An Additive Latent Feature Model for

Mario Fritz
UC Berkeley

Michael Black
Brown University

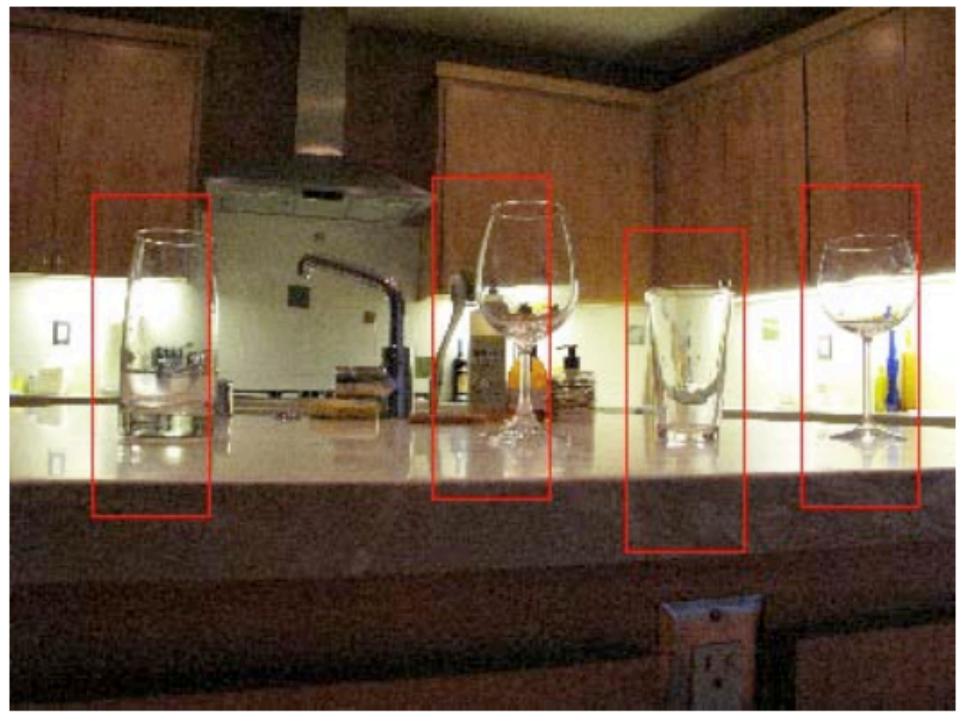
Gary Bradski
Willow Garage

Sergey Karayev
UC Berkeley

Trevor Darrell
UC Berkeley

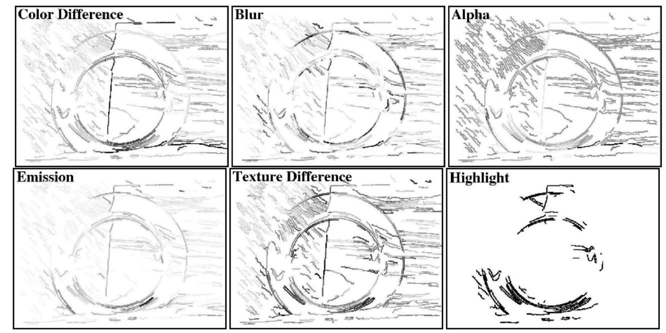
Motivation

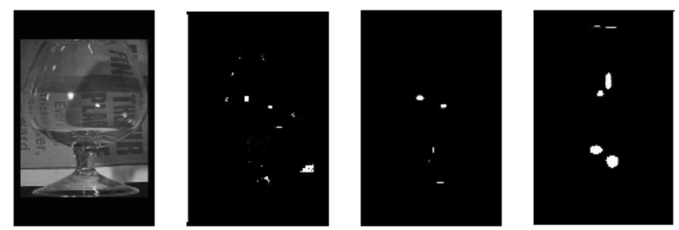
- Transparent objects are ubiquitous in domestic environments
- Relevant to domestic service robots
- Traditional local feature approach inappropriate
- Full physical model intractable

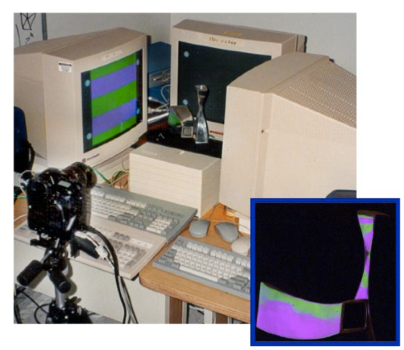


Related Work

- Material recognition:
  - Finding Glass [McHenry@CVPR05/06]
  - Detecting Specular Surfaces in Natural Images [DelPozo@CVPR07]
  - Classifying Materials from their Reflectance Properties [Nillius@ECCV04]
  - Low-Level Image Cues in the Perception of Translucent Materials [Fleming, Transactions on Applied Perception ’05]
- Recognition by Specularities:
  - Using Specularities for Recognition [Osadchy@ICCV03]
- Transparent Motion and Layered Phenomena
  - E.g. [Roth@CVPR06], [Ben-Ezra@ICCV03], [Darrell@CVPR93] …
- Acquisition and rendering of refractive patterns
  - Environment Matting and Composition [Zongker@Siggraph99]

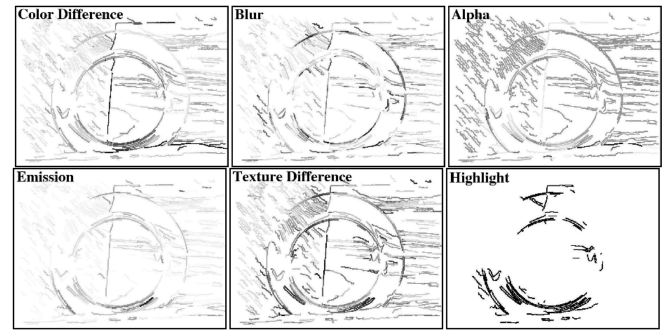




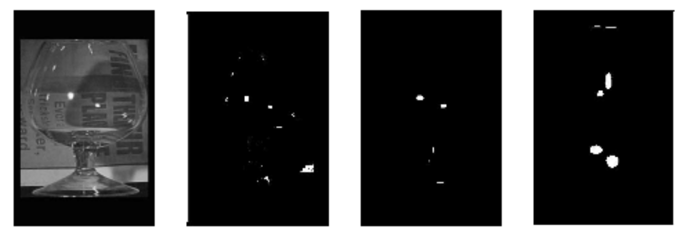


Related Work

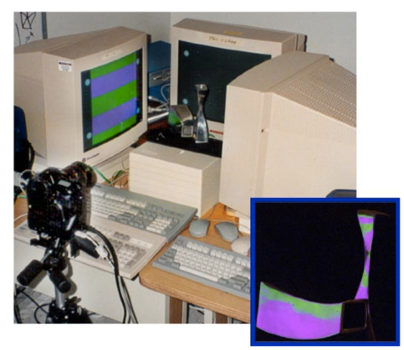
- **Material recognition:**
  - **Finding Glass [McHenry@CVPR05/06]**
  - **Detecting Specular Surfaces in Natural Images [DelPozo@CVPR07]**
  - **Classifying Materials from their Reflectance Properties [Nillius@ECCV04]**
  - **Low-Level Image Cues in the Perception of Translucent Materials [Fleming, Transactions on Applied Perception '05]**



- Recognition by Specularities:
  - Using Specularities for Recognition [Osadchy@ICCV03]

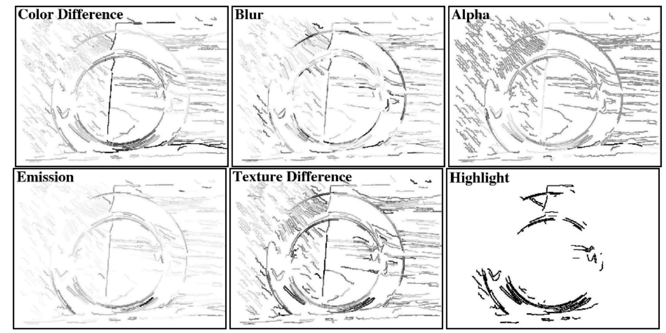


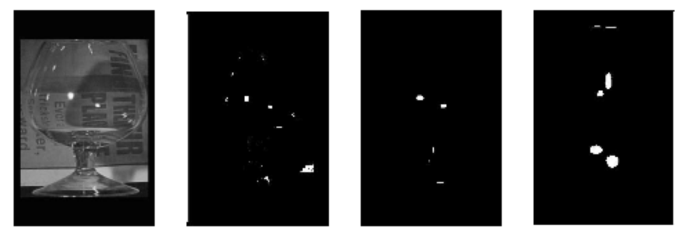
- Transparent Motion and Layered Phenomena
  - E.g. [Roth@CVPR06], [Ben-Ezra@ICCV03], [Darrell@CVPR93] …
- Acquisition and rendering of refractive patterns
  - Environment Matting and Composition [Zongker@Siggraph99]

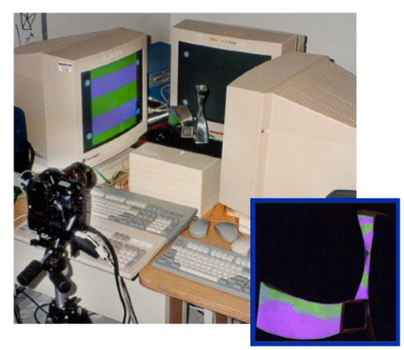


Related Work

- Material recognition:
  - Finding Glass [McHenry@CVPR05/06]
  - Detecting Specular Surfaces in Natural Images [DelPozo@CVPR07]
  - Classifying Materials from their Reflectance Properties [Nillius@ECCV04]
  - Low-Level Image Cues in the Perception of Translucent Materials [Fleming, Transactions on Applied Perception ’05]
- **Recognition by Specularities:**
  - **Using Specularities for Recognition [Osadchy@ICCV03]**
- Transparent Motion and Layered Phenomena
  - E.g. [Roth@CVPR06], [Ben-Ezra@ICCV03], [Darrell@CVPR93] …
- Acquisition and rendering of refractive patterns
  - Environment Matting and Composition [Zongker@Siggraph99]

