Universitatea Transilvania din Braşov
Laboratorul de Vedere Artificială Robustă şi Control

Sisteme de Vedere Artificială

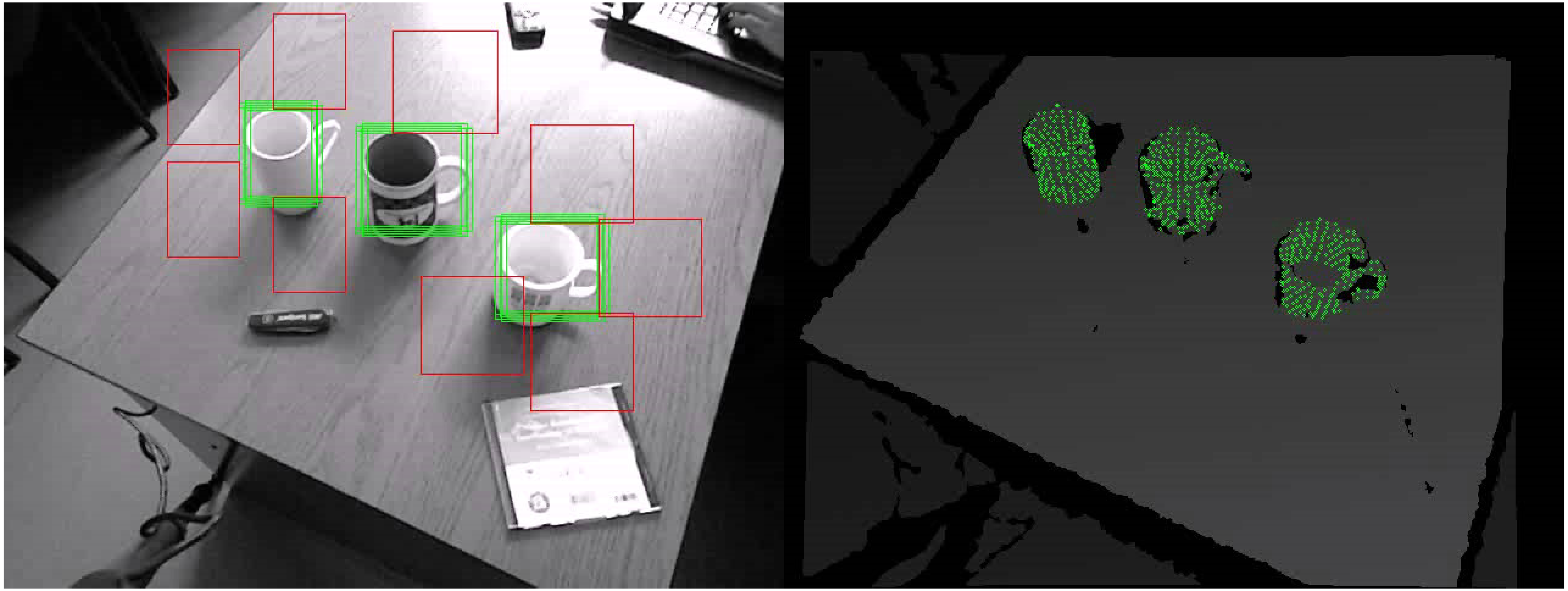
Urmărirea formelor

Sorin M. Grigorescu

Cuprins

- Fluxul optic
- Modele dinamice în urmărirea formelor

Exemplu



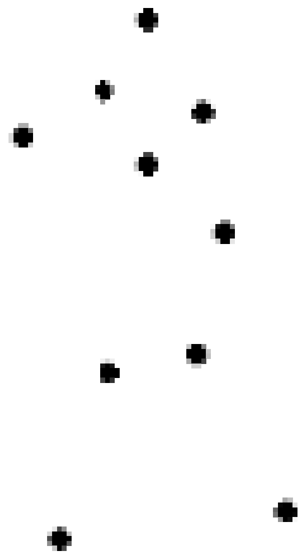
Urmărirea formelor: definiţii

Definiții

- Într-un mare număr de aplicaţii, informaţiile de intrare ce trebuiesc procesate sunt formate din secvenţe video şi nu din imagini singulare.
- În acest caz, obiectivul sistemului de vedere artificială este de a urmări mişcarea obiectului (sau a obiectelor) în secveţa de imagini video.
- Această operaţie este compusă din două componente principale: ***măsurarea*** **şi** ***modelare*****.**

