

Chapter 1

Histograma de orient'ari a gradientului

Histograma de orient'ari a gradientului (HOG) a fost introdus'a de c'atre Dalal 'si Triggs¹ 'si este un alt mod de a reprezenta o imagine de intrare printr-un vector de descriptori vizuali. HOG-ul utilizeaz'a magnitudinea 'si orientarea gradientului 'in jurul loca'tiei unui punct cheie pentru a construi o histogram'a. O imagine (sau regiune) de descris este 'imp'ar'tit'a 'in zone disjuncte numite celule. Pentru fiecare celul'a se va calcula o histogram'a de tip HOG; ulterior, histogramele celulelor se vor concatena pentru a forma descriptorul final.

'In mod tradi'tional, descriptorul HOG se aplic'a pe imagini cu un singur plan de culoare. Dac'a totu'si se dore'ste utilizarea pe o imagine color, atunci gradientul cu magnitudine maxim'a ('si orientarea corespunz'atoare) este ales.

Algoritmul tipic de calcul a HOG cuprinde urm'atorii pa'si:

1. *Preprocesare.* Acest pas este unul facultativ. Se poate utiliza o transformare neliniar'a a contrastului pentru a da mai mult'a importan't'a gradien'tilor dintr-o anumit'a gam'a de intensitate.
2. *Calculul gradientului.* Algoritmul necesit'a determinarea at'zt a gradientului vertical c'zt 'si a celui orizontal. Acest lucru se poate realiza prin convolu'tie cu:

- Nuclee centrate (figura 1.1 subgraficul (a)). Acesta este modul recomandat.
- Nuclee necentrate (figura 1.1 subgraficul (b)).
- M'a'sti de tip Sobel (figura 1.1 subgraficul (c)). 'In mod normal avantajul acestui solu'tii pentru identificarea gradientului este o mai mare robuste'te la zgomot. Totu'si 'in condi'tiile 'in care HOG implic'a calculul unei histograme, deci a unei medieri statistice, acest avantaj devine derizoriu.

'In acest moment, pentru fiecare pixel (i, j) , av'znd valoarea gradientului pe orizontal'a, $G_H(i, j)$ 'si respectiv pe vertical'a $G_V(i, j)$, se poate calcula magnitudinea M_G 'si respectiv orientarea O_G :

$$M_G(i, j) = \sqrt{G_H^2(i, j) + G_V^2(i, j)} \quad (1.1)$$

¹Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proceedings of the Computer Vision and Pattern Recognition, p. 886-893 (2005)

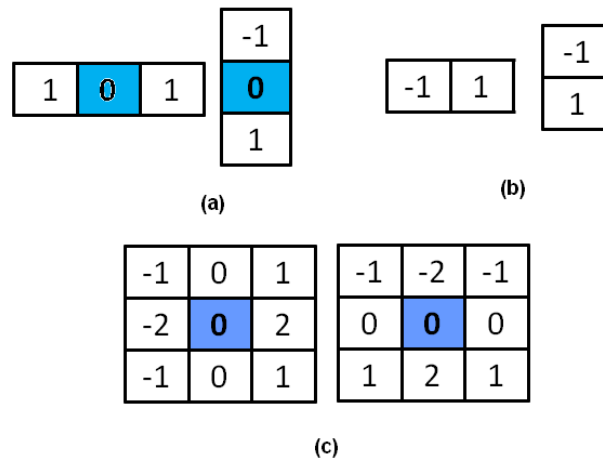


Figure 1.1: Cele trei tipuri standard de nuclee ce pot fi utilizate pentru determinarea gradien'tilor 'in descriptorul HOG: nuclee centrate - (a), nuclee necentrate (b) 'si m'a'sti de tip Sobel (c).

