University of New Hampshire
College of Engineering and Physical Sciences
Electrical and Computer Engineering Department

BIOLOGICALLY INSPIRED FEATURE
EXTRACTION FOR ROTATION AND SCALE
TOLERANT PATTERN ANALYSIS

Ph. D. Dissertation by
Dragan Vidacic
Presentation Outline

Introduction
- Problem statement
- Importance of the problem
- Specific project/research objectives

Background
- Solutions of recurrent linear networks applied in log-polar domain
- Construction of biologically inspired filters utilizing spectral properties of Toeplitz-Block-Toeplitz (TBT) matrices
- Semi-blind pre-whitening by spatially invariant networks with lateral inhibition/excitation
- Biologically inspired information pre-processing and object recognition

Conclusion and directions

Questions
The Problem

Computationally efficient feature extraction suitable for pattern analysis – how can one teach the machine to recognize the surrounding environment as well as humans do?
The Goal

- Attempt to develop computationally efficient, biologically inspired, image processing framework suitable for effective pattern recognition
The Approach

Mimic operations performed by the Human Visual System (HVS) through integrated data processing chain

– Spatially non-uniform architecture (foveation and warping)
– Lateral feed backward connectivity (recurrent neural information processing)
– Cortical information processing - sparse signal coding
Human Visual System
Log-Polar Mapping - Motivation

- Data reduction at the retinal sensory level through non-uniform sampling
  - Ganglion cell receptive field size increases linearly with eccentricity

- Retino-Cortical transform
  - inverse of cortical magnification factor is linearly dependent on eccentricity
  - Complex-log mapping as a model for retino-cortical projection
  - Rotation and size invariance in the resulting domain
Complex Conformal Log Mapping

The coordinate transform (continuum):

Input Space

Cortical Space

\[ x + iy \]

\[ u = \ln\left(\frac{r}{a}\right) \]
Log-Polar Mapping

Discrete transform (polar exponential grid):

Input Space

Cortical Space
Recurrent neural processing

- Lateral inhibition/excitation between receptor-neuron (sensory) fields
- Subtractive model can be fitted into data taken from Limulus

\[ f_{ij} = e_{ij} - \sum_{k=1}^{N} \sum_{l=1}^{N} b_{ijkl} f_{kl} \]
Recurrent neural processing (2)

The network solution is expressed as:

$$ \bar{F} = (I + B)^{-1} \bar{E} $$

- $B$ is symmetric Toeplitz-Block-Toeplitz (TBT) matrix describing the inter-neuron coupling; various models of coupling possible
- Edge detection/enhancement property of LSI
- Preserves relative intensity levels
Recurrent linear networks and Log-Polar Mapping

- How to implement the lateral inhibition/excitation operation consistent with log-polar mapping?
- The role of neural signal pre-processing in information encoding?
- How can the obtained results be used for effective feature extraction and pattern recognition?
The Proposed System

The system block diagram:

Original Image → Non-Uniform Sampling / Log-Polar Mapping → Retino-Cortical Filter → Cortical Processing (Simple Cells + ?) → Feature Extraction & Pattern Recognition

Lateral inhibition/excitation is implemented consistent with log-polar mapping through a retino-cortical filter.
Importance of the Problem

- Efficient image processing techniques are of great importance for computer vision applications.
- Spatially variant vision systems allow for reduction of computational costs while preserving the interesting signal detail.
- Contributions in image processing in foveated and warped environment.
- Reducing the gap between biology and manmade machines.
- Effective pattern recognition.
Specific Objectives

- Describe the response of the receptor array with pre-determined inter-neuron coupling as accurately as possible.
- Development of retino-cortical filter through uniform cortical processing.
- Find and evaluate the model of non-uniform LSI consistent with log-polar mapping.
- Explore the role of the frontal HVS – like pre-processing filters within the broader context of a signal processing and pattern recognition framework.
Background (1)

Physiology of vision
- P. M. Daniel and D. Whitteridge;
- D. M. Hubel and T. N. Wiesel;
- E. Schwartz
- B. Fisher
- H. K. Hartline, H. G. Wagner and F. Ratliff

Computer Vision, Image Processing and Log-Polar Mapping
- M. Bolduk and M. D. Levine
- A. Tabernero, J. Portilla and R. Navarro
- B. B. Bederson, R. S. Wallace and E. Schwartz,
- R. A. Messner and H. H. Szu
- C. F. Weiman and R. D. Juday
- E. Nattel and Y. Yeshurun

Lateral inhibition
- G. G. Furman
- R. A. Messner and H. H. Szu
- Y. Yu, T. Yamauchi and Y. Choe
- M.G. Luniewicz and R.A. Messner
Background (2)

Cortical signal processing – decorrelation, sparse coding and Independent Component Analysis (ICA)

