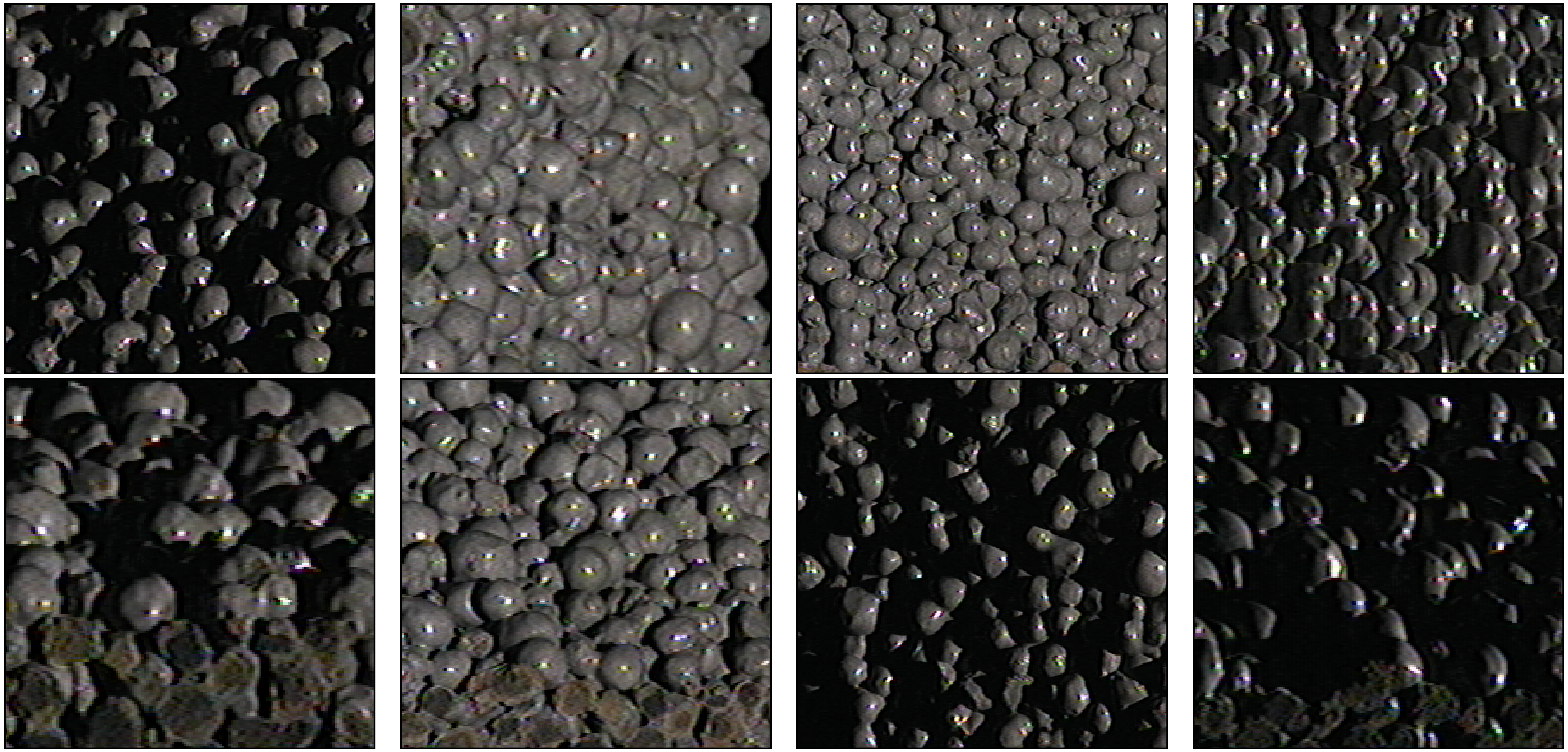
3D Texture Classification Using the Belief Net of a Segmentation Tree

Sinisa Todorovic and Narendra Ahuja

ICPR 2006

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

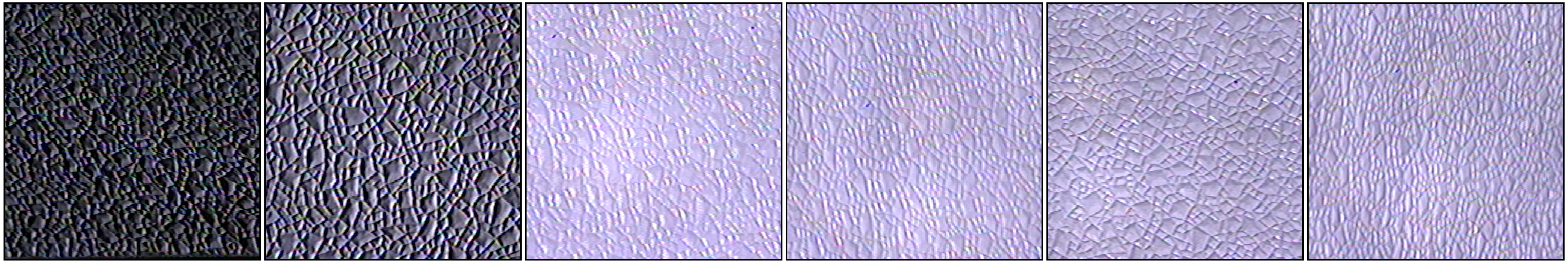
PROBLEM: 3D TEXTURE CLASSIFICATION



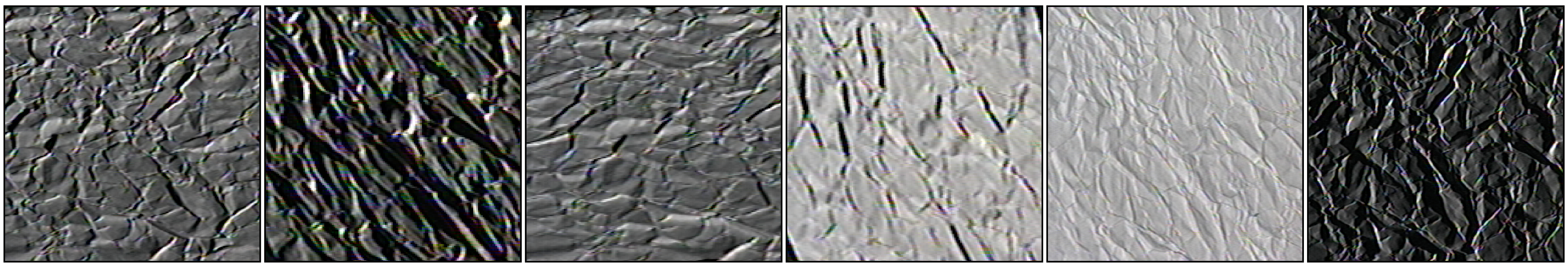
3D texture: variations in local height, color, and reflectance

