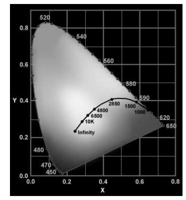
Color and Texture

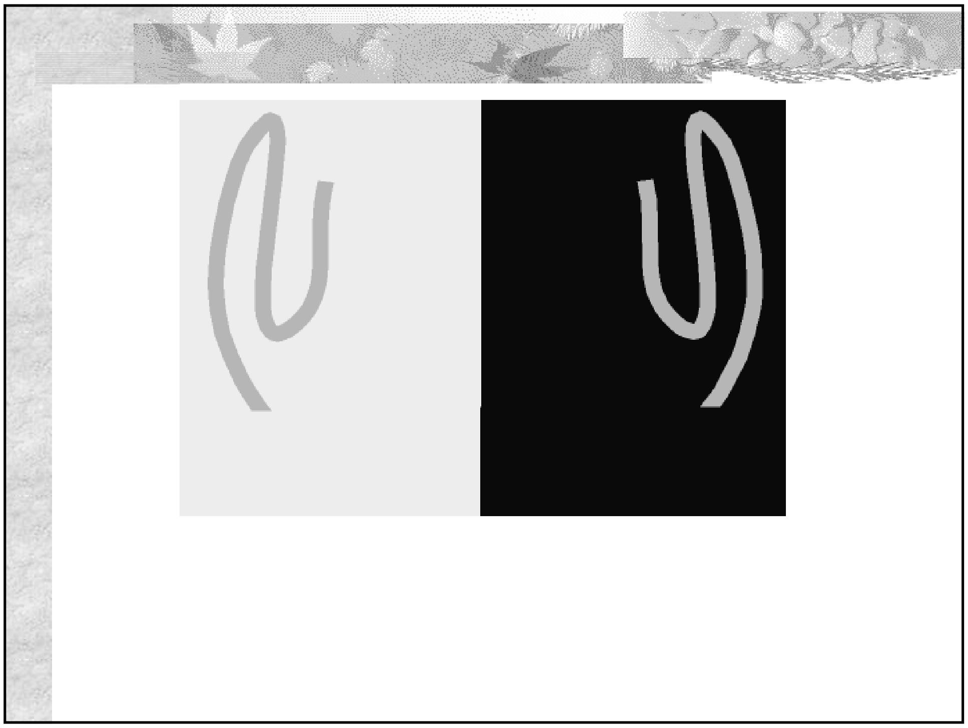


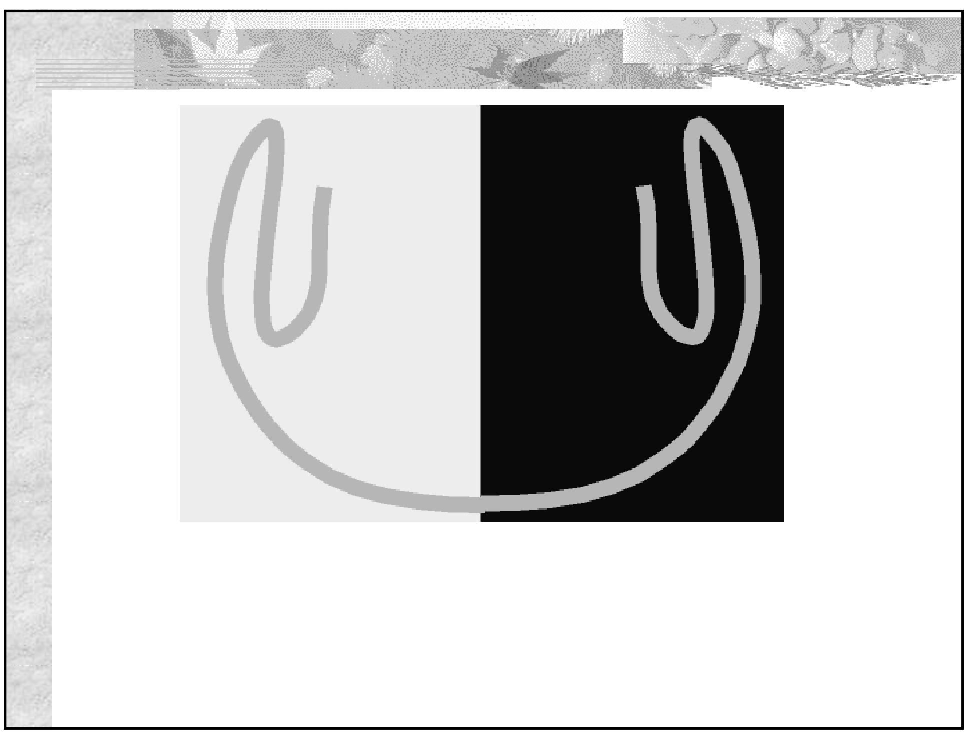
The Appearance of Colors

- Color appearance is strongly affected by (at least):
  - other nearby colors,
  - adaptation to previous views
  - “state of mind”
- We show several demonstrations in what follows.

