Visual Texture (in human and machine)
What is Texture?

• Texture depicts spatially repeating patterns
• Many natural phenomena are textures
Texture as “stuff”
Texture and Material

http://www-cvr.ai.uiuc.edu/ponce_grp/data/texture_database/samples/
Texture Analysis

Compare textures and decide if they’re made of the same “stuff”.

True (infinite) texture

ANALYSIS

“Same” or “different”
When are two textures similar?
Béla Julesz, father of texture
Texton Discrimination (Julesz)

Human vision is sensitive to the difference of some types of elements and appears to be “numb” on other types of differences.
Search Experiment I

The subject is told to detect a target element in a number of background elements. In this example, the detection time is independent of the number of background elements.
In this example, the detection time is proportional to the number of background elements, and thus suggests that the subject is doing element-by-element scrutiny.
Julesz then conjectured the following axiom:

Human vision operates in two distinct modes:

1. Preattentive vision
   parallel, instantaneous (~100--200ms), without scrutiny,
   independent of the number of patterns, covering a large visual field.

2. Attentive vision
   serial search by focal attention in 50ms steps limited to small aperture.
Examples

Pre-attentive vision is sensitive to size/width, orientation changes
Julesz Conjecture

Textures cannot be spontaneously discriminated if they have the same first-order and second-order statistics and differ only in their third-order or higher-order statistics. (not quite true)
1\textsuperscript{st} Order Statistics

5% white

20% white
2nd Order Statistics

10% white
Single Cell Recording
Single Cell Recording

Microelectrode

Electrical response
(action potentials)

mV

Time

© Stephen E. Palmer, 2002
Retinal Receptive Fields

Receptive field structure in ganglion cells: On-center Off-surround
Receptive field structure in **ganglion cells**: On-center Off-surround

Stimulus condition  Electrical response

Response | Time
Retinal Receptive Fields

Receptive field structure in *ganglion cells*: On-center Off-surround

Stimulus condition

Electrical response

Response | Time
Retinal Receptive Fields

Receptive field structure in ganglion cells: On-center Off-surrond

Stimulus condition

Electrical response

Response

Time

© Stephen E. Palmer, 2002
Retinal Receptive Fields

Receptive field structure in ganglion cells: On-center Off-surround

Stimulus condition

Electrical response

© Stephen E. Palmer, 2002
Receptive field structure in ganglion cells:
On-center Off-surround

Stimulus condition

Response

Time

Electrical response
Retinal Receptive Fields

RF of On-center Off-surround cells

Neural Response

- Center
- Surround

On  Off

Receptive Field

Response Profile

Firing Rate

- on-center
- off-surround

Horizontal Position

© Stephen E. Palmer, 2002
RF of Off-center On-surround cells

Retinal Receptive Fields

Neural Response

Surround

Center

On

Off

Receptive Field

Response Profile

Firing Rate

on-surround

off-center

Horizontal Position

© Stephen E. Palmer, 2002
Retinal Receptive Fields

- Horizontal cell
- Bipolar cells
- Retinal ganglion cells
- Amacrine cell
- Dendro-dendritic synapse
- Receptors
Receptive field structure in bipolar cells

A. WIRING DIAGRAM

B. RECEPTIVE FIELD PROFILES

Direct excitatory component (D)

Indirect inhibitory component (I)

D + I
The receptive field of a retinal ganglion cell can be modeled as a “Difference of Gaussians”

\[ G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \]
Figure 6.16 Receptive fields
The receptive field of a receptor is simply the area of the visual field from which light strikes that receptor. For any other cell in the visual system, the receptive field is determined by which receptors connect to the cell in question.
Anatomy of Pathway to Visual Cortex
Visual Cortex

Cortical Area V1

aka:
Primary visual cortex
Striate cortex
Brodman’s area 17

Diagram:
- Thalamus
- Optic nerve
- Eye
- Temporal visual cortex
- Ventral Stream
- Dorsal Stream
- Parietal visual cortex
- Striate cortex (V1)
- Extrastriate cortex

© Stephen E. Palmer, 2002
Cortical Receptive Fields

Single-cell recording from visual cortex

David Hubel & Thorston Wiesel
Cortical Receptive Fields

Single-cell recording from visual cortex

© Stephen E. Palmer, 2002
Cortical Receptive Fields

Simple Cells: “Line Detectors”

