Probabilistic Models for Symmetric Object Detection in Images

Radim Tyleček

PhD Thesis
April 2016

Center for Machine Perception
Czech Technical University in Prague

Motivation

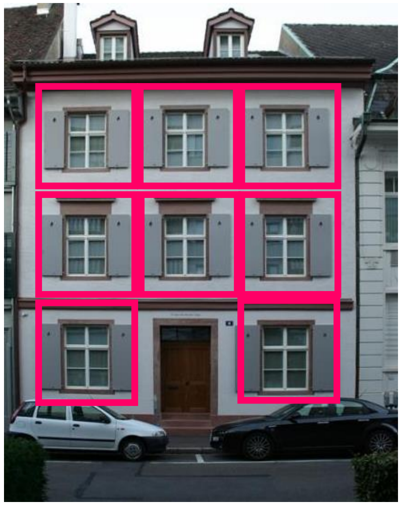
- Where standard object detection fails
- Context helps – regular structures
- Some challenges:

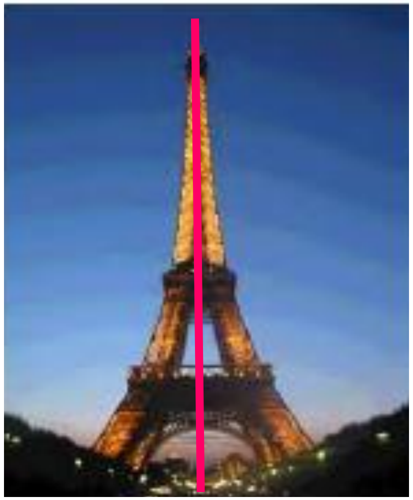
Types of Symmetry

In 2D images we deal with

› Translation, reflection, rotation

› Groups – wallpaper, dihedral

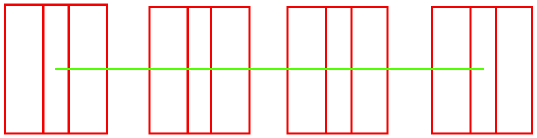


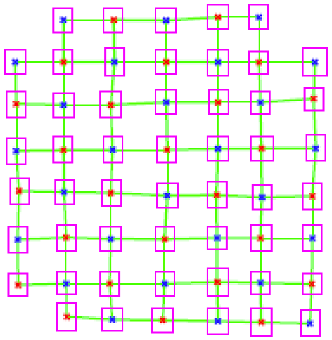


Grouping Principles

Human perception priors are based on

- Proximity
- Similarity
- Reflection
- Continuation





Also known as Gestalt laws or **symmetry** in general.

We seek a **language** to describe such structures for computer vision.

Complexity – unknown number of components

Thesis Progress

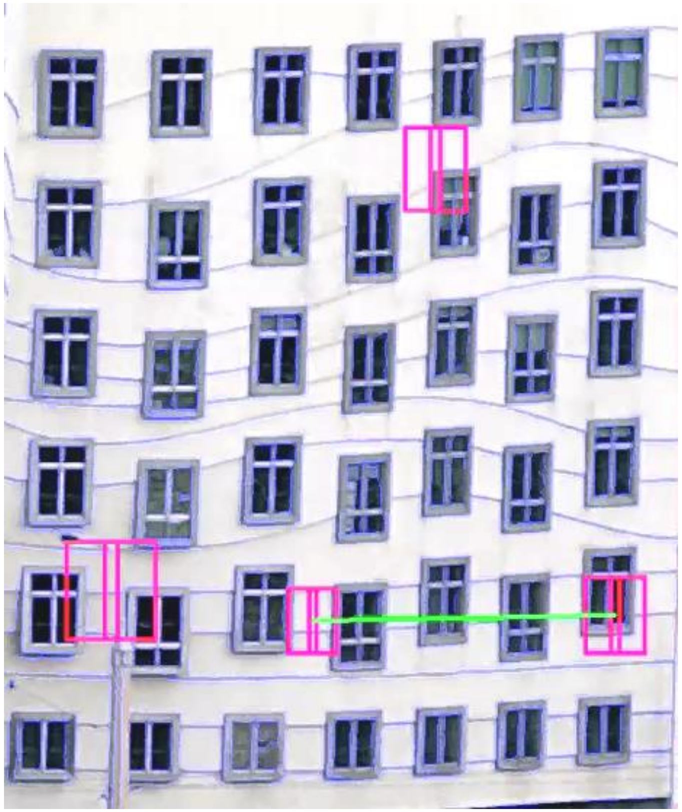
› Weak Structure Model
  – **Simple** model implementing grouping principles
  – Window detection, sampling

