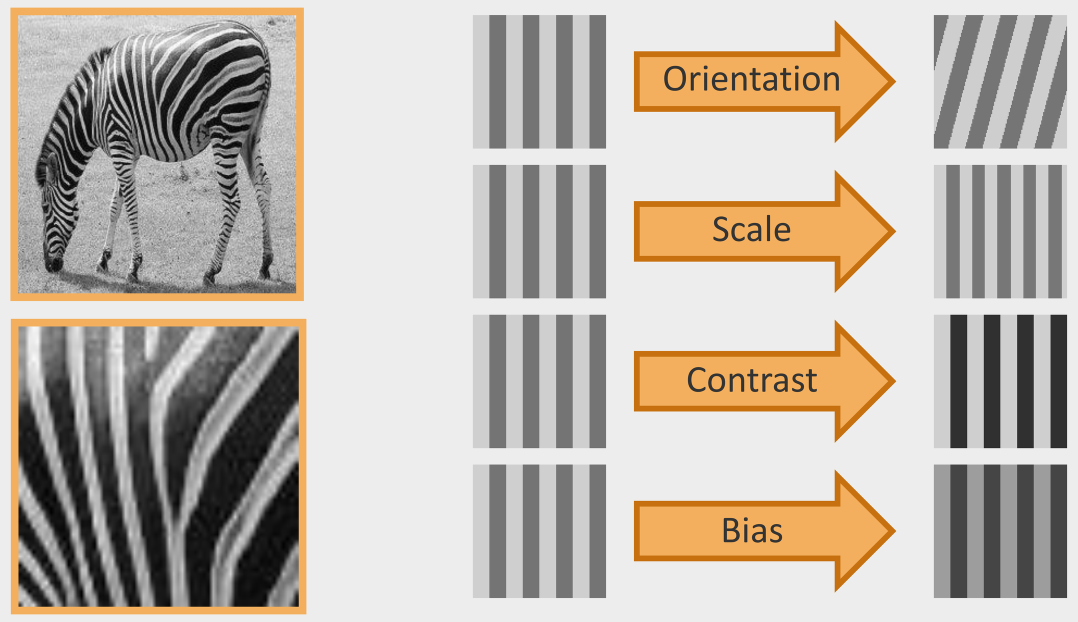
Analysis of Smoothly Varying Textures Applied to Segmentation

Jason Chang & John W. Fisher III

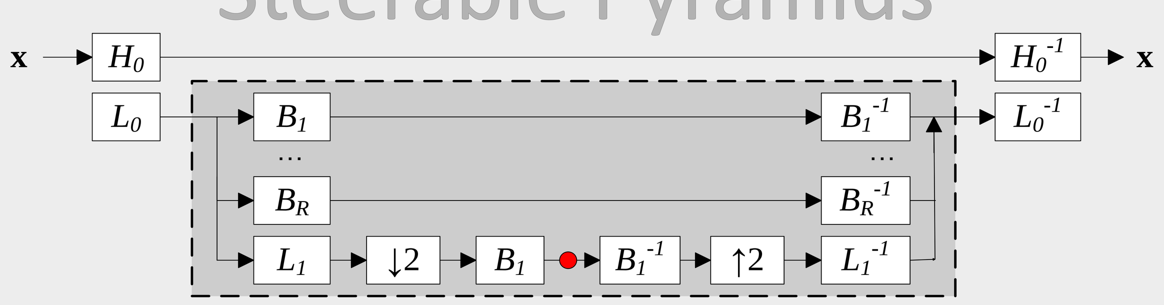
CSAIL
MIT Computer Science and Artificial Intelligence Laboratory