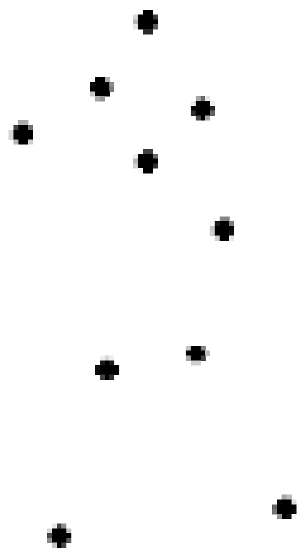
Mișcarea și organizarea datelor

- **Chiar și dintr-un număr redus de date, se pot extrage informații relevante**



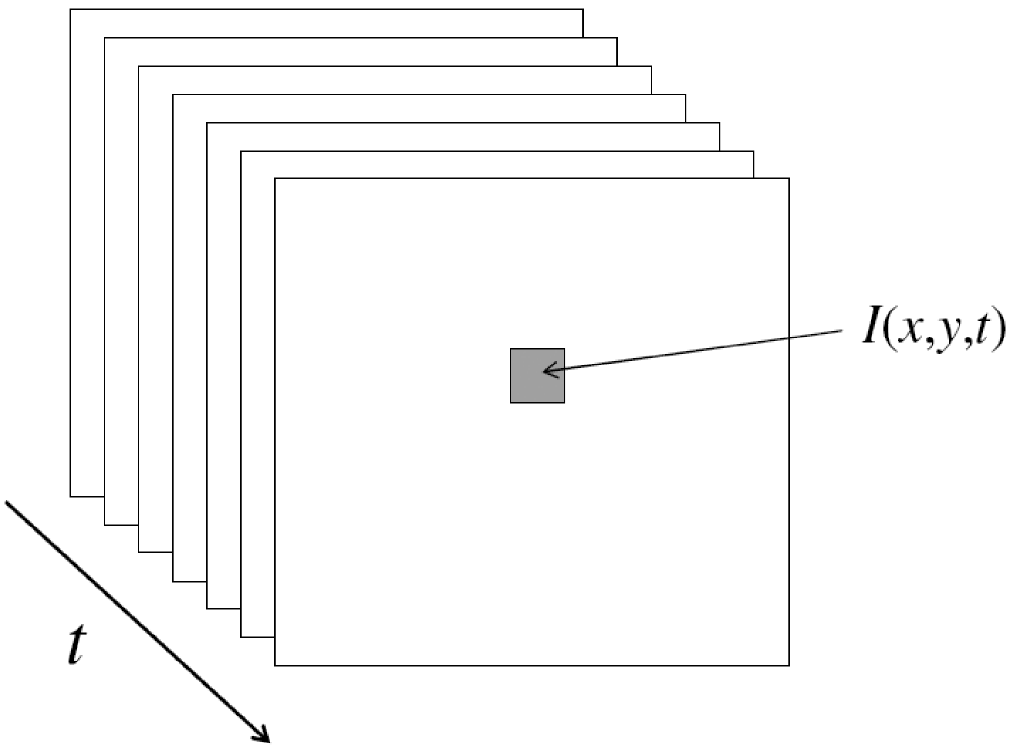
Mișcarea și organizarea datelor

- **Chiar și dintr-un număr redus de date, se pot extrage informații relevante**



Reprezentarea datelor

- O secvență video este reprezentată de un număr de imagini achiziționare de-a lungul unei perioade de timp.
- Informația vizuală este reprezentată de date spațiale (x,y) și timp t.

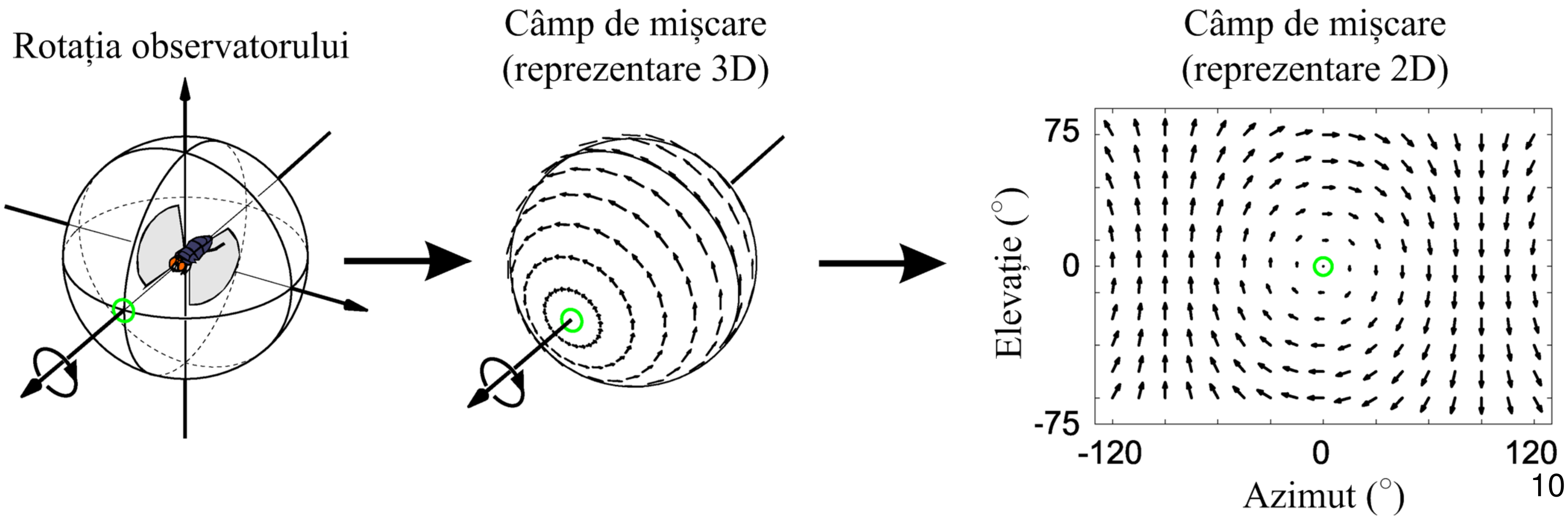


Aplicații

- Analiza traficului pe autostrăzi
- Numărarea persoanelor (în metrou)
- Analiza testelor de accident (crash tests)
- Detecția traiectoriei pietonilor în sistemele de asistență auto
- Etc.

