



Technical University of Cluj - Napoca
Computer Science Department

Interactiune Om-Calculator

Curs 5

**Detectia miscarii in imagini
(flux optic si camp de miscare).**

Detectia prezentei (background subtraction).



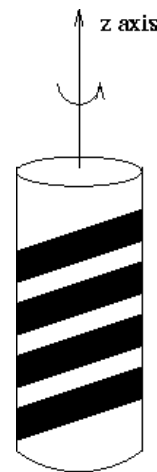
Campul de miscare si fluxul optic [1]

Camp de miscare := setul vectorilor (vitezelor) de miscare ale punctelor din imagine (2D) induse de miscarea relativa dintre scena si camera

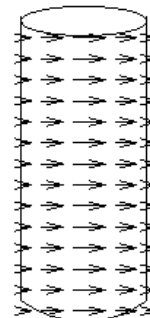
- nu este masurabil direct din imagine !

Fluxul optic := miscarea aparenta a paternurilor de intensitate din imagine

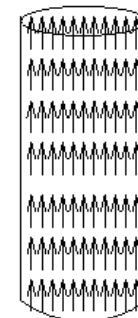
- se poate masura direct din imagine
- este o aproximare a campului de miscare cu o rata de eroare mica in puncte cu gradient mare (daca directia gradientului si directia miscarii coincid – problema de apertura)



Barber's pole



Motion field



Optical flow



Fluxul optic

Ecuatia constantei intensitatii imaginii (in timp) - constrangerea Horn-Schunk:

$$(\nabla E)^T v + E_t = 0$$

Unde:

$$\nabla E = \begin{bmatrix} \frac{\partial E}{\partial x} \\ \frac{\partial E}{\partial y} \end{bmatrix} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} \quad \text{- gradientul imaginii in punctul curent studiat}$$

$$E_t = E(t) - E(t-1) \quad \text{- derivata temporală a intensității imaginii în punctul curent studiat}$$

E – intensitatea imaginii în punctul curent studiat

$$v = \begin{bmatrix} v_x \\ v_y \end{bmatrix} \quad \text{- vector de deplasament al poziției punctului curent studiat (între cele două imagini succesive)}$$

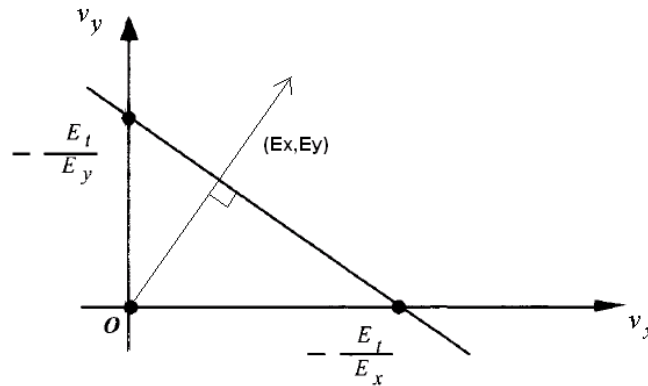
$$v_x \cdot E_x + v_y \cdot E_y = -E_t$$



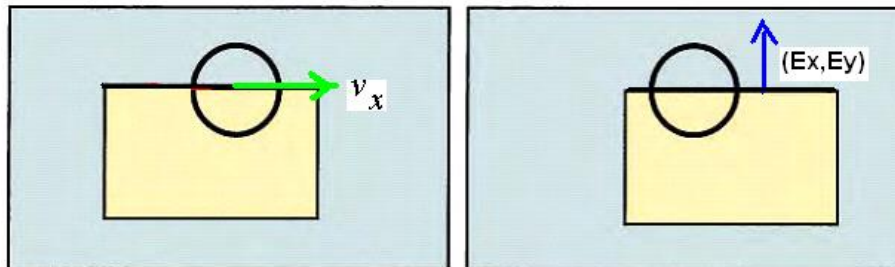
Fluxul optic

Problema de apertura [2]:

Fluxul optic se poate estima corect doar pe o directie paralela cu directia gradientului



⇒ Nu se poate estima daca in zona de interes gradientii sunt paraleli si miscarea este perpendiculara pe directia gradientilor (zona cu o singura muchie dreapta in zona de interes)





Fluxul optic

Estimarea fluxului optic

Majoritatea metodelor de estimare a fluxului optic se bazeaza pe algoritmi iterativi care incerca sa gaseasca pentru fiecare pixel din imagine un vector de deplasamanet care minimizeaza urmatoarea functie reziduala:

$$\epsilon(\mathbf{v}) = \epsilon(\mathbf{v}_x, \mathbf{v}_y) = \sum_{x=u_x-\omega_x}^{u_x+\omega_x} \sum_{y=u_y-\omega_y}^{u_y+\omega_y} (I(x, y) - J(x + \mathbf{v}_x, y + \mathbf{v}_y))^2.$$

Unde:

$I(x,y)$ – punct in imaginea I (t)

$J(x+v_x, y+v_y)$ este noua locatie a aceluiasi punct (trasaturi) in imaginea J de la mometul ($t+\Delta t$).