The Appearance of Colors

- Hering, Helmholtz: Color appearance is strongly affected by other nearby colors, by adaptation to previous views, and by "state of mind"
- Film color mode:
  View a colored surface through a hole in a sheet, so that the colour looks like a film in space; controls for nearby colors, and state of mind.
  - Other modes:
    - Surface colour
    - Volume colour
    - Mirror colour
    - Illuminant colour
- By experience, it is possible to match almost all colors, viewed in film mode using only three primary sources - the principle of **trichromacy**.
  - Other modes may have more dimensions
    - Glossy-matte
    - Rough-smooth
- Most of what follows discusses film mode.

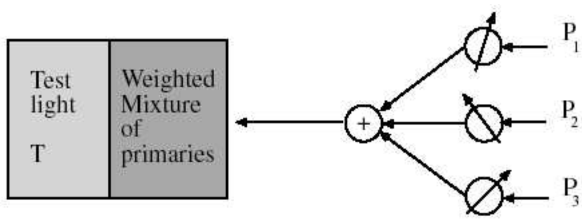




Why specify color numerically?

- Accurate color reproduction is commercially valuable
- Few color names are widely recognized by English speakers -
  - About 10; other languages have fewer/more, but not many more.
  - It's common to disagree on appropriate color names.
- Color reproduction problems increased by prevalence of digital imaging - eg. digital libraries of art.
  - How do we ensure that everyone sees the same color?

Color Matching Experiments - I



- Show a split field to subjects; one side shows the light whose color one wants to measure, the other a weighted mixture of primaries (fixed lights).
- Each light is seen in film color mode.

Color matching experiments - II

- Many colors can be represented as a mixture of A, B, C
- write

  $$M = a\,A + b\,B + c\,C$$

  where the = sign should be read as “matches”
- This is **additive** matching.
- Gives a color description system - two people who agree on A, B, C need only supply (a, b, c) to describe a color.

Subtractive matching

- Some colors can’t be matched like this: instead, must write

  $$M + a\,A = b\,B + c\,C$$

- This is **subtractive** matching.
- Interpret this as (-a, b, c)
- Problem for building monitors: Choose R, G, B such that positive linear combinations match a large set of colors

The principle of trichromacy

- Experimental facts:
  - Three primaries will work for most people if we allow subtractive matching
    - Exceptional people can match with two or only one primary.
    - This could be caused by a variety of deficiencies.
  - Most people make the same matches.
    - There are some anomalous trichromats, who use three primaries but make different combinations to match.

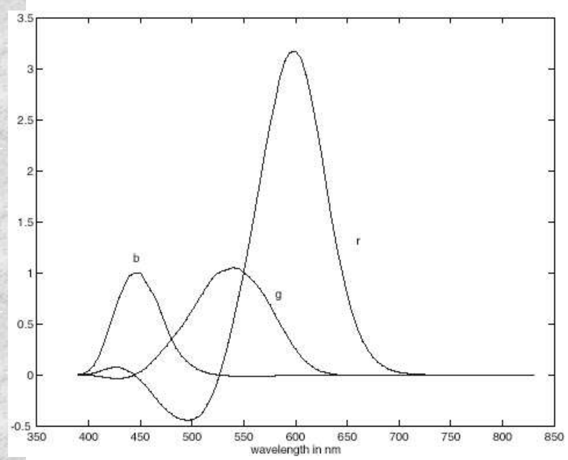
Grassman's Laws

- For colour matches made in film colour mode:
  - symmetry: U=V <=>V=U
  - transitivity: U=V and V=W => U=W
  - proportionality: U=V <=> tU=tV
  - additivity: if any two (or more) of the statements
    U=V,
    W=X,
    (U+W)=(V+X) are true, then so is the third
- These statements are as true as any biological law. They mean that color matching in film color mode is linear.

Linear color spaces

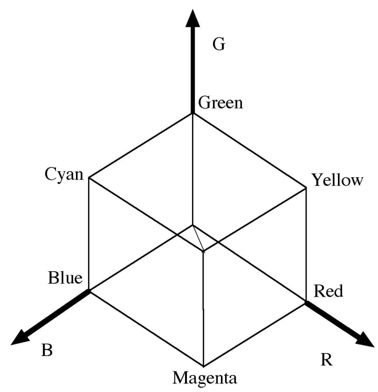
- A choice of primaries yields a linear color space --- the coordinates of a color are given by the weights of the primaries used to match it.
- Choice of primaries is equivalent to choice of color space.

