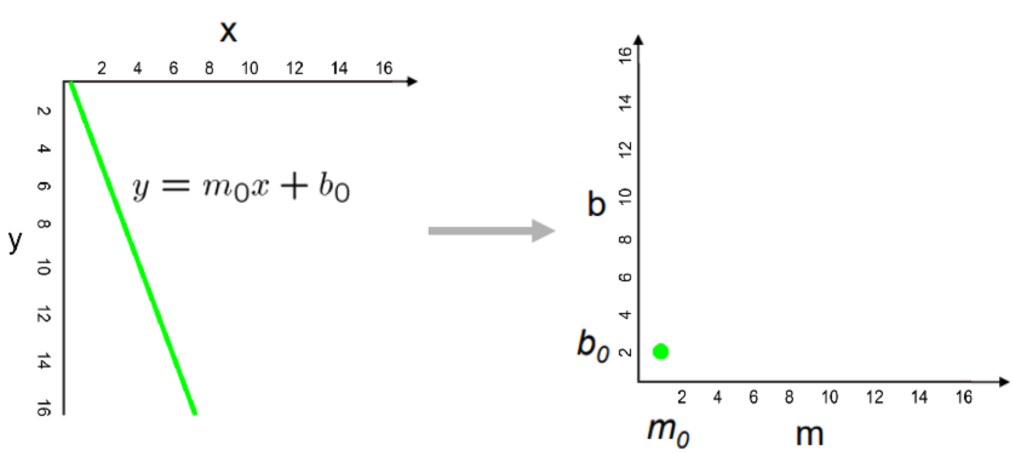
Aspect teoretic

Transformata Hough este o tehnică des utilizată pentru detectarea oricărei forme geometrice dintr-o imagine, dacă se poate reprezenta acea formă în format matematic. Este o tehnică de extracție caracteristică utilizată în viziunea artificială, în analiza imaginii și în procesarea imaginii digitale. Poate detecta forma chiar dacă este defectă sau puțin distorsionată. Obiectivul tehnicii este de a găsi instanțe imperfecte ale obiectelor dintr-o anumită clasă de forme printr-o procedură de vot. Această procedură de vot se realizează într-un spațiu parametric, din care candidații obiect sunt obținuți ca maximele locale într-un așa-numit spațiu acumulator, care este construit în mod explicit de algoritmul pentru calculul transformatei Hough.

Inițial, conceptul transformatei Hough clasice a fost de a identifica liniile din imagine, dar mai târziu transformata Hough a fost extinsă pentru identificarea pozițiilor de forme arbitrare, cele mai frecvente fiind cercurile sau elipsele.

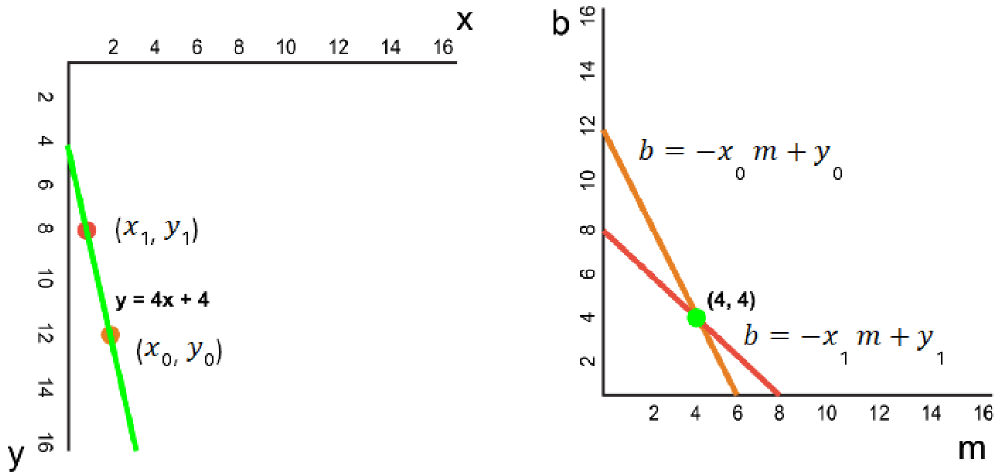
În analiza automată a imaginilor digitale, reiese de multe ori o subproblemă de detectare a formelor simple, cum ar fi liniile drepte, cercurile sau elipsele. În multe situații, un detector de margine poate fi folosit ca etapă de pre-procesare pentru a obține puncte de imagine sau pixeli de imagine care se găsesc pe curba dorită în spațiul imaginii. Cu toate acestea, datorită imperfecțiunilor datelor imaginii sau ale detectorului de cante, pot exista pixeli sau puncte lipsă pe curbele dorite, precum și deviații spațiale între linia linia ideală, cercul ideal sau elipsa ideală și punctele de margine cu zgomot, pe masură ce sunt obținute de la detectorul de cante. Datorită acestor cauze, adeseori este dificil (netrivial) să se grupeze caracteristicile marginilor extrase la un set specific de linii, cercuri sau elipse.

Cel mai simplu caz al transformatei Hough este detectarea liniilor drepte. O linie poate fi reprezentată matematic în moduri diferite, cum ar fi, de exemplu, prin ecuația $y = mx + b$, iar aceasta poate fi reprezentată ca punct (b, m) în spațiul parametric.



