Multidimensional Signal Processing
Image processing fundamentals

- Ms. Vidya Manian
- Dept. of Electrical & Computer Engineering
- University of Puerto Rico at Mayaguez

- http://www.ece.uprm.edu/~manian
- Email: manian@ece.uprm.edu

- Oct. 2003

Topics

- Image representation
- Image Processing Tools
- Algorithms for image processing
  - Image enhancement – contrast stretching & filtering
  - Zero-crossing filters -Edge detection
- Color image processing
- Classification techniques
  - Texture based
  - Nonparametric classifiers

Image representation

- Image – function of 2 variables-a(x,y)
- Digitization – from 2D continuous space to 2D discrete space a[m,n]
- Value = a(x,y,z,λ,t)
- Z-depth, λ-color, t-time

Image processing tools

- Convolution
- Image transforms – Fourier transform
- Discrete cosine transform
- Properties – real and orthogonal, fast transform
- Excellent energy compaction properties for highly correlated images (JPG compression)

Image Processing

- Transforms - FFT
- The 2D Fourier transform is used to obtain the Fourier power spectrum
- The Fourier spectrum is an intensity function where brightness is proportional to the amplitude of |F(u,v)|

$$F\{f(x,y)\} = F(u,v) = \iint_{-\infty}^{\infty} f(x,y)\exp[-j2p(ux+vy)]dxdy$$

Fourier Transform

- The spectrum is

$$|F(u,v)| = AXY\left|\frac{\sin(puX)}{(puX)}\right|\left|\frac{\sin(pvY)}{(pvY)}\right|$$

- The FT has useful properties such as separability
- Due to separability a 2D transform can be done as two 1D transforms
- Properties - translation, periodicity and conjugate symmetry, rotation, distributivity and scaling

Image Transforms

- The Walsh, Hadamard, Fourier, Haar, Slant, Cosine transforms are orthogonal transforms
- They can be inverted and the kernels are separable
- Just like Fast Fourier Transform (FFT) there are Fast Walsh Transform (FWT) and Fast Hadamard Transform (FHT)

Image Transforms

- The Hotelling or KL Transform is based on statistical properties of vector representations
- The KLT is given by $y=A(x-m_x)$ where the input vectors x are transformed into the y vectors using the transformation matrix A
- m_x is the mean of the input vectors
- A is a matrix whose rows are formed from the eigenvectors of C_x ordered so that the first row of A is the eigenvector corresponding to the largest eigenvalue and the last row is the eigenvector corresponding to the smallest eigenvalue, C_x is the covariance matrix of x

Image Transforms

- Important property of Hotelling transform is that x can be reconstructed from y
- As the rows of A are orthonormal vectors, $A^{-1}=A^T$ and $x = A^T y+m_x$
- Instead of using all eigenvectors of C_x we form matrix A_K from K eigenvectors with K largest eigenvalues we have

$$\hat{x} = A_k^T y + m_x$$

- The transform minimizes the mean square error between the vectors x and their approximation $\hat{x}$
- Useful for processing remote sensing images - six spectral channels can be converted into 1 or 2 components with largest eigenvalues

Algorithms for image processing

- Histogram based operations
- Mathematics based operations
- Arithmetic based operations
- Convolution based operations
- Smoothing operations
- Derivative based operations
- Morphology based operations

Histogram Processing

- The histogram of a digital image with gray levels in the range [0,L-1] is a discrete function $p(r_k)=n_k/n$, where r_k is the kth gray level, n_k is the number of pixels in the image with that gray level, n is the total number of pixels in the image and k=0,1,..L-1
- For histogram equalization a transformation function equal to the cumulative distribution of the gray level - r is used which produces an image whose gray levels have uniform density
- In terms of enhancement, this result implies an increase in the dynamic range of the pixels and produces an increase in image contrast

