PROBABILITY MODELS FOR IMAGE UNDERSTANDING

Anuj Srivastava

Laboratory for Computational Vision (http://lcv.stat.fsu.edu)
Department of Statistics, Florida State University

Collaborators: Ulf Grenander, Brown University
Xiuwen Liu, Florida State University

This work is supported by:
ARO DAAD19-99-0267, NIMA NMA201-01-2010, NSF FRG DMS-0101429.
OUTLINE OF THE TALK

1. Image Models:
   (a) Low Level models: Markov Random Fields
   (b) High Level Models: Deformable templates
   (c) Intermediate Levels: ??

2. Image Spectra & Bessel K forms

3. Applications of Bessel K forms
   • Homogeneous Texture Synthesis
   • Clutter Classification
   • Target Recognition
FRAMEWORK

- Computational vision is treated as a problem in statistical inference.
- Bayesian framework is fundamental:
  \( R \) is a representation of the unknown, \( I \) is the observation
  \[
  P(R|I) = \frac{P(I|R)P(R)}{P(I)}.
  \]
- The choice of \( R \) and the probability model depends upon the goal: image analysis, compression, denoising, segmentation, classification, or understanding.
- \( I \) is an element of a high dimensional space. The biggest challenge is to:
  build probability models while reducing dimensions.
PYRAMIDS of REPRESENTATIONS

Primitives: Bright pixels, edges, junctions, Curves, regions, 2D shapes, diffeomorphic groups

3D shapes, texture, articulation, transformation groups

Histograms of primitives, Histograms + locations

Structure, Dimensions

Image

Primitives: Bright pixels, edges, junctions

Curves, regions, 2D shapes, diffeomorphic groups

3D shapes, texture, articulation, transformation groups

Histograms + locations

Histograms of primitives

Primitives: Bright pixels, edges, junctions

Image
COMMON IDEA: DIMENSION REDUCTION

- Images are considered arrays of numbers, no physical considerations are attached.
- Seek a low-dimensional subspace that best represents those numbers (under some chosen criterion).
- Examples: Principal components, independent components, sparse coding, Fisher’s discriminant, Karhunen-Loeve, etc.

**Advantage**: Speed, Computational Efficiency

**Drawback**: Knowledge deficient.
- No physical or contextual information about the imaged objects
- Performance is limited, specially in clutter, low SNR.
DEFORMABLE TEMPLATES

Grenander’s School (*General Pattern Theory* 1993), Kendall’s school etc.

- Explicitly model shapes, textures, illuminations ...

![Images]

- Typical occurrences (templates) + Variations (deformations)

- Model the imaging process: a perspective projection map from \( \mathbb{R}^3 \) to \( \mathbb{R}^2 \).

- Probability(Images) \( \equiv \) Probability(physical variables) and Probability(Imaging)

**Advantage**: Detailed, good performance

**Drawback**: Computationally expensive to implement, “analysis by synthesis”
PROBABILITY MODELS ON $\mathbb{R}$

Having chosen $\mathbb{R}$, what are possible $P(\mathbb{R})$?

1. Markov random field models
2. Gaussian models, other parametric forms
3. Non-parametric forms
4. Minimax models
5. Probability measures on groups

We propose a two-parameter family for modeling image spectra.
MOTIVATION: EMPIRICAL STUDIES

Natural Image:
Seeking probability models, assume stationarity and start with a histogram of pixel values
Mumford and some others:
Take spatial derivatives and study their histograms
horizontal and vertical

Similar histograms for directional derivatives, higher derivatives, Laplacian, other differential operators.
Marr’s work on modeling animal vision suggested certain specific operators:

(1) Gabor Filters:

(2) Laplacian-Gaussian operators:
EXPLAINING THESE PATTERNS

Empirical studies show: (i) symmetry, (ii) heavier tails and sharper cusps at mode (leptokurtic), (iii) correlations across scales.

Possible Explanations:

1. Dead-Leaves Model: Mumford et al.

2. "Ubiquity of egdes" (curvelets transforms): Donoho et al.

3. Gaussian scale mixtures: Simoncelli et al.

4. “1/r^3 - law” for scale invariances: Geman et al

5. "Images are made of objects": Grenander and Srivastava
GENERATOR MODEL

An image is made up of weighted superposition of randomly placed object profiles.

\[(\log)I(z) = \sum_{i}^{n} a_ig(z - z_i), \; z, z_i \in \mathbb{R}^2, \; a_i \in \mathbb{R} .\]

\(a_i\)'s are standard normal, \(z_i\)'s are according to a homogeneous Poisson process.

What probability does this lead to?
SPECTRAL COMPONENTS OF IMAGES

- For a bank of filters \( \{F^{(j)}, j = 1, 2, \ldots, K\} \), a spectral component is
  \[
  I^{(j)} = I \ast F^{(j)}, \quad \ast \text{ denotes 2D convolution}
  \]

- Example: \( F \) is Gabor filter with orientation \( \theta \) and scale \( \sigma \):

- Decompose image \( I \) into its spectral components:
  \[
  I \Rightarrow \{I^{(1)}, I^{(2)}, \ldots, I^{(K)}\}.
  \]

- A spectral component is modeled as:
  \[
  I^{(j)}(z) = \sum_i a_i g^{(j)}(z - z_i), \quad \text{where } g^{(j)} = F^{(j)} \ast g.
  \]
PROBABILITY DENSITY of \( I^{(j)}(z) \)

- The **conditional** density of \( I^{(j)}(z) \),
  given the Poisson points \( \{z_i\} \) and the profile \( g \),
  is **Gaussian** with mean zero and variance \( u \), where

  \[
  u \equiv \sum_i (g^{(j)}(z - z_i)^2 .
  \]

- \( I^{(j)}(z) \) is a **normal variance mixture** with mixing variable \( u \).

