Multiscale Conditional Random Fields for Image Labeling

Xuming He, Richard S. Zemel, and Miguel Á. Carreira-Perpiñán

Presented by: Andrew F. Dreher
CS 395T - Spring 2007

Contributions

1) Generalization of conditional random fields (CRF) to multiscale conditional random fields (mCRF)
2) Learning features of the random field at multiple scales

Motivation

1) Segment and recognize each part by class
   Useful for database queries
2) Retain contextual information
   a) Local regions have ambiguity; using neighboring regions can aid in accurate labeling
   b) Limited geometric relationships
      Fish in water; airplanes in sky
      Sky at top of image; water at bottom

Differences from Earlier Methods

1) Discriminative, not generative
2) Uses multiple scales
   a) Locality is a major problem for Markov random fields
   b) Limitedly solved by Hierarchical Markov random fields
3) Does not require joint probabilities

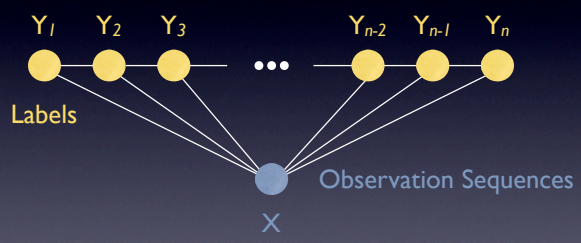
Conditional Random Fields and Restricted Boltzmann Machines

Conditional Random Field

1) Probabilistic framework for labeling, parsing, or segmenting structured data
2) Uses a conditional distribution over label sequences given an observation sequence, not the joint distribution over label and observation sequences.

More: Hanna M. Wallach (http://www.inference.phy.cam.ac.uk/hmw26/crf/)

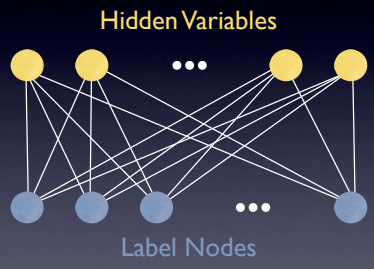
Conditional Random Field



Restricted Boltzmann Machine

1) Type of simulated annealing stochastic recurrent neural network
   Invented by G. Hinton and T. Sejnowski
2) Does not allow connections between hidden nodes
3) Can be organized into multiple layers
   Example: Handwritten digit recognition

Restricted Boltzmann Machine



Multiscale Conditional Random Fields

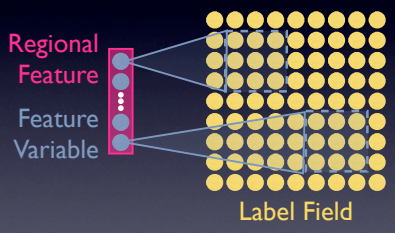
Local Features

1) Classify site using a statistical classifier
2) Limited performance due to noise, class overlap, etc.
3) This looks much like the standard conditional random field diagram

Regional Features

1) Represent geometric relationships between objects
   Corners
   Edges
   T-Junctions
2) Separate hidden variables; shared conditional probability table with other regions

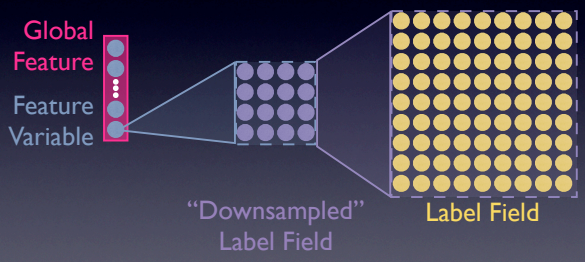
Regional Features



Global Features

1) Either whole image or large local patches
2) Like region, specifies a joint distribution over the labels given the hidden variables
3) Specifies a multinomial distribution over each label node by their parameters

Global Features



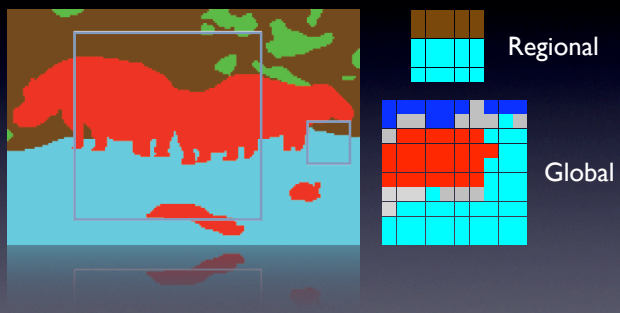
Example

X_i

Rhino / Hippo
Polar Bear
Water
Snow
Vegetation
Ground
Sky

l_i

Example



Combining Components

1) Probability distributions are combined multiplicatively
2) Many unconfident, but similar predictions, can yield a confident prediction
3) Should behave like a cascade; components should focus on aspects where previous components fail

Image Labeling

Image Labeling

1) Given a new image, what is the optimal label configuration?
2) Paper uses maximal posterior marginals
   Minimizes the expected number of mislabeled sites
3) Alternative: maximum a posteriori
   Difficult to compute for high dimensional and discrete domains