Different lighting and viewpoint conditions → Different appearances of the same 3D texture

Prior Work: Benchmark CUReT Database



Examples of texture #4 in CUReT



Examples of texture #28 in CUReT

- Dana et. al, CVPR '97, ACM Trans. Graphics '99
- Dana & Nayar, ICCV '99
  - features = autocorrelation coefficients
  - model captures spatial relations among features for varying lighting conditions and viewpoints

Prior Work: Filter-Based 3D Texture Modeling

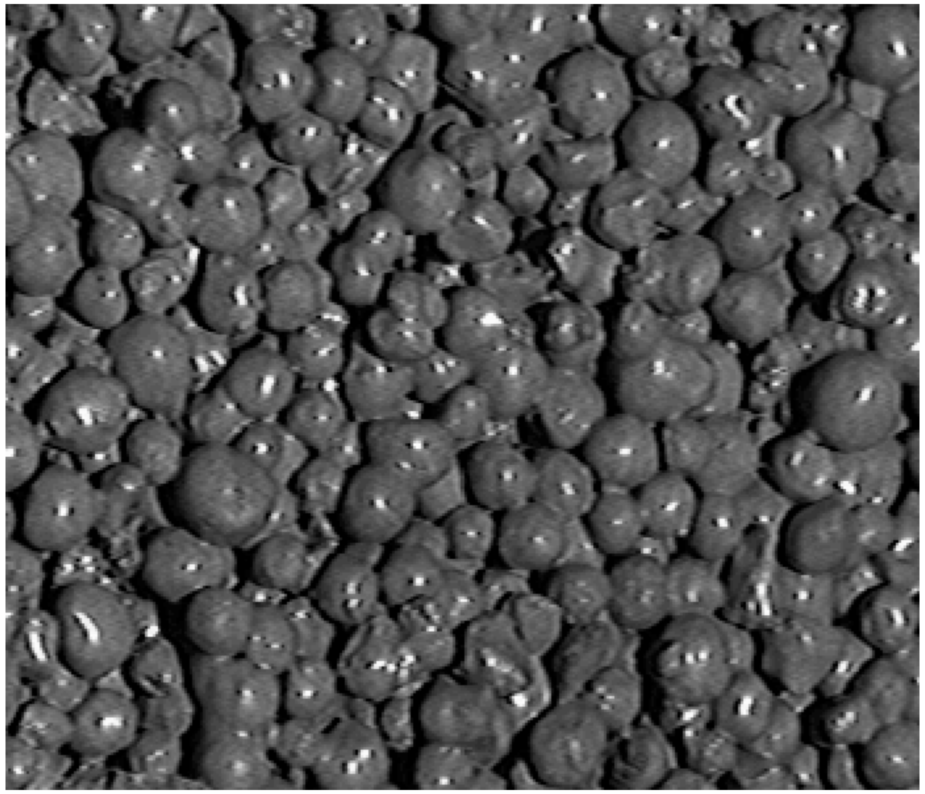
- Texture model = histograms of extracted features
- Modeling redundancy: multiple models per single texture class for varying illumination conditions and viewpoints

- Leung & Malik, IJCV '01: 3D textons in the filter space
- Cula & Dana, IJCV '04, and Varma & Zisserman, IJCV '05
    - images filtered by a bank of filters
    - universal vocabulary of filter-response textons
    - multiple frequency histograms of textons appearing in a class

Our Approach: Region-Based Texture Modeling

Find texture model that captures:

- **TEXTURE PRIMITIVES**
- **REGION properties of texture primitives for varying ILLUMINATION and VIEWPOINTS**
  - **Geometric (area, boundary shape)**
  - **Photometric (gray-level intensity)**
  - **Topology**
    - **Recursive containment of regions**
    - **Layout - relative region locations**



Outline of Our Approach

GIVEN

Training images, each with different known illumination and viewpoint parameters

SEGMENT

Images at all photometric scales present

GENERATE SEGMENTATION TREE

Nodes in the tree are samples of texture primitives

LEARN FOR EACH TEXTURE CLASS

Models of texture primitives on the given training set, parameterized by illumination and viewpoints

GIVEN

A new texture image obtained under unknown illumination and viewpoint conditions

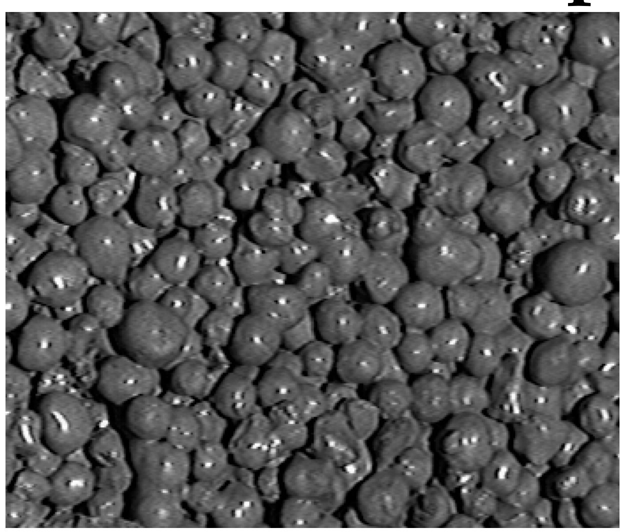
CLASSIFY

Multiscale Image Segmentation

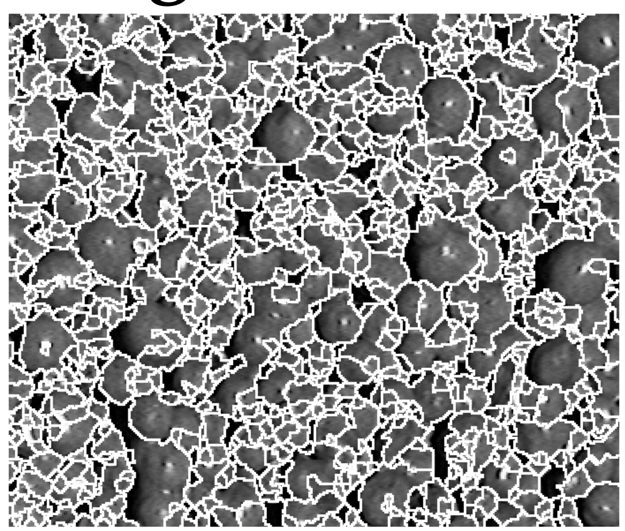
Image $\xrightarrow{\text{segmentation}}$ Homogeneous regions at ALL contrasts and sizes

