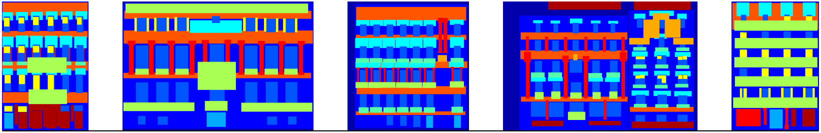
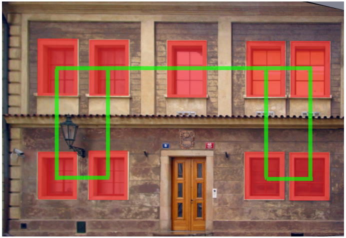
Spatial Pattern Templates for Recognition of Objects with Regular Structure

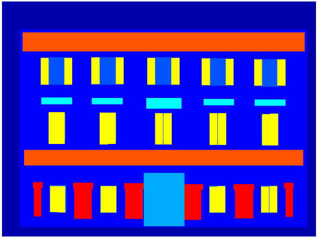
GCPR 2013

Radim Tyleček and Radim Šára

Center for Machine Perception
Czech Technical University in Prague
cmp.felk.cvut.cz

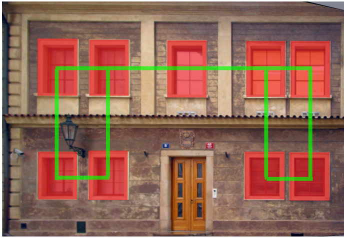


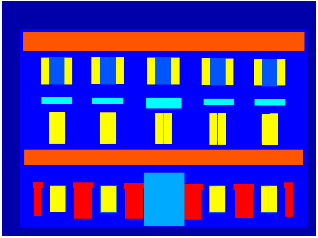




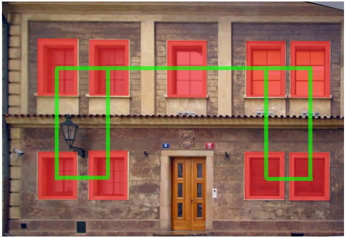
- **Observation:** Some objects in images have regular layout.

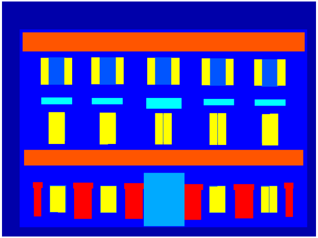
Introduction



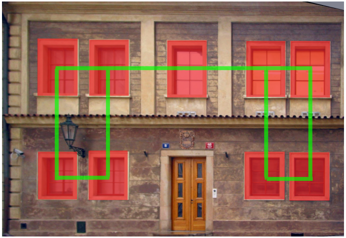
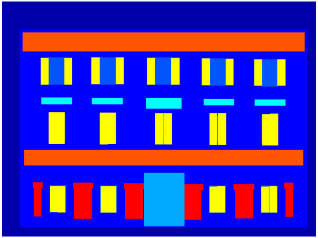


- **Observation:** Some objects in images have regular layout.
- **Regularity:** Repetition of elements according to simple rules (symmetry).





- **Observation:** Some objects in images have regular layout.
- **Regularity:** Repetition of elements according to simple rules (symmetry).
- **Task:** Incorporate regular contextual cues as a structure prior for recognition.

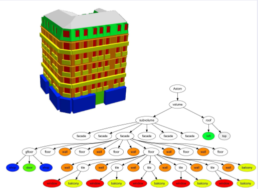
- **Observation:** Some objects in images have regular layout.
- **Regularity:** Repetition of elements according to simple rules (symmetry).
- **Task:** Incorporate regular contextual cues as a structure prior for recognition.
- **Problem:** How to specify a language for complex relations between many object instances of many classes?

Related Work

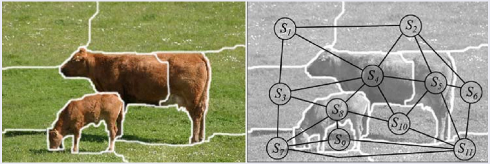
Shape Grammars

[Simon 2011]

- set of production rules
- restrictive split layout
- grammar specification



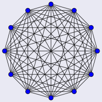
Sparse Graphical Models [Gould 2008]



