جَامعة المَنارة
MANARA UNIVERSITY

وسائط متعددة
المحاضرة العاشرة

د. أحمد أحمد

أ.د. فادي غصنه

الفيديو

خطوات تحويل الفيديو التشابهي إلى رقمي:

١ـ الترشيح (Filtering) لتجنب تداخل الأطياف.

٢ـأخذ العينات وهنا نميز بين مرحلتين:

ـأخذ العينات المكاني أو الفراغي (Spatial Sampling) ويعني تمثيل الصورة بمستطيل مقسم إلى أجزاء صغيرة تمثل عينات تدعى البيكسل.

ـ أخذ العينات الزمني (Temporal Sampling) ويعني أخذ لقطات سريعة و بفواصل زمنية دورية و بإعادة تشغيل سلسلة الإطارات هذه تظهر الصورة المتحركة.

٣ـ التكميم والترميز: حيث يتم تكميم العينات وترميزها.

فوائد الضغط الفيديوي

- الفيديو هو سلسلة زمنية مرتّبة من الإطارات (الصور) تُنشئ وهم الحركة. أحد حلول ضغط الفيديو هو الترميز التنبؤي اعتمادًا على الإطارات السابقة.
- يشكل الفديو حجم كبير في حالة عدم ضغطه.
- مثال:
  - إذا كان الفيديو غير المضغوط (**video** CIF) بدقة ٣٥٢ × ٢٨٨ فقط، فسيحتاج أكثر من ٣٥ $Mbps$. يحتاج التلفزيون عالي الدقة **(HDTV)**: ١٩٢٠ × ١٠٨٠، ٣٠ اطار بالثانية **(fps)**، ٨ بت لكل لون من الالوان الثلاثة ، لـ **1.5 Gbps**.
  - يتم الضغط: عن طريق استخدام التكرارات المكانية والزمانية. التكرار الزمني: الإطارات في نفس المشهد متشابهة جدًا. أما المكاني فيعني أنه يوجد تشابه في نفس الاطار.

فوائد الضغط الفيديوي

- تقنية شائعة:
  - ترميز الإطار الأول باستخدام تقنية ضغط الصور الثابتة.
  - ثم ترميز كل إطار لاحق عبر:
    - الترميز بين الإطارات **(Inter frame encoding)**: تحديد الفروقات بين الإطار الحالي والإطار السابق وترميزها (خطأ التنبؤ).
    - الترميز داخل الإطار **(Intra frame encoding)**: إذا كان الإطار مختلفًا جدًا عن سابقه، يتم ترميزه بشكل مستقل، مشابه لتقنية **JPEG**.
  - **H.261** و **H.263** هما معياران أساسيان لضغط الفيديو. وهما مخصصان بشكل أساسي للمؤتمرات عن طريق الفيديو.

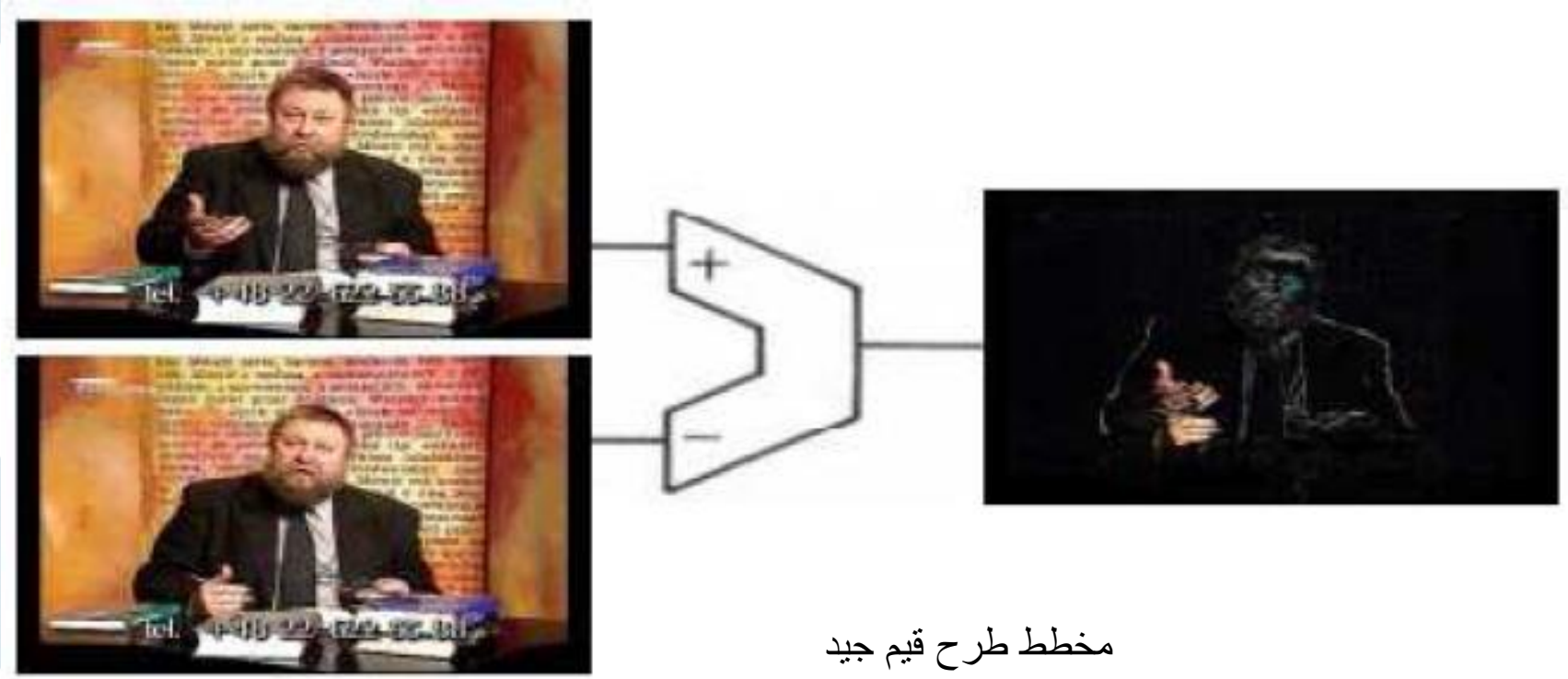
كيفية تشكيل التنبؤ

- **الهدف:** تقدير قيم البكسل في الإطار الحالي (الهدف) استنادًا إلى القيم الموجودة في إطار واحد أو أكثر (إطارات مرجعية).
- التكرار الزمني **(Temporal redundancy):** معظم الفيديو (الاطارات) لا يتغير بمرور الوقت.
- التكرار المكاني **(Spacial redundancy):** يوجد تكرار في قيم البكسلات المتجاورة من الاطار نفسه.
- هناك طريقتان رئيسيتان لتشكيل التنبؤ:

١. اختلاف الإطارات **(Frame Difference FD):** لا يوجد نموذج حركة، يعرف خطأ التنبؤ على أنه الفرق بين قيم البكسل في الإطار الحالي وتلك الموجود في الإطار المرجعي.
   - مناسبة عندما عدد قليل من العناصر يغير موضعه بين الاطارات، أما عندما تكون التغييرات كبيرة فسنحصل على قيم فروقات كبيرة.