[N. Ahuja TPAMI '96, Tobb & Ahuja TIP '97, Arora & Ahuja ICPR '06]

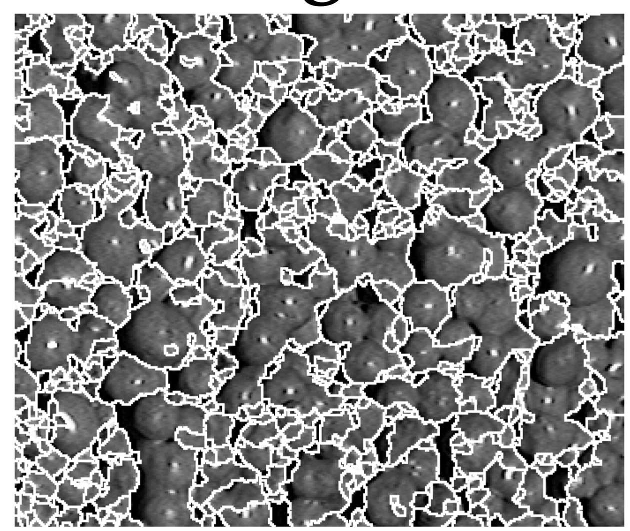
Example segmentations for a range of contrasts σ



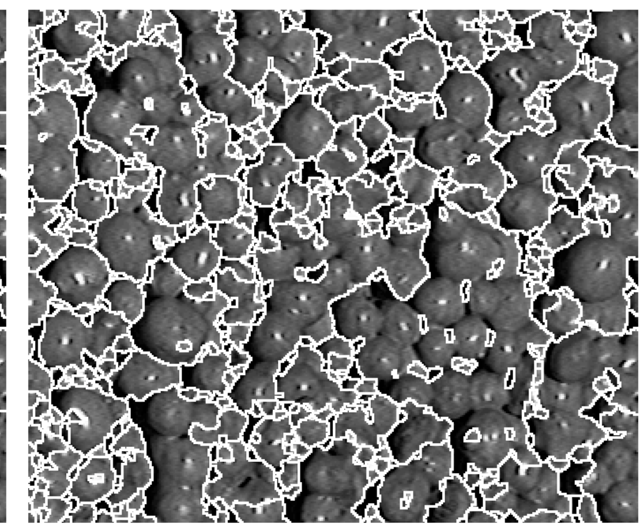
original image



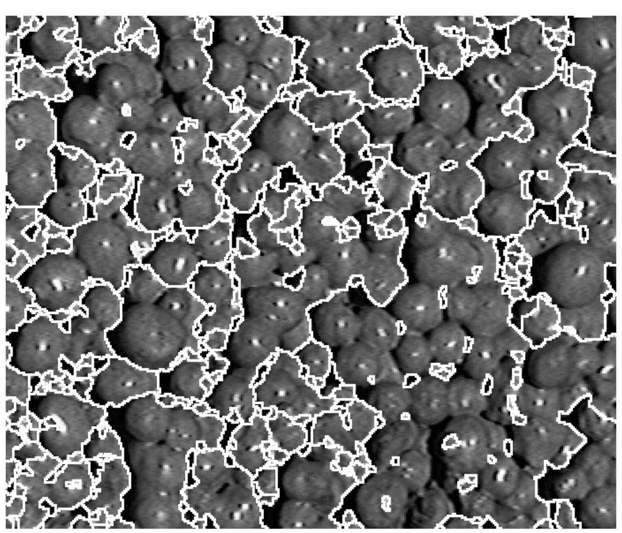
$\sigma = 10$



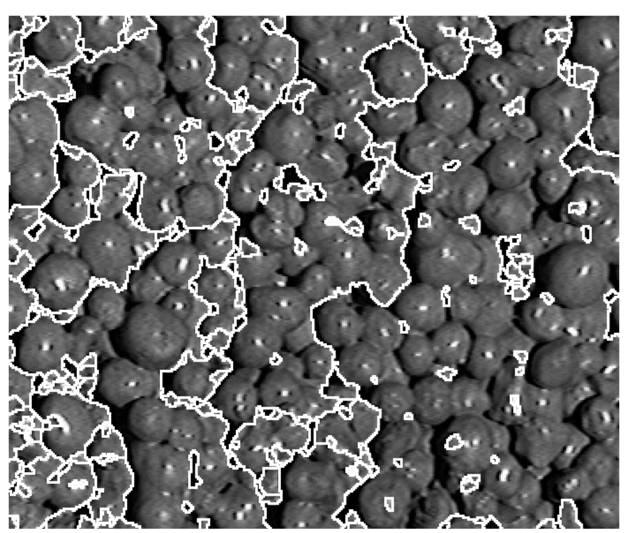