Histogram equalized image

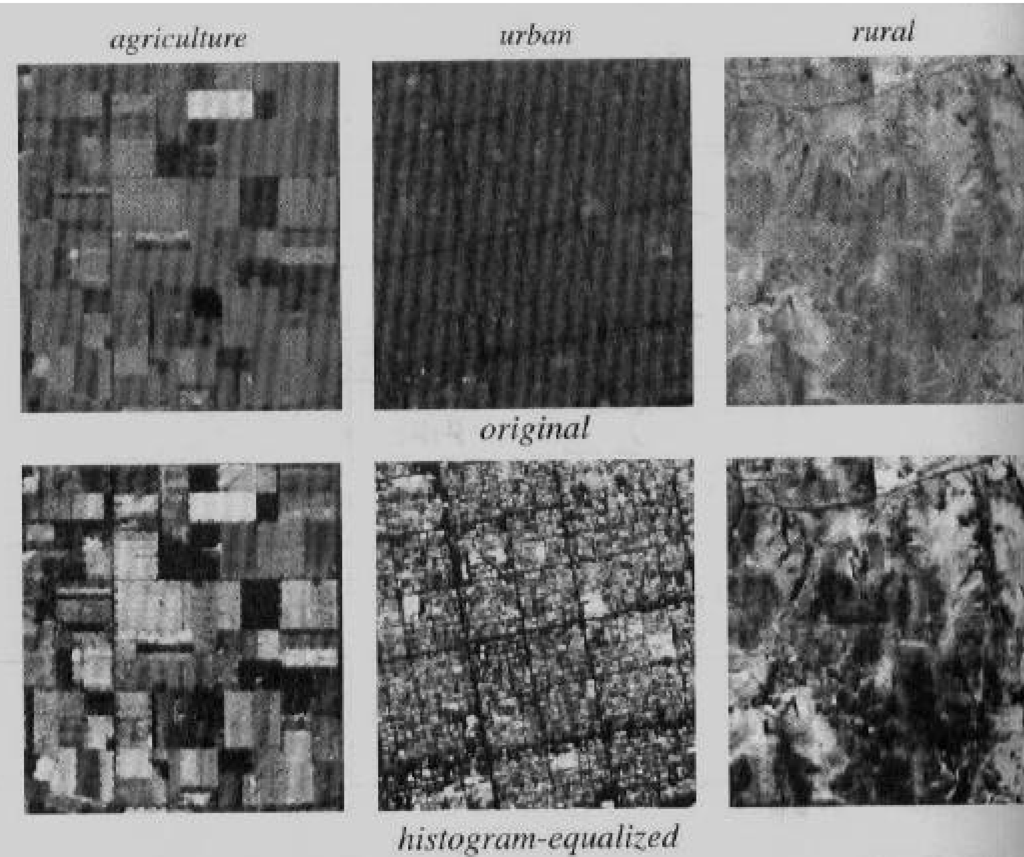


FIGURE 4-28. Three sample TM images used for co-occurrence matrix analysis. The histogram equalization process produces an image with an approximately equal pixel count-per-DN density in all parts of the DN range. It is described in Chapter 6.

Mathematics based operations

- *On binary images (0 – black, 1 – white)*
- *NOT* $c = \sim a$
- *OR* $c = a + b$
- *AND* $c = a * b$
- *XOR* $c = a * \sim b + \sim a * b$
- *SUB* $c = a * \sim b$
- *Point operations, SUB operation useful for image analysis, in region of interest finding*

Arithmetic based operations

Operation	definition	preferred data type
ADD	c=a+b	integer
SUB	c=a-b	integer
MUL	c=a*b	integer or floating point
DIV	c=a/b	floating point
LOG	c=log(a)	floating point
EXP	c=exp(a)	floating point
SQRT	c=sqrt(a)	floating point
TRIG	c=sin/cos/tan(a)	floating point
INVERT	$c=(2^B-1)-a$	integer

Morphology based operations

- *Study of forms-topology or structure of objects from their images*
- *Certain operations where an object is hit with a structuring element and thereby reduced to a more revealing shape*
- *Basic operations: Object X, structuring element-B, Bx – translation of B with origin located at x*
- *Erosion –shrinking : set of all points x such that Bx is included in X*

XQB = {x:Bx Ì X}

- *Dilation-expansion : set of all points x such that Bx hits X, they have a nonempty intersection*

XÅB={x:BxÇX¹ f}

Image Enhancement

- *Histogram equalization* can be done globally or locally
- Histogram equalization produces an approximation to the uniform density for the histogram
- By histogram specification we can specify any other density to the histogram such as Gaussian density
- Min-max method
- Normalization
- Threshold
- Reference

Image Enhancement

- <u>Spatial Domain</u>

◆ Image negatives are useful for displaying medical images and photographing a screen with monochrome positive film

◆ Low contrast images are due to poor illumination, lack of dynamic range in the imaging sensor

◆ <u>Contrast stretching</u> corrects the above by producing various degrees of spread in the gray levels of the output image

◆ Thresholding to 0s and 1s results in a binary image

Some type of DN-to-GL transformations for contrast enhancement

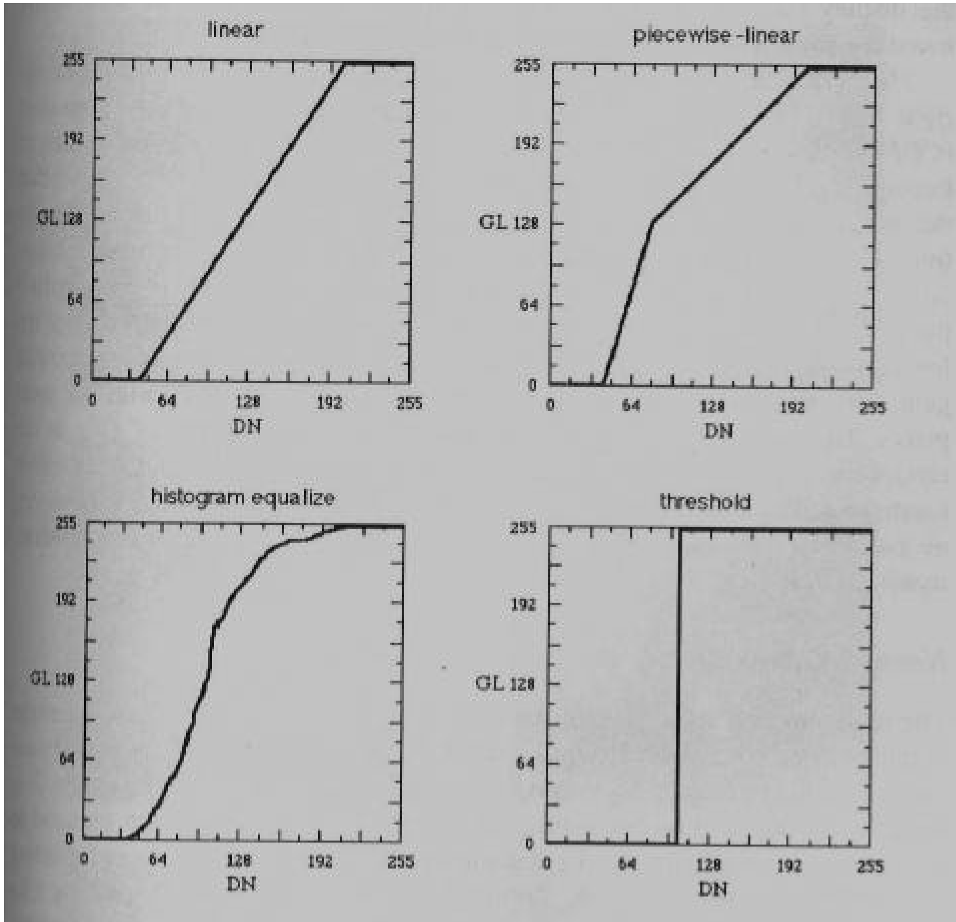


FIGURE 5-19. Some types of DN-to-GL transformations for contrast enhancement.

