

SACCDMM - Curs 09

Codarea intercadre; Estimarea mișcării intercadre

Sl.Dr.Ing. Camelia FLOREA

Compresia secvențelor video

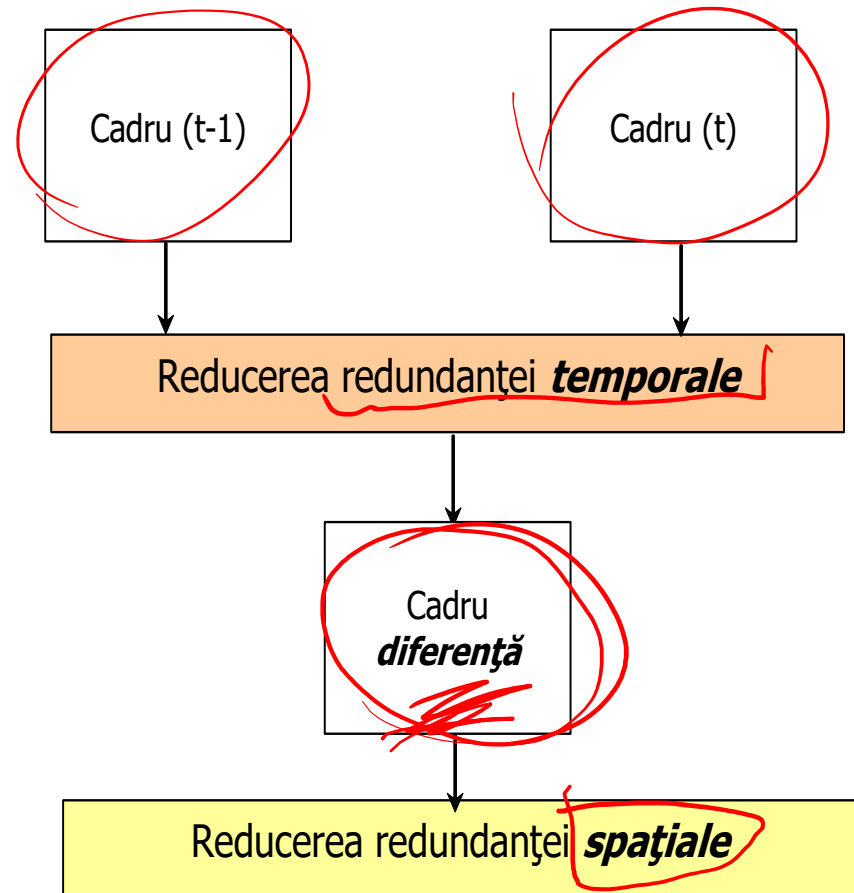
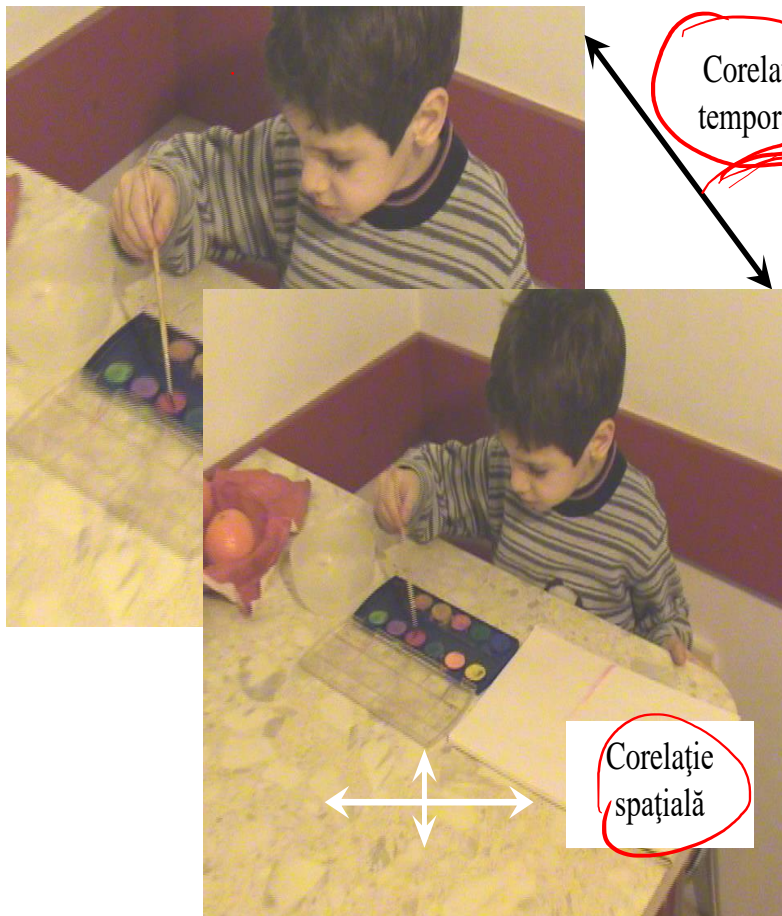
- Necesitatea măririi raportului de compresie pentru video
- Secvență: 512x515x8bițix3planex25imagini/sec rezultă 150Mbiți/sec
 - raport compresie imagini în general 12:1
 - rezultă 12,5 Mbiți/sec PREA MARE
 - este necesar un raport de compresie 100:1

