CS512 Homework 3

Lin Zhuo

A20379609

1. Corner detection

a) Explain the basic principle of corner detection. How is the number of principal directions assessed?

Check the local neighborhood to see if there are more than one direction or orientation of gradient. More specifically, it includes:

- Find the correlation matrix in local neighborhood

- Find the eigenvalues of this matrix

- Check if the multiplication of the first two eigenvalues is larger than a certain threshold

The number of principal directions can be assessed by checking throught the histogram of the gradient direction.

b) Explain how PCA is used to find principal directions of gradient orientations in a local patch.

It looks for a direction v such that the sum of the projections from all the points P_i onto it is the maximum.

The objective function is $E(v) = \Sigma_i(P_i v)(P_i v) = \Sigma_i v^T P_i P_i^T v) = v^T \Sigma(P_i P_i^T)v$

The principal direction is $v* = argmax_v E(v)$

c) Given the gradient vectors{(0,0), (0,1), (0,2), (0,3), (0,4), (1,0), (1,1), (1,2), (1,3)}, compute the coorelation matrix that can be used for corner detection.

The correlation matrix is $A = \Sigma_{i=1}^n P_i P_i^T = \begin{bmatrix} \Sigma x_i^2 & \Sigma x_i y_i \\ \Sigma y_i x_i & \Sigma y_i^2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 44 \end{bmatrix}$

d) Write down the condition on the eigenvalues of the gradient correlation matrix that is used for corner detection.

The multiplication of both the first two eigenvalues should be sufficiently large, that is, $\lambda_1 \lambda_2 > \tau$ where τ is a threshold.

e) Explain how non-maximum suppression works for corner detection.

In every loop of the corner detection, after sorting the pixels based on $\lambda_1 \lambda_2$ and during the selection of the corner from strongest on, when a corner has been

selected, the non-maximum suppression principle delete the corners in vinicity of this detected corner.

f) Explain how Harris corner detection avoids computing the eigenvalues of the gradient correlation matrix directly.

The Harris corner detection employs the formular to check if its value is beyond the threshold:

$C(G) = det(G) - k * tr^2(G)$

where G is the correlation matrix, k is an coefficient between 0 and 0.5.

$det(G) = \lambda_1 \lambda_2, \ tr(G) = \lambda_1 + \lambda_2$

g) Write and explain the formula for computing better localization of a corner. What is the condition for the solution in the formula to exist?

$(\Sigma \nabla I(P_i) \nabla(P_i)^T)P = \Sigma \nabla I(P_i) \nabla I(P_i)^T P_i$

P is the position of the corner.

The formula comes from the minimization of the error function: $\Sigma_i [(P_i - P)\nabla I(P_i)]^2$, which means the sum of the projections of all the gradients onto the edge induced by the corner point P.

It can be written as $P = C^{-1}V$, where $C = \Sigma \nabla I(P_i) \nabla(P_i)^T$ and $V = \Sigma \nabla I(P_i) \nabla I(P_i)^T P_i$.

The condition for the solution to exist is that C is not singular.

h) Explain how feature points can be characterized using HOG (Histogram of Oriented Gradients). What are the requirements from a good characterization of feature points?

In HOG the image are divided into blocks and the histogram of gradient orientations in each block is calculated, then the histograms are concatenated.

The requirements from a good characterization of feature points are:

- Translation invariant
- Rotation invariant
- Scale invariant
- Illumination invariant

i) Explain how SIFT features are computed.

- Scale space construction: create internal representations of the original image to ensure scale invariance
- LoG approximation: for detect points of interest in the image
- Find the keypoints by checking the minima and maxima of the difference Gaussian image
- Assign the orientation to the keypoints
- Generate the SIFT features

2

2. Line Detection

a) Explain the problem of using the slope and y-intercept as line parameters when using the Hough Transform.

Due to the fact that Hough Transform maps the line in the image to a point in the parameter space, this approach may lead to an unbounded transform space since the slope can go to infinity.

b) Given a line with a slope of 45 degrees that passes at a distance of 10 from the origin, write the value of the coefficients a,b,c in the explicit line equation ax+by+c=0. To verify the result, draw the line, find points on it and see that they satisfy the explicit line equation.

