Superpixel Gridization for Fast Object Localization

Wei Feng
wfeng@tju.edu.cn

Computer Vision and Media Computing Center
Tianjin University

Outline

- Background & Objective
- Superpixel Gridization
  - Criterion
  - Optimization
  - The Algorithm
- Experiments
- Conclusion

Background

- Superpixel
  - Atomic pixel group with similar low-level feature
  - Usually very fast to obtain
  - Well preserving important edges and structures
  - Irregular (near-regular) sizes, shapes & topologies

- Superpixel-level image analysis
  - Efficiency, Accuracy, Scalable
  - Widely used in many recent CV applications
    - Segmentation
    - Tracking
    - Object detection
    - …

Recent Developments

- Superpixel segmentation and application
  - Meanshift [Comaniciu et al. TPAMI02]
  - EGS [Felzenszwalb et al. IJCV04]
  - SLIC [Achanta et al. TPAMI12]
  - SEEDS [Bergh et al. ECCV12]
  - TurboPixel [Levinshtein et al. TPAMI09]
  - SuperLattice [Moore et al. CVPR08]
  - LatticeCut [Moore et al. CVPR10]
  - Importance of spatial regularity of SPs

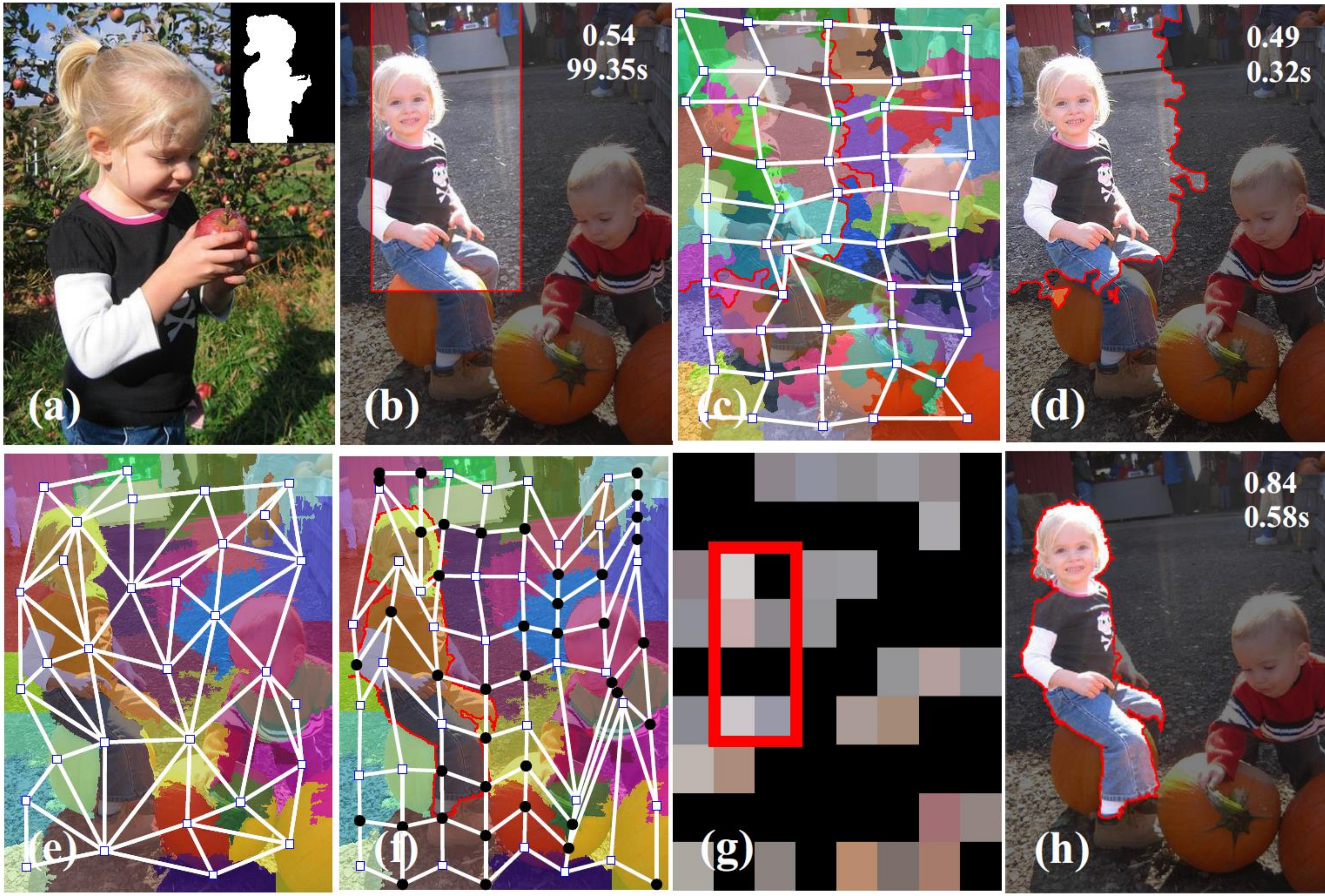
Major Drawback

- Many successful methods available for pixel-level image analysis
  - Integral image [RC Tuzel et al. ECCV06, Integral histogram Porikli CVPR05, etc.]
  - Efficient sub-window search (ESS) [Lampert et al. TPAMI09, ICCV09]
  - Fast optimization [Felzenszwalb et al. TPAMI11]
  - ...
  - Regular grid structure
- But, we cannot use SP as freely as what we do for pixels, due to topological irregularity

Superpixel Gridization

- Objective: To regularize the topology of any SP-graphs into a regular 4 (or 8)-connected grid structure

