MO815/MC861 - Análise de Imagem Orientada a um Problema do Mundo Real

Alexandre Xavier Falcão

Instituto de Computação - UNICAMP

afalcao@ic.unicamp.br

Introdução

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.

Introdução

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.
- Neste caso, o par $(\mathcal{A}, K)$ é denominado **kernel**.

Introdução

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.
- Neste caso, o par $(\mathcal{A}, K)$ é denominado **kernel**.
- Sendo $\mathcal{A}$ uma relação de adjacência invariante à translação, o kernel $\hat{K} = (\mathcal{A}, K)$ pode ser visto como uma imagem móvel.

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.
- Neste caso, o par $(\mathcal{A}, K)$ é denominado **kernel**.
- Sendo $\mathcal{A}$ uma relação de adjacência invariante à translação, o kernel $\hat{K} = (\mathcal{A}, K)$ pode ser visto como uma imagem móvel.
- Ao considerar cada pixel $p \in D_I$, estamos essencialmente deslocando uma imagem $\hat{K}$ sobre a imagem $\hat{I}$ para realizar a operação.

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.
- Neste caso, o par $(\mathcal{A}, K)$ é denominado **kernel**.
- Sendo $\mathcal{A}$ uma relação de adjacência invariante à translação, o kernel $\hat{K} = (\mathcal{A}, K)$ pode ser visto como uma imagem móvel.
- Ao considerar cada pixel $p \in D_I$, estamos essencialmente deslocando uma imagem $\hat{K}$ sobre a imagem $\hat{I}$ para realizar a operação.
- Nesta aula, vamos estudar o caso particular de filtros lineares, onde $\hat{I} * \hat{K}$ é uma operação de "convolução".

Introdução

- Operadores baseados em adjacência podem considerar um **peso** $K(i)$ associado a cada adjacente $q_i \in \mathcal{A}(p)$, $i = 1, 2, \ldots, |\mathcal{A}(p)|$.
- Neste caso, o par $(\mathcal{A}, K)$ é denominado **kernel**.
- Sendo $\mathcal{A}$ uma relação de adjacência invariante à translação, o kernel $\hat{K} = (\mathcal{A}, K)$ pode ser visto como uma imagem móvel.
- Ao considerar cada pixel $p \in D_I$, estamos essencialmente deslocando uma imagem $\hat{K}$ sobre a imagem $\hat{I}$ para realizar a operação.
- Nesta aula, vamos estudar o caso particular de filtros lineares, onde $\hat{I} * \hat{K}$ é uma operação de "convolução".
- No código da aula veremos suas aplicações no realce para seleção de regiões candidatas a placas de veículos.

- A relação de adjacência $\mathcal{A}$ em $\hat{K}$ tem sempre uma origem $(0,0)$ referente ao pixel p.

Kernel e Convolução

- A relação de adjacência $\mathcal{A}$ em $\hat{K}$ tem sempre uma origem $(0,0)$ referente ao pixel p.
- A diferença entre correlação e convolução está na reflexão $\hat{K}'$ de $\hat{K}$ em relação à origem.

Kernel e Convolução

- A relação de adjacência $\mathcal{A}$ em $\hat{K}$ tem sempre uma origem $(0,0)$ referente ao pixel p.
- A diferença entre correlação e convolução está na reflexão $\hat{K}'$ de $\hat{K}$ em relação à origem.
- A rigor, a convolução $\hat{I} * \hat{K}$ requer a reflexão $\hat{K}'$ de $\hat{K}$:

$$
\begin{aligned}
\hat{K}' &= (\mathcal{A}', K) \\
\mathcal{A}' &= \{(-dx_1, -dy_1), (-dx_2, -dy_2), \ldots, (-dx_d, -dy_d)\}
\end{aligned}
$$

para $d = |\mathcal{A}|$.