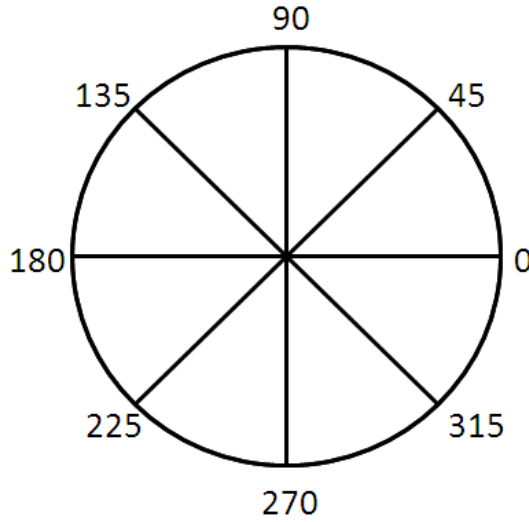


Figure 1.2: 'Imp'ar'tirea domeniului unghiular 'in 8 intervale echidistante ce vor forma binurile din histograma de orient'ari din descriptorul HOG.

$$O_G(i, j) = \arctan \frac{G_V(i, j)}{G_H(i, j)} \quad (1.2)$$

3. *Histograma de orient'ari.* Conform ecua'tiei 1.2, ini'tial orientarea se poate calcula cu precizie continu'a. Totu'si pentru calculul histogramei, se define'ste un numar, N_o 'si un set de intervale unghiulare. Acestea pot fi cu semn (acoper'a intervalul $[0^0; 360^0]$) sau f'ar'a semn (sunt 'in intervalul $[0^0; 180^0]$). Tipic, intervalele sunt de l'a'timi egale, a'sa cum se exemplific'a 'si 'in figura 1.2.

'In varianta original'a, valoarea continu'a a orient'arii este 'imp'ar'tit'a la cele mai apropiate dou'a binuri. Votul fiec'arui pixel poate fi:

- 'In ceea ce prive'ste valoarea:
 - Este 1, dac'a gradientul este suficient de puternic (adic'a magnitudinea M_G este peste un prag).
 - Depinde continuu de M_G adic'a poate fi egal sau propor'tional sau extras cu ajutorul unei func'tii din valoarea magnitudinii.
- 'In ceea ce prive'ste distribu'tia unui vot c'atre binurile histogramei, aceasta poate fi:
 - Invers propor'tional'a cu distan'ta p'zn'a la centrul binului 'si pixelul curent contribuie la cele mai apropiate dou'a.
 - Contribuie doar la binul 'in interiorul c'arui interval pic'a. Trebuie observat, c'a 'in special 'in acest caz, se poate evita calculul func'tiei arctan din ecua'tia 1.2 (care este destul de costisitor) prin utilizarea unei tabele de asociere 'intre intervalul unghiular 'si valoarea raportului 'intre gradientul pe vertical'a 'si cel pe orizontal'a (adic'a s'a se aproximeze func'tia arctan).

4. *Blocul descriptiv.* Imaginea sau regiunea de descris se 'imparte 'in celule astfel 'inc'zt fiecare pixel s'a apar'tin'a unei celule 'si numai uneia singure. Pentru fiecare asemenea celul'a se poate calcula histograma de orient'ari. Scopul celulei nu este altul dec'zt de a facilita calculul histogramei de orient'ari. Celulele pot fi rectangulare sau radiale (log-polare).

Pentru descrierea de orient'ari se consider'a un num'ar de blocuri. Spre deosebire de celule, blocurile pot fi suprapuse sau pot evita complet o anumit'a por'tiune din imagine. Celulele fiec'arui bloc sunt puse 'impreun'a 'si formeaz'a histograma de orient'ari a blocului. "Pus 'impreun'a", aici, semnific'a utilizarea pixelilor din celulele constituyente ca 'si cum ar face parte dintr-o singur'a celul'a; 'in condi'tiile 'in care histograma de orient'ari pentru o celul'a este nenormalizat'a, histograma de orient'ari a unui bloc se ob'tine prin 'insumarea, component'a cu component'a, a histogramelor celulelor constituyente. Un exemplu pentru a ar'ata semnifica'tia blocurilor 'si respectiv a celulelor se poate vedea 'in figura 1.3.