$\sigma = 12$



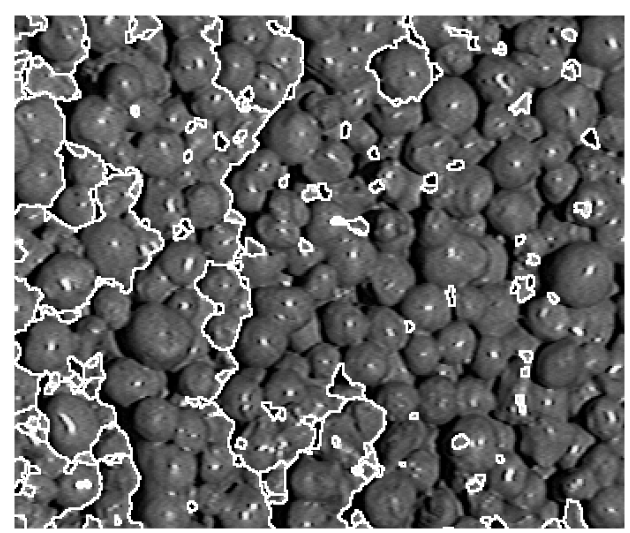
$\sigma = 14$



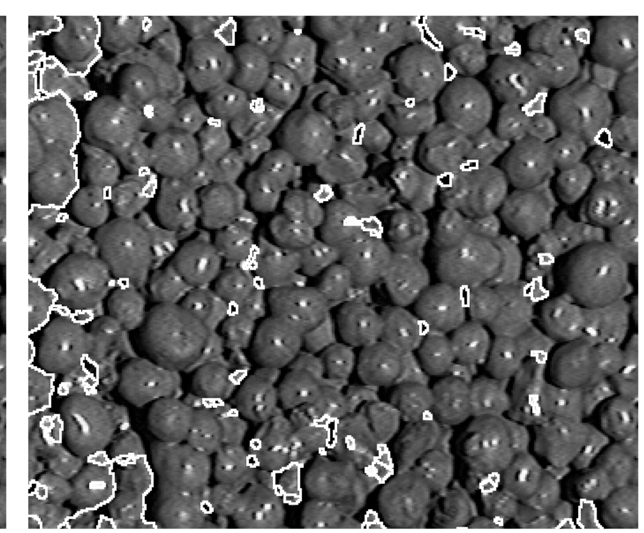
$\sigma = 16$



$\sigma = 18$



$\sigma = 20$



$\sigma = 22$

Multiscale Segmentation to Segmentation Tree

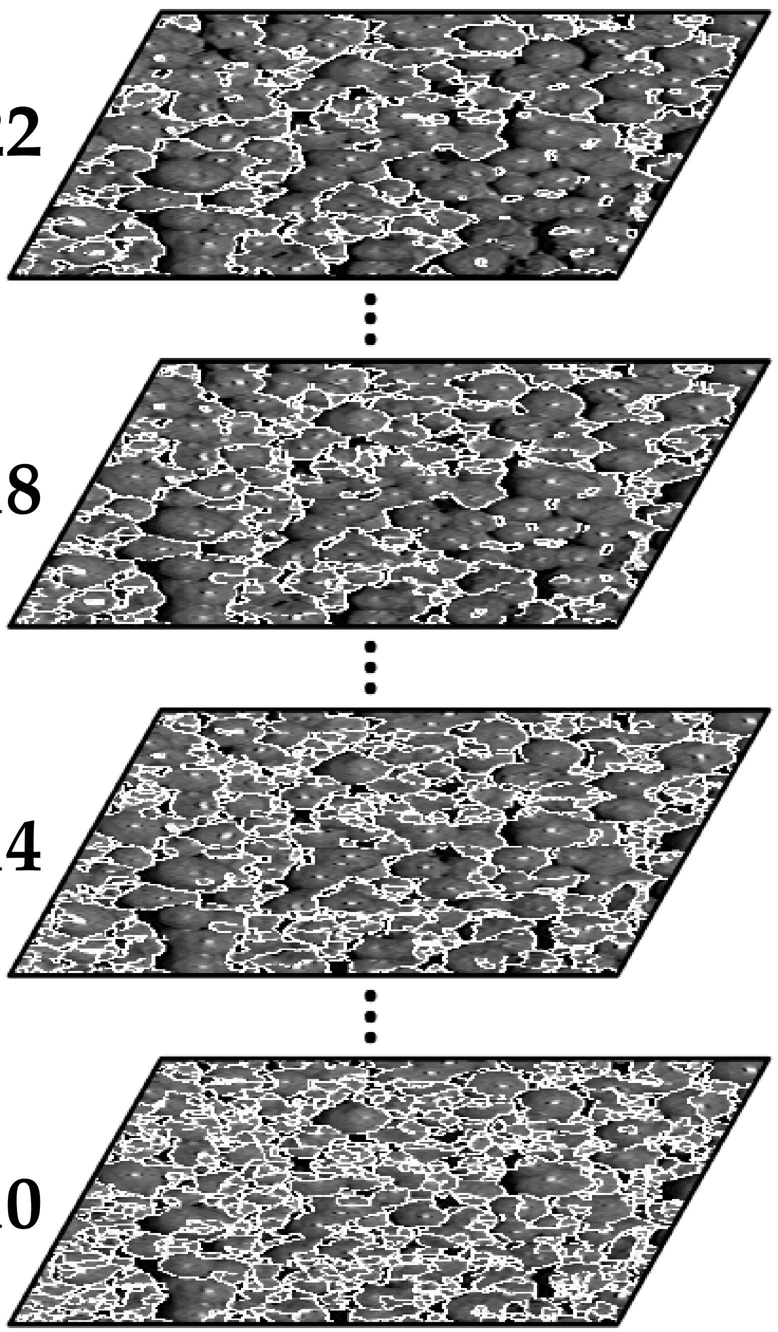
Example segmentations for a few selected contrasts σ