– H. Barlow and P. Földiák
– M. D. Plumbley
– B. A. Olshausen and D. J. Field
– A. Hyvärinen and E. Oja
– A. J. Bell and T. J. Sejnowski
Solutions of recurrent linear networks applied in log-polar domain

- Computationally efficient implementation of recurrent network as a complement to non-uniform process of log-polar mapping
- General recurrent network solutions
  - Direct network solution - requires large matrix inverse
  \[ \overline{F} = \left( I + B \right)^{-1} \overline{E} \]
  - Iterative solution – very slow
  \[ f_{ij}^n = e_{ij}^n - \sum_{k=1}^{N} \sum_{l=1}^{N} b_{ijkl} f_{kl}^p \]
  - Convolution based model: find the network impulse response and convolve the input
Solutions of recurrent linear networks applied in log-polar domain (2)

Two practical neural coupling models (taken from literature):

Model A: \[ b_{ijkl} = 0.25 \frac{1}{d^3} \]

Model B: \[ b_{ijkl} = e^{-1.55d} \]
Convolution kernels

- Model A and Model B kernels are center-surround

- Linear combination of four Difference of Gaussian (DOG) kernels
Convolution in Log-Polar Domain - Filter Warping

- Perform the convolution in log-polar domain by uniform kernels and achieve the result equivalent to non-uniform filtering at input.
- Linear scaling of receptive fields with distance from fovea results in uniform cortical processing.
- Signal energy has to be preserved (Jacobian).
- Warping the real discrete network response is problem.
Non-Uniform Lateral Inhibition - the Exact Network Solution

- Concept of non-uniform inter-neuron coupling based on distanced in PEG
- Computations in warped log-polar space

\[ b_{ijmn} = e^{-kd} \]
Non-Uniform Lateral Inhibition - Results

Original Image

Processed Image
Characterization of the Space Invariant Network with Lateral Connectivity in Log-Polar Space

- Center-surround property of discrete network impulse response is combined with averaging receptive field in PEG => spatially variant receptive field
- Example: Gaussian PEG receptive field and kernel from model B coupling form Difference of Offset Gaussian (DOOG) center-surround filter

Network implementation achieved through uniform operations in log-polar space
- Convolution based for large size problems
- Direct network solution for small size problems
Construction of Biologically Inspired Filters Utilizing Spectral Properties of Toeplitz-Block-Toeplitz (TBT) Matrices

- Spatially invariant connection model
- Direct network solution substituted by the solution based on Taylor series expansion
- TBT network connectivity matrix structure

\[
B_{n,m} = \begin{bmatrix}
B_0 & B_1 & \cdots & B_{n-1} \\
B_{-1} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & B_1 \\
B_{-n+1} & \cdots & B_{-1} & B_0
\end{bmatrix}
\]

\[
B_l = \begin{bmatrix}
b_{l,0} & b_{l,1} & \cdots & b_{l,m-1} \\
b_{l,-1} & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & b_{l,1} \\
b_{l,-m+1} & \cdots & b_{l,-1} & b_{l,0}
\end{bmatrix}
\]

TBT matrix generating function:

\[
b_{p,q} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(\xi, \eta) e^{-j(p\xi + q\eta)} d\xi d\eta
\]

- Eigenvalues of the TBT matrix are bounded by the generating function extremes
Construction of Biologically Inspired Filters Utilizing Spectral Properties of Toeplitz-Block-Toeplitz (TBT) Matrices (2)

- If the generating function is bounded by unity the solution of the network can be found as direct implementation of Taylor series

\[
\max_{-\pi,\pi} \left| f(\xi, y) \right| \leq 1 \quad \Rightarrow \quad y = \sum_{i=-\infty}^\infty B^i x = \left( -B + B^2 - B^3 + \cdots + 1 \right) x
\]

- Matrix inversion is not necessary
- Multiplication of a vectorised 2-D signal with TBT matrix represents convolution with kernel:

\[
b(n, n) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(\xi, \eta) e^{-jn\xi + n\eta} \, dx \, dy
\]

Discretised connection function

- Network impulse response as cascade of N+1 filters:

\[
h_t(n, n) = \delta(n, n) + \sum_{i=1}^N \sum_{j=1}^n b(j) \delta(n-j, n)
\]
Construction of Biologically Inspired Filters Utilizing Spectral Properties of Toeplitz-Block-Toeplitz (TBT) Matrices (3)