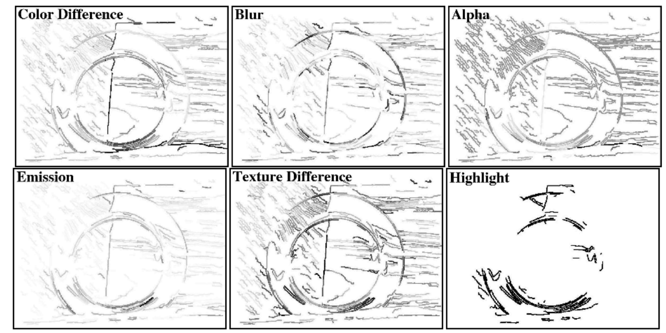




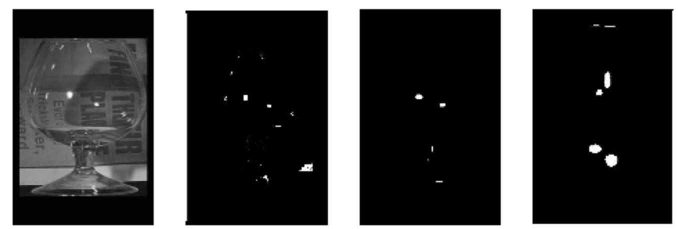


Related Work

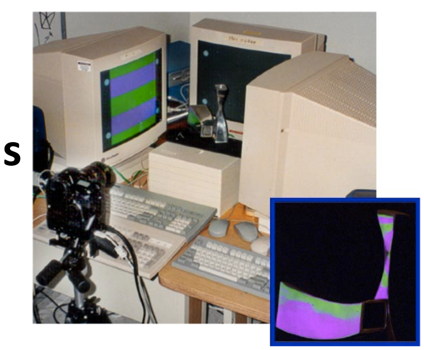
- Material recognition:
  - Finding Glass [McHenry@CVPR05/06]
  - Detecting Specular Surfaces in Natural Images [DelPozo@CVPR07]
  - Classifying Materials from their Reflectance Properties [Nillius@ECCV04]
  - Low-Level Image Cues in the Perception of Translucent Materials [Fleming, Transactions on Applied Perception '05]



- Recognition by Specularities:
  - Using Specularities for Recognition [Osadchy@ICCV03]

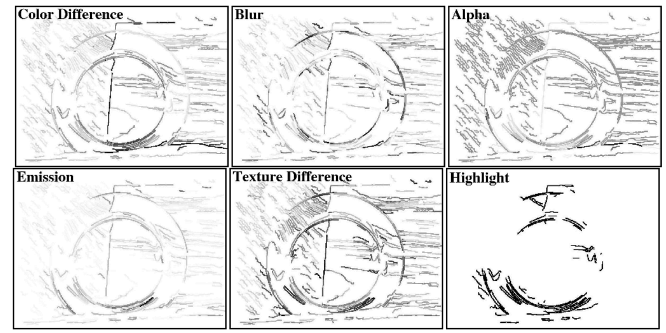


- **Transparent Motion and Layered Phenomena**
  - **E.g. [Roth@CVPR06], [Ben-Ezra@ICCV03], [Darrell@CVPR93] …**
- **Acquisition and rendering of refractive patterns**
  - **Environment Matting and Composition [Zongker@Siggraph99]**

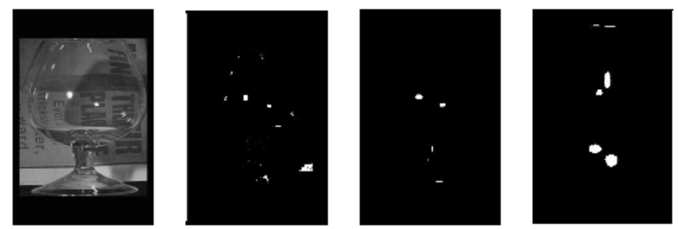


Related Work

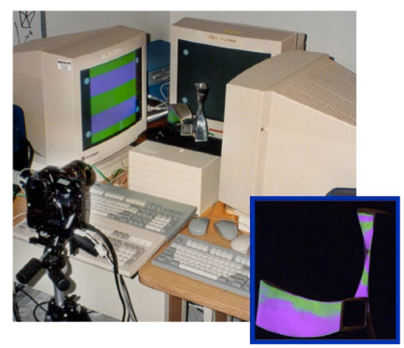
- Material recognition:
  - Finding Glass [McHenry@CVPR05/06]
  - Detecting Specular Surfaces in Natural Images [DelPozo@CVPR07]
  - Classifying Materials from their Reflectance Properties [Nillius@ECCV04]
  - Low-Level Image Cues in the Perception of Translucent Materials [Fleming, Transactions on Applied Perception '05]



- Recognition by Specularities:
  - Using Specularities for Recognition [Osadchy@ICCV03]



- Transparent Motion and Layered Phenomena
  - E.g. [Roth@CVPR06], [Ben-Ezra@ICCV03], [Darrell@CVPR93] …
- Acquisition and rendering of refractive patterns
  - Environment Matting and Composition [Zongker@Siggraph99]

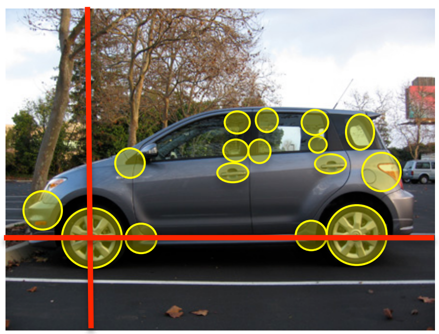


- **Non of these approaches addresses transparent objects recognition in real-world conditions**

Traditional Local Feature-based Recognition

Codebook: $\begin{pmatrix} \text{cluster}\vec{\text{C}}\text{enter}_1 \\ \vdots \\ \text{cluster}\vec{\text{C}}\text{enter}_N \end{pmatrix}$

x_1



y_1

$$\begin{pmatrix} x_1 & y_1 & f(\vec{I}_1) \\ \vdots & \vdots & \vdots \\ x_N & y_N & f(\vec{I}_N) \end{pmatrix} \xrightarrow{\text{quantize}} \begin{pmatrix} \hat{x_1} & \hat{y_1} & \text{clusterId}_1 \\ \vdots & \vdots & \vdots \\ \hat{x_N} & \hat{y_N} & \text{clusterId}_N \end{pmatrix}$$

