

Sisteme de Recunoastere a Formelor – Laborator 3-4

Histograma Orientarilor Gradientilor

1. Obiective

Descriptorii de tip **histograma a orientarii gradientilor**, sau descriptori HOG, sunt descriptori de trasatori folositi in viziunea artificiala si in procesarea imaginilor in scopul detectiei obiectelor. Aceasta tehnica numara de cate ori apare o orientare anume a gradientului intr-o anumita regiune din imagine.

Scopul primei lucrari de laborator (lab 3) este de a implementa algoritmul de extragere a trasaturilor de tip HOG. Aceste trasaturi vor fi utilizate in urmatoarea lucrare de laborator (lab 4) pentru clasificarea si recunoasterea obiectelor.

2. Fundamente teoretice

Trasaturile HOG au fost introduse de Navneed Dalal si Bill Triggs [1] care au dezvoltat si au testat mai multe variante de descriptori HOG, cu diferite organizari spatiale, diferite metode de calcul a gradientilor, si diferite metode de normalizare.

Ideea esentiala care sta la baza descriptorilor HOG este ca aspectul local si forma unui obiect intr-o imagine poate fi descrisa de distributia intensitatii gradientilor si de distributia orientarii muchiilor. Implementarea acestor descriptori poate fi obtinuta prin impartirea imaginii in regiuni conexe mai mici, numite celule, si calcularea, pentru fiecare celula, a histogramei orientarii muchiilor (sau a directiei gradientilor) pe baza pixelilor apartinand celulei. Combinarea histogramelor din celule reprezinta descriptorul imaginii. Pentru o mai buna performanta, histogramele locale pot fi normalizate prin calcularea unei masuri a intensitatii pe o regiune mai mare a imaginii, numita bloc, si apoi folosind aceasta masura pentru normalizarea tuturor celulelor din acel bloc. Aceasta normalizare are ca efect o invarianta sporita la schimbarea iluminarii, sau la umbre.

3. Implementarea algoritmilor

3.1. Calculul gradientului

Primul pas pentru obtinerea HOG este calculul valorilor gradientului. Cea mai folosita metoda este aplicarea mastilor de derivare unidimensionala pe directie verticala si pe directie orizontala, adica filtrarea imaginii cu urmatoarele doua nuclee de convolutie:

$$D_x = [-1 \ 0 \ 1] \text{ and } D_y = [-1 \ 0 \ 1]^T$$



Figura 3.1 Imagine initiala

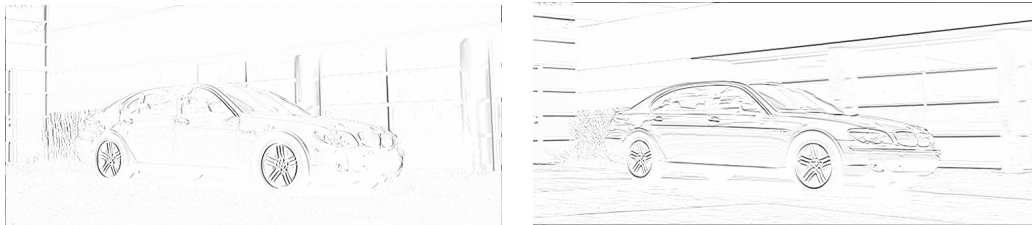


Figura 3.2 stanga: derivate pe orizontala; dreapta – derivate pe verticala

Dandu-se o imagine I , vom obtine derivate orizontala si derivate verticala folosind un operator de convolutie: $I_x = I * D_x$ si $I_y = I * D_y$

Magnitudinea gradientului este $|G| = \sqrt{I_x^2 + I_y^2}$

Orientarea gradientului este: $\theta = \arctan \frac{I_y}{I_x}$

3.2. Discretizarea orientarii

Al doilea pas consta in crearea histogramelor pentru fiecare celula. Fiecare pixel din fiecare celula are un vot in histograma. Pozitia in histograma care va fi incrementata de un anume pixel este data de valoarea orientarii gradientului acestui pixel. Dalal si Triggs au gasit ca este suficient ca histograma sa aiba 9 pozitii, iar daca orientarea gradientului ia valori intre 0 si 360 grade, pozitia in histograma este calculata prin impartirea cu 40. In procesul de votare fiecare vot poate fi ponderat cu valoarea magnitudinii gradientului, cu radacina patrata sau cu patratul magnitudinii.