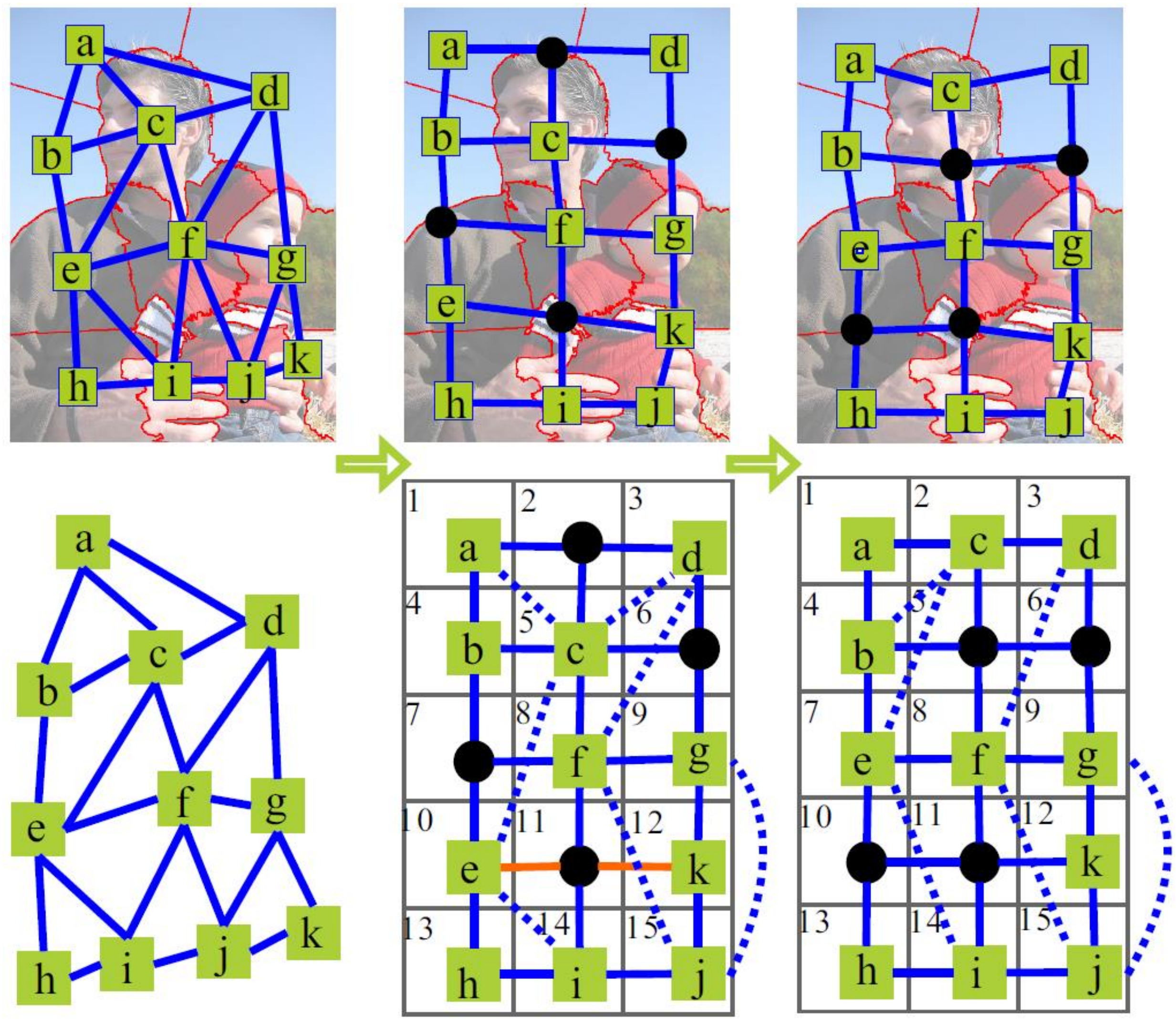
- Advantages:
  - To directly use integral image and ESS on SP-level (very fast and more accurate)
  - Not limited to a particular SP generation method
  - A new angle to preserve spatial regularity in SP segmentation



(a) Query image; (b) Object localization by pixel-level region covariance (RC) searching; (c) & (d) SEEDs SPs and RC searching result; (e) Irregular SLIC SPs; (f)-(h) Regularized SLIC by our method and detected object.

Superpixel Gridization

- Methodology: adding dummy nodes into SP-graphs to regularize them into SP-grids



Superpixel Gridization

- SP Gridization is not an equivalent transform
  - Some information must be lost
  - We only preserve most important image structures

- SP-graph: $G = \langle P_G, W_G \rangle$
  - Real SP nodes: P_G
  - Pairwise SP edge weights: W_G

- SP-grid: $L = \langle P_L, W_L \rangle$
  - $P_L = P_G \cup D$
  - Dummy nodes: D
  - Pairwise SP edge weights: W_L

Criterion

- Maximum cohesive correlation (MCC) [CVPR13]: to maximally preserve the most important d-linked SP-connections
  - Preserving larger d-linked SP-edges and discarding smaller ones in SP-grid

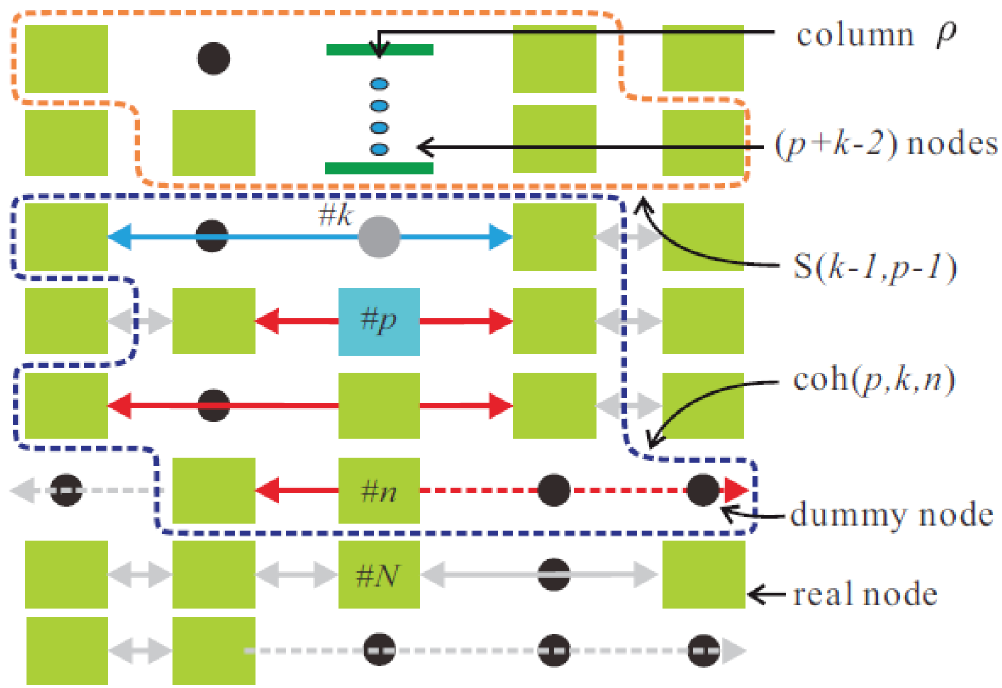
$$\hat{L}_{\mathrm{MCC}} = \arg_L \max \mathrm{Coh}(L|G)$$

$$\mathrm{Coh}(L|G) = \sum_{p \sim q} \mathrm{sim}(p, q)$$

$$\mathrm{sim}(p, q) = \exp\left(-\frac{\left\|o_p - o_q\right\|_2}{w_{\mathrm{pos}}} - \frac{\mathrm{g}(p, q)}{w_{\mathrm{conn}}} + \frac{\mathrm{Ba}\left(H_p, H_q\right)}{w_{\mathrm{app}}}\right)$$

