

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

Alexandre Xavier Falcão

Instituto de Computação - UNICAMP

afalcao@ic.unicamp.br

Histograma de Gradientes Orientados

Nesta aula vamos aprender a extrair o descritor HoG para detectar placas de veículos.

Histograma de Gradientes Orientados

Nesta aula vamos aprender a extrair o descritor HoG para detectar placas de veículos.

- O primeiro passo é normalizar a imagem de entrada entre $[0, L]$ (e.g., $L = 4095$), para reduzir os efeitos da iluminação.

Histograma de Gradientes Orientados

Nesta aula vamos aprender a extrair o descritor HoG para detectar placas de veículos.

- O primeiro passo é normalizar a imagem de entrada entre $[0, L]$ (e.g., $L = 4095$), para reduzir os efeitos da iluminação.
- O segundo passo é calcular uma imagem real de magnitude de gradiente e outra imagem real de orientação de gradiente a partir da imagem normalizada.

Histograma de Gradientes Orientados

Nesta aula vamos aprender a extrair o descritor HoG para detectar placas de veículos.

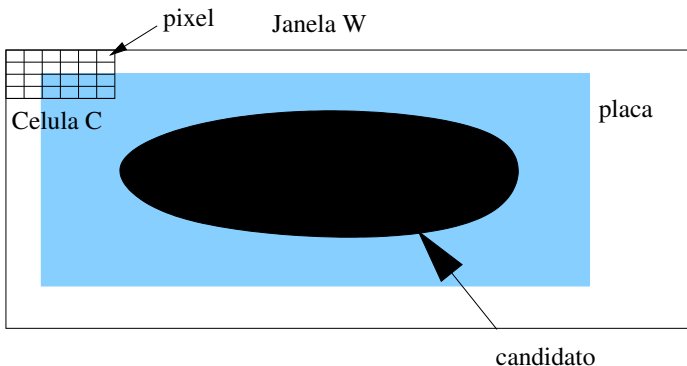
- O primeiro passo é normalizar a imagem de entrada entre $[0, L]$ (e.g., $L = 4095$), para reduzir os efeitos da iluminação.
- O segundo passo é calcular uma imagem real de magnitude de gradiente e outra imagem real de orientação de gradiente a partir da imagem normalizada.
- O terceiro passo é definir uma subimagem de $n_1 \times m_1$ pixels (**janela de detecção W**) em torno do centro de cada componente detectado anteriormente como candidato à localização da placa. Os valores de n_1 e m_1 devem ser baseados nas proporções entre largura e altura máximas das placas de treinamento, e as respectivas dimensões da imagem. Pixels adicionais de fundo em ambas direções devem ser considerados.

Histograma de Gradientes Orientados

- Cada janela de detecção deve ainda ser dividida em células C de $n_2 \times m_2$ pixels de modo que W contenha um número inteiro de células.

Histograma de Gradientes Orientados

- Cada janela de detecção deve ainda ser dividida em células C de $n_2 \times m_2$ pixels de modo que W contenha um número inteiro de células.



Histograma de Gradientes Orientados

- O passo seguinte gera um histograma de b orientações (bins) de gradientes para cada célula, de acordo com intervalos θ graus. Normalmente, $b = 9$, onde o primeiro bin indica magnitude de gradiente zero e os 8 restantes indicam os intervalos de orientação $\theta = 45$ graus:
 $0 - 44, 45 - 89, \dots, 315 - 359$.

Histograma de Gradientes Orientados

- O passo seguinte gera um histograma de b orientações (bins) de gradientes para cada célula, de acordo com intervalos θ graus. Normalmente, $b = 9$, onde o primeiro bin indica magnitude de gradiente zero e os 8 restantes indicam os intervalos de orientação $\theta = 45$ graus:
 $0 - 44, 45 - 89, \dots, 315 - 359$.
- O valor associado a um bin do histograma é denominado voto e provém das magnitudes dos gradientes dos pixels cujas orientações contribuem para o valor do bin.