Figura 3.3 Imaginea initiala



Figura 3.4 Magnitudinea gradientului

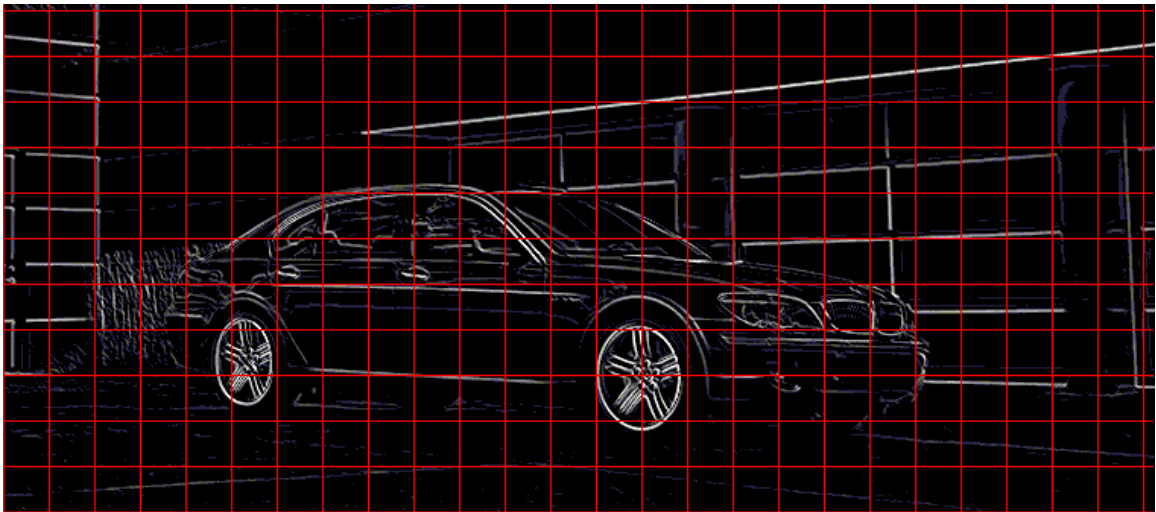


Figura 3.5 Impartirea in celule

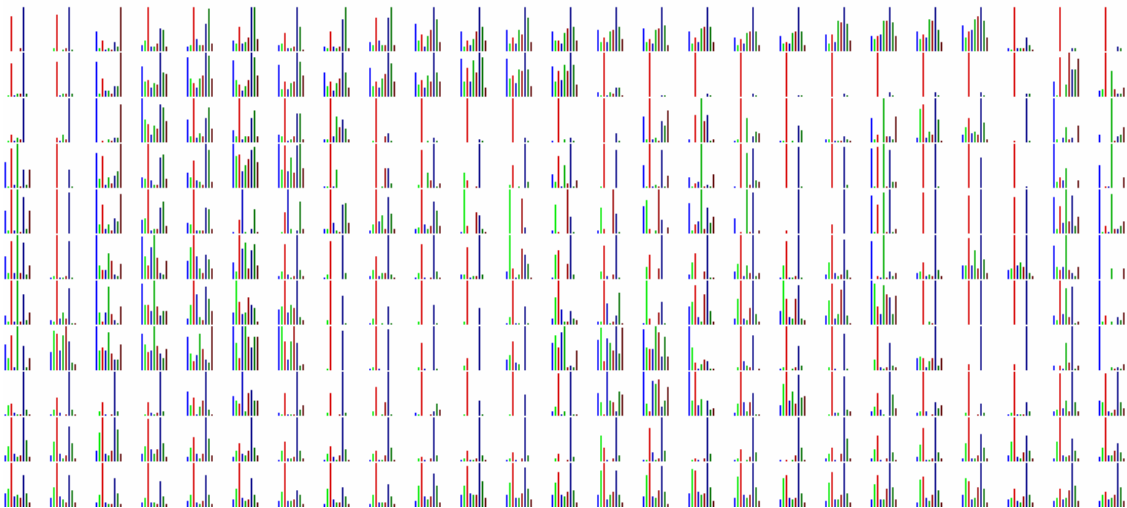


Figura 3.6 Histograma orientarii gradientului

3.3. Blocuri de descriptori

Pentru a tine cont de variatiile de iluminare si contrast, magnitudinea gradientilor trebuie normalizata local, ceea ce inseamna gruparea celulelor in structuri spatiale mai mari, numite blocuri. Descriptorul HOG este vectorul format din componentele normalizate ale histogramelor din fiecare celula din bloc. De obicei blocurile se suprapun partial, astfel ca fiecare celula contribuie de mai multe ori la descriptorul final. Exista doua tipuri principale de blocuri: blocul rectangular, R-HOG, si blocul circular, C-HOG. Blocurile R-HOG sunt grile patrate, definite de trei parametri: numarul de celule din bloc, numarul de pixeli intr-o celula, si numarul de pozitii in histograma unei celule.

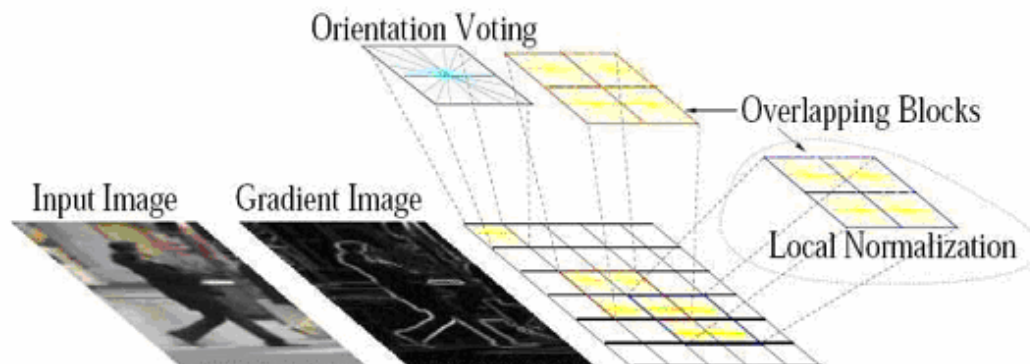
3.4. Normalizarea blocurilor

Exista mai multe metode de normalizare a blocurilor. Daca notam cu v vectorul ne-normalizat ce contine toate histogramele unui bloc, iar $\|v_k\|$ este norma k pentru $k = 1, 2$ si e este o constanta de valoare foarte mica (ce nu va influenta rezultatele), factorul de normalizare se calculeaza cu una din ecuatiile de mai jos:

$$\text{L2-norm: } f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$$

$$\text{L1-norm: } f = \frac{v}{\|v\|_1 + e}$$

$$\text{L1-sqrt: } f = \sqrt{\frac{v}{\|v\|_1 + e}}$$



4. Masura de similaritate

Pentru a se masura similaritatea dintre doi vectori, se pot utiliza mai multe metrici.

1. Distanța Euclidiană

Având doi vectori $P = (p_1, p_2, \dots, p_n)$ și $Q = (q_1, q_2, \dots, q_n)$ distanța este:

$$d = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}$$

2. **Cosinusul unghiului dintre vectori:** asemanarea dintre doi vectori este data de cosinusul unghiului dintre ei. Având doi vectori A și B , cosinusul unghiului este dat de produsul lor scalar, normalizat cu produsul magnitudinilor lor.