Criterion

- SP nodes *p* and *q* of SP-grid are d-linked, iff:
  - they lie in the same column (or row)
  - all inbetween nodes from *p* to *q* in the column (or row) are all dummy nodes



Criterion

- Minimum topological discrepancy (MTD) [in submission to TPAMI]: to optimally transforming arbitrary given SP-graph G into regular SP-grid L by minimizing their TD,

$$\hat{L}_{\mathrm{MTD}} = \arg_L \max \mathrm{TD}(L|G) = \arg_L \max \ \mathrm{P}(L|G) + \alpha\, \mathrm{S}(L|G)$$

  - $\mathrm{P}(L|G)$: position discrepancy
  - $\mathrm{S}(L|G)$: structure discrepancy

$$\mathrm{S}(L|G) = \sum_{p \sim q} |W_L(p,q) - W_G(p,q)|$$

Criterion

- MCC is a special case of MTD criterion

$$\hat{L}_{\mathrm{MCC}} = \hat{L}_{\mathrm{MTD}}(G_{\mathrm{full}})$$

- Optimization of MCC and MTD are NP hard
  - SP gridization is a grid coordinates assignment problem, the solution space is of exponential size
  - Both MCC and MTD consider d-linked SP pairs
  - d-linked SPs induce high-order terms, whose energy function is very hard to be optimized

Optimization

- Feasible two-step solution
  - SP-grid initialization
  - SP-grid refinement by Cascade Dynamic Programming (CDP)

- Observation: given a specific 2D grid coordinates assignment (SP-grid initialization), for any column (or row), if fixing the order of real SP nodes, the optimal 1D coordinate assignment of dummy nodes can be exactly optimized by DP

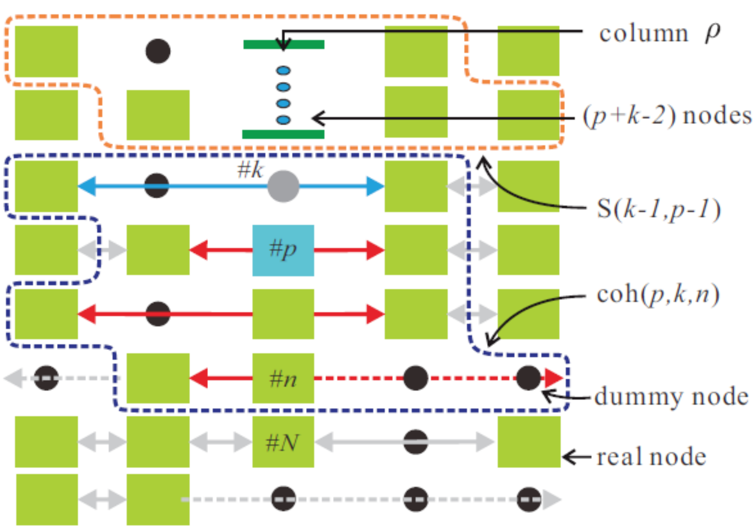
Dummy nodes assignment of a column (row) via DP under MCC criterion

DP process:

$$S(k,n) = \max_{k \le p \le n+1} [S(k-1, p-1) + coh(p,k,n)]$$

$$\mathrm{B_S}(k,n) = \arg\max_{k \le p \le n+1} [\mathrm{S}(k-1, p-1) + \mathrm{coh}(p,k,n)]$$

$$\mathrm{coh}(p,k,n) = \sum_{i=p}^{n} \mathrm{Coh}(i, i_r) + \mathrm{Coh}(i, i_l) + \mathrm{Coh}(k_l, k_r)$$

