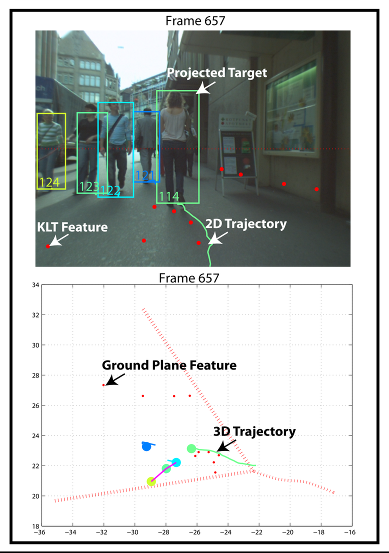
Multiple Targets Tracking in World Coordinate with a Single, Minimally Calibrated Camera.

Wongun Choi Silvio Savarese

Department of Electrical Engineering and Computer Science, University of Michigan

Overview

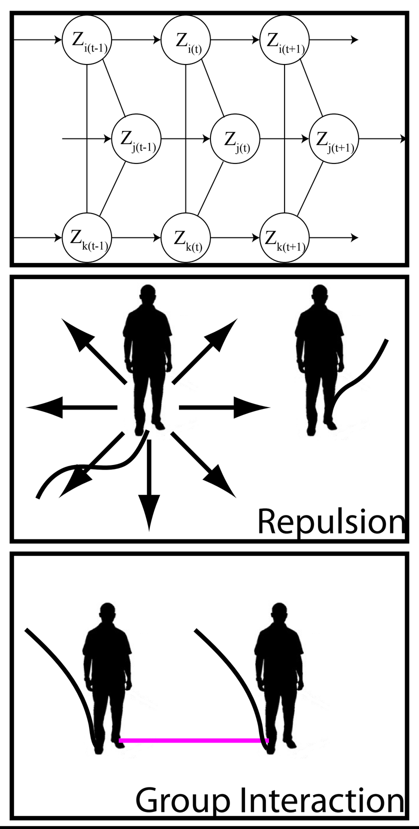


- Objective : robustly track multiple targets in 3D.
- Joint estimation of camera and target's motion.
- Modeling interaction between targets.
- Ground feature points for robust camera estimation.
- MCMC particle filtering for efficient estimation.

Challenges

- Unknow (uncalibrated) camera motion.
- Occlusion between targets.
- Background clutter.
- Detector failure (missed detection).

Interaction Model



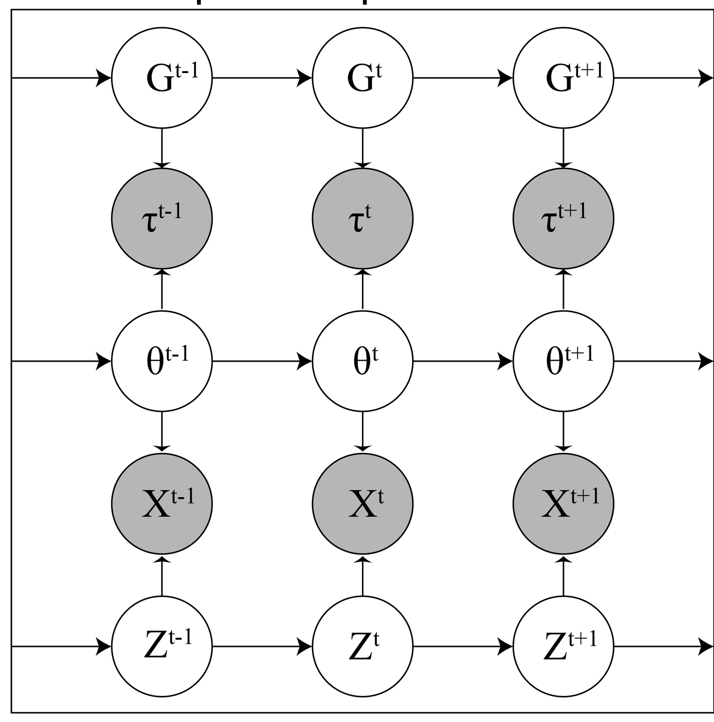
- Pairwise interaction Model
  - Targets' motions are dependent.
- Two exclusive mode of interactions
  - Repulsion: people want to keep a distance from others.
  - Group Interaction: People moving as a group tend to move together.
  - Switch variables to select one of the above models.