Fluxul optic

Metoda celor mai mici patrate

- Punct p si o vecinatate de $N \times N$ in jurul punctului curent

$$A \cdot v = B$$

$$A = \begin{bmatrix} E_x(p_1) & E_y(p_1) \\ E_x(p_2) & E_y(p_2) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ E_x(p_{N \times N}) & E_y(p_{N \times N}) \end{bmatrix}$$

$$B = \begin{bmatrix} E_t(p_1) \\ E_t(p_2) \\ \cdot \\ \cdot \\ \cdot \\ E_t(p_{N \times N}) \end{bmatrix}$$

$$v = (A^T A)^{-1} A^T B$$

Imbunatatiri (eliminare zgomote)

- Filtrare gausiana (spatiala a imaginii si temporala la nivel de pixel)
- Weighted least squares:

$$v = (A^T W^2 A)^{-1} A^T W^2 B$$

W – o matrice de ponderi (formă gaussiană)



Fluxul optic

Limitări:

Regiuni în care gradientii sunt nuli (zone cu intensitate uniformă) sau paraleli (exista o singură muchie dreaptă)

⇒ $(A^T A)$ este singulară ⇔ determinantul ei este nul

⇒ soluția la ec:

$$v = (A^T A)^{-1} A^T B$$

nu se poate calcula.

În acest caz din ec: $A \cdot v = B$

se va alege soluția cu **norma minimă**, iar fluxul optic în punctul respectiv se va numi **flux normal**.



Algoritmul Horn-Shunk

- Metodă iterativă

- Parametrii de intrare: n_0 – numărul maxim de iterații; λ – ponderea de corecție

Se face o primă parcurgere a imaginii. Pentru fiecare pixel $p(x,y)$ se calculează $E_x(p)$, $E_y(p)$, $E_t(p)$ și se inițializează v_x și v_y cu 0 (este necesară alocarea unor matrice de dimensiunea imaginii pentru stocarea acestor valori).

1. Se aleg valorile pt. λ (ex. $\lambda = 10$) și n_0 (ex. $n_0 = 8$).

2. Pentru $n = 1 .. n_0$:

Se parcurge imaginea. Pentru fiecare pixel $p(x,y)$ se calculează valorile medii ale v_x și v_y (din vecinii de pe direcțiile cardinale):

$$\bar{v}_x = \frac{1}{4} [v_x(i-1, j) + v_x(i+1, j) + v_x(i, j-1) + v_x(i, j+1)] \quad \bar{v}_y = \frac{1}{4} [v_y(i-1, j) + v_y(i+1, j) + v_y(i, j-1) + v_y(i, j+1)]$$

Se calculează coeficientul de corecție α :

$$\alpha = \frac{E_x \bar{v}_x + E_y \bar{v}_y + E_t}{1 + \lambda(E_x^2 + E_y^2)}$$

Se actualizează (corectează) valorile v_x și v_y :

$$v_x = \bar{v}_x - \alpha \cdot E_x$$

$$v_y = \bar{v}_y - \alpha \cdot E_y$$



Fluxul optic

Algoritmul Horn-Shunk- etapa finala:

Thresholding + non-maxima suppression pe o vecinatate a fiecarui pixel current \Rightarrow eliminarea zgomotelor (vectorii de flux optic foarte mici \approx zgomote)

\Rightarrow Flux optic discret ('sparse' / lb. engleza)



Ilustrarea vectorilor de flux optic estimati prin metoda Horn-Shunk folosind implementarea din OpenCV (vezi L6)



Fluxul optic

Fluxul optic pt. imaginii color [2]

