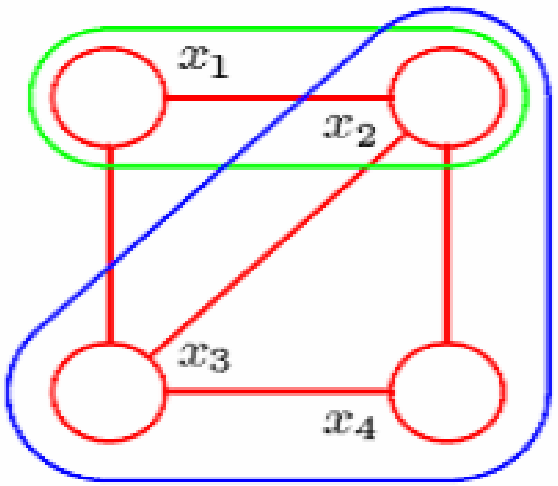
Graphical models in image understanding

Olga Barinova

Lomonosov Moscow State University

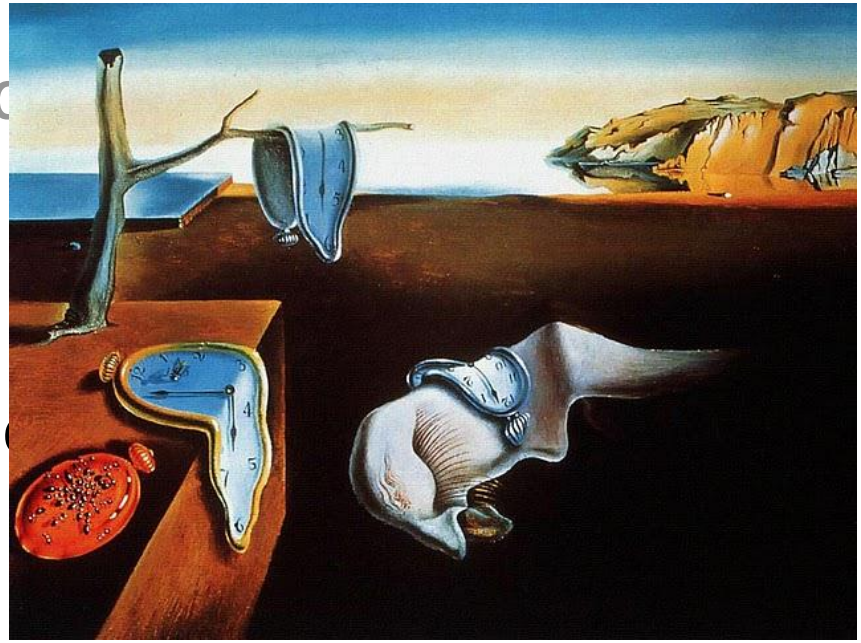
Expressing constraints with graphical models

- Outline of the talk
  - **The idea of graphical models**
  - Examples:
    - Limiting the set of allowed deformations
    - Occlusion constraint
    - Depth ordering constraint
    - Modeling the rules of perspective geometry



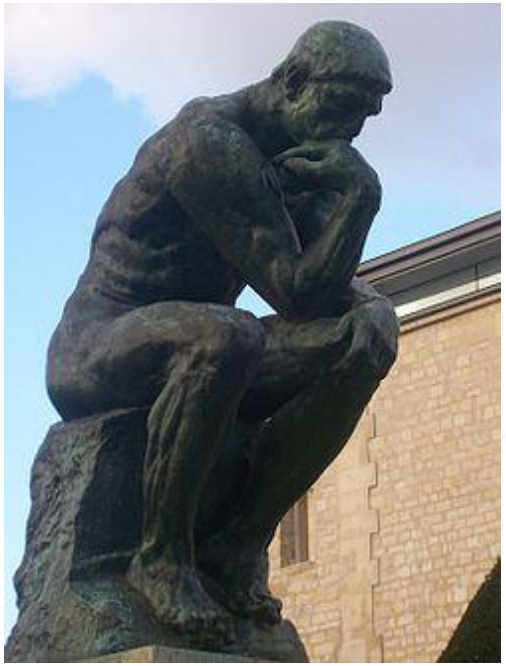
Expressing constraints with graphical models

- Outline of the talk
  - The idea of graphical mod
  - **Examples:**
    - **Limiting the set of allowe deformations**
    - Occlusion constraint
    - Depth ordering constraint
    - Modeling the rules of perspective geometry



Expressing constraints with graphical models

- Limiting the set of allowed deformations



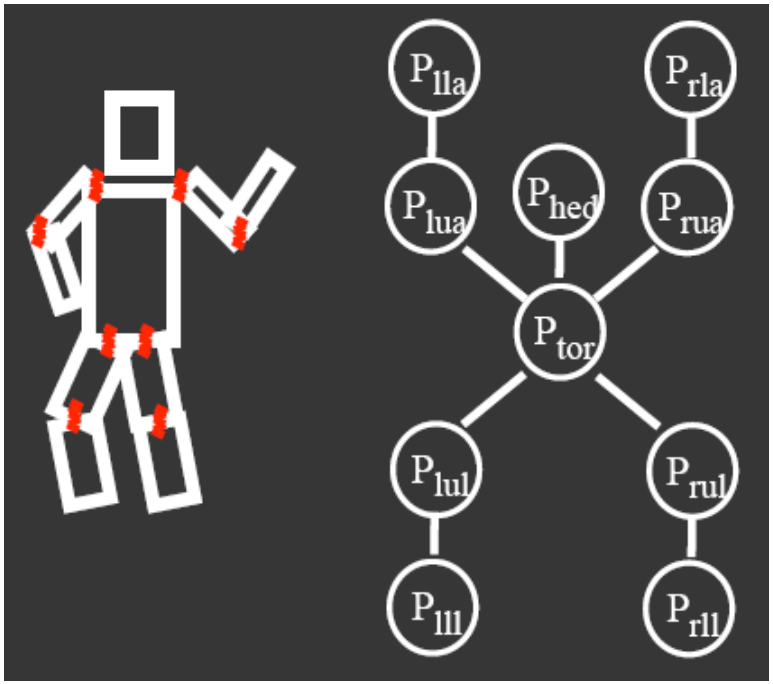
- Model should be flexible enough, but constrain the allowed deformations of an object

Limiting the set of allowed deformations

- Pictorial structures model

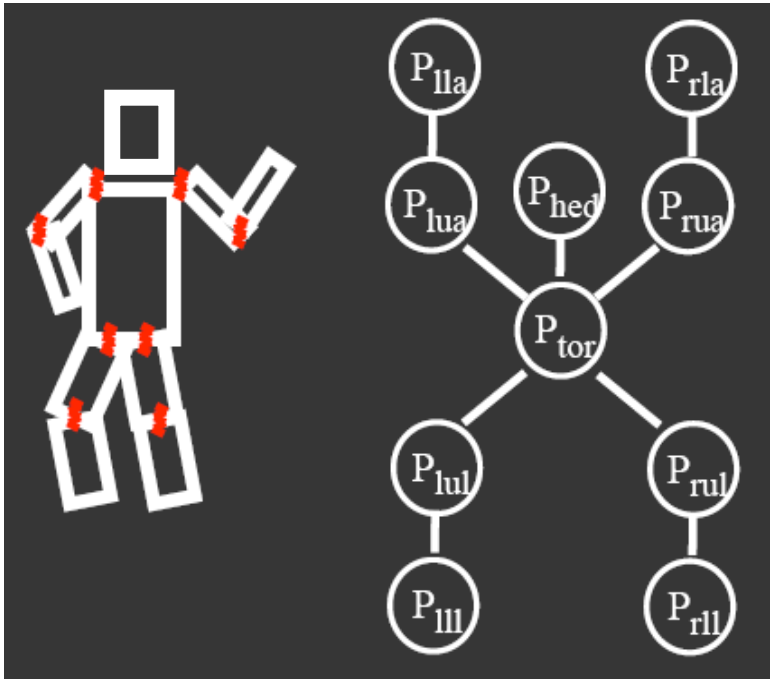
  - **Pictorial structures** strike a good balance between flexibility and tractability

  [Fischler & Elschlager 73],
  [Felzenshwalb & Huttenlocher 00]



Limiting the set of allowed deformations

- Pictorial structures model
  - Each vertex corresponds to a part of a person: *'Head', 'Torso', 'Legs', 'Arms'*
  - Edges form a tree
  - Person detector - for each vertex find a corresponding position from the set of valid positions



