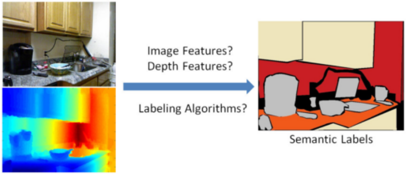
Paper Reading Seminar

Yan Wang

August 30, 2012

RGB-(D) Scene Labeling: Features and Algorithms

- Problem: indoor scene, optical photo + depth image $\Rightarrow$ pixel-wise label



- Evaluation: NYU Depth Dataset (13 categories), Stanford Background Dataset (8 categories, no depth info), Mean AP.

Intuition

- Kernel Descriptor + Efficient Matching Kernel: pixel level features in different domains $\Rightarrow$ superpixel level feature
- Segmentation tree: different scales of superpixel
- Contextual refinement

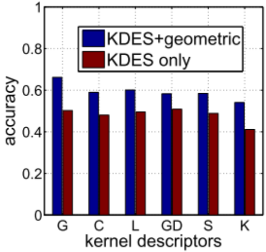
Approach

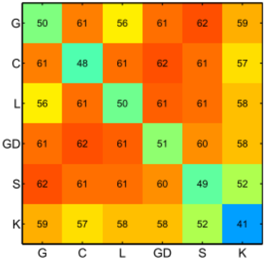
- Segmentation trees
  - gPb: local + global contrast cues ⇒ pixel-level probability-of-boundary map
  - Extend to depth frames
  - Linear fusion for RGB-D frames
- Feature design
  - Gradient, color, local binary pattern, depth gradient, spin/surface normal, KPCA/self-similarity

Approach

- Kernel descriptors
  - Intuition: pixel features $\Rightarrow$ superpixel

$$F^t_{\text{grad}} = \sum_{z \in Z} \tilde{m}_z k_o(\tilde{\theta}_z, p_i) k_s(z, q_j)$$





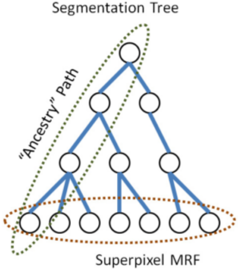
  - Use image gradient + spin/normal

Approach

- Classification
  - Efficient Match Kernel for fixed-length features on superpixels
  - Linear SVM
  - Normalize on superpixel area (A_s)

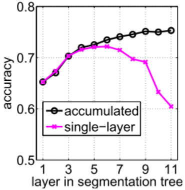
$$A_s / (\sum_{q \in Q_c} A_q)^p$$

- Segmentation tree
  - Different level (t) of segmentation tree $\Leftrightarrow$ different scale of superpixels
  - Tree(s) $= \{f_{t,c}(s_t)\}, t, c$

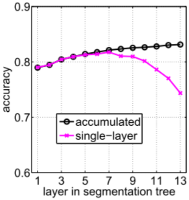


Approach

- Segmentation tree
  - Accumulate features along paths for better accuracy



(a) NYU Depth



(b) Stanford

- Superpixel MRF with gPb
  - Data term: $-f_{c,t}$
  - Smoothing term

$$V_{s,r} = \beta \exp(-\gamma \cdot \mathrm{gPb}_{\mathrm{rgbd}}(s, r))$$

  - Solve with graph-cut