A. Light Line Detector

B. Dark Line Detector

Firing Rate

Horizontal Position

© Stephen E. Palmer, 2002
Cortical Receptive Fields

Simple Cells: “Edge Detectors”

C. Dark-to-light Edge Detector

D. Light-to-dark Edge Detector

Firing Rate

Horizontal Position

© Stephen E. Palmer, 2002
Constructing a line detector
The 1D Gaussian and its derivatives

\[ G_\sigma(x) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{x^2}{2\sigma^2}} \]

\[ G'_\sigma(x) = \frac{d}{dx} G_\sigma(x) = -\frac{1}{\sigma} \left( \frac{x}{\sigma} \right) G_\sigma(x) \]

\[ G''_\sigma(x) = \frac{d^2}{dx^2} G_\sigma(x) = \frac{1}{\sigma^2} \left( \frac{x^2}{\sigma^2} - 1 \right) G_\sigma(x) \]

\( G'_\sigma(x) \)'s maxima/minima occur at \( G''_\sigma(x) \)'s zeros. And, we can see that \( G'_\sigma(x) \) is an odd symmetric function and \( G''_\sigma(x) \) is an even symmetric function.
Oriented Gaussian Derivatives in 2D

\begin{align}
  f_1(x, y) &= G'_{\sigma_1}(x) G_{\sigma_2}(y) \\  f_2(x, y) &= G''_{\sigma_1}(x) G_{\sigma_2}(y)
\end{align}

We also consider rotated versions of these Gaussian derivative functions.

\begin{align}
  \text{Rot}_\theta f_1 &= G''_{\sigma_1}(u) G_{\sigma_2}(v) \\  \text{Rot}_\theta f_2 &= G'''_{\sigma_1}(u) G_{\sigma_2}(v)
\end{align}

where we set

\[
\begin{pmatrix}
  u \\
  v
\end{pmatrix} = \begin{pmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
  x \\
  y
\end{pmatrix}
\]

These are useful when we convolve with 2D images, e.g. to detect edges at different orientations.
Oriented Gaussian First and Second Derivatives
Hypercolumns in visual cortex

Model of Striate Module in Monkeys
Modeling hypercolumns

- Elongated directional Gaussian derivatives
- Gabor filters could be used instead
- Multiple orientations, scales
Overcomplete representation: filter banks

LM Filter Bank

Code for filter banks: www.robots.ox.ac.uk/~vgg/research/texclass/filters.html
Kristen Grauman
How can we represent texture?

• Idea 1: Record simple statistics (e.g., mean, std.) of absolute filter responses
Can you match the texture to the response?

Filters

Mean abs responses

A

B

C
How can we represent texture?

• Can be thought of as an single “orientation histogram”

• **Idea 2:** Marginal histograms of filter responses
  – one histogram per filter
Multi-scale filter decomposition

Filter bank

Input image
Filter response histograms
Start with a noise image as output

Main loop:

- Match pixel histogram of output image to input
- Decompose input and output images using multi-scale filter bank (Steerable Pyramid)
- Match subband histograms of input and output pyramids
- Reconstruct input and output images (collapse the pyramids)
Figure 7: (Left pair) Inhomogeneous input texture produces blotchy synthetic texture. (Right pair) Homogenous input.

Figure 8: Examples of failures: wood grain and red coral.

Figure 9: More failures: hay and marble.
Marginal statistics are not enough

Neighboring filter responses are highly correlated
  • an edge at low-res will cause an edge at high-res

Let’s match 2\textsuperscript{nd} order statistics too!

Match joint histograms of pairs of filter responses at adjacent spatial locations, orientations, and scales.

Optimize using repeated projections onto statistical constraint surfaces.
How can we represent texture?

- Marginal filter response histograms don’t talk to each other (in a direct way)
- **Idea 3**: Histograms of joint responses (textons)
We can form a feature vector from the list of responses at each pixel.