RGB Color Space

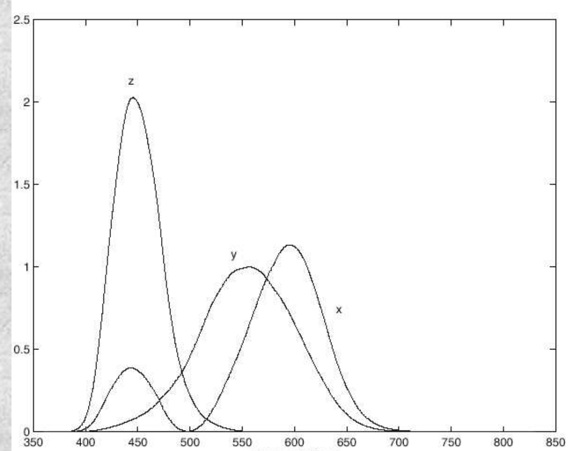


RGB: primaries are monochromatic, energies are 645.2nm, 526.3nm, 444.4nm. Color matching functions have negative parts -> some colors can be matched only subtractively.

RGB Color Space



CIE XYZ Color Space



CIE XYZ: Color matching functions are positive everywhere, but primaries are imaginary. Usually draw x, y, where
x=X/(X+Y+Z)
y=Y/(X+Y+Z)

CIE XYZ Color Space

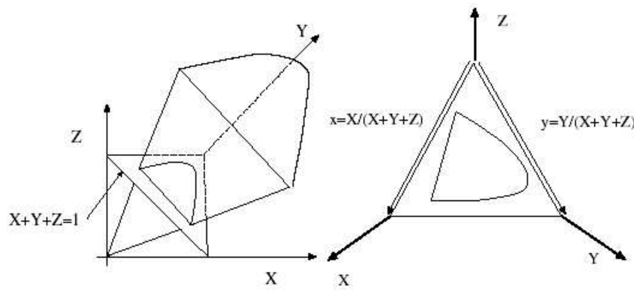
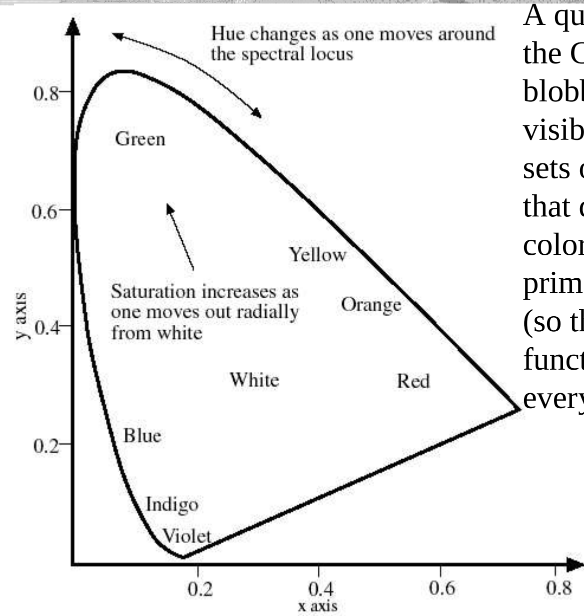
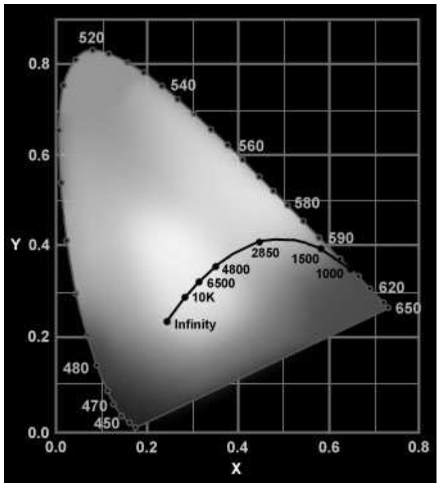


Figure 4.7. The volume of all visible colours in CIE XYZ coordinate space is a cone whose vertex is at the origin. Usually, it is easier to suppress the brightness of a colour — which we can do because to a good approximation perception of colour is linear — and we do this by intersecting the cone with the plane $X + Y + Z = 1$ to get the CIE xy space



A qualitative rendering of the CIE (x,y) space. The blobby region represents visible colors. There are sets of (x, y) coordinates that don't represent real colors, because the primaries are not real lights (so that the color matching functions could be positive everywhere).

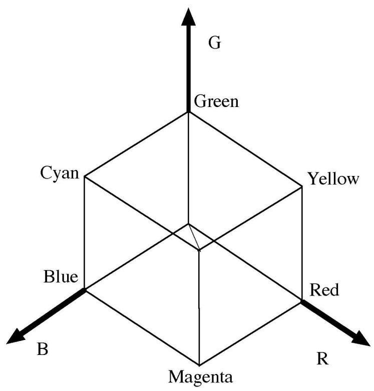


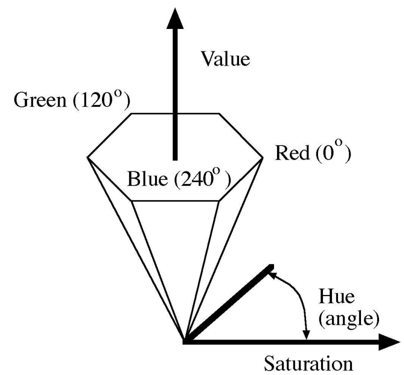
A qualitative rendering of the CIE (x,y) space. The blobby region represents visible colors. There are sets of (x, y) coordinates that don't represent real colors, because the primaries are not real lights (so that the color matching functions could be positive everywhere).