Motivation

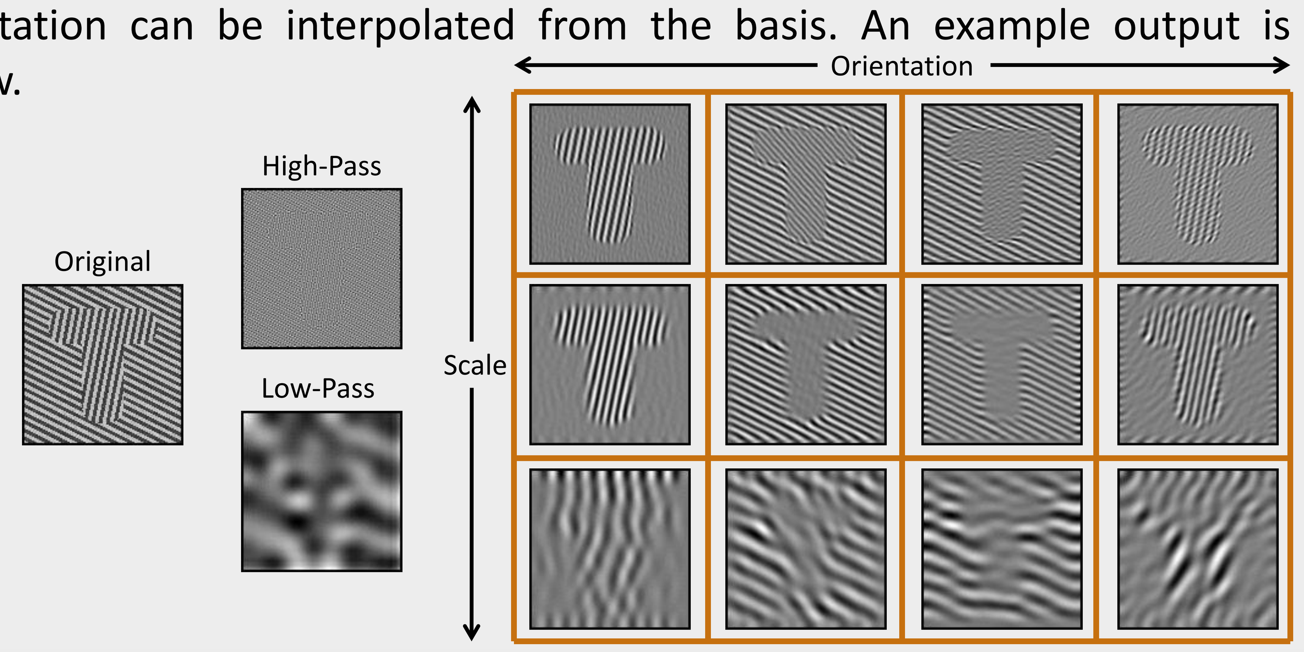


We aim to capture four texture features: orientation, scale, contrast, and bias. In natural textures, each feature can change due to the underlying geometric shape of the object or a physical change in the texture. We measure these features and impose a smooth Markov random field to capture the spatial correlations.

Steerable Pyramids



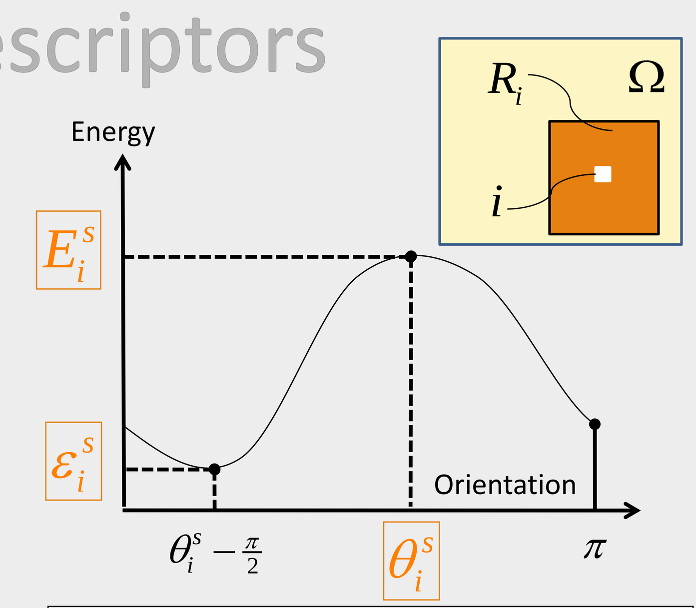
The steerable pyramid [1] is a multi-scale and orientation decomposition of an image. The red dot is recursively replaced with the gray box. The output at any orientation can be interpolated from the basis. An example output is shown below.