$$P(Z_t|Z_{t-1}) = \prod_{i<j} \psi(Z_{it}, Z_{jt}; \beta_{ijt}) \prod_{i<j} P(\beta_{ijt}|\beta_{ij(t-1)}) \prod_{i=1}^{N} P(Z_{it}|Z_{i(t-1)})$$

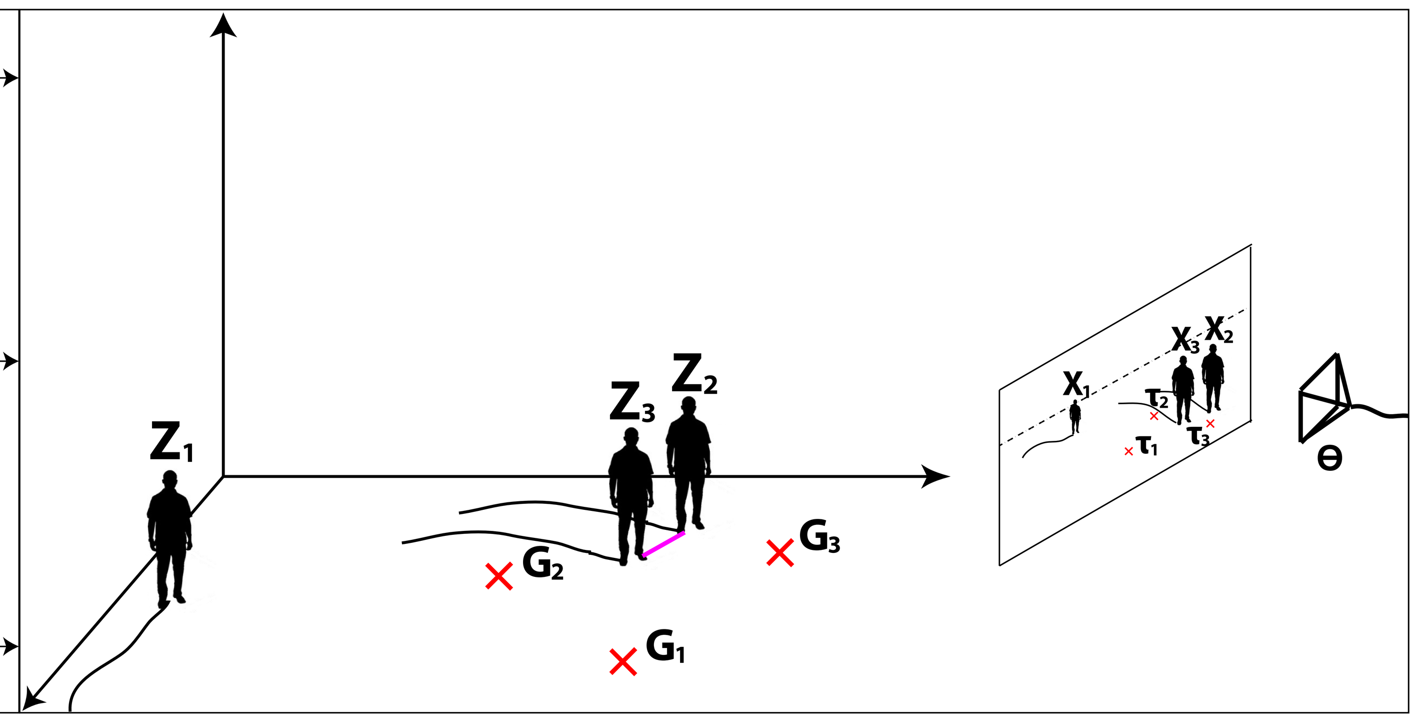
$$\psi(Z_{it}, Z_{jt}; \beta_{ijt}) = \begin{cases} \psi_g(Z_{it}, Z_{jt}), & if\ \beta_{ijt} = 1 \\ \psi_r(Z_{it}, Z_{jt}), & otherwise \end{cases}$$

