Wavelet Correlation Signatures for Color Texture Characterization

Wouwer, Scheunders, Livens, Dyck

Pattern Recognition
1999

Presented 2006-11-01 for COSI175 by

Ross Shaull

rshaull@cs.brandeis.edu

Brandeis University

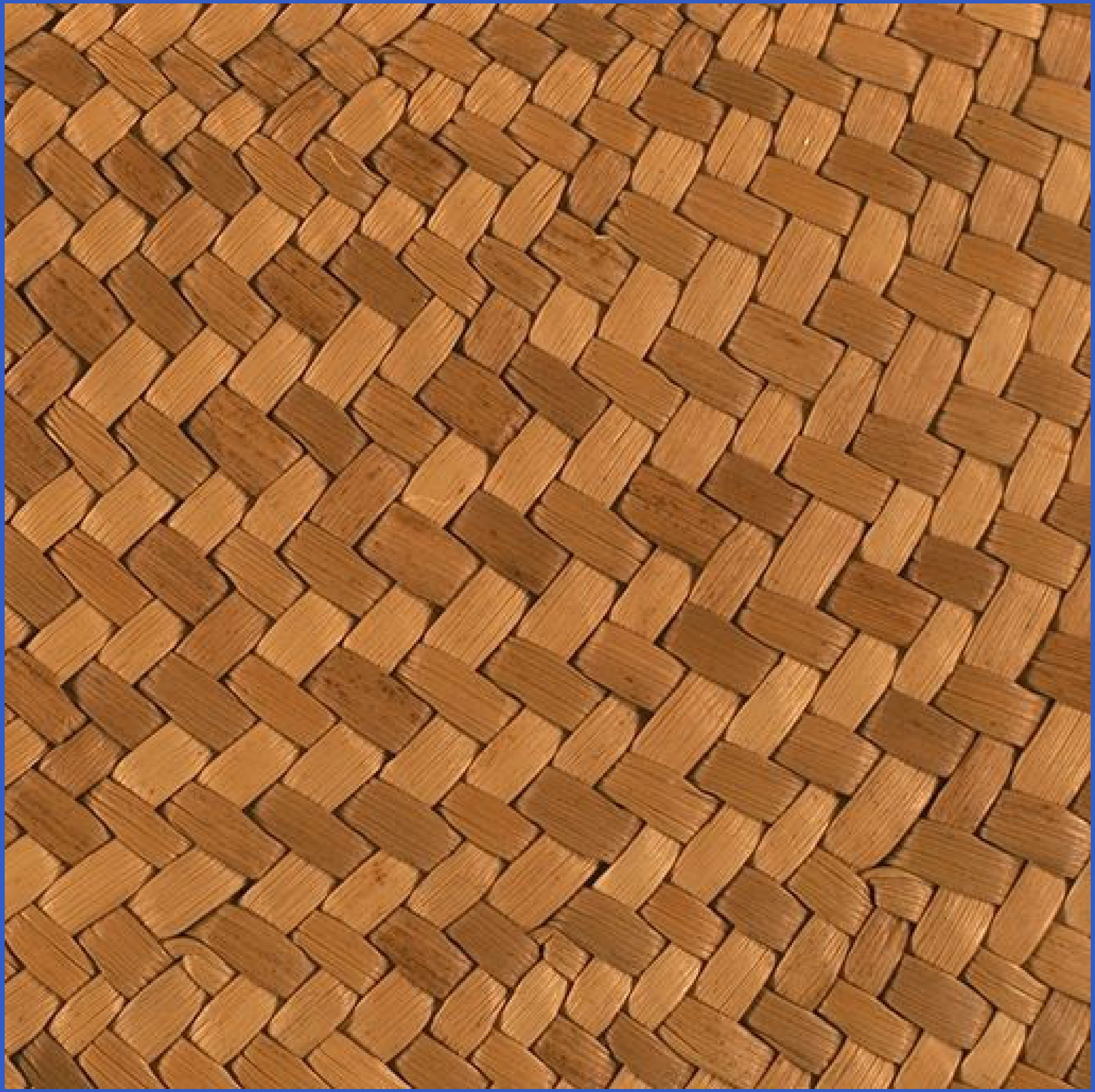
Outline

- What is texture?
- Wavelet transform
- Multiresolution analysis
- k-nearest neighbor classification
- Texture classification
- Conclusions

What is texture?