Experiments

Data Sets

1) Corel images of African and Arctic Wildlife
   100 images (60 training / 40 test)
   Image size: 180 x 120 pixels
2) Sowerby Image of British Aerospace
   Color scenes of rural & suburban roads
   104 images (60 training / 44 test)
   Image size: 96 x 64 pixels

Image Statistics (X_i)

30 image statistics per pixel

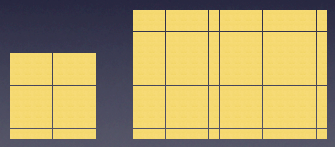
1) Color: CIE colorspace
2) Edge & Texture
   a) Difference-of-Gaussian (3 scales)
   b) Quadrature pairs of even-symmetric and odd-symmetric filters (3 scales; 4 orientations)
      Orientations: 0, $\pi/4$, $\pi/2$, $3\pi/4$

Performance Evaluation

1) Compare against generative method (Markov random field)

Corel Dataset

1) Local features: 3-layer multilayer perceptron with 80 hidden nodes
2) Regional features: 8x8 patch; 30 total
3) Global features: 18x12 patch; 15 total



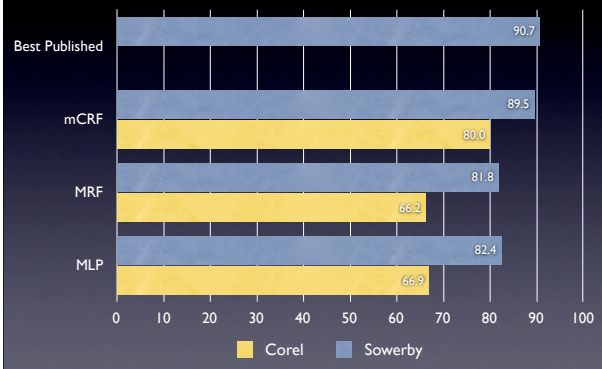
25

Sowerby Dataset

1) Local features: 3-layer multilayer perceptron with 50 hidden nodes
2) Regional features: 6x4 patch; 20 total
3) Global features: 8x8 patch; 10 total

26

Classification Rates



27

Corel Confusion Matrix

	Rhino/ Hippo	Polar Bear	Water	Snow	Vegetation	Ground	Sky
Rhino/ Hippo	9.27	0.14	0.53	0.01	1.01	1.00	0.00
Polar Bear	0.08	8.06	0.01	0.52	0.12	0.63	0.00
Water	0.33	0.00	12.87	0.00	0.42	0.76	0.05
Snow	0.00	0.82	0.00	12.83	0.23	0.09	0.04
Vegetation	0.95	0.55	0.09	3.18	15.06	2.99	0.06
Ground	1.13	1.18	1.11	0.26	1.56	21.19	0.00
Sky	0.00	0.00	0.00	0.00	0.19	0.01	0.66

28

Sowerby Confusion Matrix

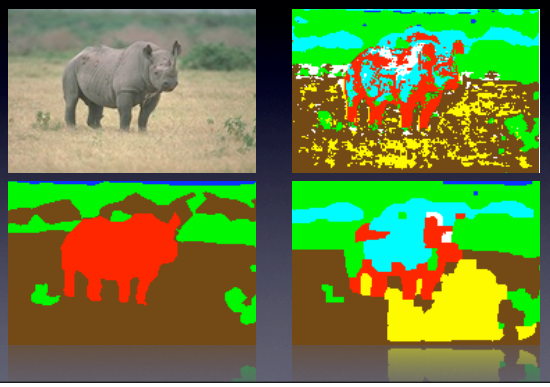
	Sky	Vegetation	Road Markings	Road Surface	Building	Street Objects	Cars
Sky	12.01	0.53	0.00	0.01	0.03	0.00	0.01
Vegetation	0.83	33.39	0.01	1.41	2.71	0.03	0.09
Road Markings	0.00	0.00	0.08	0.10	0.00	0.00	0.00
Road Surface	0.01	0.94	0.02	40.33	0.10	0.01	0.05
Building	0.06	2.60	0.02	0.30	3.05	0.01	0.05
Street Objects	0.02	0.25	0.00	0.03	0.12	0.02	0.01
Cars	0.02	0.27	0.00	0.09	0.24	0.00	0.14