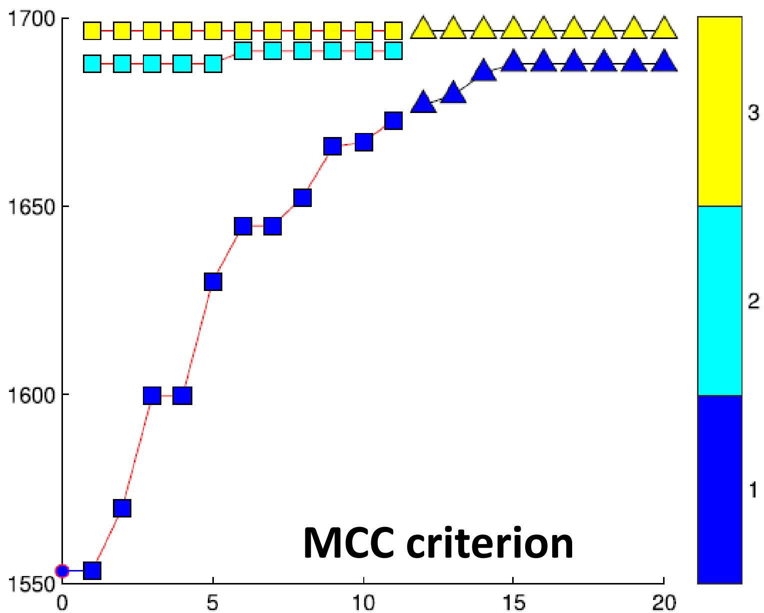


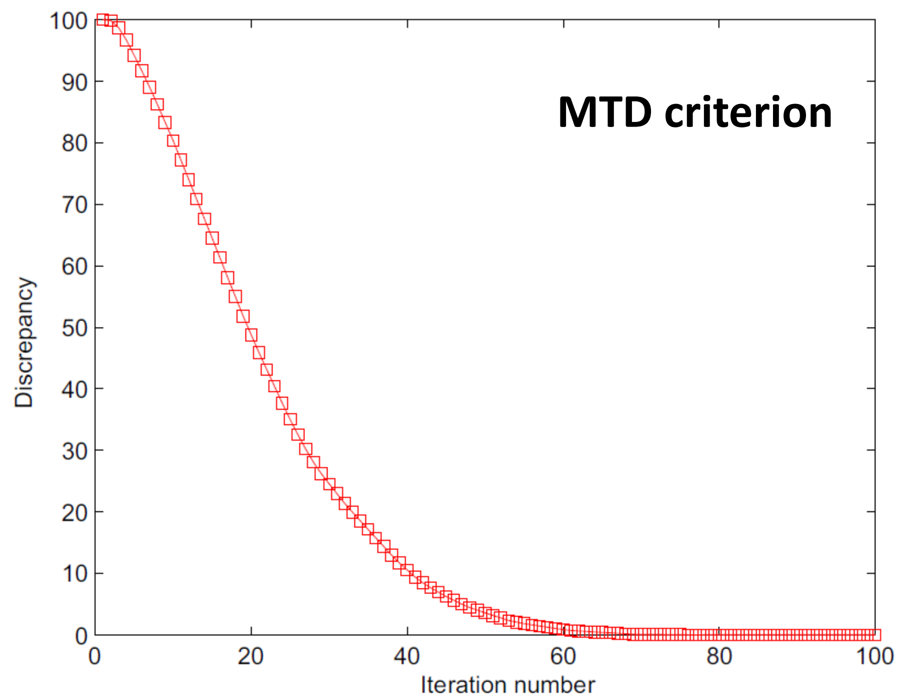
Note:

- $\mathrm{Coh}(p,q) = \mathrm{sim}(p,q)$
- $\mathrm{S}(k,n)$ is the maximum cohesion increment caused by allocating *k* dummy nodes in the first *n* real nodes of current column (row)
- $\mathrm{B_S}(k,n)$ is the optimal 1D position of the #*k* dummy node
- $\mathrm{coh}(p,k,n)$ denotes the extra coherence increment by assigning the #*k* dummy node right in front of the #*p* real node

Optimization

- Cascade DP:
  - Column-by-column optimization
  - Row-by-row optimization
  - Column-row alternative optimization





SP-Grid Initialization

- Dynamic clustering of SP nodes:

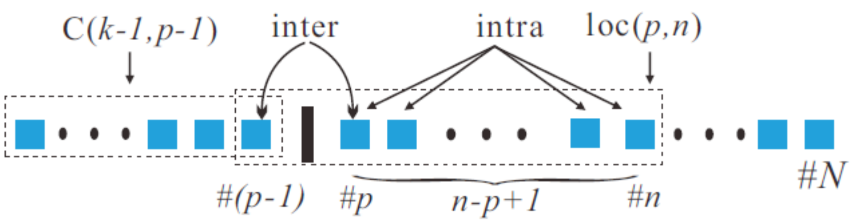
Parition all SPs $\mathcal{P}$ into columns:

1. Sort all SPs $\mathcal{P}$ into 1D $\mathcal{P}_{\mathrm{xsort}}$ with increasing x-coordinates of their centroids.
2. Define column-cuts $\mathcal{B} = \{b_i\}_{i=1}^{c-1}$: partition $\mathcal{P}$ into c columns. *e.g.*: $b_i = p$ means $\#i$ column-cut lies between $\#(p-1)$ and $\#p$ of $\mathcal{P}_{\mathrm{xsort}}$.
3. Seeking $\hat{\mathcal{B}} = \arg\min_{\mathcal{B}} \mathrm{Loc}(\mathcal{B})$, where

$$\mathrm{Loc}(\mathcal{B}) = \sum_{i=1}^{c-1} \mathrm{loc}(b_i, b_{i+1})$$

and

$$\mathrm{loc}(b_i, b_{i+1}) = \frac{\omega_{\mathrm{sep}} \cdot \mathrm{intra}_i}{\max(\mathrm{inter}_i, \epsilon)} + \omega_{\mathrm{len}} \cdot (L_i - \bar{L}).$$



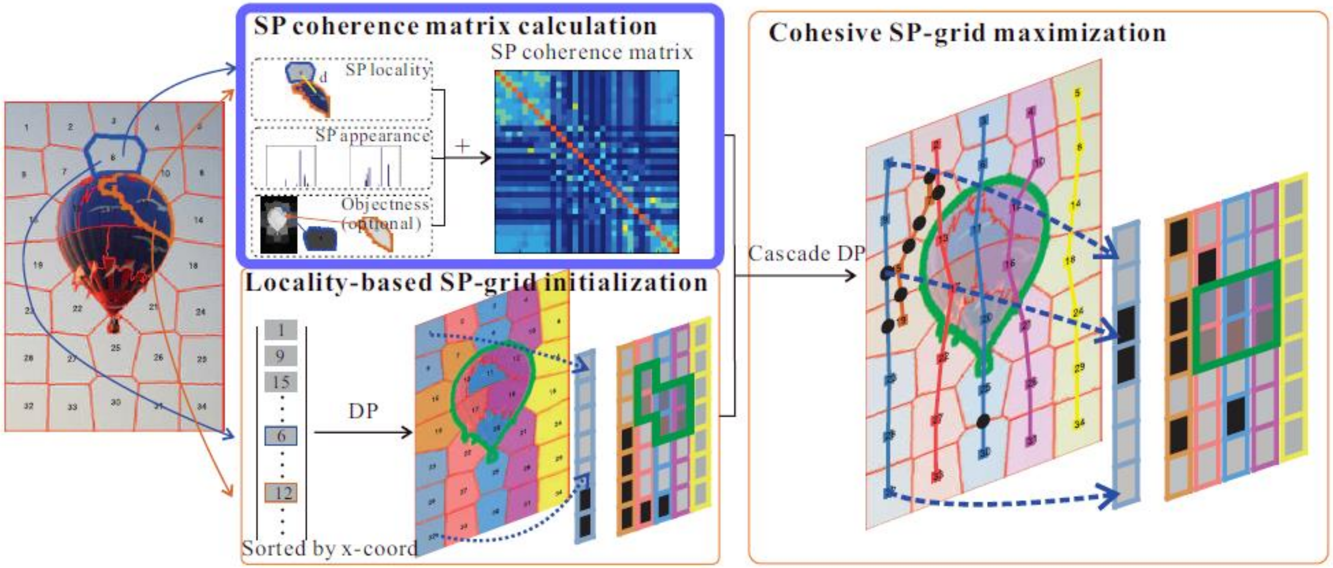
SP-Grid Initialization

- Dynamic clustering of SP nodes:

Seeking $\hat{\mathcal{B}}$ using the following DP process:

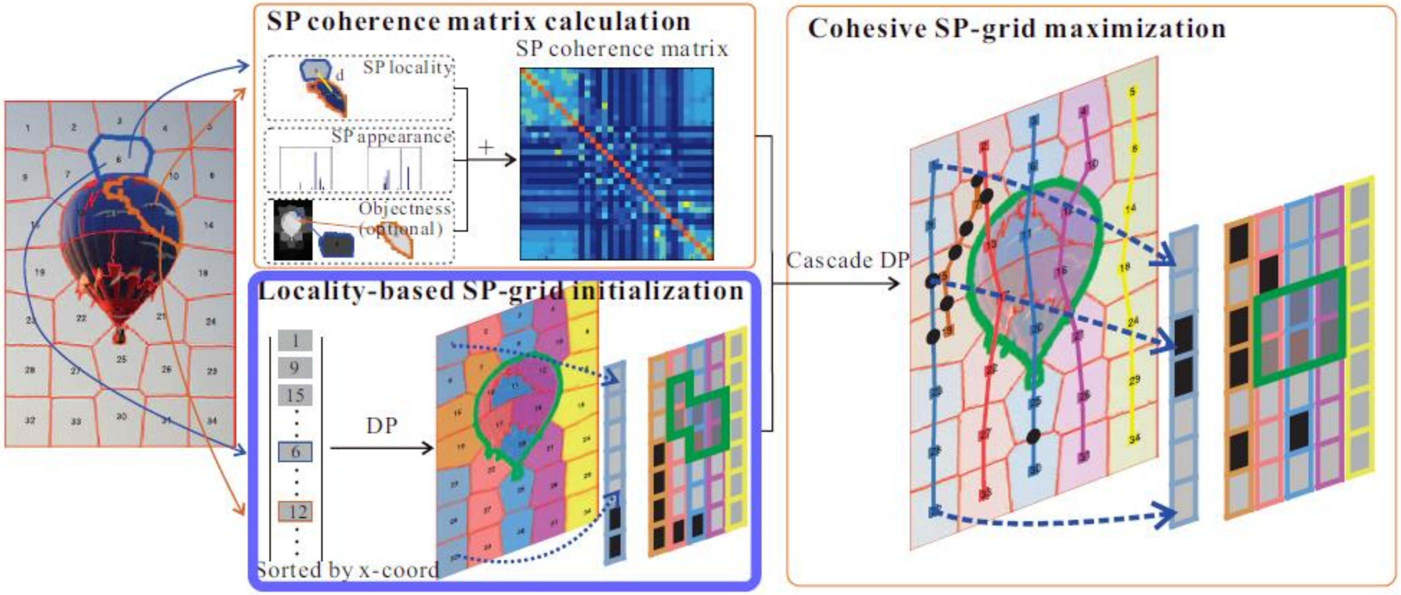
Require: The set of input SPs $\mathcal{P}$
Ensure: Optimal $\hat{\mathcal{B}} = \{\hat{b}_i\}_{i=1}^{c-1}$
//Comp. $\mathrm{C}(k, n)$ and $\mathrm{B_L}(k, n)$
for $n = 1$ to $|\mathcal{P}|$ **do**
  for $k = 2$ to c **do**
    Comp. $\mathrm{C}(k, n)$ and
    $\mathrm{B_L}(k, n)$ using Eqs. (1)–(3);
  end for
end for
// Back retrieving $\hat{\mathcal{B}}$
Set $n := |\mathcal{P}|$ and $\hat{b}_1 = 1$;
for $k = c$ to 2 **do**
  Set $\hat{b}_{k-1} := \mathrm{B_L}(k, n)$;
  Set $n := \hat{b}_{k-1} - 1$;
end for

Example: MCC SP Gridization



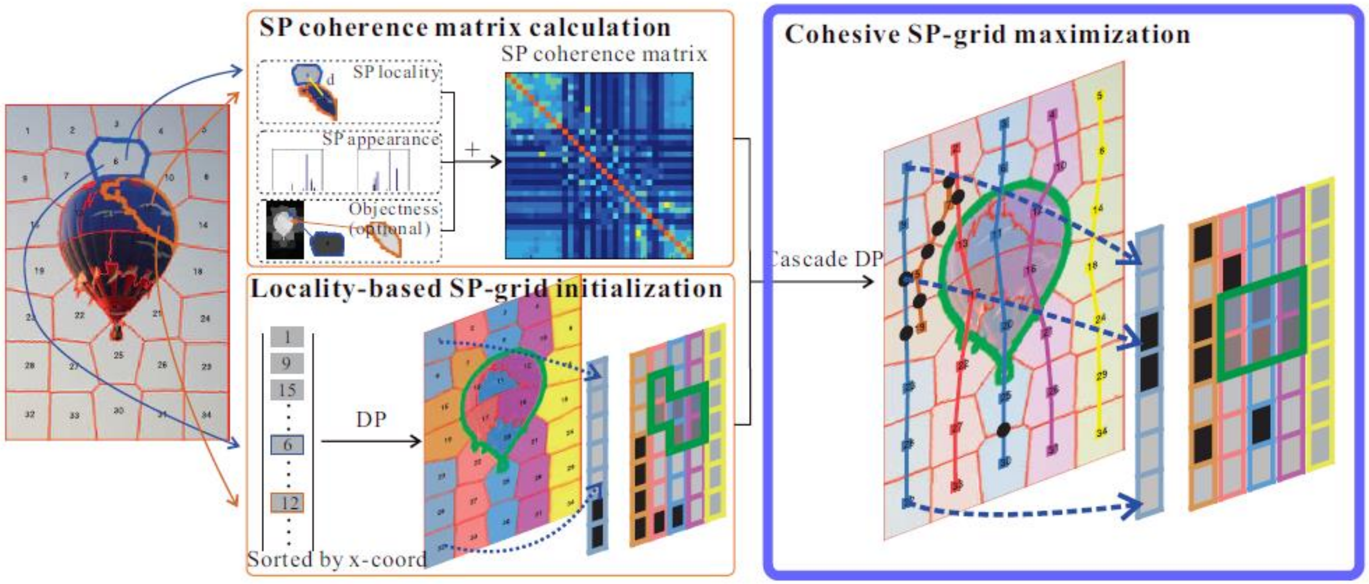
- SP – SP coherence metric: locality, appearance, objectness
- SP-grid initialization: locality-based DP
- Cohesive SP-grid maximization: cascade DP

Example: MCC SP Gridization



- SP – SP coherence metric: locality, appearance, objectness
- SP-grid initialization: locality-based DP
- Cohesive SP-grid maximization: cascade DP