- Special relationship between pixels in a neighborhood that forms a recognizable pattern. (Wouwer 1999).
- Details that are periodic and directional, localized to a particular area. (Arivazhagan et. al. 2003).
- We are often interested in classifying swatches or regions of an image by their textural content.
- Before this paper, most techniques focused only on grey-level texture analysis. This paper seeks to include color information to improve classification performance.

Texture Example



Wavelet transform 1

- Produces subsampled (low-pass) image that is 1/4 the size of the original source image.
- Also produces detail images that correspond to the information in the high-pass filter.
- Sometimes called Quadrature Mirror Filtering (QMF), splits a signal into high- and low-pass band filters which are then subsampled (usually by factor of 2).

Wavelet transform 2

$Lo_x - Lo_y$	$Hi_x - Lo_y$
$Lo_x - Hi_y$	$Hi_x - Hi_y$

- Most of the energy is in the top-left, which is essentially just a lower-resolution version of the original image.
- The other three quadrants hold the vertical, horizontal, and diagonal details that are needed to reconstruct the original resolution image.

Lenna

http://cnx.org/content/m11087/latest/

Wavelet transform 3: 1-d

- The 1-d wavelet transform is an integral over a function $f(t)$ scaled by the mother wavelet $\psi_{a,b}(t)$, where the mother wavelet has time-shift and time-scale parameters a and b, respectively.
- The mother wavelet ψ must pass some admissibility criteria concerning smoothness.
- The integral gives us the energy of the signal in terms of versions of the mother wavelet.

Wavelet transform 3: 2-d

- Essentially, we can extend the wavelet transform to 2-dimensions by applying the 1-d transform to the rows and columns. This can be done using the discretized version (DWT).
- We have very fast techniques for applying the filters separably, that is, using iterated integrals.
- It turns out that with a discrete signal, we can treat the curves as a series of impulses, and so the integration can be represented as convolutions, which are much simpler to compute.