=> NECESAR exploatarea ***codării intracadru si intercadru***



Corelația temporală și spațială

- 2 etape: redundanța intercadru, redundanța intracadru



Reducerea redundanței

- Spațiale
 - Cele utilizate și la imaginile statice
- Temporale
 - exploatează redundanța dintre cadrele succesive
 - Se codează doar modificările – “diferențele”, datorate:
 - mișcării obiectelor în imagine
 - mișcării camerei
 - zoomingului, etc
 - urmărirea fiecărui pixel
 - > compromis eficiență-calcul -> Împărțirea în **macro-blocuri** 16x16 pixeli -> Urmărirea macroblocurilor de la un cadru la altul

Segmentarea imaginii în zone staționare și “în mișcare”

- sistemului vizual uman este foarte slab în cazul imaginilor care conțin simultan frecvențe spațiale și temporale înalte.
=> zonele cu modificări rapide dintr-o imagine pot fi reprezentate cu o amplitudine și o rezoluție spațială mai redusă în comparație cu zonele staționare.
- Acest lucru permite înlocuirea rezoluției spațiale cu rezoluția temporală și se folosește în scopul obținerii unor secvențe de imagini de calitate bună, la o rată de transmisie de $2 \div 2,5$ biți/pixel.
- O asemenea metodă presupune segmentarea imaginii în zone staționare și “în mișcare”.
- *Segmentarea se poate* realiza prin evaluarea semnalului-diferența între două cadre succesive.
 - În zonele staționare se transmite diferența între cadre pentru fiecare al doilea pixel, iar pixelii rămași sunt repetați din cadrul precedent.
 - În zonele “în mișcare” se folosește o subeșantionare orizontală de 2:1, elementele netransmise fiind reconstituite prin interpolare de-a lungul liniilor.
- Cele mai importante distorsiuni pot apărea la muchiile ascuțite ce se deplasează în imagine cu viteză moderată.

Codarea “*cu înlocuire condiționată*”

- se bazează pe detecția și codarea zonelor în mișcare care se înlocuiesc de la un cadru la altul:

$$e(m,n,i) = u(m,n,i) - u^*(m,n,i-1)$$

- $u(m,n,i)$, pixelul de coordonate (m,n) în cadrul I.
 - $u^*(m,n,i-1)$ este valoarea reprodusă a pixelului de coordonate (m,n) din cadrul I-1.
-
- În cazul în care $e(m,n,i)$ depășește o valoare de prag η
=> ea este cuantizată și codată pentru transmisie.

Codarea “*cu înlocuire condiționată*”, la recepție:

- valoarea unui pixel se obține:
 - dacă acesta provine dintr-o zonă staționară
 - prin repetarea valorii pixelului din cadrul precedent,
 - dacă pixelul provine dintr-o zonă “în mișcare”
 - prin înlocuirea cu semnalul-diferență decodat

$$u(m, n, i) = \begin{cases} u^*(m, n, i-1) + e(m, n, i), & \text{daca } |e(m, n, i)| > \eta \\ u^*(m, n, i-1), & \text{in rest} \end{cases}$$

- Pentru a obține o rată de transmisie stabilă este necesară folosirea unui buffer de mărime rezonabilă și o strategie de control a bufferului adecvată.
- În aceste condiții se poate obține o rată de transmisie de 1 bit/pixel pentru nivele ale raportului semnal/zgomot mediu de 34 dB (39 dB în zonele staționare și 30 dB în zonele în mișcare).

Problema 1:

- În figura următoare sunt prezentate: un bloc de 8x8 pixeli codat și reconstruit dintr-un cadru de referință, și respectiv blocul de pe aceeași poziție spațială de 8x8 pixeli din cadrul curent (următor cadrului de referință) care urmează a fi codat. Dacă asupra secvenței video se aplică o codare cu înlocuire condiționată cu pragul =11,
- Cum arată matricea de eroare $E[8 \times 8]$ transmisă decodorului? Se va presupune o cuantizare fără pierderi a lui E .
 - Cum arată blocul de 8x8 pixeli din cadrul curent codat, reconstruit la decodor? Calculați MSE (eroarea medie pătratică) dintre blocul decodat (reconstruit) din cadrul curent și blocul original din cadrul curent.