- Filter steerability - arbitrarily rotated function can be synthesized by using the linear combination of basis kernels
- Connection function expandable in Fourier series in polar angle
  \[ b(\cdot, \rho) = \sum_{n=-N}^{N} a_n e^{jn\phi} \]
  \[ B(\cdot, \theta) = \sum_{n=-N}^{N} A_n e^{jn\theta} \]
- Cascades of these filters are still steerable
- The sum of cascades (overall network) is also steerable since it is represented as a Fourier decomposition in polar angle
- Order n directional derivative of circular symmetric function very suitable for generating steerable network
- The network response steered with 2nN+1 basis functions, N order of Taylor polynomial
Steerable network response

\[ g^{2,1}_{\alpha,\sigma}(x,y) \]

\[ g^{2,2}_{\alpha,\sigma}(x,y) \]

\[ g^{2,3}_{\alpha,\sigma}(x,y) \]

\[ \delta(x,y) \]

\[ h^\theta_t(x,y) = \sum_{i=1}^{7} k_i(\theta) h^\theta_i(x,y) \]
Number of filters in Taylor series for five practical models

![Graph showing number of filters in Taylor series for different models]

- **Gaussian - with self-inhibition**
- **Gaussian - no self-inhibition**
- **DOOG**
- **2DOG**
Generating steerable network response – synthetic pattern
Semi-Blind Pre-Whitening by Spatially Invariant Networks with lateral Inhibition/Excitation

- Linear model of network with recurrent inhibition can learn about the input signal statistics
- Anti-Hebbian learning produces decorrelated outputs

\[ y_i = x_i - \sum_j b_{i,j} y_j \]
\[ \Delta b_{i,j} = \alpha y_i y_j \quad \text{if } i \neq j \]
\[ \Delta b_{i,j} = 0 \quad \text{otherwise} \]

- Signal whitening

\[ C_Y = E \left( y y^T \right) \beta I \]

- Whitening algorithm (self-inhibition):

\[ \Delta B = \alpha \left( y y^T - \beta I \right) \]
Semi-Blind Pre-Whitening by Spatially Invariant Networks with lateral Inhibition/Excitation (2)

- ICA – signal decomposition into statistically independent events/components
- ICA captures essential structure of data
- ICA produces sparse codes
- Whitening as pre-processor for ICA
  - Easier or well posed problem
  - Improves convergence of ICA algorithms
  - Transform reduced to orthogonal matrix solution
- Decorrelation role of recurrent network with and without learning capability – adaptive model far superior
Category Specific Pre-Whitening Filters

- Constrain the whitening algorithm to produce TBT matrix with symmetric Toeplitz blocks
- Small memory footprint
- Operates on smaller image patches – fast response
- Trained on specific category of input data
- Batch training mode
Learning pre-whitening filters

1. Initialize:
   - $\alpha$ (small), $B = 0$
   - $\beta = 1, Tr$ (small), $Nm$

2. Calculate Input Data Covariance Matrix
   \[ C_x = E\{xx^T\} \]

3. Calculate:
   - $y = (I + B)^{-1}x = Sx$
   - $C_r = SC_xS^T$
   - $B = B + \alpha(C_r - \beta I)$

4. Convert $B$ to Toeplitz-Block-Toeplitz

5. Evaluate:
   \[ E_f = \sqrt{\sum_{i,j} |C_{r,i,j} - \beta \delta_i|^2} < Tr \text{ or } \text{Number of Iterations} > Nm \]

   - Yes: Stop
   - No: Repeat from step 3

---

**Data (Image Patch) Acquisition and Mean Removal**

**Calculate Input Data Covariance Matrix**

**Calculate**

**Convert $B$ to Toeplitz-Block-Toeplitz**

**Evaluate**

---

**Stop**
Input data types
Filter performance

Average value for $E_F$ - Category 2

- No pre-processing
- Uniform pre-whitening
- Complete whitening
- Uniform pre-whitening - Category 1
- Uniform pre-whitening - Category 3
Biologically Inspired Information Pre-Processing and Object Recognition

- Cortical signal representations are sparse
- Redundancy reduction
- Localized, oriented and bandpass responses of simple cells
- ICA – equivalent to sparse signal encoding

\[ I(\xi, y) = \sum_i s_i a_i(\xi, y) \]

- PDF of sparse coding coefficients
Biologically Inspired Information Pre-Processing and Object Recognition (2)

ICA decomposition of input image \( I \) by serializing the image rows into input vectors \( x \)

\[ x = As \]

- Mixing matrix (columns are basis functions)
- vector of independent components

FastICA (Fixed Point ICA)
- Learn the de-mixing matrix \( W = A^{-1} \)
- Minimization of negentropy (nongaussianity)
- Very fast algorithm
- No user defined parameters
- Computationally simple
- Readily available code…

For given data input the ICA decomposition is \( s = Wx \)

The same pre-whitening filter used for ICA learning and processing

Two stage signal whitening: pre-whitening filter/Plumbley algorithm

Feature code:
\[ s = B_1 \tilde{x} \]
\[ \tilde{x} \]
- Whitened input
\[ W = B_1 B_{WT} \quad B_{WT} \]
- Overall whitening transform
Recognition of Objects through Log-Polar Transform and Local ICA Analysis