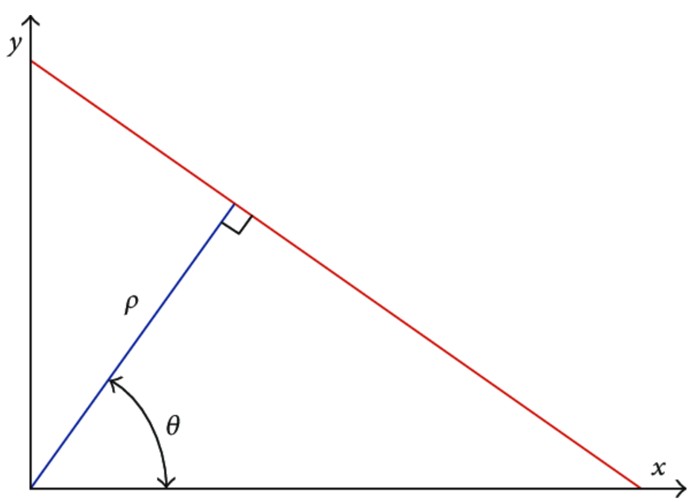
O dreaptă în spațiul imaginii => Un punct în spațiul (parametric) Hough

O altă legatură dintre spațiul imaginii (x, y) și spațiul Hough (b, m) este că unui punct (x_0, y_0) îi corespunde o linie $b = -x_0 m + y_0$ în spațiul Hough:



Un punct în spațiul imaginii => O dreaptă în spațiul parametric Hough

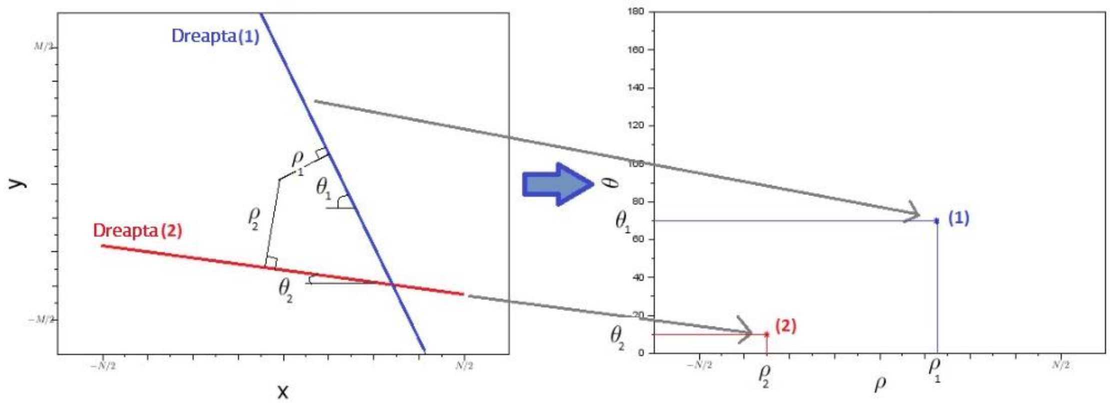
Totuși liniile verticale reprezintă o dificultate. Ele ar genera valori nemărginite ale parametrului pantă m deoarece dinamica acestui coefficient poate fi foarte mare: poate fi zero pentru o linie orizontală, dar și infinită pentru una verticală. Prin urmare, s-a propus utilizarea formei normale Hesse $\rho = x \cos\theta + y \sin\theta$,unde ρ este distanța perpendiculară de la origine la linie, și θ este unghiul format de această linie perpendiculară și axa orizontală masurată în sens trigonometric, θ reprezentând orientarea liniei. Se poate reprezenta în următoarea imagine vectorul perpendicular pe dreaptă, ce trece prin origine, ilustrând dreapta prin parametri ρ şi θ:



Descrierea parametrică a unei drepte, prin reprezentarea polară a parametrilor (ρ, θ)

Transformata Hough constă în potrivirea liniilor din planul cartezian (x, y) cu reprezentarea lor în spațiul parametric (ρ, θ). Este dificil pentru că la un anumit punct (x, y) în plan se poate potrivi cu un număr infinit de linii (toate liniile care trec prin acel punct), care este un număr infinit de puncte în spațiul parametric (ρ, θ).

$$\rho = x \sin\theta + y \cos\theta$$

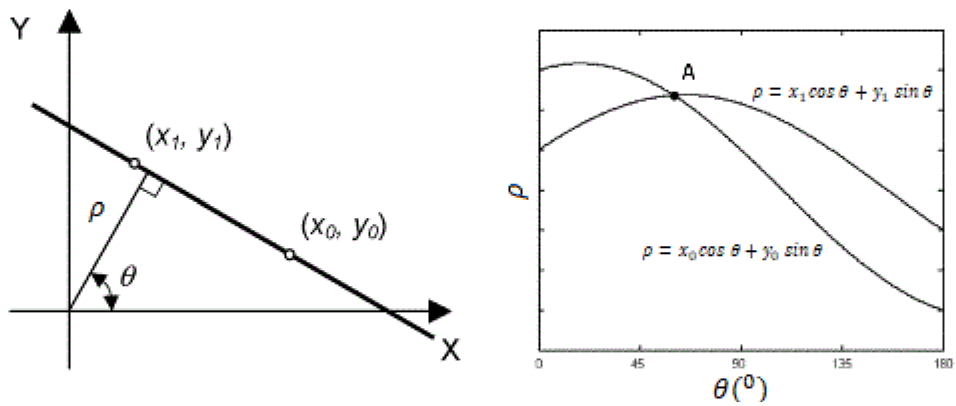


Reprezentarea a doua drepte în spațiul (x, y) Spațiul parametric (ρ, θ)

Dacă linia trece sub origine, va avea un rho pozitiv și un unghi mai mic de 180. Dacă trece peste origine, în loc să ia unghi mai mare de 180, unghiul este luat mai mic de 180, iar rho este luat negativ. Orice linie verticală va avea 0 grade, iar liniile orizontale vor avea 90 de grade.

Orice linie a imaginii poate fi reprezentată în acești doi termeni, (ρ, θ). Planul (ρ, θ) este denumit câteodată spațiul liber pentru setul de linii drepte în două dimensiuni. Această reprezentare face ca transformata Hough să fie din punct de vedere conceptual foarte apropiată de transformata Radon bidimensională. De fapt, transformata Hough este echivalentă cu transformata Radon, dar cele două transformate au interpretări de calcul diferite asociate în mod tradițional cu acestea.

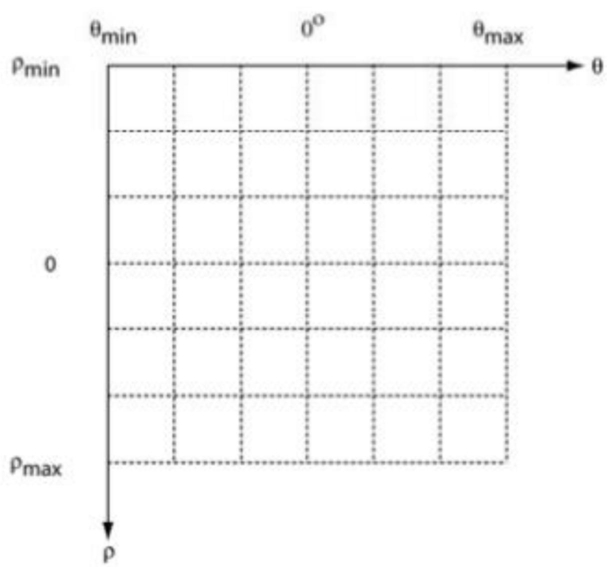
Luându-se în considerare un singur punct în plan, atunci setul tuturor liniilor drepte care trec prin acel punct corespund unei curbe sinusoidale în planul (ρ, θ) ,care este unic la acel punct. Un set de două sau mai multe puncte care formează o linie dreaptă va produce sinusoide care trec la (ρ, θ) pentru acea linie. În acest mod, problema detectării punctelor coliniare poate fi convertită la problema aflării curbelor concurente.



Reprezentarea a două puncte în spațiul (x, y) reprezentând două sinusoide în planul (ρ, θ)

Implementare

Întâi se creează o matrice 2D sau un acumulator (pentru a deține valorile celor 2 parametri) și este setat inițial la 0.



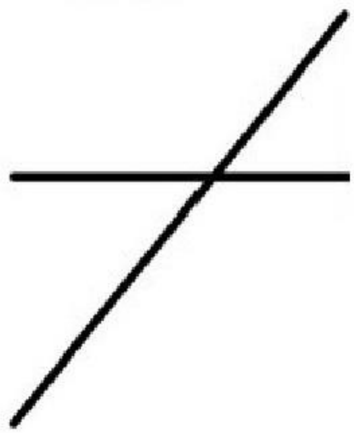
Matricea acumulator

Algoritmul linear al transformatei Hough folosește o matrice bidimensională denumită accumulator, pentru a identifica existența unei linii drescrise de $\rho = x \cos\theta + y \sin\theta$. Rândurile indică parametrul ρ iar coloanele indică parametrul θ. Dimensiunea matricei depinde de precizia de care este nevoie. Dacă se presupune că se dorește ca precizia unghiurilor să fie de 1 grad, este nevoie de 180 de coloane. Pentru ρ, distanța maximă posibilă este lungimea diagonală a imaginii. Astfel, luând o precizie de un pixel, numărul de rânduri poate fi lungimea diagonalei imaginii. Dimensiunea acumulatorului este egală cu numărul de parametri necunoscuți, adică doi, luând în considerare valorile cuantizate ale lui ρ și θ din perechea (ρ, θ). Pentru fiecare pixel la (x, y) și vecinătatea lui, algoritmul de transformare Hough determină dacă există suficiente dovezi ale unei linii drepte la acel pixel. În acest caz, se va calcula parametrii (ρ, θ) ai liniei respective, apoi se va căuta containerul acumulatorului în care parametrii se încadrează și se va mări(incrementa) valoarea acelui container. Prin găsirea containerelor cu cele mai mari valori, de obicei prin căutarea maximelor locale în spațiul acumulatorului, liniile cele mai probabile pot fi extrase, iar definițiile lor geometrice (aproximative) sunt citite. Cea mai simplă cale de a afla aceste vârfuri este prin aplicarea unei anumite forme de prag, dar alte tehnici pot produce rezultate în diferite situații – determinarea liniilor care se gasesc, precum și a numărului de linii. Datorită faptului că liniile returnate nu conțin informații despre lungime, este deseori necesar, în pasul următor, să se găsească părțile imaginii care corespund liniilor. Mai mult decât atât, din cauza erorilor de imperfecțiune în etapa de detectare a marginilor, vor exista de obicei erori în spațiul acumulatorului, ceea ce îl poate face sa nu fie ușoară aflarea vârfurile corespunzătoare și, astfel, liniile corespunzătoare.