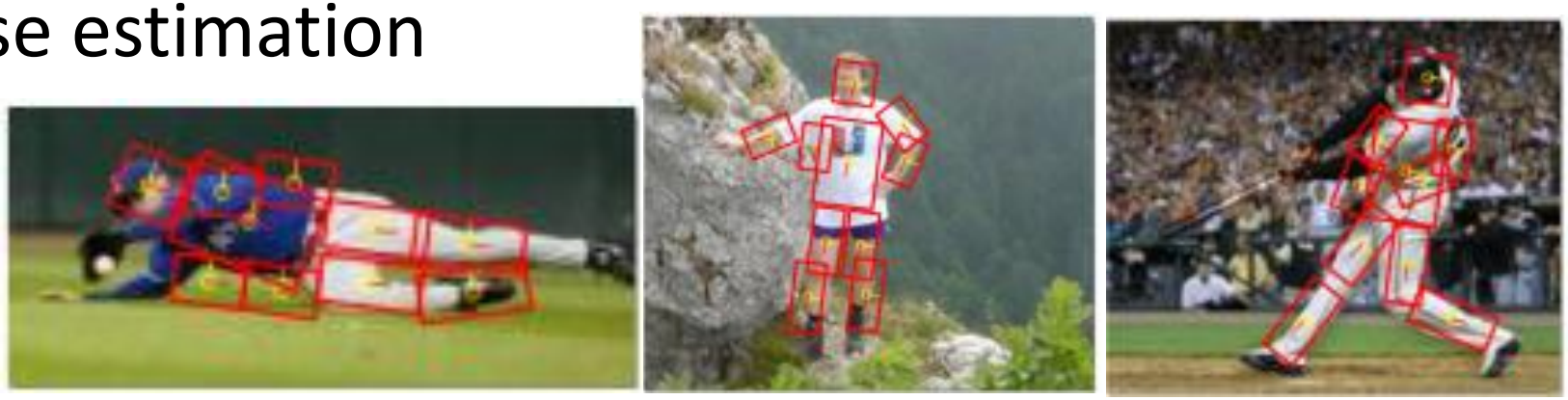
$$\Pr(P_{tor}, P_{arm}, \ldots | \mathrm{Im}) \propto \prod_{i,j} \Pr(P_i | P_j) \prod_i \Pr(\mathrm{Im}(P_i))$$

part geometry

part appearance

Limiting the set of allowed deformations

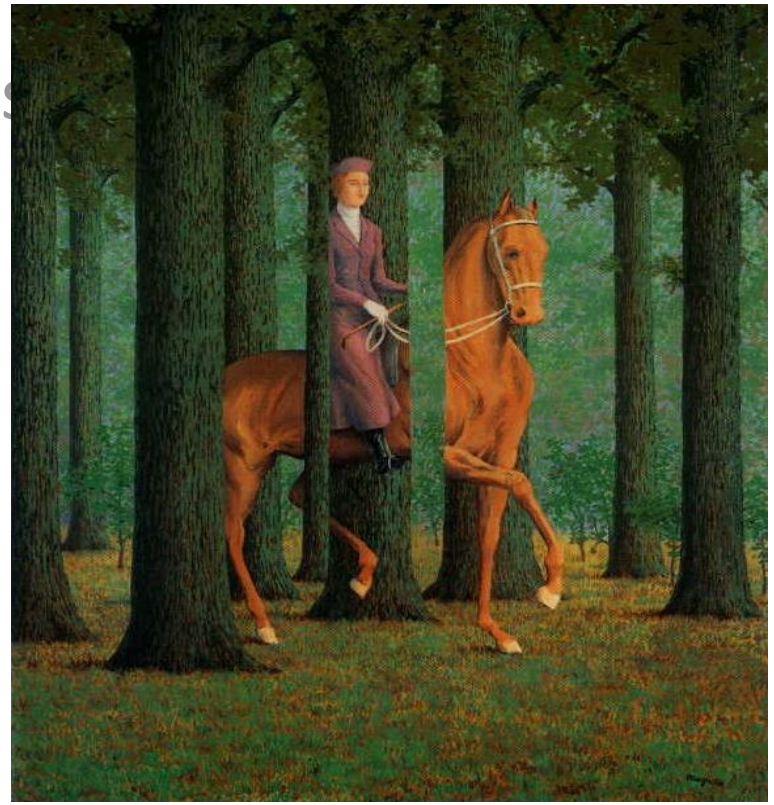
- Pose estimation



- Pedestrian detection

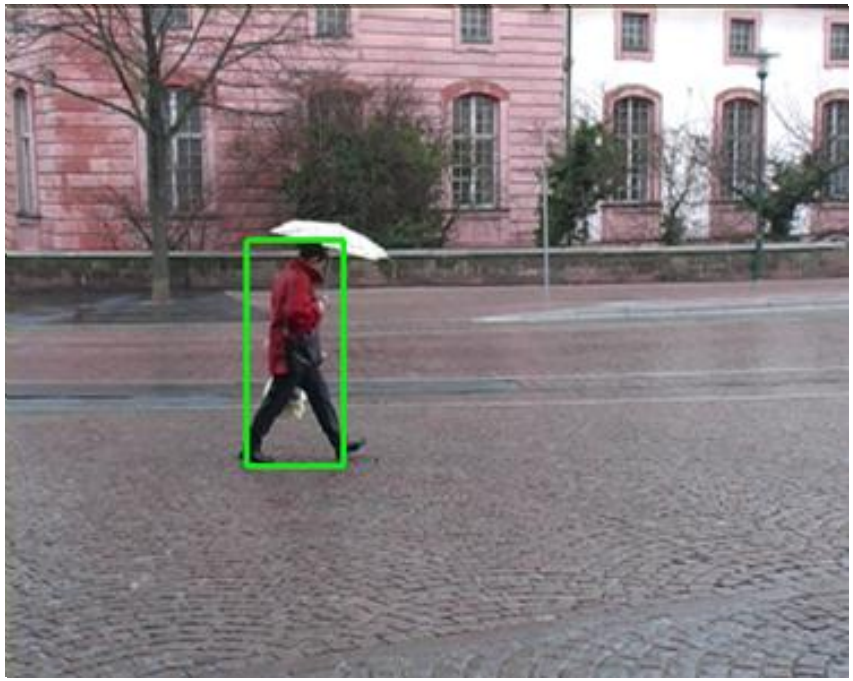
Expressing constraints with graphical models

- Outline of the talk
  - The idea of graphical models
  - **Examples:**
    - Limiting the set of allowed deformations
    - **Occlusion constraint**
    - Depth ordering constraint
    - Modeling the rules of perspective geometry



Expressing constraints with graphical models

- Occlusions

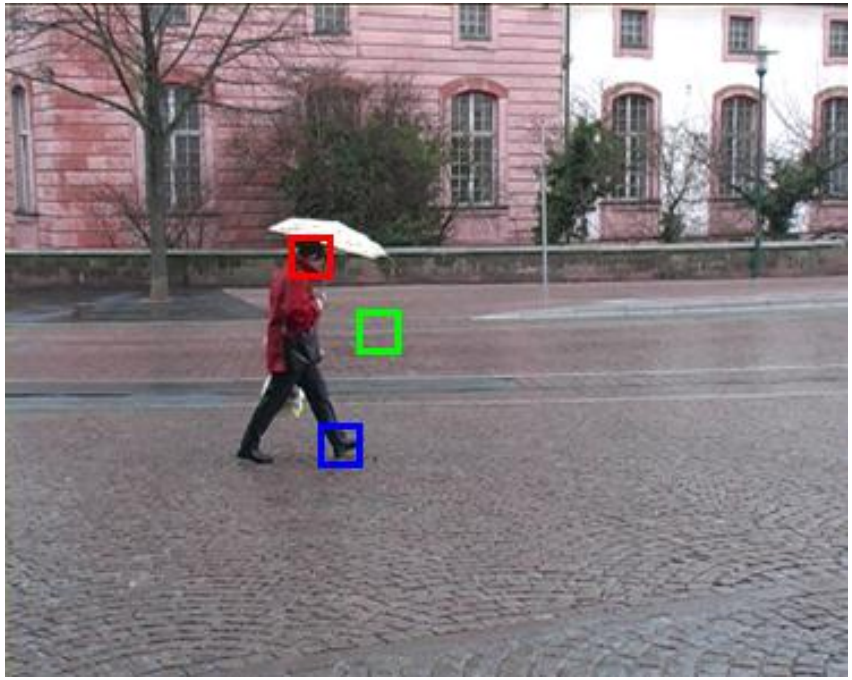


- Occlusions and self occlusions make the task of object detection even harder

Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

- Local model

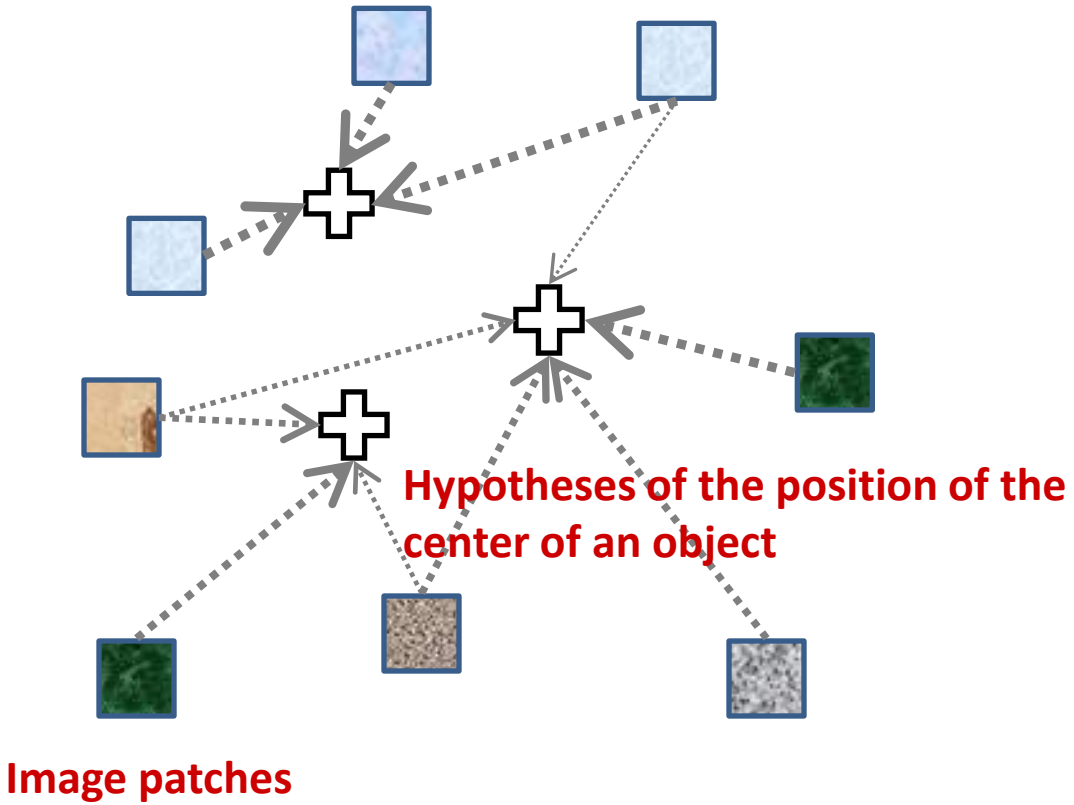


Model from [Gall & Lempitsky, CVPR09]

Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

- Each image pixel belong to no at most one object



Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

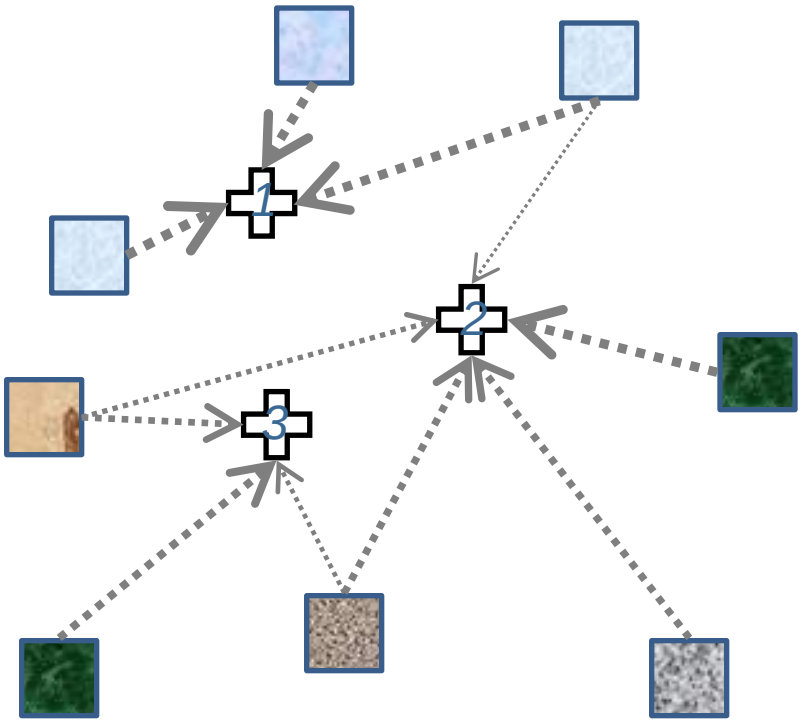
- Modeling the occlusion constraint

$\boldsymbol{x}$ – labelling of image patches,

x_i = index of hypothesis,

if the patch votes for hypothesis,

$x_i = 0$, if the patch votes for background



$\boldsymbol{y}$ – labelling of hypotheses, binary variables:

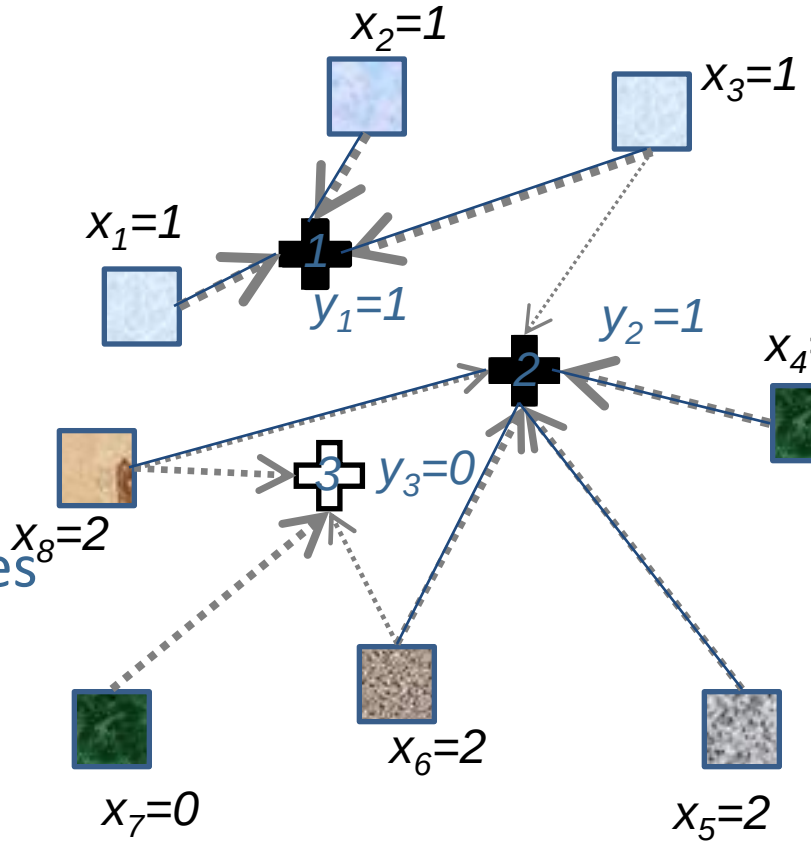
$y = 1$, if the object is present,

$y = 0$, otherwise

Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

- Modeling the occlusion constraint



$\boldsymbol{x}$ – labelling of image patches,

x_i = index of hypothesis, if the patch votes for hypothesis,

$x_i = 0$, if the patch votes for background

$\boldsymbol{y}$ – labelling of hypotheses, binary variables:

$y = 1$, if the object is present,

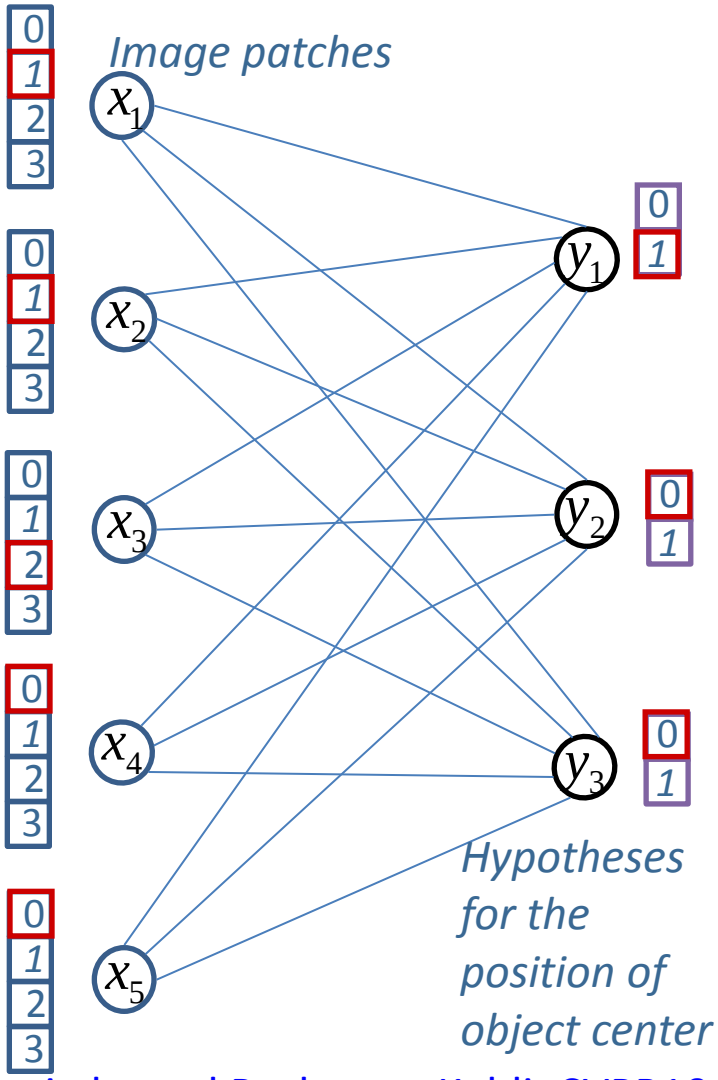
$y = 0$, otherwise

Key idea : joint MAP-inference in x and y

Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

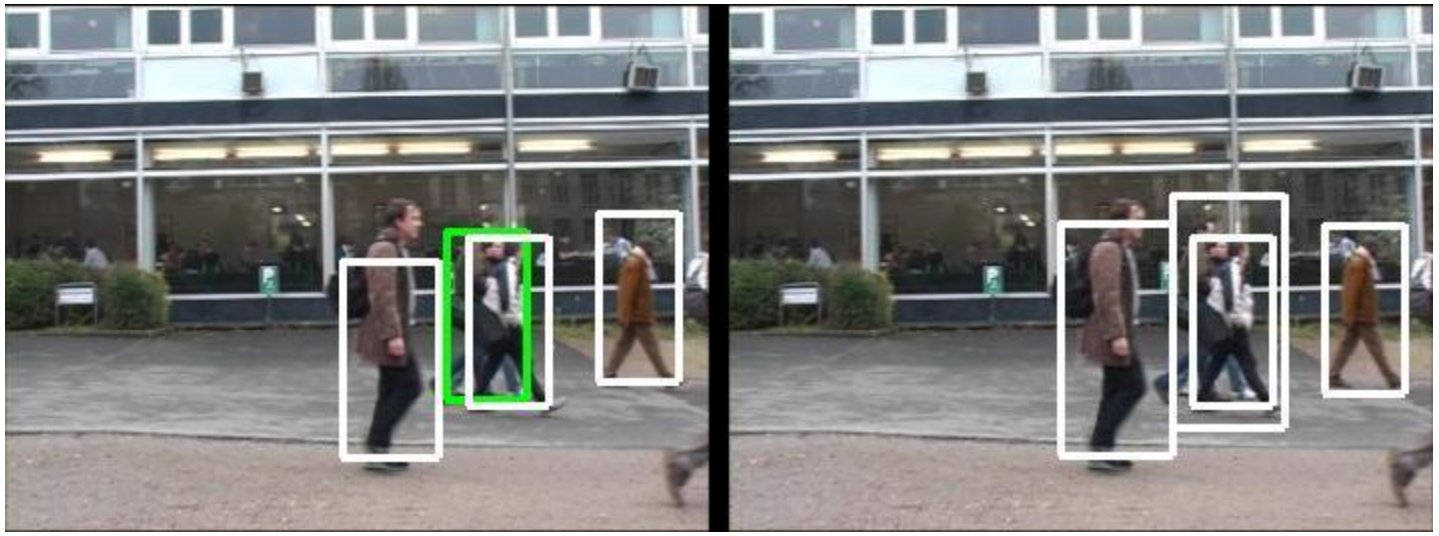
- Graphical model
  - If labeling of $\mathbf{y}$ is fixed, the values of x_i are independent
  - So we can maximize $\mathbf{x}$ out first and perform inference over $\mathbf{y}$



Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Occlusion constraint

- Comparison



Without occlusion constraint

Using occlusion constraint

White = correct detection

Green = missing object

Red = false positive

Code available online!!!