$\sigma = 22$

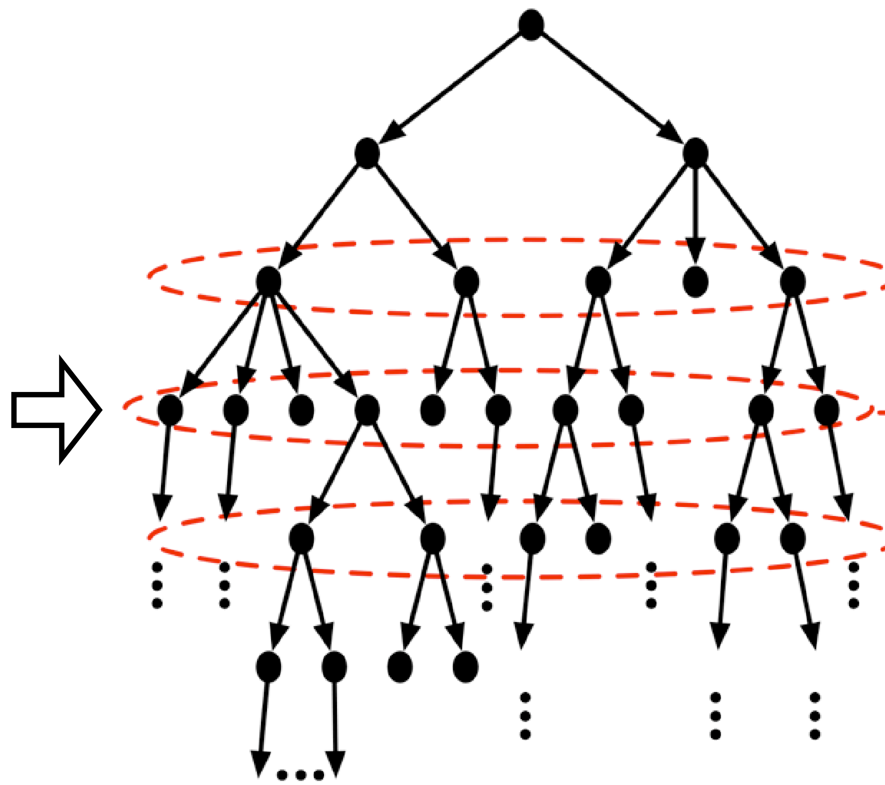
$\sigma = 18$

$\sigma = 14$

$\sigma = 10$

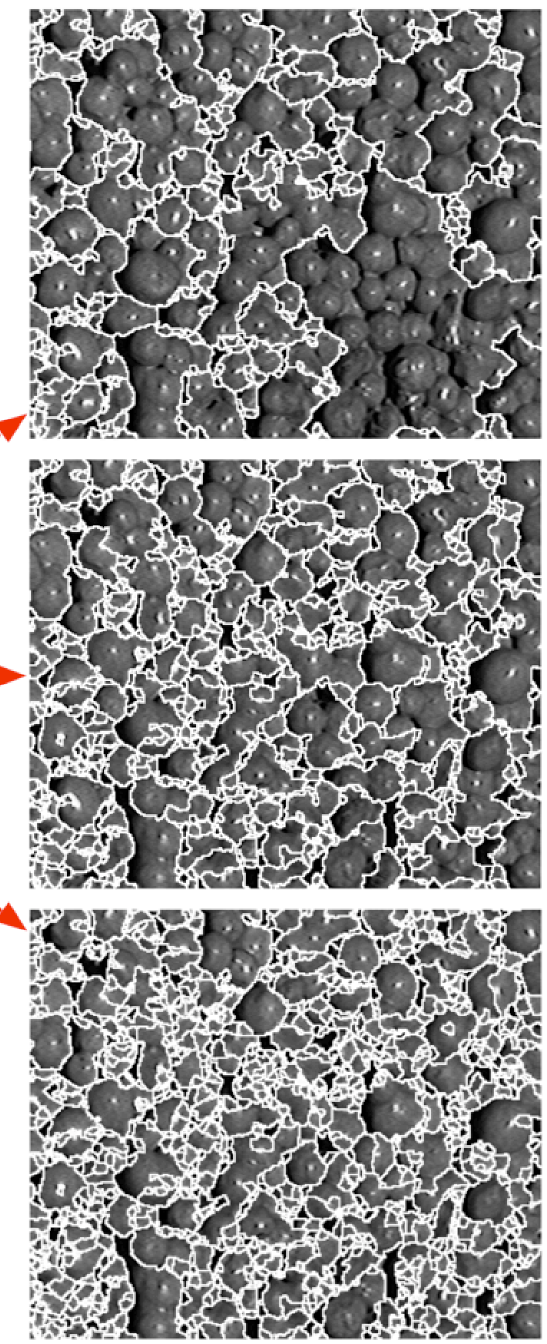


Segmentation tree



Tree level ≠ Contrast level σ

Example cutsets of segmentation tree

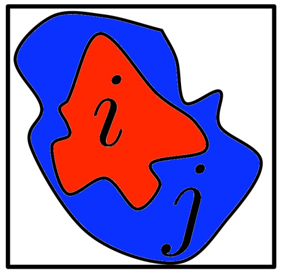


Node = Segmented Region

- Each node in the segmentation tree is associated with the properties of the corresponding segmented region:
    - Intrinsic and relative photometric properties
    - Intrinsic and relative geometric properties
- The properties of a given region are expressed relative to that region's parent properties
- Texture model:
    - invariant to texture orientation in the image plane
    - invariant to small changes in resolution (scale) at which texture is imaged

Region Properties Relative to Parent Region

- **Relative gray-level contrast**



$$\mu_i = \text{mean intensity of region } i$$

$$\sigma_i = \frac{|\mu_i - \mu_j|}{\max\{\mu_i, \mu_j\}}$$

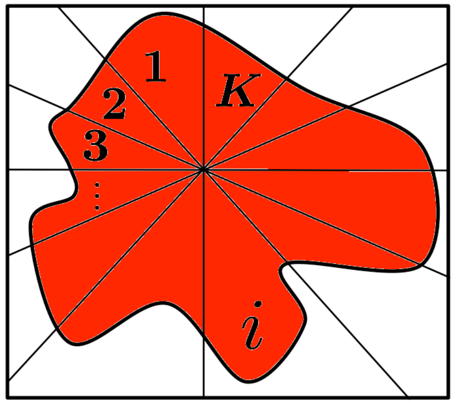
- **Relative gray-level variance**

$$\Sigma_i = \frac{\text{variance}_i}{\text{variance}_j}$$

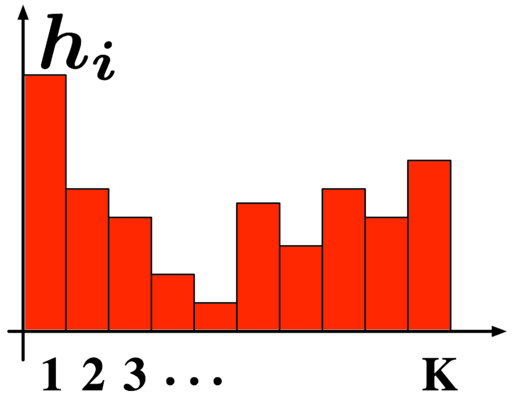
- **Relative area**

$$A_i = \frac{\text{area}_i}{\text{area}_j}$$

- **Rotation invariant shape context**



$\rightarrow$



$$\boldsymbol{y}_i = \begin{bmatrix} \sigma_i \\ \Sigma_i \\ A_i \\ h_i(1) \\ \vdots \\ h_i(K) \end{bmatrix}$$

observable attribute vector

Outline of Our Approach

GIVEN

Training images, each with different known illumination and viewpoint parameters

SEGMENT

Images at all photometric scales present

GENERATE SEGMENTATION TREE

Nodes in the tree are samples of texture primitives

LEARN FOR EACH TEXTURE CLASS

Models of texture primitives on the given training set, parameterized by illumination and viewpoints