Examples of contrast enhancement

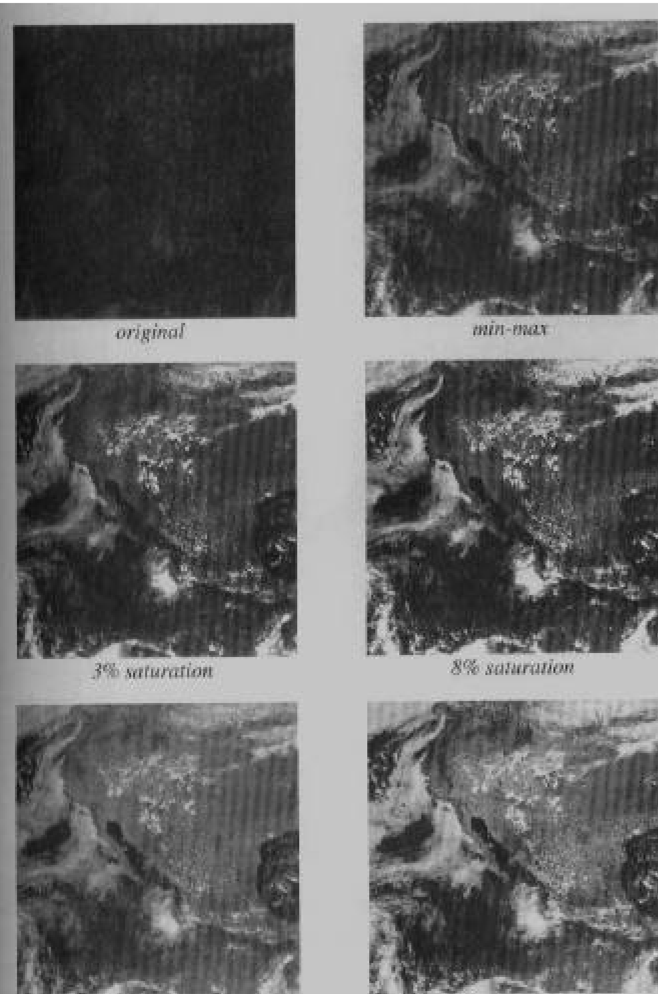


FIGURE 5-20. Examples of contrast enhancement using point transformations and global statistics. The test image is from the GOES visible wavelength sensor and shows cloud patterns over North America.

Summary of contrast enhancement algorithms

TABLE 5-3. Summary of contrast enhancement algorithms. A display GL range of [0,255] is assumed.

algorithm	equation	remarks
min-max	$GL = \frac{255}{DN_{max} - DN_{min}}(DN - DN_{min})$	sensitive to outliers
histogram equalization	$GL = 255CDF(DN)$	produces uniform histogram
normalization	1. $GL = \frac{\sigma_{ref}}{\sigma}(DN - \mu) + \mu_{ref}$ 2. $GL = 255, GL > 255$ $GL = 0, GL < 0$	matches means and variances
threshold	$GL = 255, DN \geq DN_T$ $GL = 0, DN < DN_T$	binary output
reference	$GL = CDF_{ref}^{-1}[CDF(DN)]$	matches histograms

Thresholding

- Object and background pixels have gray levels grouped into two dominant modes
- A threshold T is selected that separates the two modes
- Any point (x,y) for which f(x,y)>T is called an object point; otherwise the point is the background point
- The histogram can be used to select the threshold
- If the histogram has three dominant modes two thresholds T_1 and T_2 are selected
- Eg., two types of light objects on a dark background
- A point (x,y) belongs to: one object class if

  $T_1 < f(x,y) \leq T_2$, to the other object class if $f(x,y) > T_2$, to the background if $f(x,y) \leq T_1$

Iterative Threshold Selection Algorithm

1. Select an initial estimate of the threshold, T
2. A good initial value is the average intensity of the image
3. Partition the image into two groups, R_1 and R_2, using the threshold T
4. Calculate the mean gray value m_1 and m_2 of the partitions R_1 and R_2
5. Select a new threshold: $T=1/2(m_1+m_2)$
6. Repeat steps 2-5 until the mean values m_1 and m_2 in successive iterations do not change