$|0a + 0b + c|/\sqrt{a^2 + b^2} = 10$

$\frac{-a}{b} = tan(45°) = 1$

Solve the equation set and get:

$a = -b$; $c = \pm10\sqrt{2}a$, which could be $a = 1, b = -1, c = \pm10\sqrt{2}$

the explicit form: $x - y \pm 10\sqrt{2} = 0$

c) When using the polar representation of lines, what does the vote of each point in the image look like in the parameter plane?

The vote looks like a sine wave.

d) Explian how are lines detected by checking the parameter plane.

In the most ideal case, the cross of all the curves in the parameter plane is the representation of the line in the image. However in most cases the curves in the parameter plane are not crossing over a point. Thus non-maximum depression can be applied through counting the votes of each parameter, such as $(\theta, round(d))$ can selecting the peaks for representing the detected lines.

e) Explain the trade-off regarding the bin size in the parameter plane.

If the bin size is too small, the computational cost can be high. On the contrary if the bin size is too large, the precision of detecting the peaks cannot be guaranteed.

f) How can voting in the parameter plane be improved if the normal at each voting point is know?

The angle θ in the parameter plane does not need to loop from 0 to 360 degrees, instead in a narrower range around the known normals of the previous voting points. More specifically, compute the gradient θ of the image at the voting point first, and execute the loop from $\theta - \delta$ to $\theta + \delta$.

g) When using Hough transform for circles, what is the number of dimensions of the parameter space?

In 2-dimension space, a circle can be expressed by:

$(x - a)^2 + (y - b)^2 = r^2$

For a point (x, y) in the image the parameter space is 3-dimensional with three parameters a, b, r. However, if the radius is fixed, then the parameter space would be reduced to 2D (the position of the circle center).

3. Model Fitting

a) Explain the disvantages of using the equation y=ax+b for line fitting. What kind of lines cannot be fitted accurately using this equation?

The disvantage is the (a, b) parameters are not bounded. The lines that are parallel to y axis cannot be fitted.

b) Given a line with a normal (1,2) and the distance of 2 from the origin, write the value of the vector l representing this line in the explicit line equation $l^T x = 0$. Note that l is a 3×1 vecotr.

$l = [a, b, c]^T$ where a and b are the normal and c is the distance.

$l = [a, b, 2]^T$

It is known that the normal is $(1, 2)$ and it should be transformed to unit length, hence we get:

$b = 2a$ and $\sqrt{a^2 + b^2} = 1$

$l = [\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 2]^T$

c) Explain how to fit a line using the explicit line equation. Write the equation that has to be solved for the unkown line paramters.

The explicit line can be written as $ax + by + c = 0$, and the parameters are $l = [a, b, c]^T$.

The objective is to get the sum of all the distances of the points to the above line model $E(l) = \Sigma(l^T x_i)^2$, that is, $\partial E(l) / \partial l = 0$

$E(l) = \Sigma_i (l^T x_i)^2 = \Sigma l^T x_i x_i^T l = l^T \Sigma (x_i x_i^T) l$

$\frac{\partial E}{\partial l} = 2\Sigma(x_i x_i^T) l = 0$

The equation to solve is:

$\Sigma(x_i x_i^T) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$

d) Given the points $(0, 1), (1, 3), (2, 6)$ write the 3×3 matrix that has to be formed to find the parameters of the line that fits the points.

From c) we know the equation is:

$$\Sigma(x_i x_i^T) \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$

$\Sigma(x_i x_i^T)$ is the 3×3 matrix, more specifically, it can be written as:

$$\begin{bmatrix} \Sigma x_i^2 & \Sigma x_i y_i & \Sigma x_i \\ \Sigma x_i y_i & \Sigma y_i^2 & \Sigma y_i \\ \Sigma x_i & \Sigma y_i & \Sigma 1 \end{bmatrix} = \begin{bmatrix} 5 & 15 & 3 \\ 15 & 46 & 10 \\ 3 & 10 & 3 \end{bmatrix}$$

e) Write the explicit equations for conic curves. What is the constraint on the parameters a,b,c,d,e,f that guarantees that the model will be ellipse?