- A relação de adjacência $\mathcal{A}$ em $\hat{K}$ tem sempre uma origem $(0,0)$ referente ao pixel p.
- A diferença entre correlação e convolução está na reflexão $\hat{K}'$ de $\hat{K}$ em relação à origem.
- A rigor, a convolução $\hat{I} * \hat{K}$ requer a reflexão $\hat{K}'$ de $\hat{K}$:

$$\begin{aligned} \hat{K}' &= (\mathcal{A}', K) \\ \mathcal{A}' &= \{(-dx_1, -dy_1), (-dx_2, -dy_2), \ldots, (-dx_d, -dy_d)\} \end{aligned}$$

  para $d = |\mathcal{A}|$.
- Vamos assumir que $\hat{K}$ já vem refletido.

Filtragem Linear

A imagem $\hat{J} = (D_J, J)$ resultante da operação $\hat{I} * \hat{K}$ pode ser definida para todo $p \in D_I$:

$$J(p) \quad = \quad \sum_{i=1}^{d} I(q_i)K(i).$$

Algoritmo de Filtragem Linear

Entrada: $\hat{I} = (D_I, I)$ e $\hat{K} = (\mathcal{A}, K)$.
Saída: $\hat{J} = (D_J, J)$, $D_J = D_I$.

1. Para todo $p \in D_J$, faça
2. $\quad J(p) \leftarrow 0$.
3. $\quad$ Para todo $q_i \in \mathcal{A}(p)$, $1 \leq i \leq d$, tal que $q_i \in D_I$, faça
4. $\quad\quad J(p) \leftarrow J(p) + I(q_i)K(i)$.

Filtragem Gaussiana

Considerando uma adjacência circular, $q \in \mathcal{A}(p)$ se $\|q-p\|^2 \leq r^2$, de raio r, os coeficientes do filtro Gaussiano podem ser calculados por:

$$K(i) \quad = \quad \frac{\exp\left(-\frac{\|q_i-p\|^2}{2\sigma^2}\right)}{\sum_{i=1}^{d} K(i)}$$

para $i = 1, 2, \ldots, d$, onde $\sigma = r/3$ e p é o centro do disco (origem de $\mathcal{A}(p)$).

Considerando uma adjacência circular, $q \in \mathcal{A}(p)$ se $\|q - p\|^2 \leq r^2$, de raio r, os coeficientes do filtro Gaussiano podem ser calculados por:

$$K(i) \quad = \quad \frac{\exp\left(-\frac{\|q_i - p\|^2}{2\sigma^2}\right)}{\sum_{i=1}^{d} K(i)}$$

para $i = 1, 2, \ldots, d$, onde $\sigma = r/3$ e p é o centro do disco (origem de $\mathcal{A}(p)$).
Esta operação "borra" a imagem eliminando ruídos de alta frequência.

Filtros de realce de bordas destacam bordas da imagem em uma dada direção, mas podem amplificar o ruído.

- Realce por filtros de Sobel

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

Realce de Bordas

Filtros de realce de bordas destacam bordas da imagem em uma dada direção, mas podem amplificar o ruído.

- Realce por filtros de Sobel

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

- K_x e K_y realçam bordas nas direções x (bordas verticais) e y (bordas horizontais), respectivamente.

Realce de Bordas

Filtros de realce de bordas destacam bordas da imagem em uma dada direção, mas podem amplificar o ruído.

- Realce por filtros de Sobel

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

- K_x e K_y realçam bordas nas direções x (bordas verticais) e y (bordas horizontais), respectivamente.
- Sendo $\hat{G}_x = \hat{I} * \hat{K}_x$ e $\hat{G}_y = \hat{I} * \hat{K}_y$, $\vec{G}(p) = G_x(p)\vec{i} + G_y(p)\vec{j}$ é dito **vetor gradiente** em p, o qual aponta para a direção de maior crescimento de brilho na imagem $\hat{I}$ em torno de p. Sua magnitude $|\vec{G}(p)|$ pode ser usada como imagem de realce de bordas.