Texture Descriptors

For the interpolated filter output, $y_i^{s,\theta}$ at pixel i, scale s, and orientation θ, we define the angular energy as:

$$E_i^s(\theta) = \frac{1}{|R_i|}\sum_{j\in R_i} y_j^{s,\theta_i^s}$$



A plot of the angular energy as a function of orientation is shown on the right. The extracted features are boxed in orange.

We also define the bias of the texture as:

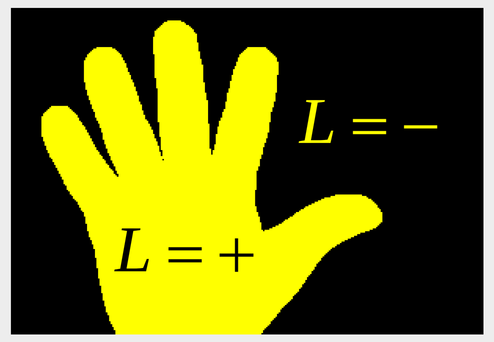
$$\mu_i^s = \frac{1}{|R_i|}\sum_{j\in R_j} x_i$$

Our final feature set is shown in the table to the right.

Feature Set	
$\eta_i = \arg\max_s E_i^s$	Scale
$E_i = E_i^{\eta_i}$	Contrast Energy
$\varepsilon_i = \varepsilon_i^{\eta_i}$	Residual (Orthogonal) Energy
$\theta_i = \theta_i^{\eta_i}$	Orientation
$\mu_i = \mu_i^{\eta_i}$	Average Intensity

Ali-Silvey Distances for Segmentation

We define a binary class label, $L_i \in \{+,-\}$, at each pixel i.



There are two popular information-theoretic approaches to segment an image:

1) Maximize the distance between the conditional distributions $d_C(p_X^+, p_X^-)$
2) Maximize the distance between the joint of the pixels and the labels and the product of the marginals $d_L(p_{XL}, p_X p_L)$

When the distance between the conditional distributions takes on a particular symmetric form, $d_C(p_X^+, p_X^-) = d(p_X^+, p_X^-) + d(p_X^-, p_X^+)$, there is an equivalent Ali-Silvey distance that can be used in Method 2 to reveal the exact same relationship as Method 1.

Distance Measures

$$\begin{cases} d_C(p_X^+, p_X^-) = \mathbb{E}_{p_X^+}\left[C_+\left(\frac{p_X^-(\bullet)}{p_X^+(\bullet)}\right)\right] + \mathbb{E}_{p_X^-}\left[C_-\left(\frac{p_X^+(\bullet)}{p_X^-(\bullet)}\right)\right] \\ d_L(p_{XL}, p_X p_L) = \sum_{\ell\in\mathcal{L}} \pi^\ell \mathbb{E}_{p_X^\ell}\left[\widetilde{C}_\ell\left(\frac{p_X(\bullet)}{p_X^\ell(\bullet)}\right)\right] \end{cases}$$