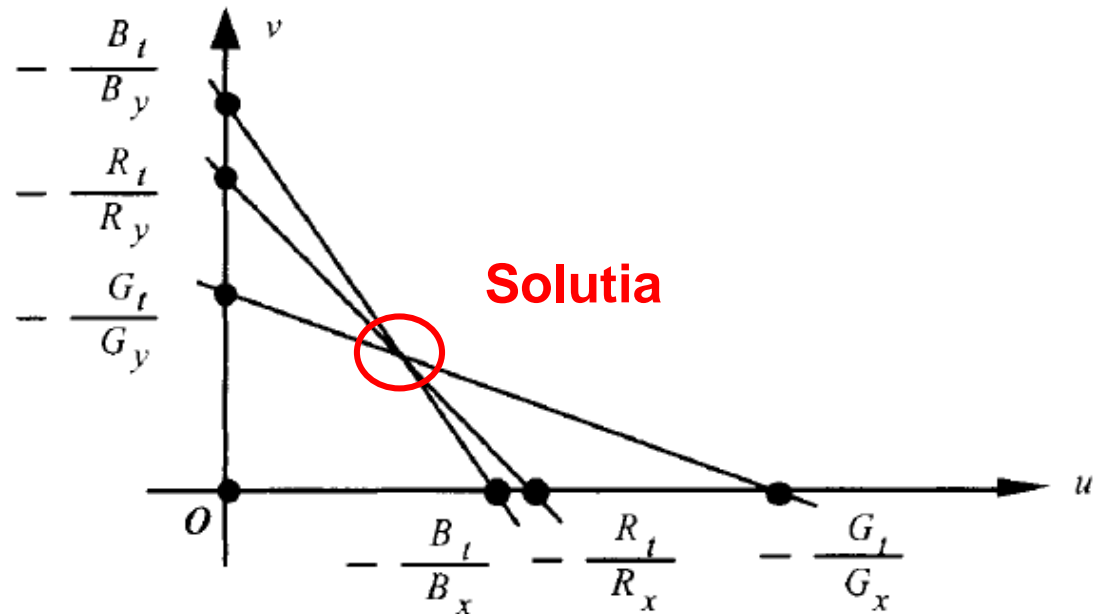
$$C_i(x, y, t) = (R_i(x, y, t), G_i(x, y, t), B_i(x, y, t)), i = 0, 1, 2, \dots$$

Constrangerea Horn-Schunk:

$$R_x \cdot u + R_y \cdot v = -R_t,$$

$$G_x \cdot u + G_y \cdot v = -G_t, \text{ and}$$

$$B_x \cdot u + B_y \cdot v = -B_t$$





Fluxul optic

$$R_x = \frac{\partial R}{\partial x} \quad \text{and} \quad R_y = \frac{\partial R}{\partial y}$$

$$\mathbf{J} \cdot \mathbf{u} = -\mathbf{C}_t$$

$$\mathbf{J} = \begin{pmatrix} R_x & R_y \\ G_x & G_y \\ B_x & B_y \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \text{and} \quad \mathbf{C}_t = \begin{pmatrix} R_t \\ G_t \\ B_t \end{pmatrix}$$

$$(\mathbf{J}^T \cdot \mathbf{J}) \cdot \mathbf{u} = -\mathbf{J}^T \cdot \mathbf{C}_t$$

$$(\mathbf{J}^T \cdot \mathbf{J})^{-1} (\mathbf{J}^T \cdot \mathbf{J}) \cdot \mathbf{u} = -(\mathbf{J}^T \cdot \mathbf{J})^{-1} \mathbf{J}^T \cdot \mathbf{C}_t$$

$$\mathbf{u} = -(\mathbf{J}^T \cdot \mathbf{J})^{-1} \cdot \mathbf{J}^T \cdot \mathbf{C}_t .$$

Nu exista solutie cand matricea $\mathbf{J}^T \mathbf{J}$ este singulara:

$$\det(\mathbf{J}^T \cdot \mathbf{J}) = 0_1$$



Fluxul optic

$$\mathbf{J}^T \cdot \mathbf{J} = \begin{pmatrix} R_x^2 + G_x^2 + B_x^2 & R_x R_y + G_x G_y + B_x B_y \\ R_x R_y + G_x G_y + B_x B_y & R_y^2 + G_y^2 + B_y^2 \end{pmatrix}$$

$$\|\mathbf{J}^T \cdot \mathbf{J}\| = (R_x^2 + G_x^2 + B_x^2)(R_y^2 + G_y^2 + B_y^2) - (R_x R_y + G_x G_y + B_x B_y)^2$$

$$\|\mathbf{J}^T \cdot \mathbf{J}\| = (R_x G_y - R_y G_x)^2 + (R_x B_y - R_y B_x)^2 + (G_x B_y - G_y B_x)^2$$

$$\det(\mathbf{J}^T \cdot \mathbf{J}) = 0 \quad \Leftrightarrow \quad R_x G_y = R_y G_x \text{ and } R_x B_y = R_y B_x \text{ and } G_x B_y = G_y B_x$$

Gradientii de culoare sunt nuli

$$R_x = G_x = B_x = 0 \text{ or}$$

$$R_y = G_y = B_y = 0 \text{ or}$$

$$R_x = R_y = G_x = G_y = 0 \text{ or}$$

$$R_x = R_y = B_x = B_y = 0 \text{ or}$$

$$G_x = G_y = B_x = B_y = 0$$

Gradientii de culoare sunt paraleli

$$R_y \neq 0, G_y \neq 0, \text{ and } B_y \neq 0$$

$$\frac{R_x}{R_y} = \frac{G_x}{G_y} = \frac{B_x}{B_y}$$



Alte metode de estimare:

[4] Simon Baker, Daniel Scharstein, J.P. Lewis, Stefan Roth, Michael J. Black, Richard Szeliski, A Database and Evaluation Methodology for Optical Flow, Int J Comput Vis (2011) 92: 1–31. <http://vision.middlebury.edu/flow/floweval-ijcv2011.pdf>

- Survey cu metode de calcul al fluxului optic + evaluare + afisare flux optic dens in cod de culori

[5] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, Proceedings of Imaging Understanding Workshop, pp. 121-130 (1981).