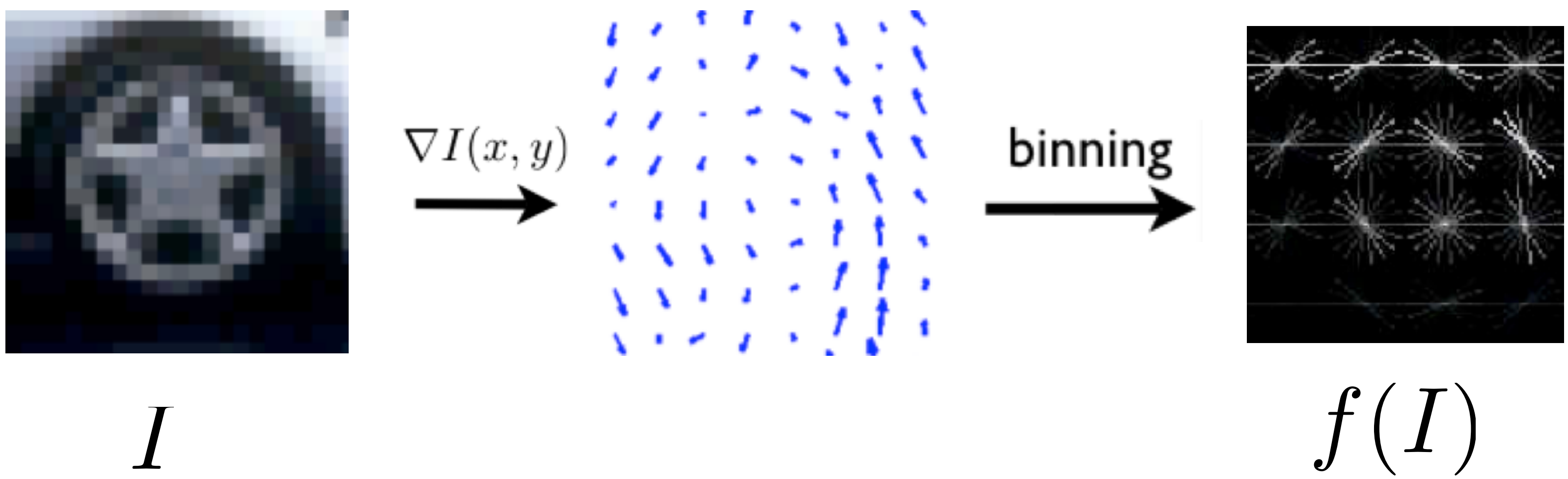
histogram

$\vec{d}$

classifier

- Codebook clusters assume prototypical global patch appearance

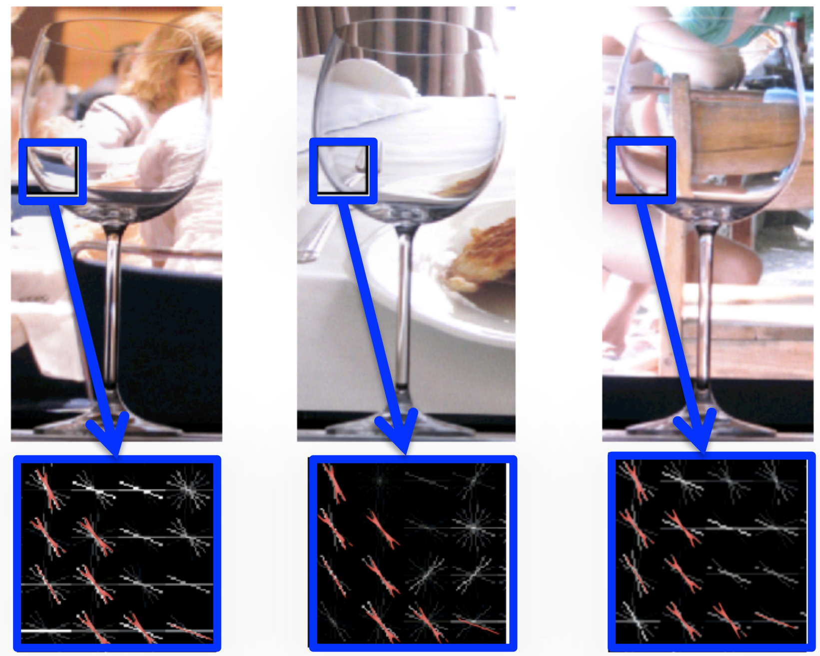
SIFT-type Descriptors



- SIFT is popular choice for local feature computation f
- It performs spatial binning of orientation quantized gradient information
- Unnormalized distribution over local gradient statistics
- We will use the a particular visualization as proposed for the related HOG method

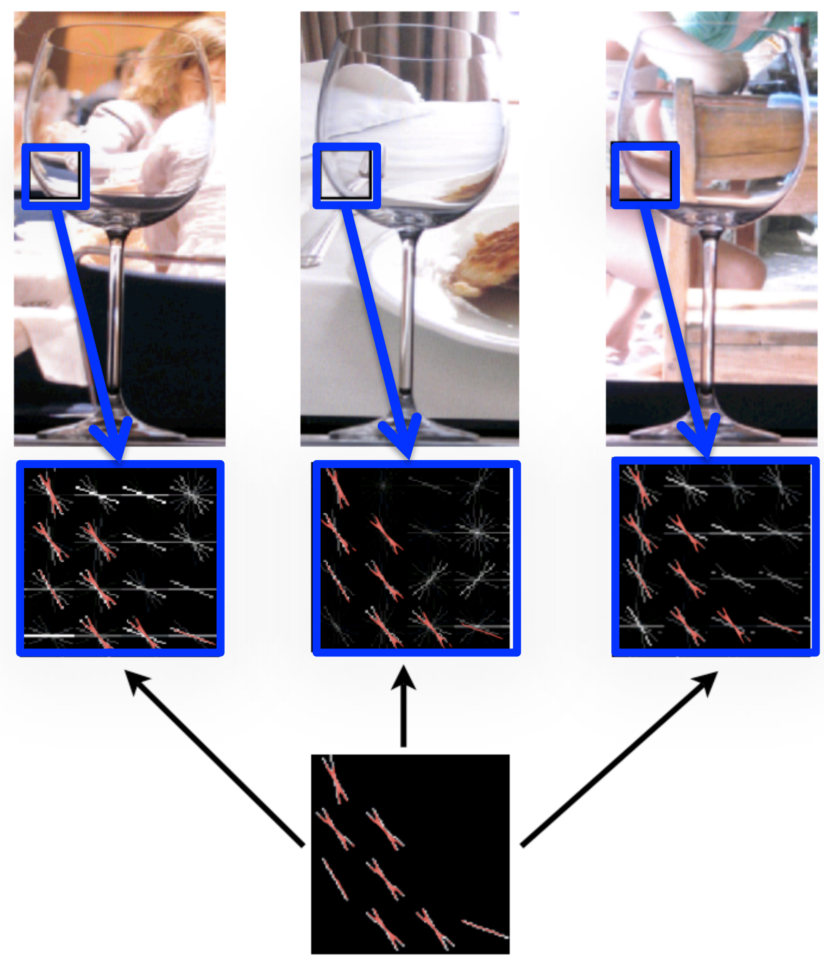
The Problem of Transparency

- Significant variation in patch appearance
- Often gradient energy is dominated by background

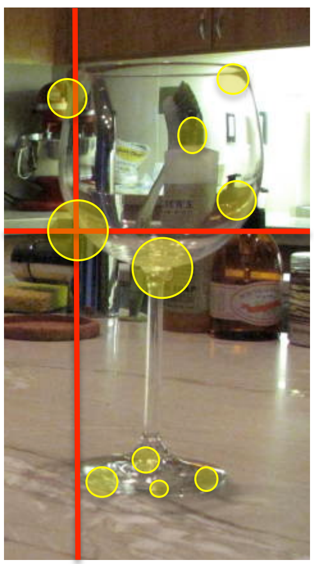


The Problem of Transparency

- Significant variation in patch appearance
- Often gradient energy is dominated by background
- ... but common latent structure



The Problem of Transparency



Codebook: $\begin{pmatrix} \mathrm{cluster}\vec{\mathrm{C}}\mathrm{enter}_1 \\ \vdots \\ \mathrm{cluster}\vec{\mathrm{C}}\mathrm{enter}_N \end{pmatrix}$

$$\begin{pmatrix} x_1 & y_1 & f(\vec{I}_1) \\ \vdots & \vdots & \vdots \\ x_N & y_N & f(\vec{I}_N) \end{pmatrix} \xrightarrow{\text{quantize}} \begin{pmatrix} \hat{x_1} & \hat{y_1} & \mathrm{clusterId}_1 \\ \vdots & \vdots & \vdots \\ \hat{x_N} & \hat{y_N} & \mathrm{clusterId}_N \end{pmatrix}$$

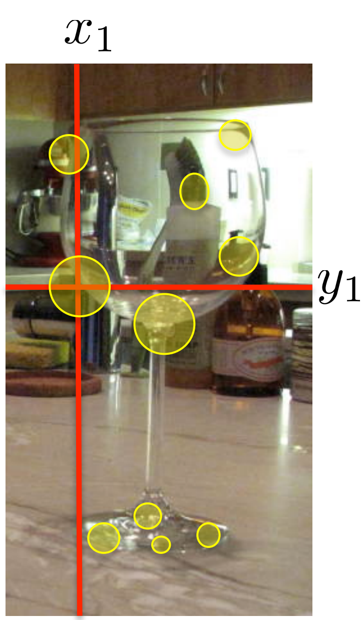
histogram

$\vec{d}$

classifier

- Codebook clusters assume prototypical global patch appearance

The Problem of Transparency



Codebook: $\begin{pmatrix} \text{cluster}\vec{\text{C}}\text{enter}_1 \\ \vdots \\ \text{cluster}\vec{\text{C}}\text{enter}_N \end{pmatrix}$

$$\begin{pmatrix} x_1 & y_1 & f(\vec{I}_1) \\ \vdots & \vdots & \vdots \\ x_N & y_N & f(\vec{I}_N) \end{pmatrix} \xrightarrow{\text{quantize}} \begin{pmatrix} \hat{x_1} & \hat{y_1} & \text{clusterId}_1 \\ \vdots & \vdots & \vdots \\ \hat{x_N} & \hat{y_N} & \text{clusterId}_N \end{pmatrix}$$

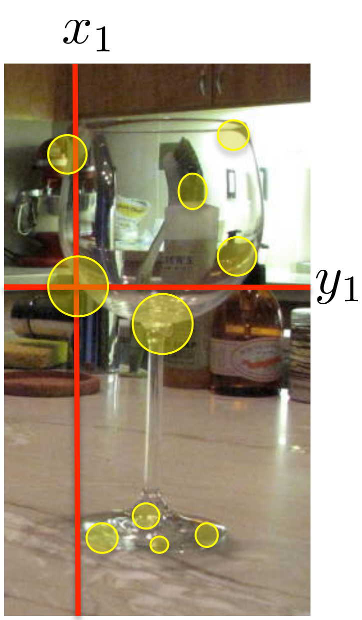
histogram

$\vec{d}$

classifier

- Codebook clusters assume prototypical global patch appearance

Key Idea: Local Latent Factorization



Components: $\begin{pmatrix} \text{component}_1 \\ \vdots \\ \text{component}_T \end{pmatrix}$