Relationship

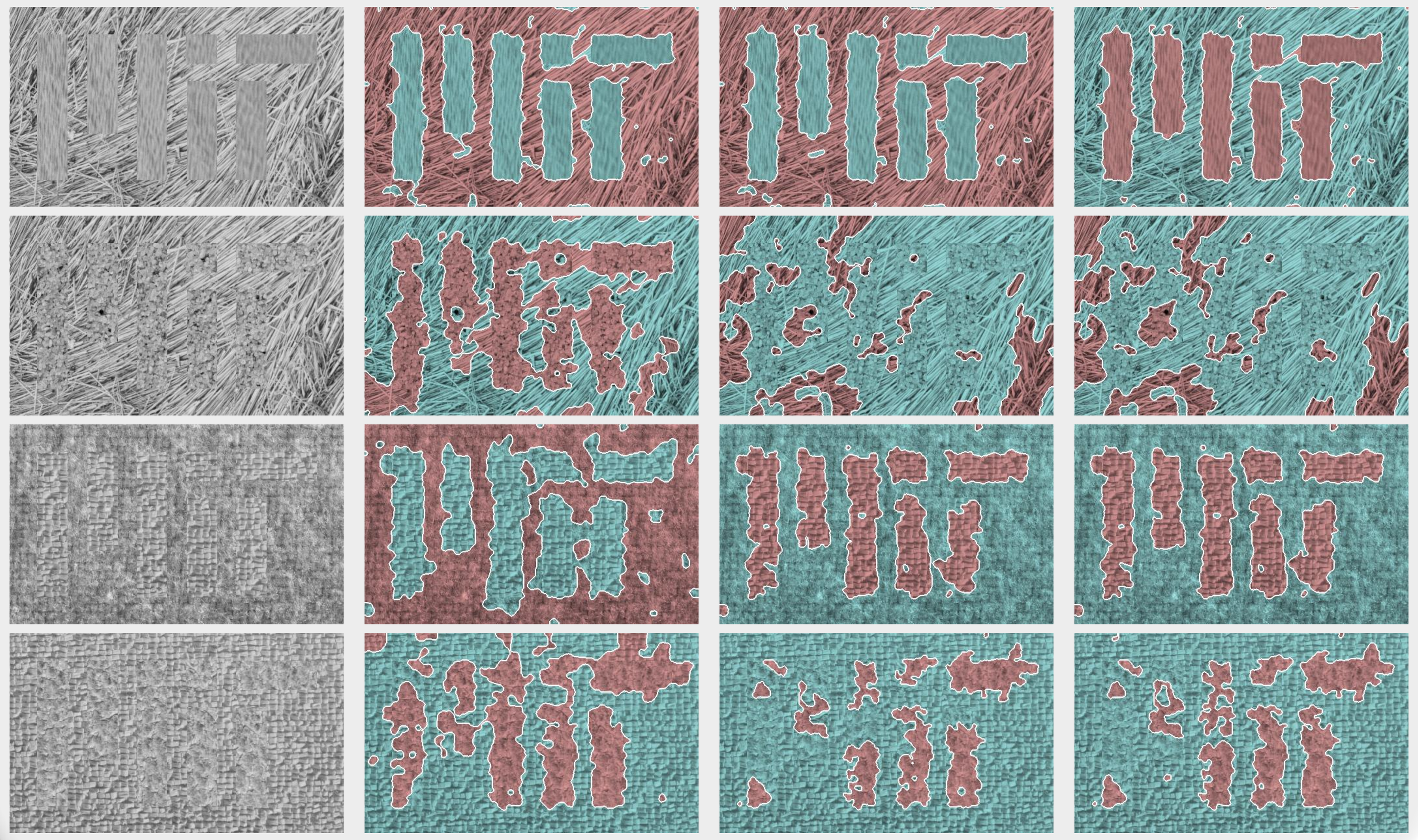
$$\begin{cases} \widetilde{C}_\ell(\bullet) = \frac{1}{\pi^\ell} C_\ell\left(\frac{1}{1-\pi^\ell}\left[\bullet - \pi^\ell\right]\right) \\ C_\ell(\bullet) = \pi^\ell C_\ell\left((1-\pi^\ell)\bullet + \pi^\ell\right) \end{cases}$$

Ali-Silvey Distances Comparison

We compare the following information-theoretic measures for segmentation

Compared Distance Measures		
Mutual Information	$I(X;L)$	$D(p_{XL} \| p_X p_L)$
J Divergence	$J(p_X^+, p_X^-)$	$D(p_X^+ \| p_X^-) + D(p_X^- \| p_X^+)$
Balanced J Divergence	$J_B(p_X^+, p_X^-)$	$\pi^+ D(p_X^+ \| p_X^-) + \pi^- D(p_X^- \| p_X^+)$

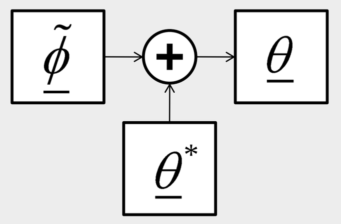
Using our defined features, classification on the Brodatz textures can be computed perfectly under any of the three distance measures. Instead we segment synthetic Brodatz images, with a subset of results shown below.



$I(X;L)$ $J(p_X^+, p_X^-)$ $J_B(p_X^+, p_X^-)$

Smooth Fields

For each feature, we impose a smooth, additive MRF. In the orientation case, this is represented as the diagram to the right where $\tilde{\phi}$ is the smooth field and $\underline{\theta}^*$ is the intrinsic orientation.