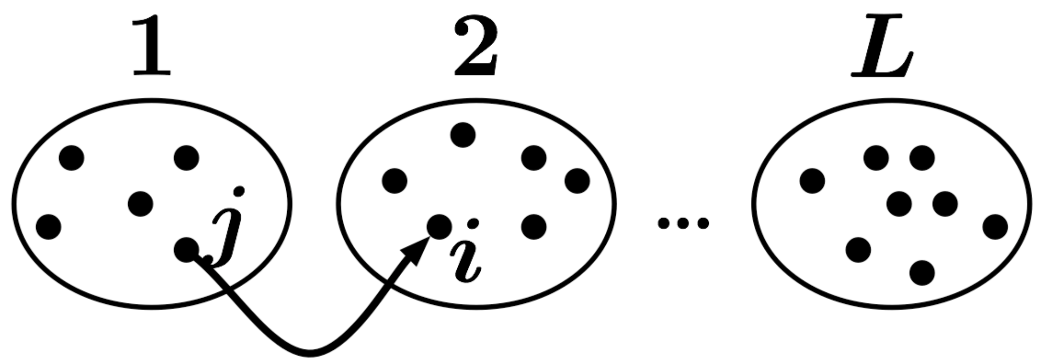
GIVEN

A new texture image obtained under unknown illumination and viewpoint conditions

CLASSIFY

Modeling Vocabulary of Texture Primitives

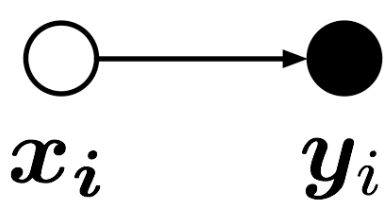
Node = Sample from possible L texture primitives



$$P(x_i\!=\!2|x_j\!=\!1) \longleftarrow$$ **Transition probability table**

Node label = Texture primitive label

Texture primitive generates observable region attributes



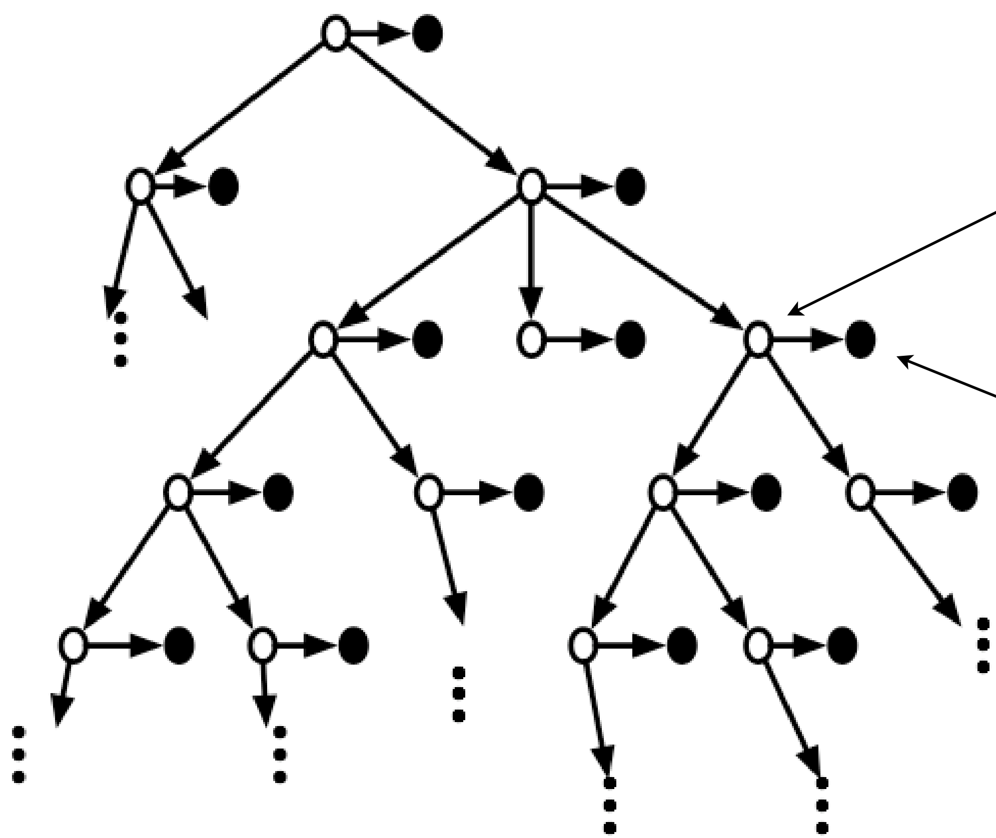
$$P(y_i|x_i) \longleftarrow$$ **Likelihood**

observable random variable → y_i

hidden random variable → x_i

Modeling Image Segmentation Tree

Belief net has the structure of segmentation tree



hidden variable = label of texture primitive

observable = regions from segmentation tree

$\mathcal{V}$ = **Viewpoint**

$\mathcal{I}$ = **Illumination**

$\mathcal{T}$ = **Texture class**

$$P(X, Y | \mathcal{T}, \mathcal{V}, \mathcal{I}) = \prod_{ij} P(y_i | x_i) P(x_i | x_j)$$

Joint probability density function of the belief net

region attributes

label of texture primitive

Training per Class

GIVEN

Images with N different viewpoints and illumination

LEARN

N models parameterized by viewpoint and illumination using the standard belief propagation on trees

$$P(X,Y|\mathcal{T},\mathcal{V}_1,\mathcal{I}_1)$$

$$P(X,Y|\mathcal{T},\mathcal{V}_2,\mathcal{I}_2)$$

$$\vdots$$

$$P(X,Y|\mathcal{T},\mathcal{V}_N,\mathcal{I}_N)$$

$\mathcal{V}$ = **Viewpoint**

$\mathcal{I}$ = **Illumination**

$\mathcal{T}$ = **Texture class**

Reducing the Number of Models per Class

N models

$$P(X,Y|\mathcal{T},\mathcal{V}_1,\mathcal{I}_1)$$

$$\boxed{P(X,Y|\mathcal{T},\mathcal{V}_2,\mathcal{I}_2)}$$

$$P(X,Y|\mathcal{T},\mathcal{V}_3,\mathcal{I}_3)$$

$$P(X,Y|\mathcal{T},\mathcal{V}_4,\mathcal{I}_4)$$

$$\boxed{P(X,Y|\mathcal{T},\mathcal{V}_5,\mathcal{I}_5)}$$

$$\vdots$$

$$P(X,Y|\mathcal{T},\mathcal{V}_N,\mathcal{I}_N)$$

K-Medoid Clustering of Models

M models, M << N

$$P(X,Y|\mathcal{T},\mathcal{V}_2,\mathcal{I}_2)$$

$$P(X,Y|\mathcal{T},\mathcal{V}_5,\mathcal{I}_5)$$

Classification

GIVEN

Image captured from <u>unknown viewpoint</u>, and under <u>unknown lighting</u> conditions

SEGMENT

Images at all photometric scales present

GENERATE SEGMENTATION TREE

Nodes in the tree are samples of texture primitives

COMPUTE

Using the standard belief propagation on trees

$$\hat{\mathcal{T}} = \arg \max_{\mathcal{T},\mathcal{V},\mathcal{I}} P(X, Y | \mathcal{T}, \mathcal{V}, \mathcal{I})$$