Examples of thresholding on the GOES image

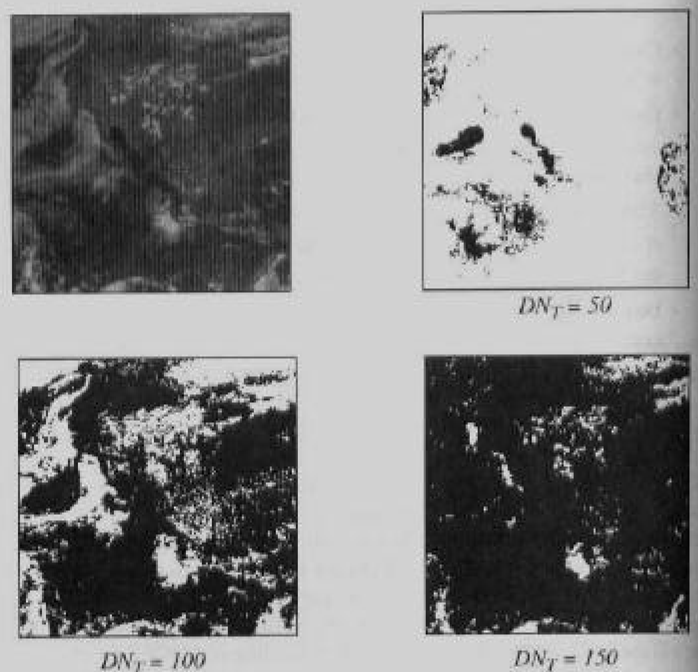


FIGURE 5-26. *Examples of thresholding on the GOES image of Fig. 5-20.*

enhancement may therefore be achieved by using an *adaptive* algorithm whose parameters change from pixel-to-pixel according to the local image contrast. We will describe a robust algorithm, *Local Range Modification (LRM)*, which illustrates the nature of adaptive processing (Fahnestock and Schowengerdt, 1983).

The essential idea is to partition the image into adjoining blocks (designated *A*, *B*, etc.) and derive a contrast stretch, different at each pixel, which is dependent on the local contrast within the corresponding block and surrounding blocks (Fig. 5-27). The stretch must change smoothly from pixel-to-pixel; otherwise brightness discontinuities occur at the boundaries between blocks (Fig. 5-28). Also, one premise in LRM is that the final *GL* range of the enhanced image should be predictable and not exceed specified minimum and maximum *GLs*.

Spatial Filtering

- *Image averaging* is used to remove noise in the image
- Noise reduction is achieved after 2, 8, 16, 32 or 128 noisy images
- Smoothing filters are used for blurring and for noise reduction
- Blurring is used in preprocessing, to remove small details from an image prior to object extraction and bridging of small gaps in lines or curves
- Filters are used as 3x3, 5x5 or 7x7 masks
- Low pass filters results in a response which is the average of all the pixels in the area of the mask
- For noise reduction <u>median filters</u> are used where the gray levels of each pixel is replaced by the median of the gray levels in a neighborhood of that pixel, instead of by the average

Spatial Filtering

- Smoothing filters are used for blurring and for noise reduction
- Blurring is used in preprocessing, to remove small details from an image prior to object extraction and bridging of small gaps in lines or curves
- Filters are used as 3x3, 5x5 or 7x7 masks
- Low pass filters results in a response which is the average of all the pixels in the area of the mask
- For noise reduction median filters are used where the gray levels of each pixel is replaced by the median of the gray levels in a neighborhood of that pixel, instead of by the average

Spatial Filtering

- High pass filters are sharpening filters which have masks with positive coefficients in the center and negative coefficients in the outer periphery
- A high pass filtered image may be computed as Highpass = Original - Lowpass
- Multiplying the original image by an amplification factor A, yields the high-boost or high-frequency emphasis filter
- High boost = (A-1)(Original) + Highpass
- While averaging analogous to integrations blurs details in an image, derivative filters use differentiation to sharpen an image e.g., Roberts cross-gradient, Sobel and Prewitts operators

Image Restoration

- An image degradation model is given by g(x,y)=H[f(x,y)]=n(x,y) where g(x,y) is the degraded image, f(x,y) is the original image and n(x,y) is the noise
- *Inverse filtering* is done by

$$\hat{F}(u,v) = \frac{G(u,v)}{H(u,v)}$$

- The transfer function H(u,v) can be approximated by the Fourier transform of the degraded image

Image Restoration

- Least mean square (Wiener filter) is given by

$$\hat{F}(u,v) = \left[\frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + g\left[S_n(u,v)/S_f(u,v)\right]}\right] G(u,v)$$

for u,v = 0,1,2,...N-1 where $|H(u,v)|^2 = H^*(u,v)H(u,v)$

- If g is variable the expression is called the parametric Wiener filter
- g has to be adjusted to satisfy the constraint

$$\|g - H\hat{f}\|^2 = \|n\|^2$$

- $S_f(u,v)$ and $S_n(u,v)$ are the spectral densities of the image and noise respectively

Edge Detection

- Gradient operators - the gradient is used for image differentiation
- The gradient of an image f(x,y) at location (x,y) is the vector

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{bmatrix}$$

- The magnitude of the gradient is

$$\nabla f = mag(\nabla f) = [G_x^2 + G_y^2]^{1/2}$$

- The direction of the gradient is

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

Edge Detection

- The Sobel operator provides both a differencing and a smoothing effect
- The smoothing effect is a particularly attractive feature as derivatives enhance noise
- $G_x=(z_7+2z_8+z_9)-(z_1+2z_2+z_3)$
- $G_y=(z_3+2z_6+z_9)-(z_1+2z_4+z_7)$ where z's are the gray levels of the pixels

<u>Mask to compute G_x</u>

-1	**-1**	**-1**
0	**0**	**0**
1	**2**	**1**

<u>Mask to compute G_y</u>

-1	**0**	**1**
-2	**0**	**2**
-1	**0**	**1**

Edge Detection

- The Laplacian of a 2-D function f(x,y) is a second-order derivative defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- It is implemented in digital form as a 3x3 region, and is given by $\nabla^2 f = 4z_5 - (z_2 + z_4 + z_6 + z_8)$