Histograma de Gradientes Orientados

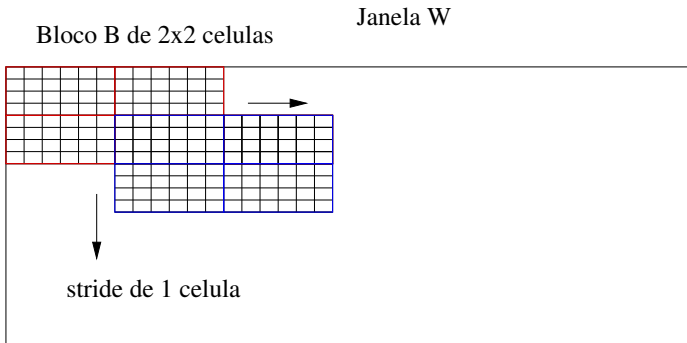
- O passo seguinte gera um histograma de b orientações (bins) de gradientes para cada célula, de acordo com intervalos θ graus. Normalmente, $b = 9$, onde o primeiro bin indica magnitude de gradiente zero e os 8 restantes indicam os intervalos de orientação $\theta = 45$ graus:
 $0 - 44, 45 - 89, \dots, 315 - 359$.
- O valor associado a um bin do histograma é denominado voto e provém das magnitudes dos gradientes dos pixels cujas orientações contribuem para o valor do bin.
- O histograma de cada célula deve ainda ser normalizado com relação à energia total dos histogramas de grupos de células, denominados **blocos**, para tratar variações de contraste entre imagens.

Histograma de Gradientes Orientados

- Isto é, cada bloco consiste de $n_3 \times m_3$ células e o stride (deslocamentos em x e y) entre blocos adjacentes pode implicar em sobreposição de células.

Histograma de Gradientes Orientados

- Isto é, cada bloco consiste de $n_3 \times m_3$ células e o stride (deslocamentos em x e y) entre blocos adjacentes pode implicar em sobreposição de células.



Histograma de Gradientes Orientados

- Ao final, os histogramas das células normalizados em cada bloco são concatenados para todos os blocos gerados por deslocamentos na janela de detecção, formando o **descritor HoG**.

Histograma de Gradientes Orientados

- Ao final, os histogramas das células normalizados em cada bloco são concatenados para todos os blocos gerados por deslocamentos na janela de detecção, formando o **descritor HoG**.
- Por exemplo, se a janela for de 126×36 pixels, as células forem de 6×6 pixels, então cada janela terá 21×6 células.

Histograma de Gradientes Orientados

- Ao final, os histogramas das células normalizados em cada bloco são concatenados para todos os blocos gerados por deslocamentos na janela de detecção, formando o **descritor HoG**.
- Por exemplo, se a janela for de 126×36 pixels, as células forem de 6×6 pixels, então cada janela terá 21×6 células.
- Se os blocos forem de 2×2 células e os deslocamentos dos blocos forem de 1 célula nas direções x e y , cada janela produzirá 20×5 blocos.

Histograma de Gradientes Orientados

- Ao final, os histogramas das células normalizados em cada bloco são concatenados para todos os blocos gerados por deslocamentos na janela de detecção, formando o **descritor HoG**.
- Por exemplo, se a janela for de 126×36 pixels, as células forem de 6×6 pixels, então cada janela terá 21×6 células.
- Se os blocos forem de 2×2 células e os deslocamentos dos blocos forem de 1 célula nas direções x e y , cada janela produzirá 20×5 blocos.
- Os 4 histogramas de 9 bins de cada bloco são concatenados e normalizados para formar um descritor de 36 características por bloco.

Histograma de Gradientes Orientados

- Ao final, os histogramas das células normalizados em cada bloco são concatenados para todos os blocos gerados por deslocamentos na janela de detecção, formando o **descritor HoG**.
- Por exemplo, se a janela for de 126×36 pixels, as células forem de 6×6 pixels, então cada janela terá 21×6 células.
- Se os blocos forem de 2×2 células e os deslocamentos dos blocos forem de 1 célula nas direções x e y , cada janela produzirá 20×5 blocos.
- Os 4 histogramas de 9 bins de cada bloco são concatenados e normalizados para formar um descritor de 36 características por bloco.
- Depois os descritores dos blocos são concatenados para formarem o descritor da janela. Portanto, o descritor HoG da janela W terá $20 \times 5 \times 36$ características.