5. *Normalizarea blocurilor.* Pentru fiecare bloc, 'in acest moment, exist'a un descriptor de tip histogram'a de orient'ari, pe care 'il vom nota cu ν . Num'arul de elemente al acestui descriptor este egal cu num'arul de orient'ari N_o considerate la pasul 3. Pentru o mai bun'a reprezentare a imaginii sau, mai precis, pentru cre'sterea robuste'tii la iluminare lateral'a, pe fiecare bloc se aplic'a normalizarea contrastului descriptorului:

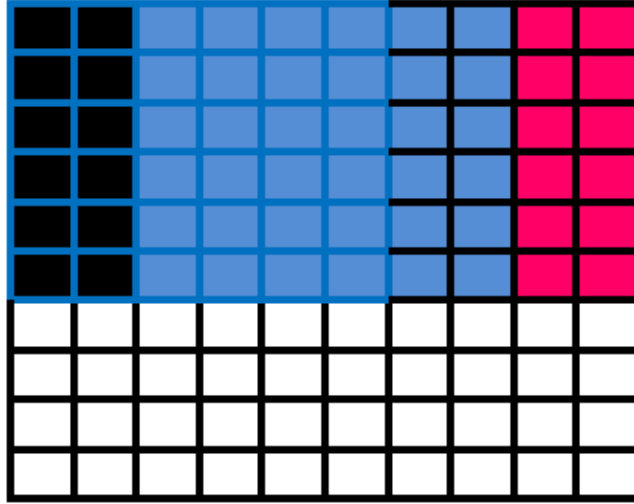


Figure 1.3: Blocuri și celule în calculul HOG. Imaginea are 3×3 blocuri ce se suprapun puternic. Imaginea are 10×10 celule disjuncte. Un bloc are 6×6 celule și fiecare bloc are 6×4 celule comune cu vecinii din stânga sau dreapta și 4×6 celule comune cu vecinii de sus și de jos.

$$\nu(i) = \frac{\nu(i)}{\sqrt{L(\nu) + \epsilon}}, \quad i = 1, \dots, N_o \quad (1.3)$$

unde ϵ este o constantă adăugată pentru a evita împărțirea la zero, iar funcția $L(\nu)$ are semnificație de normă și variante recomandate [?] sunt:

- Norma L_2 ;
- Valoarea curentă trunchiată (de exemplu limitată superior la 0.2);
- Norma L_1 .

Descriptorul final al imaginii sau al regiunii considerate este vectorul format din concatenarea histogramelor de orientări ν pentru toate blocurile considerate. De exemplu, pentru cazul ilustrat în figura 1.3, dacă se consideră $N_o = 8$ orientări (cum este exemplificat în figura 1.2), descriptorul HOG va avea o lungime de 72 elemente: 9 blocuri fiecare descris de o histogramă cu 8 valori normalizate.

Fiind bazat pe contururi, descriptorul HOG este prin definiție invariant la modificarea monotonă (aditivă) a iluminării. În condițiile în care histograma presupune o mediere statistică, descriptorul HOG este de asemenea robust la zgomot. Normalizarea valorilor în cadrul blocului, respectiv pre- și post-procesările îi pot crește robustețea la iluminări laterale, iar utilizarea blocurilor, care la rândul lor pot avea ponderi diferite, facilitează descrierea selectivă a zonelor de interes dintr-o imagine.

În ceea ce privește calculul, se poate observa că determinarea valorii unei histogramme presupune o însumare, deci putem accelera descrierea unei celule cu histograma de orientări prin utilizarea conceptului de imagine integrală.

Aplicațiile histogrammei de orientări a gradientului sunt extrem de variate, pornind de la recunoașterea gesturilor, persoanei sau posturii, la detecția pietonilor sau descrierea tablourilor. Varianta inițială de HOG a fost extinsă fie prin considerarea unei structuri piramidale a imaginii de interes, fie prin complementare cu alți descriptori vizuali cum ar fi histograma de curbatură a gradientului.

ACTIVITATE: Scrieți o funcție care primind ca parametrii o zonă din imagine (fereastră), tipul nucleului (Sobel, etc) și un prag al magnitudinii întoarce histograma HOG asociată acelui bloc.

1.1 Recunoașterea fețelor utilizând HOG

Se consideră o bază de date de mici dimensiuni: 4 personaje \times 3 fețe / personaj a 300×300 pixeli fiecare imagine. În vederea recunoașterii se vor urma pașii:

- Fiecare imagine se descompune 7×7 blocuri disjuncte precum în figura 1.4 (a). Un bloc va avea 43×43 pixeli.

