Behavioural Pedestrian Tracking (II)

Javier Cruz

École Polytechnique Fédérale de Lausanne

Outline

- Visual tracking
- Pedestrian Visual Tracking
- Evolution of the approach
- Future work

Visual Tracking

Ross, D., Lim, J. and Lin, R.-S. (2008), *Incremental Learning for Robust Visual Tracking*. International Journal of Computer Vision 77: 125-141.

- On-Line low-dimensional subspace representation (incremental PCA)
- Gaussian variables
- (Particle filtering)

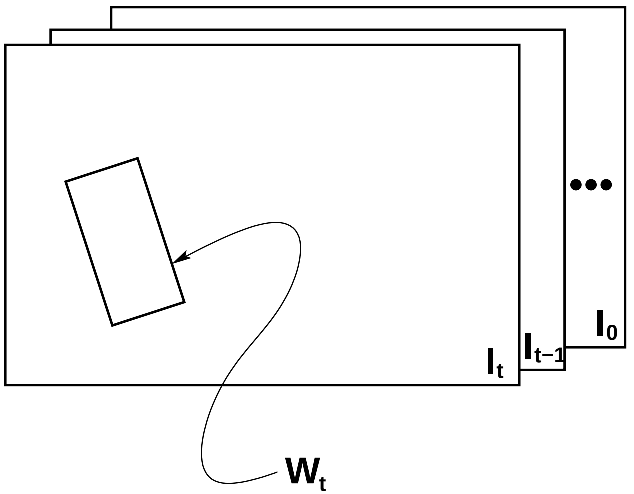
TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Visual Tracking

Definitions:

- I_t: frame t
- $X_t = (x_t, y_t)$: position
- φ_t: rotation
- γ_t: skewness
- s_t: scale
- $r_t = \frac{w_t}{h_t}$: aspect ratio
- $W_t = f(I_t, X_t, \varphi_t, \gamma_t, s_t, r_t)$: a patch in I_t
- $\Psi_t = \{X_t, \theta_t\} = \{X_t, \varphi_t, \gamma_t, s_t, r_t\} \in \Theta_t$



TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Visual Tracking

Everything is supposed to be Gaussian:

- $X_t \sim \mathcal{N}(X_{t-1}, \sigma_X)$
- $\varphi_t \sim \mathcal{N}(\varphi_{t-1}, \sigma_\varphi)$
- $\gamma_t \sim \mathcal{N}(\gamma_{t-1}, \sigma_\gamma)$
- $s_t \sim \mathcal{N}(s_{t-1}, \sigma_s)$
- $r_t \sim \mathcal{N}(r_{t-1}, \sigma_r)$

$$\Psi_t = \{X_t, \varphi_t, \gamma_t, s_t, r_t\} \in \Theta_t$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Visual Tracking

$$\Psi_t^* = \arg\max_{\Psi_t \in \Theta_t} p(I_t|\Psi_t)p(\Psi_t|\Psi_{t-1})$$

Observation:

$$p(I_t|\Psi_t) \sim \mathcal{N}(W_t; \mu, UU^\top + \varepsilon I)\mathcal{N}(W_t; \mu, U\Sigma_o^{-2}U^\top)$$

Dynamics:

$$p(\Psi_t|\Psi_{t-1}) \sim \mathcal{N}(\Psi_t; \Psi_{t-1}, \Sigma_\Psi)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Visual Tracking

(Visual Tracking demo)

TRANSP-OR

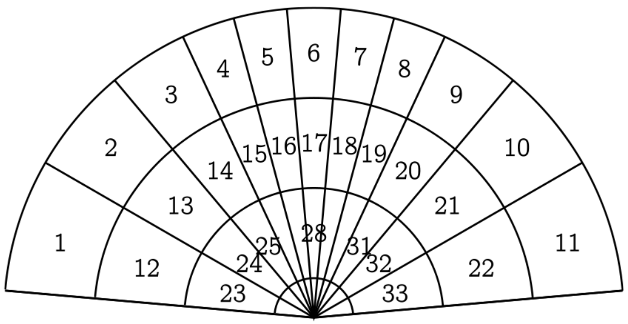
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Behavioural Visual Tracking

Robin, T., Antonini, G., Bierlaire, M., and Cruz, J. (2009), *Specification, estimation and validation of a pedestrian walking behavior model*. Transportation Research Part B: Methodological 43(1):36-56.