29

Pictorial Results

30

Select Rhino

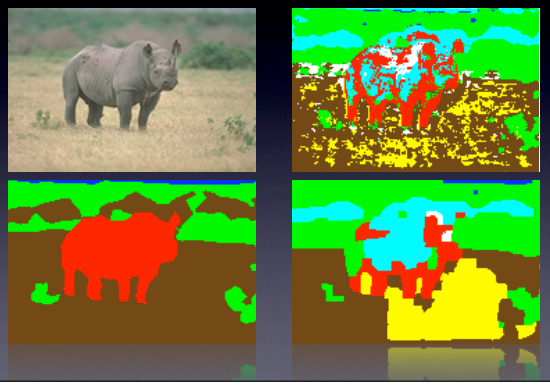


31

Select Rhino

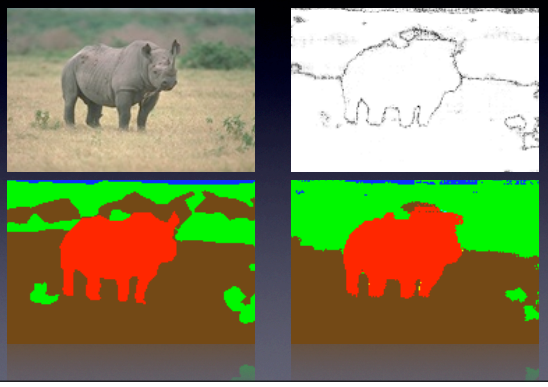
31

Select Rhino



31

Select Rhino

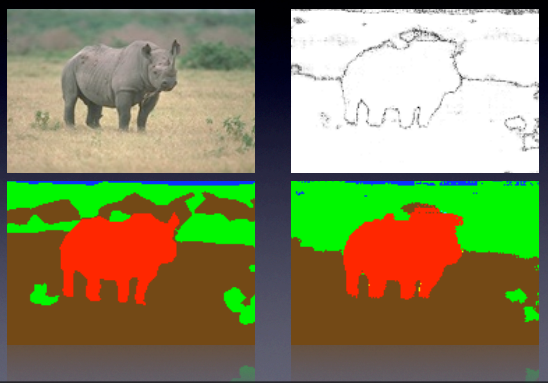


32

Select Rhino

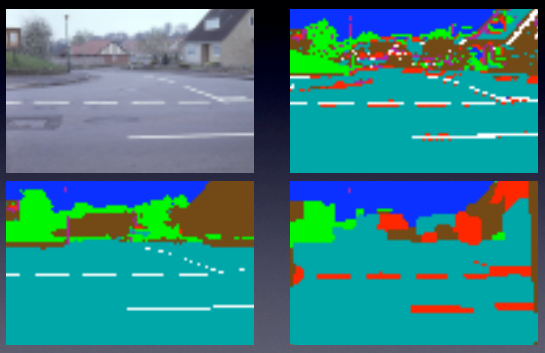
32

Select Rhino



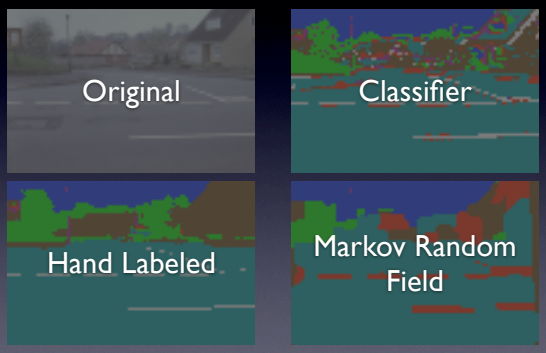
32

Select Street Scene



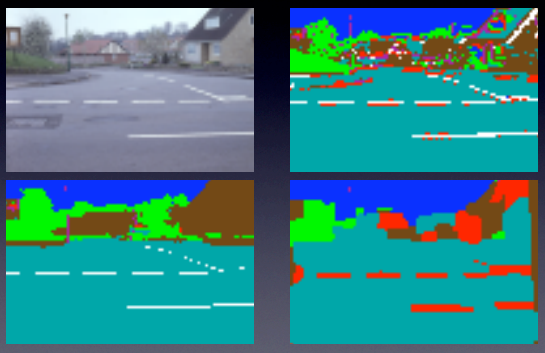
33

Select Street Scene



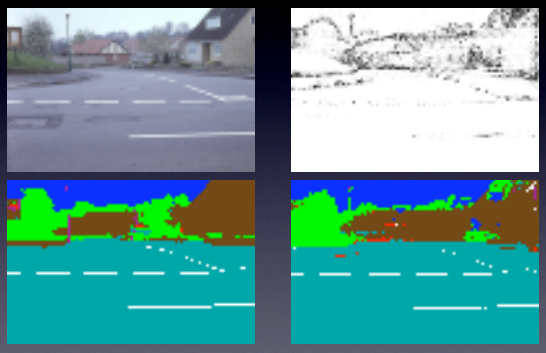
33

Select Street Scene



33

Select Street Scene

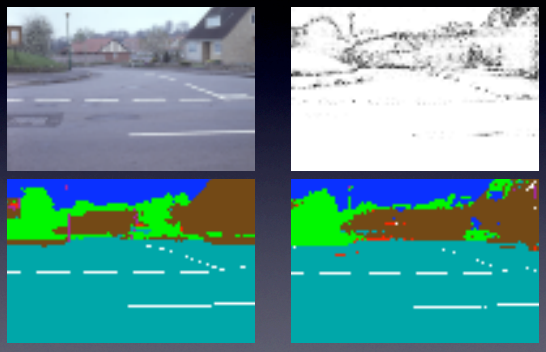


34

Select Street Scene

34

Select Street Scene



34

Thank You