Computer Vision

Edge and Texture

Jianbo Shi
how our brain acts when we see this:
how our brain acts when we see this:
What’s in front?
What is it?
HUMAN VISUAL PATHWAY begins with the eyes and extends through several interior brain structures before ascending to the various regions of the visual cortex (V1, and so on). At the optic chiasm, the optic nerves cross over partially so that each hemisphere of the brain receives input from both eyes. The information is filtered by the lateral geniculate nucleus, which consists of layers of nerve cells that each respond only to stimuli from one eye. The inferior temporal cortex is important for seeing forms. Researchers have found that some cells from each area are active only when a person or monkey becomes conscious of a given stimulus.
Figure 1

The Human Eye

- Cornea
- Iris
- Pupil
- Lens
- Retina
- Eye Muscle
- Optic Nerve
Human Visual System

[Diagram of the human visual system, showing the eye, light energy, and various parts of the brain.]
LGN cells

1) Left and Right LGN, each has 6 Layers
2) 0.9 Million retinal optic fiber enters each LGN, 0.1 M goes to eye movement controller
3) LGN also gets input from visual cortex,
4) For every 10 nerve inputs from retina, only 4 got send to cortex
5) The optics from the same side of eye goes to layer 2, 3, 5, and opposite to 1, 4, 6
6) Retinal map is organized in layer 6
The Human Visual System

Dorsal Stream

Second level of visual association cortex in parietal lobe

Dorsal lateral geniculate nucleus

Thalamus

Eye

Optic nerve

Inferior temporal cortex: Second level of visual association cortex

Dorsal Stream

Striate cortex (primary visual cortex)

Extrastriate cortex

Ventral Stream
$P_b$ Images

<table>
<thead>
<tr>
<th>Image</th>
<th>Canny</th>
<th>2MM</th>
<th>Us</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image.png" alt="Image 1" /></td>
<td><img src="image.png" alt="Canny 1" /></td>
<td><img src="image.png" alt="2MM 1" /></td>
<td><img src="image.png" alt="Us 1" /></td>
<td><img src="image.png" alt="Human 1" /></td>
</tr>
<tr>
<td><img src="image.png" alt="Image 2" /></td>
<td><img src="image.png" alt="Canny 2" /></td>
<td><img src="image.png" alt="2MM 2" /></td>
<td><img src="image.png" alt="Us 2" /></td>
<td><img src="image.png" alt="Human 2" /></td>
</tr>
</tbody>
</table>
$P_b$ Images II

<table>
<thead>
<tr>
<th>Image</th>
<th>Canny</th>
<th>2MM</th>
<th>Us</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="canny1.png" alt="Canny" /></td>
<td><img src="2mm1.png" alt="2MM" /></td>
<td><img src="us1.png" alt="Us" /></td>
<td><img src="human1.png" alt="Human" /></td>
</tr>
<tr>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="canny2.png" alt="Canny" /></td>
<td><img src="2mm2.png" alt="2MM" /></td>
<td><img src="us2.png" alt="Us" /></td>
<td><img src="human2.png" alt="Human" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="canny3.png" alt="Canny" /></td>
<td><img src="2mm3.png" alt="2MM" /></td>
<td><img src="us3.png" alt="Us" /></td>
<td><img src="human3.png" alt="Human" /></td>
</tr>
<tr>
<td>Image</td>
<td>Canny</td>
<td>2MM</td>
<td>Us</td>
<td>Human</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-----</td>
<td>----</td>
<td>-------</td>
</tr>
<tr>
<td><img src="image1.jpg" alt="Image" /></td>
<td><img src="canny1.jpg" alt="Canny" /></td>
<td><img src="2mm1.jpg" alt="2MM" /></td>
<td><img src="us1.jpg" alt="Us" /></td>
<td><img src="human1.jpg" alt="Human" /></td>
</tr>
<tr>
<td><img src="image2.jpg" alt="Image" /></td>
<td><img src="canny2.jpg" alt="Canny" /></td>
<td><img src="2mm2.jpg" alt="2MM" /></td>
<td><img src="us2.jpg" alt="Us" /></td>
<td><img src="human2.jpg" alt="Human" /></td>
</tr>
</tbody>
</table>
Readings

• Puzzicha, Buhmann, Rubner, Tomasi
  • Empirical Evaluation of Dissimilarity measure for color and texture

• Rubner, Tomasi, Guibas
  • The earth moving distance as a metric for image retrieval

• Martin, Fowlkes, Malik
  • Learning to Detect Natural Image Boundaries using Brightness and Texture
Image Feature: Texture