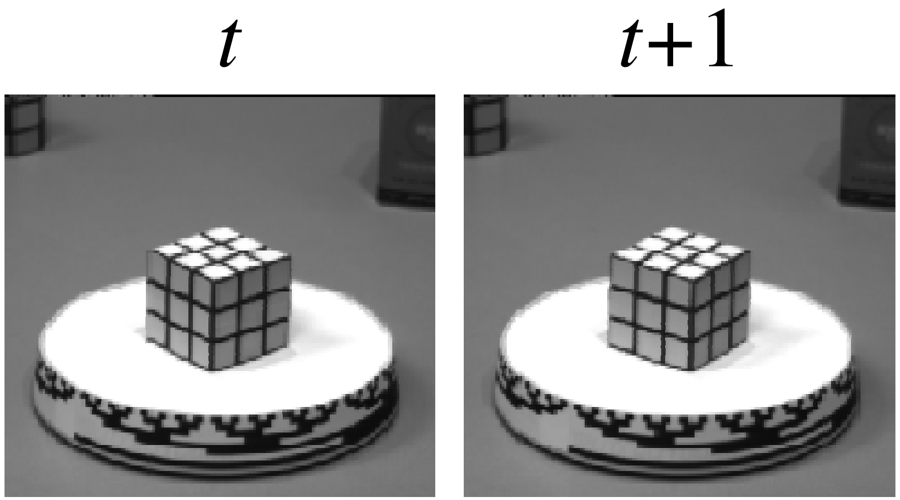
Câmp de mişcare

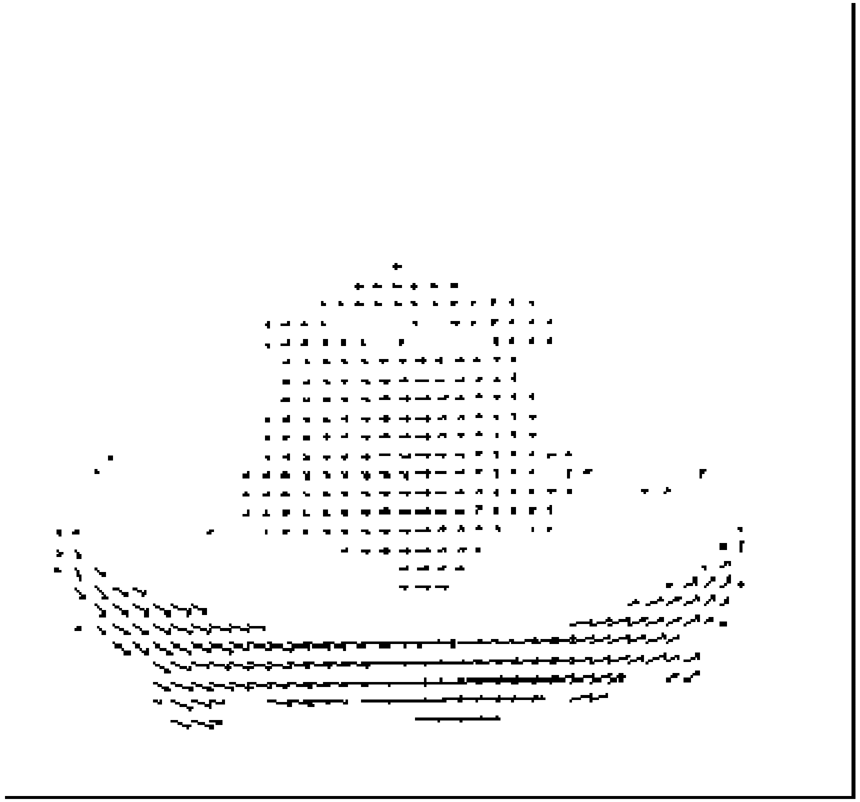
- ***Câmpul de mişcare*** **reprezintă deplasarea relativă într-o scenă, sau secvenţă video, a obiectelor, suprafeţelor şi cantelor, cauzată de mişcarea relativă dintre un observator (cameră) şi scena vizualizată.**
- **Direcţia şi intensitatea (magnitudinea) fluxului optic la orice locaţie sunt indicate de direcţia şi lungimea fiecărui vector.**



Câmp de mișcare

- Câmpul de mișcare reprezintă proiecția mișcării 3D în imagine



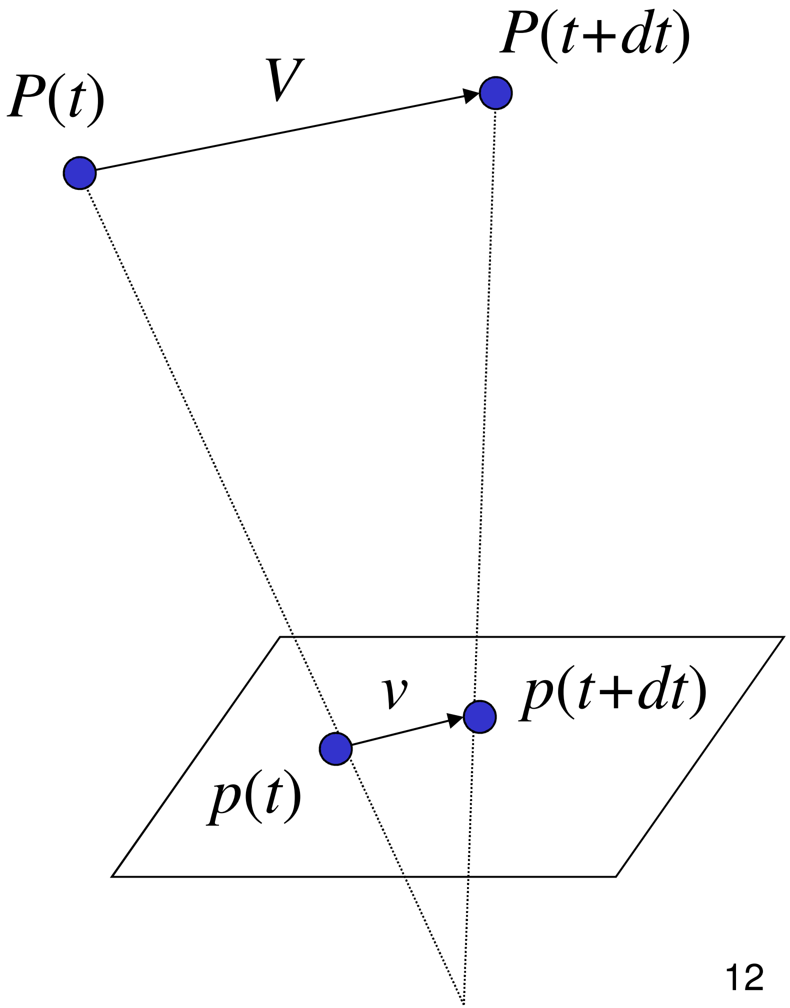


Câmp de mișcare

- $P(t)$ **este un punct 3D în mișcare**
- **Viteza unui punct** $V = \dfrac{dP}{dt}$
- $p(t) = [x(t), y(t)]$ **reprezintă proiecția lui** P **în imagine**
- **Viteza relativă** v **în imagine dată de componentele**

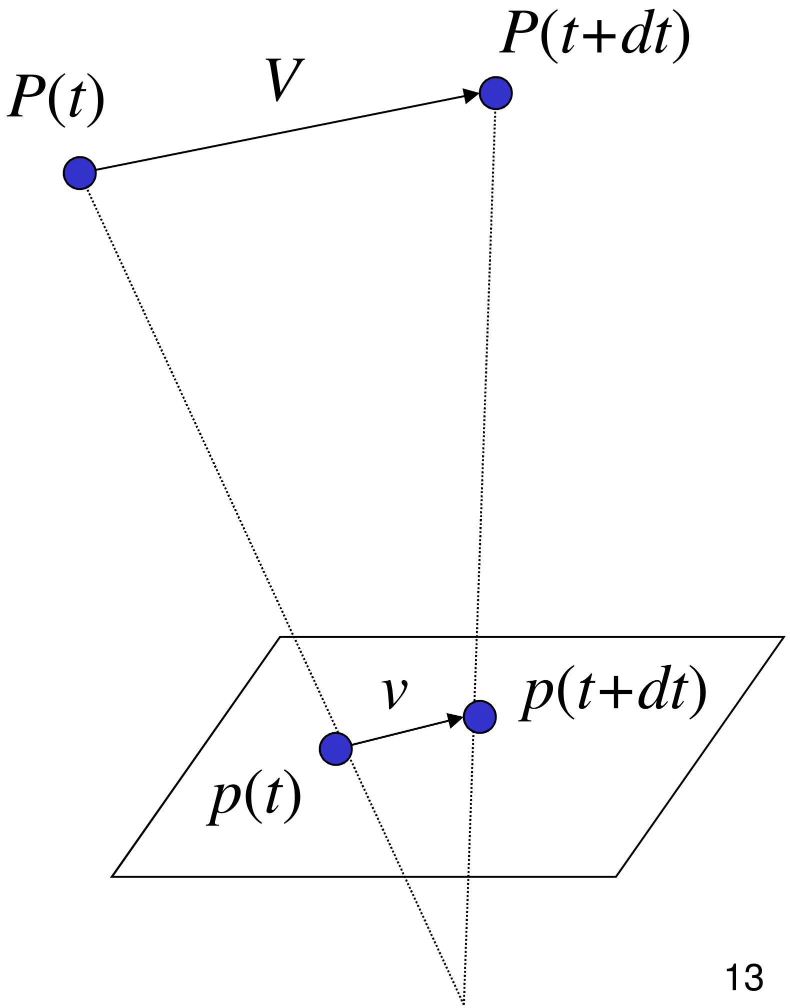
$$v_x = \frac{dx}{dt} \qquad v_y = \frac{dy}{dt}$$