Convolutions to get the values in the four quadrants

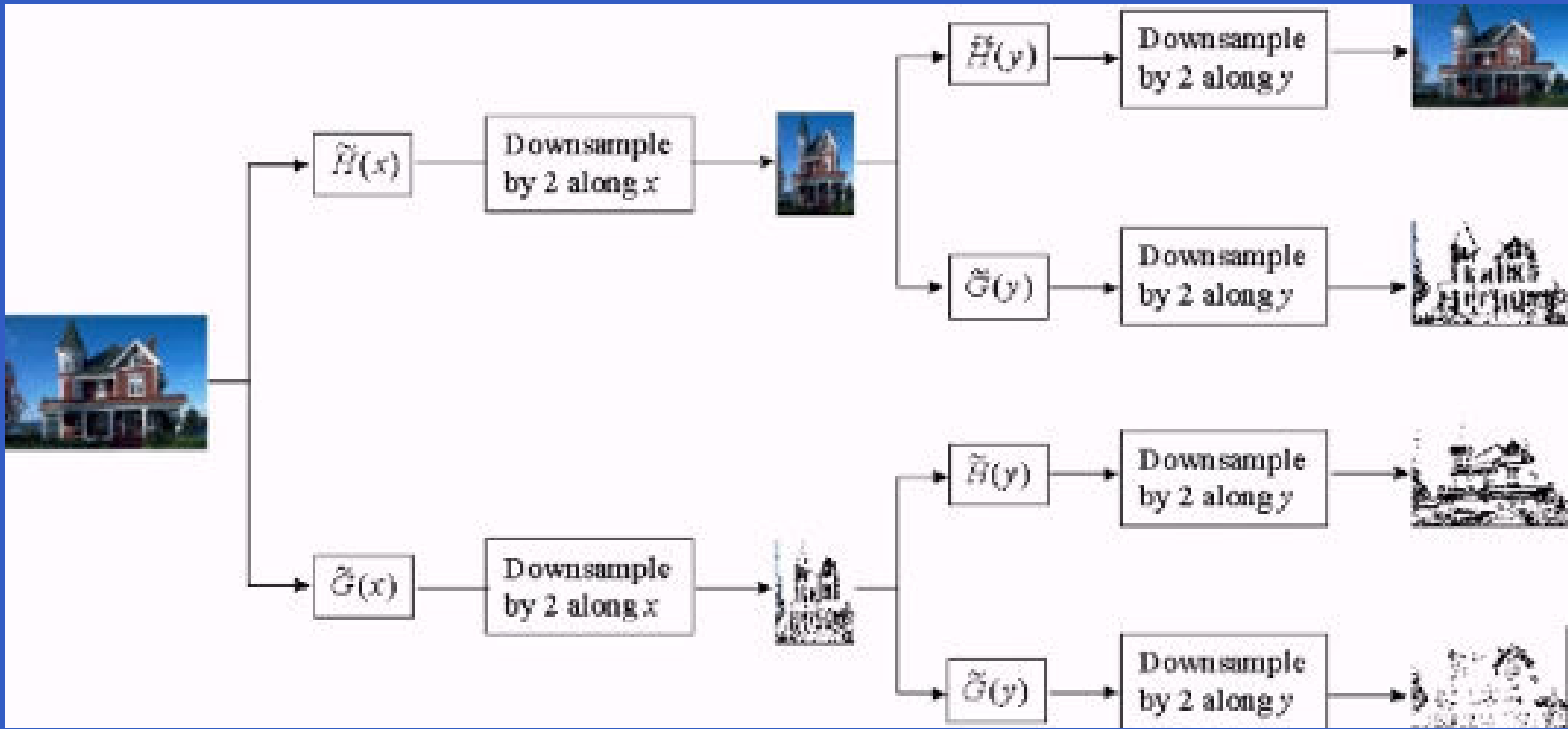
- $L_n(b) = [L_{n,x} * [L_{n,y} * L_{n-1}]_{\downarrow rows}]_{\downarrow cols}$
- $D_{n1}(b) = [L_{n,x} * [H_{n,y} * L_{n-1}]_{\downarrow rows}]_{\downarrow cols}$
- $D_{n2}(b) = [H_{n,x} * [L_{n,y} * L_{n-1}]_{\downarrow rows}]_{\downarrow cols}$
- $D_{n3}(b) = [H_{n,x} * [H_{n,y} * L_{n-1}]_{\downarrow rows}]_{\downarrow cols}$
- L_{n-1} is the low-pass image from the previous resolution.
- L_0 is the source image.
- $H_{n,x}$ is the high-pass filter at level n of the rows.
- $L_{n,x}$ is the low-pass filter at level n of the rows.
- y indicates over columns.
- The $\downarrow$ means to sub-sample over rows or columns.
- The $*$ operator means **convolution**, which for us is just the sum of the impulse responses (technically it is the area under the product of the two input curves).

Filter Banks

We compute a series of high-pass filters on subsampled images, representing the full-resolution image as a series of “detail images”, which are just transform of the same image at several scales, each time using a different scale.

Filter Bank Step

A single filter step generates four images:



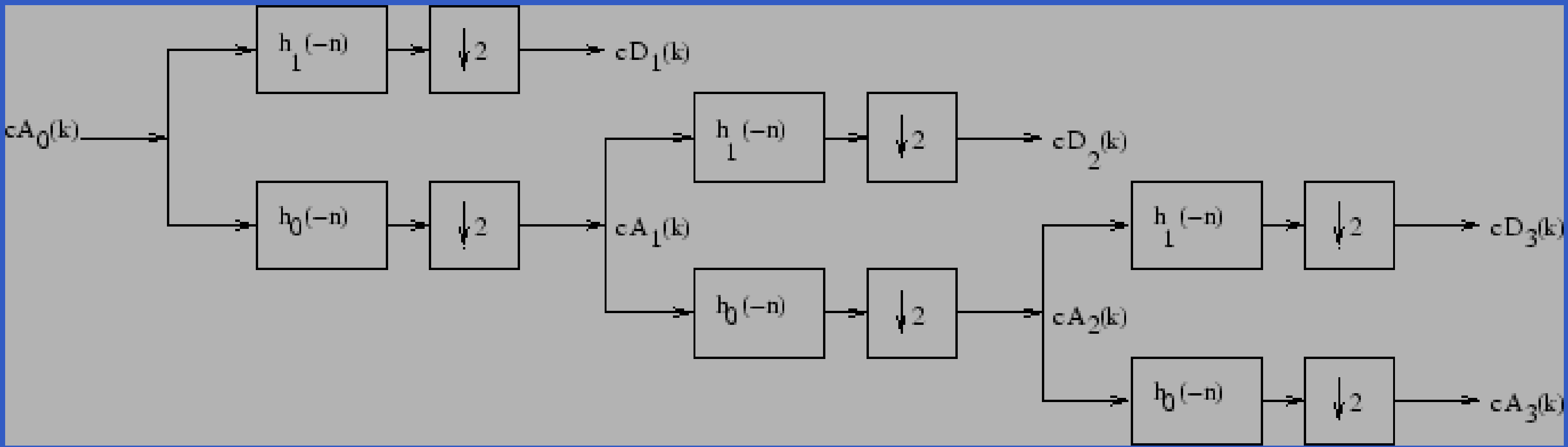
Stolen from Sharon Shen's presentation on the DWT at University of Maryland, Baltimore City.

