The University of Auckland

CBIR: Texture Features - 2

COMPSCI.708.S1.C

A/P Georgy Gimel'farb

Gabor Features

- **Multi-resolution texture representation** based on Gabor filters
  – Image representation using Gabor filter responses minimises the joint space– frequency uncertainty
  – The filters are orientation- and scale-tunable edge and line detectors
  – Statistics of these local features in a region relate to the underlying texture information

Gabor Filters / Wavelets

- 1D Gabor function and its Fourier transform:

$$\gamma(x \mid x_0, \sigma_x, \xi_x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp\left[-\frac{1}{2} \left(\frac{(x - x_0)^2}{\sigma_x^2} \right) + 2\pi j \xi_x x \right]$$

$$\Gamma(u \mid \sigma_u, \xi_x) = \exp\left[-\frac{1}{2} \left(\frac{(u - \xi_x)^2}{\sigma_u^2} \right) \right] \quad \text{where} \quad \sigma_u = \frac{1}{2} \pi \sigma_x$$

- Product of a Gaussian and a complex-plane wave
  - Minimal joint uncertainty in both spatial and frequency domain

The University of Auckland

Gabor Filters / Wavelets

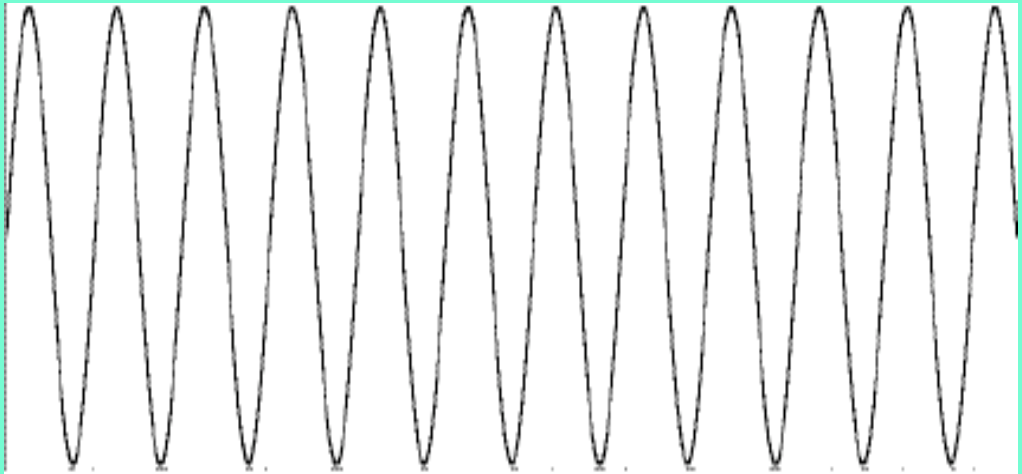
- Signal s is encoded by its projection onto these functions
- Decomposition is equivalent to a Fourier transform of the signal $s(x)$ in a Gaussian window:

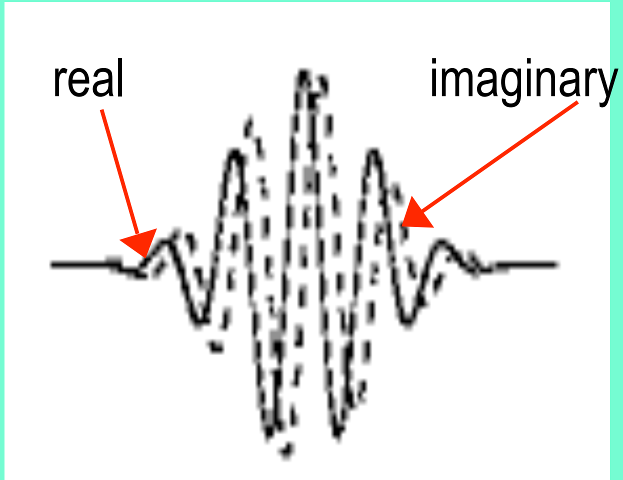
$$S(x \mid \sigma_x, \xi_x) = \int s(u)\gamma^*(x - u \mid \sigma_x, \xi_x)du$$

where * denotes the complex conjugate

- The filters form a complete but non-orthogonal basis for expanding a signal and gertting its localised description in terms of spatial frequencies

Gabor Filters / Wavelets





Wave: an oscillating function over infinite temporal or spatial domain having infinite energy