- v_x **și** v_y **reprezintă** ***câmpul de mișcare*** **în imagine**



Câmp de mișcare

$$V = (V_x, V_y, V_Z) \qquad p = f\frac{P}{Z}$$



- **Pentru a determina viteza** v**, se calculează derivata lui** p **față de** t

$$v = f\frac{ZV - V_z P}{Z^2}$$

$$v_x = \frac{fV_x - V_z x}{Z} \qquad v_y = \frac{fV_y - V_z y}{Z}$$

- **Mișcarea în imagine este o funcție dependentă de mișcarea 3D (**V**), și de distanța 3D (**Z**) față de cameră**

Câmp de mișcare + mișcarea camerei

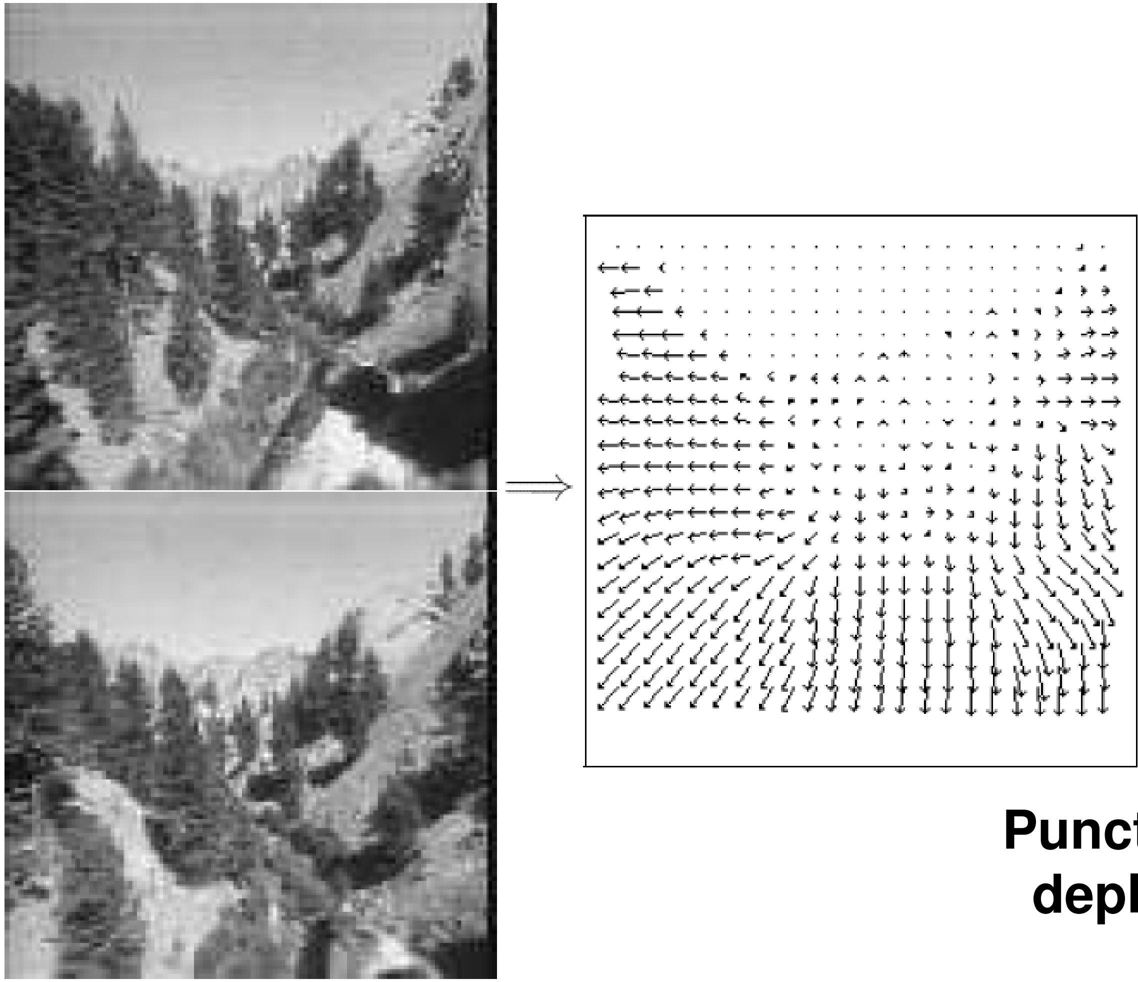
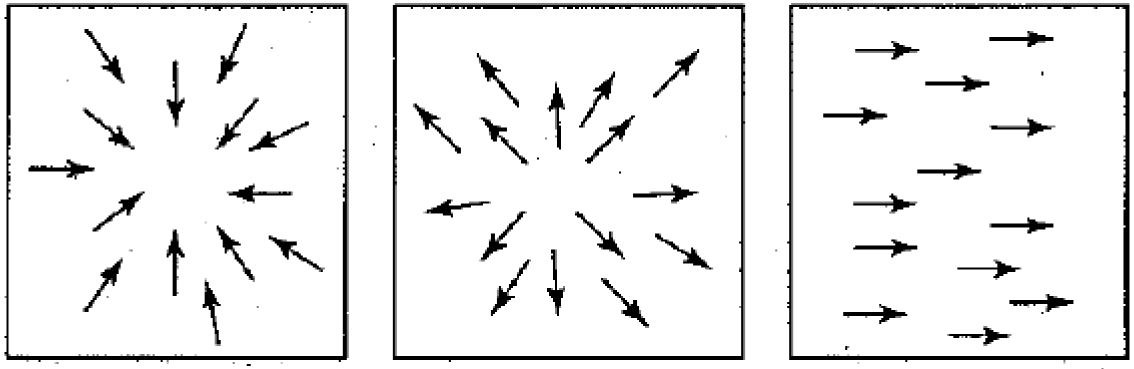


Figure 1.2: Two images taken from a helicopter flying through a canyon and the computed optical flow field.

