ASSIGNMENT 3

1)

a) Basic principle of corner detection:

→ Find correlation matrix of gradients in local window.
→ Find eigen values of correlation matrix
→ Detect ~~the~~ corner in window if eigenvalues are sufficiently large.

If in a ~~g~~ given window, two or more principal directions exist, that means there is ~~an~~ a corner in that window.

b) We need to find direction V such that projection of $\{g_i\}$ onto V is minimised

$$E(V) = \sum_i (g_i . V)^2 = \sum_i (g_i^T V)(g_i^T V) = \sum_i (V^T g_i)(g_i^T V)$$
$$= \sum_i V^T g_i g_i^T V = V^T \left(\sum_i g_i g_i^T\right) V = V^T C V$$

Additional directions minimize projection such that being orthogonal to previous direction.

$$\begin{cases} E(V) = V^T C V & C = \sum_i g_i g_i^T \\ V^* = \underset{V}{\arg\min}\ E(V) \end{cases}$$

$\Rightarrow \nabla E(V) = 0$

$2CV = 0$ ⇒ solution is eigen vector belonging to smallest eigen value.

eigen value = variance in corresponding principal direction

c) Gradient vectors = $\{(0,0), (0,1), (0,2), (0,3), (0,4), (1,0), (1,1), (1,2), (1,3)\}$

$C = D^T D$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 & 0 & 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 2 \\ 0 & 3 \\ 0 & 4 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & 6 \\ 6 & 44 \end{bmatrix}$$ ← Correlation matrix

d) If we get two eigen values that are sufficiently large, then we can say that there is a corner in that window

$$\lambda_1 . \lambda_2 > \tau$$

where λ_1, λ_2 = eigen values of correlation matrix

τ = Threshold

e) Compute $\lambda_1 . \lambda_2$ for all windows
Select windows with $\lambda_1 . \lambda_2 > \tau$ and sort in descending order
~~Select~~ Select the top of the list as corner, and delete all other corners in it's neighbourhood from the list
Stop once detecting k% of the points as corners.

f) For Harris corner detection, the cornerness measure is calculated by

$$C_i(c) = \det(c) - k\, tr^2(c)$$

where $\det(c)$ = determinant of c
$tr(c)$ = trace of C
c = correlation matrix

$\det(c)$ is equivalent to $\lambda_1 . \lambda_2$
$tr^2(C)$ is equivalent to $(\lambda_1 + \lambda_2)^2$

and the corner is detected where $C_i(c)$ is high.

1) g) Objective function for localizing a corner:

$$E(p) = \sum_i \left((p_i - p) \cdot \nabla I(p_i) \right)^2$$

when objective function is minimized solution for p is,

$$p = C^{-1} \sum_i \nabla I(p_i) \nabla I(p_i)^T p_i$$

~~[illegible]~~ $p = C^{-1} V$

$$C = \sum_i \nabla I(p_i) \nabla I(p_i)^T$$

Since we detected ~~that~~ corner in the window $\lambda_1 . \lambda_2 > \tau$

C must be non singular and hence the solution exists and C is invertible.

h) Split each patch into cells (possibly overlapping)
Create orientation histogram in each cell (using edge or gradient directions, possibly weighted by distance from center or gradient magnitude)
Concatenate orientation histograms.

Requirements for good characterization are translation invariance, rotation invariance, scale invariance and illumination invariance.

i) SIFT features are computed by using weighted sum to create orientation ~~hist~~ histograms in cells. and then concatenating
Then the histograms are aligned based on dominant direction.

2) a) Problem of using slope and y-intercept is that both of them have a range $-\infty$ to ∞.

b) $ax + by + c = 0$

$\Rightarrow \cos\theta \cdot x + \sin\theta \cdot y - d = 0$

$\theta = 45 \quad d = 10$

$\cos\theta = \cos 45 = 1/\sqrt{2}$

$\sin\theta = \sin 45 = 1/\sqrt{2}$

$$\frac{1}{\sqrt{2}} x + \frac{1}{\sqrt{2}} y - 10 = 0$$

c) The vote of each point in the image looks like a sinusoidal curve in the parameter plane.

d) For line detection, using parameter space Each point in image space makes a vote in parameter space which is defined and bound by the parameters of the line being used to describe the line.

Once votes are made in the parameter space, paramaters, with maximum votes are used to define the equation in image space.

In case of multiple points of intersection, non-maximum supression is performed.

2) e) In the parameter plane, larger bins are more efficient since there are fewer bins because of which ~~less~~ casting of votes is ~~le~~ less and the process is quicker. However, this provides less localization which is ~~less accurate~~. ~~no~~ not very accurate.

f) If the normal at each ~~v~~ voting point is known, a ~~s~~ single point needs ~~to~~ to cast only a single vote in the parameter plane rather than having to cast multiple votes which ~~m~~ makes the ~~proc~~ ~~voting~~ ~~very efficient~~ process very efficient.

g) When using Hough transform for circles ~~the~~ three parameters are considered (a, b) ~~as~~ as the center and r as the radius.

therefore, there are 3 dimensions in the parameter space.

~~3) b) ax + by + c = 0~~

~~1.x + 2.y + 2 = 0~~

3)

a) Equation $y = ax+b$ only ~~...~~ minimizes the algebraic distance between the actual points and the predicted points on the line, that is, it does not lead to an optimal solution.

Lines whose slopes are bigger cannot fit accurately with this ~~equ~~ equation.

b) $l^T x = 0$

l^T has three coefficients $\rightarrow$ a, b, c

$1.x + 2.y + 2 = 0$ where $a = 1, b = 2, c = 2$

$\therefore l = [1, 2, 2]$

c) The explicit line equation to minimize geometric distance is $l^T x = 0$, where all points 'x' on line 'l' should satisfy their objective function to be minimized

$$E(l) = \sum_{i=1} (l^T p_i)^2$$

$$= l^T \underbrace{\left(\sum_{i=1} (p_i p_i^T)\right)}_{S \,=\, \text{correlation matrix}} l = l^T S l$$

~~wher~~ $\cancel{S = \sum x_i^2 \; \sum x_i y_i \; \sum x_i}$

where $S = \begin{bmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i \\ \sum x_i & \sum y_i & n \end{bmatrix}$

d) Given points = $\{(0,1),(1,3),(2,6)\}$

$S = D^T D$

$$= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 3 & 6 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 1 \\ 2 & 6 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 15 & 3 \\ 15 & 46 & 10 \\ 3 & 10 & 3 \end{bmatrix}$$

e) The explicit equations for conic curves are

$$P_i = (x_i^2, x_i y_i, y_i^2, x_i, y_i, 1)$$

and $S = \sum_i P_i P_i^T$

The implicit equation is

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

$b^2 - 4ac < 0$, guarantees that the model will be an ellipse.

f) Given $\{x_i, y_i\}_{i=1}^{n}$, we have to solve

$$E(l) = \sum_i (l^T P_i)^2 \quad \text{where } P_i = (x_i^2, x_i y_i, y_i^2, x_i, y_i, 1)$$

to fit an ellipse using algebraic distance

$\Rightarrow S.l = 0$ where

$$S = \sum_i P_i P_i^T$$

Points ~~of~~ close to the short-axis of the ellipse have more effect on the fittings as these points get more weight considering the algebraic distance of these are lesser than those closer to the long axis of the ellipse.

g) Objective function to be minimized when fitting an ellipse using geometric distance

$$E(l) = \sum_i \frac{|f(P_i ; l)|}{|\nabla f(P_i ; l)|} \quad \text{where } f(P_i ; l) = l^T P_i$$

The complication is that this does not result in a quadratic equation, so we don't get an explicit solution and a linear solver cannot be used. So, we have to do an iterative approach like gradient descent.

h) Objective function for contours is

$$E[\phi(s)] = \int_{\phi(s)} \Big(\underbrace{\alpha(s)E_{continuity} + \beta(s)E_{curvature}}_{①} + \underbrace{\gamma(s)E_{img}}_{②} \Big) d_s$$

where $E_{continuity}$, $E_{curvature}$ and E_{img} are energy terms,

$\alpha(s)$, $\beta(s)$, $\gamma(s)$ are coefficients of the different energy terms

① is the ~~ind~~ internal energy and
② is the external energy

We want E_{cont} and E_{curv} to be ~~large~~ small and E_{img} to be high

$$E_{cont} = \left|\frac{\partial \phi}{\partial s}\right|^2, \quad E_{curv} = \left|\frac{\partial^2 \phi}{\partial s^2}\right|^2 \text{ and } E_{img} = -|\nabla I|^2$$

i) When the curve is discrete and using active contours:

$\cdot E_{cont}$ is estimated as distance between neighbouring points.

$E_{cont} = |P_i - P_{i-1}|^2$

E_{curv} is estimated as difference of tangents at neighbouring points

$$E_{curv} = \sum |(P_{i+1} - P_i) - (P_i - P_{i-1})|^2$$

$$= \sum |P_{i+1} - 2P_i + P_{i-1}|^2$$

j) The continuity of ~~sharp~~ active contours may be relaxed or to allow discontinuity we find high curvature points and set $\beta_i = 0$ i.e. if $|P_{i+1} - 2P_i + P_{i-1}| > \tau$ then $\beta_i = 0$.