<http://cseweb.ucsd.edu/classes/sp02/cse252/lucaskanade81.pdf>

[6] Gunnar Farneback, Two-frame motion estimation based on polynomial expansion, Lecture Notes in Computer Science, 2003, (2749), 363-370.

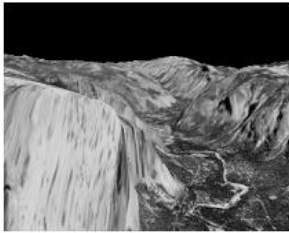
<http://www.diva-portal.org/smash/get/diva2:273847/FULLTEXT01.pdf>

- Calcul flux optic dens, cu implementare in OpenCV (se va integra in lucrarea de laborator cu nr. 7)

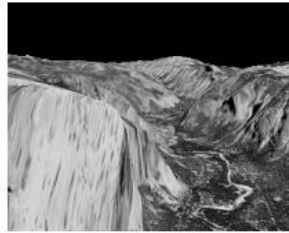


Fluxul optic

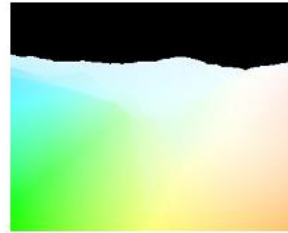
Vizualizarea fluxului optic dens:



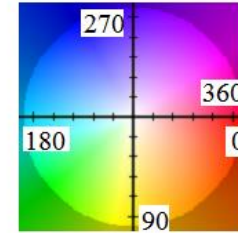
Yosemite frame 0



Yosemite frame 1



Yosemite GT flow



flow color coding

Exemplificare a rezultatului codificarii fluxului de miscare in conformitate cu conventia de culoare Middleburry [4]:
Hue – directia / Saturatia – amplitudinea (similar cu codificarea culorilor in modelul HIS)





Detectia miscarii / prezentei [3]

Modelare background (fundal) \Rightarrow background subtraction

Input: secventa de imagini, camera/ imagine de fundal fixe

Scop: detectia / segmentarea obiectelor in miscare (tranzitorii in scena) care nu fac parte din fundal

Segmentare pixeli $p(x,y)$ ca apartinand unui obiect (foreground):

$$|frame_i - background_i| > Th$$

Probleme:

- Cine este $background_i$? \Rightarrow Imagine a modelului fundalului
- Variatii de iluminare ale scenei / umbre
- Miscari/vibratii ale camerei sau ale scenei (copaci/frunze)
- Schimbari ale geometriei scenei (apar obiecte care devin statice - ex. masini parcate).



Background subtraction

Solutii:

1. Diferenta între frame-uri

(se face pt. fiecare pixel $p(x,y)$ din imagine)

$$| frame_i - frame_{i-1} | > Th$$

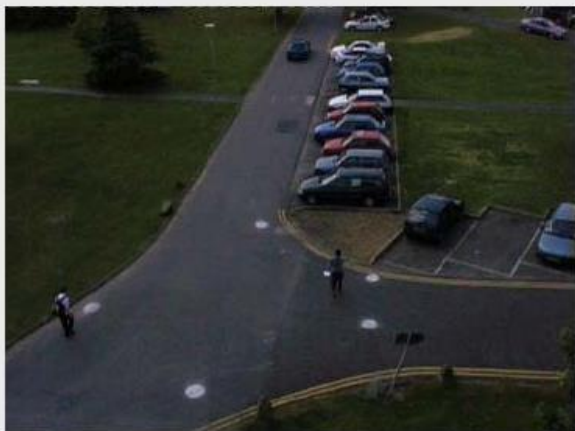
- Daca pt. un pixel $p(x,y)$ este satisfacuta conditia de mai sus el se clasifica ca si pixel de obiect (foreground) in imaginea curenta
- “Background” (fundalul) este frame-ul anterior
- Limitari legate de fps camera si viteza obiectelor
- Sensibilitate mare la valoarea Th



Background subtraction [3]

Frame difference: an example

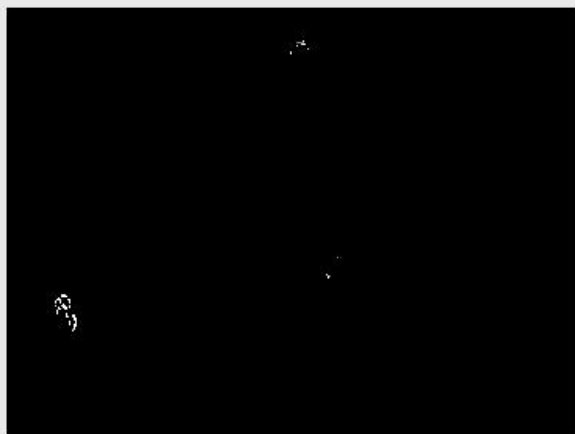
the frame



absolute
difference



threshold:
too high



threshold:
too low





Background subtraction [3]

