Phuong Vu
30873216
Prof. Chenliang Xu

Homework #1
(Analytical Questions)

Problem 3 (20pts): Part One:

① Box filter:

$$\frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{9}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}\right)$$

② Sobel filter

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

Part Two:

Convolving a separable filter of size $k \times k$ can be done by separating that filter into 2 vectors of size $k \times 1$ and $1 \times k$, then convolve them individually to the image. Each such convolution costs $O(kHW)$ in terms of time complexity, so overall they cost $O(kHW + kHW) = O(2kHW) = O(kHW)$ (because of big-O's characteristics, a constant factor can be dropped)

Problem 4 (10pts) : Edges

Laplacian of Gaussian means blurring the image using a Gaussian filter before applying (convolving) the Laplacian kernel to the same image. (then edges can be detected using zero-crossing points)

Compared with Laplacian of Gaussian, the Laplacian filter is not a good edge detector as it is very sensitive to noise.

Problem 5 (20pts) : Bonus for CSC 249

① Derive the form of the 3D structure tensor from the sum of squared diffs (SSD) error function (for window W):

$$E(u, v, w) = \sum_{x,y,t \in W} [I(x+u, y+v, t+w) - I(x,y,t)]^2$$

Use Taylor series to approximate $I(x+u, y+v, t+w)$:

$$I(x+u, y+v, t+w) = I(x, y, t) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} w + \text{higher order terms}$$

Because (u, v) is small, the first order approximation is good and the effect of the higher order terms is negligible.

$$I(x+u, y+v, t+w) \approx I(x,y,t) + \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} w$$

$$\approx I(x,y,t) + [I_x \ I_y \ I_t] \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

where $I_x = \frac{\partial I}{\partial x}$

$I_y = \frac{\partial I}{\partial y}$

$I_t = \frac{\partial I}{\partial t}$

We have the following:

$$E(u,v,w) = \sum_{x,y,t \in W} [I(x+u, y+v, t+w) - I(x,y,t)]^2$$

$$\approx \sum_{x,y,t \in W} \left[I(x,y,t) + I_x u + I_y v + I_t w - I(x,y,t) \right]^2$$

$$\approx \sum_{x,y,t \in W} \left[I_x u + I_y v + I_t w \right]^2$$

$$= \sum_{x,y,t \in W} \left(I_x^2 u^2 + I_y^2 v^2 + I_t^2 w^2 + 2 I_x u I_y v + 2 I_y v I_t w + 2 I_x u I_t w \right)$$

$$= A u^2 + B v^2 + C w^2 + 2 D uv + 2 E vw + 2 F uw$$

for $A = \sum_{x,y,t \in W} I_x^2$, $B = \sum_{x,y,t \in W} I_y^2$, $C = \sum_{x,y,t \in W} I_t^2$, $D = \sum_{x,y,t \in W} I_x I_y$,

$E = \sum_{x,y,t \in W} I_y I_t$, $F = \sum_{x,y,t \in W} I_x I_t$

$$= \begin{bmatrix} u & v & w \end{bmatrix} \underbrace{\begin{bmatrix} A & D & F \\ D & B & E \\ F & E & C \end{bmatrix}}_{H} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

Propose a criterion to extract "3D corners":

Let the eigenvalues of matrix H be $\lambda_1, \lambda_2, \lambda_3$, where λ_1, λ_2 and λ_3 is the change in horizontal, vertical and time direction respectively. So the criterion to extract "3D corners" is have large $\lambda_1, \lambda_2, \lambda_3$.

② Different types of 3D structures that results from the variations between the eigenvalues:

* If both λ_1, λ_2 are small, it's a flat region
  - If λ_3 is small, that flat region does not change with time. Otherwise, if λ_3 is large, that flat region changes with time.
* If λ_1 is small, λ_2 is large OR λ_1 is large, λ_2 is small, it's an edge
  - If λ_3 is small, that edge does not change with time. Otherwise, if λ_3 is large, that edge changes with time

* If both λ_1, λ_2 are large, then it's a corner.
  - If λ_3 is small, that corner does not change with time. Otherwise if λ_3 is large, that corner changes with time.

My definition of a 3D corner (in a video): a 3D corner is essentially a 2D corner in terms of geometry, but the difference is that its shape and location can vary with respect to time.

For example, imagine a video of a flashlight pointing at a white cube placed on a table. At first, the flash light is turned off, so the whole video is black and nothing can be seen. Then the flash light is turned on and we can see all the corners in the video. That's an example of corner changes with respect to time. Also, imagine a video of a car moving and the camera is standing in one place, not moving with the car. The corners observed in the video is of the same shape but they vary in terms of location with respect to time.