<u>Mask used to compute Laplacian</u>

0	**-1**	**0**
-1	**4**	**-1**
0	**-1**	**0**

Orthogonal Masks

- Orthogonal masks are used for formulating multimasks

<u>Basis of edge subspace</u>

1	$\sqrt{2}$	-1
0	0	0
-1	$-\sqrt{2}$	-1

1	0	-1
$\sqrt{2}$	0	$-\sqrt{2}$
1	0	-1

0	-1	$\sqrt{2}$
1	0	-1
$-\sqrt{2}$	1	0

$\sqrt{2}$	-1	0
-1	0	1
0	1	$-\sqrt{2}$

Orthogonal Masks

Basis of line subspace

0	1	0
-1	0	-1
0	-1	0

-1	0	1
0	0	0
1	0	-1

1	-2	1
-2	4	-2
1	-2	1

-2	1	-2
1	4	1
-2	1	-2

“Average” subspace

1	1	1
1	1	1
1	1	1

Examples of gradient magnitude images produced by common gradient filters

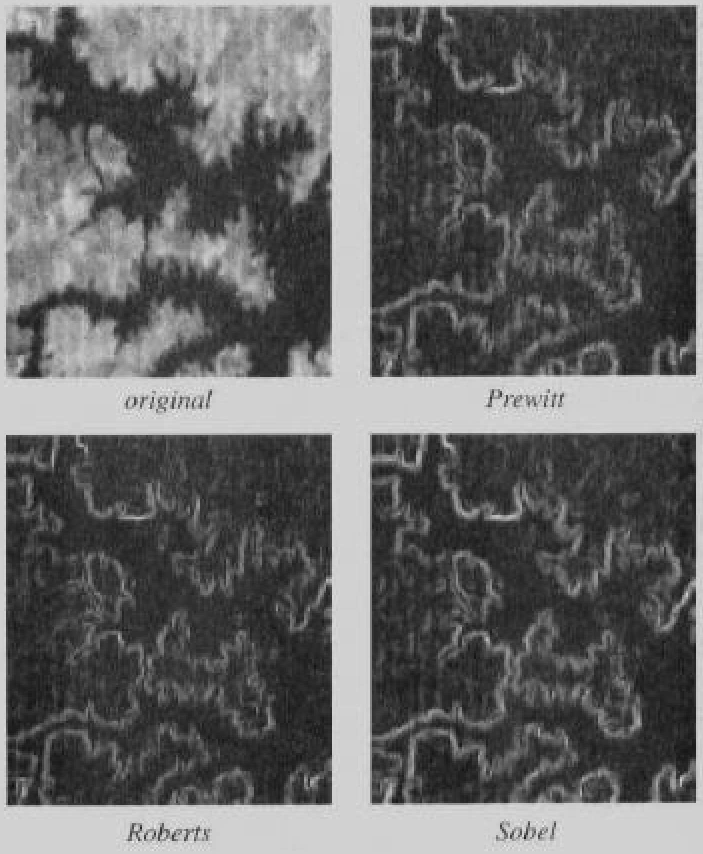


FIGURE 6-13. Comparison of the gradient magnitude images produced by common gradient filters.

Edge maps produced by thresholding the Roberts gradient magnitude at different levels

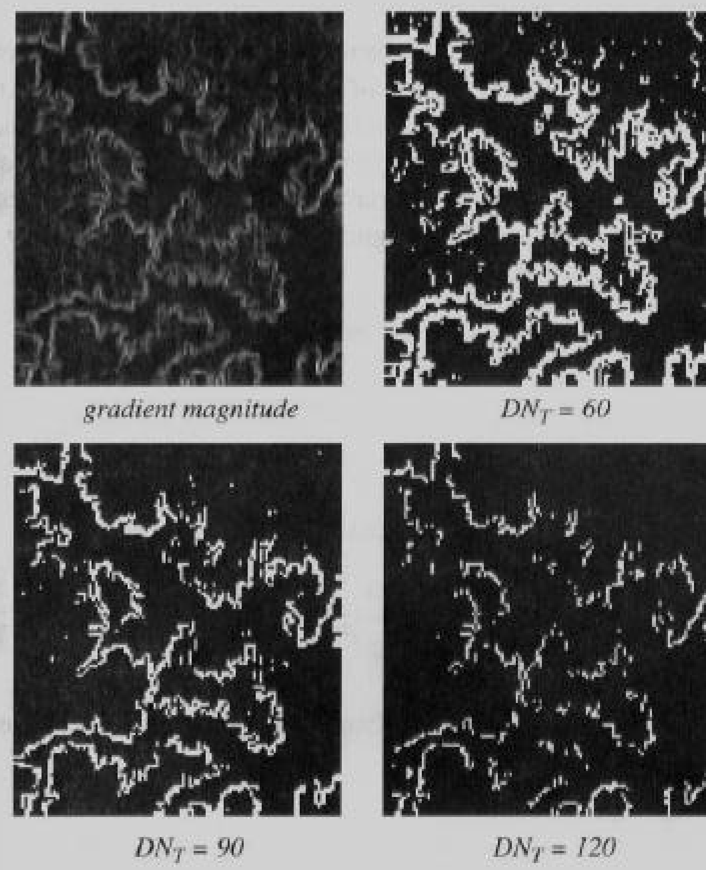


FIGURE 6-14. Edge maps produced by thresholding the Roberts gradient magnitude at different levels.