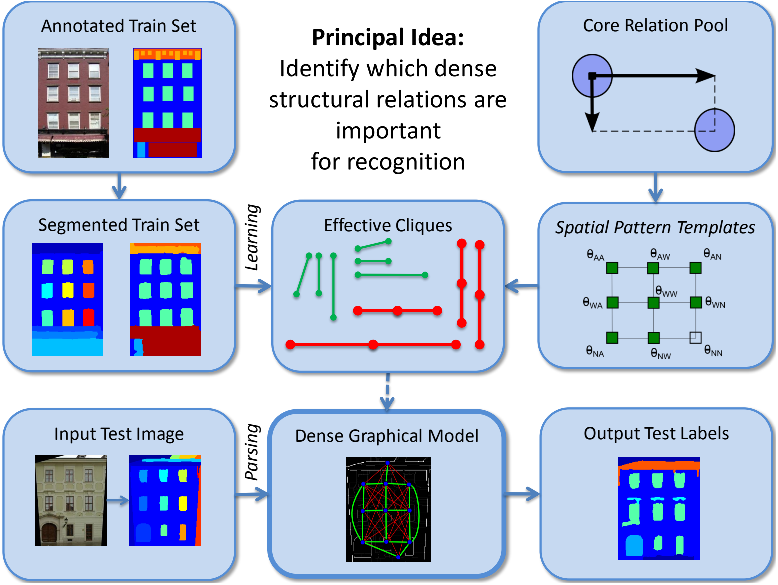
- adjacency, associative potentials
- complex relations not captured

Dense Graphical Models [Schmidt 2010]

- complete, high order cliques
- learning weights jointly in large graphs intractable



[Simon 2011] Simon, Teboul, Koutsourakis, Paragios: Random exploration of the procedural space. IJCV (2011)
[Gould 2008] Gould et al.: Multi-class segmentation with relative location prior. IJCV (2008)
[Schmidt 2010] Schmidt, Murphy: Convex structure learning: Beyond pairwise potentials. AISTATS (2010)


Annotated Train Set
Principal Idea:
Identify which dense structural relations are important for recognition
Core Relation Pool
Segmented Train Set
Learning
Effective Cliques
Spatial Pattern Templates
θAA
θAW
θAN
θWA
θWW
θWN
θNA
θNW
θNN
Input Test Image
Parsing
Dense Graphical Model
Output Test Labels

Problem Formulation

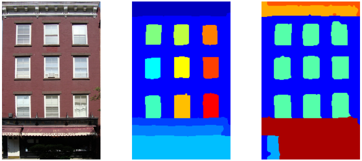
Input

X – rectified image data

S – unsupervised segmentation

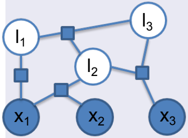
Output

L – class labels for segments in S



X S L

Probability Model: Conditional Random Field [CRF]



General CRF

find labeling $L : S \to \mathcal{C}$ that maximizes

$$p(L|X,S) = \frac{1}{Z} \prod_{q \in Q(S)} exp\Big(\sum_{j \in \Phi_q} \theta_j \varphi_j(\mathbf{l}_q, \mathbf{x}_q, \mathbf{s}_q) \Big)$$

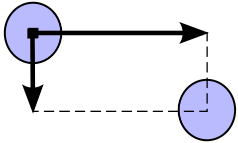
topology given by cliques (factors) in $Q(S)$

[CRF] Lafferty et al.: CRF: Probabilistic models for segmenting and labeling sequence data. ICML (2001)

SPT = representation for learning dense structural relations

1. Specify core attribute relation functions

relations act on attributes of segments tuples



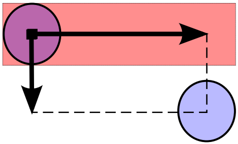
SPT = representation for learning dense structural relations

1. Specify core attribute relation functions

relations act on attributes of segments tuples

2. Discretize relation functions

value range is split into discrete intervals



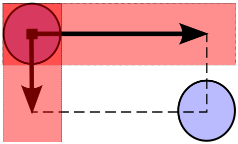
SPT = representation for learning dense structural relations

1. Specify core attribute relation functions

relations act on attributes of segments tuples

2. Discretize relation functions

value range is split into discrete intervals



SPT = representation for learning dense structural relations

1. Specify core attribute relation functions

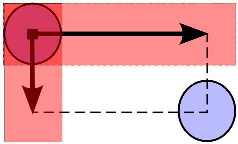
relations act on attributes of segments tuples

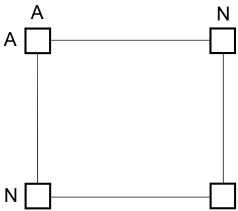
2. Discretize relation functions

value range is split into discrete intervals

3. Create composite relations

subsets in Cartesian product of relations





SPT = representation for learning dense structural relations

1. Specify core attribute relation functions

relations act on attributes of segments tuples

2. Discretize relation functions

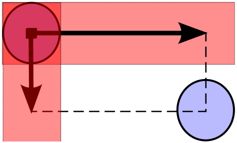
value range is split into discrete intervals

