Applications of New Stochastic Image Models to Materials Engineering

Mary Comer, Huixi Zhao, Dae Woo Kim, Shruthi Kubatur, Marc De Graef, and Jeff Simmons

School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN, USA

This research is supported by the Air Force Office of Scientific Research MURI contract # FA9550-12-1-0458.
Stochastic Image Models: The MRF

- The Markov random field (MRF) model is perhaps the most commonly used stochastic image model in the image processing community
  - Form used here is equivalent to Ising or Potts models from statistical mechanics
  - Penalizes interfaces by imposing a cost for different spins at neighboring lattice sites (ferromagnet) or different pixel classifications at neighboring pixels in an image (image segmentation)

NiCrAl image (from Marc De Graef)  Segmentation using MRF model
The marked point process (MPP) image model is a newer model that describes an image at a coarser scale than the MRF

- Prior information at the mesoscale is modeled more easily using an MPP than an MRF
- A superellipse model has been shown to fit the NiCrAl particles well
- Each particle is represented by a random spatial location and random superellipse parameters

New Image Model: The MPP/MRF

NiCrAl image

Mesoscale model (MPP) only

Microscale model (MRF) only

Joint MPP/MRF model
The Marked Point Process: NiCrAl Example

An MPP is a collection of objects represented by their random spatial locations (points) and random values describing the objects (marks).

A superellipse model is used as the object model to describe the geometric shape of a NiCrAl particle\cite{1}

\[ x_0, y_0, a, b, \text{ and } \theta \text{ are all random variables} \]

\[ \text{NiCrAl particle image} \quad \text{superellipse model} \quad \text{marks of the model} \]

The Marked Point Process: The Math

- Let \( w \) be a random field consisting of objects \( w_1, \ldots, w_n \), where each \( w_i \) represents a single object (e.g., particle).
- Each \( w_i \) is a random vector containing the spatial location and marks for the \( i \)th object.
- The math becomes complicated because \( w \) is a non-ordered collection of points, unlike other image models.
  - The density function \( p(w) \) is the representation of the probability distribution that we optimize to solve real problems.
  - The MPP density function cannot be defined as usual.
    - Most density functions used in practice are defined with respect to the Lebesgue measure.
    - Since the MPP is a non-ordered collection of points, the Lebesgue measure cannot be used.
    - We have to resort back to measure theory.
    - The measure used to address this problem is the Lebesgue-Poisson measure (INRIA).
  - Practical implication: Typical derivatives do not exist, so MCMC sampling is used for optimization.
The Missing Link: The Gibbs distribution

• For a configuration of random variables \( v \) (e.g., an image), a Gibbs distribution has the form

\[
p(v) = \frac{1}{z} \exp(-H(v)),
\]

where \( H(v) \) is the energy for configuration \( v \), and \( z \) is the partition function.

• All of the image models described in this presentation have Gibbs distributions.

• Physics models can often be written in terms of an energy function \( H(x) \):
  – For the Potts model, for example,
    \[
    H(x) = \beta \sum_{(i,j)} \delta(x_i, x_j),
    \]
    where the summation is over all pairs of neighboring sites \( i \) and \( j \), \( \beta = 1/kT \), and \( \delta() \) is the Kronecker delta function.
  – A configuration \( x \) that minimizes the energy \( H(x) \) also maximizes the probability distribution \( p(x) \).
We begin by defining random variables in our system.

Random variables in previous examples:
- $Y$: Experimental image data, modeled using a Gaussian model
- $X$: Pixel classifications (segmentation) modeled with Potts model
- $W$: Spatial locations and shape parameters of particles modeled using MPP

It is not clear how to combine all of these into a single energy function $H$.

We model smaller parts, for example the MPP and the MRF, and combine them mathematically to get, for example, the MPP/MRF.

- Probability theory provides a rigorous, time-tested, elegant way to combine the smaller parts, using
  - Conditional probability
  - Bayes Theorem
  - The total probability law
Combining the MPP and MRF models

denote \( w \) the object configuration, \( x \) the segmentation and \( y \) the observed image. The new model is defined as:

\[
p(w, x | y) = \frac{f(y | w, x)p(w, x)}{f(y)}
\]

\( f(y | w, x) \) is the conditional probability density function of \( y \) given \( w, x \). \( p(w, x) \) is a joint probability mass function.
Model for the observed image \( f(y|w, x) \)

We define the **Silhouette** \( S_w \) as the projection of \( w \) onto the image region, where \( w = (w_1, w_2, \ldots, w_n) \) is an object realization.