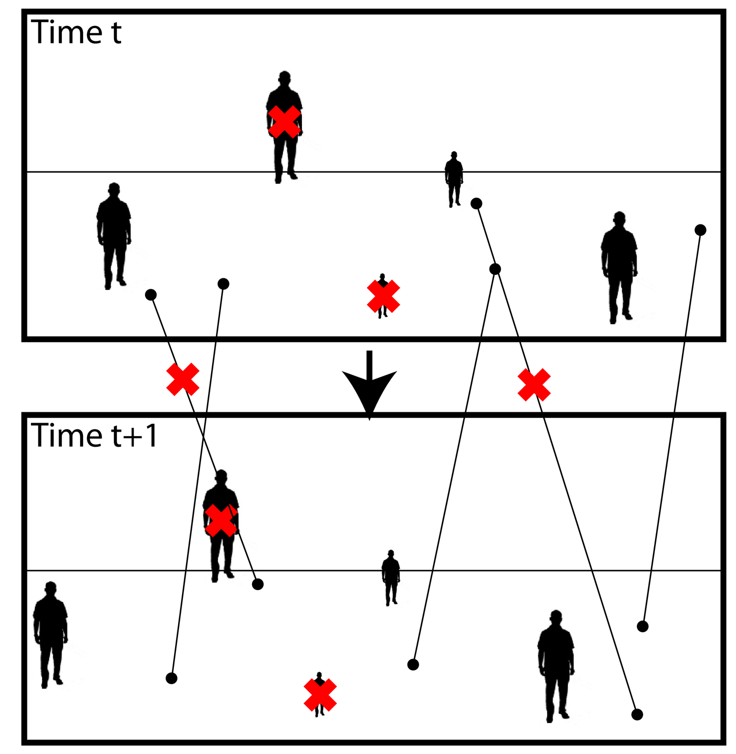
Joint Model

Graphical Representation



World Illustration





- Sequential bayesian formulation

$$P(\Omega_t|\chi^t) \propto P(\Omega_t, \chi_t|\chi^{t-1}) = P(\chi_t|\Omega_t) \int P(\Omega_t|\Omega_{t-1}) P(\Omega_{t-1}|\chi^{t-1}) d\Omega_{t-1}$$

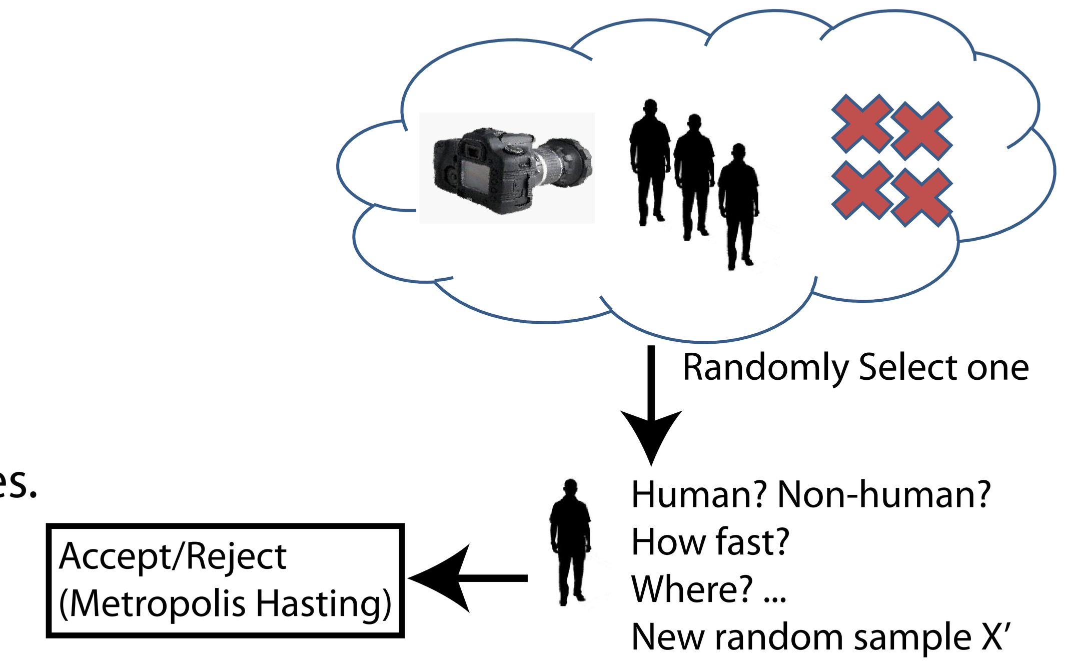
$$P(\chi_t|\Omega_t) = P(X_t, Y_t|Z_t, \Theta_t) P(\tau_t|G_t, \Theta_t)$$

