Basics and Hands-on of Computer Vision

Intensive course at University of Helsinki – day 4

Jorma Laaksonen

Aalto University
Department of Computer Science
Espoo

17.5.2018

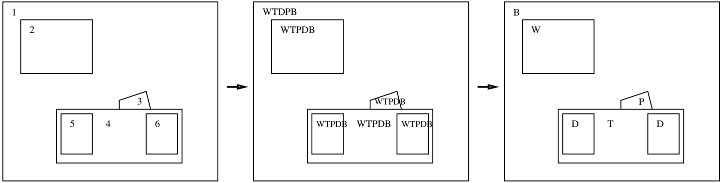
Contents

Contents

14.7 Scene labeling and constraint propagation (10.7/8.5)

- aiming at consistent interpretation of the image
- discrete / probabilistic labeling
- regions, attributes, relations
- regions $R_i,\ i = 1, \cdots, N$, labels $\Omega = \{\omega_1, \cdots, \omega_R\}$
- moving from local constraints to image level
- relaxation in constraint propagation
- discrete relaxation
  - attributes are discrete Boolean values: is / is not
  - first all regions are given all labels
  - impossible labels are removed one by one

Discrete relaxation: example (10.7.1/8.5.1)



a. window (W) is rectangular

b. table (T) is rectangular

c. drawer (D) is rectangular

d. phone (P) is above table

e. drawer is inside table

f. background (B) is adjacent to the border

Probabilistic relaxation (10.7.2/8.5.2)

- produces always some solution
- support for label ω_k in region θ_i at iteration step s:

$$Q^s(\theta_i = \omega_k) = \sum_{j=1}^{N} c_{ij} q_j^s(\theta_i = \omega_k) \quad , \quad \sum_{j=1}^{N} c_{ij} = 1$$

$$= \sum_{j=1}^{N} c_{ij} \sum_{l=1}^{R} r(\theta_i = \omega_k, \theta_j = \omega_l) P^s(\theta_j = \omega_l)$$

- linear relaxation

$$P^0(\theta_i = \omega_k) = P(\theta_i = \omega_k \mid X_i)$$

$$P^{s+1}(\theta_i = \omega_k) = Q^s(\theta_i = \omega_k) \;\; \forall i, k$$

- non-linear relaxation

$$P^{s+1}(\theta_i = \omega_k) = \frac{1}{K} P^s(\theta_i = \omega_k) Q^s(\theta_i = \omega_k)$$

$$K = \sum_{l=1}^{R} P^s(\theta_i = \omega_l) Q^s(\theta_i = \omega_l)$$

Relaxation as optimization problem

Maximization F:

$$F = \sum_{k=1}^{R} \sum_{i=1}^{N} P(\theta_i = \omega_k) \sum_{j=1}^{N} c_{ij} \sum_{l=1}^{R} r(\theta_i = \omega_k, \theta_j = \omega_l) P(\theta_j = \omega_l)$$

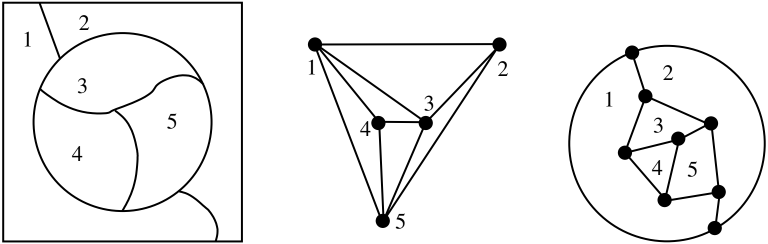
$$\sum_{k=1}^{R} P(\theta_i = \omega_k) = 1 \ \ \forall i, \qquad P(\theta_i = \omega_l) > 0 \ \ \forall i, k$$

Image interpretation as tree search (10.7.3/8.5.3)

- number of image regions = number of layers in search tree
- leaves of the tree correspond to different full image labelings

14.8 Semantic image segmentation (10.8/8.6)

- region adjacency graph and its dual



- iterative updating of data structures
- semantic region growing
- merging of adjacent regions
- aiming at maximizing objective function F
- always the most probable interpretation is fixed

Contents

SIFT, SURF and ORB

- SIFT, SURF and ORB are techniques for interest point detection and description
- interest points are called also keypoints or feature points
- detectors detect details that might be good in distinguishing the image content
- descriptors produce fixed-length feature vectors that describe a local area around the keypoint
- illumination, rotation and scale invariance in both detection and description

SIFT, SURF and ORB

- can be used for 1-to-many and 1-to-1 point matching
- can be used for object tracking in videos
- can be collected in bag-of-words (BoW) histogram features
- BoW requires creation of a codebook or reference samples for 1-NN mapping
- descriptors can be used without detectors in dense sampling
- similarly, detectors can be used alone

Contents

Hands-on work and home assignment

- Instructions in Google Drive
- Also hands-on work can be done at home
- Report all code and images of the hands-on and the home assignment in the same PDF
- Report also how long time the assignments took
- Email the PDF before midnight to the lecturer