Cadrul de referință =

120	128	120	140	160	200	210	192
120	120	130	164	164	200	200	180
120	100	80	80	80	90	120	164
100	70	60	64	64	60	60	100
70	60	32	32	32	32	60	70
100	80	80	70	70	100	132	132
100	90	90	100	100	100	100	132
100	100	90	80	80	90	100	132

Cadrul curent, de codat =

114	120	120	120	140	180	200	180
114	120	120	160	160	200	210	180
114	100	80	80	92	92	120	160
90	64	70	90	80	80	60	70
50	40	48	70	64	50	60	60
90	80	80	70	80	120	140	140
110	80	90	114	114	110	100	140
104	100	80	80	80	90	108	128

Cadrul de referinta =

120	128	120	<u>140</u>	160	200	210	192
120	120	130	164	164	200	200	180
120	100	80	80	80	90	120	164
100	70	60	64	64	60	60	100
70	60	32	32	32	32	60	70
100	80	80	70	70	100	132	132
100	90	90	100	100	100	100	132
100	100	90	80	80	90	100	132

~~110~~

Cadrul curent, de codat =

114	120	120	<u>120</u>	<u>140</u>	180	200	180
114	120	120	160	160	200	210	180
114	100	80	80	92	92	120	160
90	64	70	90	80	80	60	70
50	40	48	70	64	50	60	60
90	80	80	70	80	120	140	140
110	80	90	114	114	110	100	140
104	100	80	80	80	90	108	128

$$e = \text{Cadrul Curent} - \text{Cadrul Referinta}$$

$$e_{+1(m,n)} = \begin{cases} 2(m,n), & \text{dacă } 2(m,n) > m_{\max} \\ 0, & \text{altfel} \end{cases}$$

$m_{\max} = 11$

$E =$

-6	8	0	-20	-20	-20	-10	-12
-6	0	-10	-4	-4	0	10	0
-6	0	0	0	12	2	0	-6
-10	-6	10	26	16	20	0	-30
-20	-20	16	38	32	18	0	-12
-10	0	0	0	10	20	8	8
10	-10	0	14	14	10	0	8
8	0	-10	0	0	0	8	-4

$E_t =$

0	0	0	-20	-20	-20	0	-12
0	0	0	0	0	0	0	0
0	0	0	0	12	0	0	0
0	0	0	26	16	20	0	-30
-20	-20	16	38	32	18	0	0
0	0	0	0	0	20	0	0
0	0	0	14	14	0	0	0
0	0	0	0	0	0	0	0

$$MSE = \frac{1}{64} ((-6)^2 \cdot 4 + 8^2 \cdot 5 + 12^2 \cdot 10^2 + 4^2 \cdot 5 + 2^2) = 27.34$$

Cadruul reconstituit = Cadruul Referință + E_t

120	128	120	120	140	180	210	180
120	120	130	164	164	200	200	180
120	160	80	80	68	90	120	164
100	70	60	98	48	40	60	70
50	40	48	70	64	50	60	70
100	80	80	70	70	120	132	132
160	90	90	114	114	100	100	132
100	100	90	80	80	90	100	132

$$\eta = ?$$

$$E_t = 0$$

$$\eta = 39$$

\Rightarrow MSE - mare

$$MSE E = \frac{\sum_{i=1}^n (C_{rec.} - C_{currente\ de\ codat})^2}{(8 \times 8)} \quad \eta = ?$$

$$MSE = 0$$

$$\Rightarrow \eta = 1$$

$$E_t = E$$

Problema 2:

- În figura următoare sunt reprezentate cadrele 10 și 11 dintr-o secvență video. Ce valoare trebuie să aibă valoarea de prag de codare a erorii, η , pentru a coda cadrul 11 fără a transmite nici un semnal de eroare, folosind tehnica de codare intercadre cu înlocuire condiționată, considerând ca referință cadrul 10? Calculați în acest caz MSE la redarea cadrului 11.

12	8	7	9	5	6	6	3
4	5	3	0	2	1	4	5
1	2	3	5	3	2	3	2
4	0	0	0	2	3	0	4
8	0	2	2	2	3	1	2
4	2	3	4	5	6	3	2
1	7	4	8	9	6	3	9
9	8	7	6	8	9	8	9

Cadrul 10

12	8	7	9	5	6	6	3
4	5	3	0	2	1	4	5
1	3	0	0	2	2	3	2
4	3	2	3	5	3	2	4
8	4	0	0	0	2	3	2
4	3	0	2	2	2	3	2
1	7	2	3	4	5	6	9
9	8	7	6	8	9	8	9

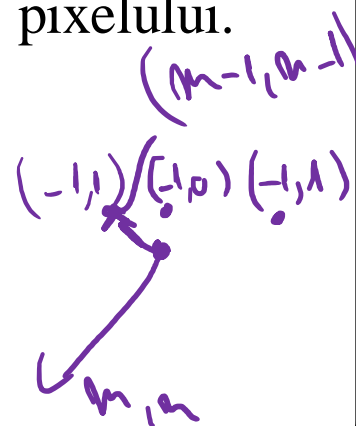
Cadrul 11

Codarea predictivă adaptivă

- se bazează pe adaptarea codării la caracteristicile mișcării din imagine.
- pixelii sunt clasificați ca aparținând:
 - unei zone staționare (CS),
 - în mișcare înceată/moderată (CM), sau
 - în mișcare rapidă (CR).
- Clasificarea se bazează pe un indice de activitate $\alpha(m,n,i)$, care este suma absolută a diferențelor intercadre dintr-o vecinătate N a pixelilor codați anterior, adică:
- O valoare mare a lui $\alpha(m,n,i)$ indică mișcare în vecinătatea pixelului.