E.g. a pure sinusoid $\sin(2\pi\omega x)$

Wavelet: a small wave with localised in time or space finite energy

The University of Auckland

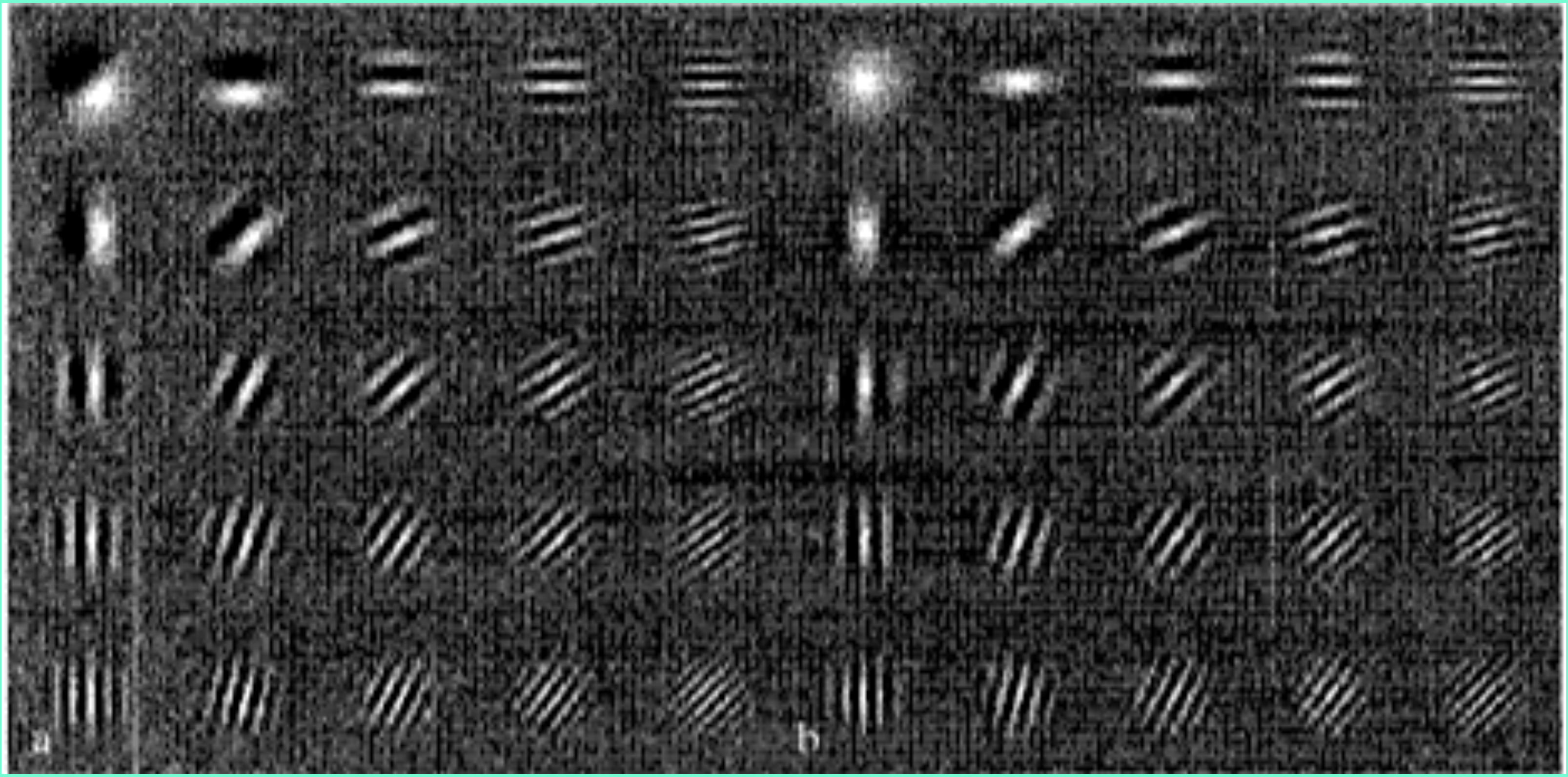
Gabor Filters / Wavelets

- 2D Gabor function: $\gamma(x,y) = \gamma(x|x_0,\sigma_x,\xi_x)\cdot\gamma(y|y_0,s_y,\xi_y)$
- Self-similar **Gabor wavelets**
  - By dilating and rotating the ***mother*** function $g(x,y)$
- KM wavelets with K orientations θ_n and M scales m:

$$\gamma_{mn}(x,y) = a^{-m}\gamma\left(\frac{x\cos\theta_n + y\sin\theta_n}{a^m}, \frac{-x\sin\theta_n + y\cos\theta_n}{a^m}\right);$$