Non-linear color spaces

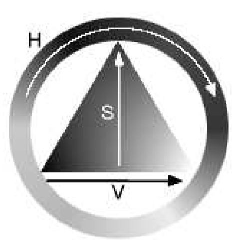
- HSV: Hue, Saturation, Value are non-linear functions of XYZ.
  - because hue relations are naturally expressed in a circle
- Uniform: equal (small!) steps give the same perceived color changes.

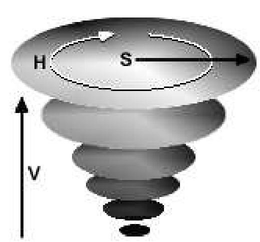
HSV hexcone





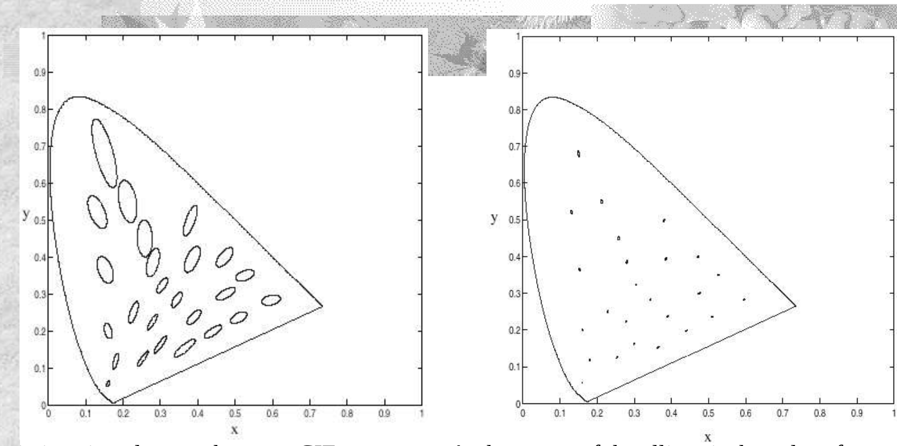
HSV Color Space





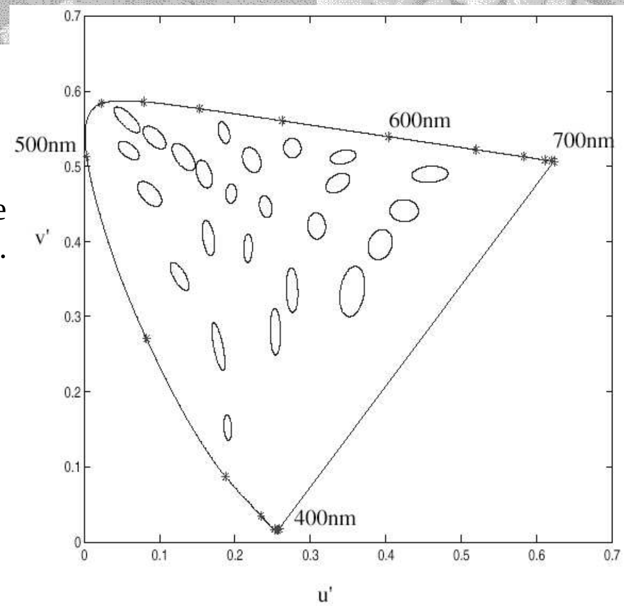
Uniform color spaces

- McAdam ellipses (next slide) demonstrate that differences in x,y are a poor guide to differences in color
- Construct color spaces so that differences in coordinates are a good guide to differences in color.



Variations in color matches on a CIE x, y space. At the center of the ellipse is the color of a test light; the size of the ellipse represents the scatter of lights that the human observers tested would match to the test color; the boundary shows where the just noticeable difference is. The ellipses on the left have been magnified 10x for clarity; on the right they are plotted to scale. The ellipses are known as MacAdam ellipses after their inventor. The ellipses at the top are larger than those at the bottom of the figure, and that they rotate as they move up. This means that the magnitude of the difference in x, y coordinates is a poor guide to the difference in color.

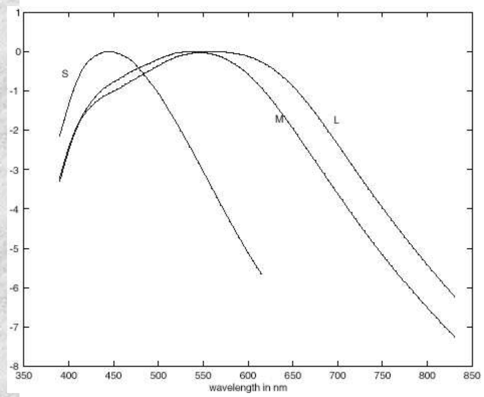
CIE u'v' which is a projective transform of x, y. We transform x,y so that ellipses are most like one another. Figure shows the transformed ellipses.