Example: MCC SP Gridization

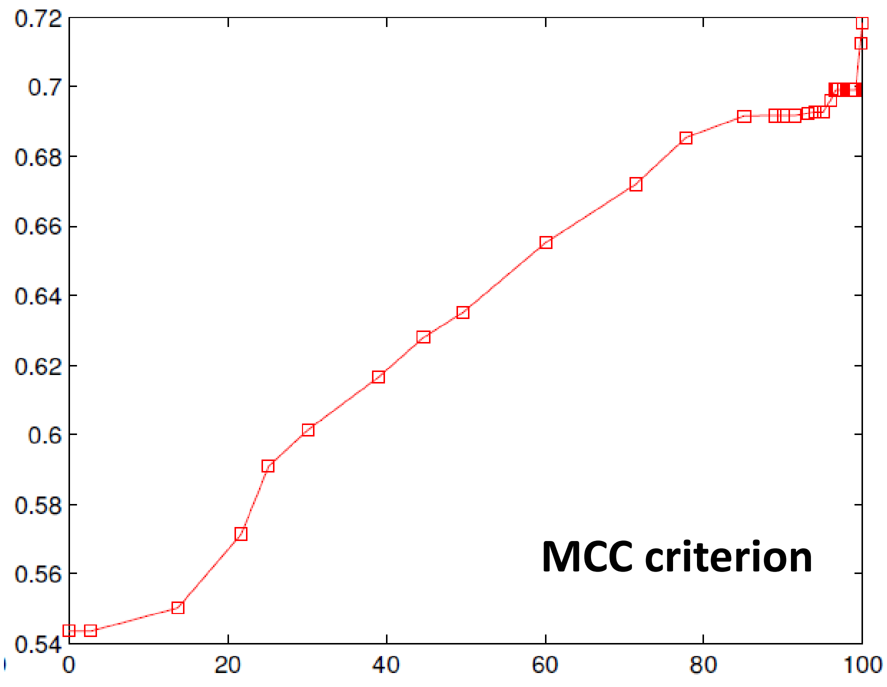


- SP – SP coherence metric: locality, appearance, objectness
- SP-grid initialization: locality-based DP
- Cohesive SP-grid maximization: cascade DP

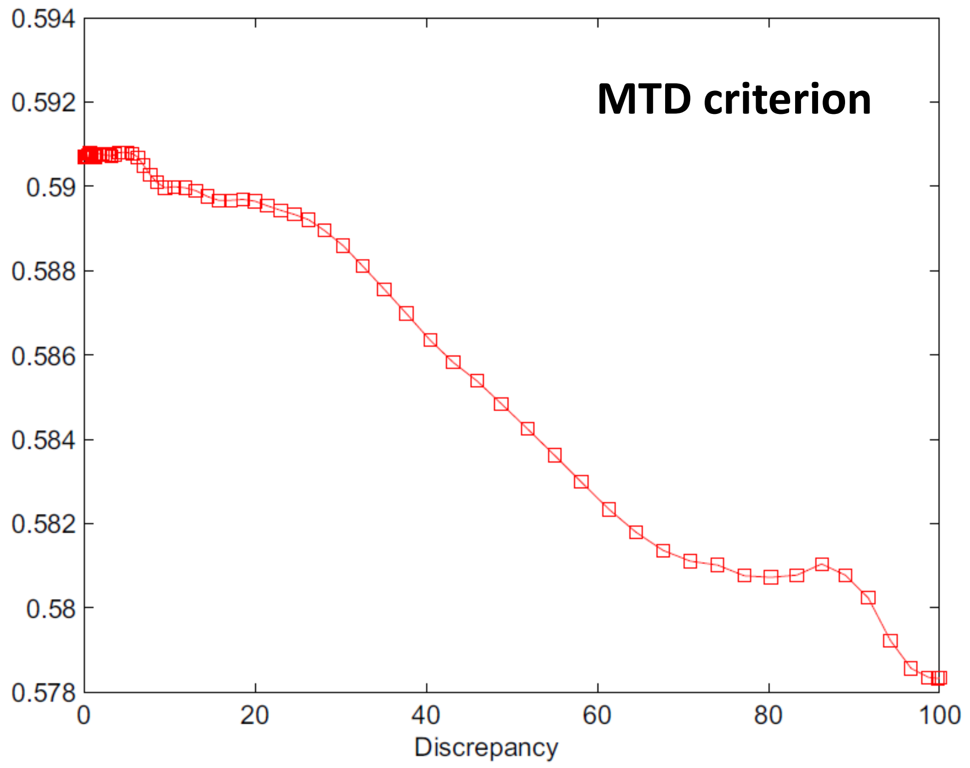
Experiments

- **Object localization**: segment object from a target image using an query object template
- **Dataset**: extended MFC dataset [Kim & Xing, CVPR12] with more than 2200 query cases (50% tuning, 50% testing)
- **Evaluation**: all methods are tuned on tuning set by maximizing F1-measure via DE and compare to each other on testing set
- **Baselines**: STOA regular/near-regular SPs, pixel-level searching
- **Pipeline**: SP generation -> SP gridizatoin -> RC feature extraction -> Exhaustive searching