2. Background : running average:

$$B_{i+1} = \alpha * F_i + (1 - \alpha) * B_i$$

α – rata de invatare (ex. 0.05)

- Nu necesita memorie suplimentara (doar o matrice pt. background / fundal)
- Se considera de obicei istoria ultimelor n frame-uri
- Se poate face mediere cu pondere variabila (frame-urile recente sa aiba o pondere mai mare)
- Detectia pixelilor de obiect (foreground):

pixeli $p(x,y)$ care satisfac conditia : $|F_i - B_i| > Th$



Background subtraction [3]

Problema selectivitatii

In fiecare imagine, fiecare pixel este clasificat ca obiect sau fundal \Rightarrow feedback in constructia modelului pt. fundal:

daca pixelul este clasificat ca si obiect /foreground se va ignora in actualizarea modelului pt. fundal !!!

\Rightarrow Se previne “coruperea” modelului cu pixeli care nu apartin fundalului

2'. Background : running average with selectivity

Actualizarea pixelilor fundal se va face in mod selectiv dupa urmatoarele reguli:

$$B_{i+1}(x, y) = \alpha F_t(x, y) + (1 - \alpha) B_t(x, y) \quad \text{if } F_t(x, y) \text{ background}$$

$$B_{i+1}(x, y) = B_t(x, y) \quad \text{if } F_t(x, y) \text{ foreground}$$

Abordarea selectiva se poate aplica la toate metodele de background

subtraction !



Background subtraction [3]

3. Folosirea histogramelor de culoare (pixel)

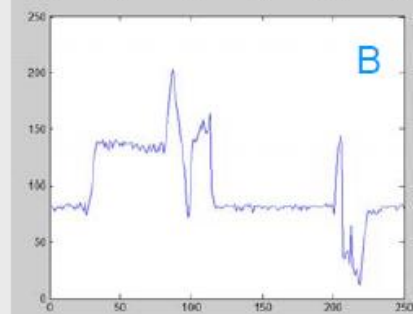
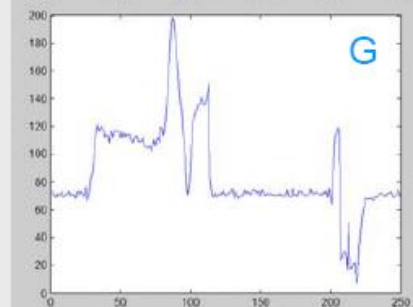
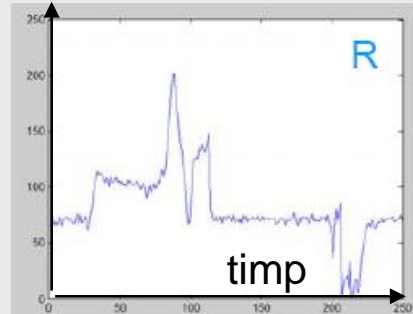
● Example:



pixel location

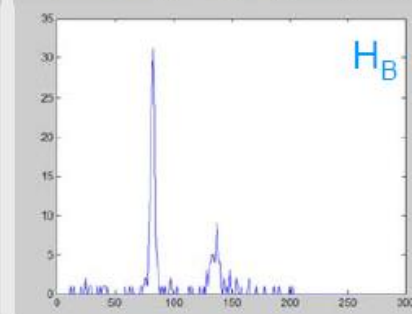
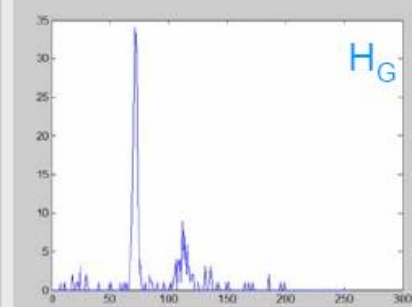
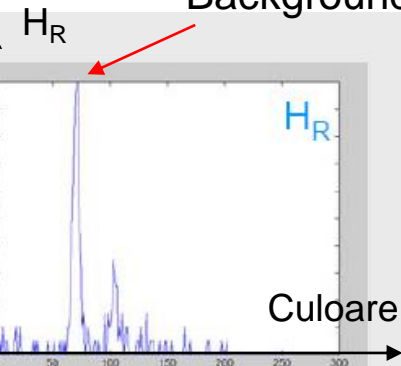
Variatia temporală a
culturilor în punctul selectat

Culoare



sequence of
pixel values

Background



histograms



Background subtraction [3]

