TRACKING VIA LOCAL PATCHES AND HIERARCHICAL SAMPLING

Yi K. M., Kim S. W., Jeong H., Choi J. Y. - Seoul National University
{kmyi, soowankim, hwjeong, jychoi}@snu.ac.kr

SEOUL NATIONAL UNIVERSITY

Abstract

To track objects showing partial occlusions and non-rigid deformations in real-time, we propose a tracking method based on sequential Bayesian inference. The proposed method is consisted of two parts: (1) modeling the target object using elastic structure of local patches for robust performance; and (2) efficient hierarchical sampling method to obtain an acceptable solution in real-time. The method is tested on a number of image sequences with occlusion and non-rigid deformation.

Problem Definition

Visual tracking works well, but still has problems when tracking objects which undergo partial occlusions, or which show non-affine motions.

Inference Framework

The proposed method is based on particle filtering (sequential Bayesian inference).

$$P(\mathbf{X}_t|\mathbf{Y}_{1:t}) \propto P(\mathbf{Y}_t|\mathbf{X}_t) \times \int P(\mathbf{X}_t|\mathbf{X}_{t-1}) P(\mathbf{X}_{t-1}|\mathbf{Y}_{1:t-1})$$

Elastic Structure of Local Patches (Likelihood)

Hierarchical Sampling (Prior)

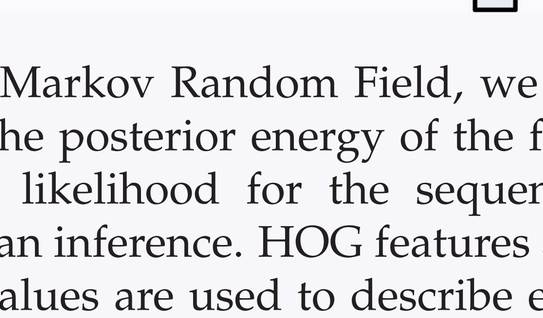
References

[1] K. M. Yi, S. W. Kim, H. Jeong, and J. Y. Choi, "Non-Rigid Object Tracking with Elastic Structure of Local Patches and Hierarchical Sampling", In IVCNZ 2010.

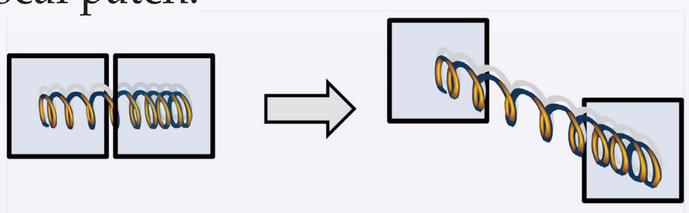
Local Patch Structure

$$P(\mathbf{Y}_t|\mathbf{X}_t) \propto \prod_{k=1}^{m} P(\mathbf{Y}_t|\mathbf{X}_t^k) \prod_{j \in N_k} P(\mathbf{X}_t^k|\mathbf{X}_t^j)$$

$$E(\mathbf{Y}_t;\mathbf{X}_t) \equiv Z + \sum_{k=1}^{m} \left[E(\mathbf{Y}_t;\mathbf{X}_t^k) + \sum_{j \in N_k} E(\mathbf{X}_t^k, \mathbf{X}_t^j) \right]$$



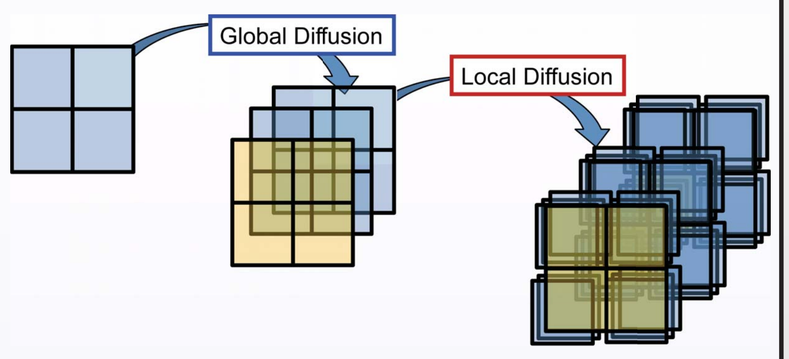
Using Markov Random Field, we defined the posterior energy of the filed as the likelihood for the sequential Bayesian inference. HOG features and RGB values are used to describe each local patch.



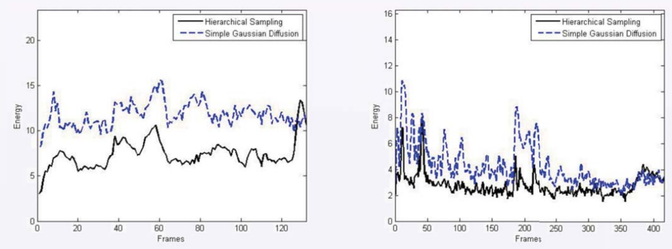
$$E(\mathbf{X}_t^k, \mathbf{X}_t^j) = \beta \frac{\|l_c(j,k) - l_m(j,k)\|_2^2}{\|l_m(j,k)\|_2^2}$$

The energy for each connection are modeled as elastic energy of springs.

Hierarchical Sampling



To focus the samples to be in regions which are likely to contain the MAP solution, we sample without changing the relative positions among each local patches and then allow each individual patches to move freely.



(a) **caviar** (b) **pedestrian**

Above is the energy obtained using hierarchical sampling and simple random walk, with same number of samples. With simple random walk, tracker fails to track the target.

Experimental Results

The proposed method is implemented in C++, using openCV and VXL libraries. Our implementation of the proposed method runs over 30 to 50 FPS, ensuring real-time performance.

Tracking results for an partially occluded person. For comparison with other methods, refer to [1].

Tracking results for a person performing high jump. The target person shows non-rigid deformations. For comparison with other methods, refer to [1].

Acknowledgements

This research was sponsored by Samsung Techwin Co.,Ltd. and SNU Brain Korea 21 IT program.