Realce de Bordas: Caso Geral

A magnitude de um vetor gradiente deve realçar bordas em todas as direções. De um modo mais geral:

A magnitude de um vetor gradiente deve realçar bordas em todas as direções. De um modo mais geral:
Considerando uma adjacência circular, $q \in \mathcal{A}(p)$ se $\|q - p\|^2 \leq r^2$ de raio r, o vetor gradiente $\vec{G}(p)$ pode ser calculado por

A magnitude de um vetor gradiente deve realçar bordas em todas as direções. De um modo mais geral:
Considerando uma adjacência circular, $q \in \mathcal{A}(p)$ se $\|q-p\|^2 \leq r^2$ de raio r, o vetor gradiente $\vec{G}(p)$ pode ser calculado por

$$\vec{G}(p) \quad = \quad \sum_{\forall q \in \mathcal{A}(p)} [I(q) - I(p)] \exp\left(-\frac{\|q-p\|^2}{2\sigma^2}\right) \vec{pq},$$

onde $\sigma = r/3$ e $\vec{pq} = (\frac{x_q - x_p}{\sqrt{(x_q-x_p)^2+(y_q-y_p)^2}}, \frac{y_q-y_p}{\sqrt{(x_q-x_p)^2+(y_q-y_p)^2}})$.

Normalização

A normalização de uma imagem $\hat{I} = (D_I, I)$ trata a variação de iluminação ressaltando bordas e gerando $\hat{J} = (D_J, J)$, $D_J = D_I$.

$$J(p) \;=\; \frac{I(p)}{\sqrt{\sum_{i=1}^{d} I(q_i) I(q_i)}},$$

onde $q_i \in \mathcal{A}(p)$. A relação $\mathcal{A}$ pode ser circular de raio r (e.g., $r = \sqrt{2}, \sqrt{5}, \ldots$).

Normalização

A normalização de uma imagem $\hat{I} = (D_I, I)$ trata a variação de iluminação ressaltando bordas e gerando $\hat{J} = (D_J, J)$, $D_J = D_I$.

$$J(p) \quad = \quad \frac{I(p)}{\sqrt{\sum_{i=1}^{d} I(q_i)I(q_i)}},$$

onde $q_i \in \mathcal{A}(p)$. A relação $\mathcal{A}$ pode ser circular de raio r (e.g., $r = \sqrt{2}, \sqrt{5}, \ldots$).
O denominador da normalização pode ser visto como uma convolução entre uma imagem $\hat{X}$ com os valores quadráticos de brilho da imagem $\hat{I}$ e um kernel uniforme $\hat{K}$, onde $K(i) = 1$, $i = 1, 2, \ldots, d$.

Pooling

O pooling visa ressaltar regiões com altos valores em uma dada adjacência. O pooling de uma imagem $\hat{I} = (D_I, I)$ pode considerar uma subamostragem da imagem (stride) para fins de acumular valores mais espaçados, seguida de reamostragem. Neste caso,

$$J(p) = \sqrt[\alpha]{\sum_{i=1}^{d} I(q_i)^\alpha},$$

onde $q_i \in \mathcal{A}(p)$ e $\alpha \geq 1$ controla a intensidade da operação, *i.e.*, quanto maior, mais importância será dada ao maior valor de entrada.

Pooling

O pooling visa ressaltar regiões com altos valores em uma dada adjacência. O pooling de uma imagem $\hat{I} = (D_I, I)$ pode considerar uma subamostragem da imagem (stride) para fins de acumular valores mais espaçados, seguida de reamostragem. Neste caso,

$$J(p) = \sqrt[\alpha]{\sum_{i=1}^{d} I(q_i)^{\alpha}},$$

onde $q_i \in \mathcal{A}(p)$ e $\alpha \geq 1$ controla a intensidade da operação, *i.e.*, quanto maior, mais importância será dada ao maior valor de entrada.

Similarmente, a somatória do pooling pode ser vista como uma convolução entre uma imagem $\hat{X}$ com os valores de $\hat{I}$ elevados à α e o kernel uniforme $\hat{K}$, onde $K(i) = 1$, $i = 1, 2, \ldots, d$.