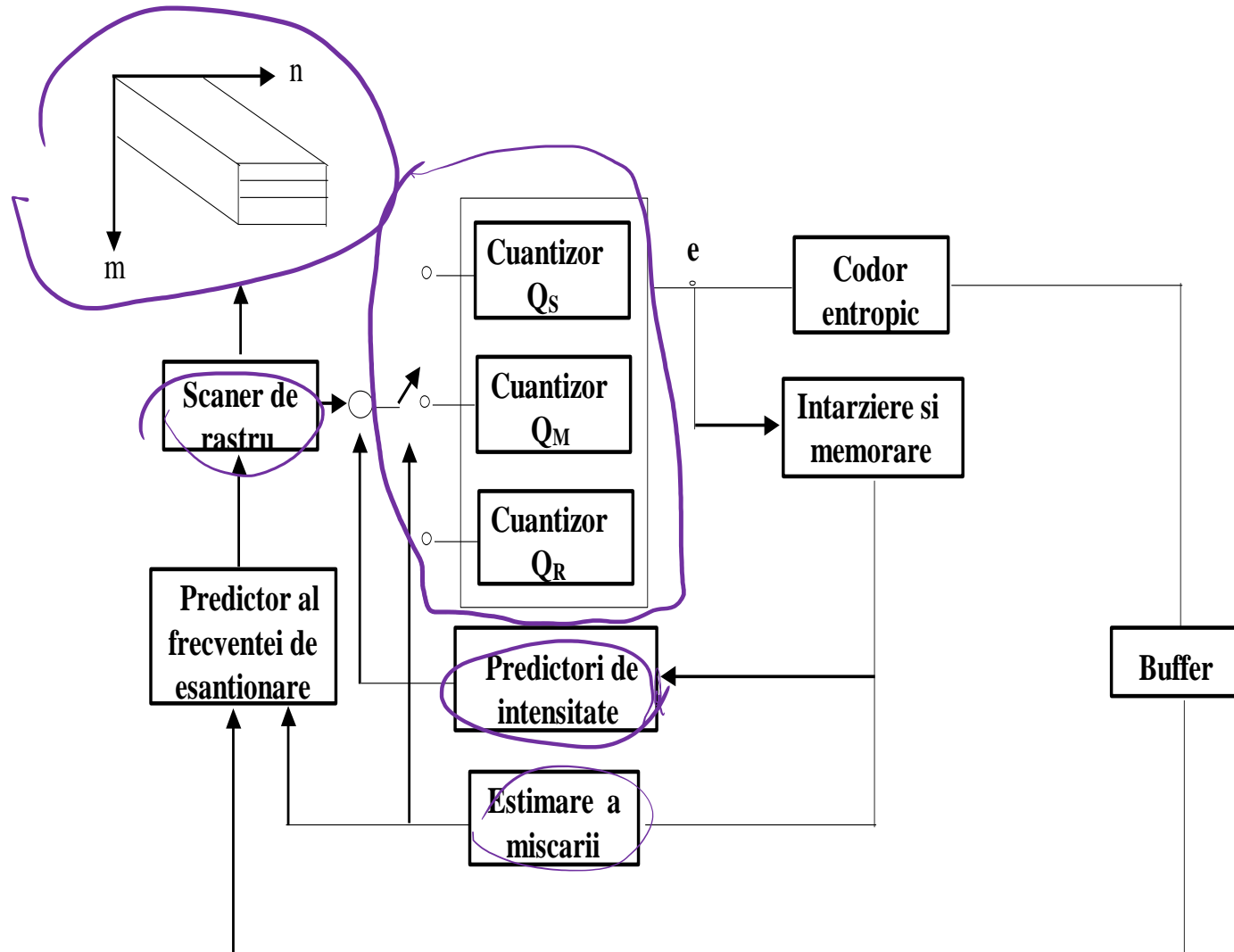
$$\alpha(m,n,i) = \sum_{(x,y) \in N} |u^*(m+x, n+y, i) - u^*(m+x, n+y, i-1)|$$

$$N = \{(0, -s), (-1, -1), (-1, 0), (-1, 1)\}$$



Codarea adaptiva predictiva intercadre

- schema bloc



- Valoarea prezisă a pixelului $u(m,n,i)$ este:

$$\bar{u}^*(m,n,i) = \begin{cases} u^*(m,n,i-1), & (m,n) \in C_s \\ u^*(m-p,n-q,i-1), & (m,n) \in C_M \\ \rho_1^s u^*(m,n-1,i) + \rho_2 u^*(m-1,n,i) - \rho_1 \rho_2 u^*(m-1,n-1,i), & (m,n) \in C_R \end{cases}$$

- unde ρ_1 și ρ_2 sunt coeficienții de corelație cu pixelii vecini pe direcțiile m și n.
- valorile p și q se aleg după estimarea deplasării vecinătății N care dă activitatea minimă

- De exemplu, daca se ignora cadrele de ordin par, $u(m,n,2i)$, $i=1,2,\dots$, prin folosirea compensarii miscarii se obtine:
 - prin repetarea cadrelor:

$$u^\bullet(m,n,2i) = u^\bullet(\underline{m-p}, \underline{n-q}, 2i-1)$$
 - prin interpolarea cadrelor:

$$u^\bullet(m,n,2i) = 1/2 [u^\bullet(\underline{m-p}, \underline{n-q}, 2i-1) + u^\bullet(\underline{m-p'}, \underline{n-q'}, 2i+1)]$$
- unde (p,q) si (p',q') sunt vectorii de deplasare relativ la cadrul precedent si, respectiv, urmator.

