Shape Context Matching Notes.

Let the two images taken as input be Im1 and Im2.

Points are sampled from internal and external contours of an image using canny edge detector.

The kind of sampling used in this paper is known as the Jitendra Sampling where initially many random samples are taken on an image, then one sample point out of every closest pair is eliminated until the total number of samples remaining is equal to the number of samples required (represented by 'n').

Let the contours be represented by the row and column pixel co-ordinates of the image. Therefore, vector x1 and y1 will represent the horizontal and vertical pixel co-ordinate vector of the contours of Im1.

Gradient is calculated using the Sobel operator. Let the horizontal and vertical gradients be G1 and G2 respectively.

$$t1(i) = \tan^{-1}\left(\frac{G2\big(x1(i), y1(i)\big)}{G1\big(x1(i), y1(i)\big)}\right) \qquad \text{for } i = 1, 2, 3, \ldots, n.$$

Computing shape context:

Let X = $[x1\ y1]^T$,

Computing the Euclidean distance between each sample point on the contour with respect to every other sample point on the contour and representing in a matrix form.

$$X_r = \sqrt{||X||} = \sqrt{(X^T, X)}$$

Also compute the slopes of each sample point w.r.t. every other sample point.

Let A be vector of ones of dimensions (1 x n)

$$\mathrm{T}_r = \tan^{-1}\frac{(y1 * A) - (A^T * y1)}{(x1 * A) - (A^T * x1)}$$

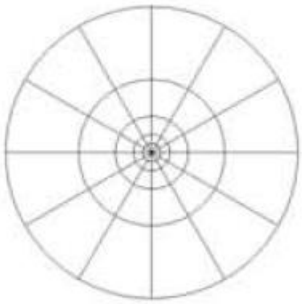
Normalizing X_r

$\mu = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}$ where μ is the mean distance

$\mathrm{X}_{rn} = \frac{X_r}{\mu}$ where X_{rn} is the normalized distance matrix

For creating a log polar histogram, 5 distance bins and 12 angular bins will be used as shown in the below figure.

$N_{dBins} = 5$ and $N_{thetaBins} = 12$



As per the paper [1] we'll be using a logarithmically spaced vector of distance 2 units of length 5 since N_{dBins} = 5.

B_{edges} = [0.125   0.25   0.5   1   2]

$$
x_{rq}(k) \begin{cases} 5; & \text{if } x_{rn}(k) < 0.125 \\ 4; & \text{if } x_{rn}(k) < 0.25 \\ 3; & \text{if } x_{rn}(k) < 0.5 \\ 2; & \text{if } x_{rn}(k) < 1 \\ 1; & \text{if } x_{rn}(k) < 2 \\ 0; & \text{if } x_{rn}(k) > 2 \end{cases}
$$

Eliminating all the outliers, distances greater than 2, and flagging the remaining sample points.

$x_{rn}(k) \in X_{rn}$ and $x_{rq}(k) \in X_{rq}$ for k = (i, j), where i = 1, 2, 3, …, n and j = 1, 2, 3, …, n

Converting all the angles in the range of [0,2π). Quantizing to a fixed set of angles.

$T_q = ((T_r \; mod \; 2\pi) \; mod \; 2\pi)/(\frac{2\pi}{N_{thetaBins}})$ here "*mod*" is the modulo operator.

$T = \lfloor T_q \rfloor$ getting all slopes in the range of [1,12] for all the 12 $N_{thetaBins}$.

$N_{bins} = N_{dBins} * N_{thetaBins}$

Finally, creating a matrix of dimensions (N_{dBins} x $N_{thetaBins}$) and placing each sample point in its appropriate bin in the log polar histogram forming the shape descriptor for that sample point.

$$N_{bins}^{i}(k,l) = \begin{cases} \text{increment by 1;} & \text{for } k = x_{rq}(i,j) \text{ and } l = T(i,j)\text{, where i = 1,2,...,n and j = 1,2,...,n} \\ \text{Increment by 0;} & \text{otherwise} \end{cases}$$

Here (k, l) together represents one of the 60 bins of the log polar histogram.

Therefore, a shape descriptor is created for every i^{th} sample point.

Shape descriptors of each sample point are reshaped and stacked together in a matrix forming the shape context of the image.

$S_{context}^{Im1}$ is of dimensions n * (k+l).

Similarly computing the shape context of 2nd image ($S_{context}^{Im2}$).

Next step is cost calculation for matching points on the 2 shapes.

Consider a point p_i on the first shape and a point q_j on the second shape. Let

$C_{ij} = C(p_i, q_j)$ denote the cost of matching these two points.

Let $h_i(k)$ represent the

$$C_{ij} = \frac{1}{2}\sum_{k=1}^{K} \frac{\left[h_i(k) - h_j(k)\right]^2}{\left[h_i(k) + h_j(k)\right]}$$