› Spatial Pattern Templates
  – Learn **where** grouping principles apply
  – Facade parsing: semantic labels

› Reflection Symmetry Detection
  – More **general** approach resembling clustering
  – Improved inference engine, dihedral group

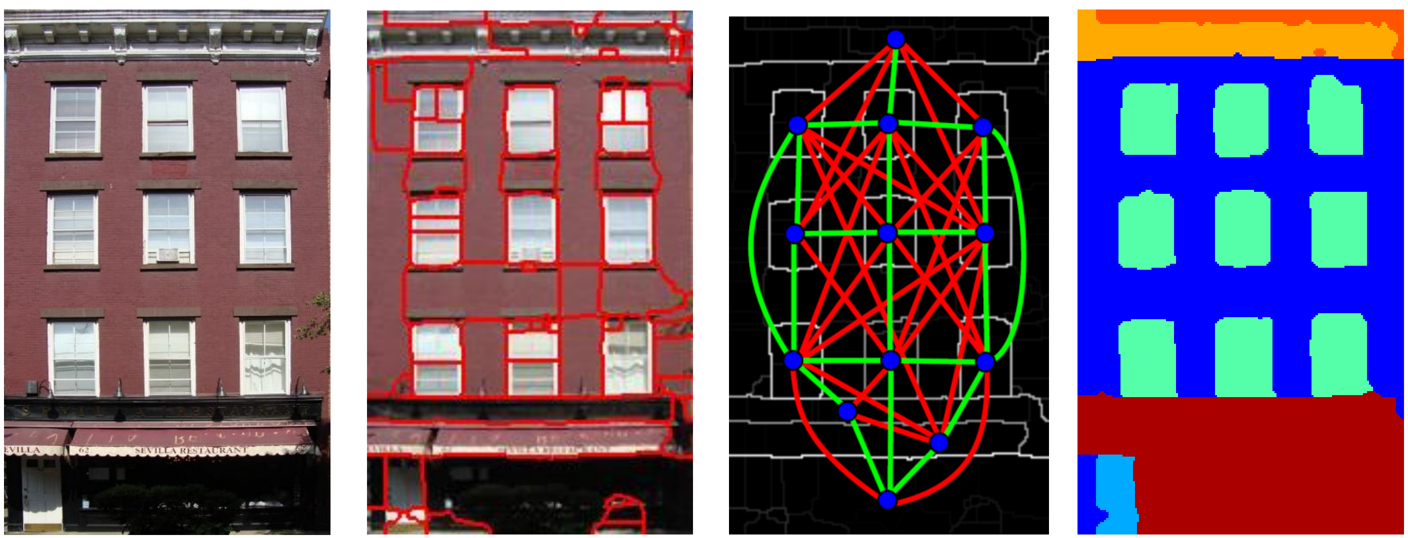
Weak Structure Model

- *Can we infer global structure from local interactions?*
- Markov Chain Monte Carlo sampling to find MAP solution
- Random Walk
- Reversible Jump
- Proposal Efficiency
- Convergence



Facade Image Parsing

- *Can we learn where grouping principles should be applied?*
- Dense Graphical Model
- More semantic labels and context
- New database for learning