Lungimea vectorilor este invers proporțională cu distanța Z a punctelor 3D

Punctele apropiate de cameră se deplasează mai rapid în planul imaginii

Câmp de mișcare + mișcarea camerei



Zoom out | Zoom in | Translație către dreapta

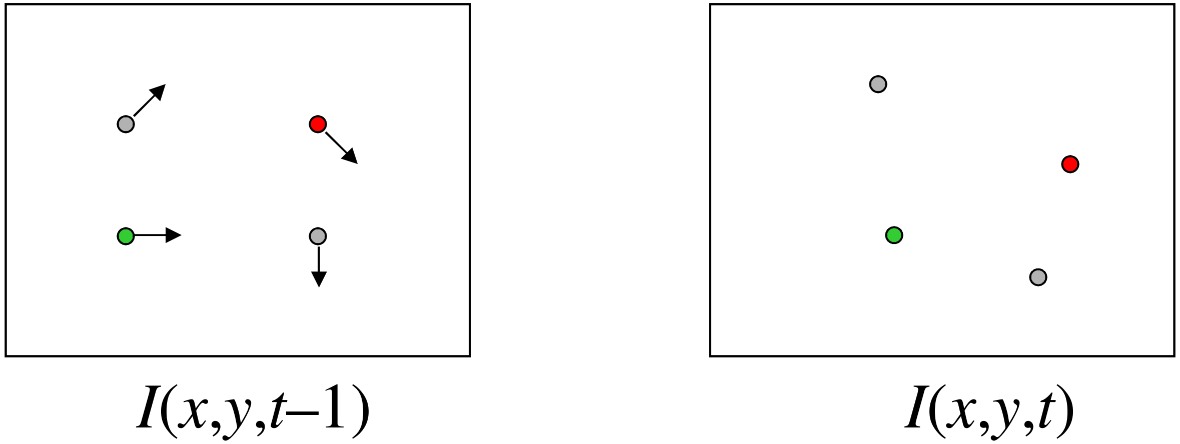
Fluxul optic

- ***Fluxul optic*** **reprezintă mișcarea aparentă a regiunilor de o anumită intensitate luminoasă în imagine.**
- **Ideal, fluxul optic ar trebui sa fie identic câmpului de mișcare.**
- **Pot apărea mișcări aparente de intensitate chiar și când câmpul de mișcare este zero:**

Fluxul optic

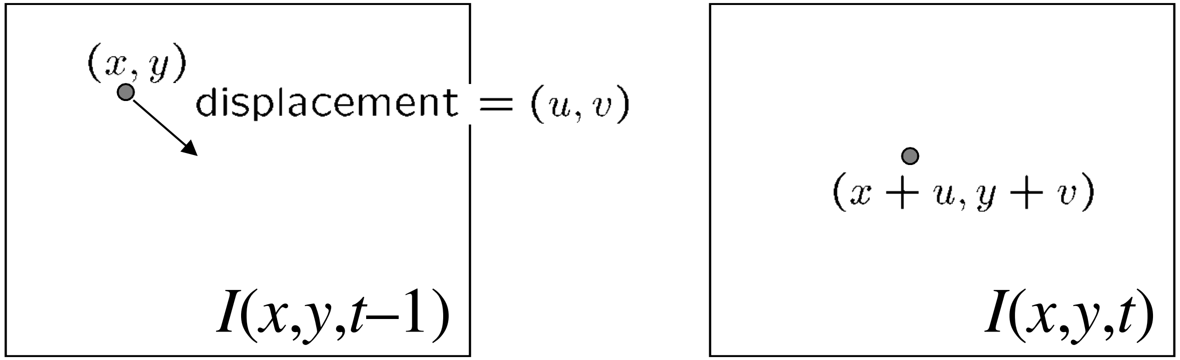
- ***Fluxul optic*** **reprezintă mişcarea aparentă a regiunilor de o anumită intensitate luminoasă în imagine.**
- **Ideal, fluxul optic ar trebui sa fie identic câmpului de mişcare.**
- **Pot apărea mişcări aparente de intensitate chiar şi când câmpul de mişcare este zero:**
  - **Atunci când obiectul şi camera rămân nemişcate, dar sursa luminoasă se deplasează**

Estimarea fluxului optic



- **Estimarea câmpului de mișcare din două imagini.**
- **Presupuneri:**
  - **Intensitatea luminoasă este similară în ambele cadre (același punct din $t-1$ este similar cu cel din t)**
  - **Deplasare mică (punctele nu se deplasează mult)**
  - **Consistență spațială (punctele se deplasează asemenea vecinilor)**

Ecuația intensității luminoase constante



$$I(x, y, t-1) = I(x + u(x, y), y + v(x, y), t)$$

$$I(x, y, t-1) \approx I(x, y, t) + I_x \cdot u(x, y) + I_y \cdot v(x, y)$$

$$I_x \cdot u + I_y \cdot v + I_t \approx 0$$

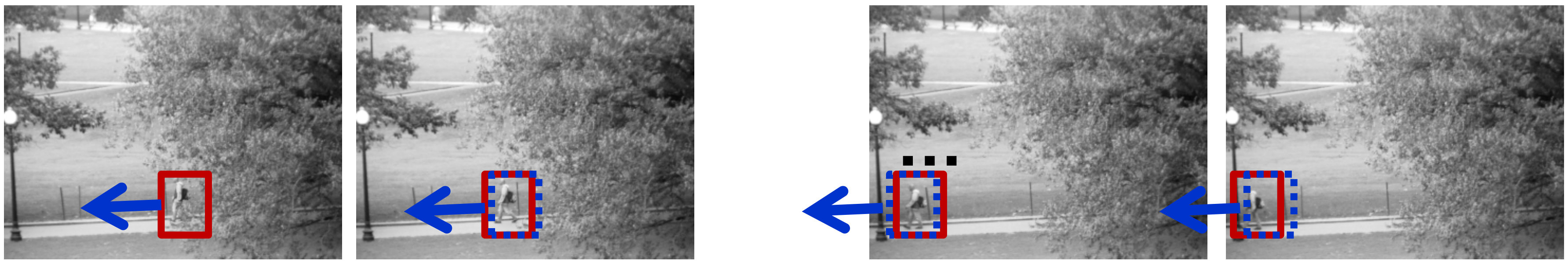
$$I_x = \frac{\partial I}{\partial x}$$

Modele dinamice în urmărirea formelor

Detecție vs. Tracking

- **Detecția: Obiectul în mișcare este detectat independent în fiecare imagine, iar poziția sa este înregistrată.**