Octave Filtering

- We will focus on “octave”-based filtering, which successively downsamples and transforms the low-pass results. This gives the features considered most useful to texture analysis (the detail images) at successively lower resolutions for a set of n resolutions.
- Typically we have four levels in the filter.

Octave Filtering in Excruciating Detail

- At each level, one or more detail images are produced, along with a low-pass image that results in a lower resolution version.
- Each level in the octave filter operates on the low-pass image from the previous level.
- Resolution decreases from 2^i to 2^{i-4} after passing through four levels in the filter.



http://www.engmath.dal.ca/courses/engm6610/notes/node6.html

Multiresolution Analysis

- It's a part of how the wavelet transform works with a multi-level filter bank.
- When we take the wavelet transform, we examine resolutions from 2^i to 2^j.
- We generate successively lower-resolution images from the original image, and at each step we find the detail images for that particular resolution.

Quick aside: the trouble with orientation 1

- Applying the filters separably produces strong vertical and horizontal bias in the detail images.
- This paper doesn't deal with orientation directly, but it is important to note that there are techniques for dealing with it.

Let's look at Lenna again

http://cnx.org/content/m11087/latest/

Quick aside: the trouble with orientation 2

- We can instead do the double integral using Lebesgue integration, which looks at volumes under surfaces to do the integration “all at once”, instead of iterating over each direction separately.
- This non-separable transform can be used with isotropic wavelets (wavelets that do not change when rotated).
- Another way is to use anistropic wavelets, but introduce a rotation parameter to create a periodic signal.

Wouwer, Ph.D. Thesis, University of Antwerp, 1998, http://visielab.ua.ac.be/staff/wouwer/thesis.php.

Wavelet Energy Signatures

- A wavelet energy signature at resolution n along orientation i is:

$$E_{ni} = \int (D_{ni}(\mathbf{b}))^2 d\mathbf{b}$$

- $\mathbf{b} \in R^2$.
- The energy signatures reflect how the image is distributed along the frequency axis over scale and orientation. We will use the notion of how energetic the signal is as a scaling factor later.

Wavelet Covariance Signatures 1

- The energy signature is useful for grayscale texture analysis. What about for color?
- Apply same filtering process to each color plane separately. We will have three sets of coefficients instead of just one. We call the new detail images we generate:

$$D_{ni}^{X}(\mathbf{b})$$

- Where X is the color plane being filtered.

Wavelet Covariance Signatures 2

- We can find a **covariance** signature between any two color planes by:

$$C_{ni}^{X_j X_k} = \int D_{ni}^{X_j}(\mathbf{b}) D_{ni}^{X_i}(\mathbf{b}) d\mathbf{b}$$

- This is cool because it tells us the coupling of energy (color and texture properties) between two planes.

Wavelet Correlation Signatures

From the title of the paper! Must be important...

- Scale the covariance signatures by the energy signatures (normalize by the grayscale transform to get that information which is only available in each color plane).
- The **wavelet correlation signatures** are what will be used to classify textures, because they offer a way to measure how texture features correlate between different planes and in different orientations, instead of flattening all three planes to get grayscale texture analysis.

How does this relate to texture analysis?

- Detail images store texture information at different resolutions.
- This allows recognition of micro- and macro-features of texture. (Arivazhagan et. al. 2003).