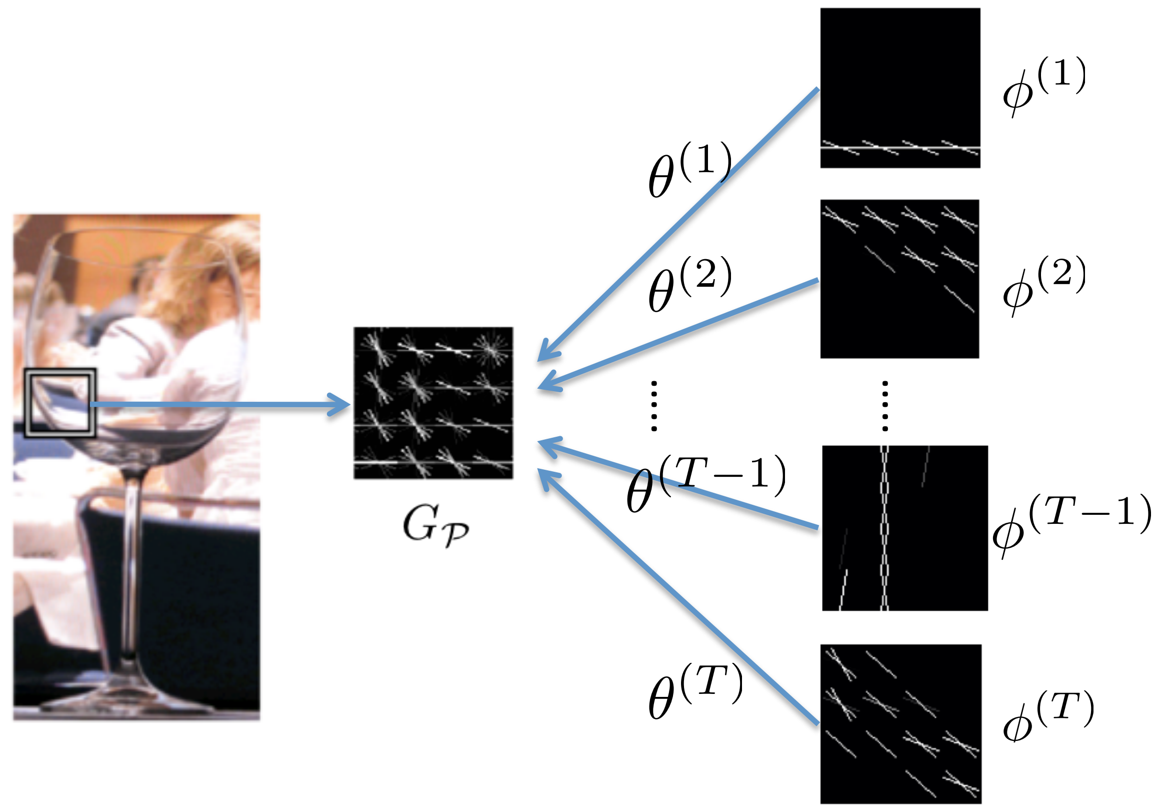
$$\begin{pmatrix} x_1 & y_1 & f(\vec{I}_1) \\ \vdots & \vdots & \vdots \\ x_N & y_N & f(\vec{I}_N) \end{pmatrix} \rightarrow \boxed{\text{Latent component model}} \rightarrow \begin{pmatrix} \hat{x_1} & \hat{y_1} & \text{componentIds}_1 \\ \vdots & \vdots & \vdots \\ \hat{x_N} & \hat{y_N} & \text{componentIds}_N \end{pmatrix}$$

histogram → $\vec{d}$ → classifier

- Codebook is replaced by a set of latent components

Local Additive Feature Model

- Factor gradient descriptor into
  - Unknown non-negative mixture weights
  - Unknown mixture components
- Regularize with sparsity assumption
- Advantages vs. e.g. VQ, PCA:
  - Additive model allows for superimposed structures
  - Appropriate model for factorizing local gradient distribution
  - No reliance on global patch appearance



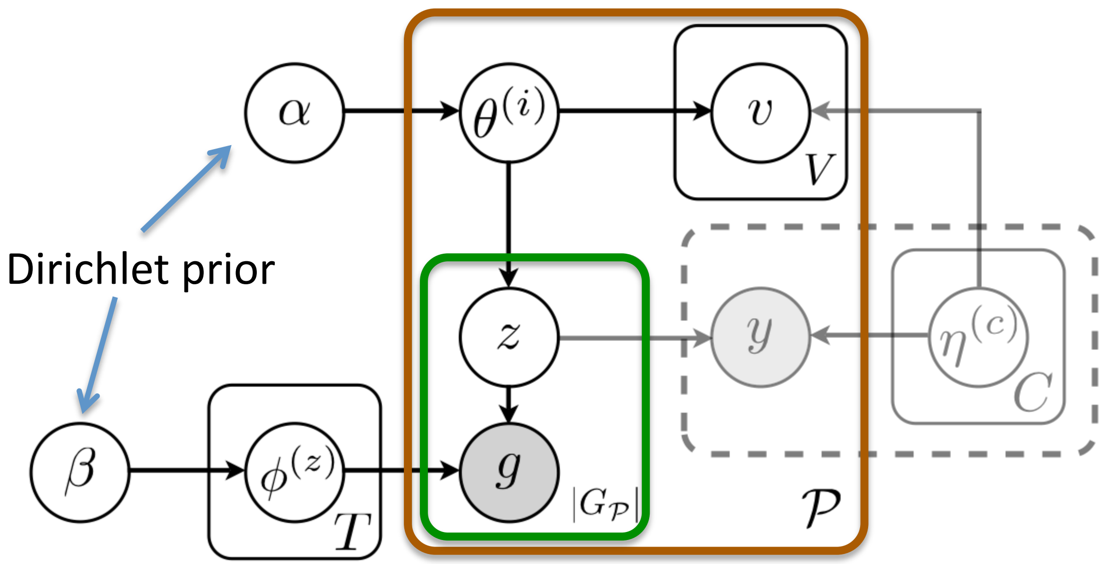
$$G_{\mathcal{P}} = \sum_i \theta^{(i)} \phi^{(i)}$$

LDA-SIFT

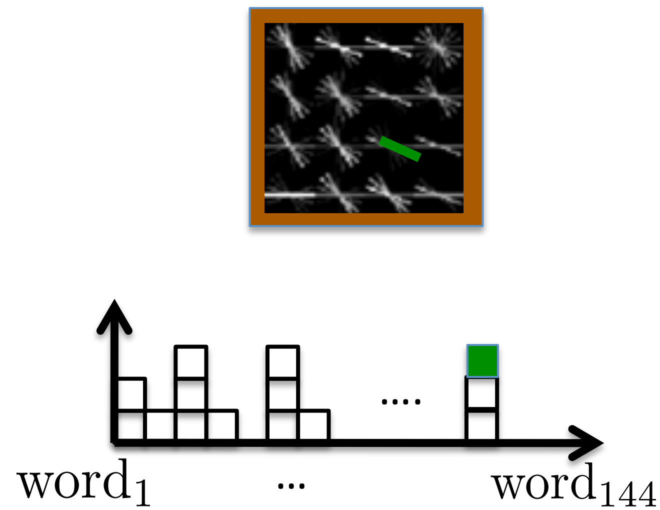
Factor SIFT descriptor into latent components using LDA/sLDA [Blei03,Griffiths04,Blei07]:

- additivity is realized as multinomial mixture model
- sparsity assumption is implemented as Dirichlet priors

Graphical model



Document = Patch



Transparent

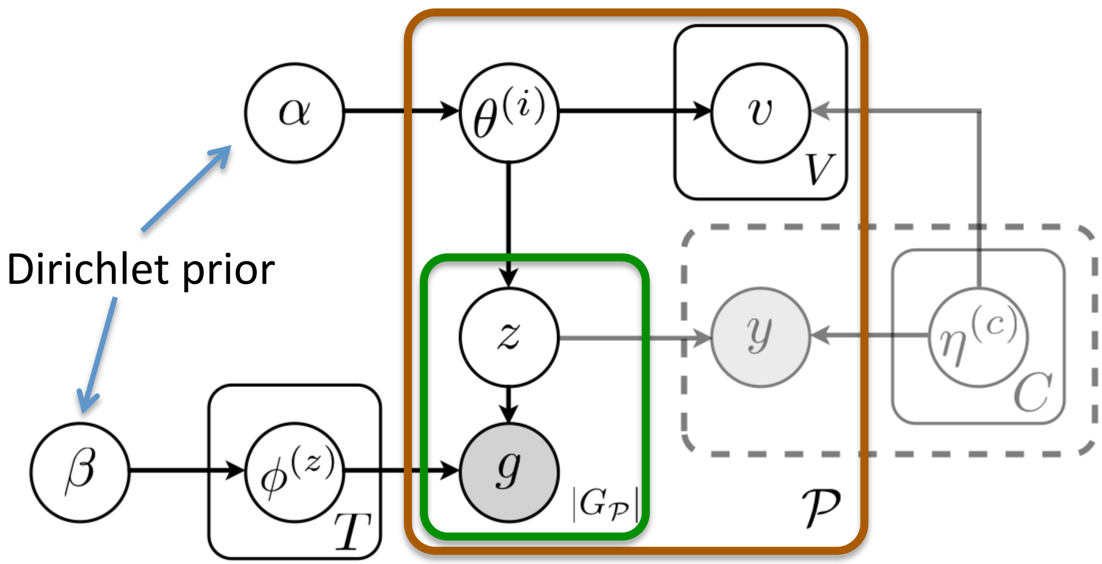
Different from projections like PCA -> Inhibition effects

LDA-SIFT

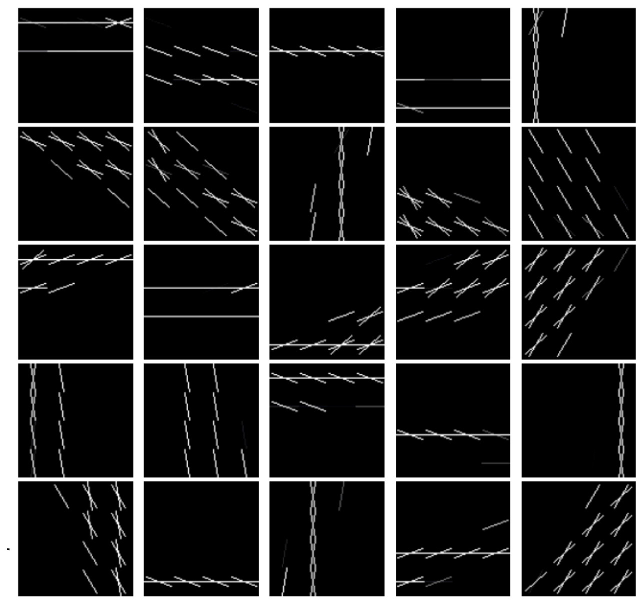
Factor SIFT descriptor into latent components using LDA/sLDA [Blei03,Griffiths04,Blei07]:

- additivity is realized as multinomial mixture model
- sparsity assumption is implemented as Dirichlet priors

