Tracking Multiple People with Illumination Maps

Gloria Zen, Oswald Lanz, Stefano Messelodi, Elisa Ricci

{gzen,lanz,messelod,eliricci}@fbk.eu

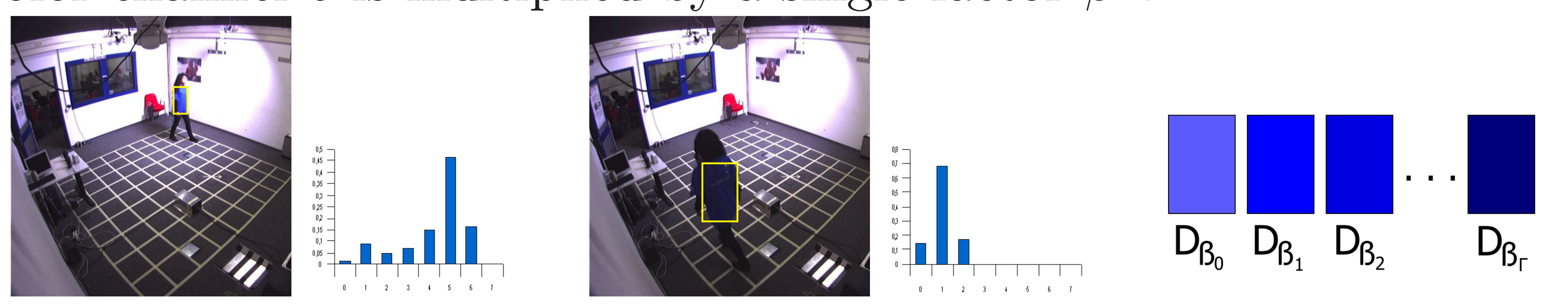
FONDAZIONE BRUNO KESSLER

Contribution

Motivation: developing a visual tracking system that is robust and adaptive to changing illumination conditions.
Our Scenario: *(i)* non-homogeneous and non-stationary (slowly varying in time) illumination conditions, *(ii)* occlusions due to other targets or neighbouring objects.
Color-based tracking: cues other than color (motion, edges or speech) are not always reliable; color histograms are sensitive to illumination.
Idea: learn a global illumination map describing the lighting conditions of the scene .
Our approach: *(i)* joint estimation of the position of the target and its illumination condition within a Bayesian framework, *(ii)* implement the map as a hierchical Markov Random Field (MRF) with a novel efficient inference approach.
Experiments. We tested our method on two different scenarios: our laboratory (LAB) and a public dataset (PETS2007).

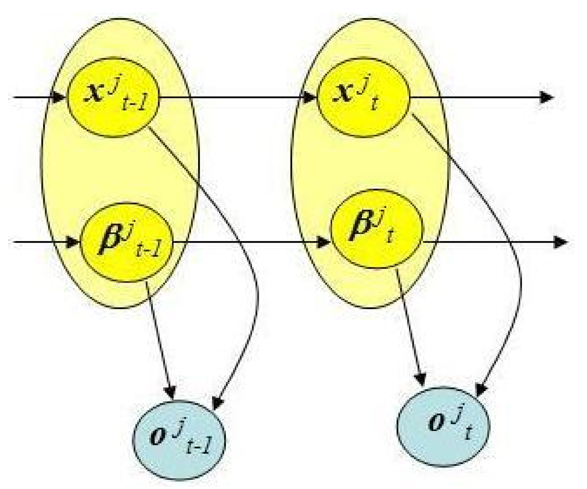
Modelling the illumination changes

The color change under changing illumination can be described by the so-called diagonal (or von Kries) model [3], according to which each color channel c is multiplied by a single factor β^c.