Normalização da imagem original

A normalização de uma imagem $\hat{I} = (D_I, I)$ deve gerar uma imagem $\hat{J} = (D_J, J)$, $D_J = D_I$, cujos valores podem ser definidos por

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

onde $q_i \in \mathcal{A}(p)$ e $d = |\mathcal{A}(p)|$ (**A adjacência \mathcal{A} pode ser circular de raio $r \geq 1$.**).

Normalização da imagem original

A normalização de uma imagem $\hat{I} = (D_I, I)$ deve gerar uma imagem $\hat{J} = (D_J, J)$, $D_J = D_I$, cujos valores podem ser definidos por

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

onde $q_i \in \mathcal{A}(p)$ e $d = |\mathcal{A}(p)|$ (**A adjacência \mathcal{A} pode ser circular de raio $r \geq 1$.**). Ou pode seguir a sugestão padrão:

$$J(p) = \frac{L[I(p) - I_{\min}]}{[I_{\max} - I_{\min}]},$$

onde I_{\min} e I_{\max} são os valores mínimo e máximo da imagem \hat{I} .

Vetor Gradiente e sua Orientação

Considerando uma adjacência circular \mathcal{A} de raio $r \geq 1$, o vetor gradiente $\vec{G}(p)$ de cada pixel $p \in D_I$ pode ser calculado por

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

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

Vetor Gradiente e sua Orientação

Considerando uma adjacência circular \mathcal{A} de raio $r \geq 1$, o vetor gradiente $\vec{G}(p)$ de cada pixel $p \in D_I$ pode ser calculado por

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

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

A opção padrão é mais simples e adota a convolução com kernels horizontal $[-1, 0, 1]$ e vertical $[-1, 0, 1]^t$ para os componentes $G_x(p)\vec{i}$ e $G_y(p)\vec{j}$ de $\vec{G}(p) = G_x(p)\vec{i} + G_y(p)\vec{j}$.

Vetor Gradiente e sua Orientação

Sejam $g_x(p) = \frac{G_x(p)}{\|\vec{G}(p)\|}$ e $g_y(p) = \frac{G_y(p)}{\|\vec{G}(p)\|}$ os componentes do vetor gradiente normalizado. A orientação $\Theta(p)$ de $\vec{G}(p)$ pode ser obtida por

$$\Theta(p) = \begin{cases} \frac{180}{\pi} \cos^{-1}(g_x(p)) & \text{se } g_y(p) \geq 0 \\ 360 - \frac{180}{\pi} \cos^{-1}(g_x(p)) & \text{se } g_y(p) < 0 \end{cases}$$

- Considere as coordenadas (x, y) dos quatro centros $q_1, q_2, q_3,$ e q_4 de células da janela mais próximas da coordenada (x_p, y_p) do pixel $p \in D_I$.

Histograma de cada célula

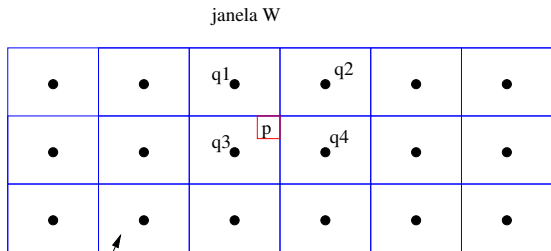
- Considere as coordenadas (x, y) dos quatro centros $q_1, q_2, q_3,$ e q_4 de células da janela mais próximas da coordenada (x_p, y_p) do pixel $p \in D_I$.
- Considere ainda as orientações b_1 e b_2 (centros de bins adjacentes) mais próximas da orientação $\Theta(p)$.