- Fiecare bloc va fi descris'ă de histograma HOG normalizat'ă.
- Histograma fiec'arui asemenea bloc va fi ponderat'ă cu ponderea din figura 1.4 (b).
- Fiecare imagine va fi descris'ă de un vector format prin concatenarea celor 49 de histograme HOG ponderate. Astfel vectorul asociat unei fe'te va fi de lungime $9 \times 49 = 441$ elemente.
- Se va utiliza un test de tipul "leave one out" cu 3 nearest neighbor. Mai precis:
 - Se va considera separat fiecare dintre cele 12 poze din baza existent'ă ("leave one out").
 - Se vor calcula distan'tele (de tip distan'te 'ntre histograme) p'zn'ă la celelalte 11 exemple din baza de date.
 - Se vor considera cele mai apropiate 3 exemple.
 - Dac'ă 'ntre cele trei mai apropiate exemple sunt celelalte dou'ă aparțin'nd fe'tei analizat'ă 'nseamn'ă ca aceasta ar fi clasificat'ă corect.
 - 'n final se consider'ă rata de recunoa'stere corect'ă ca raportul 'ntre clasific'ari corecte 'si num'arul total de teste (12).

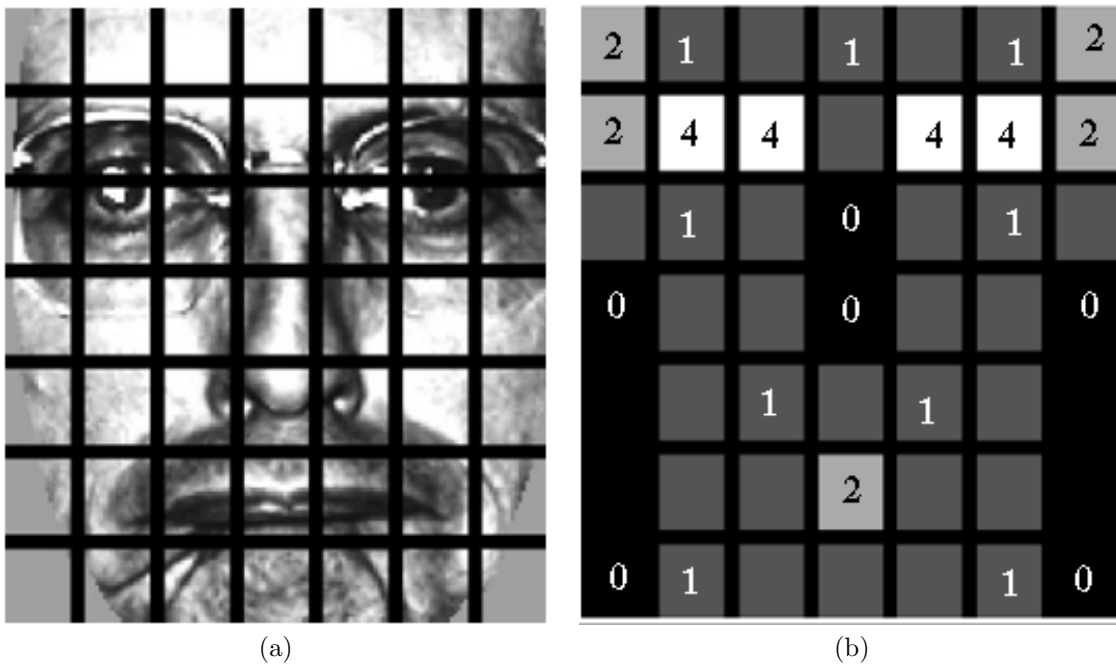


Figure 1.4: (a) Divizarea unui portret 'in blocuri disjuncte. (b) Ponderea fiec'arui bloc.

ACTIVITATE: Testa'ti algoritmul propus pentru recunoa'sterea fe'telor.

ACTIVITATE: Ce se 'int'zmpl'ă dac'ă schimba'ti ponderarea cu una uniform'ă?