- **Tracking (încorporând *dinamică*): Obiectul în mișcare este detectat, iar poziția sa determinată; Dinamica este încorporată pentru a prezice următoarea poziție a obiectului.**



Dinamica

- **Dinamica, în urmărirea formelor, utilizează un model al mişcării obiectului pentru a *prezice* unde se va afla acesta în cadrul următor (înaintea achiziției imaginii)**

Dinamica

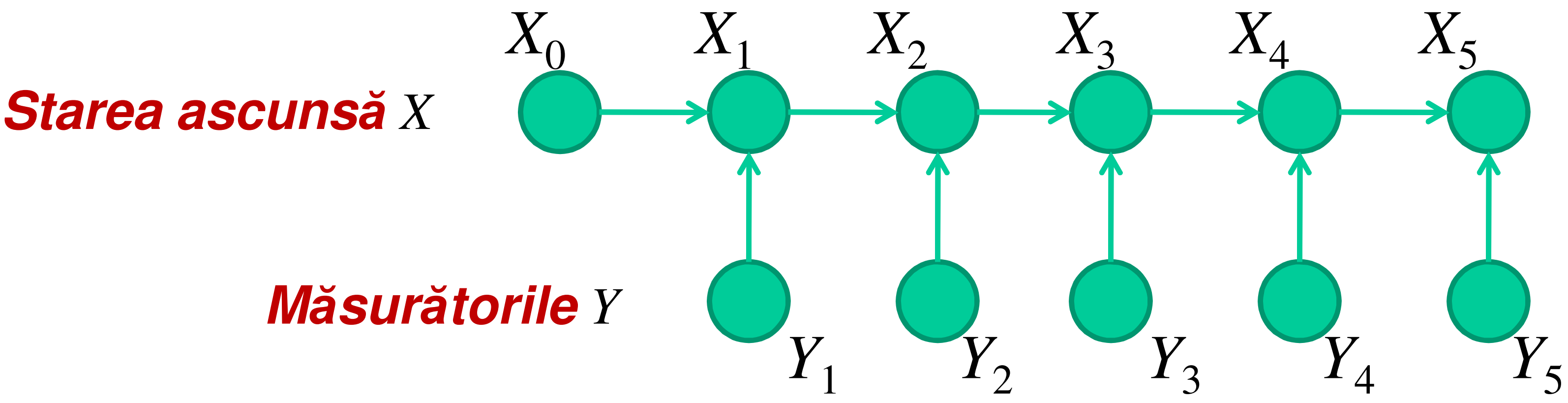
- **Avantaje:**
  - **Detecția obiectului este restricționată la anumite zone ale imaginii;**
  - **Estimatele poziției sunt îmbunătățite prin modelarea zgomotului (erorii de măsurare) în modelul dinamic.**
- **Presupuneri:**
  - **Camera nu se deplasează instant într-o nouă poziție;**
  - **Obiectele nu dispar și/sau reapar în imagini;**
  - **Obiectul și/sau camera se deplasează gradual.**

Urmărirea formelor prin inferență

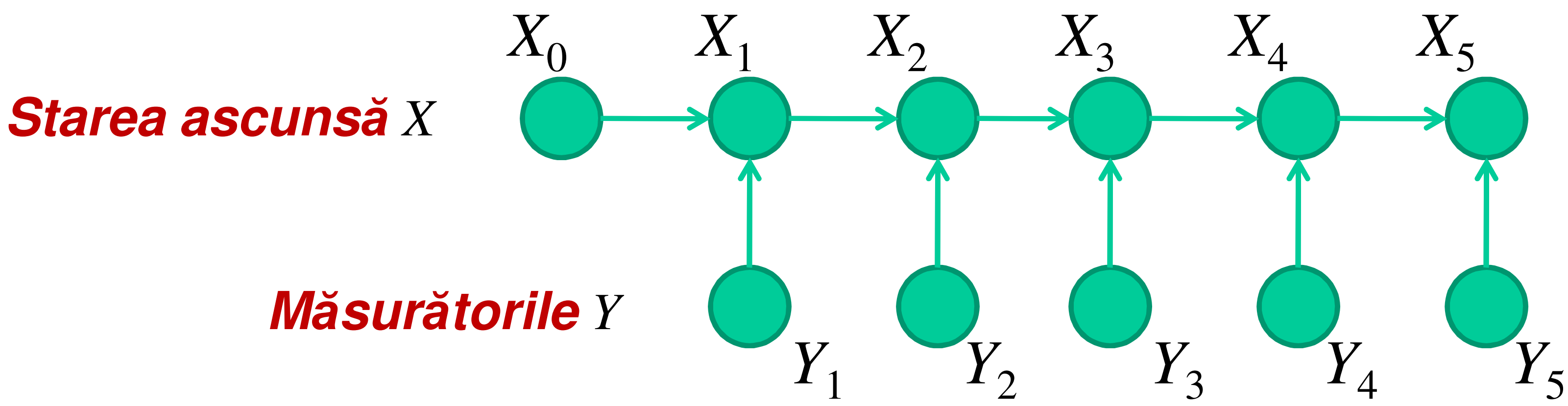
- ***Starea ascunsă*** X **este reprezentată de parametrii de interes (poziția obiectului).**
- ***Măsurătorile*** Y **(observații cu zgomot) sunt condiționate de starea curentă.**

Urmărirea formelor prin inferență

- ***Starea ascunsă*** **X este reprezentată de parametrii de interes (poziția obiectului).**
- ***Măsurătorile*** **Y (observații cu zgomot) sunt condiționate de starea curentă.**

