SEMANTIC IMAGE SEGMENTATION USING THE ICM ALGORITHM

Lazhar Khelifi & Max Mignotte
DIRO, University of Montreal
{*khelifil, mignotte*}*@iro.umontreal.ca*

Introduction: Semantic image segmentation

- Semantic image segmentation has recently become the focus of considerable interest
- This task consists in assigning a predefined class label to each pixel in an image
- Task with a complexity challenge

Methods

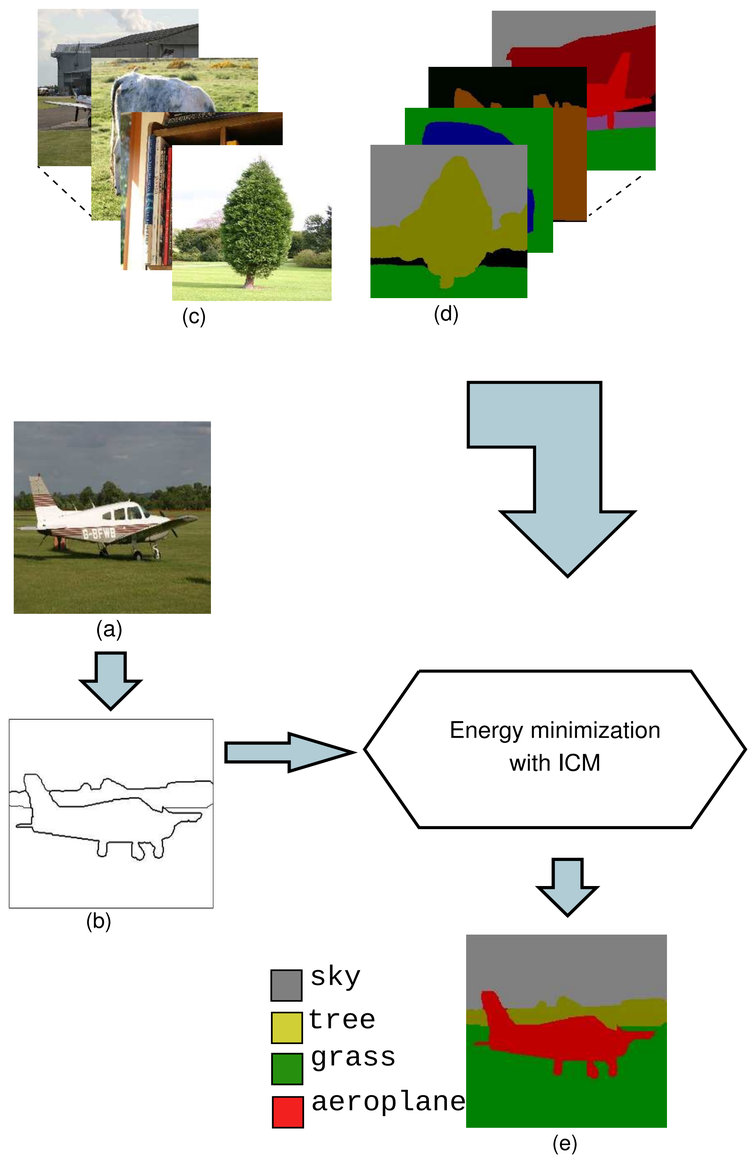
Various methods for scene parsing have been proposed in the literature

- Classifying each pixel independently
  ⇨ High computational cost
- Pairwise Markov random field (MRF)
  ⇨ Learning and inference are often expensive

Key idea

Grouping pixels into super-pixels
+
A novel and simple energy-minimization model

Proposed system overview



(a) Input image
(b) Image segmentation (achieved by the GCEBFM algorithm [1])
(c) Corpus of images
(d) Semantic segmentation of each image on the corpus
(e) Scene parsing result

Combined criteria used in our model

Criterion	Name	Dimension
Color	Color histogram	125
Texture	Oriented gradient histogram	40
Shape	Pixel area	1
Location	Top height	1
Context	Context histogram	21

Our MC-SSM model

- A single class label is assigned to each region by optimizing a global fitness function that measures the quality of the generated solution:

$$\hat{S}_{\text{MC}} = \arg\min_{S \in \mathcal{S}_\Omega} \overline{\text{MC}}\left(I, R_I, S, \{\mathcal{I}_k, \mathcal{S}_k\}_{k \leq K}\right) \quad (1)$$

with:

$$\begin{aligned}\overline{\text{MC}}\left(I, R_I, S, \{\mathcal{I}_k, \mathcal{S}_k\}_{k \leq K}\right) = \quad (2)\\ \alpha \sum_{i=1}^{m} \text{COL}\left(I, r_I^i, s_I^i, \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}\right)\\ + \beta \sum_{i=1}^{m} \text{TEX}\left(I, r_I^i, s_I^i, \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}\right)\\ + \gamma \sum_{i=1}^{m} \text{SHA}\left(r_I^i, s_I^i, \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}\right)\\ + \delta \sum_{i=1}^{m} \text{LOC}\left(r_I^i, s_I^i, \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}\right)\\ + \lambda \sum_{i=1}^{m} \frac{1}{h}\left\{\sum \text{CTX}\left(r_I^m, s_I^i, \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}\right)\right\}\end{aligned}$$

where:

I: Input image

$R_I = \{r_I^1, r_I^2, \ldots, r_I^m\}$: Region segmentation of I (to be semantically labeled)

m: Number of regions (r) in R_I

$\{\mathcal{I}_k, \mathcal{S}_k\}_{k \leq K}$: A set (or a training corpus) of K images $\mathcal{I}_k$ and their corresponding semantic segmentations $\mathcal{S}_k$, respectively

$\mathcal{S}_\Omega$: Set of all possible semantically labeled segmentation maps of I (based on its partition into regions R_I

α, β, γ, δ, λ: Weights of the different terms of the energy function

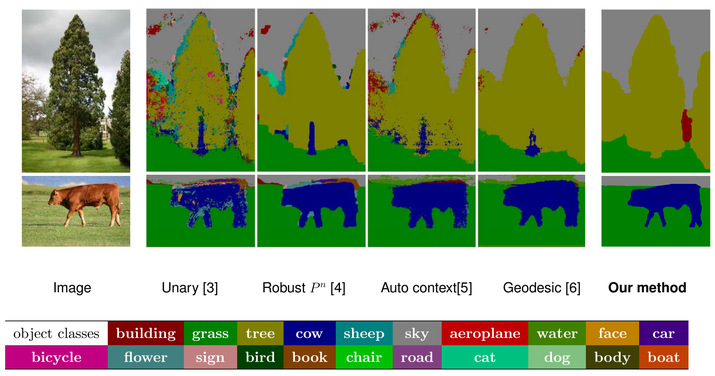
$\{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}$: Set of images and associated semantic segmentation solutions (belonging to the training corpus) that contains a region semantically labeled s_I^i and h is the total number of those semantic segmentations in the corpus $\{\mathcal{C}\}^{s_I^i} = \{\mathcal{I}_k, \mathcal{S}_k\}^{s_I^i}$

Optimization

- Global optimization problem incorporating a nonlinear objective function
- Algorithm: Iterative Conditional Modes (ICM)[2]

Results

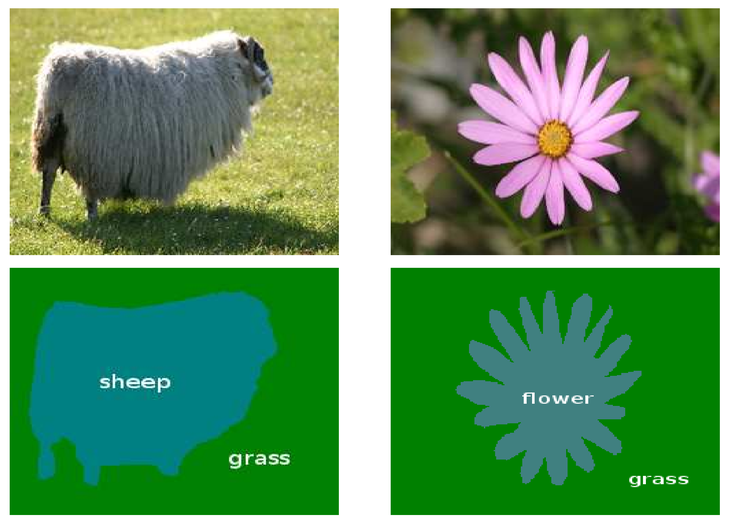
Example of segmentation results on two images from the MSRC-21 dataset compared to other algorithms



Global accuracy on the MSRC-21 dataset

Algorithms	Accuracy (%)
SuperParsing [7] in [8]	61.50
SIM [9]	69.70
SVM-BoW [10]	62.70
MC-SSM	66.07

Example results



References

[1] L. Khelifi,M. Mignotte. A novel fusion approach based on the global consistency criterion to fusing multiple segmentations. *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, Accepted, 2016.

[2] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society.*, B-48:259–302, 1986.

[3] L. Ladicky, C. Russell, P. Kohli and P.H.S. Torr. Associative hierarchical random fields. *EEE Trans. Pattern Anal. Mach. Intell.*,vol. 36, no. 6, pp. 1056–1077, 2013.

[4] P. Kohli, L. Ladicky and P.H.S. Torr. Robust higher order potentials for enforcing label consistency. *Int. J. Comput. Vis.*, vol. 82, no. 3, pp. 302–324, 2009.

[5] Z. Tu and X. Bai. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 10, pp. 1744–1757, 2010.

[6] V. Haltakov, C. Unger and S. Ilic. Geodesic pixel neighborhoods for 2D and 3D scene understanding. *Comput. Vis. Image Underst.*, vol. 148, pp. 164–180, 2016.

[7] J. Tighe and S. Lazebnik. SuperParsing: Scalable Nonparametric Image Parsing with Superpixels. *in proc. European Conf. Computer Vision (ECCV), 2010.*

[8] A. Bassiouny and M. El-Saban. Semantic segmentation as image representation for scene recognition. *in proc. IEEE Int. Conf. Image Process. (ICIP)*, 2014, pp. 981–985.

[9] F. Briggs, X. Z. Fern, and R. Raich. Rank-loss support instance machines for MIML instance annotation. *in proc. 18th ACM Int Conf. on Knowledge Discovery and Data mining (SIGKDD)*, 2012, pp. 534–542.

[10] F. Liu, G. Lin, and C. Shen. CRF learning with CNN features for image segmentation. *Pattern Recognit.*, vol. 48, no. 10, pp. 2983–2992, 2015.