Color Space 1

- Remember wavelet correlation signatures? We need to get a color space with decorrelated color planes so that we can then calculate the correlation signatures.
- Use linear transforms (as opposed to transform into something like HSV) to avoid certain kinds of degenerate cases which remove information from the signal.

$$X' = MX$$

- X is original signal (3-vector), in our case $(RGB)^\tau$.
- M is a 3x3 transformation matrix.

Color Space 2

- Covariance matrix can be used to generate a transformation matrix that decorrelates the color planes (i.e., decorrelate the R, G, and B axes).
- This will optimize the representation by eliminating redundant information, and our linear transform will maintain that property.
- This is sometimes called Principle Component Analysis (PCA), also known as Karhunen-Loève (K-L).

Color Space 3

- This transform is technically dependent on the information in a specific image (the basis vectors for the linear transform depend on the features of the dataset).
- But, the authors claim that, statistically, the color images in their database have similar enough colorspace eigenvectors to use the same transform on all of them and get *close enough* to perfectly decorrelated axes.

Color Space 4

- The bottom line is that we can transform into different colorspaces, and the colorspace configuration informs the type of features vectors that can be retrieved from the dataset.
- We want to decorrelate the color planes, then re-correlate them using texture analysis from the wavelet transform.

Classification

- Supervised learning algorithm.
- We want to classify images into categories by comparing their textures with reference textures in a pre-classified database. We will do this by comparing *feature vectors* in the to-be-classified images with those same features in images from the database.
- Each image is divided into 64 non-overlapping subimages.
- The paper looks at a variety of feature sets. I'll focus on the most effective one, K-L space. This is not too surprising since the K-L linear transform results in an effectively optimal representation (very little redundant information).

Size of Feature Set

Example: K-L colorspace

- 3 detail images per level
- 4-level wavelet transform
- 6 covariances per detail image (from 3 components in colorspace basis)

This gives us a total of 72 correlation signatures (feature vectors). As we will see in a minute, we don't use all the features during classification, we just pick the best ones.

The paper looks at a variety of feature sets, with K-L being the best. Results will discuss how they compare.

knn Classifier 1

- The k-nearest neighbor classifier (knn) uses a set of labeled feature vectors to then search the feature vectors in *unseen data* for similarities.
- In knn analysis, we plot data points in a k-dimensional space, where each dimension is a feature. Each image in the training set will be plotted in this space. New data points will be assumed to be clustered with the nearest neighbor (using Euclidean distance).

knn Classifier 2

- This is not efficient, and actually adding features does not always lead to better performance. The authors use some statistical techniques to reduce the dimensionality (floating forward feature selection, or FFFS).
- Each image in the database is classified by training the database on all other images in the database. This is done for each image, starting with a fresh classifier each time. Sometimes called *leave-one-out*.

