Texture descriptors

- Texture: Spatial arrangement of similar patches
  - A quantitative measure of the arrangement of intensities in the region

Texture descriptors:

- Edge density and direction
- Local Binary Pattern (LBP)
- Co-occurrence Matrix.

Edge density and direction

- Compute gradient at each pixel.
- The descriptor: normalized histograms of magnitudes(H(mag)) and directions of gradients(H(dir)) over a region
  - Feature Descriptor: (H(mag),H(dir))
  - Numbers of bins in histograms kept small (e.g. 10)
  - Use L1 norm between the feature vectors as a distance

Local Binary Pattern (LBP)

3	2	1
4	c	0
5	6	7

$$b(i) = \begin{cases} 1 & if\ (I(i) > I(c)) \\ 0 & Otherwise \end{cases}$$

$$LBP(c) = \sum_{i=0}^{7} b(i)2^i$$

- Values range from 0 to 255.
- Obtain normalized histogram over a region.
- Not rotational invariant.
- Invariant to illumination and contrast.

T. Ojala, M. Pietikainen, and D. Harwood, A Comparative Study of Texture Measures with Classification Based on Feature Distributions, Pattern Recognition, vol. 29, pp. 51-59, 1996.

Variations of LBP

- Making it rotational invariant.
  - A circular neighborhood of radius R, with P pixels at equal intervals of angles.
  - Use interpolation if does not belong to the discrete grid

$$LBP_{P,R}(c) = \sum_{i=0}^{P-1} b(i)2^i \qquad LBP_{8,1} \leftarrow\rightarrow LBP$$

$$LBP^{ri}_{P,R}(c) = \min\{ROR(LBP_{P,R}(c),\, i) \mid i = 0,1,2,\ldots P-1\}$$

where $ROR(x\,,\, i)$ performs a circular bit-wise right shift on the P-bit number x, i times.

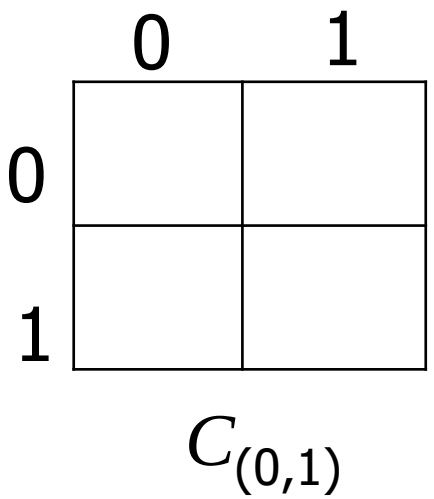
T. Ojala, M. Pietikainen, and T.i Maenpaa , Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, TPAMI, VOL. 24, NO. 7, JULY 2002, pp. 971-987

Co-occurrence Matrix (C_r)

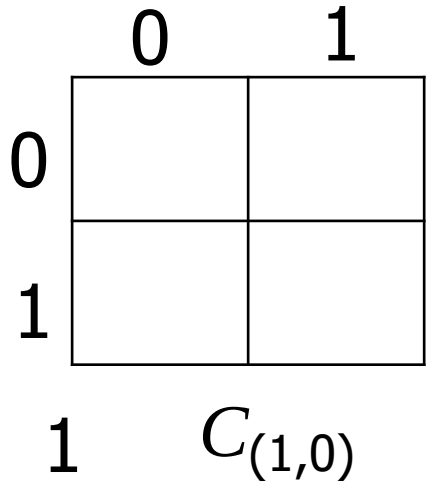
- $C_r(x,y)$: How many times elements x and y occur at a pair of pixels related spatially (designated by r in the notation).
  - e.g. **p** r **q** denotes **q** is shifted from **p** by a translation of **r**=(a,b), i.e. **q**=**p**+**r**.
    - $C_{(a,b)}(x,y)$: Number of cases in an image where I(**p**)=x and I(**p**+**r**)=y.

Co-occurrence Matrix (C_r)

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0



$C_{(0,1)}$



0 1 $C_{(1,0)}$

Co-occurrence Matrix (C_r)

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

	0	1
0	4	2
1	2	4

$C_{(0,1)}$

	0	1
0	4	2
1	2	4

$C_{(1,0)}$

	0	1
0	2	2
1	2	3

$C_{(1,1)}$

Normalized Co-occurrence Matrix (N_r)

Divide by the sum of frequencies in a matrix.

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

	0	1
0	1/3	1/6
1	1/6	1/3

$C_{(0,1)}$

		1
0	1/3	1/6
1	1/6	1/3

$C_{(1,0)}$

	0	1
0	2/9	2/9
1	2/9	1/3

$C_{(1,1)}$

Symmetric Co-occurrence Matrix (S_r)

$$S_r(x,y)=C_r(x,y)+C_{-r}(x,y$$

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

$C_{(0,1)}+C_{(0,-1)}$

	0	1
0	4+4	2+2
1	2+2	4+4

$C_{(1,0)}+C_{(-1,0)}$

	0	1
0	4+4	2+2
1	4+4	2+2

$C_{(1,1)}+C_{(-1,-1)}$

	0	1
0	2+2	2+2
1	2+2	3+3

Features from Normalized Co-occurrence Matrix

$$Energy = \sum_x \sum_y N_r^2(x,y)$$

$$Entropy = -\sum_x \sum_y N_r(x,y) \log_2 N_r(x,y)$$

$$Contrast = \sum_x \sum_y (x-y)^2 N_r(x,y)$$

$$Homogeneity = \sum_x \sum_y \frac{N_r(x,y)}{1+|x-y|}$$

Features from Normalized Co-Occurrence Matrix

$$Correlation = \frac{\sum_x \sum_y (x - \mu_x)\,(y - \mu_y)\, N_r(x,y)}{\sigma_x \sigma_y}$$

$$f(x) = \sum_y N_r(x,y)$$

$$g(y) = \sum_x N_r(x,y)$$

Mean and s.d. of row sums f(x)

Mean and s.d. of Column sums g(x)

Use of texture descriptors

- Detection of object patches represented by textured patterns.
- Segmentation of images.
- Classification / Matching
  - Generate a Library of labelled feature descriptors.
  - Detection of classes (class labels).
    - Matching to the nearest texture descriptor.

Image / Object Descriptor

- Bag of visual words
  - Compute key-point based feature descriptors over a library of images
  - Quantize them (clustering) to form a finite set of representative descriptors (visual words).
  - For an image assign the nearest visual word corresponding to the feature descriptor of a key point.
  - Represent by each image by a histogram of visual words.

Vector of locally aggregated descriptors (VLAD)

- Form the codebook of visual words as in BoVW representation
- $C_1, C_2, \ldots, C_k,$ *Cluster Centers*
- Each local descriptor x in an image is associated to one of these visual words.
- Accumulate the differences w.r.t. the corresponding cluster center.
- Form $V=[v_1\, v_2\, \ldots.\, v_k]$
    - VLAD descriptor=$V/||V||$

Application of global image Descriptor

- Content based image retrieval
  - Image search based on visual content

Query Image

Retrieved images from a database