Histograma de cada célula

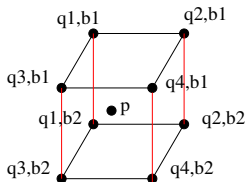
- A magnitude de $\vec{G}(p)$ deve ser distribuída (votos) entre os bins b_1 e b_2 das quatro células adjacentes $q_1, q_2, q_3,$ e q_4 por interpolação trilinear.

Histograma de cada célula

- A magnitude de $\vec{G}(p)$ deve ser distribuída (votos) entre os bins b_1 e b_2 das quatro células adjacentes $q_1, q_2, q_3,$ e q_4 por interpolação trilinear.



Celula



Histograma de cada célula

- Por exemplo, para $\Theta = 30$ graus, $b_1 = 22$ e $b_2 = 67$, já que as orientações dos 8 bins com gradientes diferentes de zero são 22, 67, 112, 157, 202, 247, 292, e 337, respectivamente.

Histograma de cada célula

- Por exemplo, para $\Theta = 30$ graus, $b_1 = 22$ e $b_2 = 67$, já que as orientações dos 8 bins com gradientes diferentes de zero são 22, 67, 112, 157, 202, 247, 292, e 337, respectivamente.
- Esta distribuição de votos procura tratar o seguinte fato: "pixels relevantes com alta magnitude de gradiente podem cair em células vizinhas, dependendo da imagem".

Histograma de cada célula

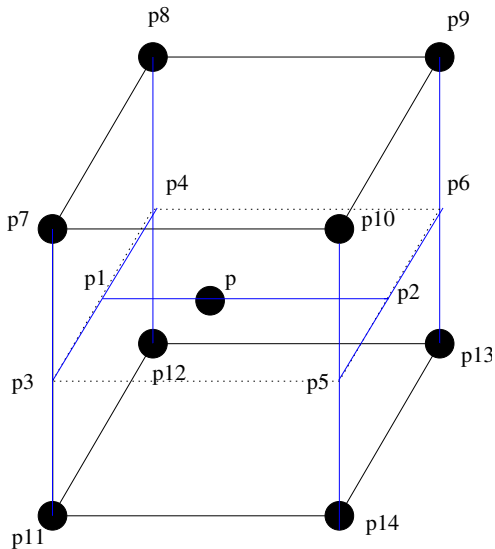
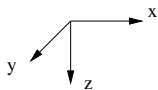
- Por exemplo, para $\Theta = 30$ graus, $b_1 = 22$ e $b_2 = 67$, já que as orientações dos 8 bins com gradientes diferentes de zero são 22, 67, 112, 157, 202, 247, 292, e 337, respectivamente.
- Esta distribuição de votos procura tratar o seguinte fato: "pixels relevantes com alta magnitude de gradiente podem cair em células vizinhas, dependendo da imagem".
- Conhecemos as coordenadas (x_i, y_i, z_i) dos 8 vértices $(q_1, b_1), (q_2, b_1), (q_3, b_1), (q_4, b_1), (q_1, b_2), (q_2, b_2), (q_3, b_2), (q_4, b_2)$, $i = 1, 2, \dots, 8$, a coordenada (x_p, y_p, z_p) , $z_p = \Theta(p)$, e a magnitude $\|\vec{G}(p)\|$. A interpolação trilinear deve distribuir o peso $w = \|\vec{G}(p)\|$ entre os 8 vértices.

- Isto é, acumular voto no bin (orientação z_i) correspondente do histograma da célula representada por (x_i, y_i) , $i = 1, 2, \dots, 8$.

Histograma de cada célula

- Isto é, acumular voto no bin (orientação z_i) correspondente do histograma da célula representada por (x_i, y_i) , $i = 1, 2, \dots, 8$.
- O peso $w = \|\vec{G}(p)\|$ deve ser distribuído entre pontos de faces opostas, p_1 e p_2 , depois os pesos nas faces devem ser distribuídos entre pontos de arestas opostas, p_3, p_4, p_5, p_6 , e por fim estes distribuídos entre os vértices da aresta correspondente, $p_7, p_8, p_9, p_{10}, p_{11}, p_{12}, p_{13}$, e p_{14} .