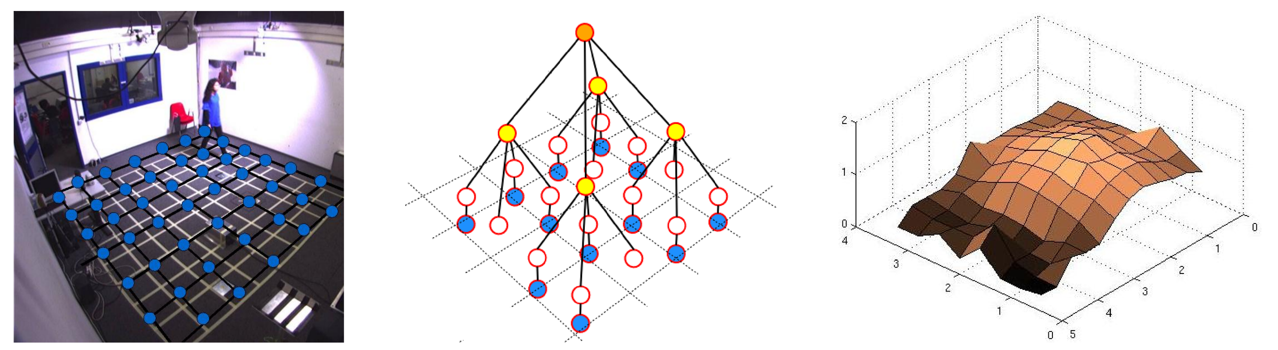
Tracking with Multiple Templates

- Multiple people tracking with the Hybrid Joint Separable (HJS) particle filter [1].
- Extending the HJS by introducing **multiple templates**. We consider for each target a discrete set of Γ templates, representing the target under different illumination conditions. The set is obtained by diagonal transforms $P_\beta = D_\beta P_0$, with $D_\beta = \mathrm{diag}(\boldsymbol{\beta}) = \mathrm{diag}(\beta^r, \beta^g, \beta^b)$.
- **Joint estimation** of target position $\boldsymbol{x}_t$ and illumination conditions $\boldsymbol{\beta}_t$.
- Tradeoff: **adaptivity vs tracking failure**. The tracker is able to adapt to target color changes but risks to drift towards background objects with similar appearance.



- **Dynamical model** of the HJS. We define $p(\boldsymbol{\beta}^q_t|\boldsymbol{\beta}^q_{t-1}, \boldsymbol{x}^q_t) = p(\boldsymbol{\beta}^q_t|\boldsymbol{\beta}^q_{t-1})p(\boldsymbol{\beta}^q_t|\boldsymbol{x}^q_t)$, where $p(\boldsymbol{\beta}^q_t|\boldsymbol{\beta}^q_{t-1})$ is modelled as a gaussian noise and $p(\boldsymbol{\beta}^q_t|\boldsymbol{x}^q_t)$ models the likelihood of certain illumination conditions given a location in the scene.
- **Building up an illumination map**. The prior $p(\boldsymbol{\beta}^q_t|\boldsymbol{x}^q_t)$ is provided by the map and limits the number of candidate templates.

Creating the Illumination Map

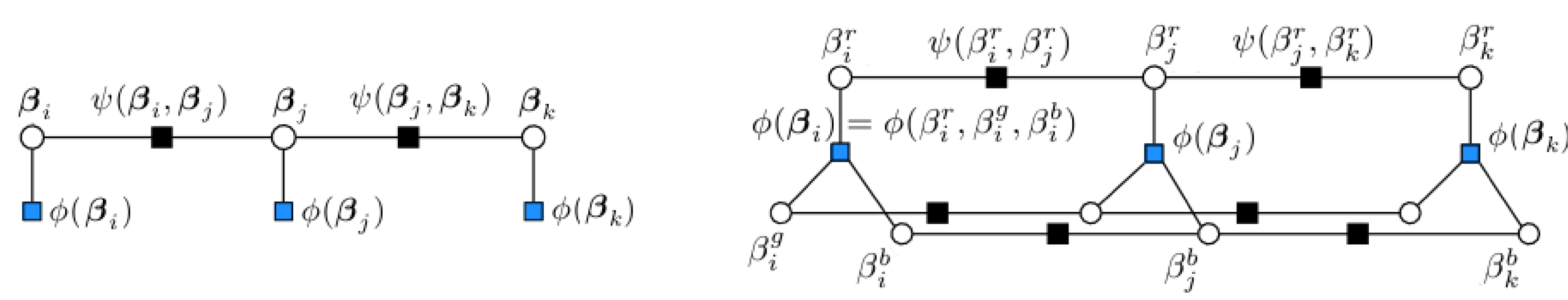


The map is obtained with a **hierarchical MRF** where:

- hidden nodes correspond to cells in the room. The illumination coefficients $\boldsymbol{\beta} = (\beta^r, \beta^g, \beta^b)$ are the hidden variables.
- observations $\boldsymbol{O_i}$ associated to the i-*th* node are histograms collected for different targets during a time interval Δ_t. A clustering method is applied to the collected histograms to reduce the effect of noisy observations.
- the likelihood $\phi_i(\boldsymbol{\beta}_i, \boldsymbol{O_i})$ for node i identify the $\boldsymbol{\beta}$ which better explains the observation $\boldsymbol{O_i}$.
- The pairwise potentials $\psi_{ij}(\boldsymbol{\beta}_i, \boldsymbol{\beta}_j)$ enforce the fact that neighboring cells should have similar illumination conditions.