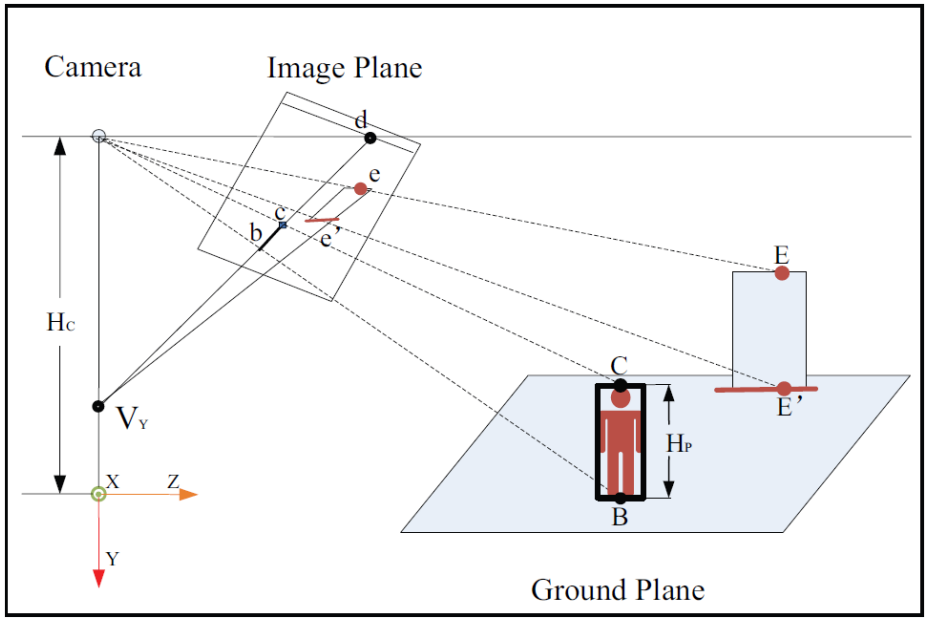
Joint work with Victor Lempitsky and Pushmeet Kohli, CVPR10

Expressing constraints with graphical models

- Outline of the talk
  - The idea of graphical mod
  - **Examples:**
    - Limiting the set of allowed deformations
    - Occlusion constraint
    - **Depth ordering constraint**
    - Modeling the rules of perspective geometry

Expressing constraints with graphical models

- Depth ordering

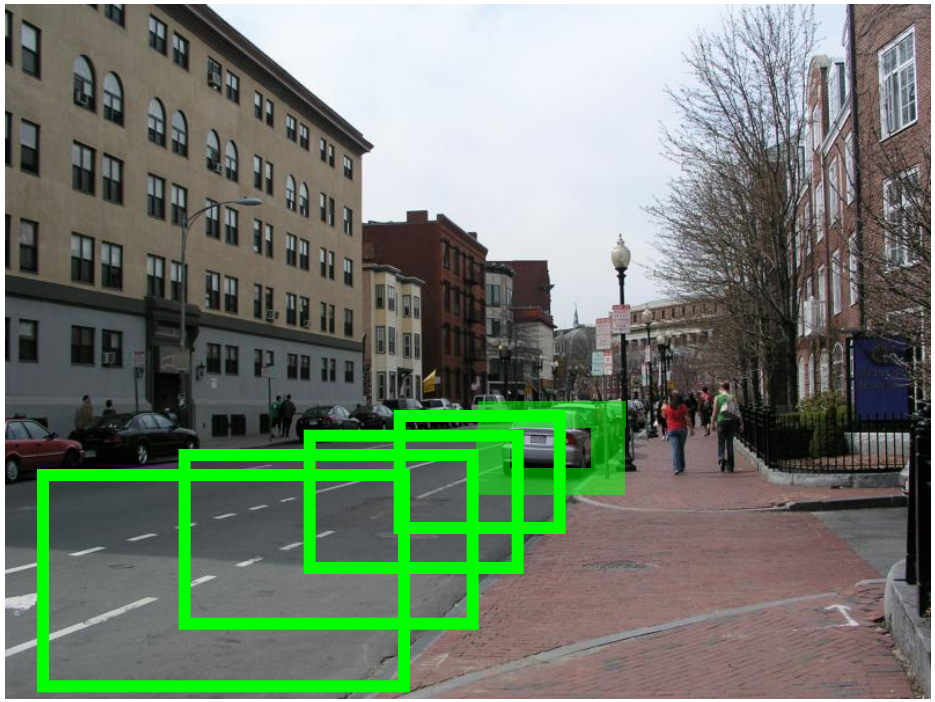


- The size of an object depends on the distance from the viewpoint
- Viewpoint is set by the position of horizon and ground plane in the image