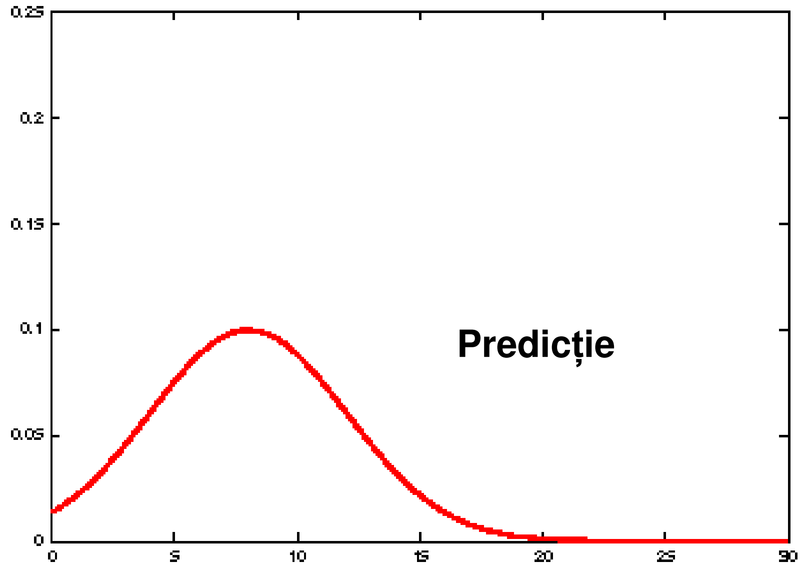


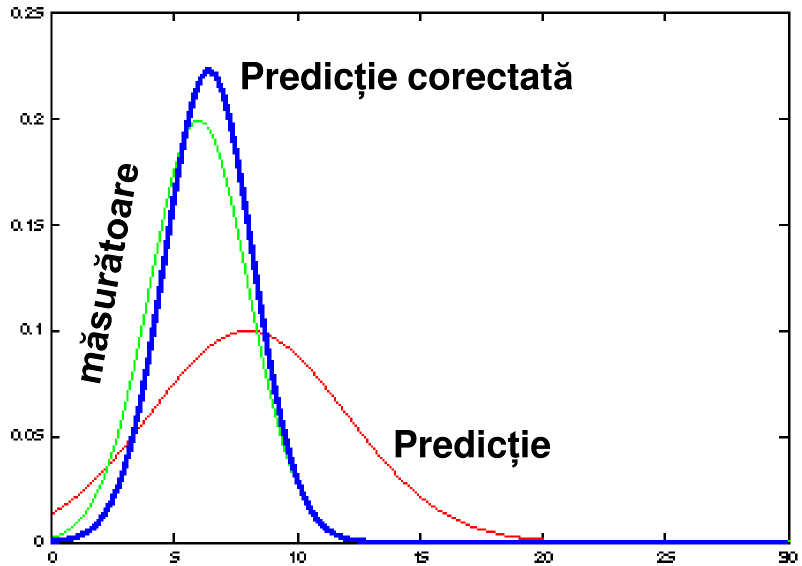
Urmărirea formelor prin inferență



- **La fiecare moment de timp, starea se modifică de la X_{t-1} la X_t, efectuând astfel noi observații Y_t**
- **Obiectiv:**
  - **Determinarea celei mai probabile stări X_t, dându-se toate observațiile efectuate.**
  - **Se cunoaște dinamica tranziției dintre stări.**

Urmărirea formelor prin inferentă





t-1 t

Presupuneri

- **Doar starea precedentă (t-1) influențează starea curentă**

$$P(\boldsymbol{X}_i|\boldsymbol{X}_1,\ldots,\boldsymbol{X}_{i-1}) = P(\boldsymbol{X}_i|\boldsymbol{X}_{i-1})$$

- **Măsurătorile la momentul de timp i depind doar de starea curentă**

$$P(\boldsymbol{Y}_i,\boldsymbol{Y}_j,\ldots\boldsymbol{Y}_k|\boldsymbol{X}_i) = P(\boldsymbol{Y}_i|\boldsymbol{X}_i)P(\boldsymbol{Y}_j,\ldots,\boldsymbol{Y}_k|\boldsymbol{X}_i)$$

Urmărirea formelor prin inferență

- **Predicție:**
  - **Dându-se măsurătorile până la momentul de timp t-1, ce stare se prezice pentru momentul t?**

$$P\left(X_t \middle| y_0, \ldots, \boxed{y_{t-1}}\right)$$

- **Corecție:**
  - **Dându-se măsurătoarea curentă (t), care este noua stare estimată pentru momentul t?**

$$P\left(X_t \middle| y_0, \ldots, \boxed{y_t}\right)$$

Proces recursiv

- **Există un estimat inițial al stării $P(X_0)$, ce va fi corectat odată cu prima măsurătoare $Y_0=y_0$.**

- **Dându-se estimatul corectat pentru cadrul t-1:**
  1) **Predicție estimat pentru cadrul t**
  2) **Corecție estimat pentru cadrul t**

Probleme deschise

- **Cum se poate reprezenta dinamica ce guvernează tranziția stărilor?**
- **Cum se pot descrie relațiile dintre stări și măsurători, luând în considerare și incertitudinea (zgomotul) măsurătorilor?**
- **Cum se poate calcula corecția stării, având noile măsurători și starea precedentă?**

- **Se vor lua în considerare *modele dinamice liniare*, reprezentate de funcții de densitate Gaussiene.**

- **Corecția stării: *Filtrul Kalman***

Model dinamic liniar

- Descrie informația apriori legată de:
  - Modelul dinamic al sistemului: evoluția stării de-a lungul timpului (incluzând zgomotul)

$$x_t = Dx_{t-1} + \varepsilon$$

  - Modelul dinamic al măsurătorilor: la fiecare moment de timp măsurătoarea obținută conține zgomot

$$y_t = Mx_t + \xi$$

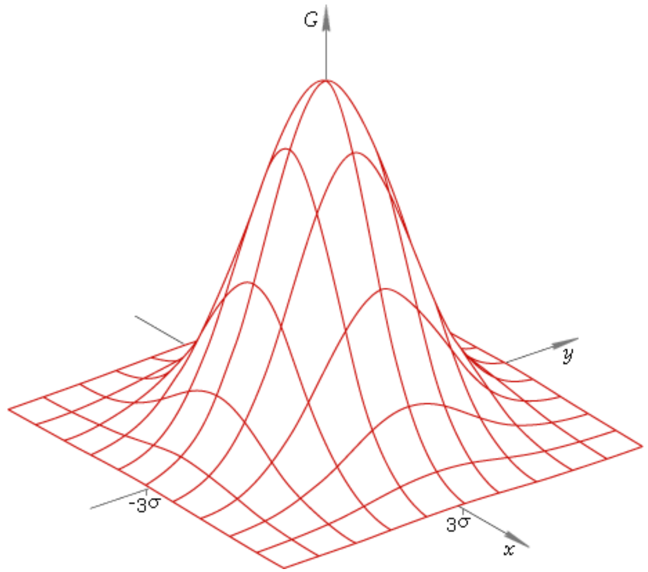
- D, M: matrici de tranziție a stărilor
- ε, ξ: zgomot Gaussian adăugat procesului (informație a-priorii)