Graphical model



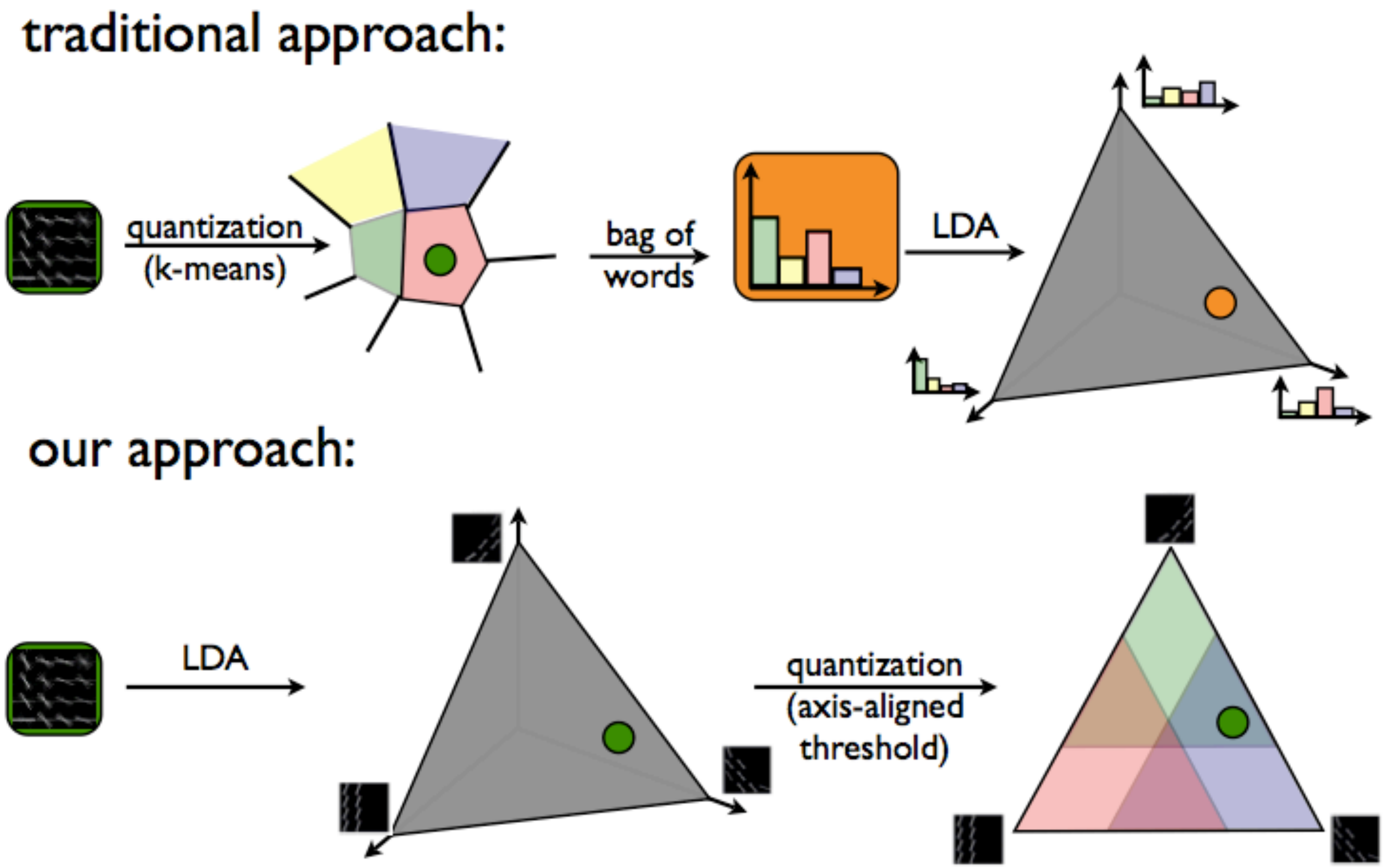
Learnt mixture components



Transparent

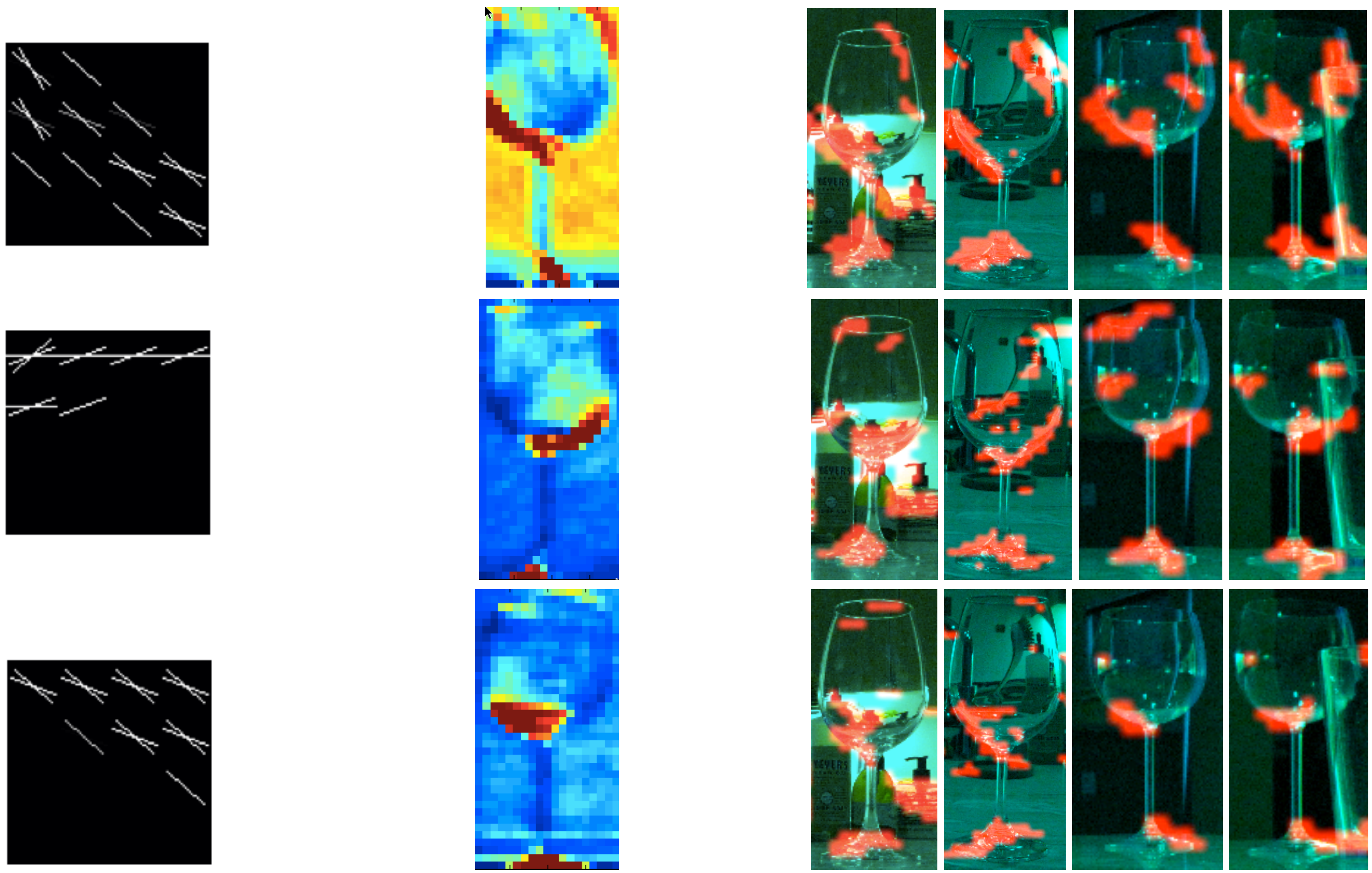
Different from projections like PCA -> Inhibition effects