$$P(\Omega_t|\Omega_{t-1}) = P(Z_t|Z_{t-1}) P(\Theta_t|\Theta_{t-1}) P(G_t|G_{t-1})$$

- Observation models
  - Simplified camera projection function [4].
- Motion models
  - Target's individual motion modeled as a first order linear dynamic model.
  - Ground features assumed to be static.
  - Camera assumed to move along the viewing direction.
- Indicator variable
  - Target : indicate whether this target is valid human or not.
  - Ground features : indicate whether the feature is on static ground plane.

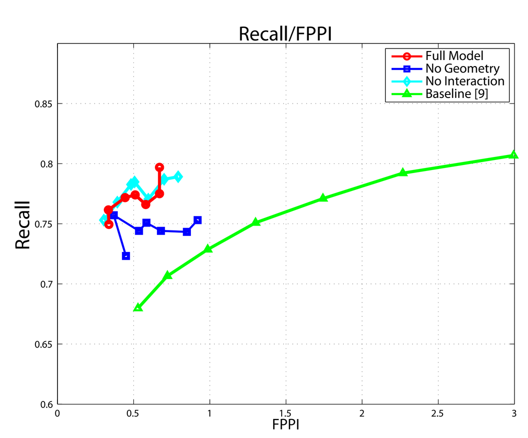
Estimation : MCMC Particle Filtering

- Motivation
  - Non-linear and non-gaussian density function.
  - Direct sampling from high dimensional space is not efficient.
- Proposal Distribution $Q(\Omega'_t; \Omega_t)$
  - Choose one of the variables : camera parameter, target states, or ground features.
  - Sample from Gaussian distribution centered on current sample.
  - Randomly change the class of a target or a ground feature.



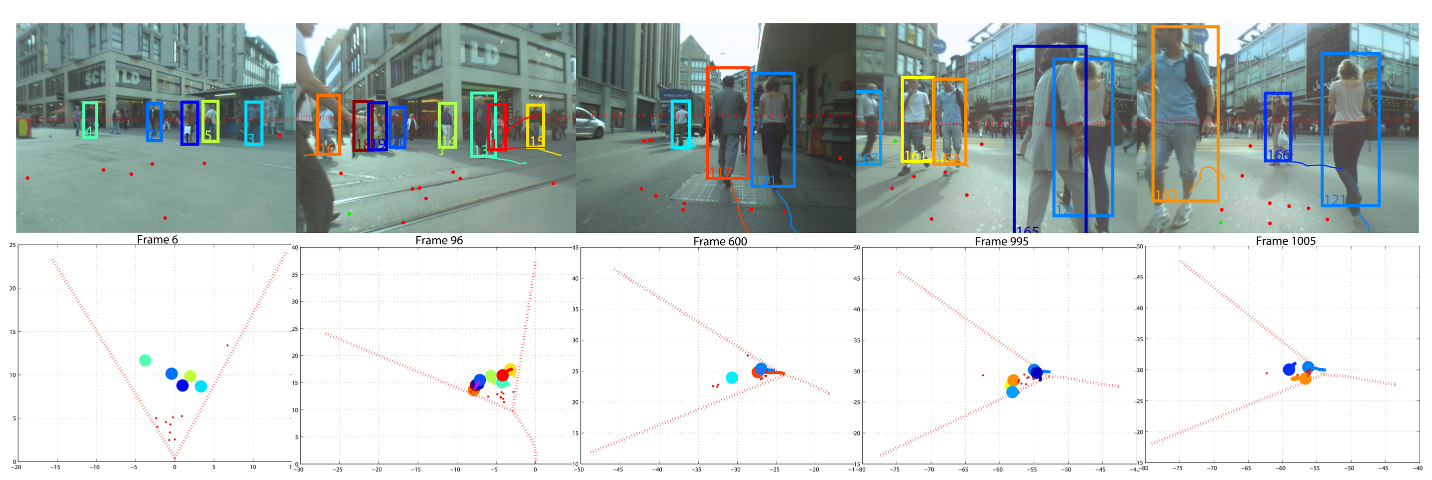
Experimental Results

- Improved detection accuracy.