Error rate

- Class error rate

$$\frac{\text{no. falsely classified test samples from class i}}{N_i}$$

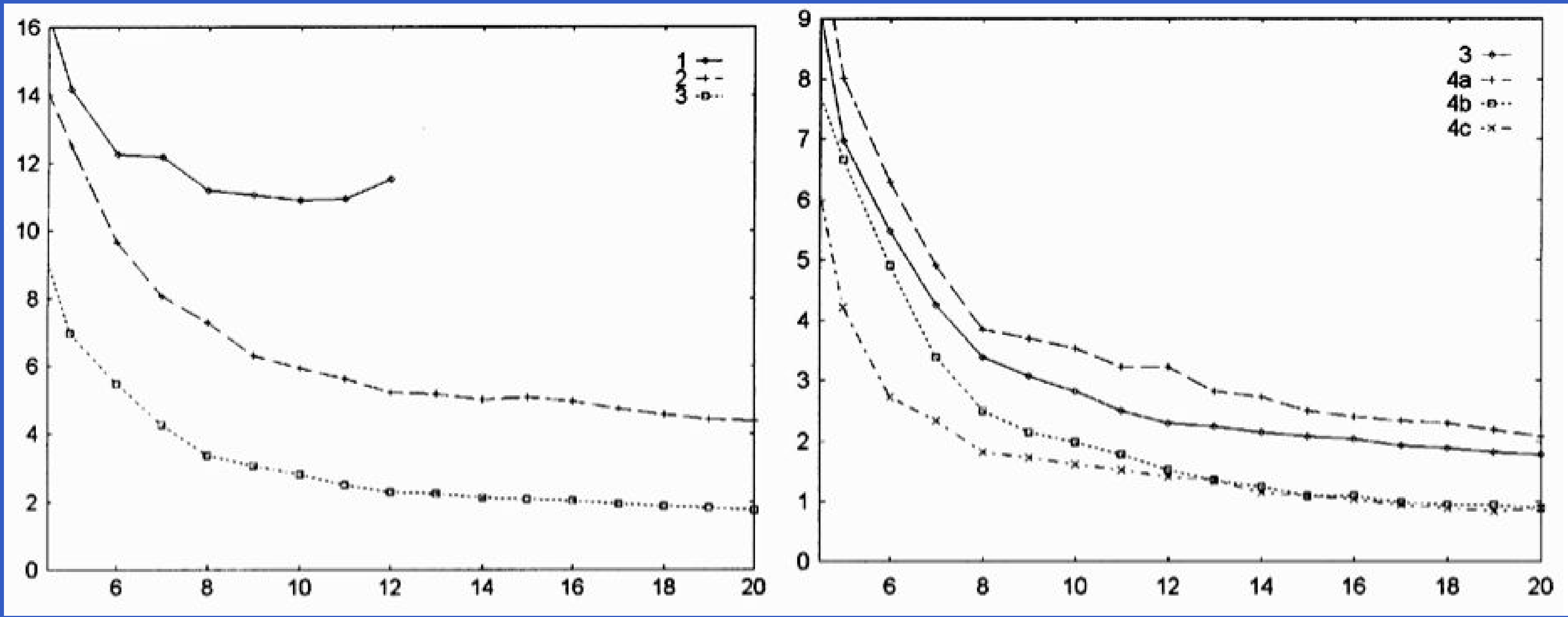
- Mean error rate

$$\frac{\text{total no. falsely classified test samples}}{N}$$

- N_i is the number of labeled feature vectors from the ith class, and;
- $N = \sum_{i=1}^{c} N_i$ is the total number of samples (the number of feature vectors in each of the c classes).
- A "class" is just the swatches taken from a single image.