Model dinamic liniar

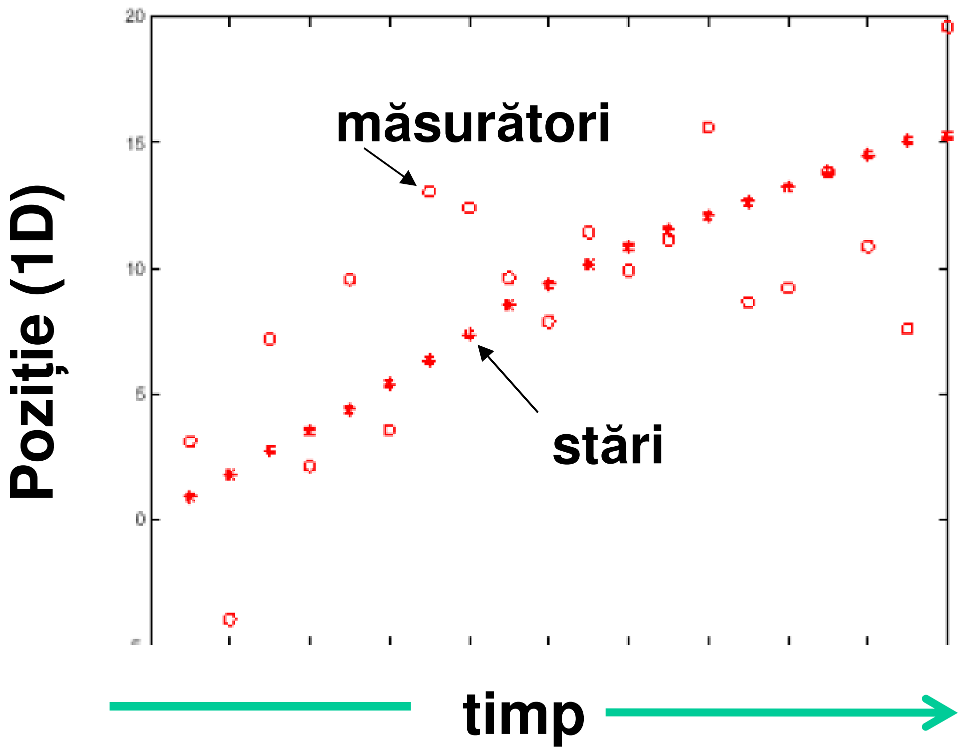
- x, y - **funcții de distribuție Gaussiană**
- ε, ξ: **zgomot Gaussian adăugat procesului (informație a-priorii)**

$$\varepsilon \sim N(0, Q)$$
$$\xi \sim N(0, R)$$

- Q, R – **matrici de covarianță a zgomotului**



Exemplu: viteză constantă



Exemplu: viteză constantă

- **Starea: poziția** p **și viteza** v

$$x_t = \begin{bmatrix} p_t \\ v_t \end{bmatrix} \qquad p_t = p_{t-1} + (\Delta t)v_{t-1} + \varepsilon$$

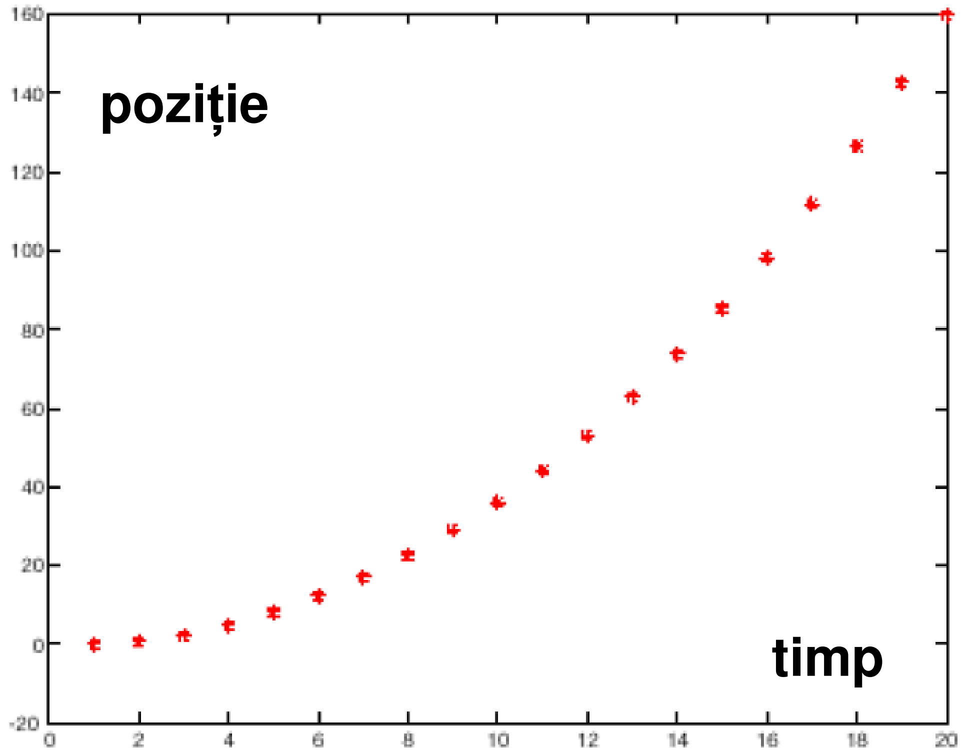
$$x_t = D_t x_{t-1} + zgomot = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_{t-1} \\ v_{t-1} \end{bmatrix} + zgomot$$

- **Măsurătoarea: poziția** p

$$y_t = Mx_t + zgomot =$$

ε, ξ **- zgomot**

Exemplu: accelerație constantă



Exemplu: accelerație constantă

- **Stare: poziție p, viteză v, accelerație a.**

$$x_t = \begin{bmatrix} p_t \\ v_t \\ a_t \end{bmatrix} \qquad \begin{aligned} p_t &= p_{t-1} + (\Delta t) v_{t-1} + \varepsilon \\ v_t &= \\ a_t &= \end{aligned}$$

$$x_t = D_t x_{t-1} + zgomot =$$

- **Măsurătoare: poziție**

$$y_t = M x_t + zgomot =$$

Corecția stării: Filtrul Kalman

$$K_{t+1} = \frac{P_{t+1|t} M^T}{M P_{t+1|t} M^T + R}$$

$$\hat{\mathbf{x}}_{t+1} = \hat{\mathbf{x}}_{t+1|t} + K_{t+1}\left(\mathbf{y}_{t+1} - M\hat{\mathbf{x}}_{t+1|t}\right)$$

$$P_{t+1} = P_{t+1|t} - K_{t+1} M P_{t+1|t}$$

- K **– matrice de amplificare Kalman**
- P **– matricea de covarianță a erorii (calculată din stările** x_k **și** x_{k+1}**)**
- **Pentru obţinerea unei stări estimate cât mai precise în raport cu starea sistemului, incertitudinea trebuie să fie cât mai mică.**
- ***Filtrul Kalman este un filtru optimal în sensul că acesta minimizează matricea de covarianţă a erorii*** P**.**