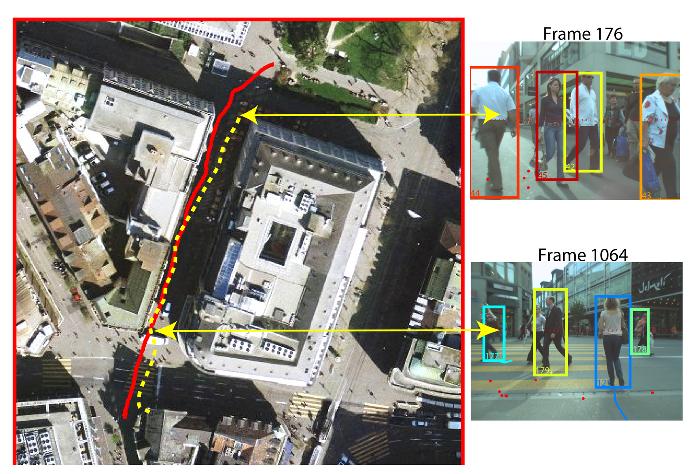


Recall/FPPI on ETH dataset			
Method		Seq.#2	Seq.#3
Our Algorithm	Recall	0.556 0.541 0.519	0.339 0.421 0.497
	FPPI	0.792 0.442 0.267	2.792 1.608 0.647
ETH [1]	Recall	0.498 0.404 0.338	0.673 0.616 0.484
	FPPI	0.781 0.431 0.262	2.772 1.593 0.638

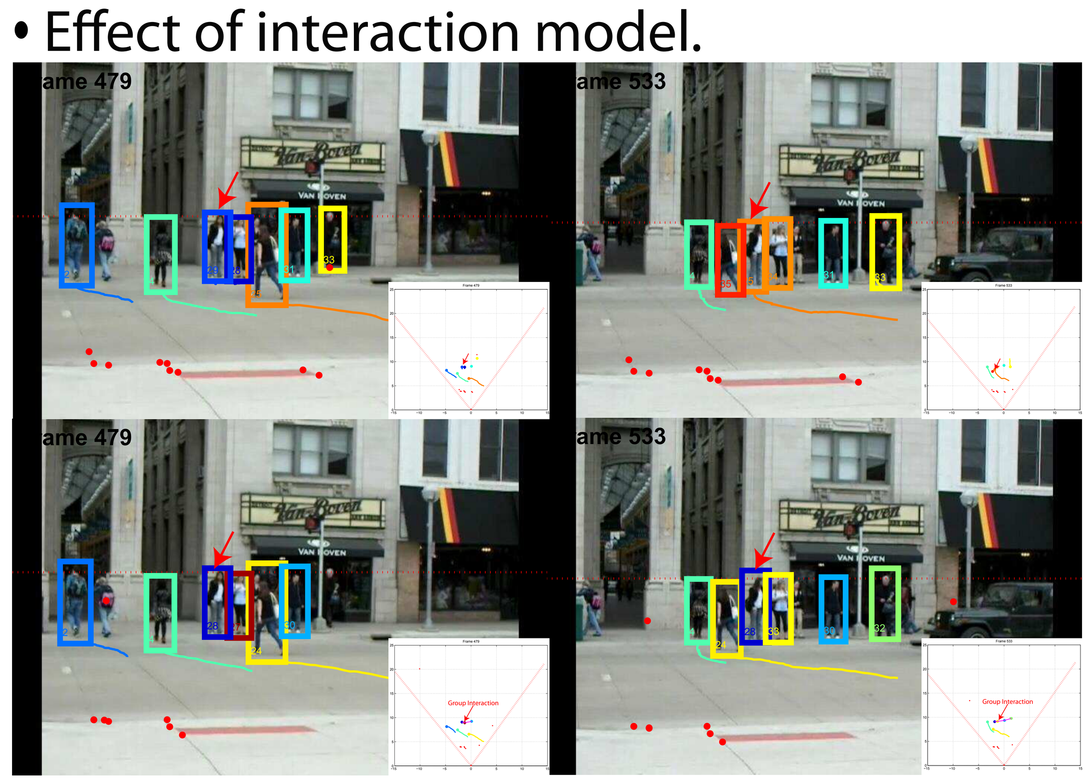
- Example tracking results on ETH sequence.