Color Image Processing

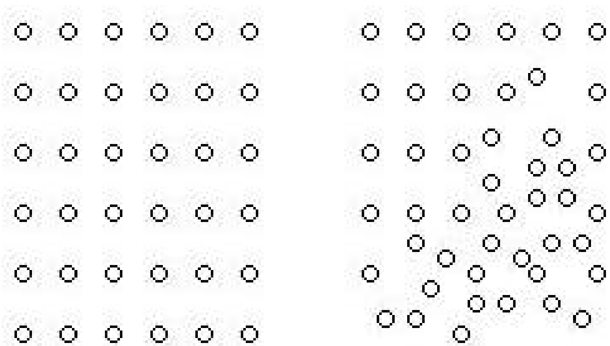
- The color models used in computers are CMY(cyan, magenta, yellow) - color printers
- YIQ (luminance, inphase and quadrature chromatic components) - color TV broadcast
- RGB (red, green, blue) and HSI (hue, saturation, intensity) models are also used
- One model can be converted to the other

Classification Techniques

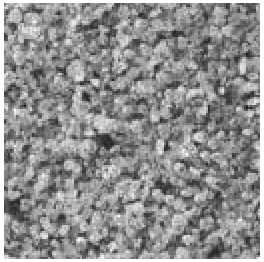
- Feature extraction
- Training
- Classification (labeling)
- Feature extraction – co-occurrence matrix method
- Parametric Classification – assumes a particular class statistical distribution (normal) and require estimates of the distribution parameters, such as the mean vector and covariance matrix for classification
- Nonparametric – no assumptions (considered robust)
  - Level-slice classifier
  - Histogram estimation classifier
  - Nearest neighbors classifier
  - Artificial neural network (ANN) classifier

Texture

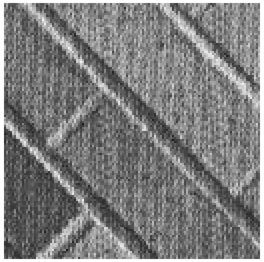
- Defined as the structural pattern of surfaces such as wood, grain, grass *which is homogeneous in spite of fluctuations in brightness and color*
- Most important visual cue in identifying regions, contributes in a number of ways to human vision-segmentation, shape detection, object recognition
- Main categories: regularity, randomness and directionality [Rao, 1993]



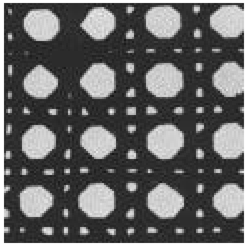
regularity



randomness



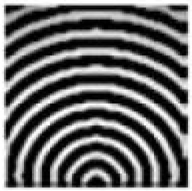
directionality



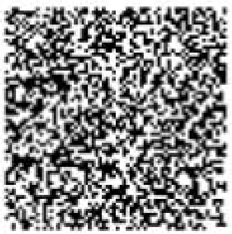
regularity

Texture

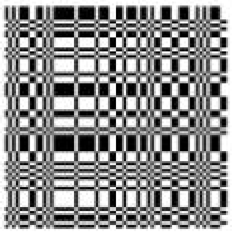
- Other types:



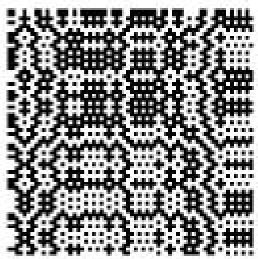
Stripe pattern



P=0.5 texture



even

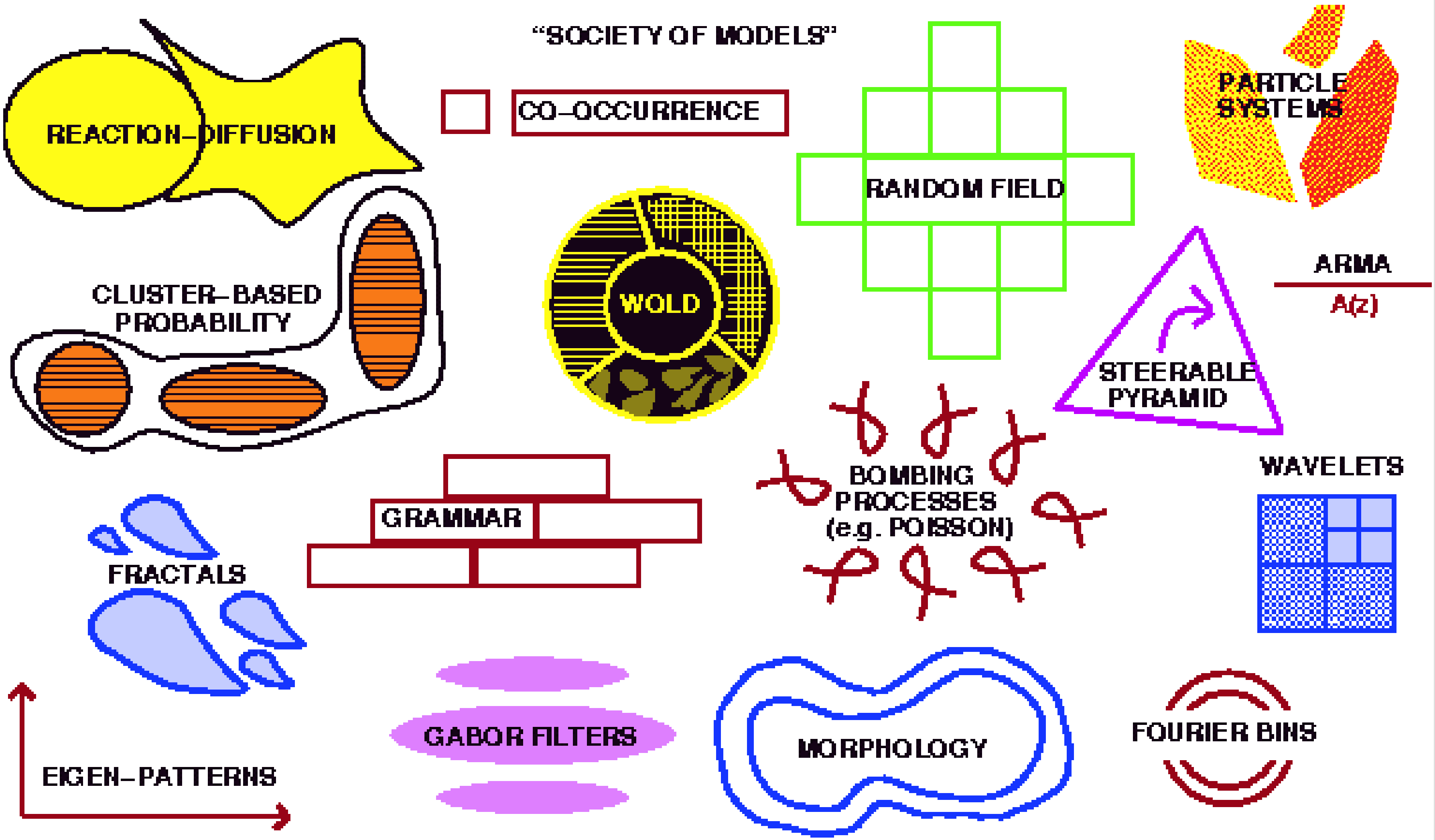


odd

Texture Analysis

Methods	Limitations
• Statistical methods –	• Not well localized, does not confirm to current vision models, best feature set is experimental and sometimes computationally intensive
	• Little useful information for texture discrimination
• Structural methods-	• Extraction of texels is a problem, applicable only to structural textures
• Filtering methods-	• Requires analysis levels, dependent on type of Gabor, wavelet filter used, some are computationally intensive
• Stochastic methods-	• Selection of model parameters
• Transform methods-	• Does not characterize textures locally
• Mask methods-	• Empirical choice of masks
• Moment based methods-	• Best set of moments is to be determined experimentally
• Other (neural network, genetic algorithm) methods-	• Optimize results for a specific applications but their convergence properties have not been established with proof.

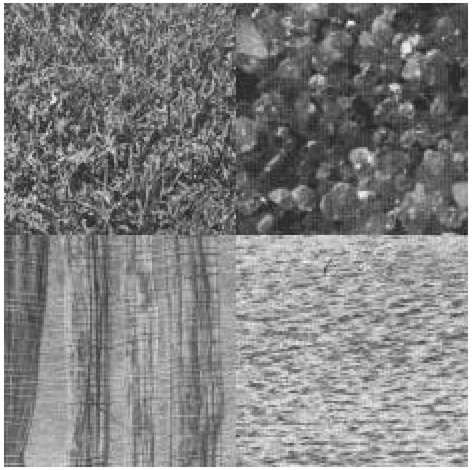
Texture Synthesis



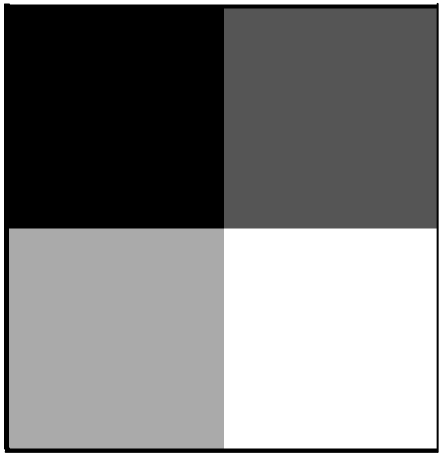
Texture in Image Analysis

- Segregation: A visible separation of features (texture)
- Discrimination: consider two entities to be *not equal* (different from segregation)
- Recognition: remembering an entity as something known
- Classification: Categorizing or sorting entities as pertaining to certain classes
- Segmentation: Using the categorized information to label the entities
- Supervised Classification: Classes are first learnt and the unknown entities are labeled as one of the known learnt classes to which they are close
- Unsupervised Classification: An initial grouping of entities into classes, is followed by further refinement of the grouping using criteria such as distance between means of the groups

Examples of Classification/Segmentation



Natural textures



Classification

Visual (Texture) Perception Theory

- Co-occurrence method - Haralick [1973]
- Caelli – an adaptive model using detectors – [1978]
- Texton theory - Julesz [1981]
- Texton - *local conspicuous features as basis for texture discrimination – analogous to phoneme in speech recognition*
- Pre-attentive – cannot process globally 3rd and higher order statistics
- Early vision and texture perception-Bergen and Adelson [1988]
- Gabor representation of biological vision - Daugman [1980] Porat and Zeevi [1989]

Problem Statement

- Many models claimed to model HVS processing- lack the robustness of HVS system
- Many methods developed heuristically pertaining to an application
- No conceptualization for developing algorithms that preserve similarity with HVS processing
- Main reason is psychophysical aspects have not been considered
- No method to model texture discrimination directly from the dynamics of perception (the dynamic nature is however a recent finding)

- Face/Vase illusion – involves higher cortical processing, stored information in brain that contains knowledge about vases and profiles

Texture Classification

- Texture is an important approach to region description
- Texture description provides measures of properties such as smoothness, coarseness and regularity
- Texture is characterized by the spatial distribution of gray levels in the neighborhood
- A texture region can be defined as a connected set of pixels satisfying a given gray-level property which occur repeatedly in an image region
- In texture classification the problem is identifying the given textured region from a given set of texture classes

Texture Classification

- Eg., a particular region in an aerial image may belong to agricultural land, forest region or an urban area
- The texture analysis algorithms extract distinguishing features from each region to facilitate classification of such patterns
- The three principal approaches for texture are statistical, structural and spectral
- The *statistical methods* are particularly useful when the texture primitives are small, resulting in *microtextures*
- Properties such as gray-level co-occurrence, contrast, entropy, and homogeneity are computed from image gray-levels