$\theta_n = \frac{n\pi}{K}$; $a > 1$; integer $m \in [1,M]$, $n \in [0,K-1]$

Gabor Filters / Wavelets



2D odd wavelets 2D even wavelets

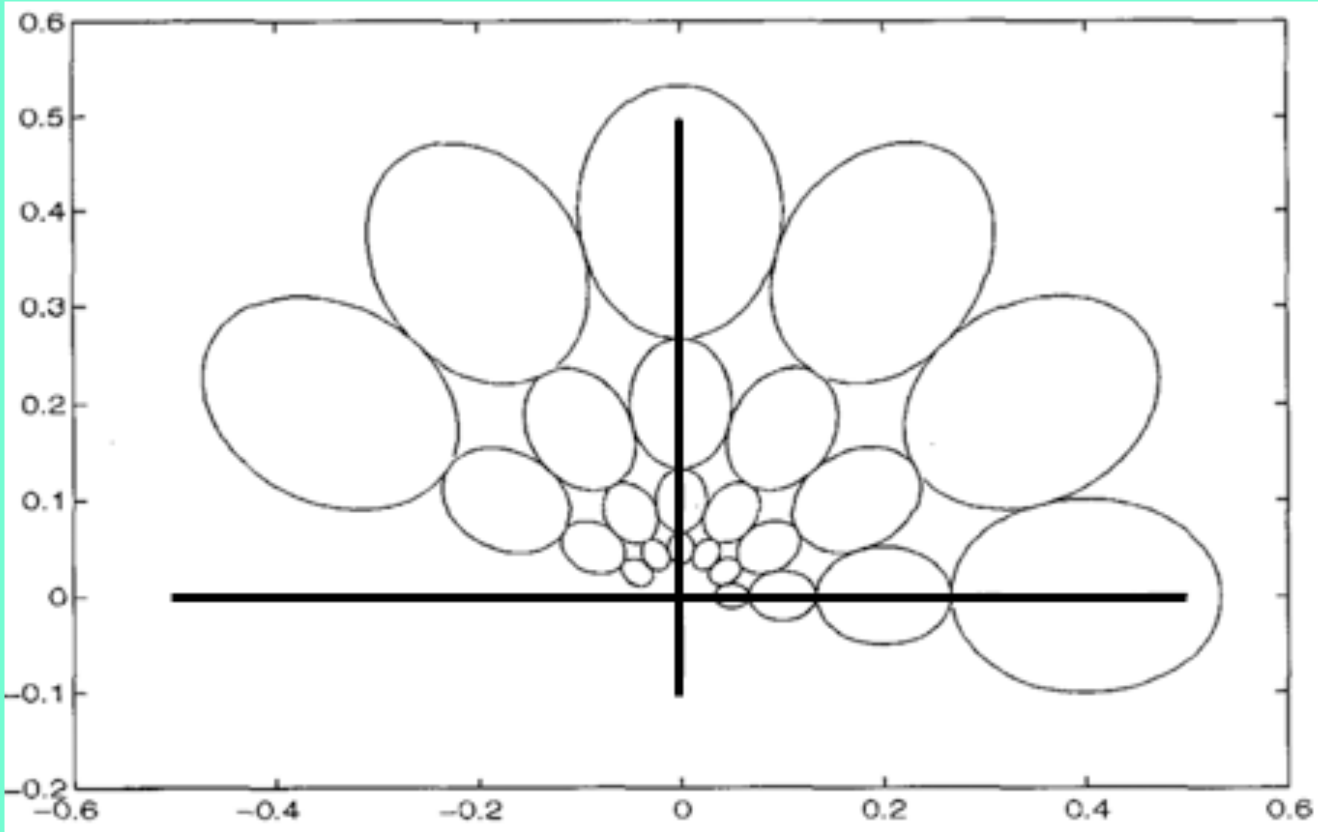
The University of Auckland

Gabor Filter Dictionary

- Gabor wavelets are not orthogonal ⇒ redundant data!
- Design: to choose a proper set (dictionary) of wavelets
  - Dictionary design ⇒ to reduce the data redundancy
- **Design strategy**: to ensure touching each other half-peak magnitude supports of the filter responses in the spatial frequency spectrum
  - Filter parameters σ_x, σ_y ⇒ from the lower and upper central frequencies ξ_x, ξ_y of the spectral domain and the number M of the scales