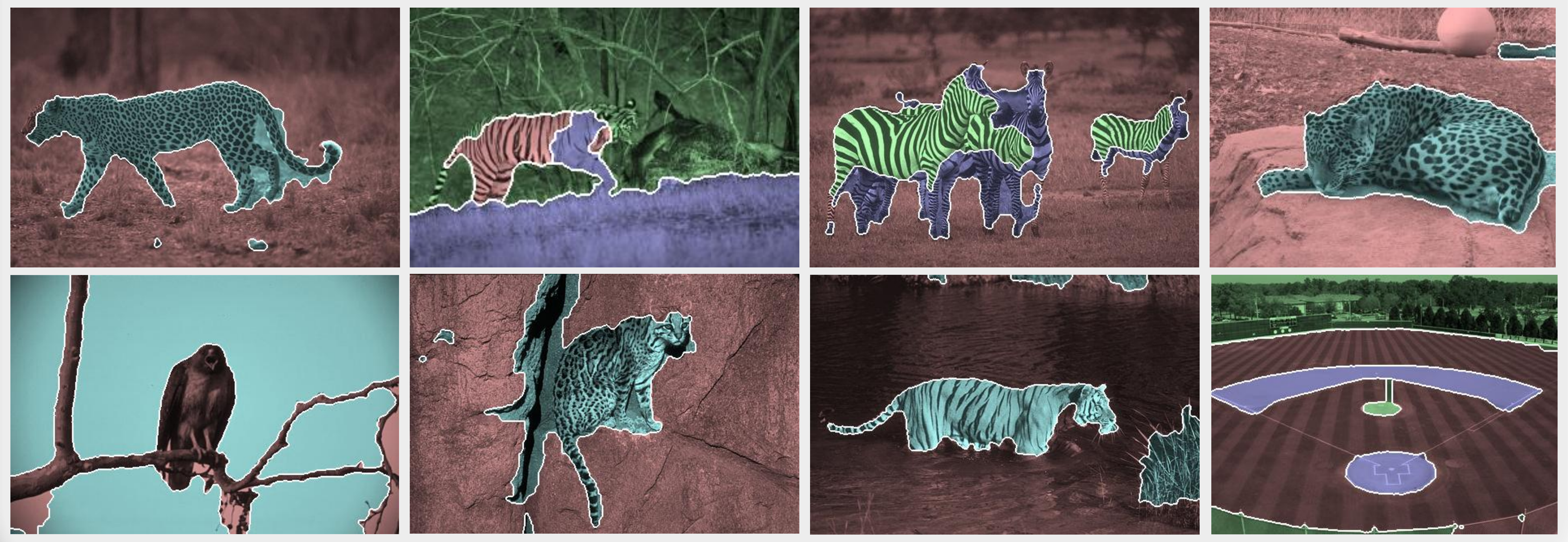


Once these fields are found, we can find the traditional intrinsic image and visualize the intrinsic texture image (equal contrast, bias, orientation, and scale).