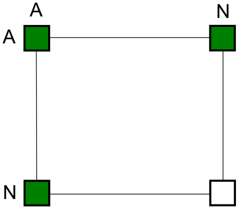
3. Create composite relations

subsets in Cartesian product of relations

4. Template domain

indicate allowed combinations





SPT = representation for learning dense structural relations

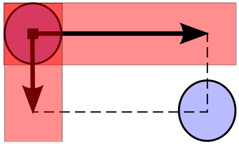
1. Specify core attribute relation functions
relations act on attributes of segments tuples

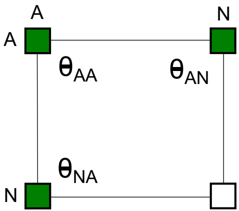
2. Discretize relation functions
value range is split into discrete intervals

3. Create composite relations
subsets in Cartesian product of relations

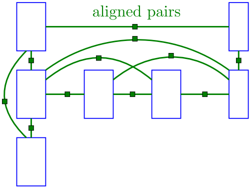
4. Template domain
indicate allowed combinations

5. Learn weights
for each allowed combination



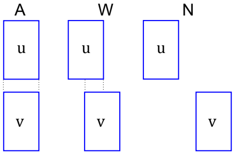


- capture **pair-wise alignment**
- connect two segments aligned horizontally or vertically
- relative position relations d_1, d_2
- nesting, overlap possible
- statistical potential [Tighe11]

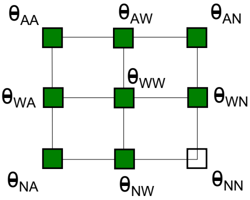


factor graph

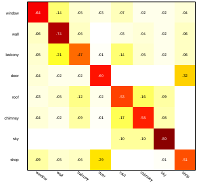
$$\varphi_2(u, v) = \theta_{d_1 d_2} \; f_c(l_u, l_v, d_1, d_2)$$



core relation function d

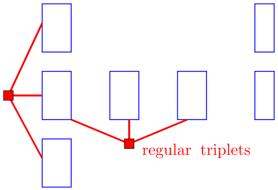


SPT domain



label co-occurence

[Tighe 2011] Tighe, Lazebnik: Understanding scenes on many levels. ICCV (2011)



factor graph

- capture basis for **repetitive structure** (rows, columns)
- connect three equally spaced segments
- relative position relation
- similarity in segment size
- generalized associative potential

[Kohli 2009]

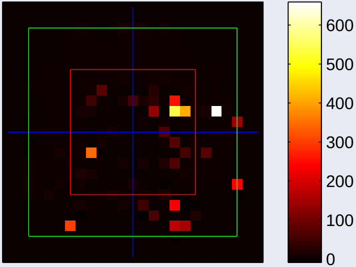
$$\varphi_3(u, v, w) = \begin{cases} \theta_c & \text{if } l_u = l_v = l_w, \\ \theta_0 & \text{if different,} \\ 0 & \text{if irregular} \end{cases}$$

- enforcing same labels in a triplet

[Kohli 2009] Kohli, Ladicky, Torr: Robust higher order potentials for enforcing label consistency. IJCV (2009)

Learning Algorithm

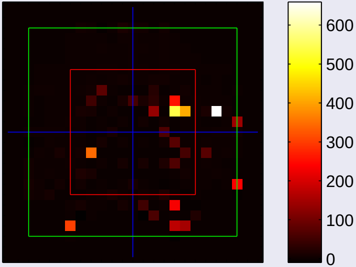
- piece-wise estimation of parameters θ
- unary SVM classifier
- CRF pseudo-likelihood maximization (50 instances)



Aligned Pairs weights $\theta_{d_1 d_2}$

Learning Algorithm

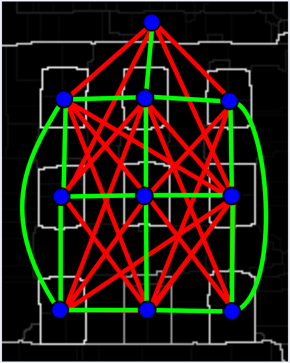
- piece-wise estimation of parameters θ
- unary SVM classifier
- CRF pseudo-likelihood maximization (50 instances)



Aligned Pairs weights $\theta_{d_1 d_2}$

Decoding Algorithm

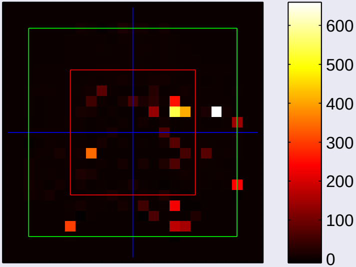
- recognition of labels L
- max-product tree-reweighted message passing
- voting scheme



