CSC249/449MachineVision: Homework2
Zeyi Pan

Problem 1:
SIFT achieves rotation invariance by assigning orientation to key points. Therefore, as orientation is also a feature of a key point, the rotation of an image will not affect the detection of a key point. We need to compute gradient magnitude and direction of key points. Then, create a weighted direction histogram in a neighborhood of a key point. Weights are gradient magnitudes. Select the peak of the histogram as the direction of the key point.

Problem 2:
Using chain rule, we get
$$\frac{\partial L}{\partial y_i} = \frac{\partial L}{\partial p_i}\frac{\partial p_i}{\partial y_i}$$
Easy to get
$$\frac{\partial L}{\partial p_i} = -\frac{1}{p_i}$$
When $i = gt$
$$\frac{\partial p_i}{\partial y_i} = \frac{e^{yi}(\sum - e^{yi})}{\sum 2} = \frac{e^{yi}}{\sum}\frac{\sum - e^{yi}}{\sum} = p_i * (1 - p_i)$$
Therefore,
$$\frac{\partial L}{\partial y_i} = \frac{\partial L}{\partial p_i}\frac{\partial p_i}{\partial y_i} = -\frac{1}{p_i}p_i * (1 - p_i) = p_i - 1$$

When $i \neq gt$, the derivative with respect to the i-th element will always be 0.
$$\frac{\partial p_i}{\partial y_i} = \frac{e^{yi}(\sum - e^{yi})}{\sum 2} = \frac{e^{yi}}{\sum}\frac{\sum}{\sum} = p_i$$
Think about get the derivative with respect to a constant
$$\frac{\partial L}{\partial p_i} = 1$$
Therefore,
$$\frac{\partial L}{\partial y_i} = \frac{\partial L}{\partial p_i}\frac{\partial p_i}{\partial y_i} = p_i$$

Problem 3:
Using chain rule, we get

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial W}$$
$$\frac{\partial y}{\partial W} = \begin{bmatrix} \frac{\partial y_1}{\partial w_1} & \cdots & \frac{\partial y_1}{\partial w_L} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_o}{\partial w_1} & \cdots & \frac{\partial y_o}{\partial w_L} \end{bmatrix}$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial x}$$

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_L} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_O}{\partial x_1} & \cdots & \frac{\partial y_O}{\partial x_L} \end{bmatrix}$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial y}\frac{\partial y}{\partial b}$$

$$\frac{\partial y}{\partial b} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{O\times 1}$$

Problem 4:

Use chain rule, we get

$$\frac{\partial L}{\partial W_K} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial W} = \sum_k \sum_i \sum_j \frac{\partial L}{\partial y(k,i,j)} \cdot x(t, i\times S+m, j\times S+n)$$

$$\frac{\partial L}{\partial b_k} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial b} = \sum_k \sum_i \sum_j \frac{\partial L}{\partial y(k,i,j)}$$

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial x} = \sum_k \sum_i \sum_j \frac{\partial L}{\partial y(k,i,j)} \cdot W_k(t,m,n)$$

Problem 5:

$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} \cdot \frac{\partial y}{\partial x}$$

for $\frac{\partial y}{\partial x}$, $\begin{cases} 1 & \text{if } x \text{ is the max in the region} \\ 0 & \text{if } x \text{ is not the max} \end{cases}$

Therefore, $\frac{\partial L}{\partial x} = \begin{cases} \frac{\partial L}{\partial y} & \text{if } x \text{ is the max} \\ 0 & \text{if } x \text{ is not the max} \end{cases}$