Comparison to previous SIFT/LDA

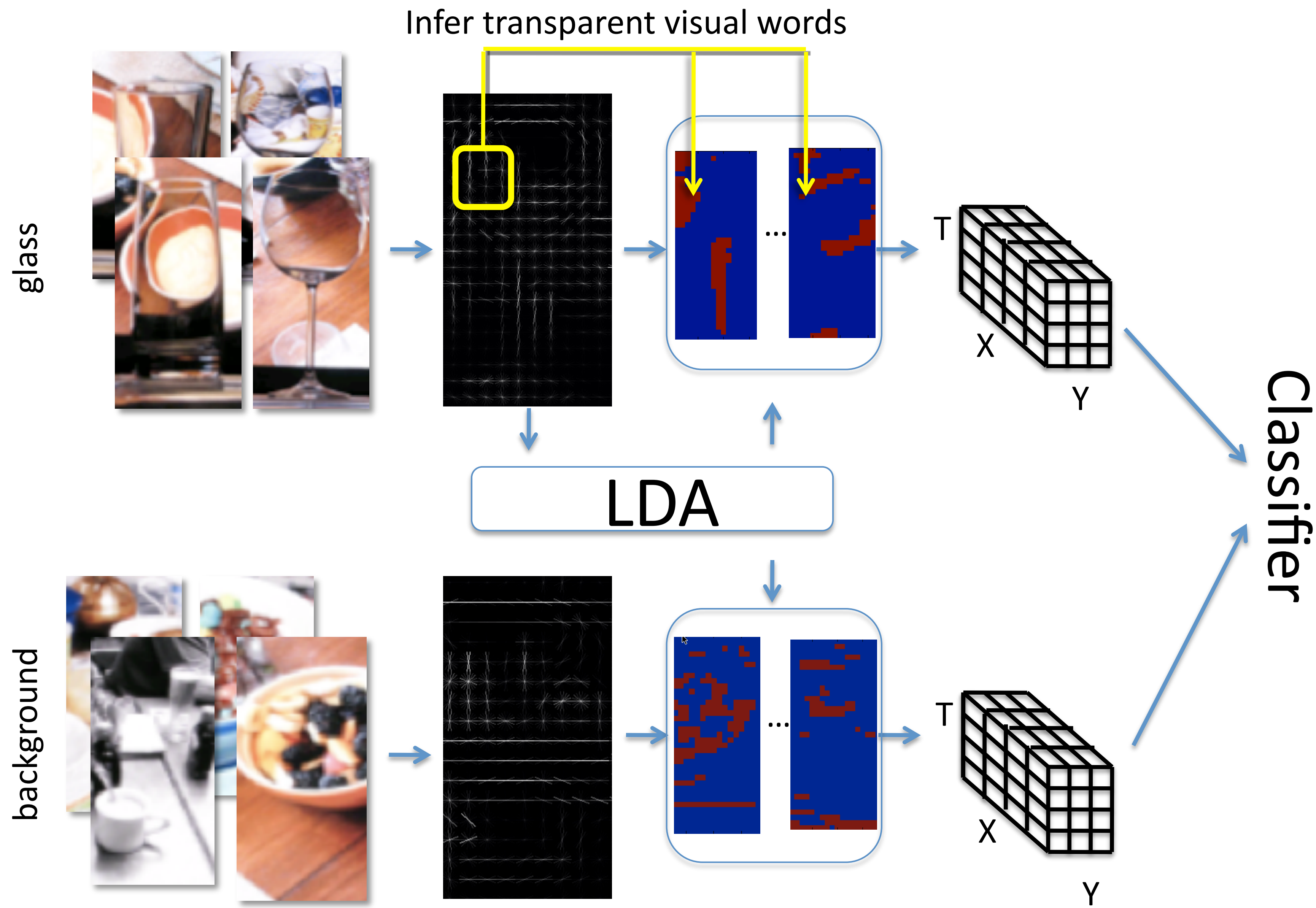


Transparent Visual Words

Latent component | Average occurrence on train | Occurrences on test



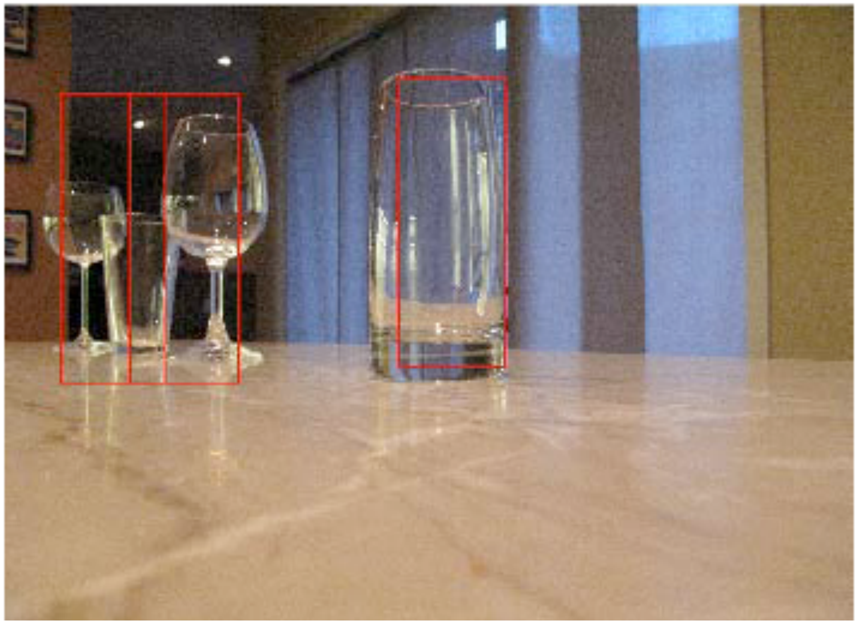
Recognition Architecture

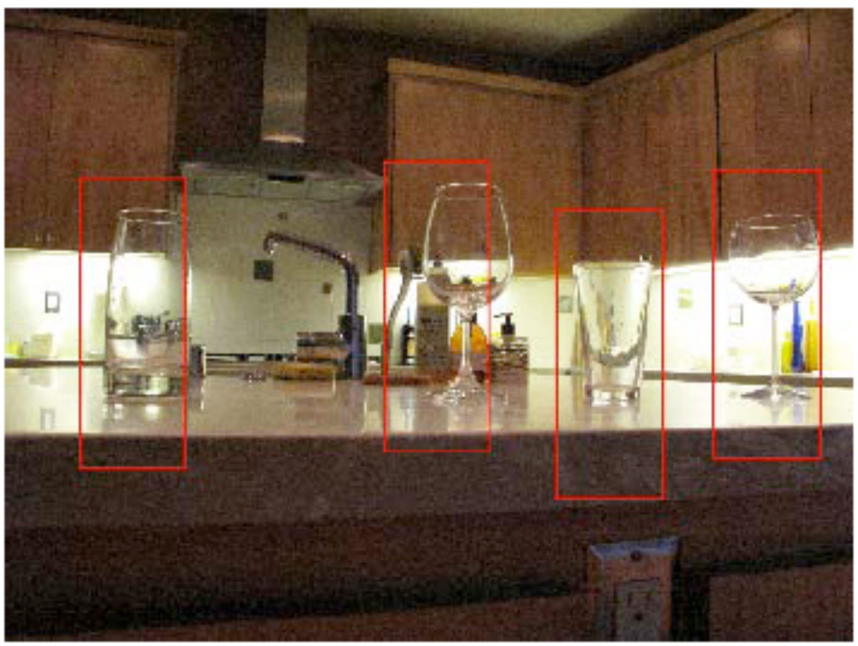


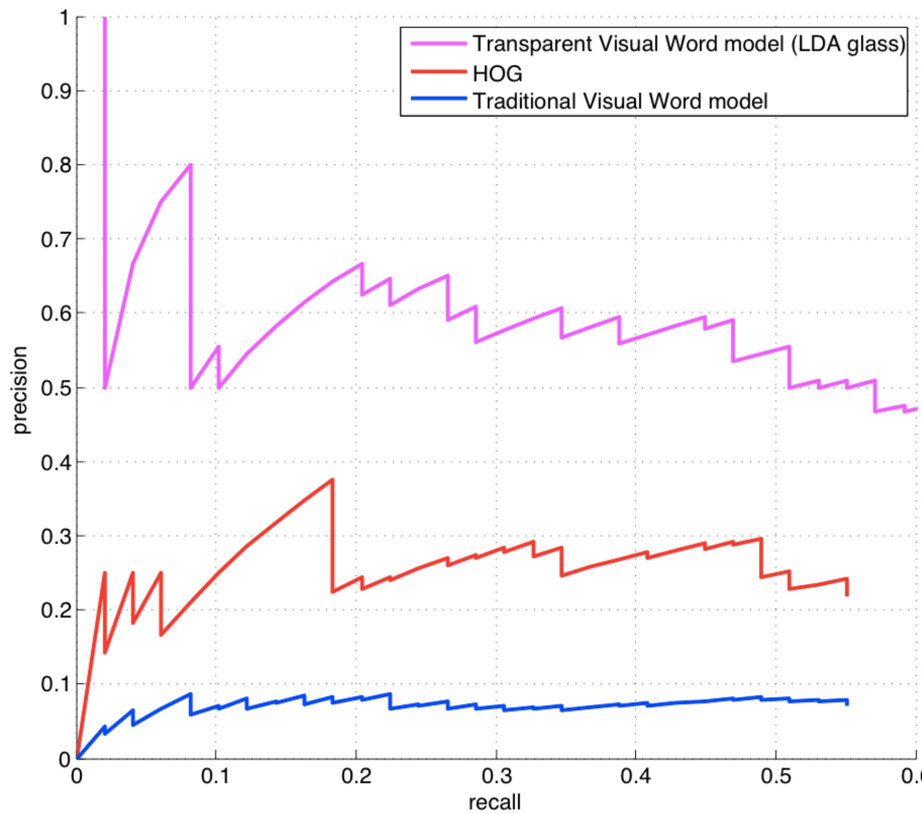