The University of Auckland

Gabor Filter Dictionary



Example: Lower frequency 0.05; upper frequency 0.40; $N = 6$ orientation angles, and $M = 4$ scales → 24 filters

Similarity of Feature Vectors

- Minkowski distance between two vectors of size L:

$$D(\mathbf{v}_1, \mathbf{v}_2) = \left[\sum_{l=1}^{L} | v_{1,l} - v_{2,l} |^p \right]^{\frac{1}{p}}$$

  – City-block (absolute): $p = 1$; Cartesian: $p = 2$

- Euclidean distance:

$$D(\mathbf{v}_1, \mathbf{v}_2) = \left[\sum_{l=1}^{L} \alpha_l (v_{1,l} - v_{2,l})^2 \right]^{\frac{1}{2}}$$

  – Weighted distance: *weight* $\alpha_l = 1 / (\text{variance } \sigma_l^2)$

The University of Auckland

Similarity of Feature Vectors

- Mahalanobis distance:

$$D(\mathbf{v}_1, \mathbf{v}_2) = \left[(\mathbf{v}_1 - \mathbf{v}_2)^{\mathsf{T}} \mathbf{S}^{-1} (\mathbf{v}_1 - \mathbf{v}_2) \right]^{\frac{1}{2}}$$

$$= \left[\sum_{l=1}^{L} \sum_{k=1}^{L} \alpha_{kl} (v_{1,k} - v_{2,k})(v_{1,l} - v_{2,l}) \right]^{\frac{1}{2}}$$

α_{kl} – the component of the inverse covariance matrix $\mathbf{A} = \mathbf{\Sigma}^{-1}$

- Accounts for variances of and statistical dependency between the vector components $v_{i,l}$; $l = 1, \ldots, L$; $i = 1, 2$

Similarity of Distributions

- **Kullback-Leibler** (KL) **divergence**

$$D(\mathbf{f}_1,\mathbf{f}_2) = \sum_{t=1}^{T} f_{1,t}\log\frac{f_{1,t}}{f_{2,t}} \equiv \sum_{t=1}^{T} f_{1,t}(\log f_{1,t} - \log f_{2,t})$$

where $\mathbf{f}. = (f_{\cdot,t}: t = 1,\ldots,T)$ is a frequency distribution

- **Chi-square (χ^2) distance**:

$$D(\mathbf{f}_1,\mathbf{f}_2) = \sum_{t=1}^{T}\left(f_{1,t} - f_{2,t}\right)^2 \Big/ f_{1,t}$$

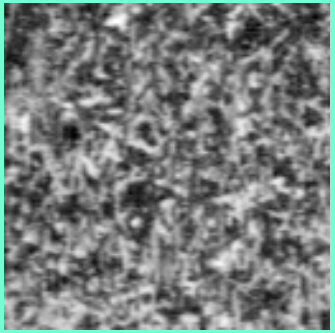
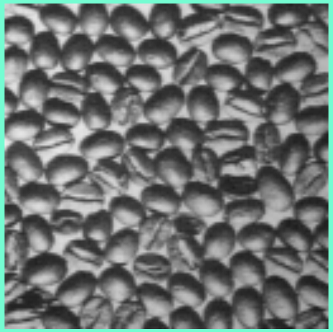
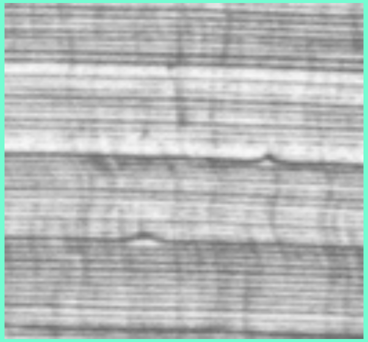
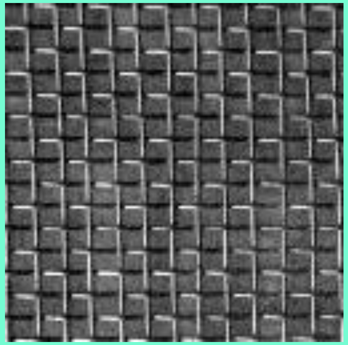
- **Symmetric distance**: $(D(\mathbf{f}_1, \mathbf{f}_2) + D(\mathbf{f}_2, \mathbf{f}_1))/2$