- Relation between parameters of \( u \) and the parameters of the filtered image.

  \[
  E[I^{(j)}(z)] = 0, \quad E[I^{(j)}(z)^2] = E[u],
  \]
  \[
  E[I^{(j)}(z)^4] = 3E[u^2], \quad \text{and kurtosis}(I^{(j)}(z)) = \frac{3\text{var}(u)}{(E[u])^2}.
  \]
PROBABILITY DENSITY of $I^{(j)}(z)$...

Model $u$ by a scaled $\Gamma$-density:

$$f_u(u) = \frac{1}{c\Gamma(p)} (u/c)^{p-1} \exp(-u/c); \quad c > 0, \ p > 0,$$

$$E[u] = pc, \ \text{var}(u) = pc^2$$

Integrate over the variation of $u$ and derive the unconditional density of $I^{(j)}(z)$.

**Theorem 1** The probability density function of $I^{(j)}(z)$ is: for $p > 0, c > 0$,

$$f(x; p, c) = \frac{1}{Z(p, c)} |x|^{p-0.5} K_{p-0.5}(\sqrt{\frac{2}{c}} |x|),$$

(1)

where $K$ is the modified Bessel function and

$$Z(p, c) = \sqrt{\pi} \Gamma(p) (2c)^{0.5p+0.25}.$$
BESSEL K FORMS for IMAGE SPECTRA

Let \( D = \{ f(x; p, c) | p > 0, c > 0 \} \).

- elements of \( D \) are called **Bessel K forms**
- the parameters \((p, c)\) are called the **shape** and **scale** parameters.

Properties of the elements of \( D \):
(i) Unimodal, (ii) Leptokurtic, (iii) infinitely-divisible, and (iii) Square-integrable only for \( p > 0.25 \).

**Estimation** of Bessel Parameters

\[
\hat{p} = \frac{3}{\text{SK}(I^{(j)}) - 3}, \quad \hat{c} = \frac{\text{SV}(I^{(j)})}{\hat{p}}
\]

where \( SK \) is sample kurtosis and \( SV \) is sample variance of pixels in \( I^{(j)} \).
EXAMPLES
EXAMPLES

- Intensity in Filtered Image (Relative Frequency)

Options:

- Intensity in Filtered Image

Parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>1.093</td>
</tr>
<tr>
<td>c</td>
<td>62386.36</td>
</tr>
<tr>
<td>p</td>
<td>0.778</td>
</tr>
<tr>
<td>c</td>
<td>81645.15</td>
</tr>
<tr>
<td>p</td>
<td>1.436</td>
</tr>
<tr>
<td>c</td>
<td>74635.87</td>
</tr>
<tr>
<td>p</td>
<td>0.254</td>
</tr>
<tr>
<td>c</td>
<td>376.38</td>
</tr>
<tr>
<td>p</td>
<td>0.946</td>
</tr>
<tr>
<td>c</td>
<td>31346.23</td>
</tr>
<tr>
<td>p</td>
<td>1.604</td>
</tr>
<tr>
<td>c</td>
<td>151002.70</td>
</tr>
</tbody>
</table>
EXAMPLES: INFRARED IMAGES

Filter angle 148

Filter angle 36

Filter angle 111

Intensity in Filtered Image

Relative Frequency
APPLICATIONS

1. Homogeneous Textures
2. Clutter classification
3. Pruning for target recognition
HOMOGENEOUS TEXTURES

- Choose a filter bank and an image \( I \).
- Estimate \( p, c \) for each filter on \( I \).
- Compute histogram of intensity values in \( I \).
- Use filtered marginals and the histogram of \( I \) to synthesize new image using Gibb’s sampling.

Examples: Real observed and synthesized
Examples: Real observed and synthesized
EXAMPLE: CLUTTER CLASSIFICATION

$I_1$  $I_2$  $I_3$  $I_4$

$I_5$  $I_6$  $I_7$  $I_8$

$I_9$  $I_{10}$
Clustering Chart:

Figure 1: Left panel: using $d_J$. Right panel: using Euclidean metric on PC.
NIGHTTIME FACE RECOGNITION

FSU IR FACE Database: Examples
<table>
<thead>
<tr>
<th>Test/training ratio</th>
<th>Correct to be the closest</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigen faces</td>
<td>Independent faces</td>
<td>Bessel forms</td>
</tr>
<tr>
<td>1:1</td>
<td>90.48%</td>
<td>89.52%</td>
<td>97.14%</td>
</tr>
<tr>
<td>3:1</td>
<td>87.42%</td>
<td>88.05%</td>
<td>91.20%</td>
</tr>
<tr>
<td>7:1</td>
<td>80.21%</td>
<td>70.05%</td>
<td>83.96%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test/training ratio</th>
<th>Correct within the closest two</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eigen faces</td>
<td>Independent faces</td>
<td>Bessel forms</td>
</tr>
<tr>
<td>1:1</td>
<td>94.29%</td>
<td>96.19%</td>
<td>99.05%</td>
</tr>
<tr>
<td>3:1</td>
<td>92.45%</td>
<td>91.20%</td>
<td>96.23%</td>
</tr>
<tr>
<td>7:1</td>
<td>87.70%</td>
<td>83.42%</td>
<td>91.98%</td>
</tr>
</tbody>
</table>
COMMENTS

- Proposed Bessel K forms for modeling image spectra.
- We have extended them to bivariate densities using eight parameters. Other extensions possible.
- Given the different $R$ and $P(R|I)$ we have (MRF models, Bessel K forms, deformable template models, shape statistics), what else is needed?
- Great need for **sequential inference procedures** that can glue these models together.