Original Image — Intrinsic Image — Intrinsic Texture Image

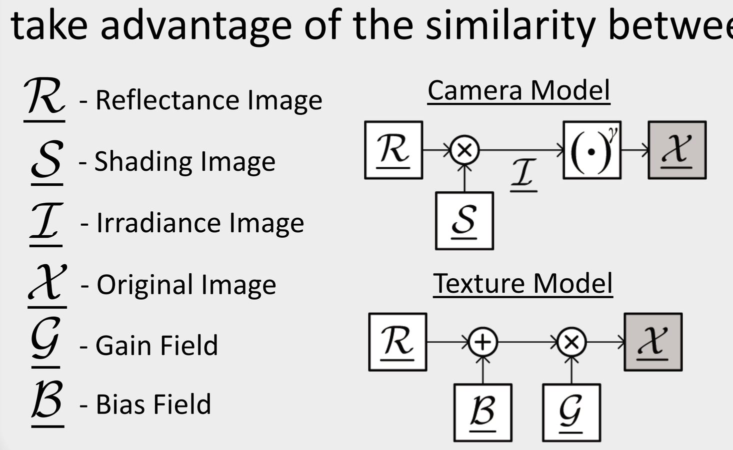
Segmentation

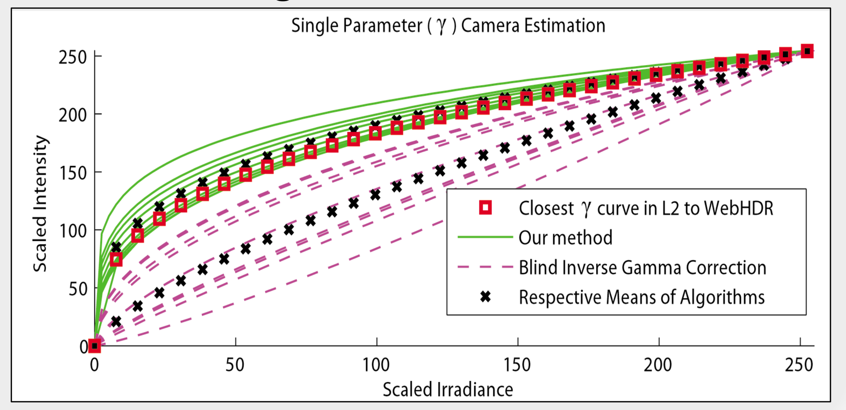
We use our feature set in the segmentation algorithm presented in [2], treating each feature as statistically independent. We incorporate the multi-region segmentation algorithm used in [3].



Nonlinear Camera Estimation

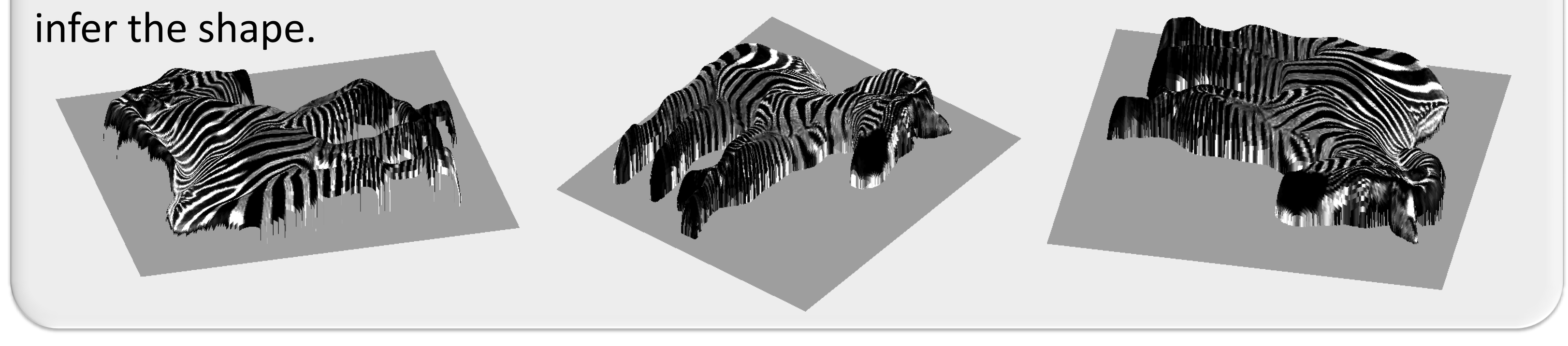
Our model allows us to estimate a one parameter camera radiometric function. We take advantage of the similarity between the following camera and texture model:





Shape from Shading

Once the nonlinear camera model has been estimated with our model, we also estimate a shading image. We use a common shape from shading algorithm [4] to infer the shape.



[1] E. Simoncelli and W. Freeman. The steerable pyramid: a flexible architecture for multi-scale derivative computation. *Image Processing, 1995. Proceedings., International Conference on*, 3:444–447 vol.3, Oct 1995.

[2] J. Kim, I. Fisher, J.W., A. Yezzi, M. Cetin, and A.Willsky. A nonparametric statistical method for image segmentation using information theory and curve evolution. *Image Processing, IEEE Transactions on*, 14(10):1486–1502, Oct. 2005.

[3] T. Brox and J. Weickert. *Level Set Based Image Segmentation with Multiple Regions.* 2004.

[4] P. sing Tsai and M. Shah. Shape from shading using linear approximation. *Image and Vision Computing*, 12:487–498, 1994.