Color receptors and color deficiency

- Trichromacy is justified - in color normal people, there are three types of color receptor, called **cones**, which vary in their sensitivity to light at different wavelengths (shown by molecular biologists).
- Deficiency can be caused by CNS, by optical problems in the eye, or by absent receptor types
  - Usually a result of absent genes.
- Some people have fewer than three types of receptor; most common deficiency is red-green color blindness in men.
- Color deficiency is less common in women; red and green receptor genes are carried on the X chromosome, and these are the ones that typically go wrong. Women need two bad X chromosomes to have a deficiency, and this is less likely.

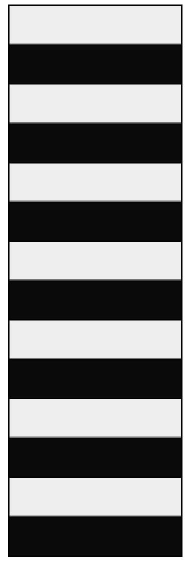
Color receptors

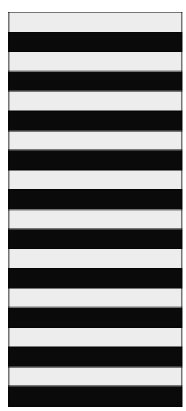


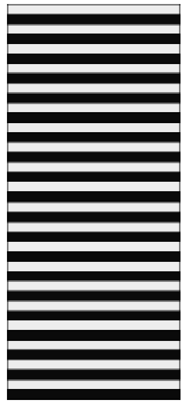
- Plot shows relative sensitivity as a function of wavelength, for the three cones. The S (for short) cone responds most strongly at short wavelengths; the M (for medium) at medium wavelengths and the L (for long) at long wavelengths.
- These are occasionally called B, G and R cones respectively, but that's misleading - you don't see red because your R cone is activated.

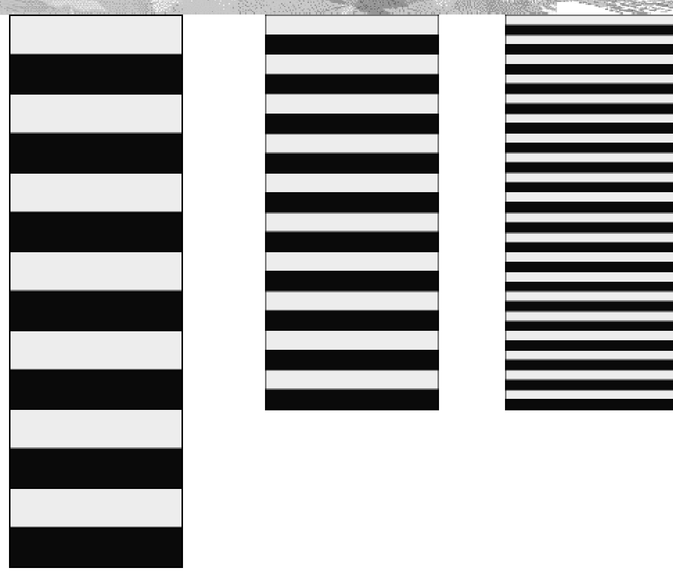
Adaptation phenomena

- The response of your color system depends both on spatial contrast and what it has seen before (adaptation)
- This seems to be a result of coding constraints --- receptors appear to have an operating point that varies slowly over time, and to signal some sort of offset. One form of adaptation involves changing this operating point.
- Common example: walk inside from a bright day; everything looks dark for a bit, then takes its conventional brightness.





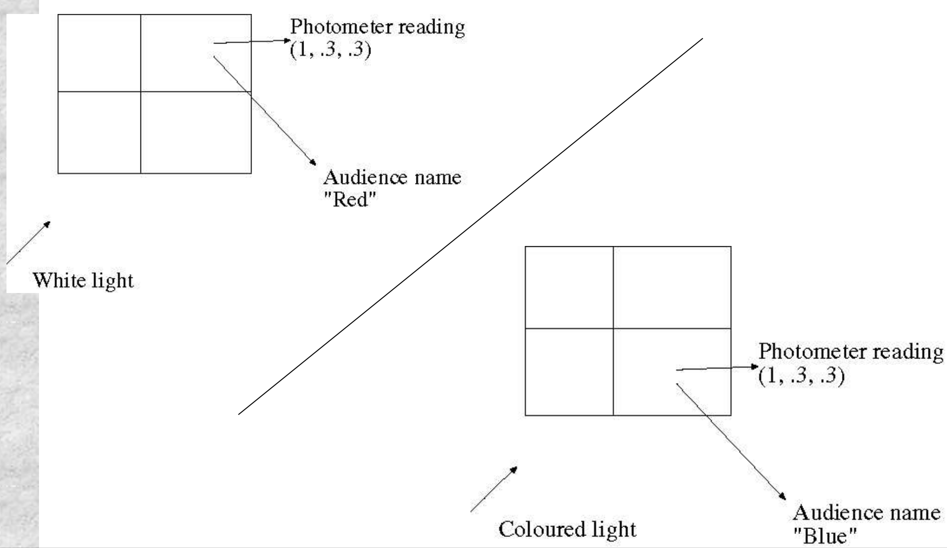




Color Constancy

- Color of reflected light depends on both illuminant and surface
- People are surprisingly good at disentangling these effects in practice (color constancy)
- This is probably where some of the spatial phenomena in colour perception come from