Next step given by a cross-nested logit model that takes into account:

- Direction
- Destination
- Speed
- “Leader-follower”
- “Collision avoidance”



Choice Set

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Behavioural Visual Tracking

$$
\begin{aligned}
V_{vdn} = &\beta_{\text{dir\_central}} \text{dir}_{dn} I_{\text{d,central}} + \\
&\beta_{\text{dir\_side}} \text{dir}_{dn} I_{\text{d,side}} + \\
&\beta_{\text{dir\_extreme}} \text{dir}_{dn} I_{\text{d,extreme}} +
\end{aligned}
\Bigg\} \quad \textit{keep direction}
$$

$$
\begin{aligned}
&\beta_{\text{ddist}} \text{ddist}_{vdn} + \\
&\beta_{\text{ddir}} \text{ddir}_{dn} +
\end{aligned}
\Bigg\} \quad \textit{toward destination}
$$

$$
\begin{aligned}
&\beta_{\text{dec}} I_{\text{v,dec}} (v_n / v_{\text{max}})^{\lambda_{\text{dec}}} + \\
&\beta_{\text{accLS}} I_{\text{n,LS}} I_{\text{v,acc}} (v_n / v_{\text{maxLS}})^{\lambda_{\text{accLS}}} + \\
&\beta_{\text{accHS}} I_{\text{n,HS}} I_{\text{v,acc}} (v_n / v_{\text{max}})^{\lambda_{\text{accHS}}} +
\end{aligned}
\Bigg\} \quad \textit{free flow acceleration}
$$

$$
\begin{aligned}
&I_{\text{v,acc}} I^{L}_{\text{d,acc}} \alpha^{L}_{\text{acc}} D_L^{\rho^L_{\text{acc}}} \Delta v_L^{\gamma^L_{\text{acc}}} \Delta \theta_L^{\delta^L_{\text{acc}}} + \\
&I_{\text{v,dec}} I^{L}_{\text{d,dec}} \alpha^{L}_{\text{dec}} D_L^{\rho^L_{\text{dec}}} \Delta v_L^{\gamma^L_{\text{dec}}} \Delta \theta_L^{\delta^L_{\text{dec}}} +
\end{aligned}
\Bigg\} \quad \textit{leader-follower}
$$

$$
I_{\text{d},d_n} I_{d,C} \alpha_C e^{\rho_C D_C} \Delta v_C^{\gamma_C} \Delta \theta_C^{\delta_C} \Big\} \quad \textit{collision avoidance}
$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Behavioural Visual Tracking

This forces some assumptions:

- Camera is calibrated
- Camera is fixed
- Pedestrians walking in normal conditions
- Destination known!!

TRANSP-OR

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Behavioural Visual Tracking

Pedestrian model + Gaussian:

- $X_t \sim$ pedestrian walking behaviour model (PWBM)
- $\varphi_t \sim \mathcal{N}(\varphi_{t-1}, \sigma_\varphi)$
- $\gamma_t \sim \mathcal{N}(\gamma_{t-1}, \sigma_\gamma)$
- $s_t \sim \mathcal{N}(s_{t-1}, \sigma_s)$
- $r_t \sim \mathcal{N}(r_{t-1}, \sigma_r)$

$$\Psi_t = \{X_t, \theta_t\} = \{X_t, \varphi_t, s_t, r_t\} \in \Theta_t$$

Note that as calibration data is known, some corrections can be done in the sizes of the windows

Behavioural Visual Tracking

$$\Psi_t^* = \arg\max_{\Psi_t \in \Theta_t} p(I_t|\Psi_t)p(\Psi_t|\Psi_{t-1})$$

Observation:

$$p(I_t|\Psi_t) \sim \mathcal{N}(W_t; \mu, UU^\top + \varepsilon I)\mathcal{N}(W_t; \mu, U\Sigma_o^{-2}U^\top)$$

Dynamics:

$$p(\Psi_t|\Psi_{t-1}) \sim \text{PWBM}(X_t; X_{t-1})\mathcal{N}(\theta_t; \theta_{t-1}, \Sigma_\theta)$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Behavioural Visual Tracking

(Behavioural Visual Tracking demo)

One step further (literally)

The idea in order to solve the problems of the current approach is to delay the decision, i.e. propagate probability distributions during several frames instead of choosing the winner for each frame.

Advantages:

- Occlusions
- Trajectory “coherence”

Drawbacks:

- Interdependence (collision avoidance and leader follower)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

One step further (literally)

Simulation-based tracking:

trajectory simulation + probability distribution of patches

Signal processing on manifolds:

- Probability distribution of patches
- Define a measure

One problem: 0.5 seconds!

Future work

- Develop formally these ideas (define the manifold and the measure)
- Implement and test
- Instead of a window, define something more similar to a pedestrian in 3D

TRANSP-OR

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE