

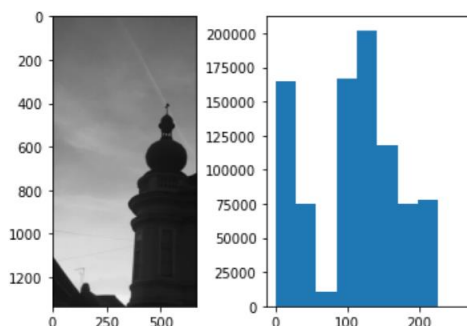
# Хистограм оријентисаних вектора

## 1.1 Дигитална фотографија

Дигитална фотографија представља темељни аспект у домену компјутерске визије. Дводимензионална фотографија је резултат перспективне пројекције тродимензионалне сцене на раван [1]. Када је дигитализована, губи своју континуалност и постаје крајњи скуп дискретних тачака, описан помоћу одговарајуће дискретне структуре података - матрице, чије димензије представљају резолуцију те фотографије. Једна бесконачно мала тачка на дигиталној фотографији (односно пресек врсте и колоне у матрици) кореспондира једном елементу названом пиксел (енгл. pixel). Пиксел је темељни градивни елемент са јединственим координатама у односу на координатни систем фотографије, где сваки ред и колона матрице одговара одређеном положају пиксела. Обично се координатни почетак налази у горњем левом углу слике, где x-оса представља хоризонтално, а y-оса вертикално кретање по фотографији. Постоје различите врсте дигиталних фотографија, које се разликују по начину на који се пиксели користе за приказ информација. Бинарне фотографије користе само две вредности интензитета - 0 и 1, где 0 представља црну, а 1 белу боју. Сиве фотографије (енгл. grayscale images) користе више нијанси сиве, обично из опсега од 0 до 255. Мултиспектралне фотографије, као што су фотографије у боји, користе комбинацију три основне боје - црвене, зелене и плаве - како би створиле различите боје које људско око препознаје.

## 1.2 Хистограм

Хистограм представља графички приказ расподеле фреквенције појављивања вредности неке променљиве, другим речима, представља учесталост појаве вредности променљиве унутар одређеног интервала, при чему се укупно динамичко подручје променљиве дели на жељени број једнаких интервала.



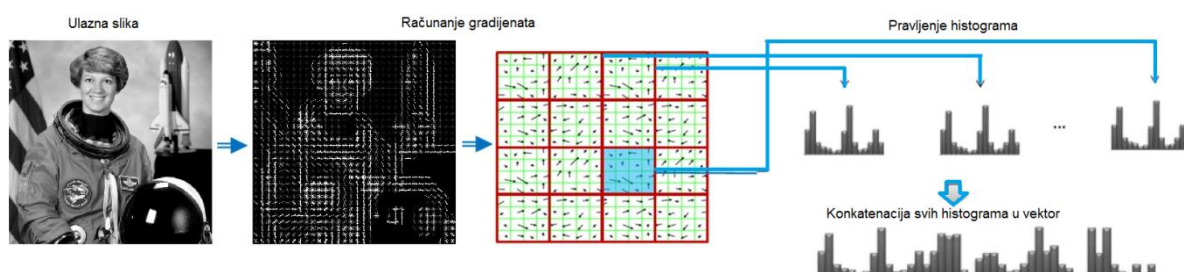
Фотографија 2.1. Хистограм интензитета сиве боје монохроматске фотографије

Фотографија је димензија 750x1334 пиксела, монохроматска је, са 8-битном дубином боје (фотографија 2.1.). Интензитети пиксела се налазе у интервалу од 0 до 255, где 0 представља црну боју, а 255 представља белу боју. Све вредности између 0 и 255 представљају различите нијансе сиве. На основу овог опсега интензитета, рачуна се хистограм појављивања одређених вредности у пикселима фотографије. Хистограм на фотографији 2.1. садржи 9 интервала (енгл. bins), и за сваки интервал апцисе

се на ординати приказује број пиксела који имају интензитет унутар вредности обухваћених тим интервалом. Из примера се види да се бела боја појављује најмање на фотографији, што се одразило и на хистограму.

