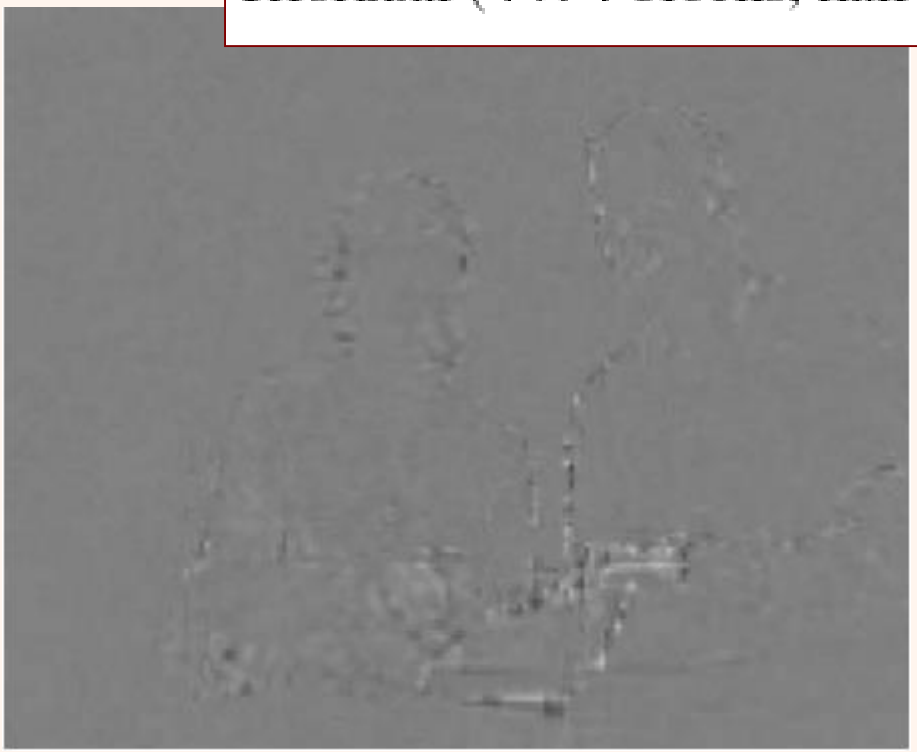
صحیح ربع تطابق بهترین

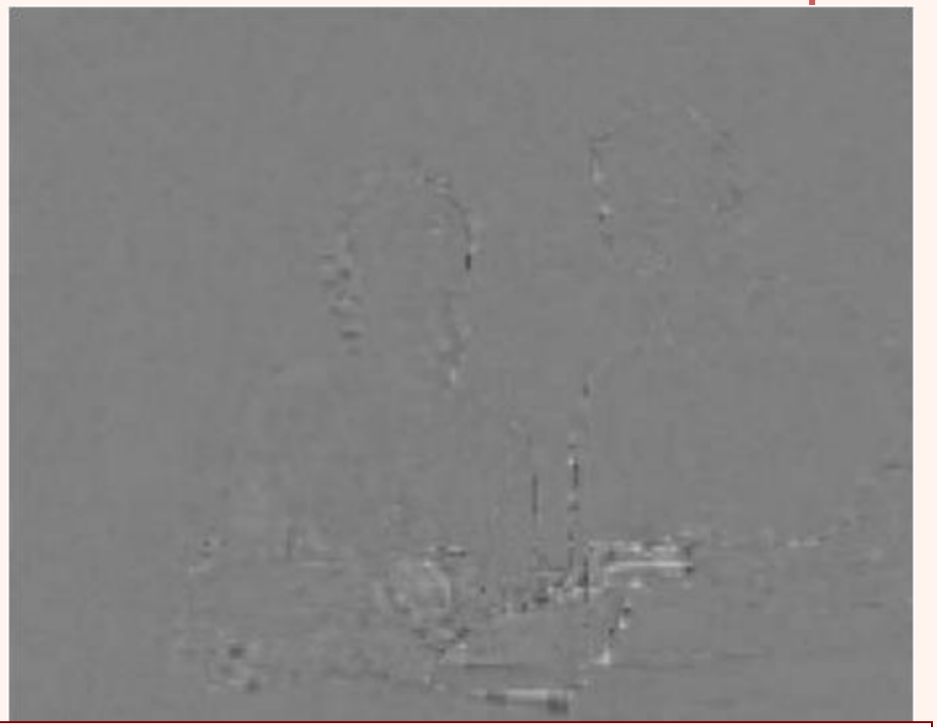
نمونه‌برداری غیرصحیح

- با توجه به آزمایش‌های صورت گرفته فشرده‌سازی مدل نیم‌نمونه‌ای نسبت به مدل صحیح، بازده بسیار بیشتری را نسبت به گزینش مدل ربع‌نمونه‌ای نسبت به نیم‌نمونه‌ای دارا می‌باشد.
  - استفاده از گام کوچک‌تر اگر چه کیفیت بالاتری را ارائه می‌نماید لیکن پیچیدگی بسیار بالایی را طلب می‌کند.



Residual (4 × 4 blocks, half-pixel compensation)





Residual (4 × 4 blocks, quarter-pixel compensation)