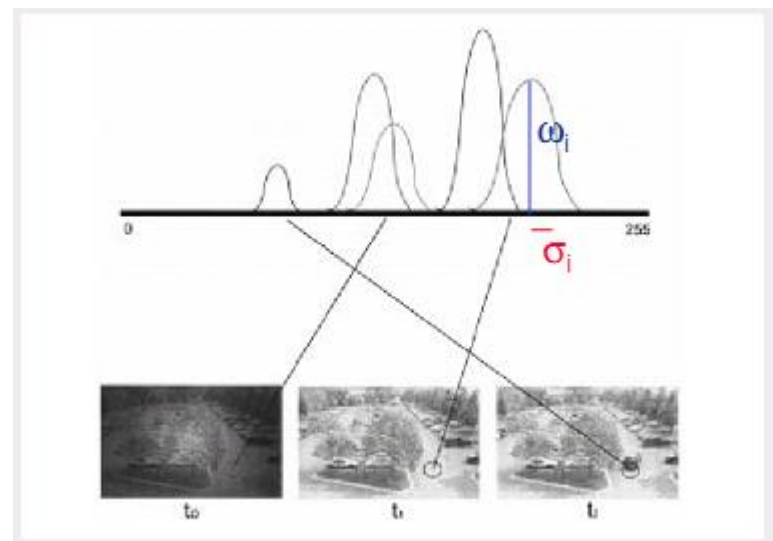
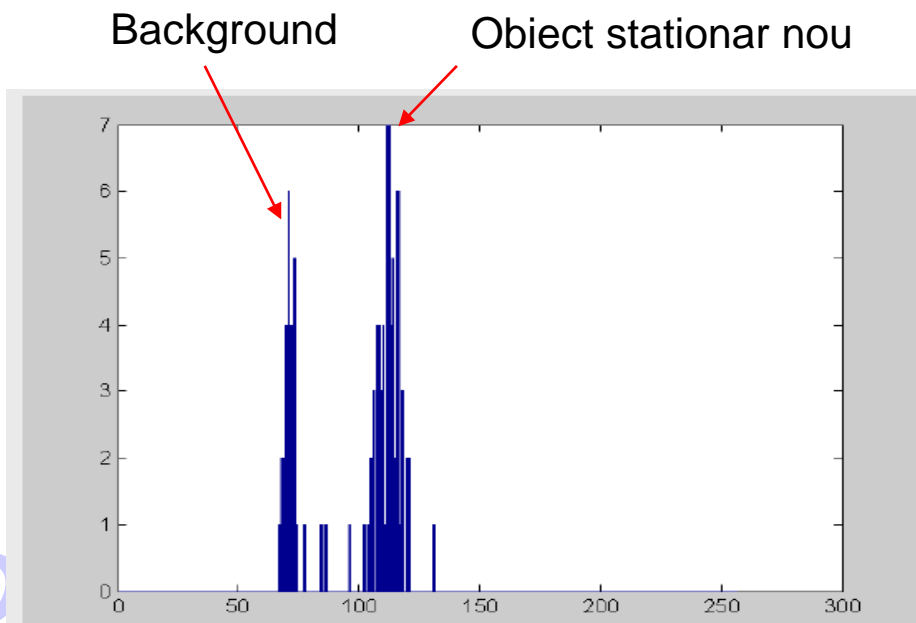
3. Folosirea histogramelor de culoare pt. fiecare pixel

Detectia pixelului de obiect (foreground):

Pt. fiecare pixel $p(x,y)$: $|F_i - B_i(\text{poz. max. histograma})| > Th$

Dezavantaje:

- Consum marit de memorie: histograme de culoare pt. fiecare pixel
- Probleme: cu histograme multimodale ale background-ului (obiecte stationare care se schimba)





Background subtraction [3]

4. Running Gaussian average

- Potrivirea unei distributii gaussiene peste “modul” corespunzator fundalului \Rightarrow FDP a fundalului (μ, σ)
- Actualizarea FDP a fundalului (pt. fiecare pixel):

$$\begin{aligned}\mu_{t+1} &= \alpha F_t + (1 - \alpha)\mu_t \\ \sigma_{t+1}^2 &= \alpha (F_t - \mu_t)^2 + (1 - \alpha)\sigma_t^2\end{aligned}$$

- Pt. fiecare pixel: daca $|F_t - \mu_t| > Th$ ($Th \sim k\sigma$)
atunci este pixel de foreground (obiect)

Metoda nu este potrivita pt. fundal (pixeli) cu histograma multimodala!