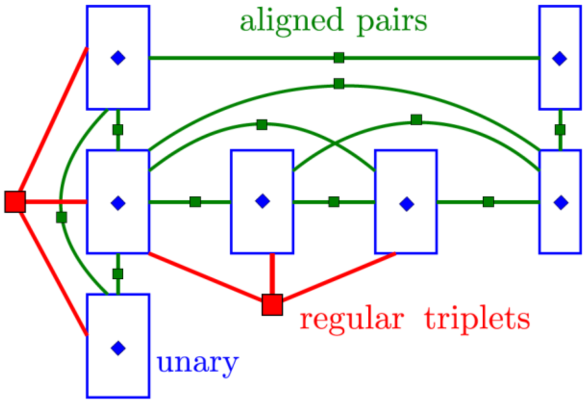


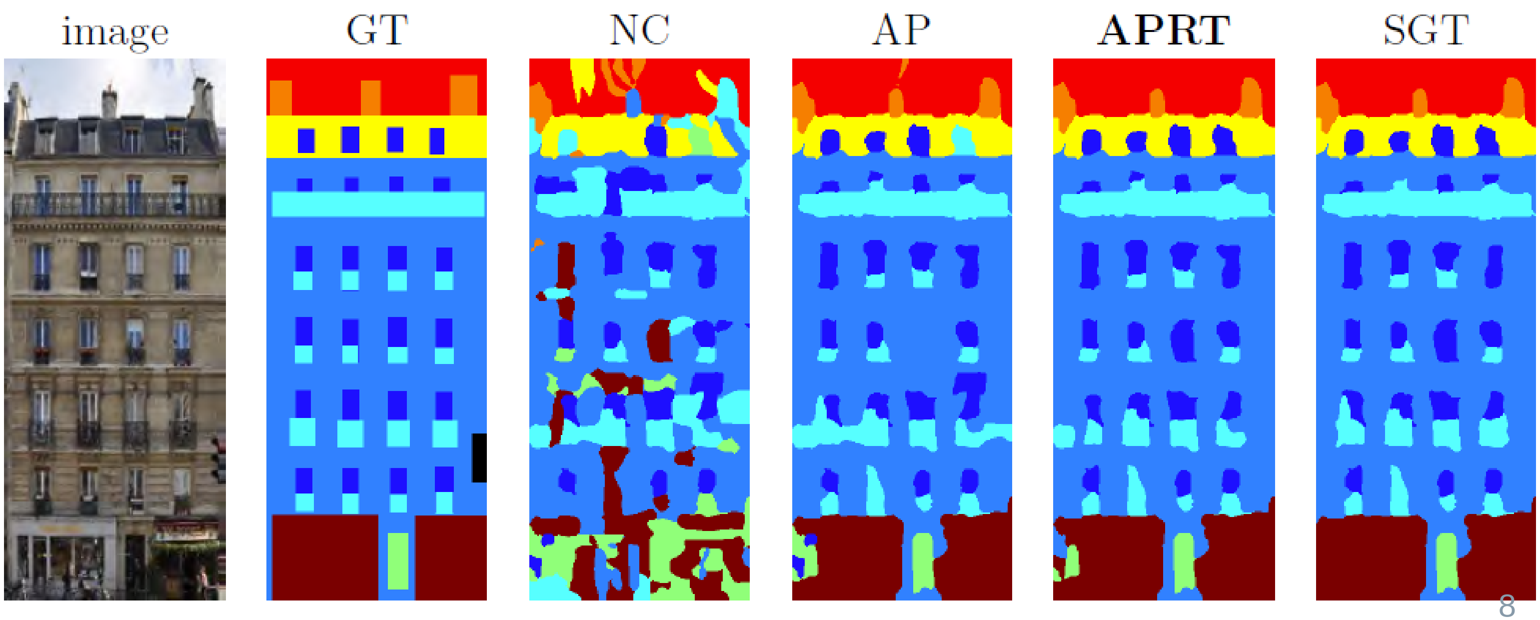
Spatial Pattern Templates

› Binary and ternary terms

› Relative spatial location

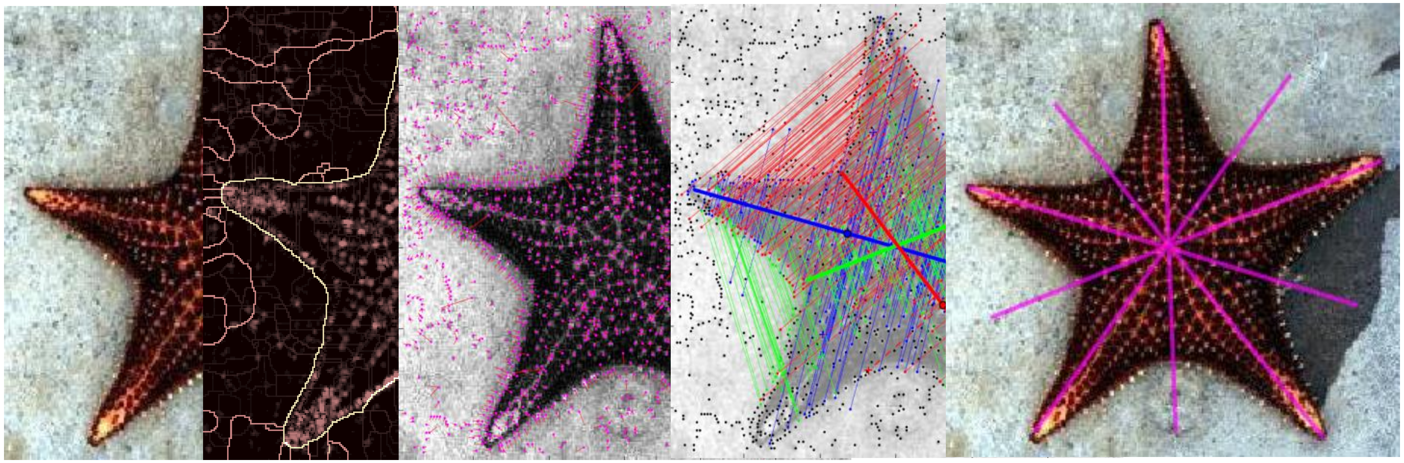
› Approximate inference



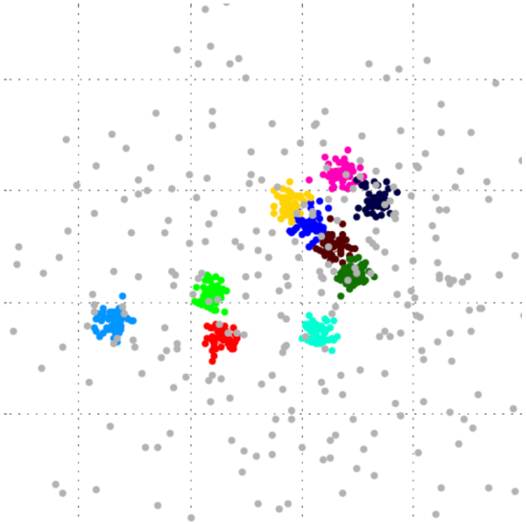


Multiple Reflection Symmetry

› Correspondence matching problem
  - **Keypoints**: detected from corners and edges
  - **Primitives**: two corresponding keypoints
  - **Components**: axes of reflection symmetry
  - **Groups**: clusters of components (dihedral)



Bayesian Modeling



› Data clustering problem
  – Gaussian mixture + outliers

› Target distribution
  = data model + allocation + priors

$$p(X, Z, \theta, k) = p(X \mid Z, \theta, k)\, p(Z \mid \theta, k)\, p(\theta|k)\, p(k)$$

  – X … data primitives with attributes
  – Z … allocation of data points to components
  – θ … component and shape parameters
  – k … complexity

› Bayesian choice
  – prior design requires some skills

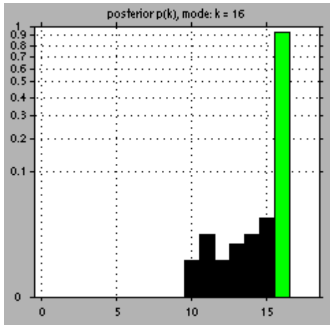
Bayesian Inference

1. Model Selection
   – Consider multiple models with different **complexity** and choose one to maximize the posterior marginal