M models per class

$\mathcal{V}$ = **Viewpoint**

$\mathcal{I}$ = **Illumination**

$\mathcal{T}$ = **Texture class**

Experimental Results

Dataset = 20 texture classes in CUReT

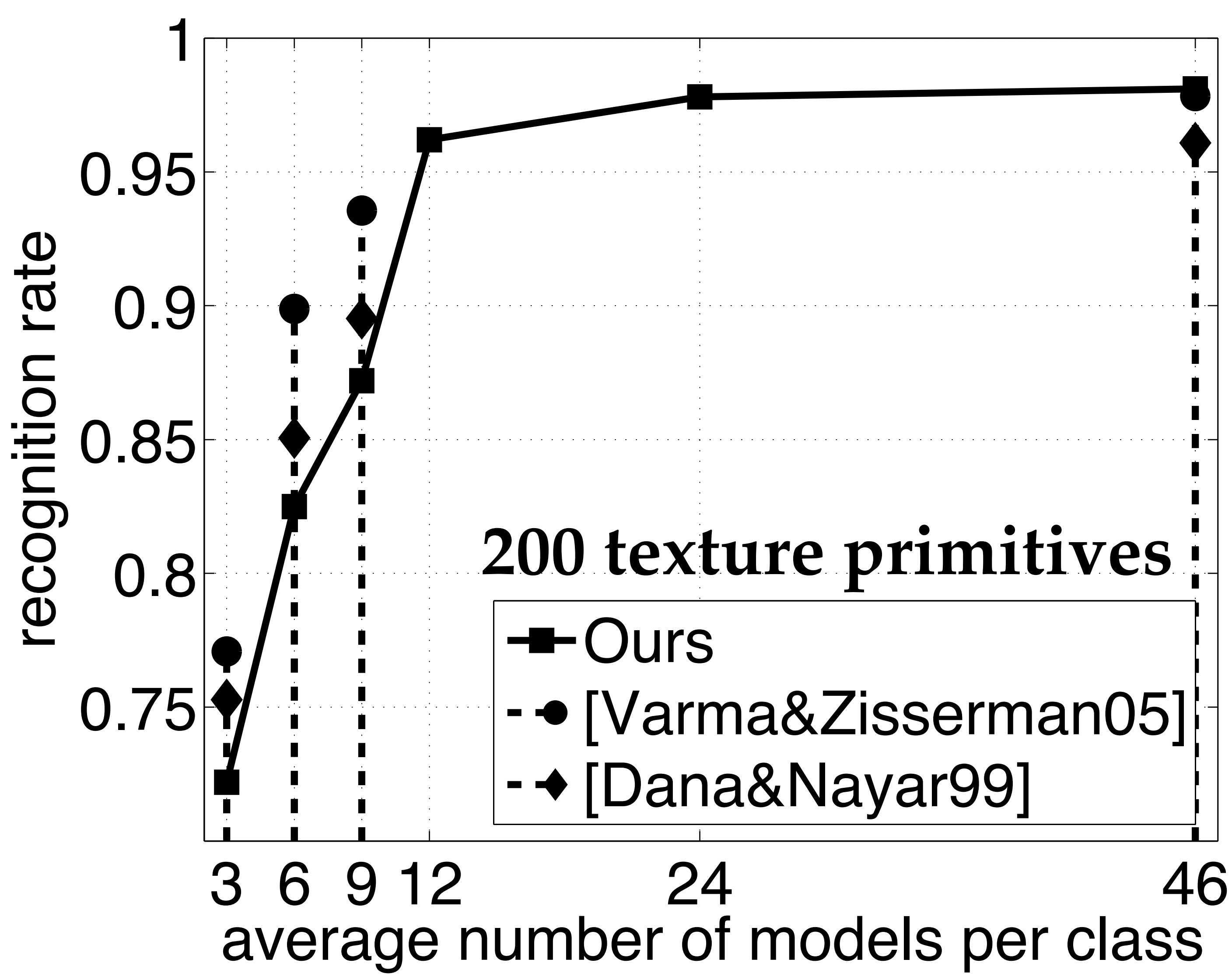
- **Performance evaluated by varying:**
  - **number of texture primitives**
  - **number of models per texture class (redundancy)**
  - **number of training images per class**