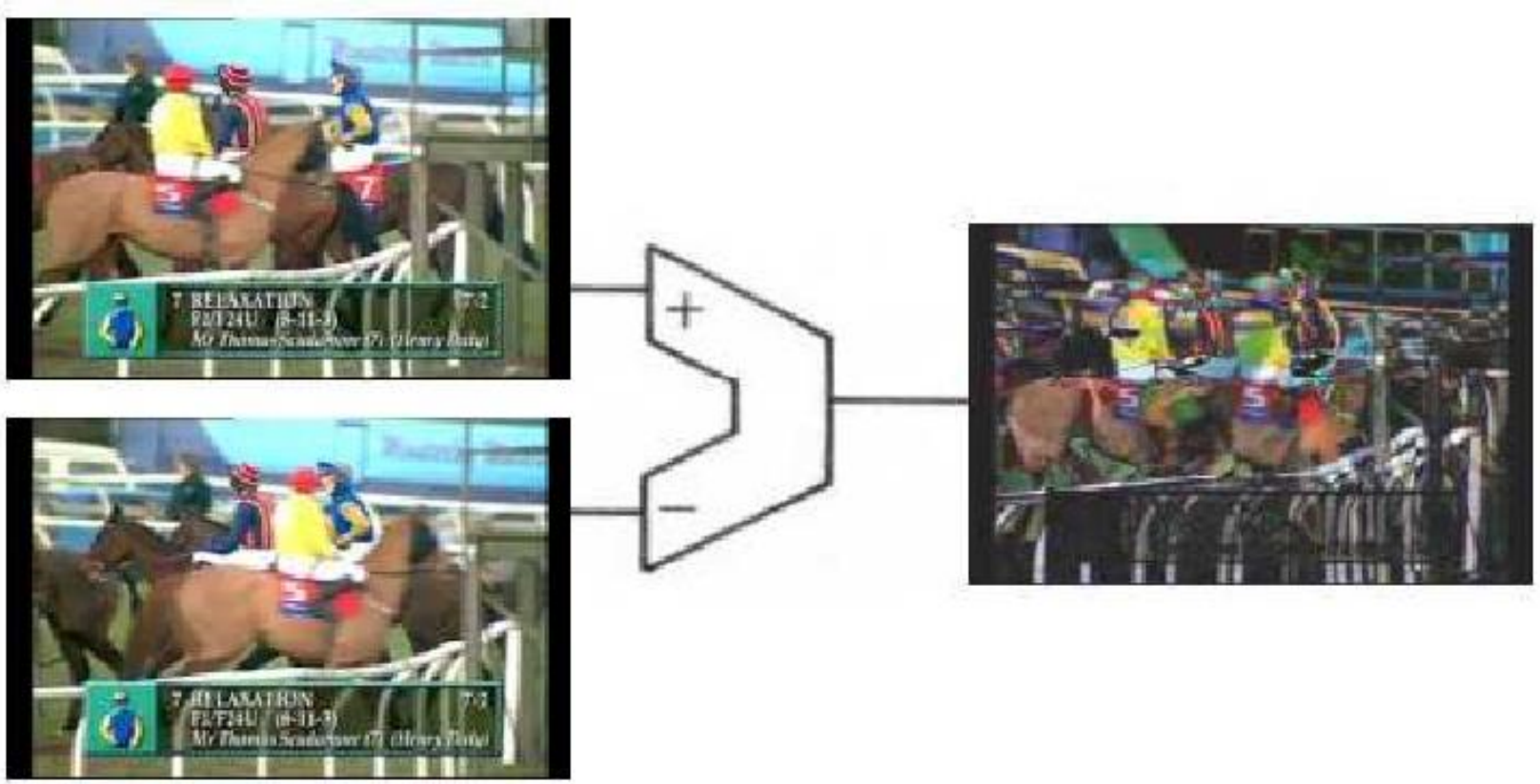
٢. التنبؤ المعوّض للحركة **(MCP: Motion-Compensated Prediction)**

التكرار الزمني: مثال



مخطط طرح قيم جيد

التكرار الزمني: مثال

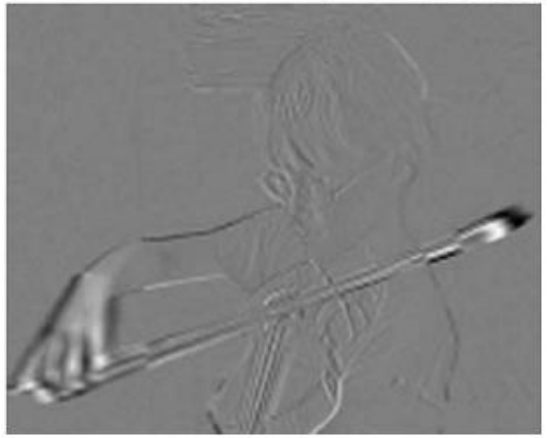


مخطط طرح قيم سيء

الضغط بين الإطارات

Frame 1

Frame 2



Difference

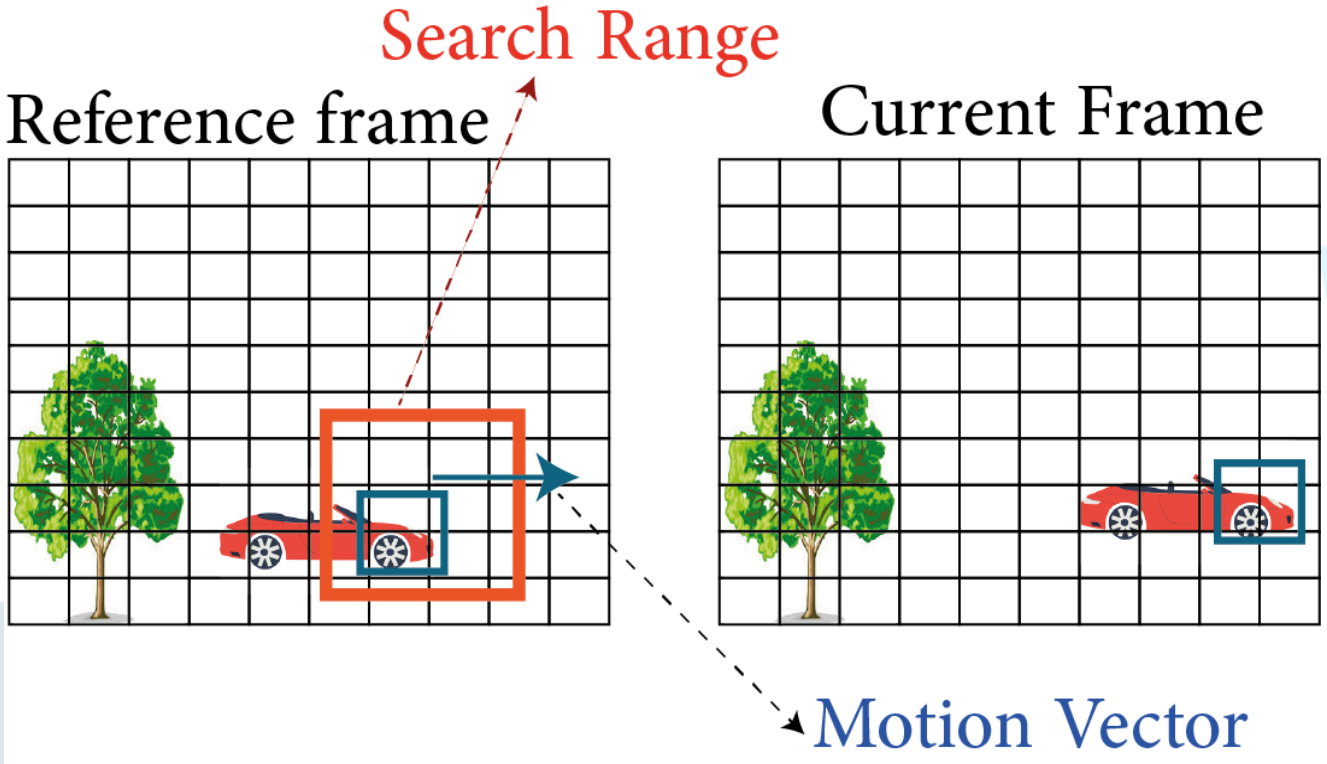
مخطط طرح قيم جيد

تقدير الحركة (ME: Motion Estimation)
وتعويض الحركة (MC: Motion Compensation)

- الهدف الأساسي:
  - تقليل متوسط عدد البتات المستخدمة لتمثيل تسلسل فيديو مع الحفاظ على جودة فيديو كافية.

- تُستخدم **ME** و **MC** على نطاق واسع في تطبيقات ترميز الفيديو بين الإطارات.
  - تقدير الحركة: هو تقدير الحركة بين الإطارات (حساب شعاع الحركة)، و الذي هو الإزاحة المكانية بين الكتلة الحالية (الكتلة الكبيرة **Macroblock**) والكتلة المحددة من إطار المرجع. يقوم المرمز بحساب الفرق بكسلًا بكسلًا بين الكتلة المحددة من إطار المرجع والكتلة الحالية.
  - الغاية إيجاد المنطقة الأكثر تطابقًا في الإطار المرجعي.
  - تعويض الحركة: هو استغلال معلومات الحركة وتحريك أجزاء الإطار حول مكانها طبقاً

تقدير الحركة ME وتعويض الحركة MC



- عند البحث و المقارنة:
  - في حال وجود تطابق، يُرسَل خطأ التنبؤ إضافة لشعاع الحركة.
  - في حال عدم وجود تطابق، يُرمَّز الـ **Macroblock** نفسه بدلاً من خطأ التنبؤ.