\[
\begin{align*}
f(y|w, x) &= \prod_{s \in S_w, x_s = f} \frac{\exp\left(-\frac{(y_s - \mu_f)^2}{2\sigma_f^2}\right)}{\sqrt{2\pi \sigma_f^2}} \prod_{s \in S_w, x_s = b} \frac{\exp\left(-\frac{(y_s - \mu_b)^2}{2\sigma_b^2}\right)}{\sqrt{2\pi \sigma_b^2}} \\
&\quad \prod_{s \in S / S_w, x_s = f} \frac{\exp\left(-\frac{(y_s - \mu_f)^2}{2\sigma_f^2}\right)}{\sqrt{2\pi \sigma_f^2}} \prod_{s \in S / S_w, x_s = b} \frac{\exp\left(-\frac{(y_s - \mu_b)^2}{2\sigma_b^2}\right)}{\sqrt{2\pi \sigma_b^2}}
\end{align*}
\]

As in the equation, we classify the pixels into four categories: 
foreground pixels in object regions 
background pixels in object regions 
foreground pixels out of object regions 
background pixels out of object regions
\[
p(w, x) = \frac{1}{z(x, w)} \exp(- \sum_{\{r,s\} \in \mathcal{C}} \beta_1 t_1(x_r, x_s) - \sum_{s \in S/S_w} \alpha_1 t_1(x_s, f) - \sum_{s \in S_w} \alpha_2 t_1(x_s, b) - \sum_{w_i \sim w_j} \beta_2 t_2(w_i, w_j))
\]

where

\[
t_1(a, b) = \begin{cases} 
0 & \text{if } a = b \\
1 & \text{if } a \neq b 
\end{cases}
\]

\(z(x, w)\) is the normalizing constant and \(t_2(w_i, w_j)\) is a function describing the interaction of object pairs. \(\beta_1\) and \(\beta_2\) are the interaction parameters for the pixel pair and object pair. \(\alpha_1\) and \(\alpha_2\) are the weighting parameters, which impose penalization when the label field and object field are not consistent.

\[\sum_{s \in S/S_w} t_1(x_s, f)\] are foreground pixels out of object regions, \[\sum_{s \in S_w} t_1(x_s, b)\] are background pixels in object regions. Although we allow these two kind of pixels in \(f(y|w, x)\), we give penalty \((\alpha_1, \alpha_2)\) on them in \(p(w, x)\).
Optimization of the Model

Our target is to find the \((w, x)\) that maximize the posterior possibility \(p(w, x|y)\). Since \(f(y)\) does not depend on \(w\) or \(x\), it is equivalent to obtain such a \((w, x)\) that minimize an energy function \(V(w, x) = V_d(w, x) + V_p(w, x)\), where \(V_d(w, x) = -\ln(f(y|w, x))\) and \(V_p(w, x) = -\ln(p(w, x))\).

Often minimizing over both \(w\) and \(x\) simultaneously is very difficult. Instead, we try to minimize \(p(w, x|y)\) with respect to \(w\) while keeping \(x\) fixed, which can be solved more easily. Then we alternate this procedure iteratively. Such an approach, usually referred to as the alternating minimization algorithm [6], is summarized as:

**Algorithm 1: Minimization of the posterior possibility**

1. Initialization: \(w^0 = \emptyset, \ x^0 = \arg\min_x V_d(w^0, x) + V_p(w^0, x)\);
2. for \(k = 1, \ldots, n\) do
3. \(w^{k+1} = \arg\min_w (V_d(w, x^k) + V_p(w, x^k))\);
4. \(x^{k+1} = \arg\min_x (V_d(w^{k+1}, x) + V_p(w^{k+1}, x))\);
Example: Silicates

Silicates (from Larry Drummy)

Mesoscale (MPP line) model only

Microscale (MRF) model only

Joint meso/microscale model
Example: Si-TiSi$_2$

Si-TiSi$_2$ (from our distinguished PM Ali)

EM/MPM segmentation

Joint MPP/MRF feature extraction and segmentation

MPP/MRF feature extraction

MPP/MRF segmentation
Impact

• NSRC workshop presentation
• Invited talk to be given at Argonne
• Invited talk to be given at Fall 2015 MRS
• Key papers
  – ...
  – ...

Future Plans

• Near-term:
  – 2D software in DREAM.3D
  – 3D implementation
  – Dynamics: 4D

• Long-term (long,long-term?)
  – Materials design

Consider a system characterized by a collection of random variables $s$ (e.g., $s = x,y,w$). For a design parameter $d$, one way to find the optimal $d$ is through maximum likelihood estimation, given by

$$d_{ML} = \underset{d}{\text{argmax}} \ p(s|d)$$
Future Outlook

• Currently have models for spheres, ellipses, superellipses, channels, and lines
  – Will work on materials systems of interest to the team using these models
  – Will develop new models as necessary
• 4D materials
• Computation
  – Processing one 2D slice takes on the order of one hour
  – Processing 3D data will require computational speedup
    • Method is inherently parallelizable
• User-friendly GUI under development
  – Plan to add to DREAM.3D