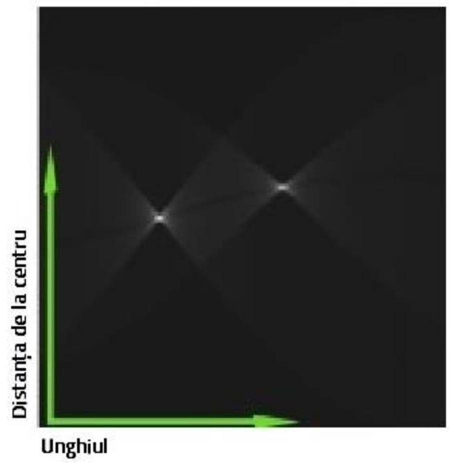
Rezultatul final al transformării Hough liniare este un vector bidimensional (o matrice) asemănător cu acumulatorul – o dimensiune a acestei matrici este unghiul cuantificat θ ,iar cealaltă dimensiune este dinstanța cuantificată ρ. Fiecare element al matricei are o valoare egală cu suma punctelor sau a pixelilor care sunt poziționați pe linia reprezentată de parametrii cuantizați (ρ, θ) .De aceea, elementul cu cea mai mare valoare indică linia dreaptă cea mai reprezentativă a imaginii de intrare.

Exemplificare

Se consider o imagine de 100x100 cu o linie orizontală la mijloc. Se ia primul punct al liniei. Se cunosc valorile (x,y) ale acestui punct. În linia ecuației se pun valorile $\theta = 0{,}1{,}2, \ldots.,180$ și se verifică ρ obținut. Pentru fiecare pereche (ρ, θ) se incrementează valoarea cu unu în acumulator în celulele sale (ρ, θ) corespunzătoare. Deci, acum în acumulator, celula (50,90) = 1 împreună cu alte celule.
Apoi se ia al doilea punct de pe linie. Se procedează la fel ca mai sus. Se incrementează valorile în celulele corespunzatoare (ρ, θ) pe care le-am obținut. De această dată, celula (50,90) = 2. Acest proces este de a vota valorile (ρ, θ). Se continuă acest proces pentru fiecare punct de pe linie. La fiecare punct, celula (50,90) va fi incrementată sau votată, în timp ce alte celule pot sau nu pot fi votate. Astfel, la final, celula (50,90) va avea voturi maxime. Deci, dacă se caută acumulatorul pentru voturi maxime, se va obține valoarea (50,90) care spune ca acolo este o linie în respectiva imagine la distanța 50 de la origine și la unghiul de 90 de grade. Urmatoarea imagine ilustrează acumulatorul. Punctele luminoase distinct din locații indică faptul ca acestea sunt parametri liniilor posibile din imagine. Din pozițiile acestor puncte, se pot determina unghiul și distanța față de centrul imaginii celor două linii din imaginea de intrare.



Doua linii în imaginea de intrare



Redarea Rezultatelor Transformatei

Utilizarea direcției gradientului pentru a reduce numărul de voturi

O îmbunătățire sugerată poate fi folosiă pentru a detecta liniile dacă se ia în considerare faptul că gradientul local al intensității imaginii va fi în mod necesar ortogonal față de margine. Deoarece detectarea muchiilor implică în general calcularea magnitudinii gradientului de intensitate, direcția gradientului este adesea găsită ca efect secundar. Dacă un punct dat de coordonate (x, y) se întâmplă să fie într-adevăr pe o linie, atunci direcția locală a gradientului dă parametrul θ corespunzător liniei respective, iar parametrul ρ este apoi obținut imediat. Direcția gradientului poate fi estimată la 20°, ceea ce scurtează semnalul sinusoid de la 180° la aproximativ 45°. Acest lucru reduce timpul de calcul și are efectul interesant de a reduce numărul de voturi inutile, sporind astfel vizibilitatea vârfurilor corespunzătoare liniilor reale din imagine.

Transformata Hough bazată pe Kernel(nucleu) (THK)

A fost sugerată o schemă de vot îmbunătățită pentru transformata Hough, care permite o implementare software pentru a obține performanțe în timp real, chiar și pe imagini relativ mari. Transformata Hough bazată pe Kernel (nucleu) utilizează aceeași parametrizare (ρ, θ), dar funcționează pe grupuri de pixeli aproximativ coliniari. Pentru fiecare grup, voturile sunt exprimate folosind un nucleu eliptic-Gaussian orientat care modelează incertitudinea asociată cu cea mai potrivită linie în raport cu grupul corespunzător. Abordarea nu numai că îmbunătățește semnificativ performanța sistemului de vot, ci produce și un acumulator mult mai curat și face transformata mai robustă pentru detectarea liniilor false.

Transformarea Hough a curbelor și generalizarea lor pentru forme analitice și neanalitice

Deși versiunea transformatei descrise mai sus se aplică numai la găsirea liniilor drepte, o transformată similară poate fi utilizată pentru găsirea oricărei forme care poate fi reprezentată de un set de parametri. Un cerc, de exemplu, poate fi transformat într-un set de trei parametri, reprezentând centrul și raza sa, astfel încât spațiul Hough devine tridimensional. Elipsele și curbele arbitrare pot fi găsite și în acest fel, la fel ca orice formă ușor exprimată ca un set de parametri.

Spre deosebire de alte abordări bazate pe transformarea Hough pentru forme analitice, tehnica nu depinde de forma pe care cineva vrea să o detecteze sau de tipul de date de intrare. Detectarea poate fi determinată de un tip de formă analitică prin schimbarea modelului de geometrie presupus în care datele au fost codificate (de exemplu, spațiu euclidan, spațiu proiectiv, geometrie conformală etc.), în timp ce formularea propusă rămâne neschimbată. De asemenea, garantează că formele propuse sunt reprezentate cu cel mai mic număr posibil de parametri și permite detectarea simultană a diferitelor tipuri de forme care se potrivesc cel mai bine unui set de intrări cu dimensiuni diferite și definiții geometrice diferite (de exemplu, detectarea concomitentă a planurilor și sferelor care se potrivesc cel mai bine unui set de puncte, linii drepte și cercuri).

Pentru forme mai complicate în plan (adică, forme care nu pot fi reprezentate analitic în unele spații 2D), este utilizată transformarea Hough generalizată, care permite unei entități să voteze o anumită poziție, orientare și/sau scalarea formei utilizând un tabel de căutare predefinit.

Transformata Hough pentru cercuri

Transformata Hough pentru Cerc (THC) este o tehnică de extracție caracteristică de bază utilizată în procesarea imaginii digitale pentru detectarea cercurilor în imagini imperfecte. Candidații cerc sunt produse prin "vot" în spațiul parametrului Hough și apoi selectându-se maxima locală într-o matrice acumulator. Este o specializare de transformare Hough.

Transformata Hough poate fi folosita la determinarea paramerilor unui cerc atunci cand exista un numar de puncte care se afla in zona acelui cerc. Un cerc de raza R cu centrul (a, b) poate fi descris cu urmatoarele ecuatii parametrice:

$x = a + Rcos(\theta)$

$y = a + Rsin(\theta)$

Când unghiul θ ia valori între 0 si 360 de grade, punctele (x, y) urmăresc perimetrul unui cerc. Dacă o imagine conține multe puncte, dintre care unele se găsesc în perimetrele unor cercuri, atunci problema este găsirea tripletelor (a, b, R) pentru a descrie fiecare cerc. Una din cele mai mari probleme este aceea a spațiului 3D, care determină o complexitate computatională foarte mare. Din acest motiv este mai bine dacă raza R este cunoscută, pentru a reduce căutarea într-un spațiu 2D. În cazul nostru razele cercurilor căutate nu sunt constante, dar se încadreaza în niște limite destul de stricte.

Procesul de detectare a cercului

Modificarea algoritmului pentru a detecta forme circulare în loc de linii este relativ simplă.

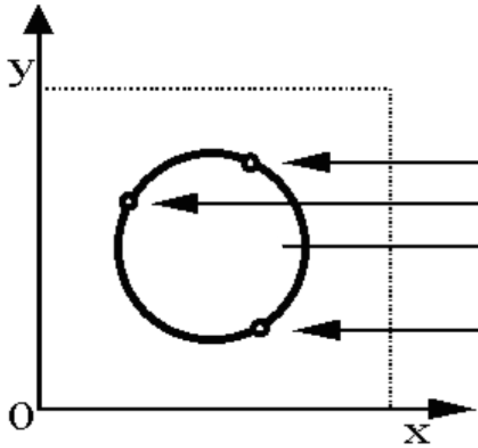
- În primul rând, creăm spațiul acumulatorului, care este format dintr-o celulă pentru fiecare pixel. Inițial, fiecare celulă este setată la 0.
- Pentru fiecare punct de margine (x, y) din imagine, se incrementează toate celulele care, conform ecuației unui cerc, $(x-a)^2+(y-b)^2 = r^2$ ar putea fi centrul unui cerc. Aceste celule sunt reprezentate prin litera a din ecuație.
- Pentru fiecare posibilitate a valorii a găsită în pasul anterior, se găsesc toate valorile posibile ale valorii b care satisfac ecuația.
- Se caută maximele locale în spațiul acumulatorului. Aceste celule reprezintă cercuri care au fost detectate de algoritm.

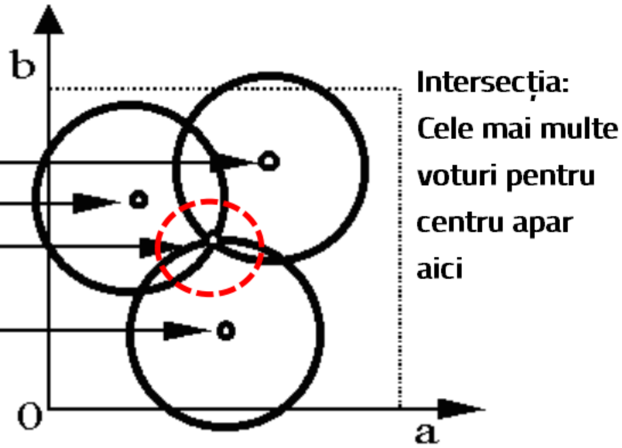
Dacă nu se cunoaște raza cercului pe care se încearcă localizarea sa în prealabil, se poate folosi un spațiu tridimensional pentru acumulator pentru a căuta cercuri cu o rază arbitrară. În mod natural, acest lucru este mai scump din punct de vedere al calculelor.

Această metodă poate detecta, de asemenea, cercuri care sunt parțial în afara spațiului acumulatorului, atâta timp cât zona cercului este încă prezentă în ea.

Aflarea parametrilor cu raza $\boldsymbol{R}$ cunoscută

Dacă raza este fixă, atunci spațiul parametrului va fi redus la 2D (poziția centrului cercului). Pentru fiecare punct (x, y) de pe cercul original, acesta poate defini un cerc centrat la(x, y) cu raza R conform ecuației $(x-a)^2+(y-b)^2 = r^2$. Punctul de intersecție al tuturor acestor cercuri din spațiul de parametri corespunde punctului central al cercului original.



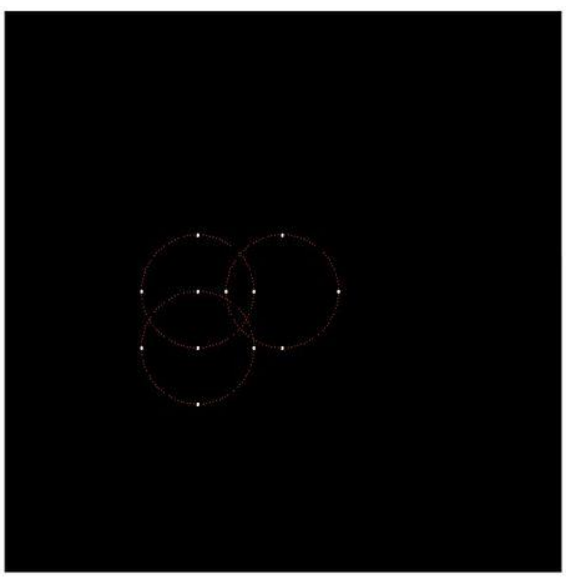


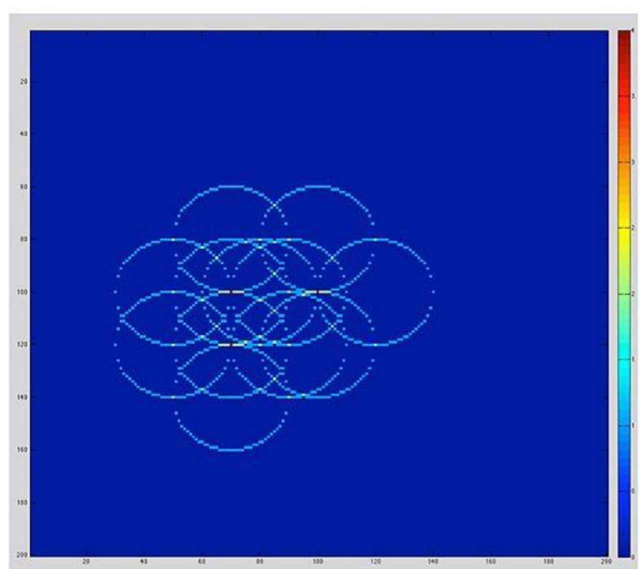
3 puncte pe cerc în imaginea original   Transformata Hough pentru Cerc

Se presupune a fi cunoscută raza. Pentru fiecare (x, y) dintre cele trei puncte (puncte albe) din imaginea originală, se poate defini un cerc în spațiul parametric Hough centrat la (x, y) cu raza R. O matrice a acumulatorului este utilizată pentru localizarea punctului de intersecție. În spațiul parametric, numărul de voturi prin care trece cercul va fi mărit cu unul.

Mai multe cercuri cu raza R cunoscută

Mai multe cercuri cu aceeași rază pot fi găsite cu aceeași tehnică.





Imaginea original cu 3 cercuri   Transformata Hough pentru Cercuri

Se observă că, în matricea acumulatorului (figura din dreapta), vor exista cel puțin 3 puncte maxime locale.

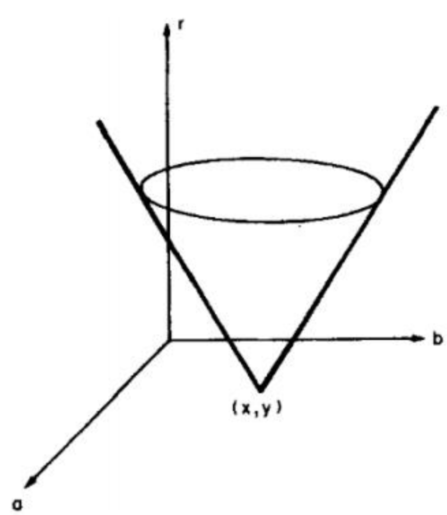
Matricea acumulatorului și votul

În practică, este introdusă o matrice a acumulatorului pentru a găsi punctul de intersecție în spațiul parametric. În primul rând, trebuie să împărțim spațiul parametric în “containere” folosind o grilă și să producem o matrice de acumulator în funcție de grilă. Elementul din matricea acumulatorului denotă numărul de “cercuri” din spațiul parametrului care trece prin celula grilei corespunzătoare din spațiul parametrului. Numărul este, de asemenea, numit “numărul de vot”. Inițial, fiecare element din matrice este zero. Apoi, pentru fiecare punct de „margine” din spațiul original, putem formula un cerc în spațiul parametric și putem crește numărul de vot al celulei de grilă prin care trece cercul. Acest proces se numește “vot”.