ضغط الفيديو بناءً على تعويض الحركة

- نهج الترميز الهجين:
- التنبؤ بالاختلافات بين إطارات الفيديو وتعويضها لإزالة التكرار الزمني.
- اكتشاف إزاحة البكسلات أو المناطق المقابلة.
- قياس فروقها (تعويض الحركة).

- تطبيق ترميز التحويل (**transform coding**) على الإشارة المتبقية (الفروقات) لتقليل التكرار المكاني.

- تعتمد جميع خوارزميات ضغط الفيديو الرقمي الحديثة (بما في ذلك **H.264** و **H.265**) على نهج الترميز الهجين

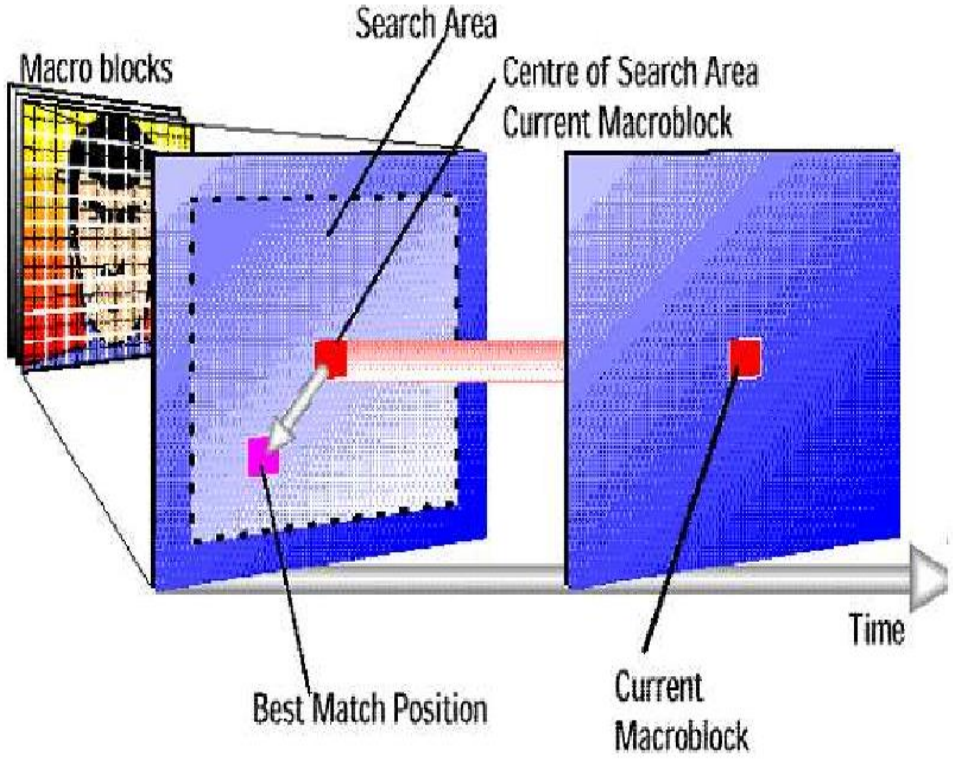
ضغط الفيديو بناءً على تعويض الحركة

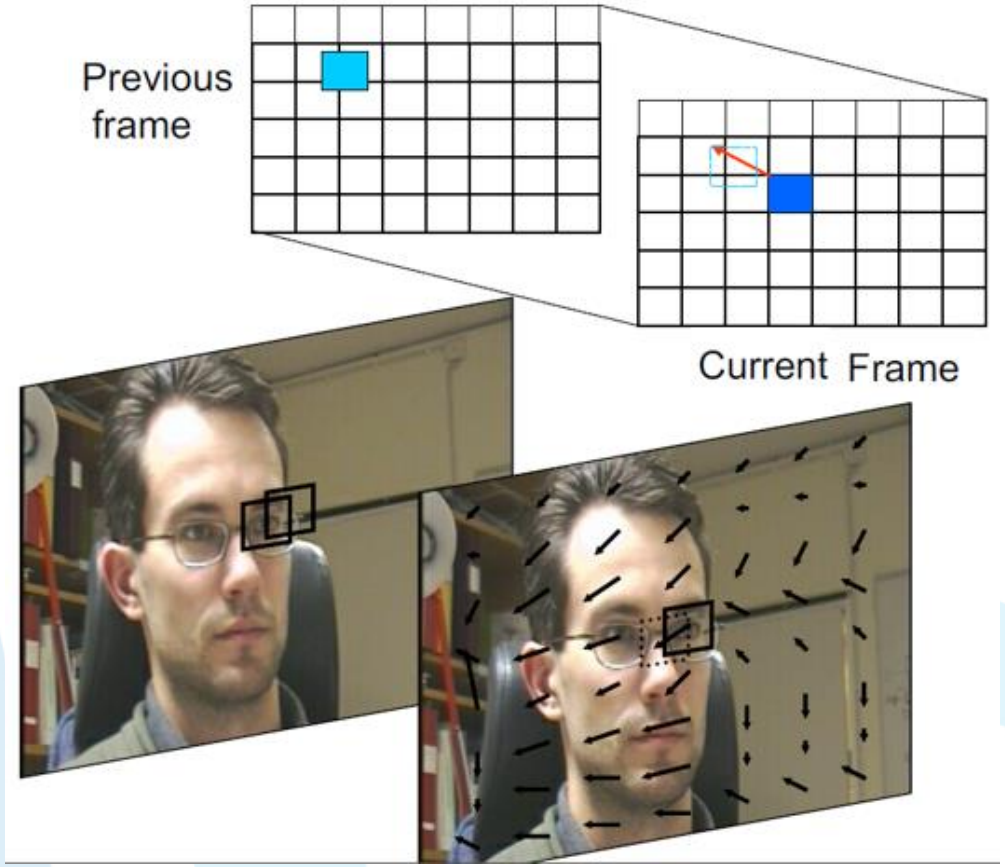
- الخطوات الثلاث الرئيسية لخوارزميات ضغط الفيديو استنادًا إلى تعويض الحركة هي:
  ١. تقدير الحركة (ايجاد شعاع الحركة).
  ٢. التنبؤ القائم على تعويض الحركة.
  ٣. اشتقاق خطأ التنبؤ - الفرق.

- تُقسّم كل صورة إلى كتل كبيرة **(macroblocks)** بأبعاد $N \times N$.
  - افتراضيًا، $N = 16$ لصور السطوع **(luminance images)** .
  - للصور الملونة ، $N = 8$.

- **ملاحظة:** يتم تعويض الحركة على مستوى الكتل الكبيرة.

ضغط الفيديو بناءً على تعويض الحركة





ضغط الفيديو بناءً على تعويض الحركة

- التنبؤ الأمامي: يُعتَبَر الإطار المرجعي إطارًا سابقًا.
- التنبؤ الخلفي: يُعتَبَر الإطار المرجعي إطارًا مستقبليًا.
- خطأ التنبؤ: الفرق بين كتلتي الماكرو المتقابلتين.

- عند وجود تطابق بين كتلتي الماكرو، نقول إن كتلة الماكرو المستهدفة مُتنبَّأة من كتلة الماكرو المرجعية.