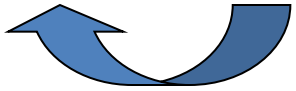
Depth ordering constraint

- Viewpoint → prior on the size of the objects

Object Position/Sizes



Viewpoint

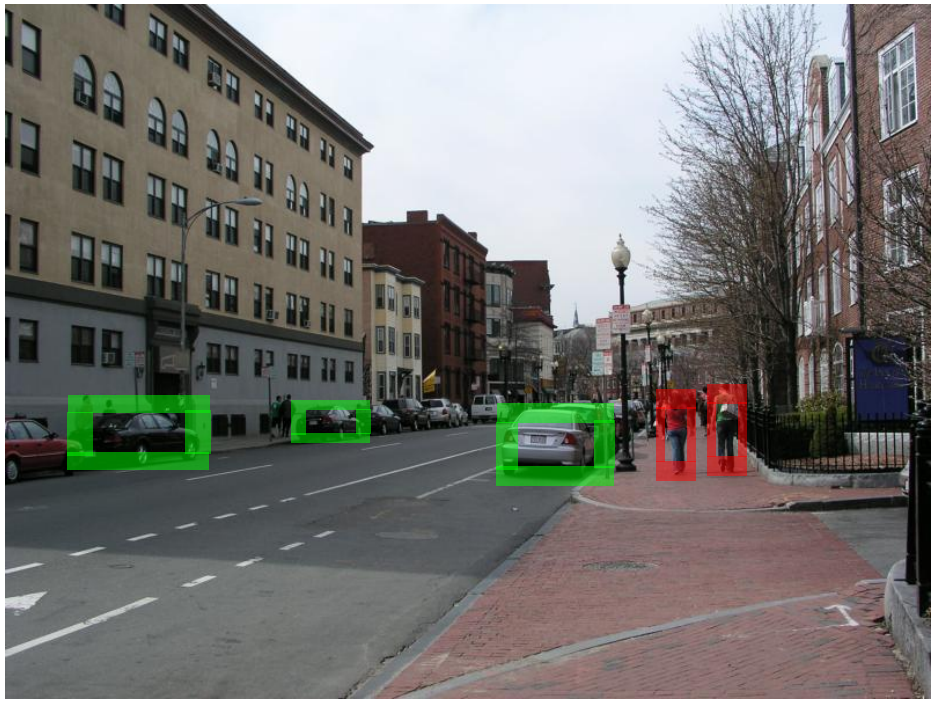


Hoiem et al. , CVPR2006

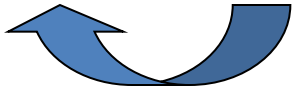
Depth ordering constraint

- Detected objects → viewpoint

Object Position/Sizes



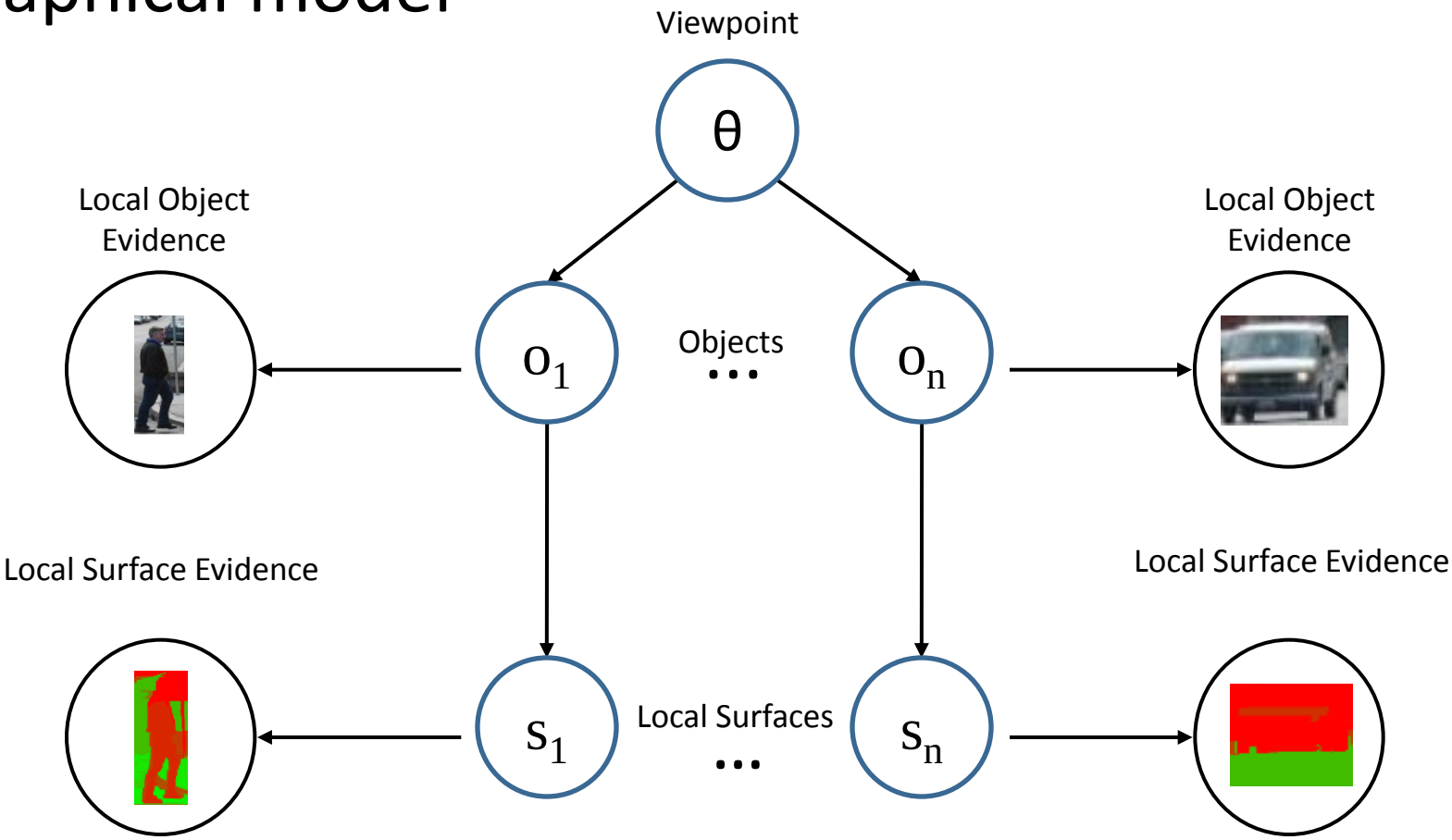
Viewpoint



Hoiem et al. , CVPR2006

Depth ordering constraint

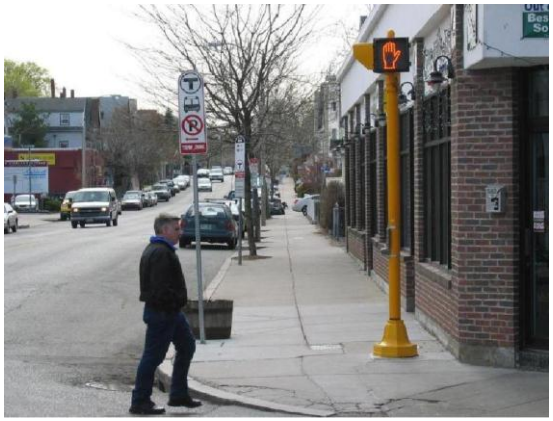
- Graphical model



Hoiem et al. , CVPR2006

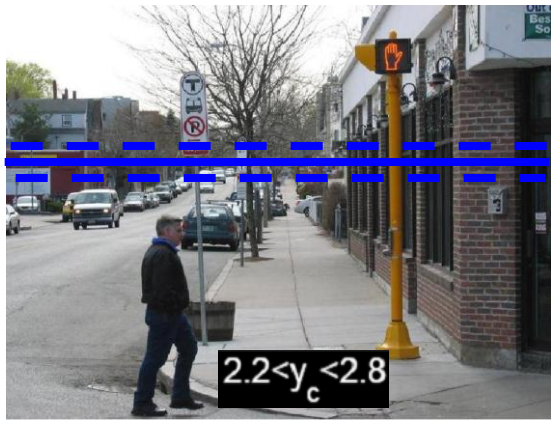
Depth ordering constraint

- Prior on the size an position of the objects

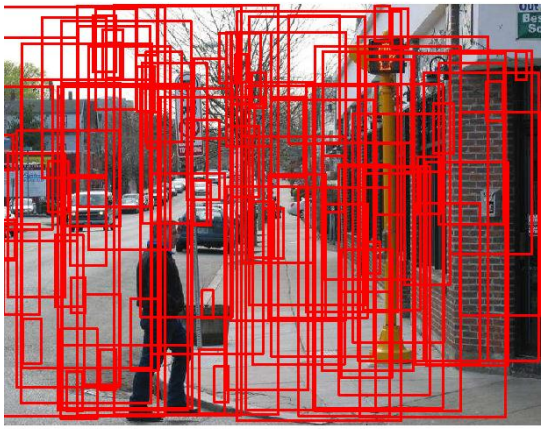


Image

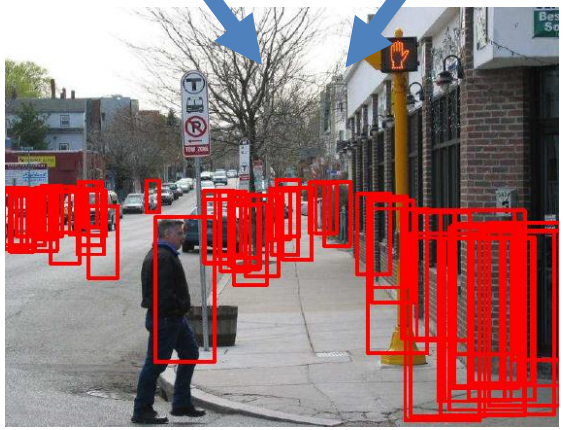
P(surfaces)



P(viewpoint)



P(object)

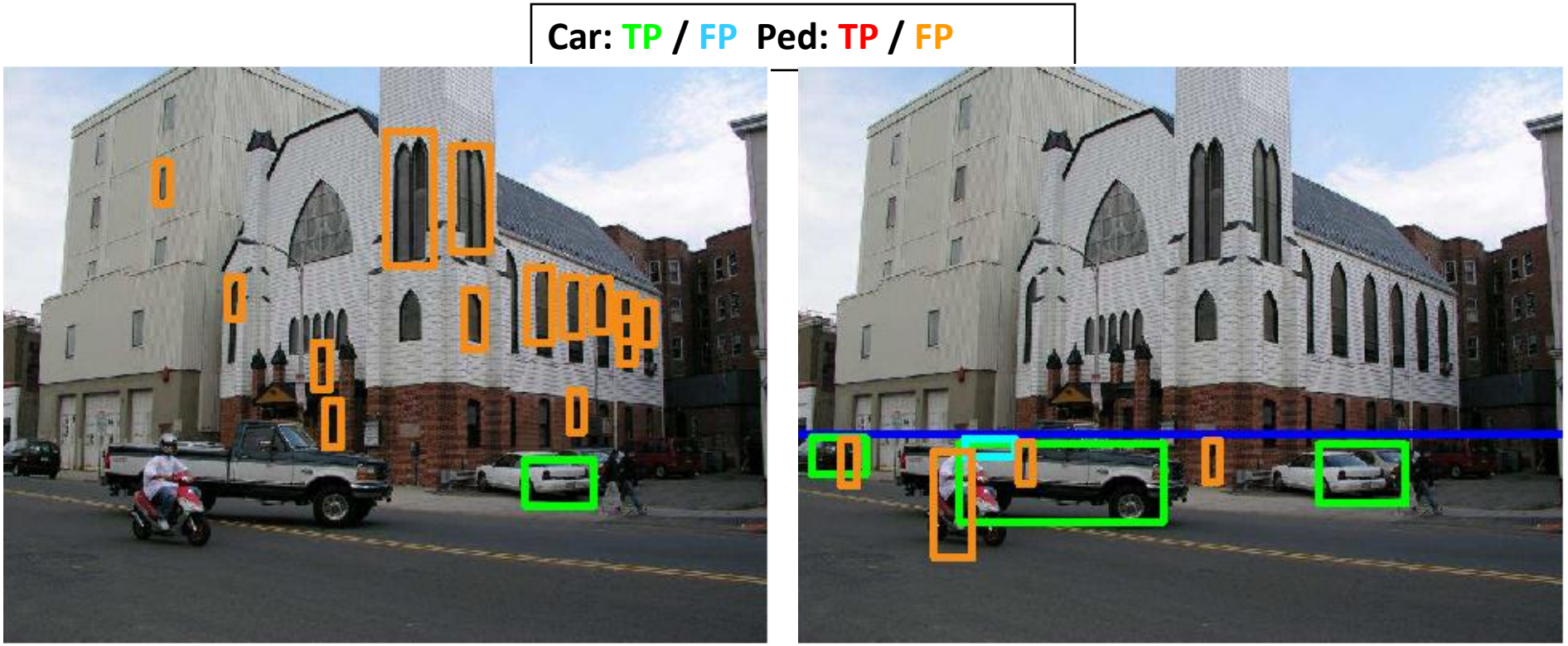


P(object | surfaces, viewpoint)

Hoiem et al. , CVPR2006

Depth ordering constraint

- Comparison



Initial: 1 TP / 14 FP

Final: 3 TP / 5 FP

Hoiem et al. , CVPR2006

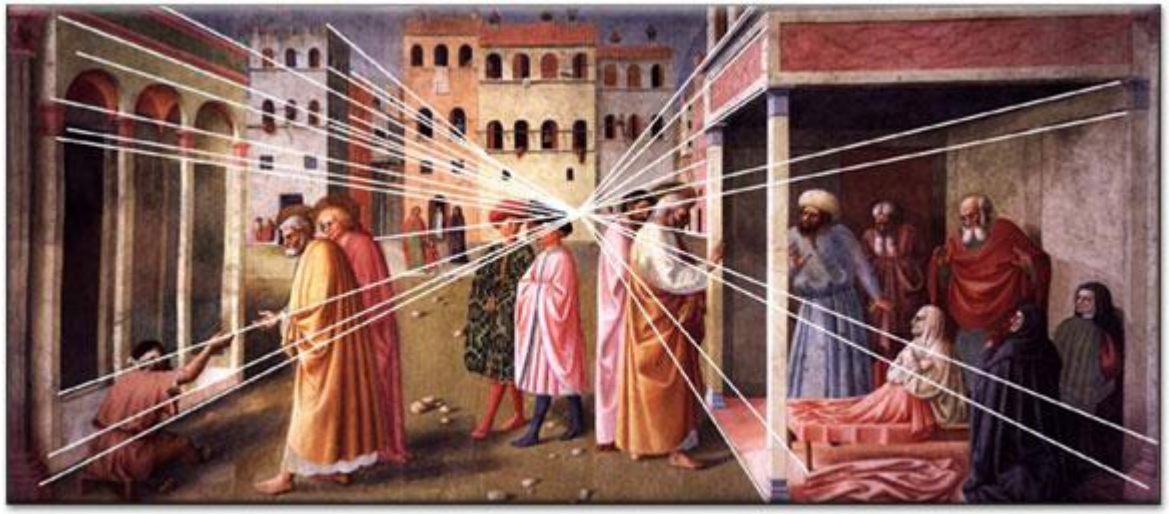
Local Detector from [Murphy-Torralba-Freeman 2003]