Codare predictivă cu compensarea mișcării

- În cazul imaginilor conținând elemente în mișcare, dacă traiectoria mișcării ar putea fi măsurată s-ar putea coda cadrul inițial și informația referitoare la traiectoria mișcării.
- La reproducerea imaginii ar trebui pur și simplu să “deplasăm” pixelii de-a lungul traiectoriilor lor.
- În practică, mișcarea obiectelor dintr-o scenă poate fi aproximată prin deplasarea treptată de la cadru la cadru.
- Din acest motiv, în codarea cu compensarea mișcării se folosește vectorul de deplasare în vederea obținerii predictorului.
- Succesul codării depinde de acuratețea, viteza și imunitatea la perturbații a metodei de estimare a mișcării.

Estimarea si compensarea mișcării

- **estimarea mișcării** - constă în determinarea blocurilor din imagine care au suferit o deplasare de la un cadru la altul, precum și a direcției și vitezei de deplasare a acestora.
- **compensarea mișcării** - procesul de reconstrucție a unei imagini, folosind secvențe de imagine din cadre anterioare, împreună cu informații referitoare la mișcare
- Algoritmii de estimare și compensare a mișcării sunt
 - relativ simpli, dar
 - necesită timpi de calcul relativ importanți.

-
- Dacă o zona de imagine reprezentată de un bloc de pixeli suferă o mișcare relativ lentă (mai mică decât lățimea blocului) și nu este distorsionată semnificativ datorită mișcării, prin schimbarea formei sau orientării, va exista o destul de bună potrivire între acest bloc și blocul de pixeli echivalent din imaginea următoare.
 - Prin suprapunerea blocului în diferite poziții ale noii imagini și prin calcularea diferenței matematice între blocuri pentru fiecare poziție, se poate obține gradul maxim de “potrivire” a blocurilor de pixeli, pentru o anumită poziție a acestuia (caz în care, evident, eroarea matematică determinată are valoarea minimă și se situează sub un anumit “prag de potrivire”).
 - Dacă s-a găsit o soluție acceptabilă de “potrivire”, noul bloc poate fi reprezentat ca o deplasare a celui anterior fără să mai fie necesară transmisia lui ca atare

Block Matching Techniques

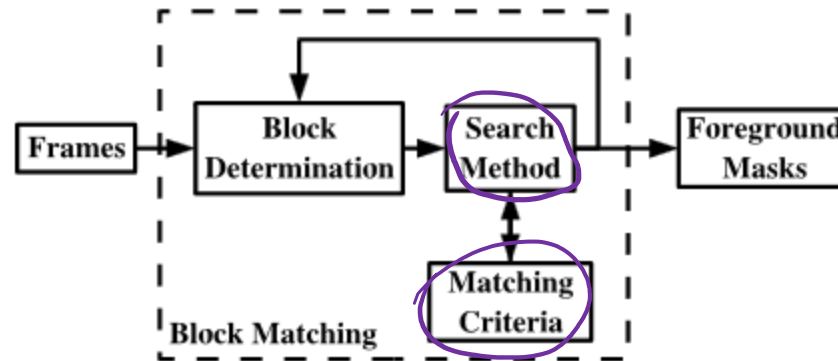


Figure 1: Block Matching Flowchart

Tehnici de potrivire a blocurilor

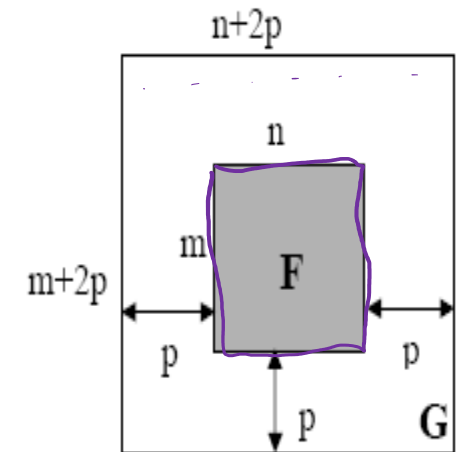
- Valoarea medie absolută a diferenței (MAD)

$$MAD(dx, dy) = \frac{1}{mn} \sum_{i=-n/2}^{n/2} \sum_{j=-m/2}^{m/2} |F(i, j) - G(i + dx, y + dy)|$$

(dx, dy)