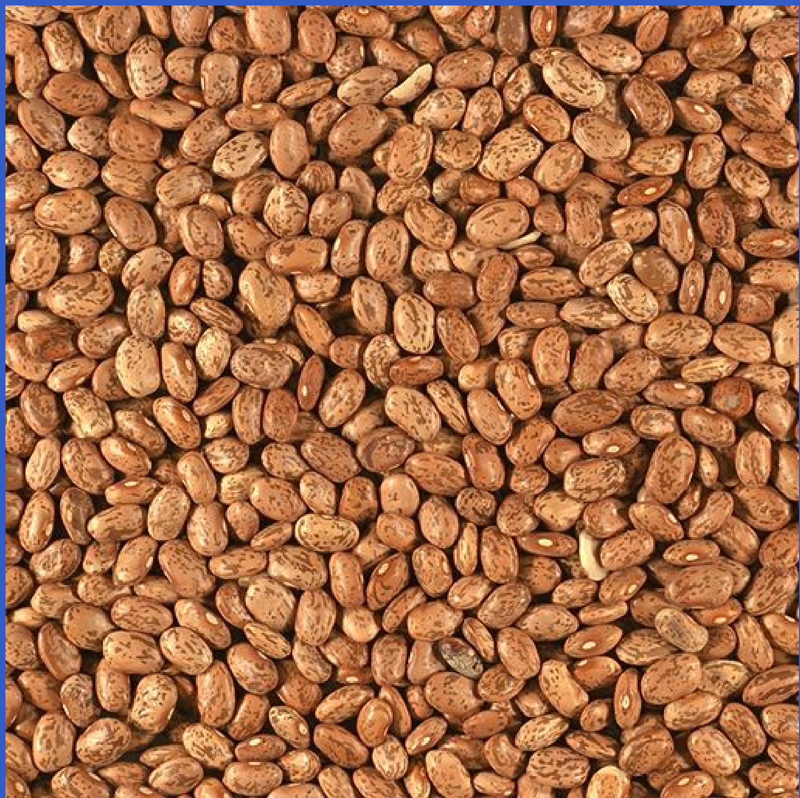
Results 1: Mean Error Rate

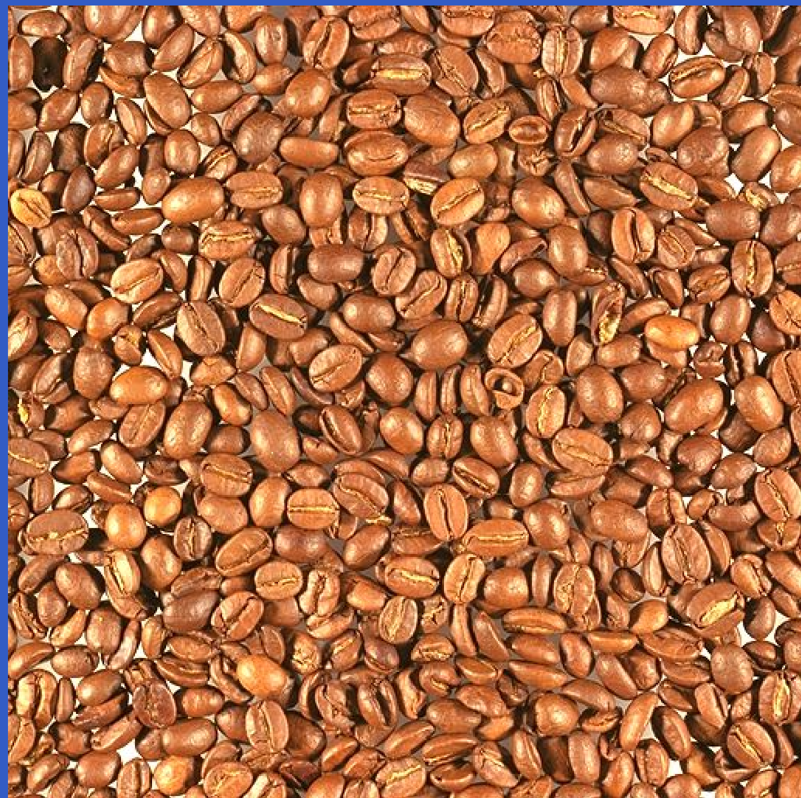
- Mean error rate (in percentage) versus feature set dimensionality.
- Left: different signatures, intensity (1), energy RGB (2), correlation RGB (3)
- Right: Correlation signatures in different color spaces, RGB (3), UVW (4a), YIQ (4b) K-L (4c).



Results 2: Intensity Alone

By analyzing with intensity alone, a mean classification error of 16% is achieved. They note that some textures are not classified well at all, such as the food textures where slight differences in color and reflectivity are more important than shape.





Results 3: Correlation Signatures

- Energy signatures from RGB better than grayscale (7% better)
- Correlation signatures are better still
- For K-L space, many of the errors were statistically not different from zero.

Results 4: Difficult Textures

These textures are among those still not perfectly classified in K-L space:

- It could be that for these textures, the 64x64 pixel blocks were too small and the main macro features of the images were lost.
- It could also be that directional variation within each image caused a problem for the classifier, this would be helped by extracting rotation invariant features as described in Wouwer's thesis.

What Did this Paper Contribute to the Literature?

- Color texture analysis. By using correlation signatures at different resolutions of the image, the wavelet decomposition can be used to accurately classify a texture against a pre-classified library of textures.
- The research shows that the correllation between color planes is an important component to identifying texture, improving classification results over simply using greyscale luminosity.

Questions?

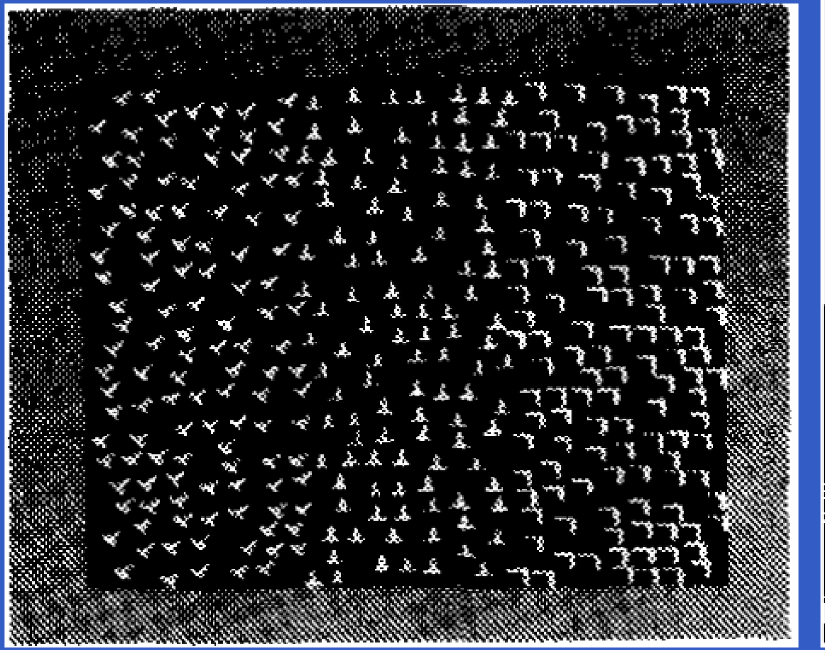
Neat stuff only partially related to the paper follows...