1. Input Grayscale Image Acquisition
2. Centroid Calculation and Object Centering
3. Log-Polar Mapping
4. Complex Log-Gabor Filtering (Pseudo-Saliency Map)
5. Collection of Image Patches
6. Precalibration with Category-Specific Filters
7. Whitening by Plumbley Algorithm
8. Feature Encoding (Multiscale ICA with Component Thresholding)

Yes: Store Feature Vector(s)
No: Training?
Yes: Classification
Pseudo-saliency

- Rigid patch sampling – problem for log-polar rotation/scale tolerance

Solution: pseudo-saliency through log-Gabor filters

Original image / salient locations

Filter response

Extracted patches
**Feature Matrix**

\[
FM = \begin{bmatrix}
    x_1 & x_2 & \cdots & x_{16} \\
    y_1 & y_2 & \cdots & y_{16} \\
    hs_{1,1} & hs_{2,1} & \cdots & hs_{16,1} \\
    \vdots & \vdots & \ddots & \vdots \\
    hs_{1,256} & hs_{2,256} & \cdots & hs_{16,256} \\
    ls_{1,1} & ls_{2,1} & \cdots & ls_{16,1} \\
    \vdots & \vdots & \ddots & \vdots \\
    ls_{1,256} & ls_{2,256} & \cdots & ls_{16,256}
\end{bmatrix}
\]

- Coordinate of the patch
- Higher scale ICA decomposition
- Lower scale ICA decomposition
Experimental results

- Image patches 16x16 pixels
- 2000 patches per image
- Rotated (artificially) and non-rotated objects: 40 objects/12 images per object/160 test images
- Feature matrix – elimination of small coefficients
- Four classification techniques – all distance based
- Pattern variations
Experimental results – rotated objects

<table>
<thead>
<tr>
<th>Number of non-zero ICA coefficients</th>
<th>Recognition result pass/fail for specific classifier</th>
<th>Hybrid Nearest Neighbour</th>
<th>Hybrid 3-Nearest Neighbour</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Hybrid Nearest Neighbour</td>
<td>Hybrid 3-Nearest Neighbour</td>
</tr>
<tr>
<td>256</td>
<td></td>
<td>153/7</td>
<td>159/1</td>
</tr>
<tr>
<td>80</td>
<td></td>
<td>153/7</td>
<td>159/1</td>
</tr>
<tr>
<td>64</td>
<td></td>
<td>152/8</td>
<td>159/1</td>
</tr>
<tr>
<td>48</td>
<td></td>
<td>152/8</td>
<td>159/1</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td>152/8</td>
<td>159/1</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>154/6</td>
<td>159/1</td>
</tr>
</tbody>
</table>
## Recognition results – normalized x and y features

<table>
<thead>
<tr>
<th>Number of non-zero ICA coefficients</th>
<th>Recognition result pass/fail – non-rotated objects</th>
<th>Recognition result pass/fail – rotated objects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hybrid Nearest Neighbour</td>
<td>Hybrid 3-Nearest Neighbour</td>
</tr>
<tr>
<td>256</td>
<td>148/12</td>
<td>146/14</td>
</tr>
<tr>
<td>80</td>
<td>147/13</td>
<td>146/14</td>
</tr>
<tr>
<td>64</td>
<td>146/14</td>
<td>147/13</td>
</tr>
<tr>
<td>48</td>
<td>145/15</td>
<td>147/13</td>
</tr>
<tr>
<td>32</td>
<td>146/14</td>
<td>148/12</td>
</tr>
<tr>
<td>16</td>
<td>145/15</td>
<td>146/14</td>
</tr>
</tbody>
</table>
Directions

- Utilization of concept of network solution without matrix inversion during the learning of pre-whitening filters
- Determine more exact criterion that describes the category of images the pre-whitening filter is trained for
- Usage of steerable networks for characterization of input data sets
- More robust classifiers – neural models?
- Improvement of signal preprocessing stages (object centering…)
- Dimensionality reduction of sparse feature set
- Test the model on more complex problems
Conclusions

Meaningful feature extraction based on biologically inspired signal processing

Uniform recurrent neural network operations in log-polar domain correlate well with center-surround space variant receptive fields of HVS

Characterization of the response of the spatially uniform inhibitory/excitatory neural models based on the synaptic connectivity of the network
  - Taylor series expansion based on bounded generating function of TBT matrix
  - Steerability of the network response

Successful development of compact semi-blind pre-whitening filters

Promising pattern recognition results based on localized ICA signal encoding
Many thanks to

- My advisor Richard A. Messner
- Ph. D. Committee members
- Friends and
- My wife Anka
Questions?