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$$

Ecuatia de mai sus se poate detalia in functie de attributele fiecarui vector, $A = (p_1, p_2, \dots, p_n)$ și $B = (q_1, q_2, \dots, q_n)$:

$$\cos(\theta) = \frac{p_1 q_1 + p_2 q_2 + \dots + p_n q_n}{\sqrt{p_1^2 + p_2^2 + \dots + p_n^2} \cdot \sqrt{q_1^2 + q_2^2 + \dots + q_n^2}}$$

5. Detalii de implementare

Având o imagine, gradientul se calculeaza gradientul cu magnitudine: $|G| = \sqrt{I_x^2 + I_y^2}$, unde:

$$I_x = I * D_x,$$

$$I_y = I * D_y,$$

$$D_x = [-1 \ 0 \ 1],$$

$$D_y = [-1 \ 0 \ 1]^T,$$

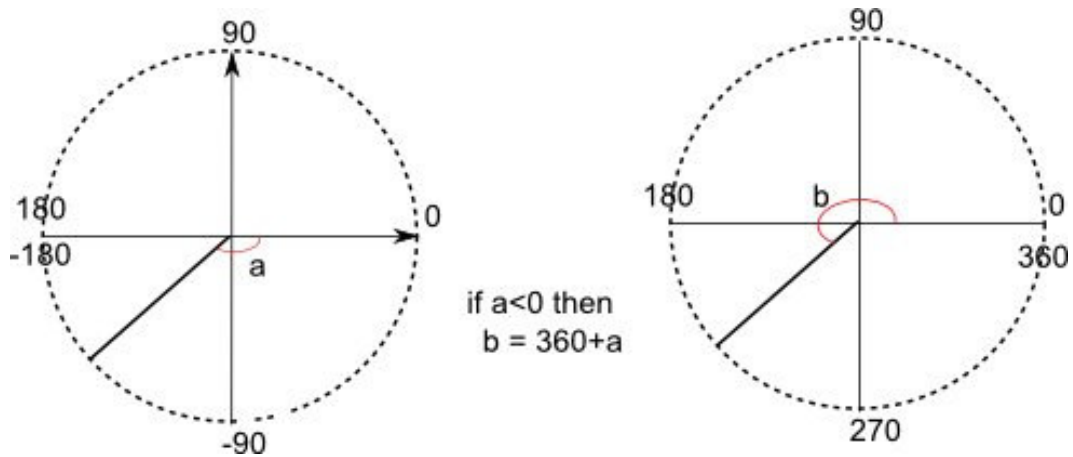
* este operatorul de convolutie

Pentru orientare, se foloseste functia *atan2* care returneaza valori in intervalul $[-\pi, \pi]$. Astfel, orientarea gradientului este $\theta = \text{atan2}(I_y, I_x)$ radiani.

Unghiul transformat in grade este $\alpha = \theta * 180 / \pi$, cu valori in intervalul $[-180, 180]$ grade.

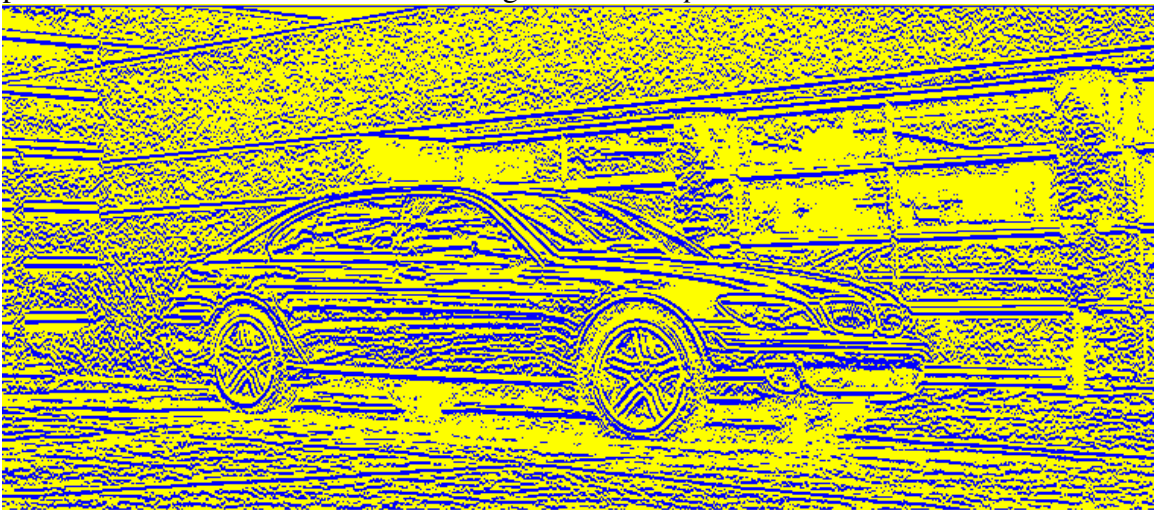
Valorile orientarii in intervalul $[-180, 180]$ se transforma in valori in intervalul $[0, 360]$ astfel:

$$\alpha_{\text{signed}} = \begin{cases} \alpha, & \text{if } \alpha \geq 0 \\ \alpha + 360, & \text{if } \alpha < 0 \end{cases}$$