Background subtraction

Exemple



Secventa originala



Background subtraction

Exemple



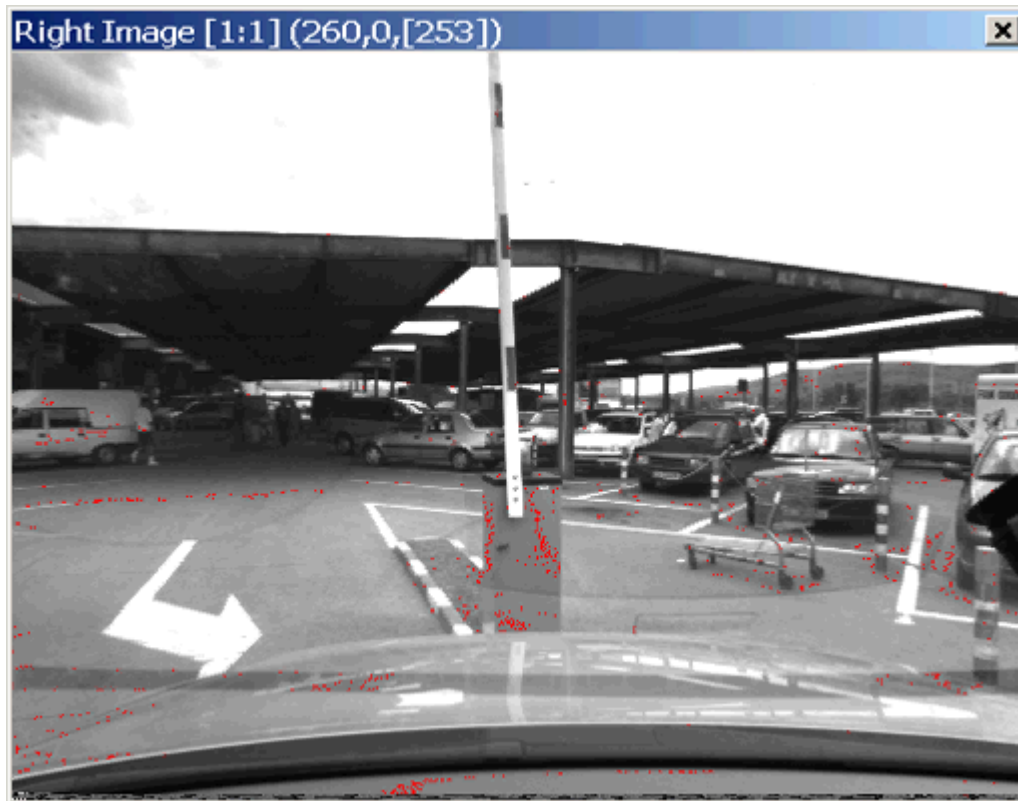
1. Diferenta intre frame-uri

($Th = 15$)



Background subtraction

Exemple



2. Background – running average (fara selectivitate)

$$(\alpha = 0.01, Th = 15)$$



Background subtraction

Exemple



3. Background - maximul histogramei

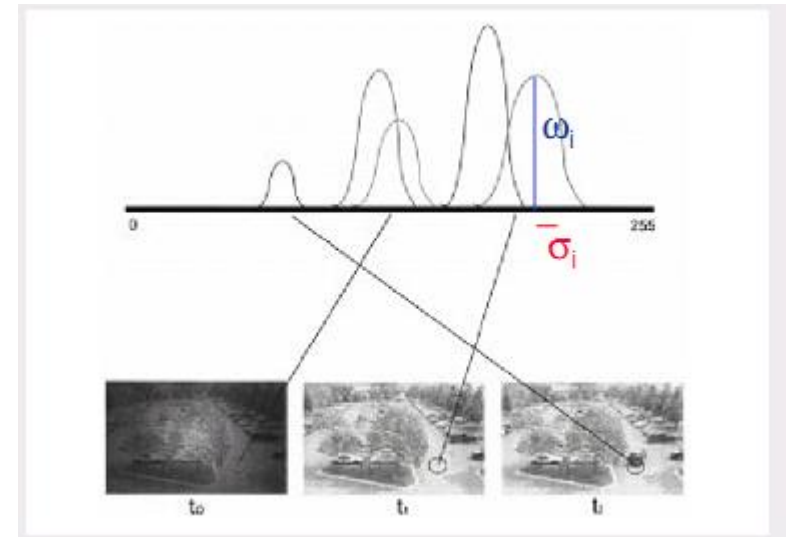
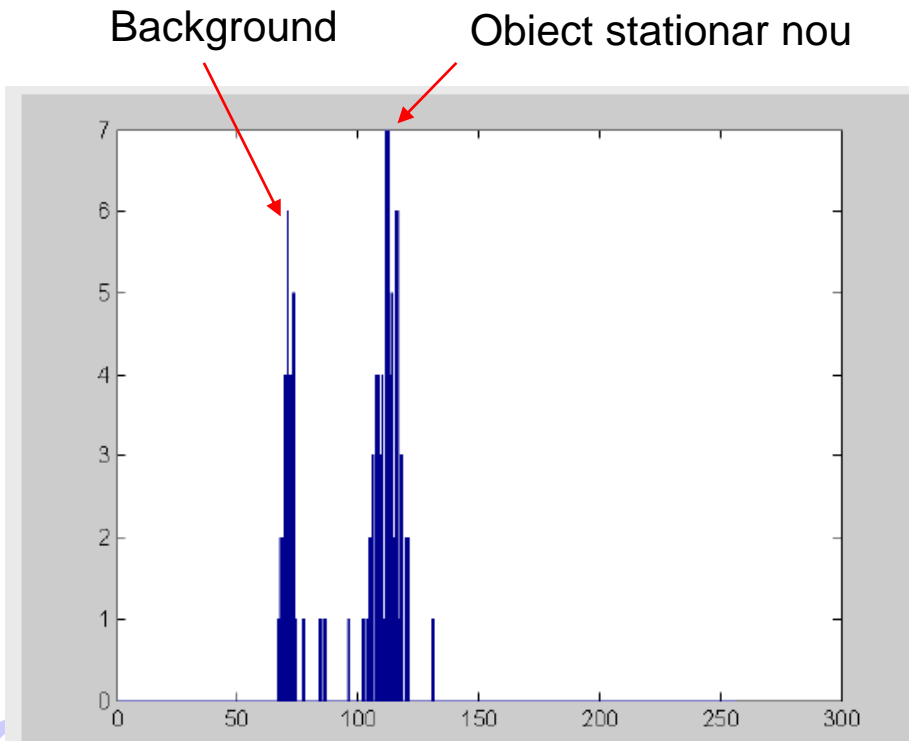
($Th = 15$)