### 1.2.1 Хистограм оријентисаних градијената

Хистограм оријентисаних градијената је један од популарнијих поступака за издвајање особина (енгл. feature extraction) са дигиталне фотографије. Фокусира се на дистрибуцију градијената интензитета пиксела. Формално, хистограм оријентисаних градијената је дескриптор фотографије (енгл. image descriptor), који представља опис визуелних особина фотографије представљен вектором. Иако се први пут помиње пре скоро две деценије [2], и дан-данас се врло често користи за проблеме детекције и препознавања објеката на фотографијама. Основна идеја је да облик и структура објеката у фотографији могу бити идентификовани на основу локалних промена интензитета пиксела, формалније речено - градијената. Разматра расподелу оријентација градијената у оквиру локалних региона како би се створио вектор карактеристика који описује изглед објеката. Хистограм оријентације градијената је инваријантан на промене осветљења, сенке, геометријске трансформације и помераје објеката. Међутим, зависан је од оријентације објекта и параметара обучавања, односно димензије ћелија и блокова за које се врше израчунавања.



Фотографија 1.2. Поступак добијања хистограма оријентисаних градијената

#### 1.2.1.1 Рачунање градијената

Први корак је припрема фотографија за анализу. Фотографија се може скалирати на одређену величину како би се осигурала конзистентност величина и да би се убрзао процес смањивањем димензионалности. Такође, фотографија се обично конвертује у нијансе сиве боје, јер информација о боји није битна за анализу облика. Градијенти интензитета пиксела играју кључну улогу у анализи.

Деривати пиксела се рачунају користећи изразе применом конволуционих кернела, као што су  $D_x$  и  $D_y$ .

$$D_x = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix}$$

$$D_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Конволуција  $L$  користећи кернеле  $D_x$  и  $D_y$  извршава се према следећим изразима.

$$L_x = L * D_x$$

$$L_y = L * D_y$$

Често се за израчунавање градијената користе Собел-Фелдман кернели. Ови кернели, представљају дискретне операторе диференцијације који омогућавају апроксимацију градијента функције интензитета слике. У хоризонталном смеру, конволуција се одвија помоћу израза:.

$$Dx = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I$$

У вертикалном смеру, конволуција се одвија помоћу израза:

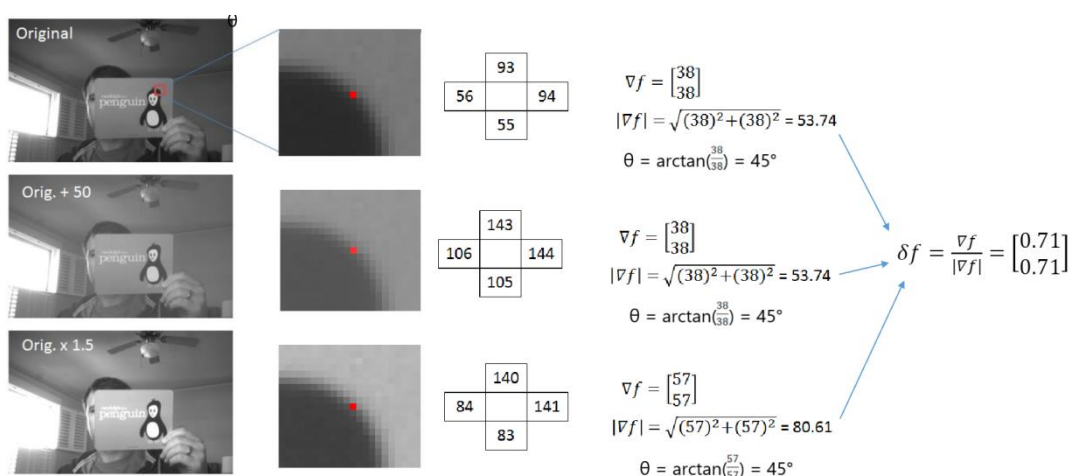
$$Dy = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I$$