- Valoarea medie pătratică a diferenței (MPD)

$$MPD(dx, dy) = \frac{1}{256} \sum_{i=-8}^8 \sum_{j=-8}^8 [F(i, j) - G(i + dx, y + dy)]^2$$



- Funcția de cross-corelație (CCF)

$$CCF(dx, dy) = \frac{\sum_i \sum_j F(i, j) G(i + dx, y + dy)}{\left(\sum_i \sum_j F^2(i, j) \right)^{1/2} \left(\sum_i \sum_j G^2(i + dx, j + dy) \right)^{1/2}}$$

$$RSZ = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2}$$

Algoritmi de estimare a mișcării

- **Tehnici de căutare** a vectorului de deplasare care asigură minimizarea distorsiunilor între un cadru de referință și cadrul curent.
 - Aceste tehnici sunt utile mai ales în cazul secvențelor de imagini în care deplasarea este constantă pentru blocuri întregi de pixeli. Pentru estimarea mișcării intercadre, căutarea se face de obicei pe ferestre de 5x5 pixeli.
- **Tehnicile diferențiale** se bazează pe faptul că valoarea nivelului de gri a pixelului rămîne constantă de-a lungul traiectoriei de mișcare

Tehnicile diferențiale

$$u((x(t), y(t), t)) = \text{const.}$$

$x(t), y(t)$ - *traiectoria de miscare*

$$\Rightarrow: \frac{\partial u}{\partial t} + v_1 \frac{\partial u}{\partial x} + v_2 \frac{\partial u}{\partial y} = 0, \quad (x, y) \in R$$

unde:

$$v_1 = \frac{\partial x(t)}{\partial t},$$

$$v_2 = \frac{\partial y(t)}{\partial t}$$

= cele doua componente ale vitezei;

R = multimea pixelilor in miscare care au aceeasi traiectorie

• **Vectorul de deplasare d** poate fi estimat ca:


$$d = \int_{t_0}^{t_0 + \tau} v dt \approx v \tau$$

unde vectorul viteza v se obtine prin minimizarea: $J = \iint_R \left[\frac{\partial u}{\partial t} + v_1 \frac{\partial u}{\partial x} + v_2 \frac{\partial u}{\partial y} \right]^2 dx dy$

Problema

În tabelul de mai jos este reprezentat cadrul 12 dintr-o secvență video, în care pixelii sunt reprezentați prin luminanțele lor, în domeniul $\{0,1,...,255\}$. Rata cadrelor secvenței video este de 100 cadre/s. Blocul colorat cu gri din cadrul 12 reprezintă un obiect rigid, care se deplasează cu viteză constantă de-a lungul secvenței, vectorul de deplasare al acestui bloc fiind specificat prin vitezele sale orizontală și verticală: 0.2 pixeli/ms pe orizontală și -0.1 pixeli/ms pe verticală.

- Dacă se consideră că variația luminanței obiectului între cadrele 12 și 20 ale secvenței video este nulă, reprezentați obiectul rigid (conținut de blocul hașurat cu gri în cadrul 12) în cadrul 14 al secvenței video.
- Cum s-ar modifica apariția acestui obiect rigid în cadrul 14 dacă luminozitatea medie a scenei se reduce liniar cu un factor de 0.5 la fiecare 10 cadre (la 10 cadre, luminozitatea în fiecare punct al scenei scade la jumătate)?



	0	1	2	3	4	5	6	7
0	5	5	5	5	5	5	50	50
1	5	5	5	23	21	20	52	50
2	50	5	5	22	24	65	53	72
3	150	25	35	35	50	75	55	70
4	150	50	25	75	70	85	130	122
5	152	175	175	75	80	90	125	140
6	250	170	170	180	124	95	124	110
7	250	170	170	185	110	120	123	120

$dx \ dy$
(4, -2)

x, y

a) Cadrul 12 \rightarrow cadrul 14 = 2 cadre

100 cadre/s

1s 100 cadre

x 2 cadre

$$x = \frac{1 \cdot 2}{100} = 0,02 \text{ s} = 20 \text{ ms}$$

deplasare pe x $\rightarrow dx = 0,2 \frac{\text{Pixeli}}{\text{ms}} \cdot 20 \text{ ms} = 4 \text{ pixeli}$

pe y $\rightarrow dy = -0,1 \frac{\text{Pixeli}}{\text{ms}} \cdot 20 \text{ ms} = -2 \text{ pixeli}$