Texture Classification

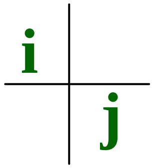
- *Structural methods* are used when the primitives are large, it is necessary to first determine the shape and properties of the primitives and then determine the rules which govern the placement of these primitives, forming *macrotextures*
- Model-based methods are used to synthesize texture
- A model for texture is first assumed and its parameters are estimated from the image region such that an image synthesized using the model closely resembles the input image region
- The parameters are then useful as discriminating features to classify the region

Gray-Level Co-occurrence Matrix

- The gray-level co-occurrence matrix P[i,j] is defined by specifying a displacement vector d=(dx,dy) and counting all pairs of pixels separated by d having gray levels i and j

2	1	2	0	1
0	2	1	1	2
0	1	2	2	0
1	2	2	0	1
2	0	1	0	1

5x5 image with 3 gray levels 0, 1, and 2



P[i,j]

$$(1/16) \times \begin{array}{c|c} \begin{bmatrix} 0 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 3 & 2 \end{bmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} \; i \\ \begin{matrix} 0 & 1 & 2 \end{matrix} & \\ j & \end{array}$$

The co-occurrence matrix for d=(1,1)

Co-occurrence Features

- The co-occurrence matrices can be computed for 0, 45, 90 and 135 degrees and at distances 1,2,3,...
- Statistical features are computed from the co-occurrence matrices
- Entropy is a feature which measures the randomness of gray-level distribution
- The angular second-moment is a measure of homogeneity of the image
- The contrast is a measure of the amount of local variations present in the image
- The correlation is a measure of linearity in the image

Contrast and entropy features from the sample images

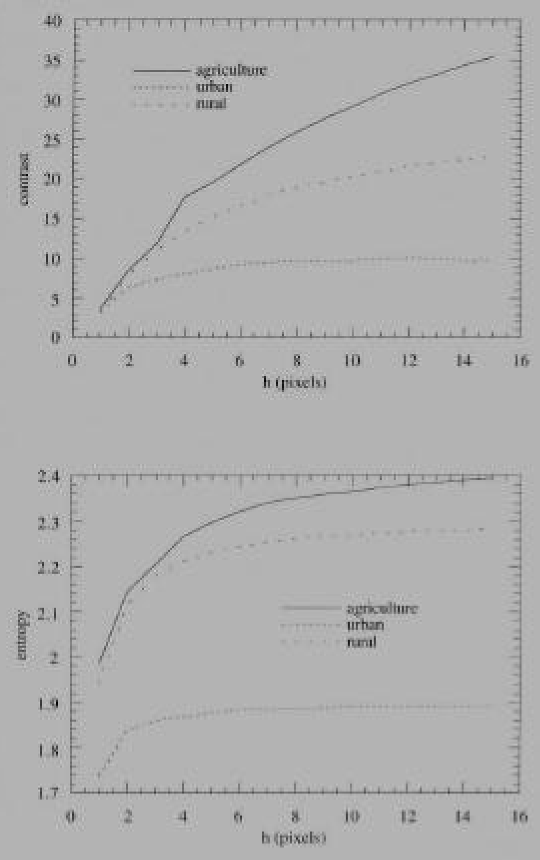


FIGURE 4-29. The contrast and entropy spatial features from the sample images of Fig. 4-28. Note the similarities to the semivariogram functions in Fig. 4-21 (which were calculated from an aerial image), namely the relatively rapid change as h increases from zero and the asymptotic behavior for large h.

Other Statistical Methods

- The difference statistics - Gray Level Difference Method (SGDM) are the distribution of probability that the gray-level difference is k between the points separated by a distance d in the image
- Higher order statistics is obtained from Spatial Gray Level Run Length Method (SGLRLM)
- The features are short runs emphasis, long runs emphasis, gray level nonuniformity, run length nonuniformity and run percentage
- Pixel based statistics such as mean, variance and distribution based features can be computed from texture regions

Level slice classifier

- Also called box classifier – a set of k-dimensional boxes, centered at the estimated class mean vectors, are placed in k-dimensional feature space
- If an unlabeled pixel vector lies within one of the boxes, it is assigned that class label
- Specification of the box limits is typically in terms of the data extent in each dimension, for example $\pm$one standard deviation about the mean

Histogram estimation classifier

- Supervised training pixels are used to construct the feature-space histogram of each defined class in K-D. The histograms are each normalized properly by the respective total number of pixels in each training class
- Every spectral vector “cell” in K-D is checked to find the class with the most histogram counts, and that class label is assigned to the cell
- A LUT that maps spectral vector to class label is created
- Unlabeled pixels are classified by the LUT

Nearest neighbor classifier

- Assigns the same label as that of the nearest training pixel
- K-NN – assigns label according to the majority label of k NN training pixels
- Distance-weighted k-NN – assigns weights to the labels of kNN training pixels, inversely proportional to their Euclidean distance from the unknown pixel, and assign the label with the highest aggregate weight

Assignment

- <u>Assignment: Fundamentals of Image processing</u>
- Write a function (functions) in Matlab to perform the Karhonen Loeve Transform (KLT) of an image. Write another function to reconstruct the original image from the transform using 1 or more eigenvectors. Compute the mean square error between the original and reconstructed image using different number of eigenvectors.
- Apply your function to the 7 channel Landsat image available in Matlab and check how the error varies when using different number of eigenvectors.
- For those interested: compute the DCT (Discrete Cosine Transform) of the image and reconstruct the original image and compare the mean square error with that of the KLT.