MPEG-7 Texture Descriptors

- The MPEG-7 ISO/IEC standard for ***multimedia content description interface*** involves descriptors for colour, texture, shape, and motion
- Three texture descriptors at this time:
  - the ***texture browsing descriptor*** $\Rightarrow$ directionality, regularity and coarseness of a texture
  - the ***homogeneous texture descriptor*** (**HTD**) $\Rightarrow$ description of homogeneous texture regions for similarity retrieval
  - the ***local edge histogram descriptor*** $\Rightarrow$ for inhomogeneous in texture properties underlying regions

Texture Browsing Descriptor

- 12 bits (maximum) to characterise a texture's regularity: 2 bits, directionality: 3 bits × 2, and coarseness: 2 bits × 2
  - The specification allows a maximum of 2 different directions and coarseness values
  - Regularity scale (2 bits): 0 (irregular, or random)…3 (periodic)

Texture Browsing Descriptor

- Directionality is quantified to 6 values: $0^o, 30^o, \ldots, 150^o$
- Bank of scale / orientation selective band-pass Gabor filters to select up to 2 dominant directions and compute the descriptor components from the filtered outputs:
  - 3 bits per each direction: **0** – no dominant directionality; **1 – 6** to code this dominant direction
  - Coarseness (2 bits per each dominant direction): 0 – a fine grain texture … 3 – a coarse texture
- Description relates to the partitioning of the frequency domain by the filter bank being used also for the HTD