Background subtraction

Solutii pt. Histograme multimodale ale fundalului [3]

- Mixture of Gaussians
- Kernel Density Estimators
- Mean-shift based estimation
- Combined estimation and propagation





Background subtraction

Postprocesare

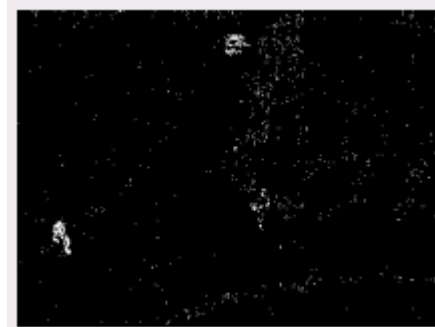
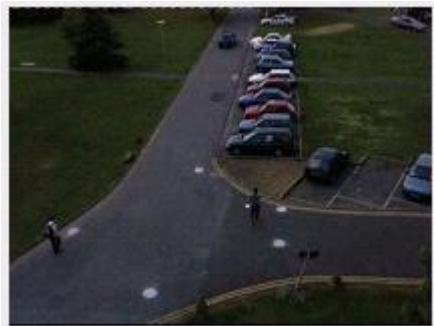
Filtrare “zgomote”

- filtre morfologice (eliminare puncte de zgomot si umplere goluri)
- etichetare obiecte si eliminare obiecte cu arie mica

Clustering (grupare puncte)

- criterii de vecinatate / densitate

Asociere de etichete / obiecte





Background subtraction

Aplicatii

Segmentare si detectie obiecte in miscare (cu prezenta tranzitorie in imagine):

- supraveghere trafic auto \Rightarrow masuratori suplimentare
- sisteme de supraveghere/securitate (on-line sau in asociere cu un dispozitiv DVR)
 - imbunatatiri la dispozitivele DVR existente

Constrangeri: camera fixa / fundal imobil relativ la camera

Limitari / probleme:

- variatii bruste de iluminare
- umbre \Rightarrow solutii particulare



Referinte:

- [1] E. Trucco, A. Verri, “Introductory Techniques for 3-D Computer Vision”, *Prentice Hall, 1998*.
- [2] A. Koschan, M. Abidi, Digital Color Image Processing, Wiley & Sons, 2008. - cap 10, pag 254 -259
- [3] M. Piccardi - Background subtraction techniques: a review, Computer Vision Research Group (CVRG), University of Technology Sydney (UTS)
- [4] Simon Baker, Daniel Scharstein, J.P. Lewis, Stefan Roth, Michael J. Black, Richard Szeliski, A Database and Evaluation Methodology for Optical Flow, *Int J Comput Vis* (2011) 92: 1–31. <http://vision.middlebury.edu/flow/floweval-ijcv2011.pdf>
- [5] B. D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, *Proceedings of Imaging Understanding Workshop*, pp. 121-130 (1981). <http://cseweb.ucsd.edu/classes/sp02/cse252/lucaskanade81.pdf>
- [6] Gunnar Farneback, Two-frame motion estimation based on polynomial expansion, *Lecture Notes in Computer Science*, 2003, (2749), 363-370. <http://www.diva-portal.org/smash/get/diva2:273847/FULLTEXT01.pdf>