$ax^2 + bxy + cy^2 + dx + ey + f = 0$ is the explicit equation for conic curves.

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

f) Write the equation that needs to be solved for fitting an ellipse using algebraic distance. Explain which points on the ellipse affect more the fitting (points close to the long axis or short axis of the ellipse).

The point (x_i, y_i) in the image is mapped to $P_i = (x_i^2, x_i y_i, y_i^2, x_i, y_i, 1)$, and the parameters vector is $l = [a, b, c, d, e, f]^T$

Similar to fitting a line, fitting an ellipse is to minimizing $\Sigma_i(l^T P_i)^2$. By taking the derivative of it we get:

$$\Sigma_i(P_i P_i^T)l = 0$$

The equation has also constraint: $b^2 - 4ac < 0$, which can be written as: $l^T C l < 0$

where $C = \begin{bmatrix} 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

After adding the penalty of the constraint, the objective function th minimize is modified to:

$$E(l) = \Sigma_i(l^T P_i)^2 + \lambda(l^T C l + 1)$$

$$\nabla E(l) = 2Sl + \lambda 2Cl = 0 \implies S^{-1}Cl = \lambda^* l, \text{ where } S = \Sigma_i P_i P_i^T$$

The solution is the eigenvector of $S^{-1}C$ belonging to the negative eigenvalue.

Because of the algebraic distance $P_i = \frac{d_i}{d_i + r_i}$, where r_i is the distance to the center of the ellipse, the points close to the short axis affect more of the fitting.

g) Write the objective function that needs to be minimized when fitting an ellipse using geometric distance. What is the additional complication involved when trying to determine the ellipse paramters by minimizing this function?

The implicit function of the ellipse is $f(x) = 0$

The objective function that needs to be minimized is: $E(l) = \Sigma_i \frac{f(P_i;l)}{\nabla f(P_i;l)}$, where $f(P_i;l)$ is the algebraic distance from point P_i to the ellipse and $\nabla f(P_i;l)$ approximates the $\nabla f(x*;l)$ where x^* is the point on the ellipse such that the tangent of x^* and $P_i - x*$ are perpendicular.

The non-linearity is involved.

h) Write the objective function for active contours, and explain the components in it.

Let the active contour have the parameter curve form $\phi(s)$.

Principally, the objective function can be expressed as:

$E(\phi(s)) = \int_{\phi(s)} \alpha(s) E_{continuity} * \beta(s) E_{curvature} * \gamma(s) E_{image}$

where $\alpha(s), \beta(s), \gamma(s)$ are the coefficients; $E_{continuity}, E_{curvature}, E_{image}$ are the three degrees of measurements about which the three objectives are: small continuity, small curvature and highly integrated gradient.

- The continuity can be represented by the magnitude of the derivative (average progression rate) of $\phi(s)$. $E_{continuity} = |\frac{\partial \phi}{\partial s}|^2$
- The curvature can be represented by the second derivative of $\phi(s)$. $E_{curvature} = |\frac{\partial^2 \phi}{\partial s^2}|^2$
- The image term can be represented by the gradient of image. $E_{image} = -|\nabla I|^2$

i) Explain how the continuity and curvature of a discrete curve can be estimated when using active contours.

In discrete space, the parameter curve $\phi(s)$ is represented by discrete points $\{P_i\}_{i=1}^n$.

The continuity can be estimated as: $E_{continuity} = |\frac{\partial \phi}{\partial s}|^2 = \Sigma|P_i - P_{i-1}|^2$;

The curvature can be estimated as:

$E_{curvature} = |\frac{\partial^2 \phi}{\partial s^2}|^2 = \Sigma|(P_{i+1} - P_i) - (P_i - P_{i-1})|^2 = \Sigma|P_{i+1} - 2P_i + P_{i-1}|^2$

j) Explain how the continuity of active contours may be relaxed to allow for sharp corners.

The formula can be changed to: $E_{continuity} = \Sigma(|P_i - P_{i-1}| - d)^2$

where d is the average distance between two points.