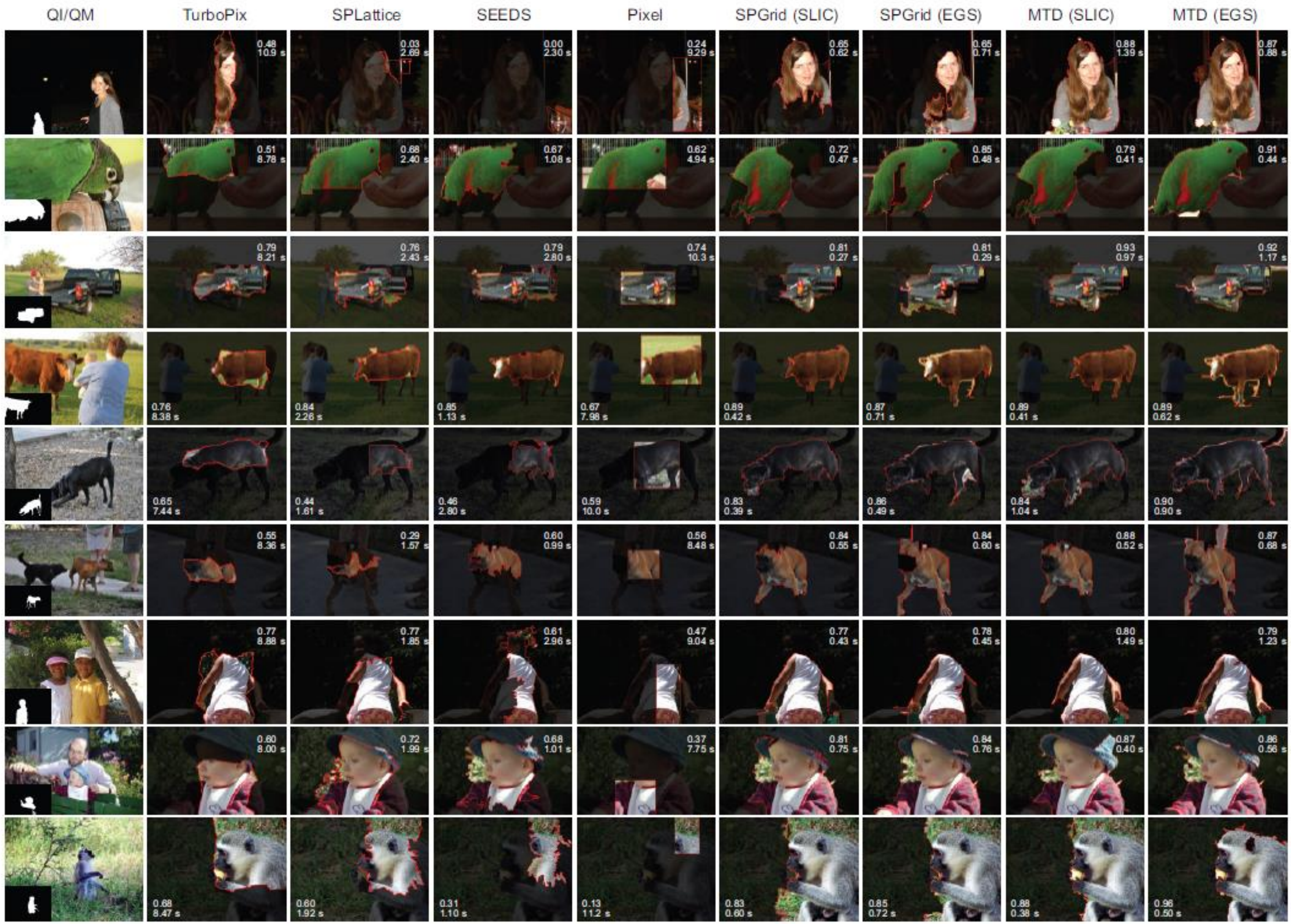
Criterion Validation



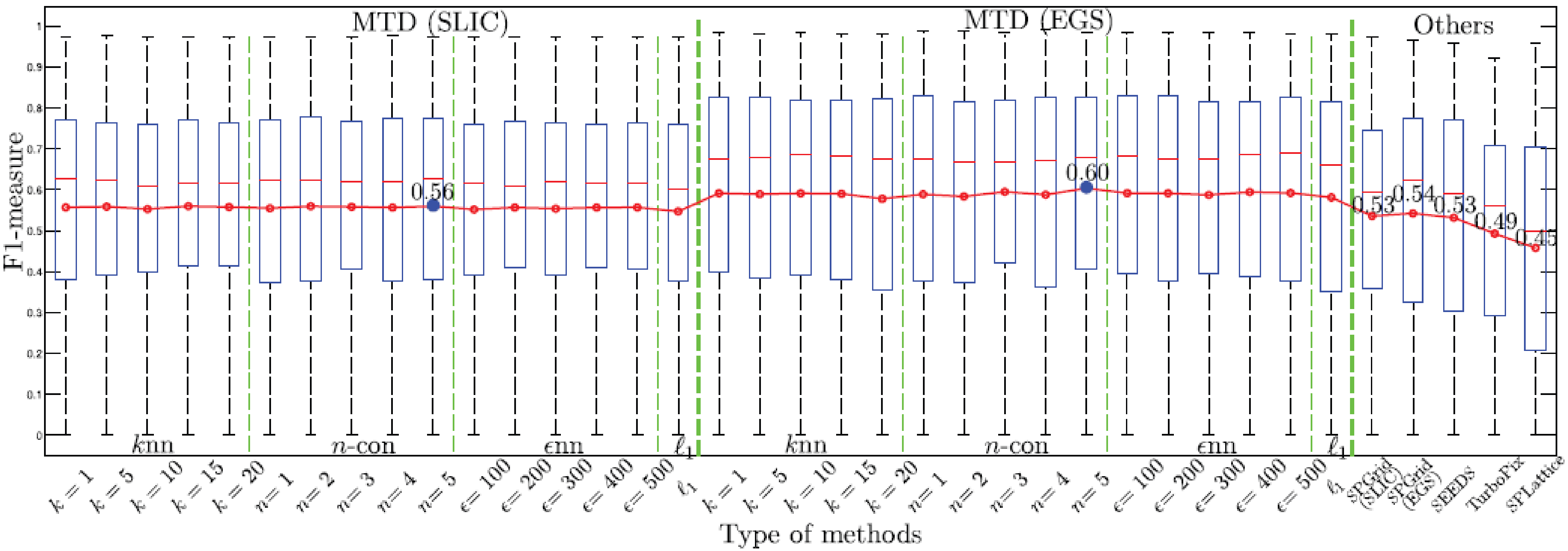
Localization accuracy vs. coherence value