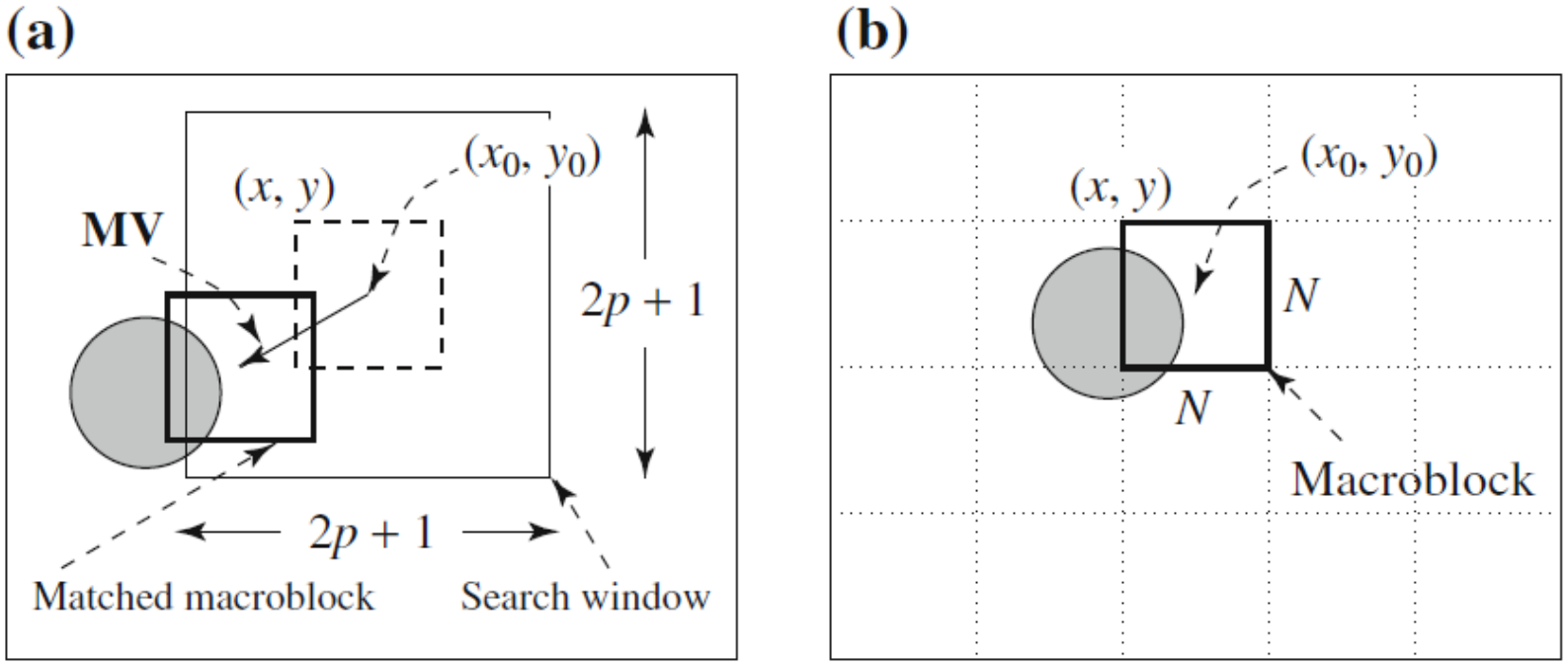
- ملاحظة: لضغط الفيديو المبني على **MC**، يلزم ترميز أشعة الحركة وكتل الماكرو المتباينة فقط.

البحث عن أشعة الحركة

- الإزاحتان الأفقية والرأسية i و j تقعان في النطاق $[-p,p]$ حيث p عدد صحيح موجب ذو قيمة صغيرة.
  - يؤدي هذا إلى إنشاء نافذة بحث بحجم $(2p + 1) \times (2p + 1)$.
- يمكن وضع مركز كتلة الماكروبلوك $(x0,y0)$ عند كل موضع من مواضع الشبكة في النافذة.

- نستخدم الزاوية العلوية اليسرى (x,y) كنقطة أصل (بداية) لكتلة الماكروبلوك في إطار الهدف.
  - $C(x + k,y + l)$ البكسلات في كتلة الماكرو في (إطار الهدف).
  - $R(x + i + k,y + j + l)$: البكسلات في كتلة الماكروبلوك (إطار المرجع).
  - k و l هما مؤشرا البكسل في كتلة الماكروبلوك.
  - i و j هما الإزاحتان الأفقية والرأسية على التوالي

البحث عن متجهات الحركة



Macroblocks and motion vector in video compression:
a reference frame; **b** target frame

معايير المطابقة

- هناك عدة معايير مطابقة:
  - خطأ المتوسط التربيعي (Mean Square Error MSE) .
  - مجموع الفرق المطلق (Sum of Absolute Difference (SAD)) .
  - مطابقة عدد البكسلات (Matching Pixel Count (MPC)).
  - متوسط الفرق المطلق (Mean Absolute Difference (MAD)).
- مثال MSE: يتم البحث عن القيمة الدنيا لـ MSE بين كتلة الماكروبلوك في إطار الهدف والكتلة داخل نافذة البحث W في إطار المرجع.

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

$$(u,v) = [(i,j)|MSE(i,j)\ is\ minimum, i \in (-p,p), j \in (-p,p)],$$

$$\begin{bmatrix}\hat{\imath}\\ \hat{\jmath}\end{bmatrix} = \arg\min MSE(i,j)$$ where N is the size of the macroblock

تقنيات البحث

- ❑ لاستخدام إحدى التقنيات:
  - اختيار معيار المطابقة (أحد المعايير الموضحة سابقًا).
  - تقدير شعاع الحركة: البحث عن الكتلة التي تُلبي المعيار المُختار ضمن نافذة البحث.

- ❑ يوجد تقنيات بحث متعددة:
  - عتبة البحث البحث التسلسلي أو الشامل أو الكامل.
  - البحث اللوغاريتمي ثنائي الأبعاد (سنعرض مثال لاحقاً)
  - البحث ثلاثي الخطوات (البحث اللوغاريتمي).
  - تقدير الحركة الهرمي أو الهرمي. ...وهكذا.

تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- **الخوارزمية:**

١. **احسب حجم الخطوة:** $s = 2^{\lfloor log_2 p \rfloor - 1}$ **(مثال:4 =[4.3])**

- **تقارن الخوارزمية B بالكتل الخمس في المواضع التالية:**

(a,b), $(a,b+s)$, $(a,b-s)$,$(a+s,b)$ and $(a-s,b)$ **في الإطار السابق.**

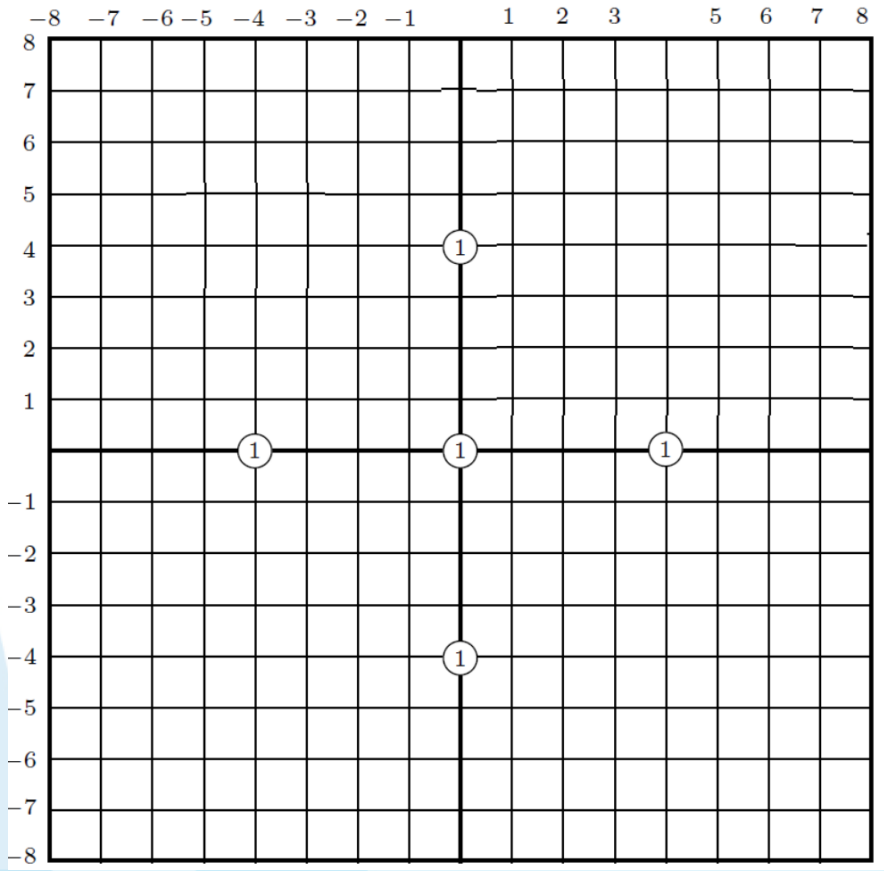
٢. **يتم اختيار أفضل تطابق بين الكتل الخمس.**

- **نرمز لموضع هذه الكتلة بـ (x,y) .**
- **إذا كان $(x,y) = (a,b)$, فإن s يُقسّم إلى النصف. وإلا، تبقى s كما هي، وينتقل مركز البحث (a,b) إلى(x,y).**

٣. **إذا كان $s = 1$، فإن: يتم البحث عن المربعات التسع حول مركز البحث (a,b)ويصبح أفضل تطابق بينها هو نتيجة الخوارزمية. وإلا، تنتقل الخوارزمية إلى الخطوة ٢.**