Land's Demonstration



Lightness Constancy

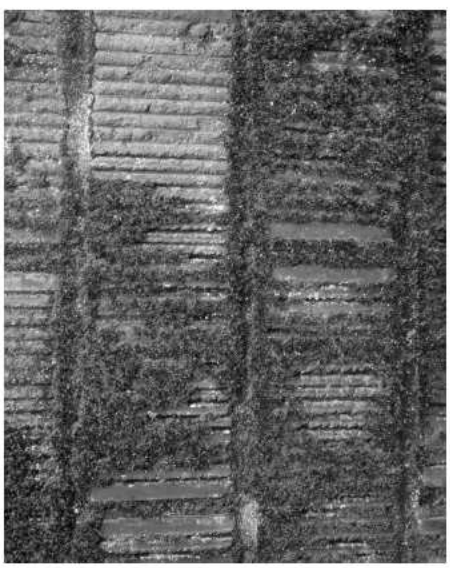
- Lightness constancy
  - how light is the surface, independent of the brightness of the illuminant
  - issues
    - spatial variation in illumination
    - absolute standard
  - Human lightness constancy is very good
- Assume
  - frontal 1D "Surface"
  - slowly varying illumination
  - quickly varying surface reflectance

Suggested Reading

- Chapter 6, David A. Forsyth and Jean Ponce, "Computer Vision: A Modern Approach"

Texture

- Key issue: representing texture
  - Texture based matching
    - little is known
  - Texture segmentation
    - key issue: representing texture
  - Texture synthesis
    - useful; also gives some insight into quality of representation
  - Shape from texture



Representing textures

- Textures are made up of quite stylised subelements, repeated in meaningful ways
- Representation:
  - find the subelements, and represent their statistics
- But what are the subelements, and how do we find them?
  - recall normalized correlation
  - find subelements by applying filters, looking at the magnitude of the response
- What filters?
  - experience suggests spots and oriented bars at a variety of different scales
  - details probably don't matter
- What statistics?
  - within reason, the more the merrier.
  - At least, mean and standard deviation
  - better, various conditional histograms.

Texture Measures

- GLCM (Gray level Co-occurrence Matrices)
- Law's Texture Energy Measures
- Wavelets
- Steerable Pyramids

GLCMs

- 2D histogram of image intensities
- P(i,j,d, θ): Count of occurrence of gray level i with j at distance d and in direction θ

GLCMs

50	51	52	50
53	51	51	52
51	50	51	52
52	53	53	52

Intensity image patch

	50	51	52	53
50	0	2	0	0
51	1	1	3	0
52	1	0	0	1
53	0	1	1	1

P(d,θ), d=1, θ=0^o

GLCMs

- Too many parameters
- Computationally Expensive
- Not suitable for coarse texture
- Susceptible to noise

Law's Texture Energy Measures

- Feature Extraction scheme based on gradient operators
- 25 Masks by convolution of 5 1-D vectors
  - Level L5 = [1 4 6 4 1]
  - Edge E5 = [-1 -2 0 2 1]
  - Spot S5 = [-1 0 2 0 -1]
  - Wave W5= [-1 2 0 -2 1]
  - Ripple R5= [1 -4 6 -4 1]

Law’s Texture Energy Measures

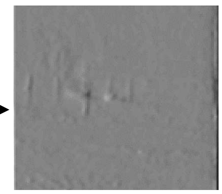
$$(L5^t)E5 = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -4 & -8 & 0 & 8 & 4 \\ -6 & -12 & 0 & 12 & 6 \\ -4 & -8 & 0 & 8 & 4 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