Efficient Inference with High Order $\phi_i()$

- Inference with sum-product belief propagation is computationally costly, $\mathcal{O}(V\Gamma^2) = \mathcal{O}(V\ell^6)$, with V the number of nodes, Γ and ℓ respectively the label space of $\boldsymbol{\beta}$ and β^c.
- The adoption of a **decomposed model** [2] reduces the cost to $\mathcal{O}(V\ell^3)$.
- A novel message passing scheme based on Taylor series reduces the cost to $\mathcal{O}(3V\ell^2)$.



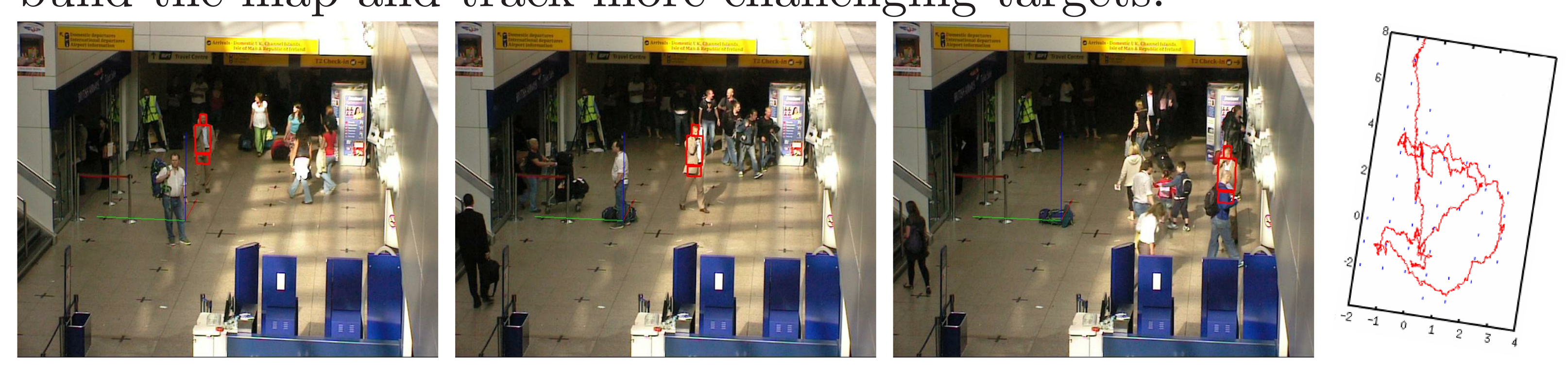
Results

LAB. Experimental results show that tracking with our method HJS-$\boldsymbol{\beta}$ allows to follow targets when subject to strong illumination changes, while adopting a $\boldsymbol{\beta}$-map reduces the risk of drifting.

The quantitative results (F-measure) calculated for 6 video sequences are reported in the table.

	s1	s2	s3	s4	s5	s6
n^o targets	1	1	2	2	3	3
HJS	0.35	0.58	0.62	0.68	0.55	0.36
HJS-β-map	0.60	0.67	0.70	0.73	0.74	0.76

PETS2007. The scenario presents strongly non-homogeneous and time-varying illumination conditions. Results show that collected information about lighting from "easy-to-track" targets can be used to build the map and track more challenging targets.



Supplementary materials of this work can be found at:
http://tev.fbk.eu/people/ricci/illumination_map.html.

References

[1] O. Lanz Approximate bayesian multibody tracking In *IEEE Trans. PAMI, 28(9):1436-1449*, 2006.
[2] A. Shekhovtsov, I. Kovtun, and V. Hlavac Efficient MRF deformation model for non-rigid image matching In *CVPR'07*.
[3] G. Finlayson, M. Drew, B. Funt Diagonal transforms suffice for color constancy In *ICCV*, 1993.