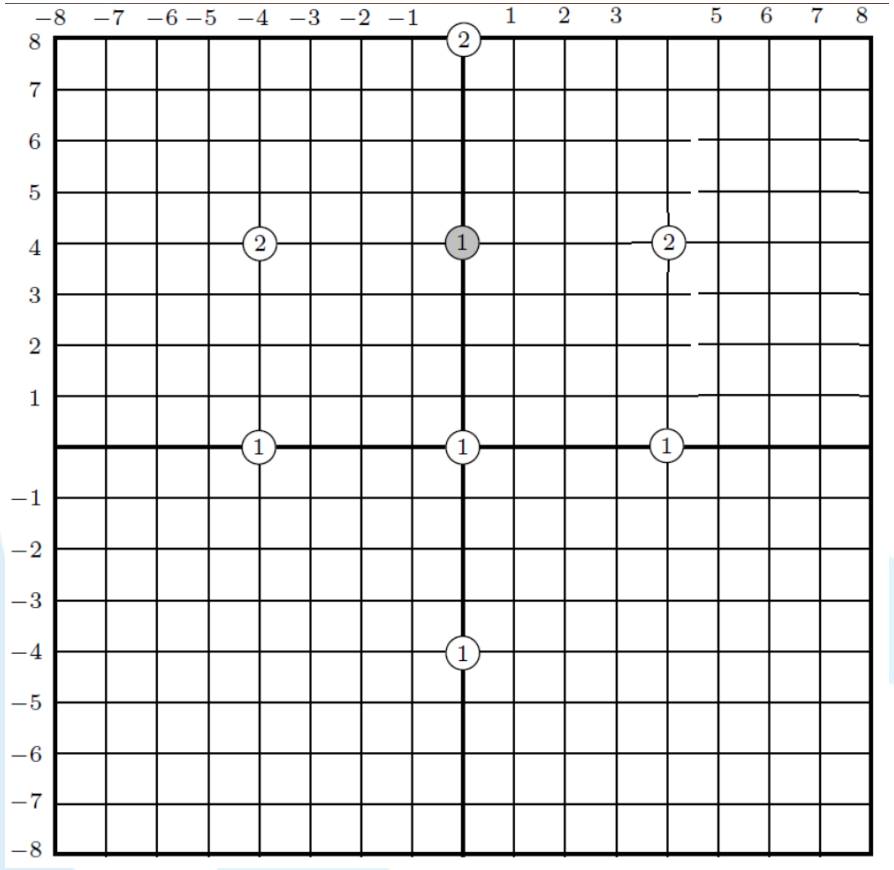
تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- افتراض أن: $p = 8$.
- حجم الخطوة $s = 2^{\lfloor log_2 8 \rfloor - 1} = 4$
- إحداثيات إطار الكتلة الحالية B هي $(0,0)$، و بالتالي يقتصر البحث على منطقة الكتلة (17 × 17 المتمركزة حول الكتلة B.
- البحث عن الكتل الخمس (المُشار إليها بالرقم ١) في المواقع $(0,0), (4,0), (4,0), (0,4), (0, -4)$ في الإطار السابق.
- نفترض أن أفضل تطابق لهذه الكتل الخمس هو عند $(0,4)$.



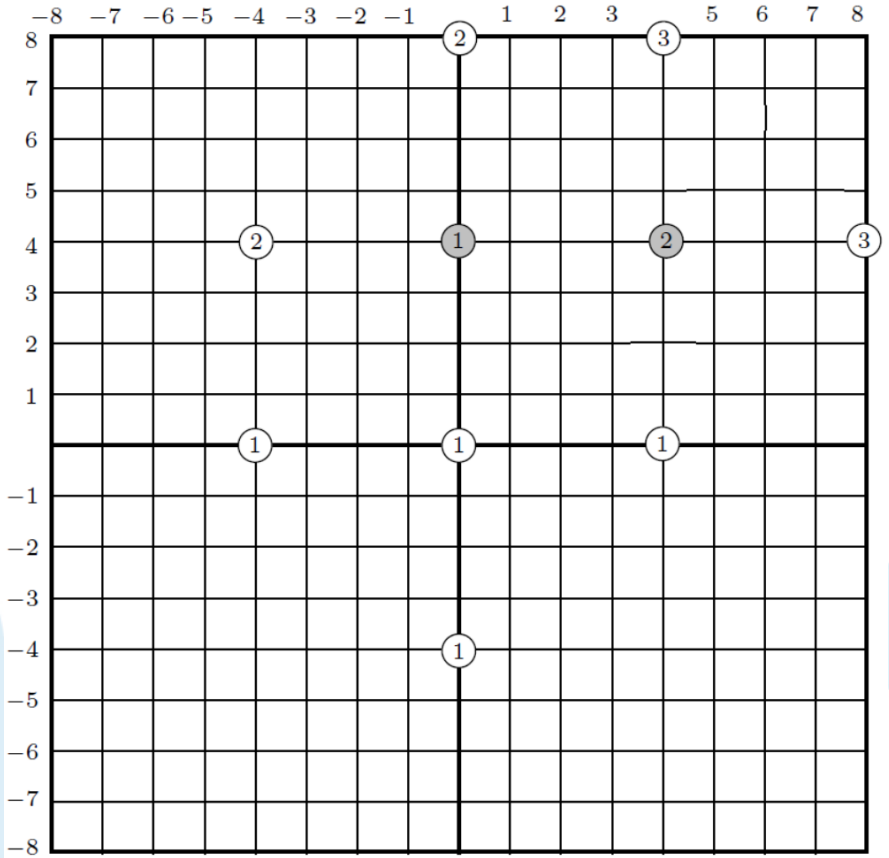
تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- يصبح هذا مركز البحث الجديد، ويتم البحث في الكتل الثلاث (المُسمّاة ٢) في المواقع (4,4), (4,4) و (8,0) في الخطوة الثانية.
- بافتراض أن أفضل تطابق بين هذه الكتل الثلاث هو في الموقع (4,4) .



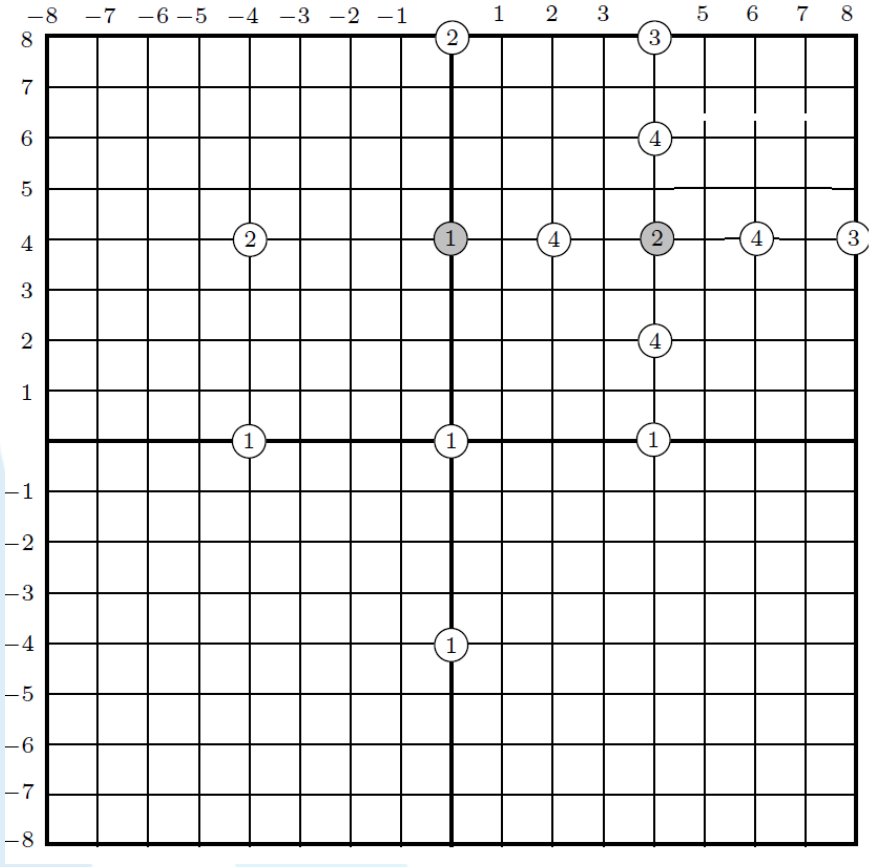
تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- الخطوة التالية هي البحث عن الكتلتين (المُسمَّيَتين بالرقم (٣) في الموقعين (8,4)و (4,8), والكتلة (المُسمَّية بالرقم (٢) في (4,4) والكتلة (١) في (0,4)و (4,0)
- بافتراض أن (4,4)هي أفضل تطابق مرة أخرى.