			150	25	35	25
			150	50	25	75
			152	175	175	25
			250	170	170	180

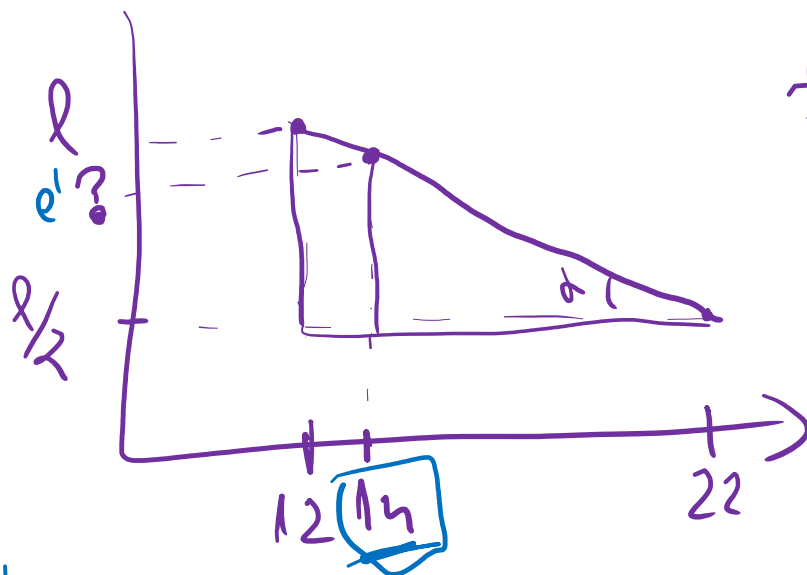
$$(0,3) \xrightarrow{(dx, dy)} (4,1)$$

(4, -2)

b) 10 cadre $0,5 \cdot l$
 2 cadre x

$$x = \frac{0,5 \cdot 2}{10} = 0,1 \Rightarrow \text{Scade cu } 0,1 \Rightarrow 0,9l$$

135	23	32	32
135	45	23	68
137	158	158	68
225	153	153	162

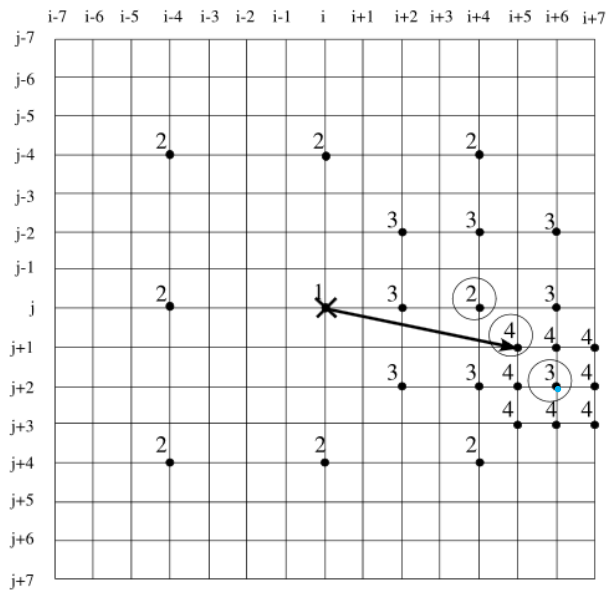


$$\text{tg} \alpha = \frac{l - l/2}{22 - 12} = \frac{l' - l/2}{22 - 14}$$

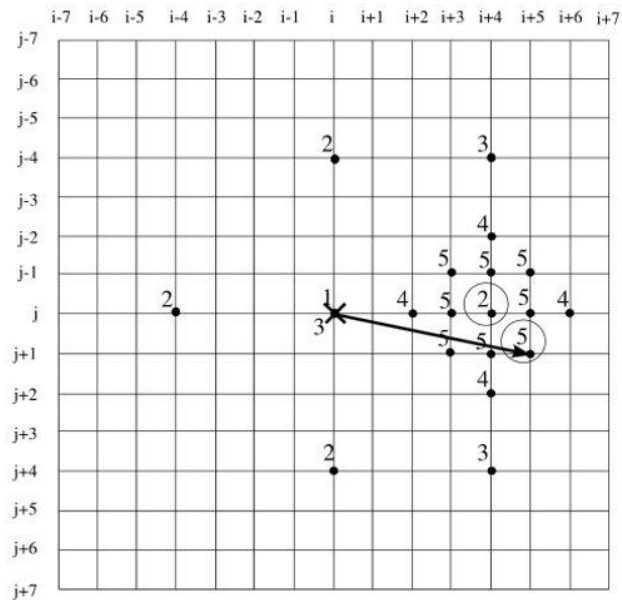
$$\Rightarrow l' - l/2 = \frac{(l - l/2)(22 - 14)}{(22 - 12)}$$

$$l' = \frac{(l - l/2) \cdot 8}{10} + \frac{l}{2} =$$

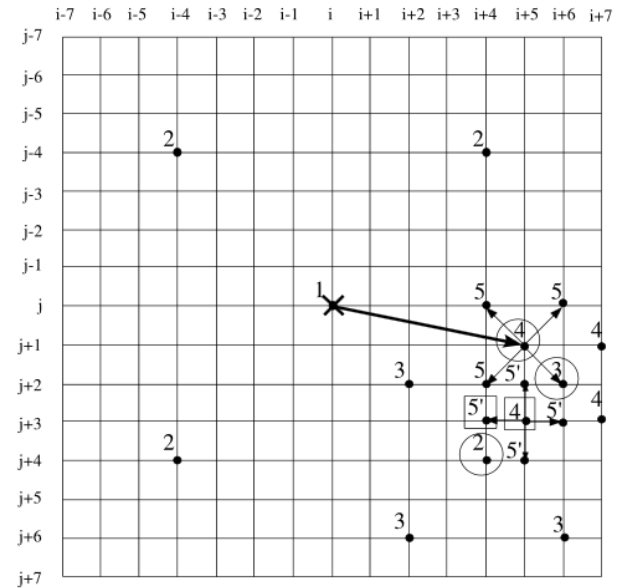
$$l' = (l - l/2) \cdot 0,8 + 0,5 \cdot l = 0,9 \cdot l$$