(Plus-ell)

(Tri-arr)

(Ti-ell)
Schools of thoughts on texture

- Texture: repeated elements, subject to randomization of their location, size, color orientation
- Julesz: Nth-order joint empirical densities of textons
- Bergen-Adelson, Malik, Tuner: Multi-scale filter banks, wavelets
Jason Salavon:
100 Special Moments
http://salavon.com/work.php
Jason Salavon:
The Top Grossing Film of All Time
http://salavon.com/work.php

The worldwide top grossing film of all time, Titanic, was digitized from video in its entirety and broken up into its constituent frames. Each of these was then averaged to a single color best representative of that frame and reformatted as a photograph mirroring the narrative sequence of the film. Reading from left-to-right and top-to-bottom, the narrative's visual rhythm is laid out in pure color.
How the brain works (v1):
How the brain works (v1):
How the brain works (v1):

Diagram of a hypercolumn, which is a small region of visual cortex containing inputs from both eyes and all visual orientations, separated spatially, as shown.
Hubel and Wiesel Cat Experiment
Visual Cortex Cell Recording
Imitation Brain:

Base filter:

\[ F(x) = G(x, \sigma) \otimes G(x, 3\sigma) \]
Filters are dedicated to a range of orientation:

Base filter: \[ F(x) = G(x, \sigma) \otimes G(x, 3\sigma) \]

Apply rotation:

\[ R(\theta) = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \]
Bank of Filters

A step in the intensity corresponding to an edge

Maximum of the first order derivate

Zero crossing for the second order derivate. The sign changes around the edge.
Odd-symmetric filters:

\[ F(x) = G_x(x, \sigma) \otimes G(x, 3\sigma) \]

\[ \frac{\delta G(x, \sigma)}{\delta x} \]
Odd-symmetric filters:

\[ F(x) = G_{xx}(x, \sigma) \otimes G(x, 3\sigma) \]
The zero-crossings correspond to magnitude peaks.
Difference of Gaussian (DOG)

\[ Dog(x, y) = \frac{\sigma^2_{ex}}{2\pi \sigma^2_{ex}} e^{-\frac{x^2 + y^2}{2\sigma^2_{ex}}} - \frac{\sigma^2_{inh}}{2\pi \sigma^2_{inh}} e^{-\frac{x^2 + y^2}{2\sigma^2_{inh}}} \]
Simple Cell Receptive Field vs. Gabor Function

Solid line: Simple Cell Receptive Field.

Dashed line: Best fitting Gabor function.

Gabor Filter

Complex
\[ g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp \left( -\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \exp \left( i \left( 2\pi \frac{x'}{\lambda} + \psi \right) \right) \]

Real
\[ g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp \left( -\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \cos \left( 2\pi \frac{x'}{\lambda} + \psi \right) \]

Imaginary
\[ g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp \left( -\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2} \right) \sin \left( 2\pi \frac{x'}{\lambda} + \psi \right) \]

where
\[ x' = x \cos \theta + y \sin \theta \]
and
\[ y' = -x \sin \theta + y \cos \theta \]
Dennis Gabor
Fourier Transform of a Gabor Filter

The Fourier Transform of a Gabor filter is a localized set of spatial frequencies.

Gabor filters are band-pass filters. They are tuned to spatial frequency.
Fourier Transform of a Gabor Filter

\[ Gabor(u, \sigma) = G(\sigma)e^{-ux} \]

\[ FFT(Gabor(u, \sigma)) = FFT(G(\sigma)) \otimes FFT(e^{-ux}) \]

\[ = \frac{1}{\sigma}G(1/\sigma) \otimes \delta(u) \]
Spatial Scales

The spatial profile of the simple cell receptive field is predicted by taking the inverse Fourier transform of the contrast sensitivity function for that cell.
Spatial Frequency Columns

As with orientation and ocular dominance, spatial frequency shows columnar organization in the cortex.
2D Gabor filter
Laplacian pyramid is doing Band-pass filtering

What is a oriented band-pass filter?
Filter Banks
Dot filter response
Odd symmetric filter outputs
Even symmetric filter
Texture Processing Pipeline
Similarity using average filter bank response
Average filter bank response

squared responses

vertical

classification

horizontal

smoothed mean
Is mean of filter outputs sufficient?
Texture Patch Types

• Simple: clean step edge

• Textured: on either side, or step with noise

• Complex: wrong scale, or just a mess

• Invisible: boundary but no edge
Simple Patches
Textured Patches
Complex Patches