Inference

Learning Algorithm

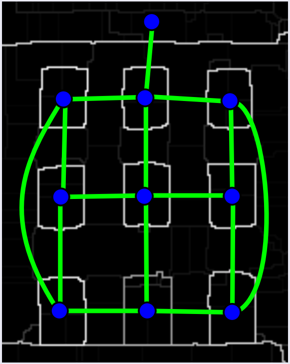
- piece-wise estimation of parameters θ
- unary SVM classifier
- CRF pseudo-likelihood maximization (50 instances)



Aligned Pairs weights $\theta_{d_1 d_2}$

Decoding Algorithm

- recognition of labels L
- max-product tree-reweighted message passing
- voting scheme



Standard Dataset Evaluation

ECP Monge – 104 images, 8 classes [3L]: 85.1%, [SG]: 74.7%

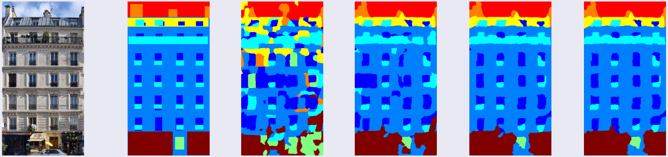


image	GT	NC	AP	**APRT**	SGT
accuracy	[%]	59.6	79.0	**84.2**	88.5

[3L] Martinovic et al.: A three-layered approach to facade parsing. ECCV (2012)
[SG] Simon, Teboul, Koutsourakis, Paragios: Random exploration of the procedural space. IJCV (2011)
[HCRF] Yang, Foerstner: A hierarchical CRF model for images of man-made scenes. ICCV (2011)

Standard Dataset Evaluation

ECP Monge – 104 images, 8 classes   [3L]: 85.1%, [SG]: 74.7%

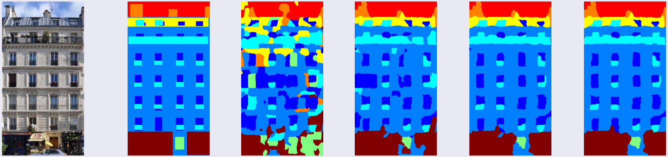


image	GT	NC	AP	**APRT**	SGT
accuracy	[%]	59.6	79.0	**84.2**	88.5

eTrims – 60 images, 8 classes   [3L]: 81.9%, [HCRF]: 65.8%

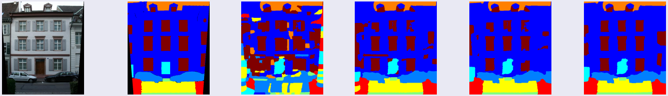


image	GT	NC	AP	**APRT**	SGT
accuracy	[%]	56.7	77.4	**82.1**	93.7

[3L] Martinovic et al.: A three-layered approach to facade parsing. ECCV (2012)
[SG] Simon, Teboul, Koutsourakis, Paragios: Random exploration of the procedural space. IJCV (2011)
[HCRF] Yang, Foerstner: A hierarchical CRF model for images of man-made scenes. ICCV (2011)

CMP Facade Database

New public dataset, 400 images, rectified, annotated
Website: `cmp.felk.cvut.cz/~tylecr1/facade/`

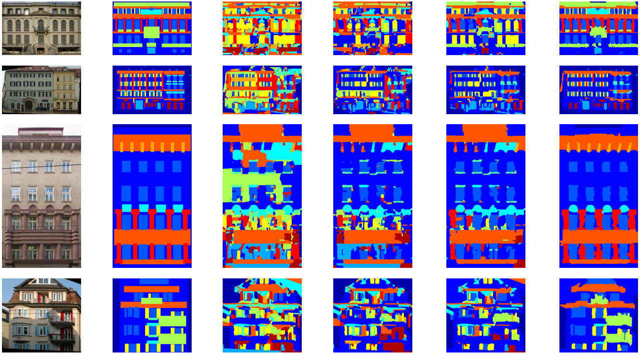


image	GT	NC	AP	**APRT**	SGT
acc [%]		33.2	54.3	**60.3**	84.8

12 classes:

- facade
- molding
- cornice
- pillar
- window
- door
- sill
- blind
- balcony
- shop
- deco

Conclusion

General

Spatial Pattern Template – new representation for learning dense structural relations

Conclusion

General

Spatial Pattern Template – new representation for learning dense structural relations

Templates for Regular Scenes

Aligned Pairs – capture pairwise alignment and co-occurrence
Regular Triplets – capture regular spacing in triplets

Conclusion

General

Spatial Pattern Template – new representation for learning dense structural relations

Templates for Regular Scenes

Aligned Pairs – capture pairwise alignment and co-occurrence
Regular Triplets – capture regular spacing in triplets

Application to Facade Parsing

Standard Datasets – performance comparable to SOA
CMP Facade Database – new 12-class challenge

Center for Machine Perception
Czech Technical University in Prague
`cmp.felk.cvut.cz/~tylecr1/facade`