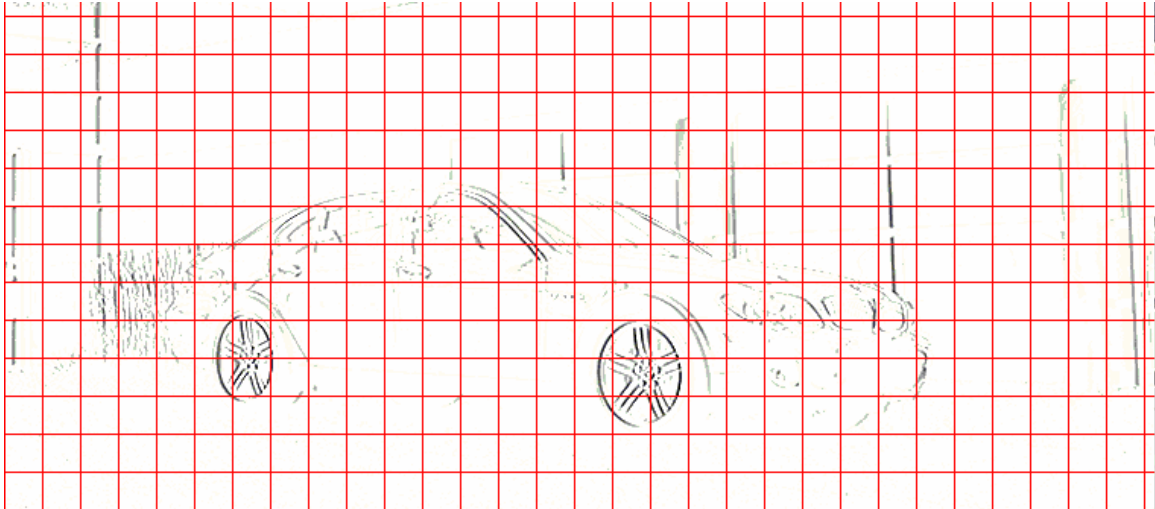


6. Activitate practica

1. Adaugati in Diblock o noua functie, care va calcula magnitudinea si orientarea gradientilor cu formulele date.
2. Adaugati o noua functie, care va atasa fiecarui pixel cu valoarea gradientului intre 0 si 180 grade culoarea galbena, si fiecarui pixel cu valoarea gradientului intre 180 si 360 grade culoarea albastra. Este posibil sa fie nevoie de modificarea paletei imaginilor pentru a reda culorile alese. Pentru imaginea *car1.bmp* rezultatul va arata cam asa:

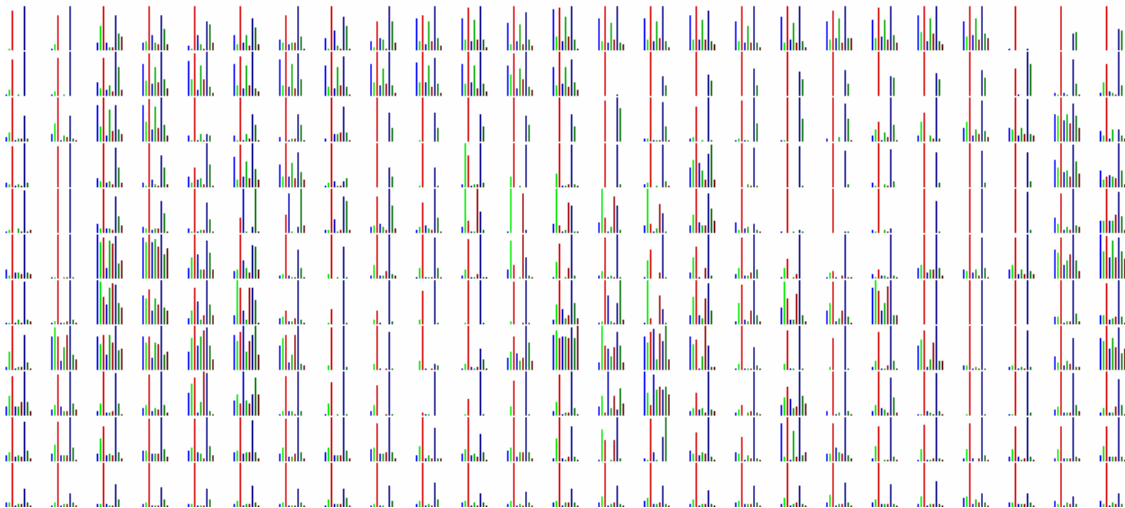


3. Implementati o functie care divide imaginea in celule (desenati liniile de divizare). Va trebui sa implementati o caseta de dialog de unde sa cititi valorile pentru dimensiunile celulelor (inaltime si latime). Pentru imaginea *car1.bmp* rezultatul asteptat este:



4. Implementati o functie care calculeaza trasaturile HOG pentru o imagine data, si le afiseaza pe imagine. Nu folositi normalizare, doar pasii 1 si 2 din algoritmul prezentat in sectiunea teoretica. Pentru aceasta operatie, implementati o caseta de dialog prin care veti citi latimea, inaltimea, si numarul de valori in histograma unei celule.

Rezultatul de mai jos este obtinut pe imaginea *car1.bmp* cu parametrii $nr_valori_histograma = 9$, $latime_celula = 30$, $inaltime_celula = 30$:



5. Fiind data imaginea *hog_similar.bmp* ce contine doua obiecte similare, calculati gradul de asemanare dintre obiecte, folosind metrica bazata pe cosinusul unghiului. Coordonatele celor doua obiecte din imagine sunt:
- obiect 1: $x_bottom = 0$, $y_bottom = 0$, $x_top = 63$, $y_top = 127$
 - obiect 2: $x_bottom = 90$, $y_bottom = 0$, $x_top = 153$, $y_top = 127$
- Folositi urmasorii parametri: $nr_valori_histograma = 12$, $latime_celula = 64$, $inaltime_celula = 128$ (dimensiunea celulei este egala cu dimensiunea obiectului).

6. Dandu-se imaginea *hog_different.bmp* ce contine doua obiecte diferite, calculati gradul de asemanare dintre ele, folosind metrica bazata pe cosinus.
Coordonatele celor doua obiecte din imagine sunt:
 - a. obiect 1: $x_{bottom} = 0$, $y_{bottom} = 0$, $x_{top} = 63$, $y_{top} = 127$
 - b. obiect 2: $x_{bottom} = 90$, $y_{bottom} = 0$, $x_{top} = 153$, $y_{top} = 127$Folositi urmasorii parametri: $nr_valori_histograma = 12$, $latime_celula = 64$, $inaltime_celula = 128$ (dimensiunea celulei este egala cu dimensiunea obiectului).
7. Aceleasi cerinte ca la punctul 5 dar folosind alt set de parametri (mai putine valori in histograma, mai multe celule pentru obiecte): $nr_valori_histograma = 9$, $latime_celula = 16$, $inaltime_celula = 16$. Pentru imaginea *hog_similar.bmp* valoarea de similaritate cosinus folosind parametri specificati este 0.785446.
8. Aceleasi cerinte ca la punctul 6 dar folosind alt set de parametri (mai putine valori in histograma, mai multe celule pentru obiecte): $nr_valori_histograma = 9$, $latime_celula = 16$, $inaltime_celula = 16$. Pentru imaginea *hog_different.bmp* valoarea de similaritate cosinus folosind parametri specificati este 0.215457.

7. Bibliografie

[1] Dalai, N. Triggs, B. : *Histograms of oriented gradients for human detection*, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.