Expressing constraints with graphical models

- Outline of the talk
  - The idea of graphical mod
  - **Examples:**
    - Limiting the set of allowed deformations
    - Occlusion constraint
    - Depth ordering constraint
    - **Modeling the rules of perspective geometry**

Expressing constraints with graphical models

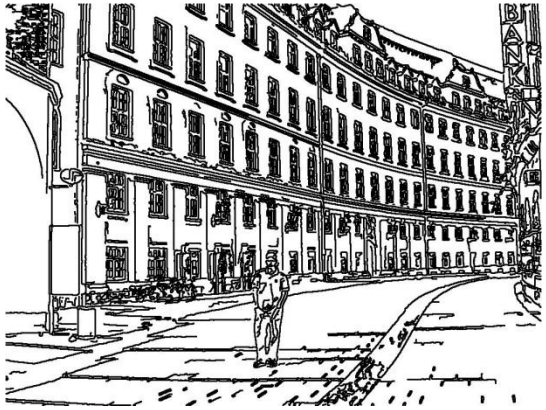
- Rules of perspective geometry
  - Straight lines lying on parallel planes in 3D intersect in the image plane in **vanishing point**
  - Vanishing point corresponding to vertical lines is called **zenith**
  - All vanishing points which correspond to horizontal lines lie on the same straight line called **horizon**



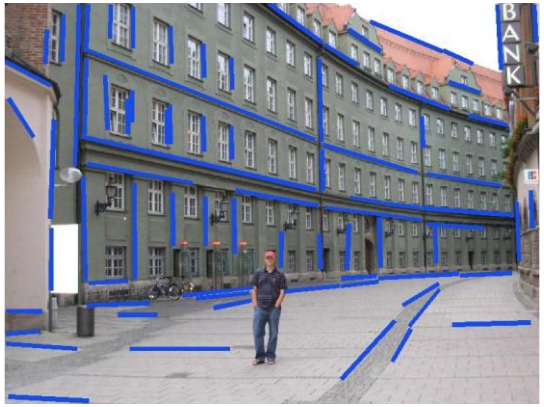
Modeling the rules of perspective geometry

- Geometric primitives and geometric parsing

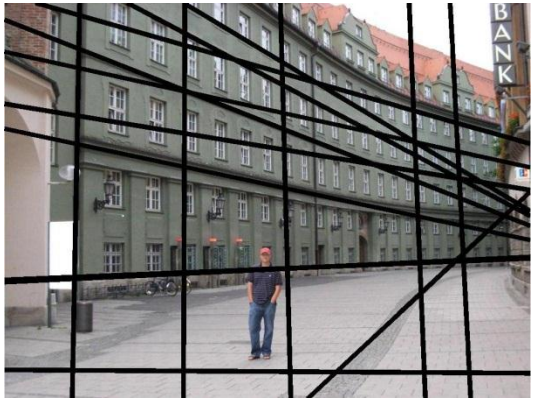
Image of man-made environment



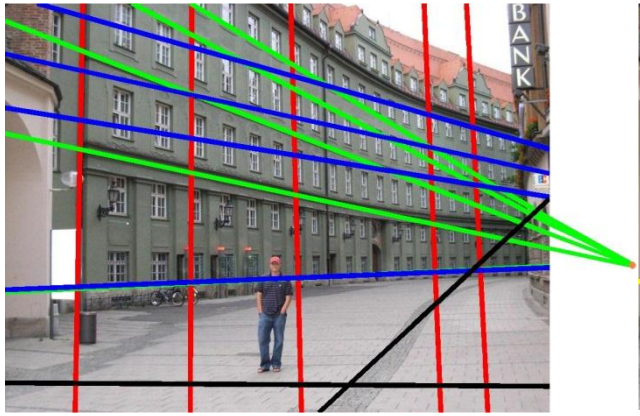
Edge pixels



Line segments



Lines



Vanishing points

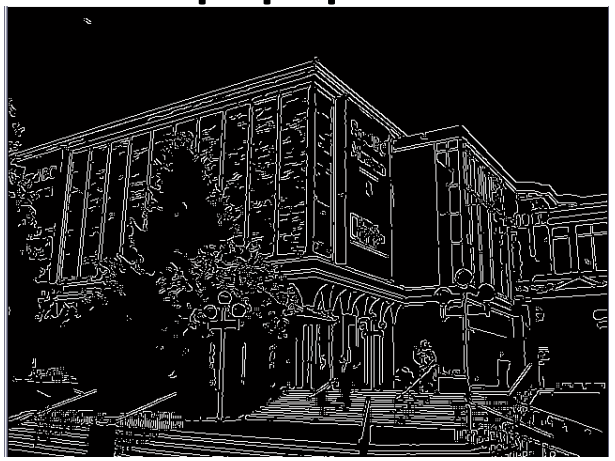
Zenith and horizon

Joint work with Elena Tretyak, Victor Lempitsky and Pushmeet Kohli, ECCV10

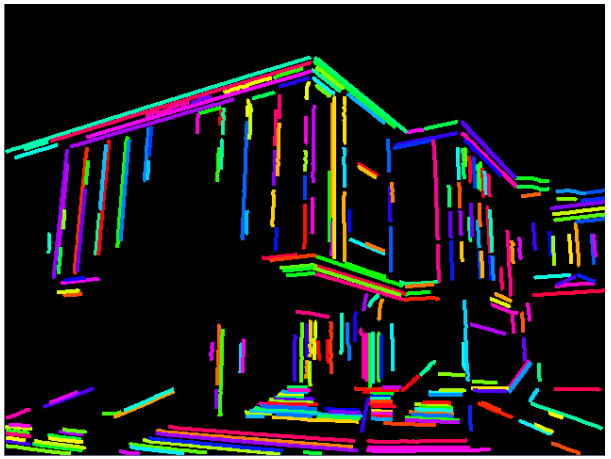
Modeling the rules of perspective geometry

- Traditional approach: bottom-up pipeline

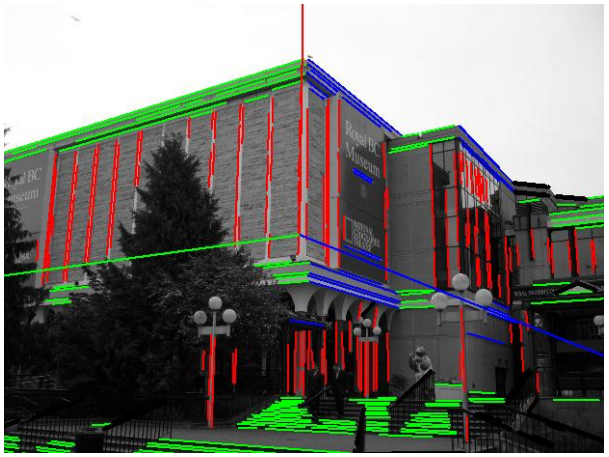
Input image



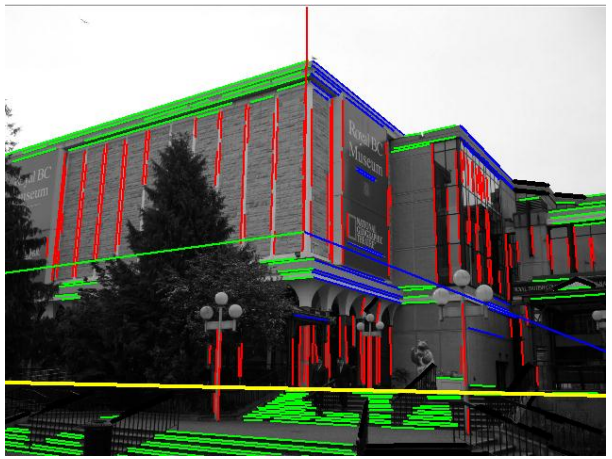
Edge map



1. *Grouping edge pixels into line segments*



2.Grouping line segments and VPs estimation



3. Horizon and zenith estimation

*Tuytelaars 1998, Antone 2000, Almansa 2003, Aguilera 2005, Boulanger 2006 , Tardif 2009

Modeling the rules of perspective geometry

- Energy function

  - All geometric primitives are detected in the simultaneously by energy minimization:

$$E_{total}(\mathbf{s},\mathbf{l},\mathbf{h},z|\mathbf{p}) = \sum_{i=1..P} E_{edge}(p_i|\mathbf{s}) + \sum_{i=1..S} E_{segment}(s_i|\mathbf{l}) + \sum_{i=1..L} E_{line}(l_i|\mathbf{h},z) + \sum_{1\le i<j\le H} E_{horizon}(h_i,h_j|z) + E_{prior}(\mathbf{s},\mathbf{l},\mathbf{h}),$$