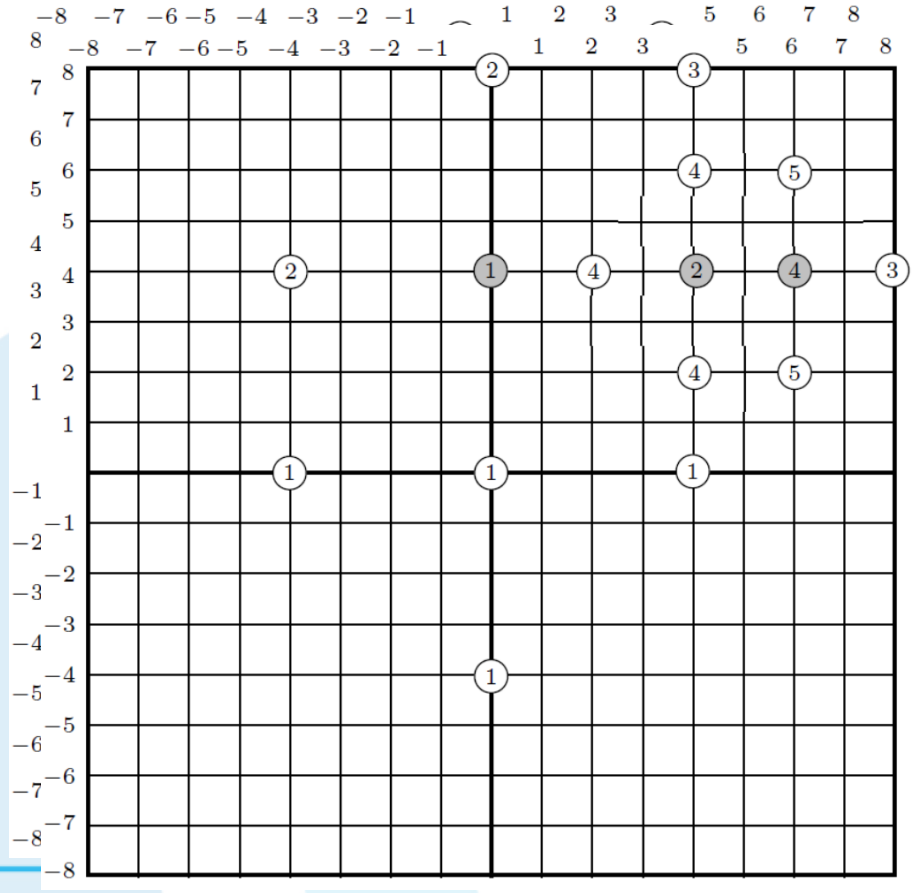
تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- بما أن (4,4)تقع في مركز منطقة البحث الحالية، فإن قيمة **(s)** تُقسّم إلى النصف، أي إلى ٢.
- الخطوة التالية هي البحث في المربعات الأربع المُسمّاة بالرقم (٤)، والمتمركزة في (4,4).
- بافتراض أن أفضل تطابق هو في (6,4).



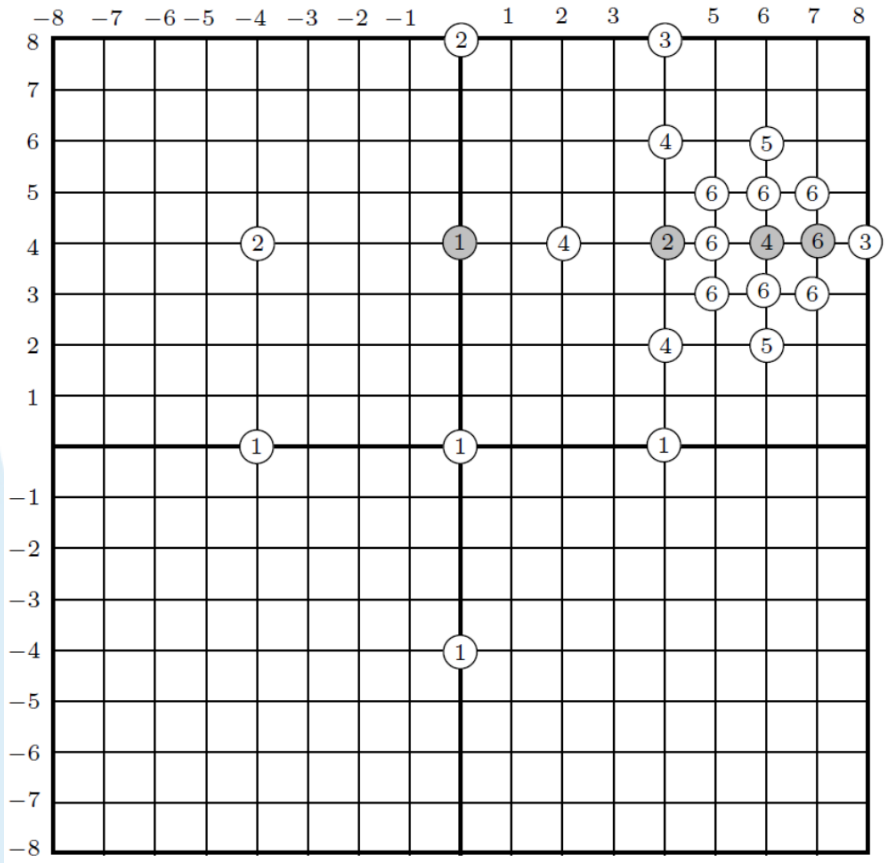
تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- تم البحث عن الكتلتين المسمى بـ (٥). بافتراض أن **(6, 4)** هي أفضل تطابق.



تقنيات البحث: البحث اللوغاريتمي ثنائي الأبعاد

- يقسَّم s إلى ١، ويتم البحث عن الكتل الثمانية المُسمَّاة (٦).
- يُظهر الرسم التخطيطي أنه تم العثور على أفضل تطابق أخيرًا في الموقع (٤،٧).



خطوات ضغط الفيديو

- يمكننا تلخيص معايير ضغط الفيديو على النحو التالي:

١. تجزئة الإطار: يُقسّم الإطار الحالي إلى كتل متساوية الحجم وغير متداخلة (كتل ماكرو). قد تكون كتل الماكرو مربعات أو مستطيلات، ولكنها عادةً ما تكون مربعات. يُعد حجم الكتلة مهمًا، لأن الكتل الكبيرة تُقلل من فرصة العثور على تطابق، بينما تُنتج الكتل الصغيرة العديد من اشعة الحركة.

٢. تحديد معايير البحث.

٣. تحديد تقنية البحث.

خطوات ضغط الفيديو

❑ ملاحظات:

✓ يُرمَّز الإطار الأول باستخدام طريقة ضغط الصورة الثابتة **(JPEG)**. تُرمَّز الإطارات المتتالية باستخدام مبادئ الاختلاف، إن أمكن.

- كلمة " إن أمكن " تعني عندما لا يوجد فرق كبير بين الإطارات المتتالية. إذا لم يكن الأمر كذلك، تُطبَّق طريقة ضغط الصورة الثابتة مجددًا على الإطار الحالي.

✓ لتحديد حركة جسم في إطار، بالنسبة للإطار السابق أو اللاحق، يجب أولًا تحديد حدود الجسم في الإطارين الحالي والمرجعي. يؤدي تحديد ذلك إلى شعاع حركة بمركبات أفقية ورأسية، تقاس بعدد البكسلات.

✓ بعد تحديد شعاع الحركة، تُفعّل عملية المحاذاة (التنبؤ بتعويض الحركة **(MC)** هذا يعني محاذاة الكائنات وتمييزها بين الإطارات المتتالية.

خطوات ضغط الفيديو

❑ ملاحظات:

✓ ينتج عن تعويض الحركة (MC) ما يُسمى خطأ التنبؤ.

- خطأ التنبؤ: يمثل تباينًا يعتمد على مقدار الحركة بين الإطارات المتتالية. الحركة الكبيرة للأجسام بين الإطارات المتجاورة زمنيًا تُنتج تباينًا كبيرًا.