Recognition Rate vs. Modeling Redundancy

92 images per class each with different illumination and viewpoint parameters

randomly selected 46 training images

randomly selected 46 test images

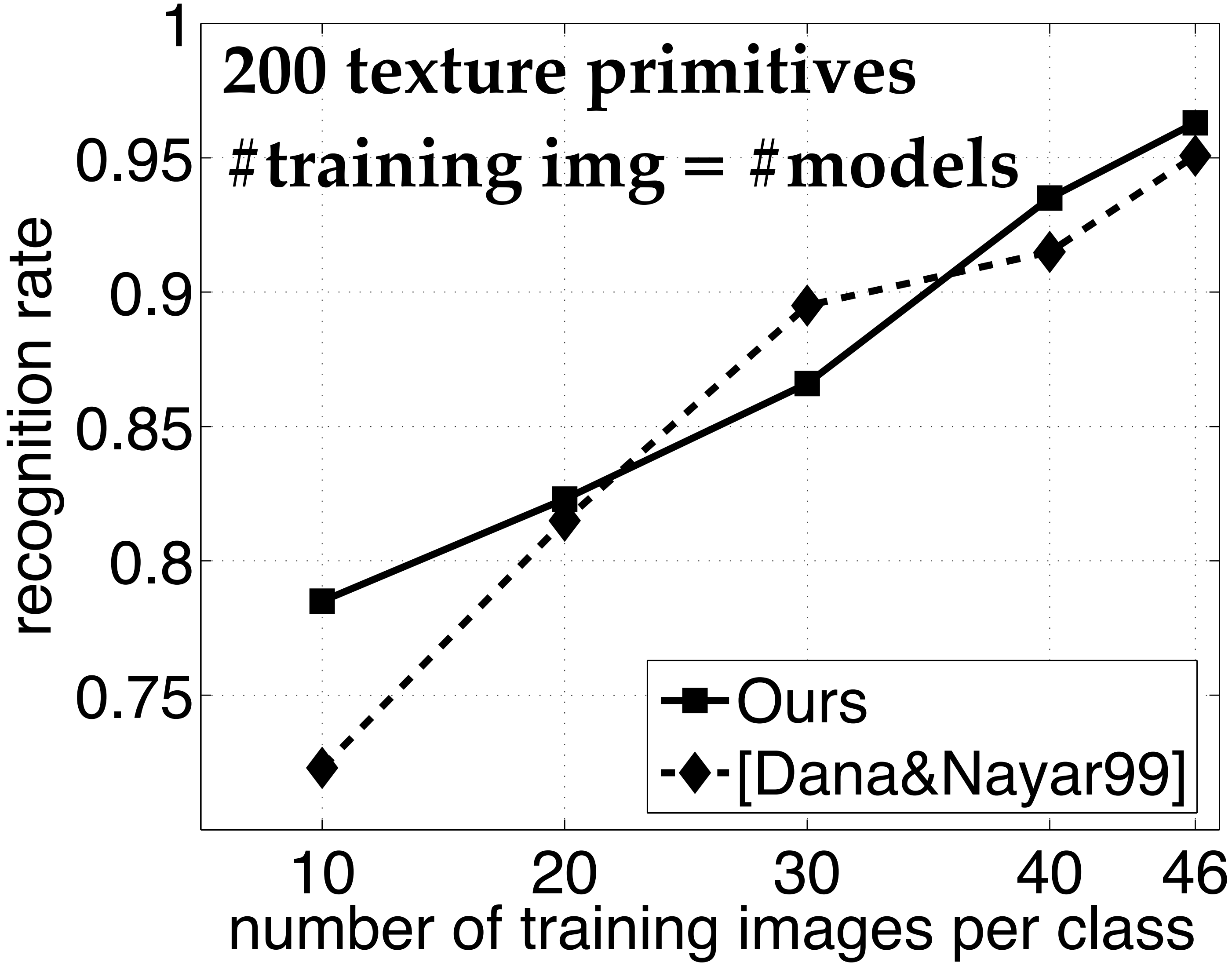


Recognition Rate vs. Size of Training Set

92 images per class each with different illumination and viewpoint parameters

randomly selected # training images

randomly selected 46 test images

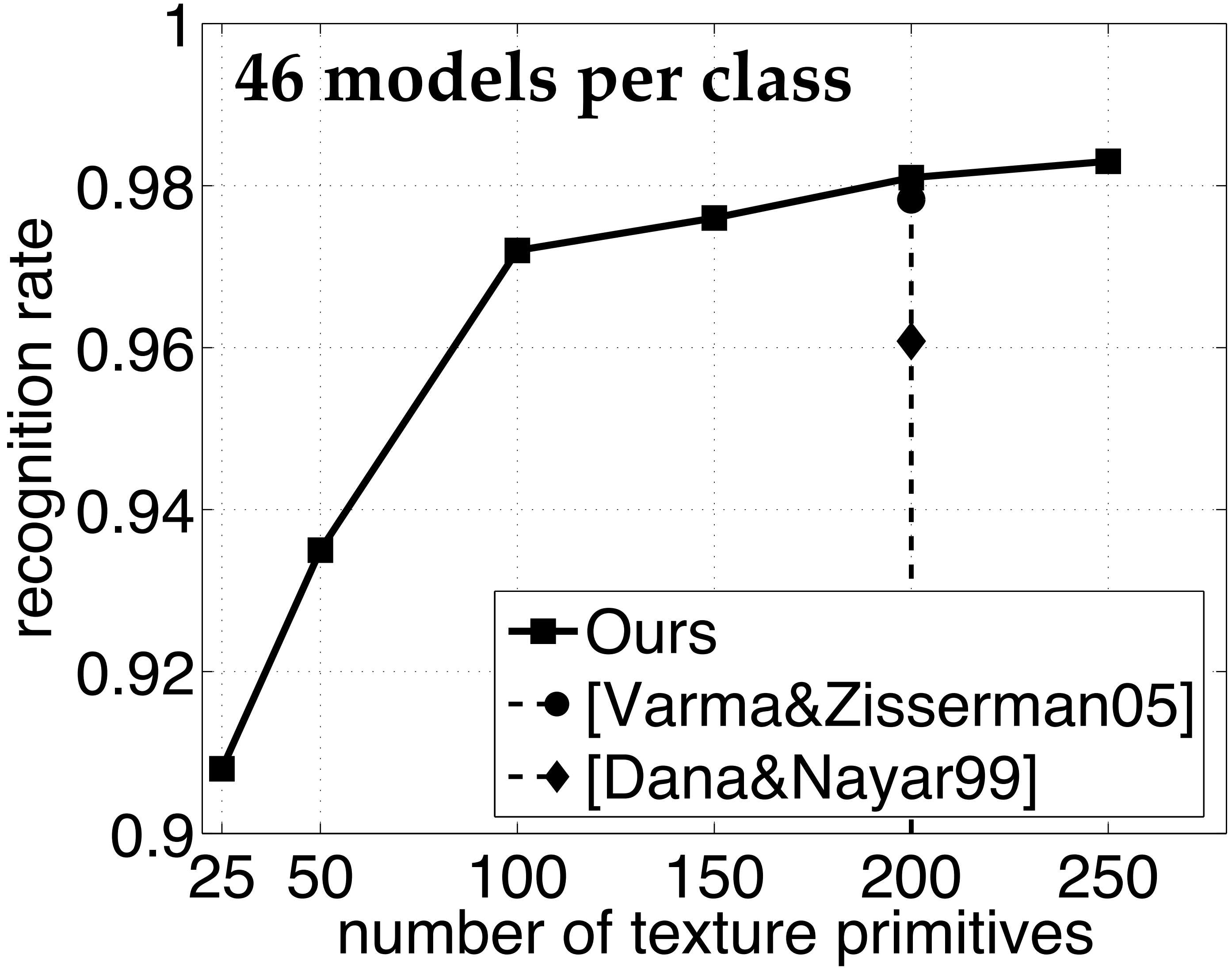


Recognition Rate vs. Number of Texture Primitives

92 images per class each with different illumination and viewpoint parameters

randomly selected 46 training images

randomly selected 46 test images



Summary and Conclusion

- **Region-based approach to 3D texture classification**
- **Feature extraction = Multiscale image segmentation**
- **Belief nets capture successfully photometric, geometric and topological properties of texture primitives for a few selected illumination and viewpoint values**
- **Belief nets have a great generalization capability to classify images with unknown illumination and viewpoint parameters**
- **Recognition rate 98.09% on 20 classes in CUReT outperforms the state of art**

THANK YOU!

{sintod, ahuja}@vision.ai.uiuc.edu

http://vision.ai.uiuc.edu/~sintod