După votare, putem găsi maximele locale în matricea acumulatorului. Pozițiile maxime locale corespund centrelor cercului din spațiul original.

Aflarea parametrului cercului cu rază necunoscută

Deoarece spațiul parametric este 3D, matricea acumulatorului va fi și ea 3D.



Spațiul Hough 3D

Putem itera prin raze posibile; pentru fiecare rază, folosim tehnica anterioară. În cele din urmă, găsiți maximele locale în matricea acumulatorului 3D. Matricea acumulatorului trebuie să fie $A[x, y, R]$ în spațiul 3D. Votarea ar trebui să fie pentru fiecare pixeli, rază și theta $A[x, y, R] += 1$.

Algoritmul

1. Pentru fiecare $A[x, y, r] = 0$;
2. Se preprocesează algoritmul de filtrare pe imaginea folosint Blurarea Gaussiană, se convertește imaginea în gri ,apoi se pune operatorul Canny, operatorul Canny oferă marginile imaginii.
3. Se votează toate cercurile posibile în acumulator.
4. Maximele locale ale cercurilor votate pentru acumulatorul A oferă spațiul Hough pentru cerc.
5. Cercul maxim votat de acumulator oferă cercul.

Limitări

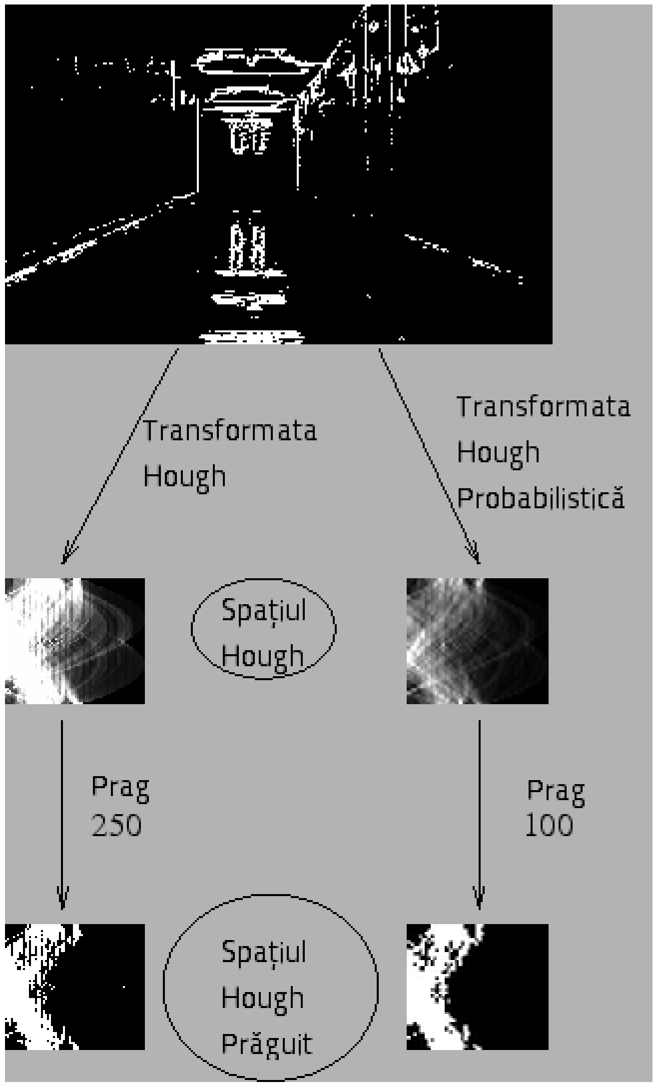
Deoarece spațiul parametrilor CHT este tridimensional, poate necesita mult spațiu de stocare și calcul. Alegerea unei dimensiuni mai mari a grilei poate ameliora această problemă.

Cu toate acestea, alegerea unei dimensiuni adecvate a grilei este dificilă. Din moment ce o grilă prea inferioară (aspră) poate duce la obținerea unor valori mari care sunt false votului, deoarece multe structuri destul de diferite corespund unui singur container. O grilă prea fină poate duce la faptul că structurile nu sunt găsite, deoarece voturile rezultate din jetoanele care nu sunt exact aliniate ajung în containere diferite și nici un container nu are un vot mare.

De asemenea, THC nu este foarte robustă la zgomot.

Transformata Hough Probabilistică

În transformata Hough, se poate observa că, chiar și pentru o linie cu două argumente, este nevoie de o mulțime de calcule. Transformata Hough Probabilistă este o optimizare a transformării Hough. Ea nu ia în considerare toate punctele, ci ia doar un subset aleatoriu de puncte și este suficient pentru detectarea liniei. Trebuie doar să reducem pragul. Se poate observa în imaginea urmatoare, în care se compară Transformata Hough și Transformata Hough Probabilistică în spațiul Hough.