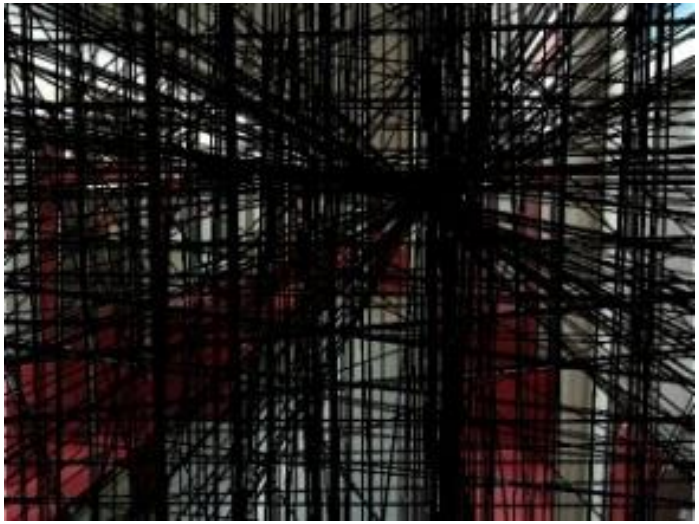
p – edge pixels, ***s*** – line segments, ***l*** – lines,
z – zenith, ***h*** – horizontal vanishing points

Joint work with Elena Tretyak, Victor Lempitsky and Pushmeet Kohli, ECCV10

Modeling the rules of perspective geometry

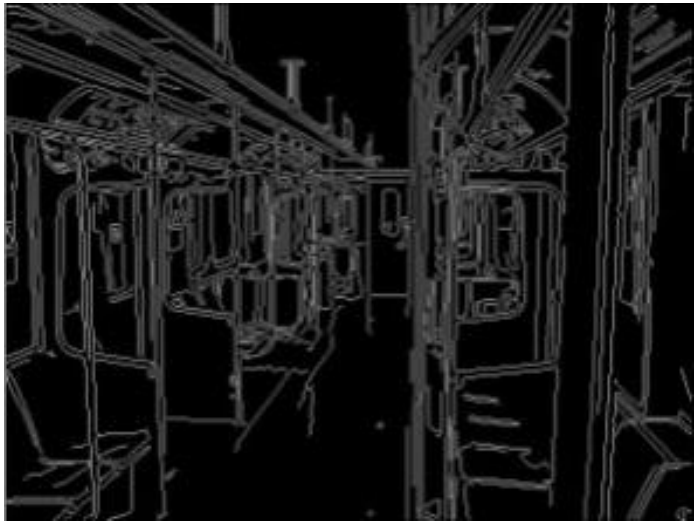
- Discretization of the model

Candidate vanishing points



Candidate lines

Candidate line segments



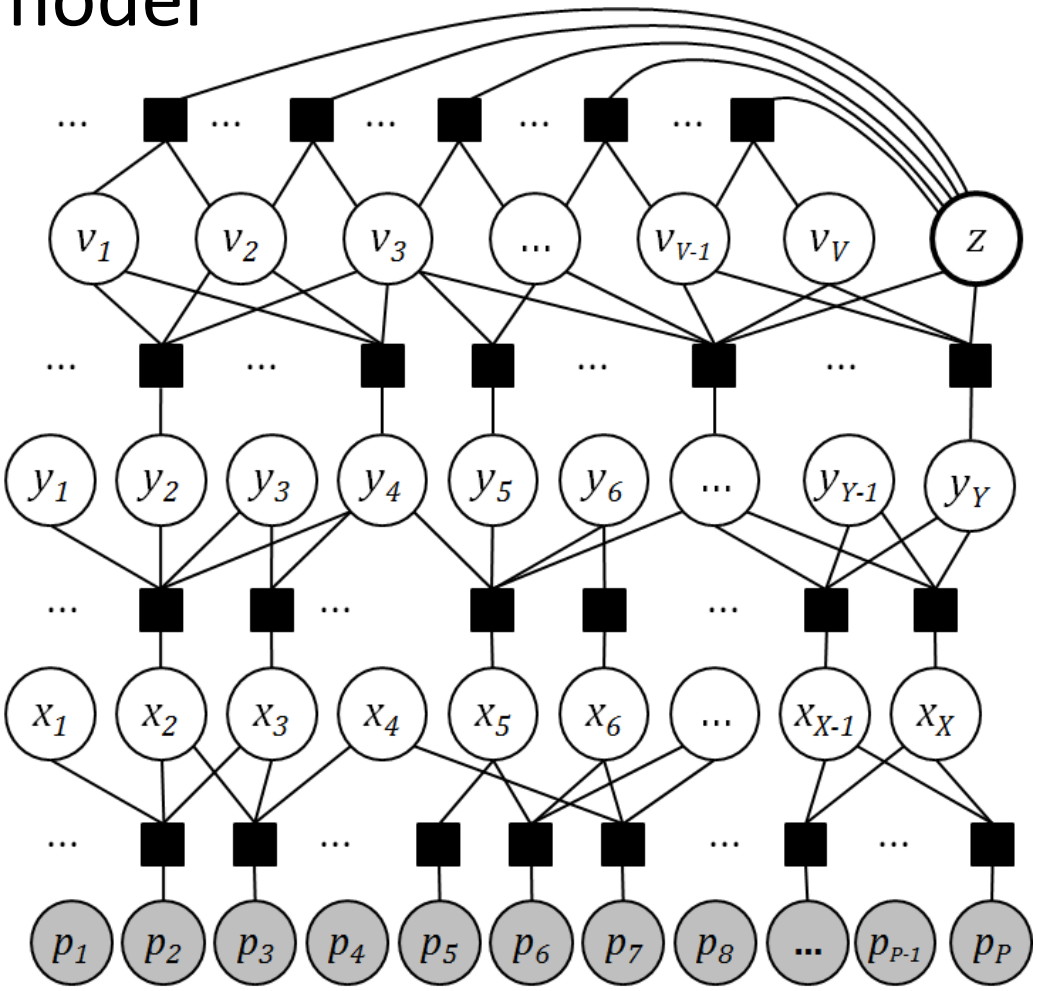
Candidate edge pixels

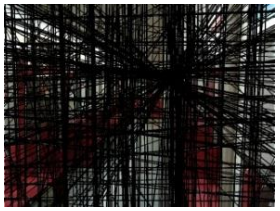
Modeling the rules of perspective geometry

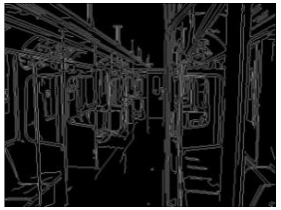
- Graphical model

■ – factors correspond to the potentials of the energy function

○ – nodes correspond to geometric primitives





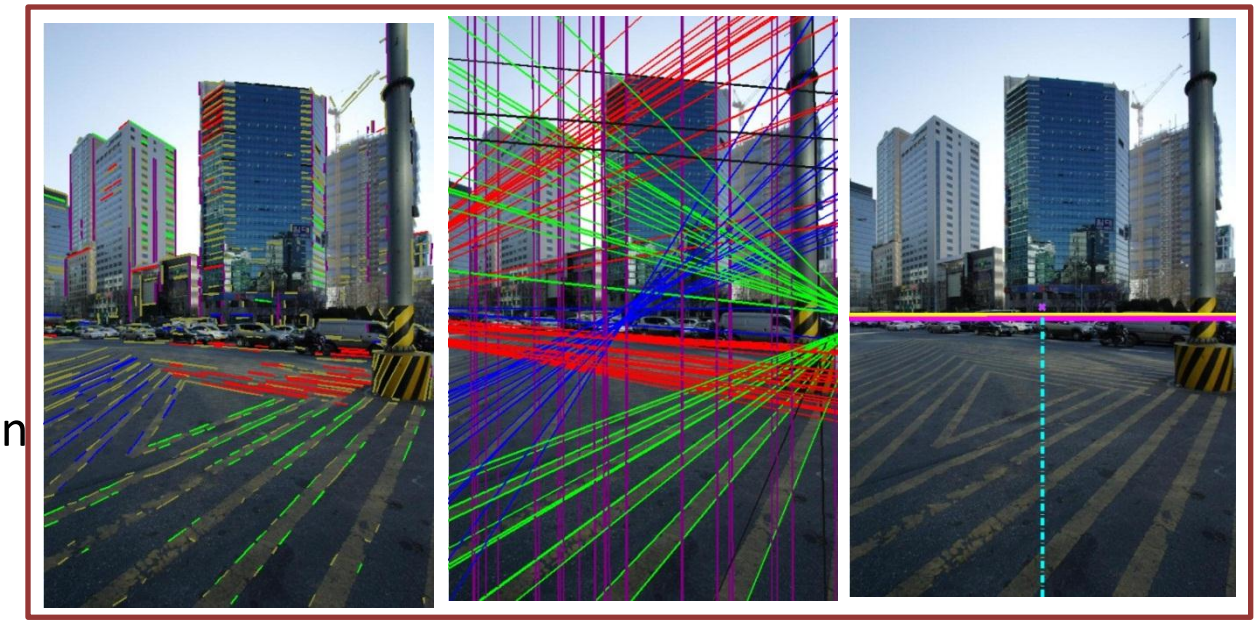


Joint work with Elena Tretyak, Victor Lempitsky and Pushmeet Kohli, ECCV10

Modeling the rules of perspective geometry

- Results of geometric parsing

Energy minimization



Input image

Detected geometric primitives

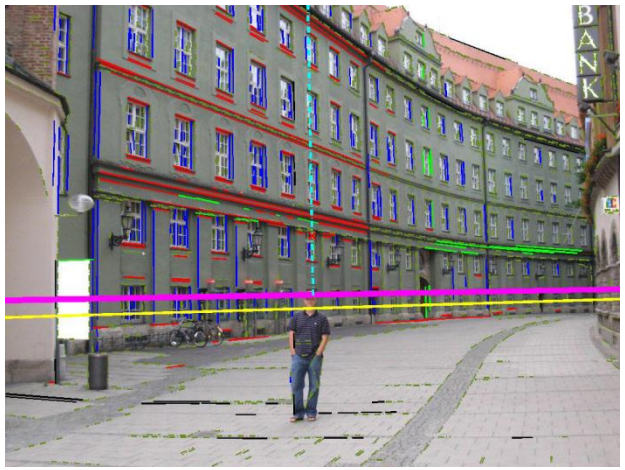
Joint work with Elena Tretyak, Victor Lempitsky and Pushmeet Kohli, ECCV10

Modeling the rules of perspective geometry

- Comparison

Result of geometric parsing with full energy

Code available online!!!