$$L5 = [1 \quad 4 \quad 6 \quad 4 \quad 1]$$

$$E5 = [-1 \quad -2 \quad 0 \quad 2 \quad 1]$$

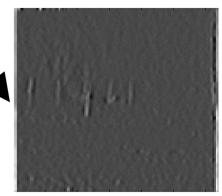
Input Image

$(L5^t)E5$



Output Image

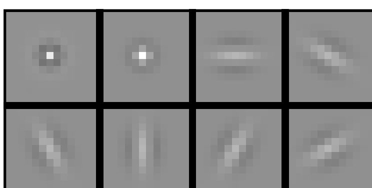
$(L5^t)S5$

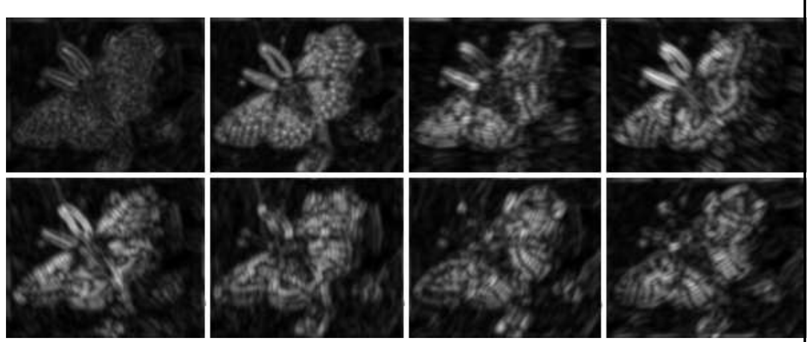


Output Image

$$(L5^t)S5 = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix}$$

$$S5 = [-1 \quad 0 \quad 2 \quad 0 \quad -1]$$





Gabor Filters of different orientations

Wavelet Analysis

- Tool for multi-resolution analysis
- Provides localization in both spatial and frequency domain
- Every decomposition contains information of a specified scale and orientation

Wavelet Analysis

- Wavelet transform decomposes $f(x)$ onto a basis of wavelet functions:

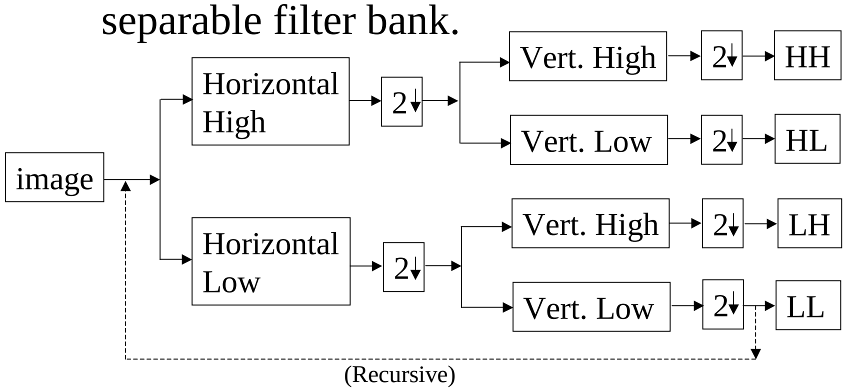
$$(W_a f)(b) = \int f(x)\varphi_{a,b}(x)dx$$

where

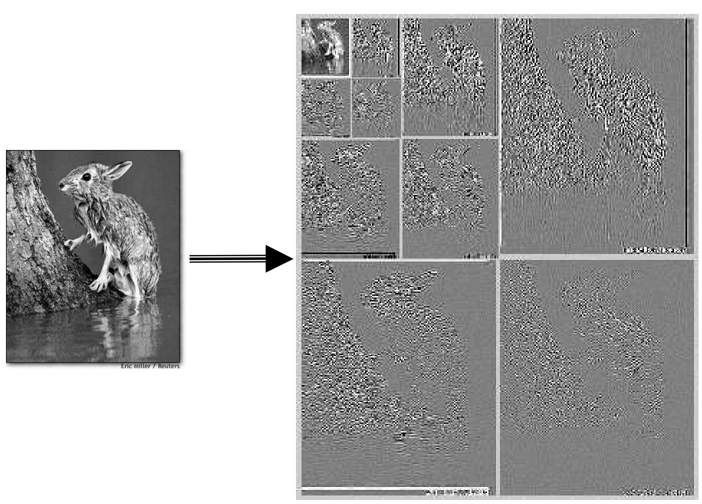
$$\varphi_{a,b}(x) = \frac{1}{\sqrt{a}}\varphi\left(\frac{x-b}{a}\right)$$

Wavelet Analysis

- 2-D Wavelet decomposition is obtained by separable filter bank.



Wavelet Analysis



Wavelet Analysis

- Desirable Properties
  - Symmetry
  - Compactness
  - Small Support
  - Shift Invariance

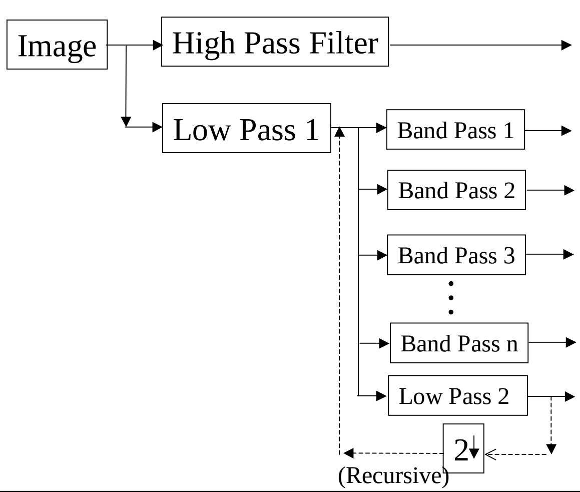
Wavelet Analysis

- Families
  - Haar
  - Debauchies (Daub)
  - Quadrature Mirror Filters (QMF)
  - Biorthogonal Filters (Biort)
  - etc