Some Results



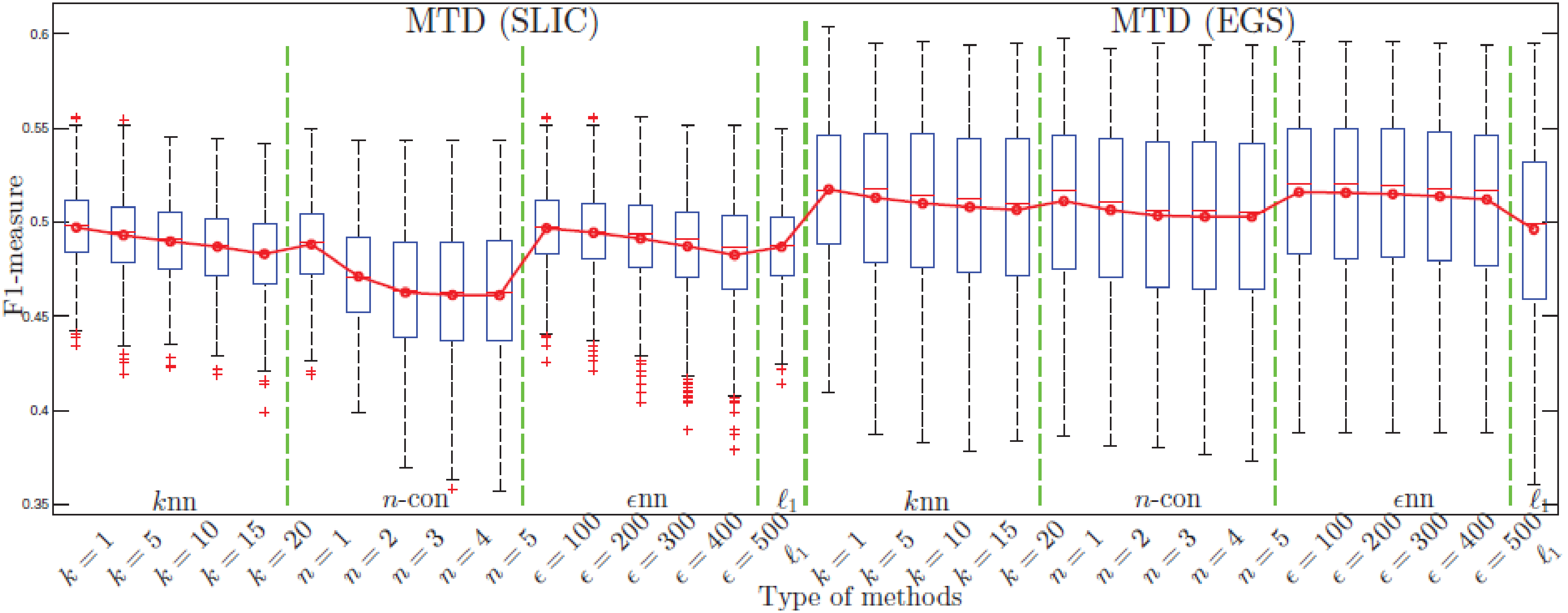
Comparison of Best Performance



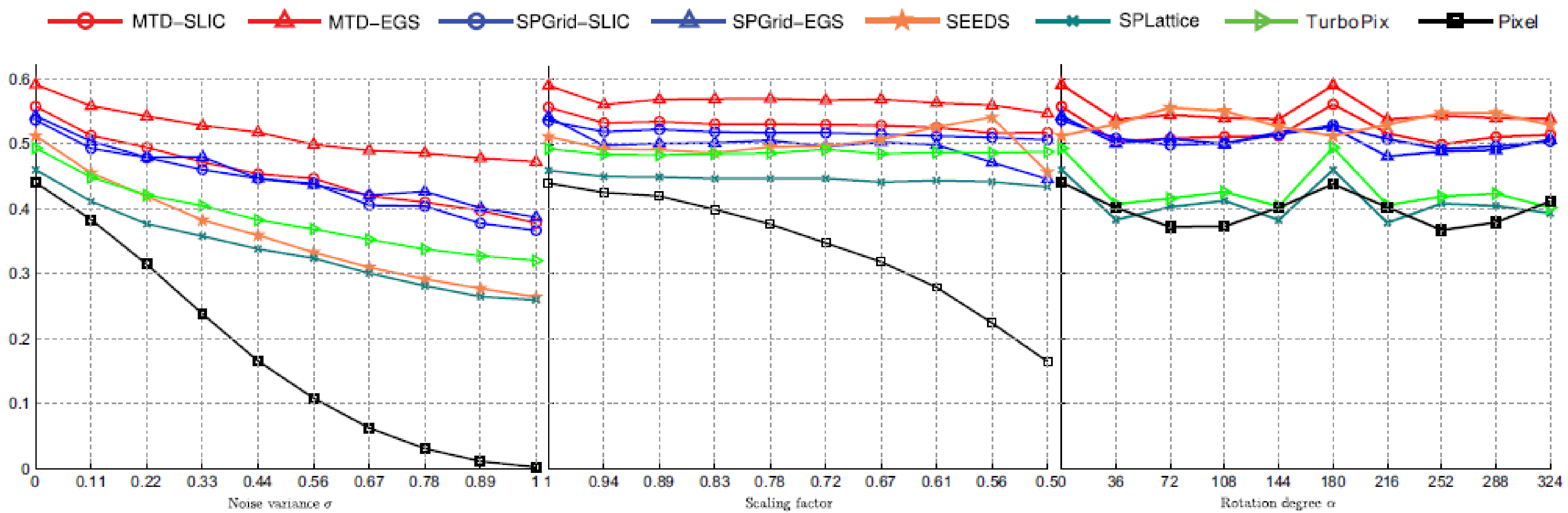
Comparison of Best Performance

Method	Precision	Recall	F_1	Precom-Time(sec)	SP-Time(sec)	Matching-Time(sec)	Total-Time(sec)
MTD $k$nn 3	0.59	0.60	0.59	-	0.06	0.20	0.26
MTD n-con 5	0.63	0.57	0.60	-	0.07	0.24	0.31
MTD ϵnn 400	0.61	0.58	0.59	-	0.06	0.22	0.28
MTD L_1-graph	0.65	0.55	0.58	-	0.06	0.92	0.98
SP-Grid (SLIC)	0.61	0.48	0.53	-	0.16	0.24	0.40
SP-Grid (EGS)	0.64	0.50	0.54	-	0.06	0.25	0.31
SEEDS	0.62	0.50	0.53	-	0.10	0.25	0.35
TurboPixel	0.61	0.44	0.49	-	3.58	0.25	3.83
SPLattice	0.57	0.42	0.45	0.17	0.19	0.29	0.48
Pixel	0.40	0.61	0.42	-	-	7.38	7.38

Results on 1000 Random Parameters



Robustness



MTD-SLIC
MTD-EGS
SPGrid-SLIC
SPGrid-EGS
SEEDS
SPLattice
TurboPix
Pixel
0.6
0.5
0.4
0.3
0.2
0.1
0
0 0.11 0.22 0.33 0.44 0.56 0.67 0.78 0.89 1
Noise variance σ
1 0.94 0.89 0.83 0.78 0.72 0.67 0.61 0.56 0.50
Scaling factor
0 36 72 108 144 180 216 252 288 324
Rotation degree α

Conclusions

- First framework for generic SP gridization
- Practical optimization solution
- With gridized SPs, much faster and more accurate object localization & segmentation can be realized
- Can we directly apply all pixel-level vision methods on SP-level?

Future Work

- Local SP gridization for fast object tracking
- Object proposal generation
- More powerful and theoretically guaranteed optimization solution
- …

- **<u>Code available upon request!</u>**

Thanks!