[r1, r2, ..., r38]
What does it capture?

\[ v = F \ast \text{Patch} \quad (\text{where } F \text{ is filter matrix}) \]

**Fig. 3.** Image reconstruction. Two example image patches (left), were reconstructed (right) from spatial filter responses at their center. Original image patches masked by a Gaussian (middle) are shown for comparison.
Textons (Malik et al, IJCV 2001)

- Cluster vectors of filter responses
Textons (cont.)
Object \rightarrow \text{Bag of ‘words’}
Of all the sensory impressions proceeding to the brain, the visual experiences are the dominant ones. Our perception of the world around us is based essentially on the messages that reach our brain from our eyes. For a long time it was thought that the retinal image was transmitted point by point to visual centers in the brain; the cerebral cortex was a movie screen, so to speak, upon which the image in the eye was projected. Through the discoveries of Hubel and Wiesel we now know that behind the origin of the visual perception in the brain there is a considerably more complicated course of events. By following the visual impulses along their path to the various cell layers of the optical cortex, Hubel and Wiesel have been able to demonstrate that the message about the image falling on the retina undergoes step-wise analysis in a system of nerve cells stored in columns. In this system each cell has its specific function and is responsible for a specific detail in the pattern of the retinal image.

China is forecasting a trade surplus of $90bn (£51bn) to $100bn this year, a threefold increase on 2004's $32bn. The Commerce Ministry said the surplus would be created by a predicted 30% jump in exports to $750bn, compared with a 18% rise in imports to $660bn. This is likely to annoy the US, which has long argued that China's exports are unfairly helped by a deliberately undervalued yuan. Beijing agrees the surplus is too high, but says the yuan is only one factor. Bank of China governor Zhou Xiaochuan said the country also needed to do more to boost domestic demand so more goods stayed within the country. China increased the value of the yuan against the dollar by 2.1% in July and permitted it to trade within a narrow band, but the US wants the yuan to be allowed to trade freely. However, Beijing has made it clear that it will take its time and tread carefully before allowing the yuan to rise further in value.
Patch Features
dictionary formation
Clustering (usually k-means)
Image representation

frequency
codewords
Scene Classification (Renninger & Malik)

beach

mountain

forest

city

street

farm

kitchen

livingroom

bedroom

bathroom
Texton Histogram Matching
Object Detection can be very fast

• On a task of judging animal vs no animal, humans can make mostly correct saccades in 150 ms (Kirchner & Thorpe, 2006)
  – Comparable to synaptic delay in the retina, LGN, V1, V2, V4, IT pathway.
  – Doesn’t rule out feed back but shows feed forward only is very powerful

• Detection and categorization are practically simultaneous (Grill-Spector & Kanwisher, 2005)
Discrimination of Basic Categories

% correct

texture model

street  bedroom  mountain  farm  bathroom  city  beach  kitchen  forest  livingroom
Discrimination of Basic Categories

% correct

texture model

chance

street
bedroom
mountain
farm
bathroom

city
beach
kitchen
forest
livingroom
Discrimination of Basic Categories

% correct

street  bedroom  mountain  farm  bathroom  city  beach  kitchen  forest  livingroom

texture model  37 ms

chance
Discrimination of Basic Categories

% correct

street  bedroom  mountain  farm  bathroom  city  beach  kitchen  forest  livingroom

* texture model

50 ms

chance
Discrimination of Basic Categories

- texture model
- 69 ms

- street
- bedroom
- mountain
- farm
- bathroom
- city
- beach
- kitchen
- forest
- livingroom

% correct

chance
Discrimination of Basic Categories

% correct

- texture model
- 37 ms
- 50 ms
- 69 ms
Scene Recognition using Texture
Why these filters?

Wavelet-like receptive fields emerge from a network that learns sparse codes for natural images.

Bruno A. Olshausen\textsuperscript{1} and David J. Field

\begin{equation}
E = -[\text{preserve information}] - \lambda [\text{sparseness of } a_i],
\end{equation}

where $\lambda$ is a positive constant that determines the importance of the second term relative to the first. The first term measures how well the code describes the image, and we choose this to be the mean square of the error between the actual image and the reconstructed image:

\begin{equation}
[preserve \text{ information}] = -\sum_{x,y} \left[ I(x, y) - \sum_i a_i \phi_i(x, y) \right]^2. 
\end{equation}
Learned filters

a.