Experiments

Evaluation Data

Results vs. baseline

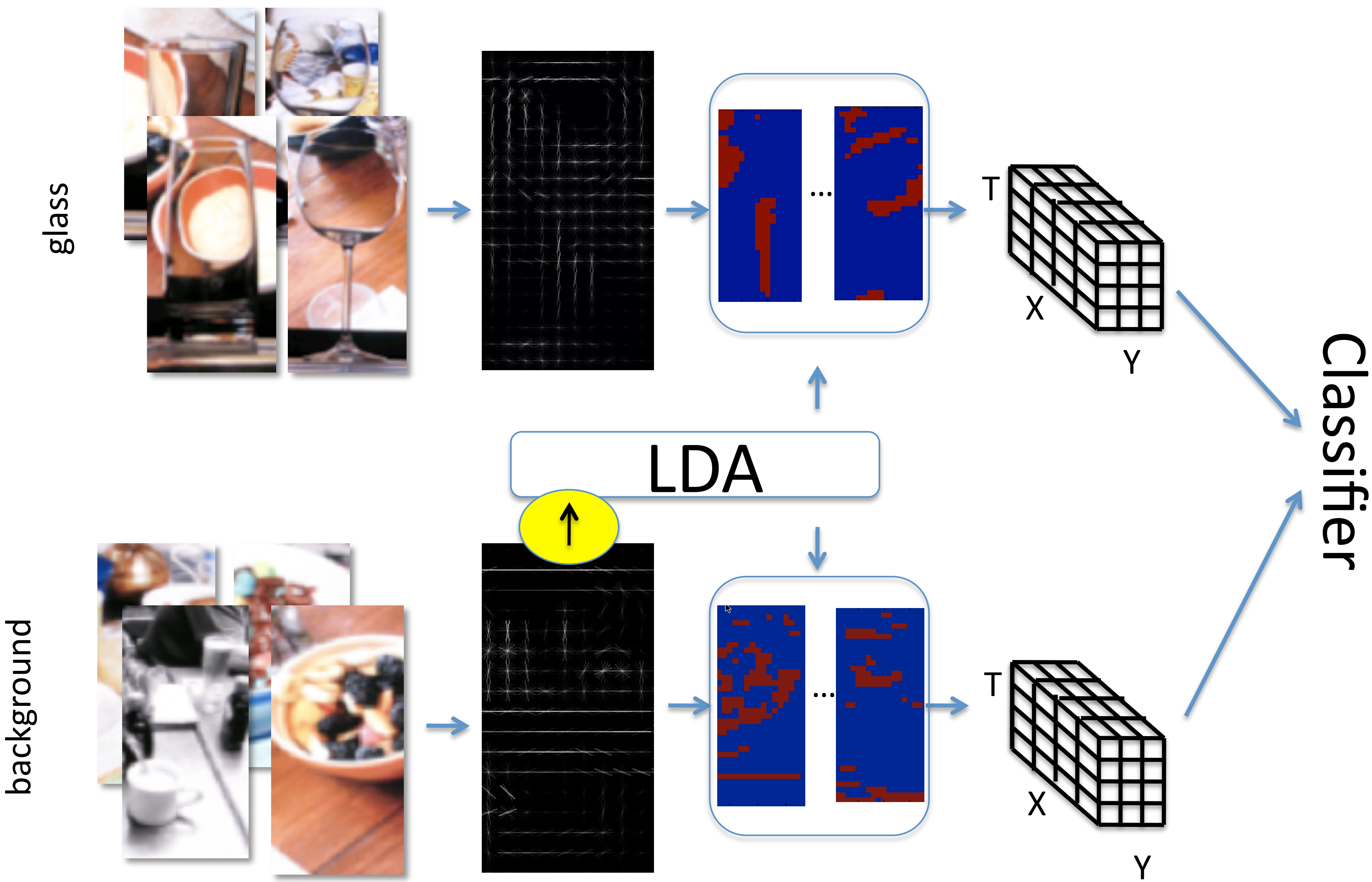




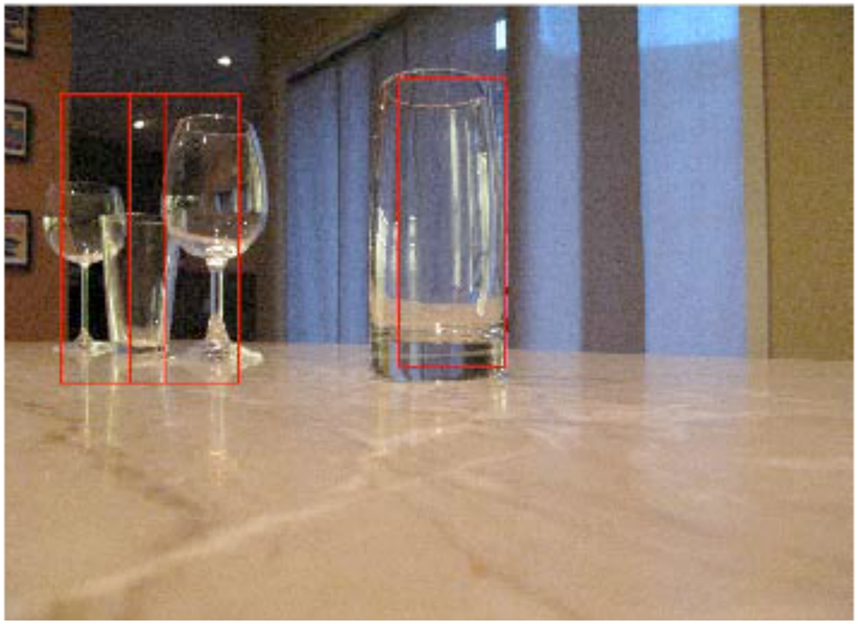


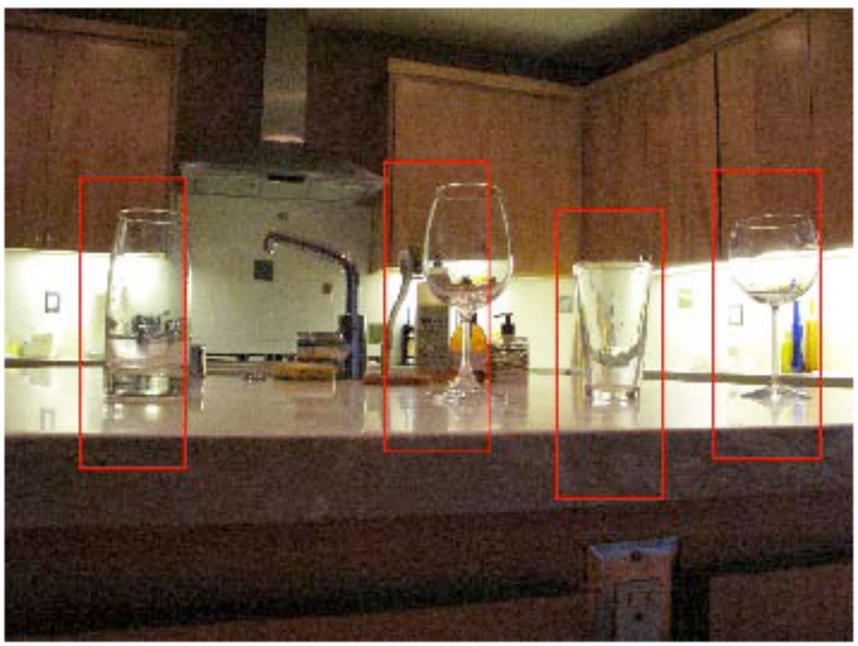
- Training on 4 different glasses in front of screen
- Testing on 49 glass instances in home environment
- Sliding window linear SVM-detection