$$k^* = \arg\max \, p(k \mid X)$$

   – Integrate over parameters by MCMC **sampling**

$$p(k \mid X) \propto \sum_{Z} \int_{\theta} p(X, Z, \theta, k) \, d\theta dZ$$



2. Parameter Estimation
   – Determine the most probable **parameters**

$$\theta^* = \arg\max \, p(X, Z, \theta \mid k^*)$$

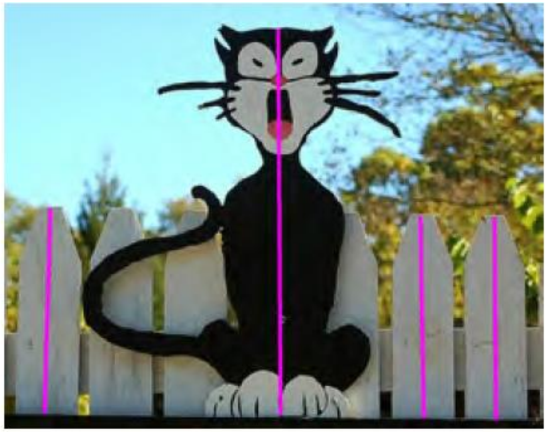
   – Use Stochastic EM to find locally optimal values

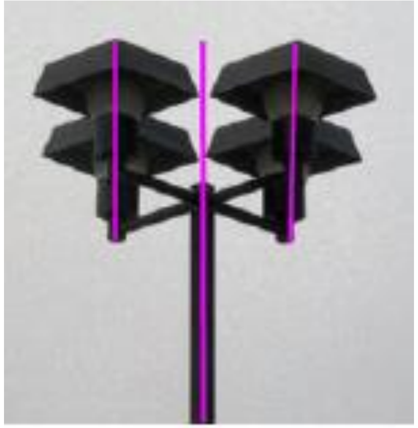
› Inference Engine: LiSAEM
   – Efficient: improved mixing rate, ~10k samples needed

Multiple Reflection Symmetry

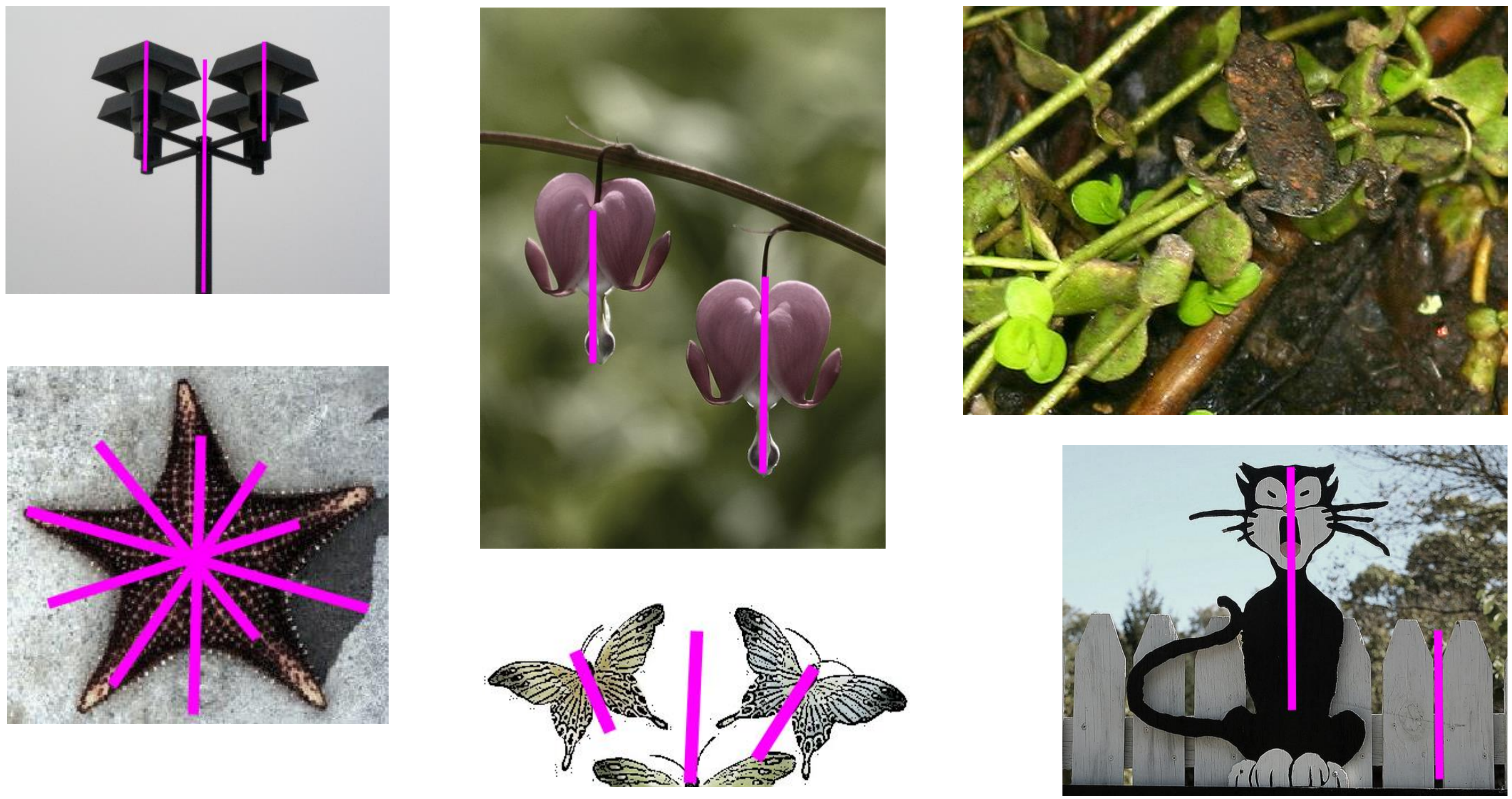
- *General difficulties*:
  - **Multiplicity**
  - **Hierarchy**
- *Domain specific ambiguities*:
  - Figure-Ground
  - Local-Global

