Markov Random Fields

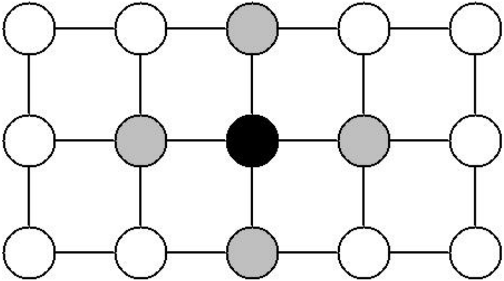
Fabio A. González Ph.D.

Depto. de Ing. de Sistemas e Industrial
Universidad Nacional de Colombia, Bogotá

September 7, 2011

Markov Property

A Markov Random Field (MRF) is a graph, (V, Ed), where each graph node, $l_i \in V$, corresponds to a random variable.

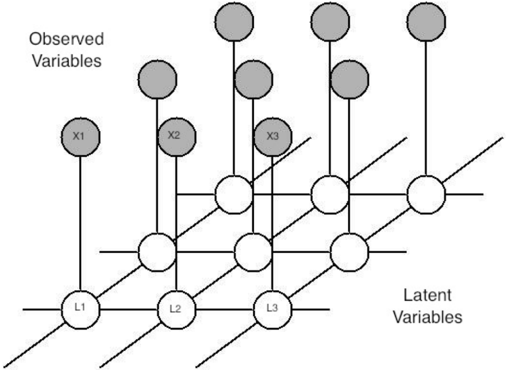


Locality/Markov property: A node (random variable) is independent of the other non-neighbor nodes given its neighbors:

$$P(l_i|V\backslash l_i) = P(l_i|\mathcal{N}_i), \forall i \in V$$

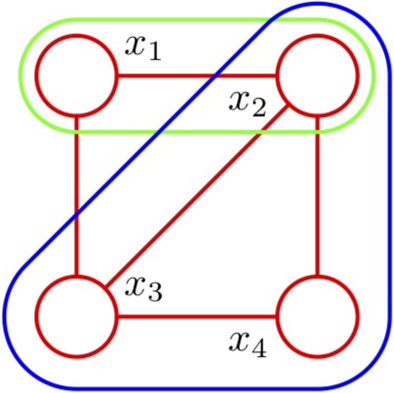
Inference Problem

$$\max_L P(L, X) = \max_L P(l_1, \ldots, l_n, x_1, \ldots, x_n)$$



Factorization Property

$$p(V) = \frac{1}{Z}\prod_{C}\psi_C(V_C)$$



Energy Function

$$\psi_C(C) = e^{-E_C(V_C)}$$

$$p(V) = \frac{1}{Z} e^{-E(V)} = \frac{1}{Z} e^{-\sum_C E_C(V_c)}$$

$$\begin{aligned}\max_L \arg P(L, X) &= \max_L \arg \frac{1}{Z} e^{-E(L.X)} \\ &= \min_L \arg E(L.X)\end{aligned}$$

Optimization of the Energy Function

- Gibbs sampling: Random samples from a probability distribution. (1984)
- Simulated annealing: MAP solution (1984)
- Iterated conditional modes (ICM) (1986)
- Loopy Belief propagation (2001)
- Graph cuts (2001)
- **Max-sum Algorithm**

Semantic Image Segmentation

- General goal: to interpret the image content
- Specific goal: to assign **semantic categories** to some image regions

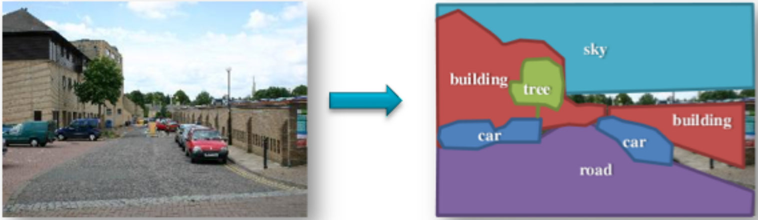


Figure: Understanding the image

Global Semantic Segmentation Process

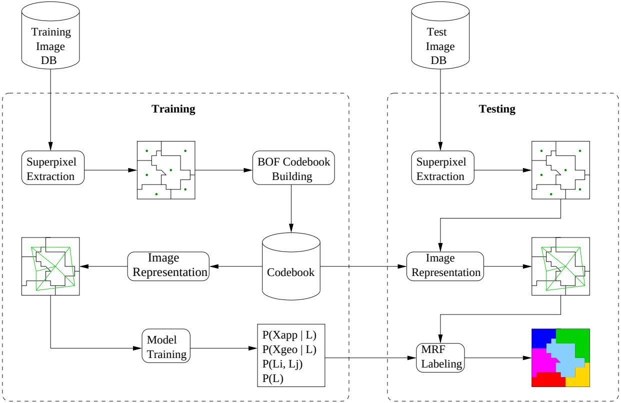
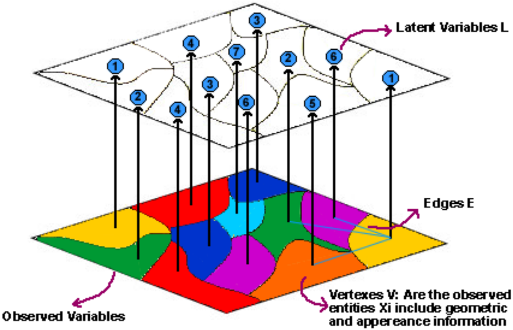


Figure: Semantic segmentation process.

MRF Model (I)



Prediction: given a set of observations to infer the most probable assignations for the latent variables

$$\max_L P(L|X) = \max_L P(l_1, \ldots, l_n | x_1, \ldots, x_n)$$

MRF Model (II)

$$P(L|X) = \frac{P(X|L)P(L)}{P(X)} = \frac{P(X^{app}|L)P(X^{geom}|L)P(L)}{P(X)}$$

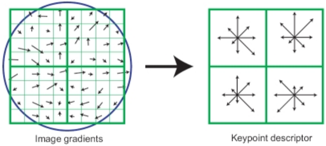
$$E(L) = \alpha E_{app}(L) + \delta E_{geom}(L) + \beta E_{edge}(L) + \gamma E_{prior}(L),$$

where:

- $E_{app}(L) = -\sum_{l_i \in V} \log P(x_i^{app}|l_i)$
- $E_{geom}(L) = -\sum_{l_i \in V} \log P(x_i^{geom}|l_i)$
- $E_{edge}(L) = -\sum_{(i,j) \in Ed} \log P(l_i, l_j)$
- $E_{prior}(L) = -\sum_{l_i \in V} \log P(l_i)$

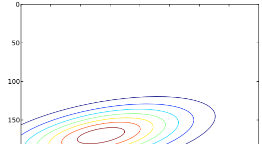
Appearance Information

- Three types of descriptors SIFT, SURFT and DCT were tested
- SIFT worked better
- Three color components were added

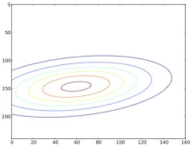


Geometric 2D Information (I)

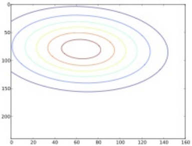
- Coordinate of the morphological center of each superpixel was recorded
- A probability distribution was estimated independently for each label class (Bivariate Gaussian)
- The vertical axis symmetry of the images was exploited to make the estimation problem easier



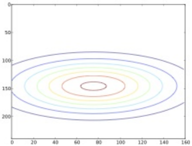
Geometric 2D Information (II)



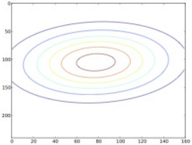
(a) Class Bicyclist



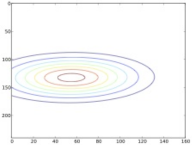
(b) Class Building



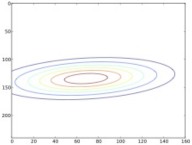
(c) Class Car



(d) Class Column_Pole

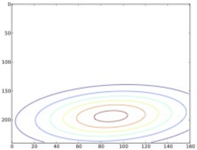


(e) Class Fence

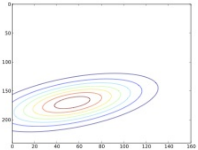


(f) Class Pedestrian

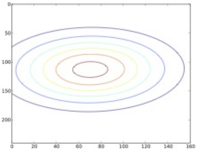
Geometric 2D Information (III)



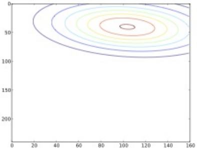
(g) Class Road



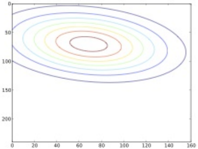
(h) Class Sidewalk



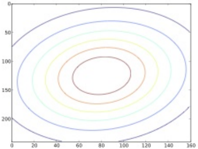
(i) Class SignSymbol



(j) Class Sky



(k) Class Tree



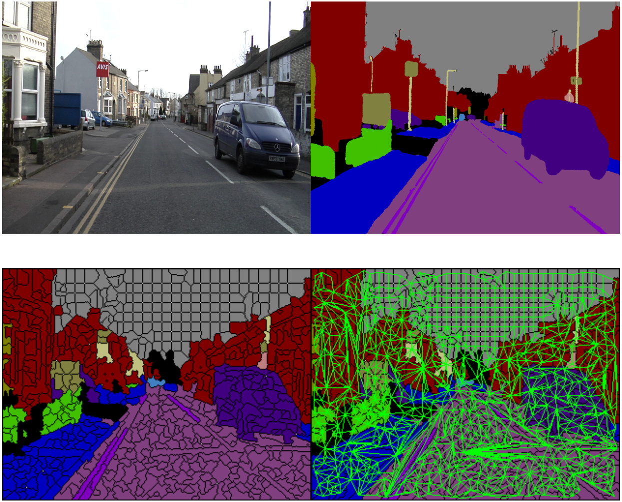
(l) Class Void

Data Set

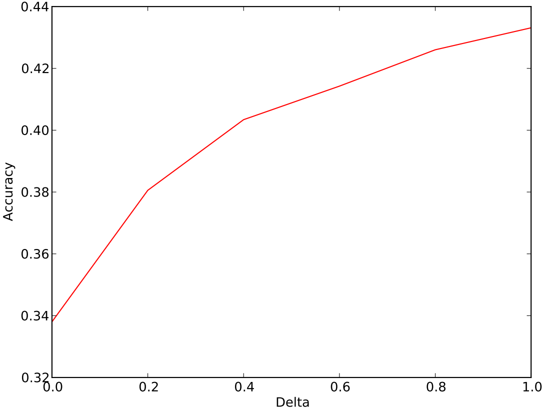
The Cambridge-driving Labeled Video Database (CadVid) is a collection of 701 labeled images (with dimension of 960 x 720 px) that associates each pixel with one of 32 semantic classes.

Figure: Image data set examples

Preprocessing



Importance of Geometric Information



Parameter Tuning

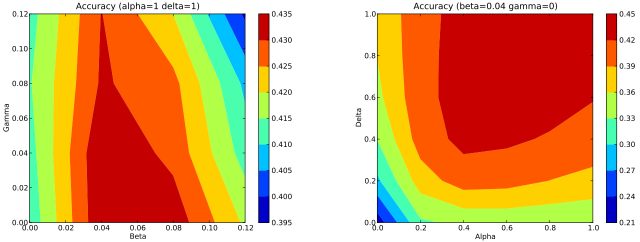
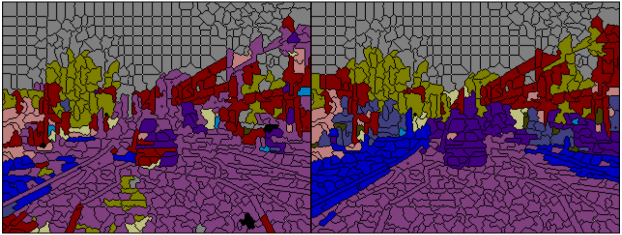


Figure: Parameter tuning on test images: left, alpha and delta parameters are set to $\alpha = \delta = 1.0$; right, beta and gamma are set to best values found ($\beta = 0.04$, $\gamma = 0$).

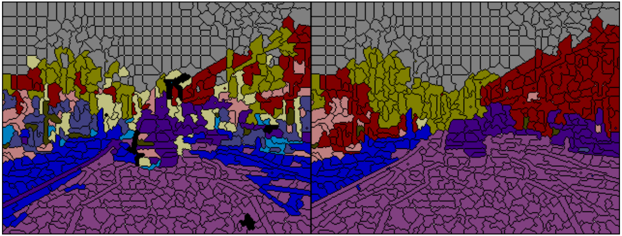
Example Segmentation (I)

$\delta = 0$ vs $\delta = 0.8$



Example Segmentation (II)

$\beta = 0$ vs $\beta = 0.12$



Experimental Results

Method	Kind	bicyclist	building	car	fence
MRF Geo+ App	2D	14.3%	55.4%	53.9%	32.50%
MRF App	-	9.8%	52.5%	37.7%	2.4%
SVM	-	1.26%	72.17%	35.15%	0.28%
co-occ&wLg Spx	3D	28.8%	71.71%	76.5%	4.8%

Method	Kind	road	sidewalk	sky	Average
MRF Geo+ App	2D	84.8%	55.1%	92%	43.09%
MRF App	-	84.5%	22.8%	93%	33.29%
SVM	-	76.76%	8.49%	84.81%	28%
co-occ&wLg Spx	3D	88.4%	84.7%	89.5%	53%

Conclusion

Geometrical Information Improves the model!