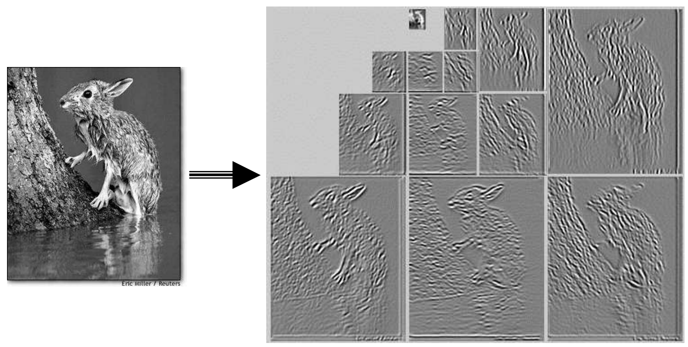
Steerable Pyramids

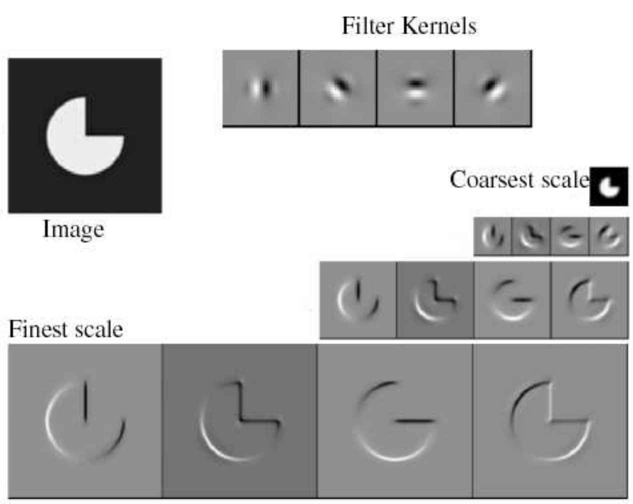
- Linear multi-scale, multi-orientation image decomposition
- Basis functions are directional derivative operators in different sizes and orientations
- Type of over-complete wavelet transform
- Steerable orientation decomposition

Steerable Pyramids



Steerable Pyramids





Reprinted from “Shiftable MultiScale Transforms,” by Simoncelli et al., IEEE Transactions on Information Theory, 1992, copyright 1992, IEEE

Steerable Pyramids vs. Wavelets

- Translation invariance
- Rotation invariance
- Aliasing
- Orthogonality
- Completeness
- Computational Efficiency

Features:

- Energy

$$e_i = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} I_i^2(x, y)$$

- Entropy

$$Entropy_i = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} I_i(x, y) \log I_i(x, y)$$

Features

- Kurtosis

$$k = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} [I_i(x, y) - \mu]^4}{M \times N \times \sigma^4} - 3$$

- Skew

$$Skew = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} [I_i(x, y) - \mu]^3}{M \times N \times \sigma^3}$$

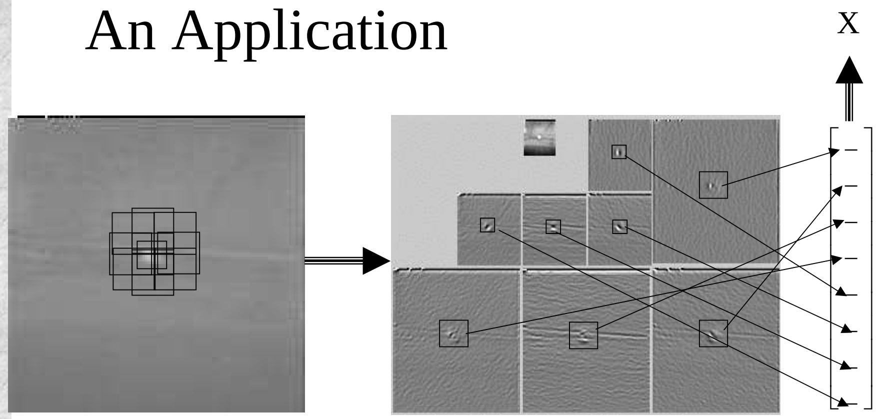
- Variance

$$\sigma^2 = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} [I_i(x, y) - \mu]^2}{M \times N}$$

Final texture representation

- Form an oriented pyramid (or equivalent set of responses to filters at different scales and orientations).
- Square the output
- Take statistics of responses
  - e.g. mean of each filter output (are there lots of spots)
  - std of each filter output
  - mean of one scale conditioned on other scale having a particular range of values (e.g. are the spots in straight rows?)

An Application



$X^{(1)}\ X^{(2)}\ X^{(3)}\ X^{(4)}\ X^{(5)}\ X^{(6)}\ X^{(7)}\ X^{(8)}$

Suggested Reading

- Chapter 9, David A. Forsyth and Jean Ponce, "Computer Vision: A Modern Approach"