Histograma de cada célula



Histograma de cada célula

Os pesos w_i em cada ponto $p_i = (x_{p_i}, y_{p_i}, z_{p_i})$, $i = 1, 2, \dots, 14$ são calculados por

$$w_1 = w \frac{(x_{p_2} - x_p)}{(x_{p_2} - x_{p_1})}$$

$$w_2 = w \frac{(x_p - x_{p_1})}{(x_{p_2} - x_{p_1})}$$

$$w_3 = w_1 \frac{(y_{p_1} - y_{p_4})}{(y_{p_3} - y_{p_4})}$$

$$w_4 = w_1 \frac{(y_{p_3} - y_{p_1})}{(y_{p_3} - y_{p_4})}$$

Histograma de cada célula

$$w_5 = w_2 \frac{(y_{p_2} - y_{p_6})}{(y_{p_5} - y_{p_6})}$$

$$w_6 = w_2 \frac{(y_{p_5} - y_{p_2})}{(y_{p_5} - y_{p_6})}$$

$$w_7 = w_3 \frac{(z_{p_{11}} - z_{p_3})}{(z_{p_{11}} - z_{p_7})}$$

$$w_{11} = w_3 \frac{(z_{p_3} - z_{p_7})}{(y_{p_{11}} - z_{p_7})}$$

$$w_8 = w_4 \frac{(z_{p_{12}} - z_{p_4})}{(z_{p_{12}} - z_{p_8})}$$

$$w_{12} = w_4 \frac{(z_{p_4} - z_{p_8})}{(z_{p_{12}} - z_{p_8})}$$

Histograma de cada célula

$$w_{10} = w_5 \frac{(z_{p_{14}} - z_{p_5})}{(z_{p_{14}} - z_{p_{10}})}$$

$$w_{14} = w_5 \frac{(z_{p_5} - z_{p_{10}})}{(z_{p_{14}} - z_{p_{10}})}$$

$$w_9 = w_6 \frac{(z_{p_{13}} - z_{p_6})}{(z_{p_{13}} - z_{p_9})}$$

$$w_{13} = w_6 \frac{(z_{p_6} - z_{p_9})}{(z_{p_{13}} - z_{p_9})}$$

Por fim, os pesos w_i são acumulados no bin correspondente do histograma da célula que contém o vértice p_i , para os casos $i = 7, 8, 9, 10, 11, 12, 13, 14$.

Descritor de bloco e sua normalização

- Seja B um bloco com $n_3 \times m_3$ células. Cada célula tem um histograma $h_k(i)$ de b bins, $i = 0, 1, \dots, b - 1$ e $k = 1, 2, \dots, n_3 \times m_3$.

Descritor de bloco e sua normalização

- Seja B um bloco com $n_3 \times m_3$ células. Cada célula tem um histograma $h_k(i)$ de b bins, $i = 0, 1, \dots, b - 1$ e $k = 1, 2, \dots, n_3 \times m_3$.
- A concatenação dos histogramas das células, varrendo elas da esquerda para direita e de cima para baixo, gera um vetor de características v_j , $j = 1, 2, \dots, b \times n_3 \times m_3$.

Descritor de bloco e sua normalização

- Seja B um bloco com $n_3 \times m_3$ células. Cada célula tem um histograma $h_k(i)$ de b bins, $i = 0, 1, \dots, b - 1$ e $k = 1, 2, \dots, n_3 \times m_3$.
- A concatenação dos histogramas das células, varrendo elas da esquerda para direita e de cima para baixo, gera um vetor de características v_j , $j = 1, 2, \dots, b \times n_3 \times m_3$.
- Este vetor deve ser normalizado para tratar variações de contraste entre imagens

$$v_j = \frac{v_j}{\sqrt{\sum_{j=1}^{b \times n_3 \times m_3} v_j v_j + \epsilon}}$$

onde ϵ é um número muito pequeno, para evitar divisão por zero.

Coleta dos descritores dos blocos

Por fim, os vetores de características de cada bloco são concatenados, varrendo-os da esquerda para direita e de cima para baixo para gerar o descritor HoG da janela.