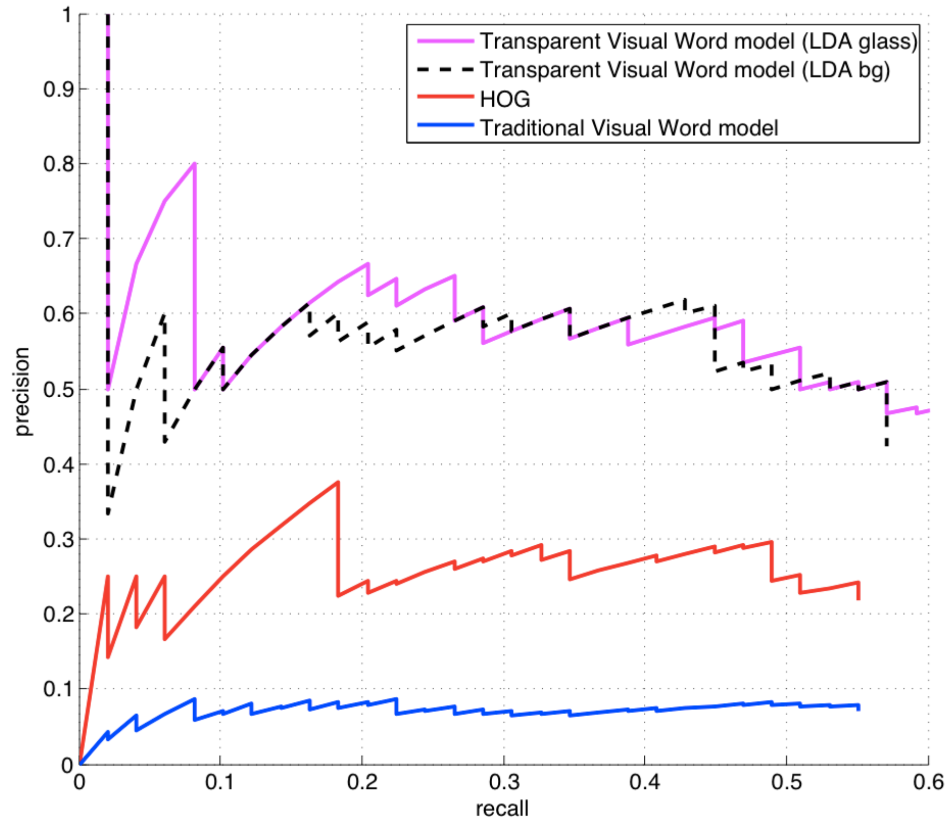
Recognition Architecture



Results: general vocabulary

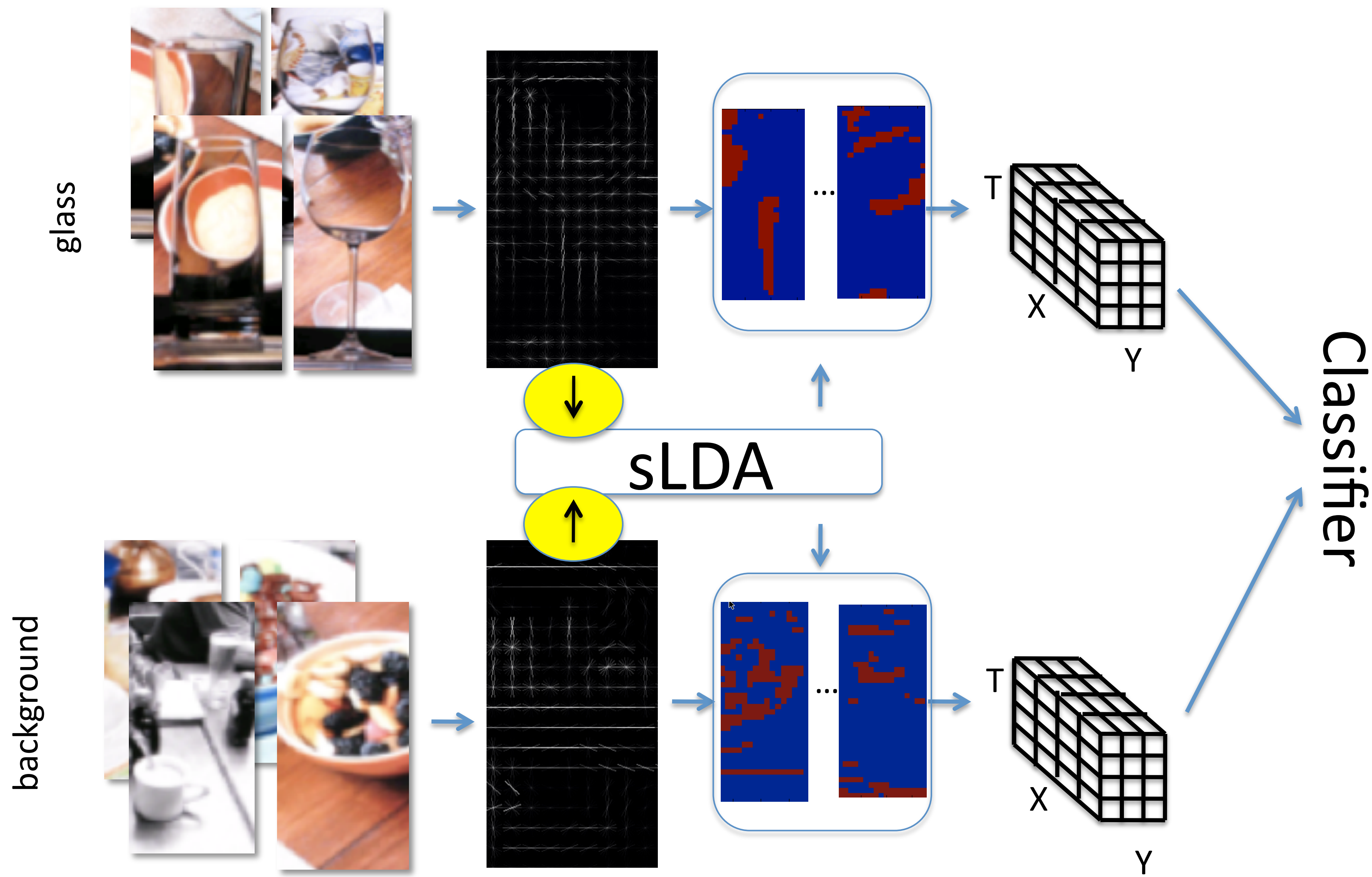




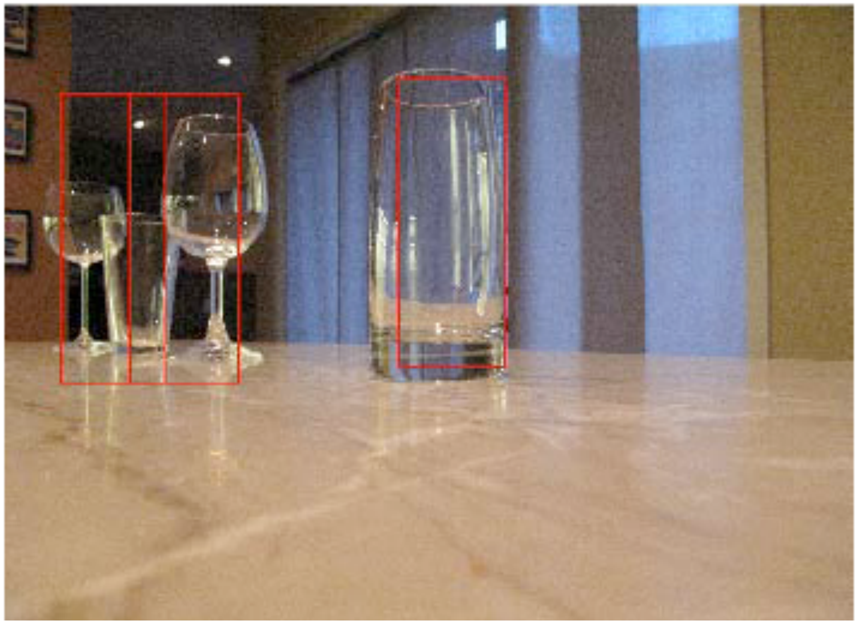


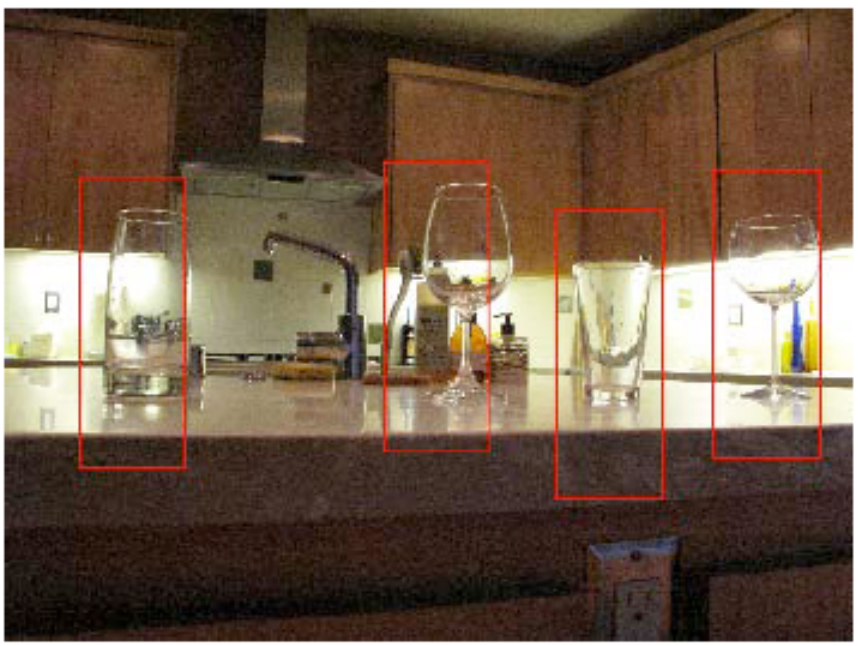
- Training on 4 different glasses in front of screen
- Testing on 49 glass instances in home environment
- Sliding window linear SVM-detection

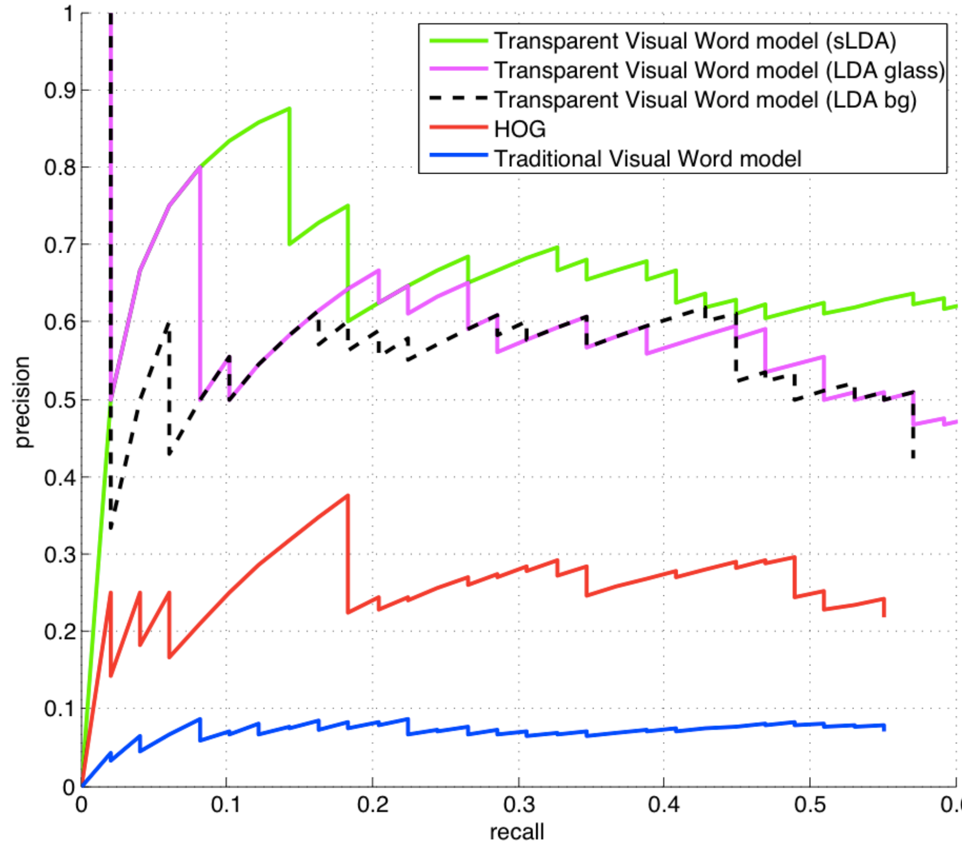
Recognition Architecture



Results: sLDA







- Training on 4 different glasses in front of screen
- Testing on 49 glass instances in home environment
- Sliding window linear SVM-detection

Conclusion

- Traditional local feature models (VQ, NN) are poorly suited for transparent object recognition
- Proposed additive local feature models can detect superimposed image structures
- Developed statistical approach to learn such representations using probabilistic topic models
- Sparse factorization of local gradient statistics
- Encouraging results on real-world data

Future Work

- Different feature representations; extend model in hierarchical fashion
- Investigate addition of material property cues; discriminative inverse local light transport models
- Explore benefits for opaque object recognition; understand relationship to sparse image coding as well as to biological motivated models

Thank you for your attention.