The University of Auckland

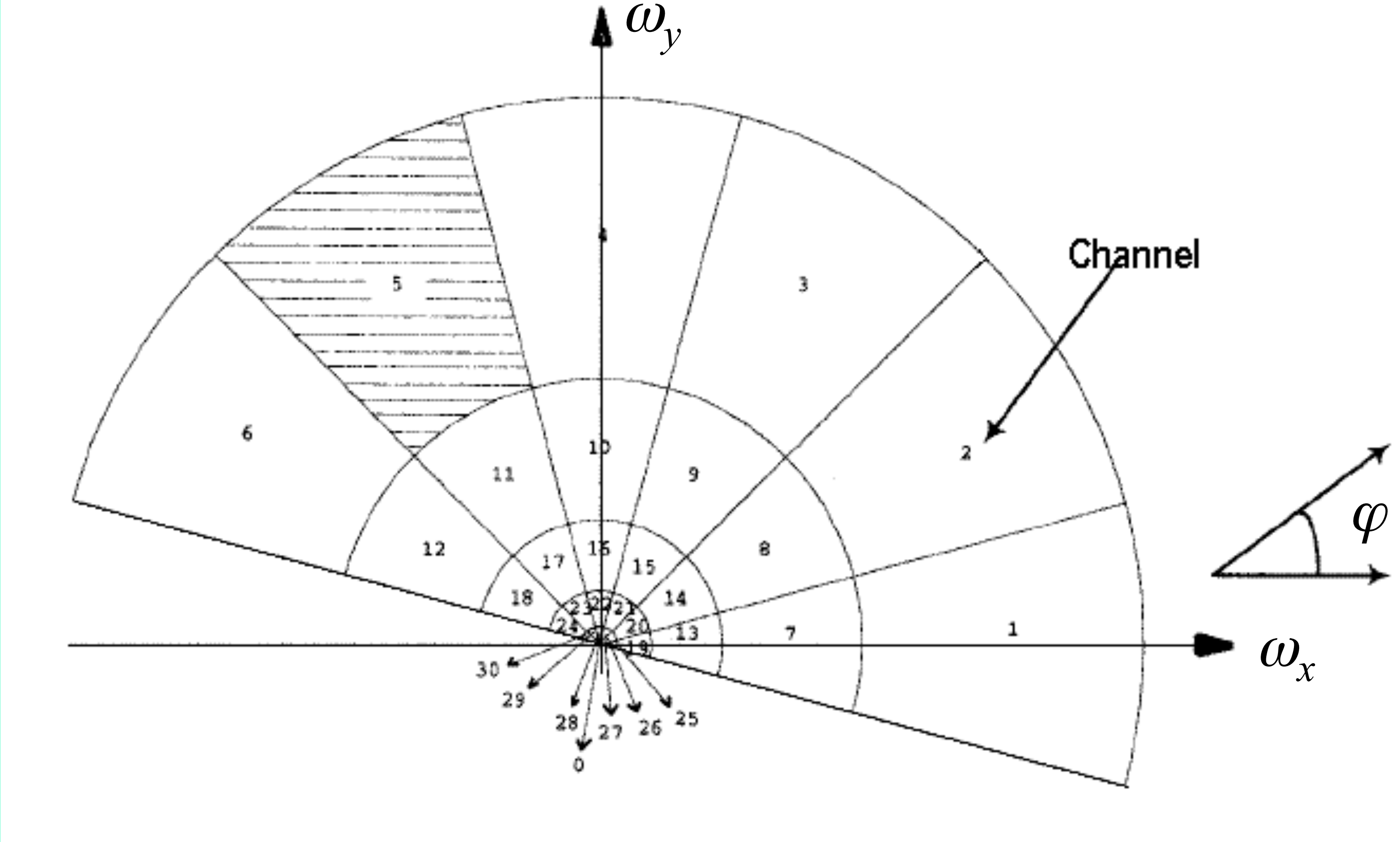
Homogeneous Texture Descriptor

- Means and standard deviations of the outputs of the bank of Gabor filters (Gaussians in the polar coordinates):

$$G_{i,j}(r_i,\varphi_j) = \exp\left[\frac{-(r-r_i)^2}{2s_{r,i}^2}\right]\exp\left[\frac{-(\varphi-\varphi_j)^2}{2s_{\varphi,j}^2}\right]$$

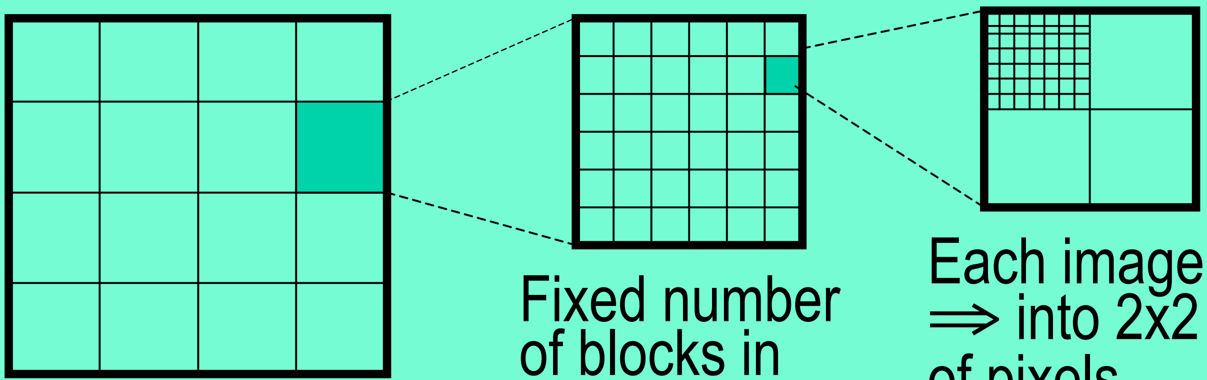
  - 30 output channels in a normalised frequency space $0 \le r \le 1$: 6 angular $\varphi_j = 30^o j$ and 5 radial divisions $r_i = r_0 2^{-i}$; $r_0 = 0.75$
  - 62 8-bit numbers per image or region: signal mean / deviation and logarithmically scaled energy / deviation in each channel

Homogeneous Texture Descriptor



The University of Auckland

Edge Histogram Descriptor



16 subimages

Fixed number of blocks in each subimage

Each image block ⇒ into 2x2 blocks of pixels

- 2x2 edge detector → to each image block by averaging signals
- Edge detectors: $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$ $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ $\begin{bmatrix} \sqrt{2} & 0 \\ 0 & -\sqrt{2} \end{bmatrix}$ $\begin{bmatrix} 0 & \sqrt{2} \\ -\sqrt{2} & 0 \end{bmatrix}$ $\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$
- **Histogram**: 5 bins×16 subimages= 80 bins; 3 bits/bin; an edge block if the maximum of the edge strengths exceed a preset threshold