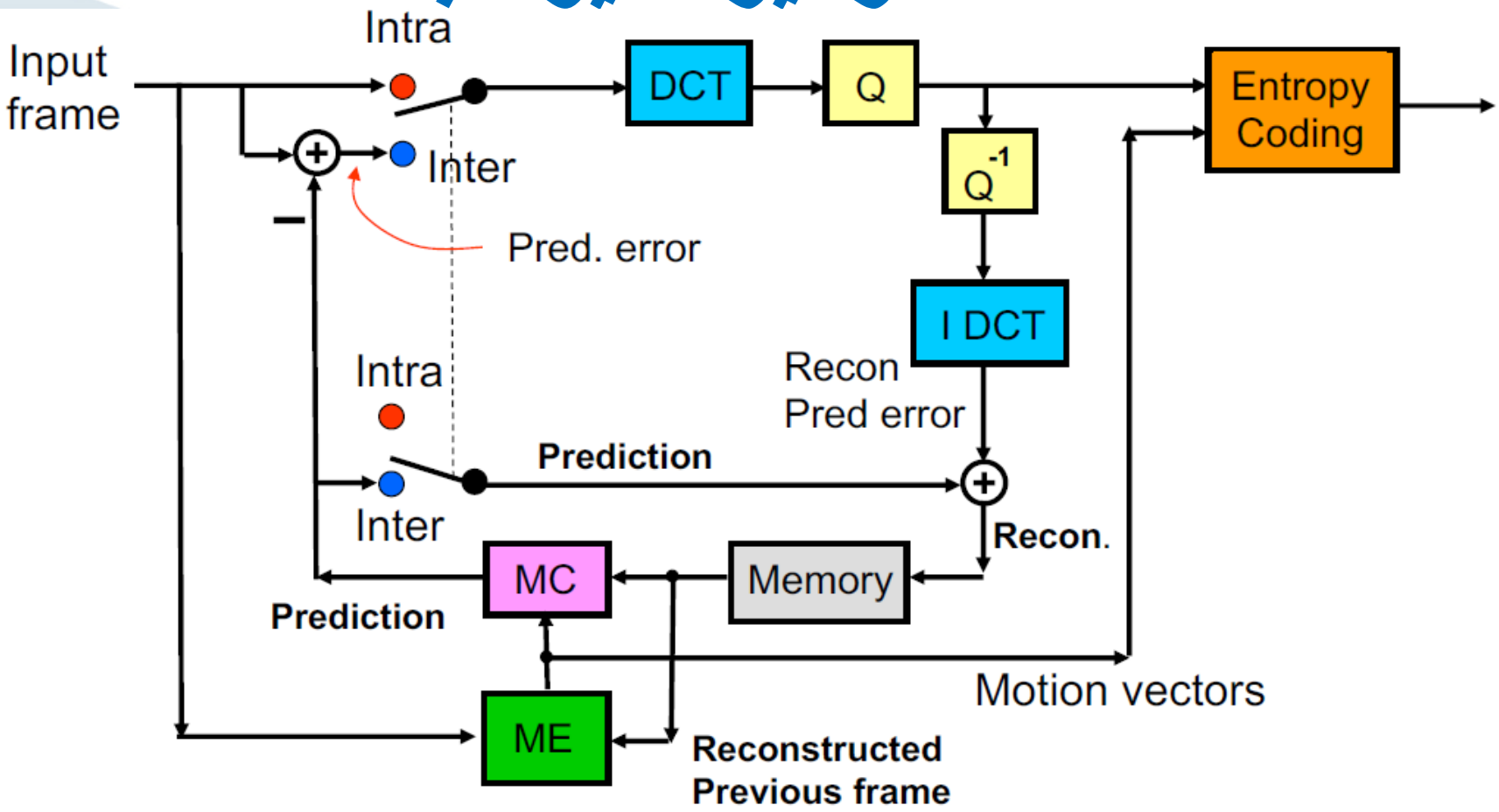
✓ نفترض أن حركة الجسم انتقالية فقط: يسارًا ويمينًا، ولأعلى وأسفل.

✓ إذا كان الإطار مختلفًا تمامًا عن سابقه، فيجب ترميزه بشكل مستقل عن أي إطار آخر.

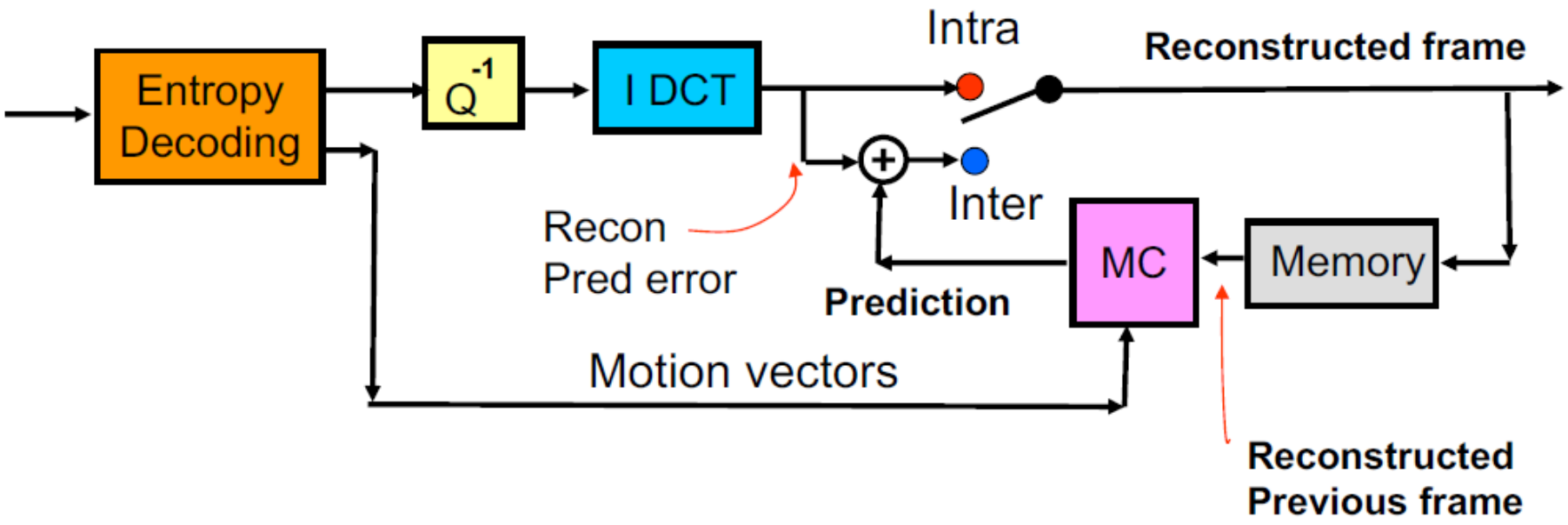
- يُسمى الإطار المُرمَّز باستخدام سابقه (inter frame).
- يُسمى الإطار المُرمَّز بشكل مستقل (intra frame).

✓ MC: معلومات تُحدد: حدود الجسم + موقعه السابق + موقعه الحالي.

مخطط ترميز فديو مبسط



مخطط فك ترميز فديو مبسط



خطوات ضغط الفيديو

❑ ملاحظات:

✓ ضغط الفيديو (MC) عادةً ما يكون فاقدا للبيانات.