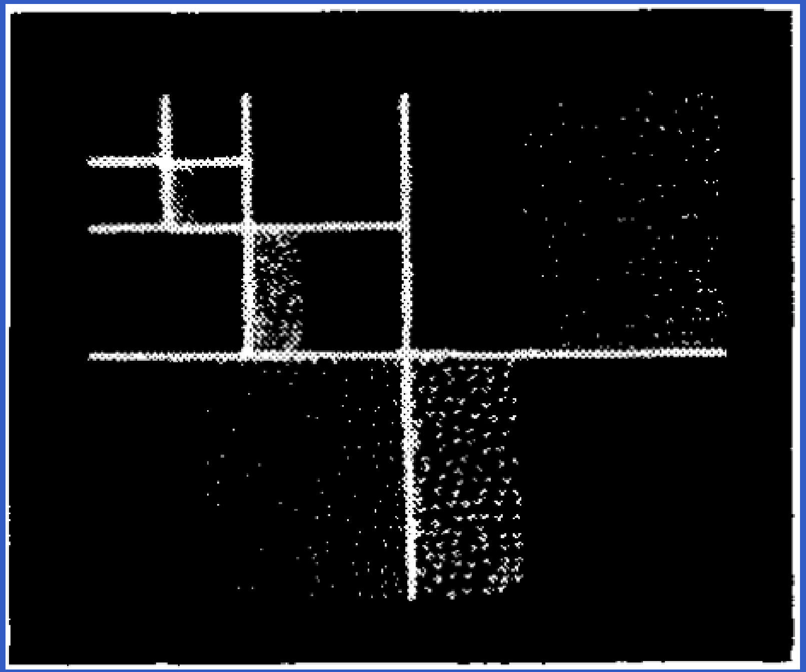
Similarity to Human Vision

Psychovisual evidence.

- There is evidence that the human visual system actually works in a multiresolution way. Some studies have shown that the classification ability of wavelet decomposition techniques provide similar discrimination abilities to that of a human.
- On the following slide, the left image has three textures. The right image shows the absolute value of the wavelet coefficients. Clearly, the first texture can be discriminated from the other two using a simple statistical measure.
- It turns out that humans have similar discrimination abilities for images like this one.

Similarity to Human Vision





Reference: Mallat 1989

What other Nifty Stuff is Out There?

- Texture descriptor in MPEG-7. MPEG-7 can classify chunks of an image or frame by their texture, which allows for semi-automated high-level semantic classification of content.
- Vision Lab UC Santa Barbara are doing neat stuff with this: http://nayana.ece.ucsb.edu/M7TextureDemo/Demo/client/M7TextureDemo.html

References and Acknowledgements

- The main paper: Wouwer et. al. Wavelet Correlation Signatures for Color Texture Characterization. Pattern Recognition Letters. 1997
- Wouwer Ph.D. Thesis. University of Antwerp. 1998.
- Arivazhagan and Ganesan. Texture Classification using Wavelet Transform. Pattern Recognition Letters. 2002.
- Websites:
  - http://vision.ece.ucsb.edu/
  - http://vismod.media.mit.edu/pub/VisTex/Images/Reference/
- Things you can learn by Googling:
  - http://en.wikipedia.org/wiki/Covariance_matrix
  - http://mathworld.wolfram.com/Convolution.html
  - Discrete-Time Convolution
  - 2-D DWT

- Special Thanks to Antonella Di Lillo for her valuable input about rotation invariance, the nature of textures, and help with slide organization.