Addressed with:
- **Model selection**
- **Grouping priors**
- Dihedral
- Objectness
- Compactness





Experimental Results



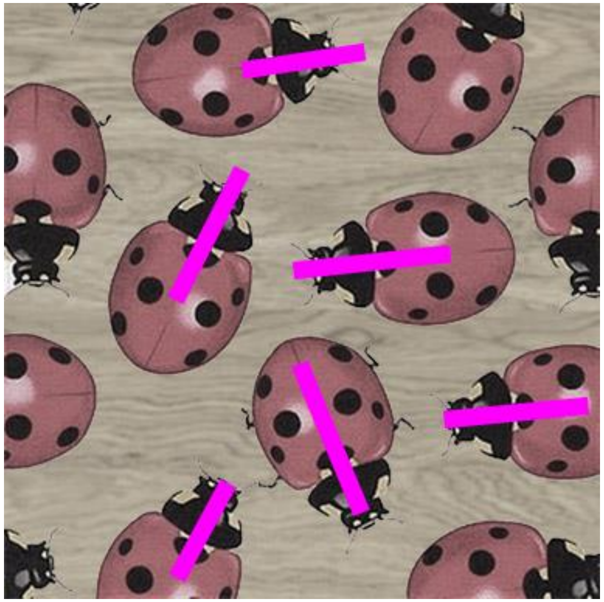
› Improved state-of-the-art results on reflection symmetry benchmarks (~10%)

Main Contributions

- Application of statistical methods for object counting new to computer vision
  - Parsimony by means of model selection
  - Learning without overfitting
- Minimal modeling principle
  - Simple language for consistent models
- Grouping priors
  - Components are not independent
  - Hierarchy of symmetries

Thank You

› Questions?



Questions

- What are the alternative models and their properties?
  - Complexity estimation
    - Bayesian Information Criterion $\mathrm{BIC} = n \cdot \ln(\widehat{\sigma_\varepsilon^2}) + k \cdot \ln(n)$
      - Fixed penalty for increase of complexity
    - Multi-RANSAC
      - complexity estimation greedy or empirical
  - Symmetry modeling
    - Near Regular Textures – element unknown
    - Grammars – strong but restricted layout
    - Sequential inference – generally suboptimal

Questions

- Would larger datasets improve the results?
  - WSM, BMRS:
    - Yes, hyper-parameter learning would be possible on the next level
  - SPT:
    - Yes, now only limited number of samples used for training (MPL)
    - Results from CNNs suggest large data are useful
  - Computationally demanding

Questions

› Hierarchical Bayesian model