- ينتج ترميز الإطار F_i من سابقه F_{i-1} بعض التشوهات. ونتيجةً لذلك، فإن ترميز الإطار التالي F_{i+1} من (المشوه بالفعل) F_i (المشوه) يزيد من التشوه. هذا يعني أنه إذا فقد الإطار F_i بعض البتات، فإن جميع الإطارات التي تليه، تُفك ترميزها بشكل غير صحيح، مما قد يؤدي إلى أخطاء متراكمة. الخلاصة: يجب استخدام ترميز الإطارات (Intra frame) من وقت لآخر داخل التسلسل، وليس فقط في بدايته.
- أنواع الاطارات بشكل عام و التي يعتمد عليها الترميز هي:

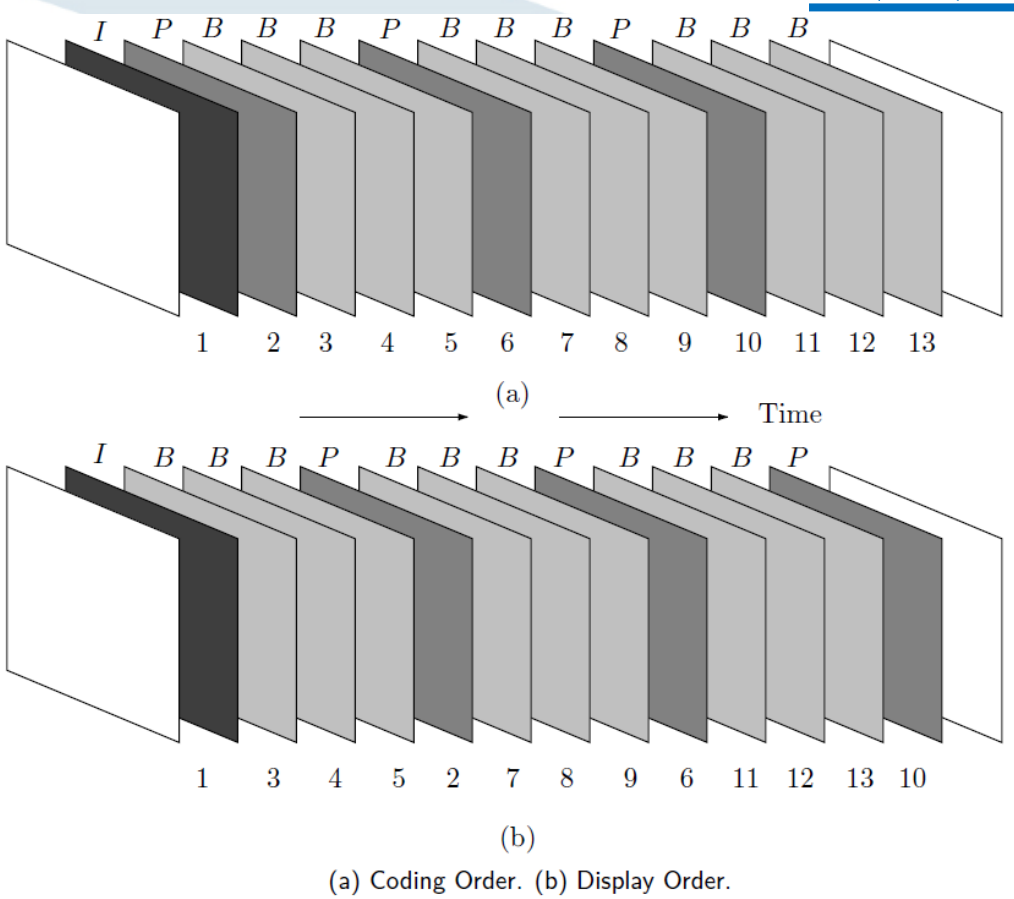
I- Frame (intra frame) ، P- Frame (inter frame) (للتنبؤ)

B- Frame مُرمَّز بناءً على الإطارات الماضية والمستقبلية (ثنائي الاتجاه).

خطوات ضغط الفيديو

- ملاحظات:
  - تكون معظم الإطارات في عرض فيديو مضغوط من النوع (I-frame, P-frame and B-frame).
  - لفك الترميز
    - يُفك ترميز I-frame بشكل مستقل عن أي إطار آخر.
    - يُفك ترميز P-frame باستخدام I-frame أو P-frame السابق. يُفك ترميز B-frame باستخدام I-frame أو P-frame السابق واللاحق.
  - ولكن... لماذا B-frame؟ في بعض الإطارات، تغطي حركة جسم في الصورة تدريجيًا منطقة خلفية. هذا يؤدي إلى أن منطقة ما قد تكون معروفة جزئيًا فقط في الإطار الحالي، ولكنها قد تكون معروفة بشكل أفضل في الإطار التالي. النتيجة: الإطار التالي هو المرشح الأمثل للتنبؤ بهذه المنطقة في الإطار الحالي.

خطوات ضغط الفيديو



(a) Coding Order. (b) Display Order.

- في الشكل التالي:
- **(a)**: سلسلة من الإطارات بالترتيب الذي تُولّد به بواسطة المرمز (ويُدخلها بواسطة مُفكّك الترميز).

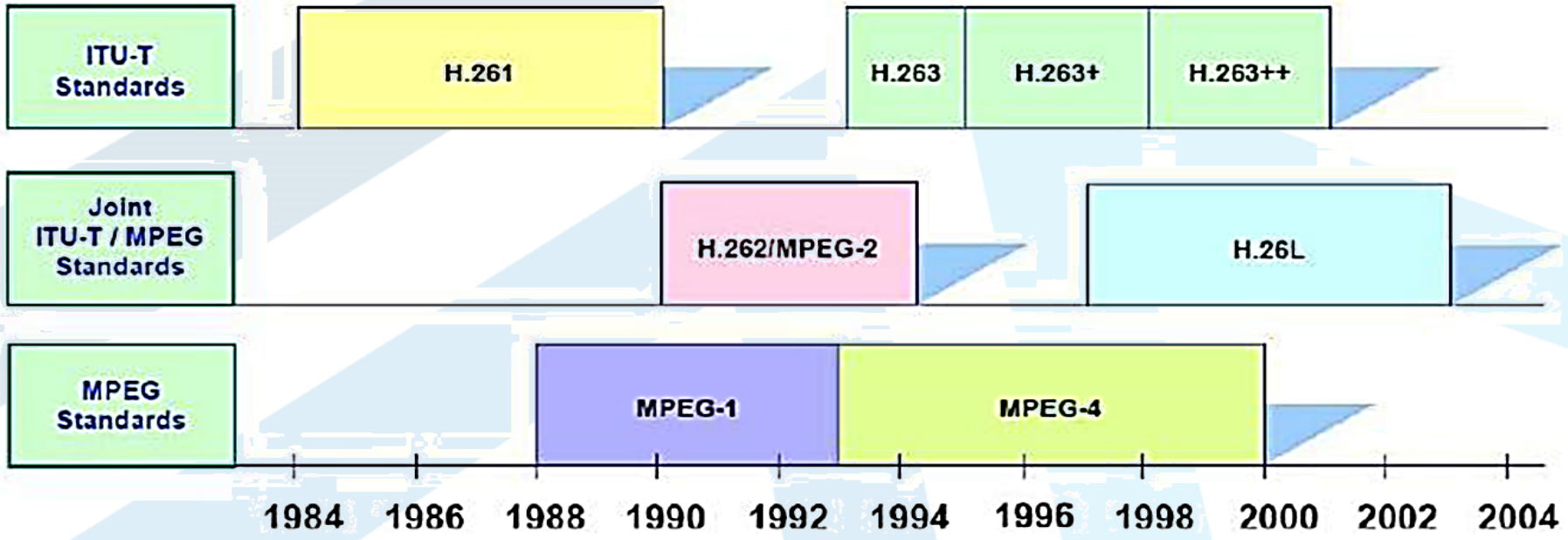
- **(B)**: نفس التسلسل بالترتيب الذي تُخرِج به الإطارات بواسطة مُفكّك الترميز وتُعرَض به.
- ولكن يجب عرض الإطار **2** بعد الإطار **5**. يجب أن يكون لكل إطار طابعان زمنيان: وقت ترميزه ووقت عرضه.

معايير الضغط الفيديوي

7- تطور المعايير العالمية للضغط الفيديوى

Development of International Standards For Video Compresses

يبين الشكل (5-1) المخطط الزمني لظهور معايير ضغط الفيديو



معايير الضغط الفيديوي

Compression format	*ISO/IEC number Issue date*	*Target bandwidth bit/s*	*Typical resolution pixels*	*Application*
H.261	1988–1990	384 k–2 M	176 × 144 or 352 × 288	Video conferencing, low delay
H.263	1992	28.8 k–768 k	128 × 96 to 720 × 480	Video conferencing
MPEG-1	11172 1993	400 k–1.5 M	352 × 288	CD-ROM
MPEG-2, MP@ML	13818 1994	1.5 M–15 M	720 × 480	Broadcast television, DVD
MPEG-4	14496 1998	28.8 k–500 k	176 × 144 or 352 × 288	Fixed and mobile web
AVC, H.264	14496–10 2002			General purpose

الجدول (2-5) أنواع الضغط وعرض الحزمة الخاص لكل منهم وتطبيقاتهم

جَـامعة
المَـنارة
MANARA UNIVERSITY

نـهايـة المحـاضرة
أسئلـة؟؟