Comparare Transformata Hough si Transformata Hough Probabilistică

Cel mai bun lucru este că, se returnează direct cele două puncte finale ale liniilor. În cazul anterior, se avea doar parametrii de linii, și a trebuit să găsească toate punctele.

Preprocesarea Imaginii

Aplicarea detectorului de canate (Canny)

Detectorul de cante(Canny) este un algoritm cu mai multe etape, optimizat pentru detectarea rapidă a marginilor în timp real. Scopul fundamental al algoritmului este de a detecta schimbările bruște de luminozitate (gradienți mari), cum ar fi o trecere de la alb la negru, și le definește ca margini, având în vedere un set de praguri. Algoritmul detectorului de cante are patru etape principale:

1. **Reducerea zgomotului:** Ca și în cazul tuturor algoritmilor de detectare a marginilor, zgomotul este o problemă crucială care duce adesea la detectarea falsă. Un filtru Gaussian 5x5 este aplicat pentru a netezi imaginea pentru a reduce sensibilitatea detectorului la zgomot. Acest lucru se face folosind un nucleu/kernel (în acest caz, un nucleu 5x5) de numere distribuite normal pentru a rula pe întreaga imagine, setând fiecare valoare a pixelilor egală cu media ponderată a pixelilor învecinați.

$$\mathbf{B} = \frac{1}{159}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} * \mathbf{A}$$

5x5 Nucleu(Kernel) Gaussian. Asteriscul indică funcționarea convoluției

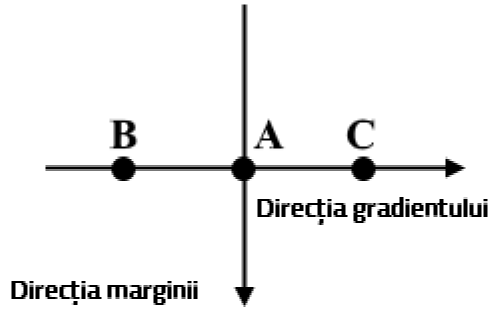
2. **Intensitatea gradientului:** Imaginea netezită este apoi îi este aplicată un nucleu Sobel, Roberts sau Prewitt (Sobel este utilizat în OpenCV) de-a lungul axei X și Y pentru a detecta dacă marginile sunt orizontale, verticale sau diagonale.

$$Edge\_Gradient\ (G) = \sqrt{G_x^2 + G_y^2}$$
$$Angle\ (\theta) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

Nucleul(Kernelul) Sobel pentru calculul primei derivate al direcției orizontalei și vertical

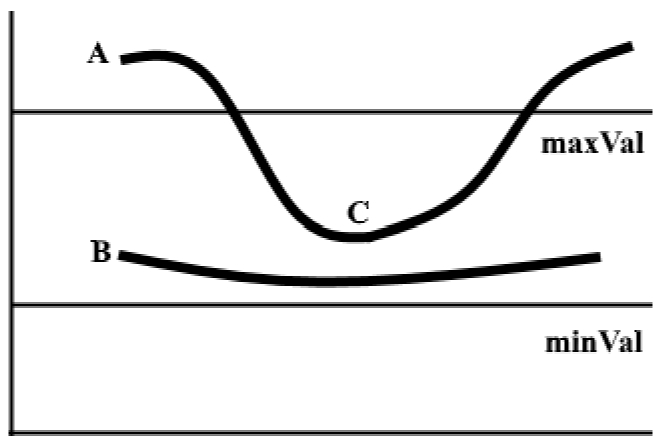
3. **Suprimarea non-maximă:** Suprimarea non-maximă este aplicată să “subțieze” și să ascută eficient marginile. Pentru fiecare pixel, valoarea este verificată dacă este o valoare maximă locală în direcția gradientului calculat anterior. În imaginea următoare A se află pe marginea cu direcție verticală. Deoarece gradientul este normal pentru direcția marginii, valorile punctelor B și C sunt comparate cu valorile punctelor A pentru a determina dacă A este maxim

local. Dacă A este valoarea maximă locală, se testează suprimarea maximă pentru următorul punct. În caz contrar, valoarea pixelului A este setată la zero și A este suprimată.



Suprimarea non-maximă pe trei puncte

4. **Pragul/Limita de histerezis:** După suprimare non-maximă, pixelii puternici sunt confirmați a fi pe harta finală a marginilor. Cu toate acestea, pixelii slabi trebuie analizați în continuare pentru a determina dacă reprezintă o margine sau un zgomot. Aplicând două valori predefinite pentru limitele(pragurile) minVal și maxVal, am stabilit că orice pixel cu gradient de intensitate mai mare decât maxVal sunt margini și orice pixel cu gradient de intensitate mai mic decât minVal nu sunt margini și este eliminat. Pixelii cu gradient de intensitate între minVal și maxVal sunt considerați muchii numai dacă sunt conectați la un pixel cu gradient de intensitate peste maxVal.



Pragul de histerezis exemplificat pe două linii

Marginea A este deasupra maxVal așa că este considerat o margine. Marginea B este între maxVal și minVal, dar nu este conectată la nici o margine deasupra maxVal, așa că este eliminată. Marginea C este între maxVal și minVal și este conectată la muchia A, o margine deasupra maxVal, deci este considerată o margine.