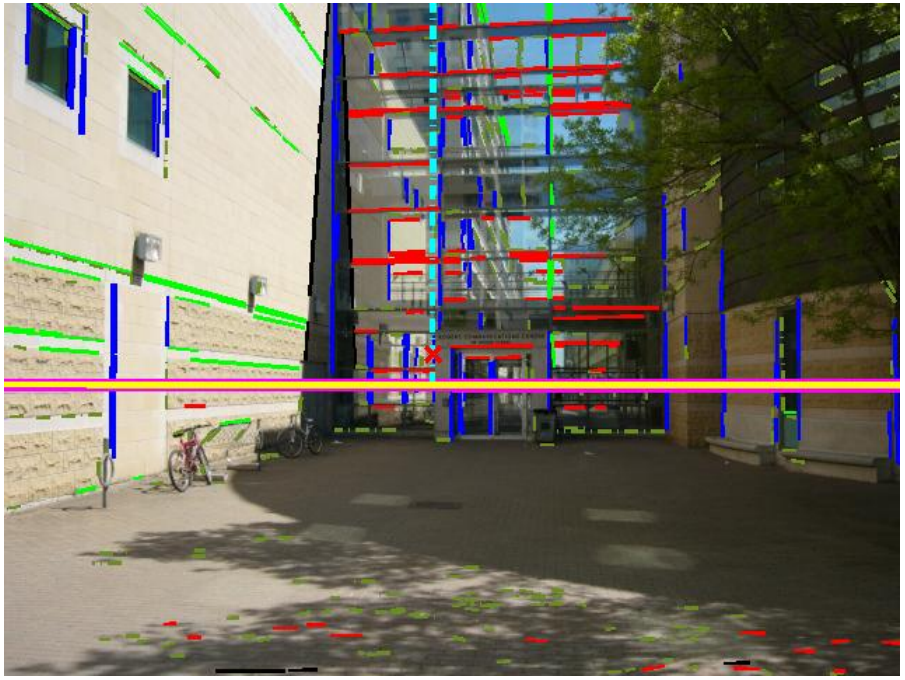


Omitting horizon constraint

Bottom-up pipeline

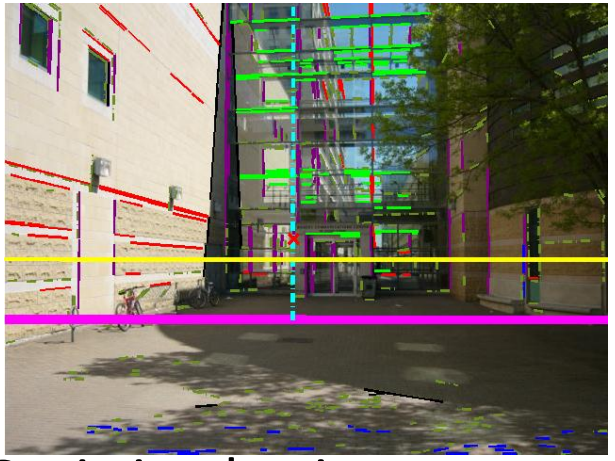
Modeling the rules of perspective geometry

- Comparison

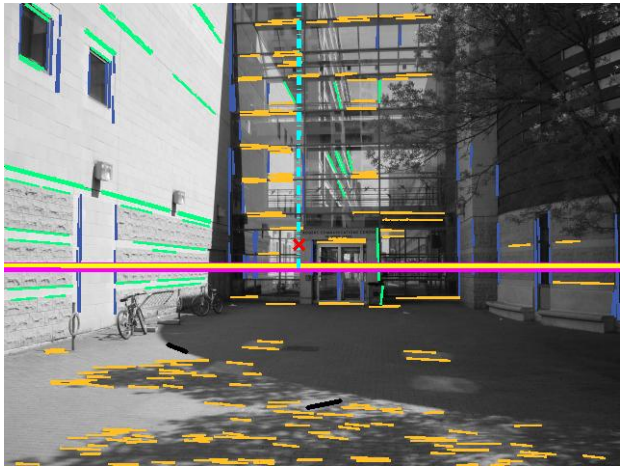


Result of geometric parsing with full energy

Code available online!!!



Omitting horizon constraint



Bottom-up pipeline

Application: single-view geometry

single image

3d model

Joint work with Vadim Konushin, Anton Yakubenko, Hwasup Lim, and Anton Konushin, ECCV08

Text recognition in natural images

Text recognition in natural images

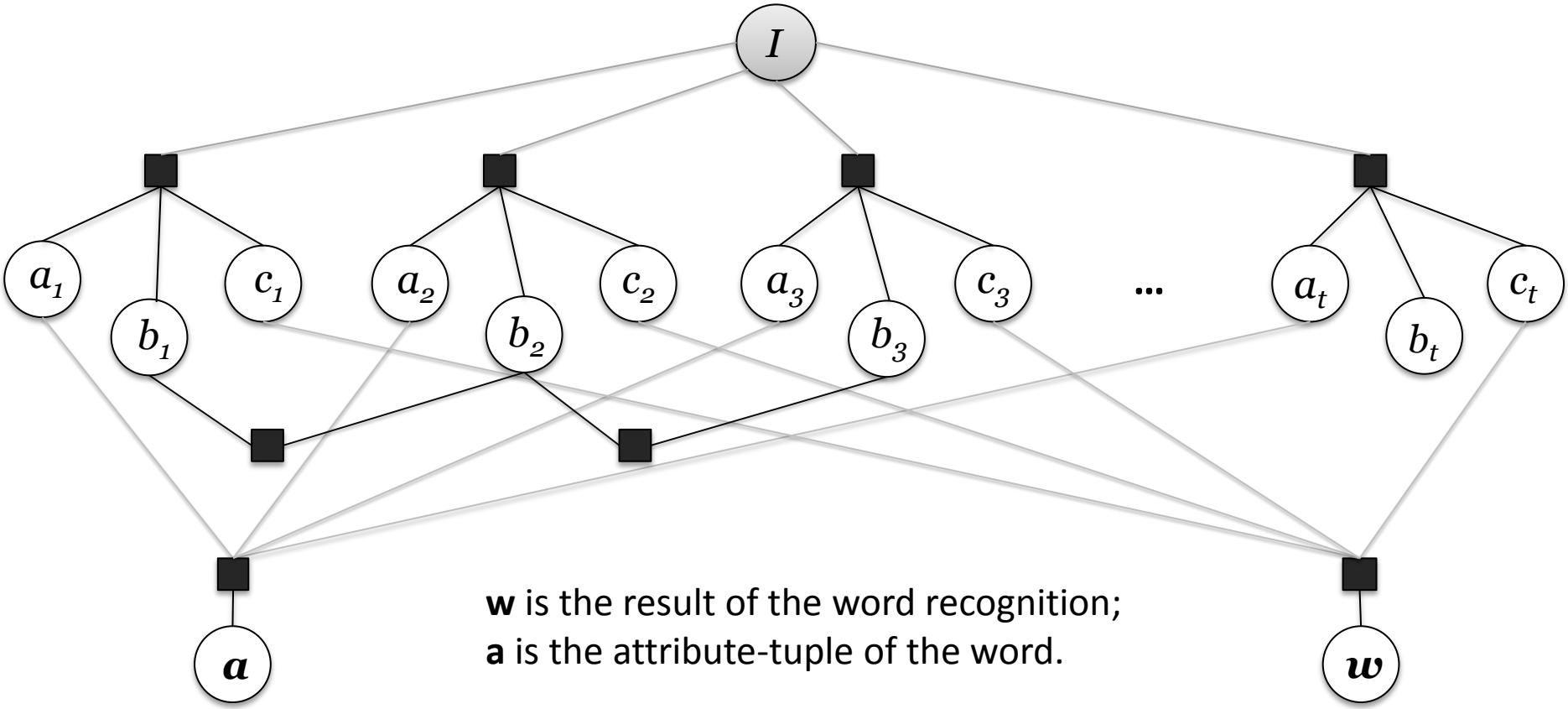
- Constraints:
  - Direction of text line
  - Letters have the same color
  - Letters have similar size
  - Letters have the same font

Text recognition in natural images

a1,…, at - attributes of the characters

b1,…, bt - character locations;

c1,…,ct- recognition results (e.g. an alphanumeric symbol).



w is the result of the word recognition;
a is the attribute-tuple of the word.

Text recognition in natural images

- Comparison to ICDAR 2011 Robust Reading Competition

	ICDAR 2011
TH-OCR System	41.2%
KAIST AIPR System	35.6%
Neumann's Method	33.11%
Our method (case sensitive)	**56.4%**
Our method (case insensitive)	*66.7%*

Text recognition in natural images

- Comparison in word spotting scenario

	ICDAR	SVT
Wang et al. [1]	59%	59%
Wang et al. [2]	62%	57%
Ours, full model	**82.8%**	**72.9%**
Ours w/o attribute consistency, w/o alternative character recognition	60%	-
Ours w/o attribute consistency	80.7%	-
Ours, full model, large lexicon	78.5%	-

Contacts

- Microsoft Summer School on Concurrency http://research.microsoft.com/en-us/events/rss2012/
  - срок регистрации до 30 марта:
  - 22–29 августа, 2012 | Санкт-Петербург, Россия

- Мой e-mail obarinova@graphics.cs.msu.ru

Thank you for attention!