3: Example of TSS with step size = 4; each num

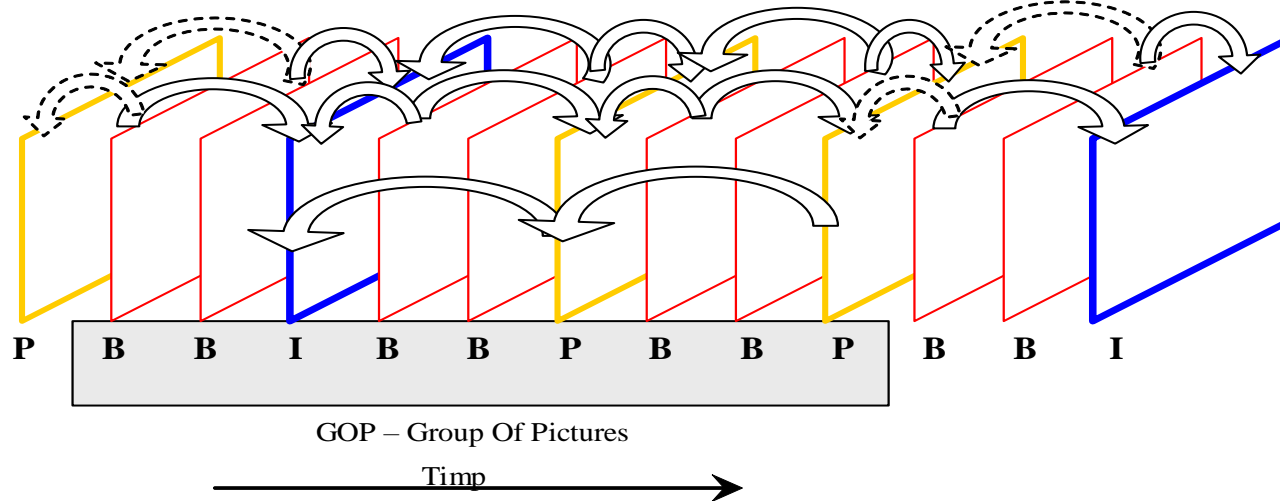


5: Example of 2D-logarithmic search with step size = 4; each number represents



4: Example of cross search with step size = 4; each number represents

Tipuri de cadre și predicția folosită



- **Cadre I**

- cadre ce se codează **fără a folosi tehnica de predicție**.
- se codează asemănător cu o imagine statică JPEG

- **Cadre P**

- care sunt **imagini precise** și folosesc predicția cu salt înainte față de un cadru de referință ce poate fi un cadru I, sau P

- **Cadre B**

- cadre interpolate, folosind două cadre de referință de tip I sau P

2.3 Matching Criterion

Given an $n \times n$ block, a matching criteria, $M(p, q)$, measures the dissimilarity of a block in the current frame, I_c , and a block in the reference frame, I_r , shifted by (p, q) . These criteria can be characterized by $M(p, q) = \sum_{i=1}^n \sum_{j=1}^n \phi(e)$, where $e = I_c(i, j) - I_r(i + p, j + q)$ and $\phi(e)$ is the criteria function. Figure 6 shows the criteria functions for a given e . We examine four matching criteria which are also known as error or matching functions.

SAD The sum of the absolute values of the differences in the two blocks:

$$M(p, q) = \sum_{i=1}^n \sum_{j=1}^n |I_c(i, j) - I_r(i + p, j + q)|$$

MAD The mean of the absolute values of the differences in the two blocks:

$$M(p, q) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |I_c(i, j) - I_r(i + p, j + q)|$$

MSD The mean of the square of the differences in the two blocks:

$$M(p, q) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (I_c(i, j) - I_r(i + p, j + q))^2$$

MPC The sum of the non-matching pixels in the two blocks, a match is determined by the absolute value of the difference being less than a threshold, t_{MPC} .

$$M(p, q) = \sum_{i=1}^n \sum_{j=1}^n D(I_c(i, j), I_r(i + p, j + q))$$

$$D(a, b) = \begin{cases} 0 & \text{if } |a - b| \leq t_{MPC} \\ 1 & \text{otherwise} \end{cases}$$

SAD and MAD only differ by a constant in the case of fixed size blocks and can be used interchangeably in our comparison. Practically, SAD is faster due to the removal of the divide operation. While MAD incorporates large differences, MSD penalizes more a block with one or more large differences. MPC on the other hand equally weights any difference above a threshold.