Након што се израчунавају производне вредности у оба смера, рачунају се градијенти. Интензитет градијента представља јачину промене осветљења у посматраном пикселу и рачуна се помоћу израза:

$$\|F\| = \sqrt{Lx^2 + Ly^2}$$

Смер градијената представља смер промене светлости и рачуна се изразом:

$$\theta = \tan^{-1}\left(\frac{Lx}{Ly}\right)$$



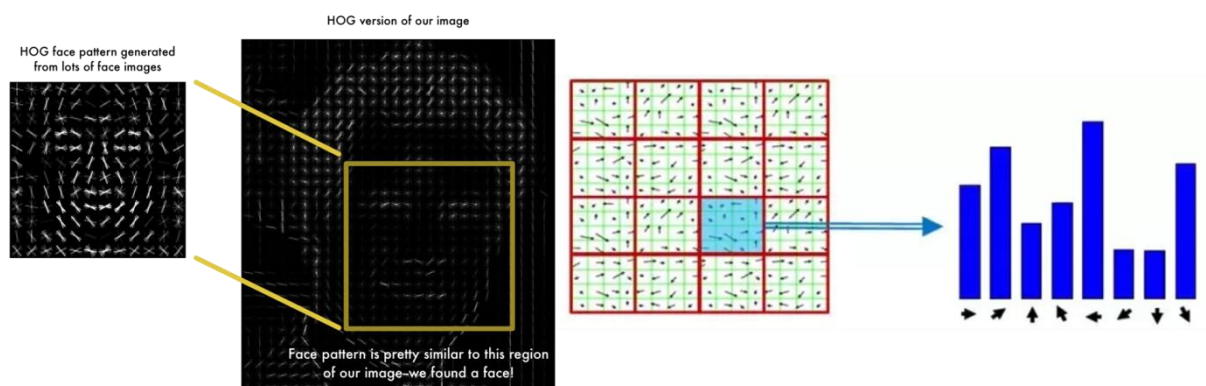
Фотографија 2.3. Пример нормализације градијента пиксела фотографија различитог контраста

Велики интензитет градијента указује на промену на фотографији, као што су ивице објекта или промена текстуре на објекту који се налази на фотографији, док смер градијента представља смер исте те промене. Како би се повећала инваријантност на измену осветљења, извршава се и нормализација градијент вектора.

### 1.2.1.2 Хистограм градијената

Фотографија се дели на мање регионе познате као ћелије (енгл. cells). За сваку ћелију, израчунава се хистограм оријентација градијената. Овај хистограм приказује расподелу оријентација градијената у тој ћелији. Смер се дискретизује у неколико интервала (енгл. bins),

најбољи резултати су добијени за цифру 9 која одговара распону од 0 до 180 степени [2], док се интензитет градијената користи за формирање хистограма. За сваки интервал на хистограму, рачуна се кумулативна вредност интензитета градијената који кореспондира смеру. [3]



Фотографија 2.2. Хистограм оријентисаних градијената једне ћелије

### 1.2.1.3 Нормализација блокова

Да би се постигла инваријантност на промене осветљења и контраста на вишем нивоу, хистограми оријентација се групишу у блокове и нормализују. Блокови се обично преклапају, а корак преклапања зависи од процене потребне величине финалног дескриптора. У [2] показано је да најбоље резултате даје  $L2$  нормализација, још позната и као Еуклидска нормализација. Ако је  $v$  ненормализовани вектор хистограма блока и ако је  $e$  произвољна мала константа чија вредност неће деловати на резултат, онда се нормализација врши на следећи начин:

$$\frac{v}{\sqrt{\|v\|^2 + e^2}}$$

Последњи корак је конкатенирање хистограма оријентисаних градијената свих блокова у један вектор карактеристика, чиме се добија дескриптор за целу фотографију.

## 2. Литература

[1] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, 3rd Edition, Toronto, Ontario: Thomson Learning, 2008.

[2] N. Dalal, B. Triggs, Histogram of Oriented Gradients for Human Detection, 2005.