- Podemos interpretar os coeficientes do kernel como elementos de um vetor $\vec{K} = (K(1), K(2), \ldots, K(d))$ e os valores dos adjacentes como elementos de outro vetor $\vec{I}(p) = (I(q_1), I(q_2), \ldots, I(q_d))$, de modo que a cada pixel p temos o produto interno $\vec{K} \cdot \vec{I}(p)$.

- Podemos interpretar os coeficientes do kernel como elementos de um vetor $\vec{K} = (K(1), K(2), \ldots, K(d))$ e os valores dos adjacentes como elementos de outro vetor $\vec{I}(p) = (I(q_1), I(q_2), \ldots, I(q_d))$, de modo que a cada pixel p temos o produto interno $\vec{K} \cdot \vec{I}(p)$.
- Seja M_K uma matriz $1 \times d$ cujos valores são os elementos de $\vec{K}$ e M_I uma matriz $d \times |D_I|$ onde os vetores $\vec{I}(p)$ são colocados em cada coluna p. A filtragem linear equivale ao produto $M_k M_I$.

Filtragem Linear: Pontos de Vista Diferentes

- Uma imagem $\hat{I} = (D_I, I)$ e uma relação de adjacência $\mathcal{A}$ definem um grafo $(D_I, \mathcal{A})$ onde os nós são os pixels e os arcos são formados por pixels adjacentes.

Filtragem Linear: Pontos de Vista Diferentes

- Uma imagem $\hat{I} = (D_I, I)$ e uma relação de adjacência $\mathcal{A}$ definem um grafo $(D_I, \mathcal{A})$ onde os nós são os pixels e os arcos são formados por pixels adjacentes.
- Neste grafo, os pesos $K(i)$ do kernel $\hat{K} = (\mathcal{A}, K)$ são os pesos $w(p, q_i)$ das arestas $(p, q_i) \in \mathcal{A}$, $i = 1, 2, \ldots, d$.

Filtragem Linear: Pontos de Vista Diferentes

- Uma imagem $\hat{I} = (D_I, I)$ e uma relação de adjacência $\mathcal{A}$ definem um grafo $(D_I, \mathcal{A})$ onde os nós são os pixels e os arcos são formados por pixels adjacentes.
- Neste grafo, os pesos $K(i)$ do kernel $\hat{K} = (\mathcal{A}, K)$ são os pesos $w(p, q_i)$ das arestas $(p, q_i) \in \mathcal{A}$, $i = 1, 2, \ldots, d$.
- Então temos que $\hat{J} = \hat{I} * \hat{K} = (D_J, J)$, onde

$$J(p) = \sum_{\forall (p,q) \in \mathcal{A}} I(q) w(p, q),$$

essencialmente calcula um peso para cada nó do grafo em função dos valores dos adjacentes e dos pesos dos arcos com esses adjacentes.

Filtragem Linear: Pontos de Vista Diferentes

- Uma imagem $\hat{I} = (D_I, I)$ e uma relação de adjacência $\mathcal{A}$ definem um grafo $(D_I, \mathcal{A})$ onde os nós são os pixels e os arcos são formados por pixels adjacentes.
- Neste grafo, os pesos $K(i)$ do kernel $\hat{K} = (\mathcal{A}, K)$ são os pesos $w(p, q_i)$ das arestas $(p, q_i) \in \mathcal{A}$, $i = 1, 2, \ldots, d$.
- Então temos que $\hat{J} = \hat{I} * \hat{K} = (D_J, J)$, onde

$$J(p) = \sum_{\forall (p,q) \in \mathcal{A}} I(q) w(p, q),$$

  essencialmente calcula um peso para cada nó do grafo em função dos valores dos adjacentes e dos pesos dos arcos com esses adjacentes.
- No caso de operadores baseados em conexidade, os caminhos entre nós do grafo são considerados.