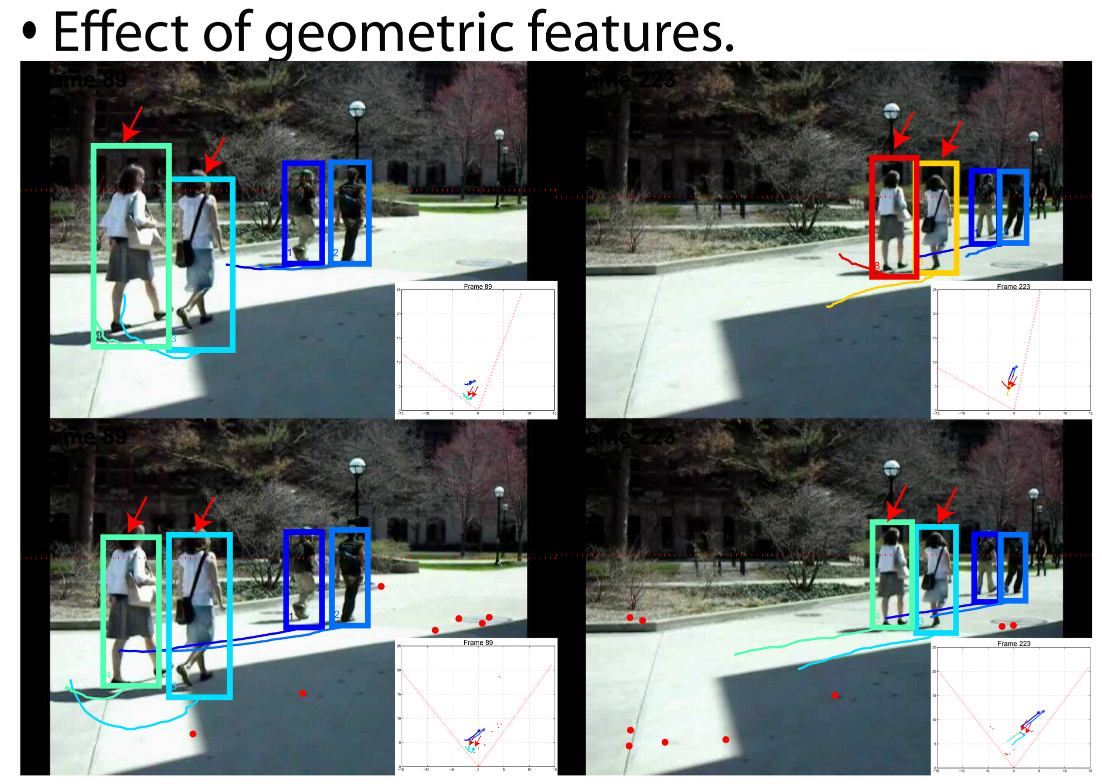
- Accurate localization of camera



- Effect of interaction model.



- Effect of geometric features.



Conclusion

- Joint estimation of camera and targets' state helps improve tracking performance.
- Other geometric cues further stabilize the system.
- Better target association can be achieved by modeling interaction between targets.

References

[1] Ess, A., Leibe, B., Schindler, K., , van Gool, L.: A mobile vision system for robust multi-person tracking. In: CVPR. (2008)

[2] Khan, Z., Balch, T., Dellaert, F.: Mcmc-based particle fltering for tracking a variable number of interacting targets. PAMI (2005)

[3] Pellegrini, S., Ess, A., Schindler, K., van Gool, L.: You'll never walk alone: Modeling social behavior for multi-target tracking. In: ICCV. (2009)

[4] Hoiem, D., Efros, A., Hebert, M.: Putting objects in perspective. In: CVPR. (2006)