Računalni vid
Analiza kretanja u slikovnoj ravnini

Izv. prof. dr. sc. Zoran Kalafatić
Zoran.Kalafatic@fer.hr

ZEMRIS, 2022/23

Sadržaj predavanja

- Dinamički računalni vid
- Slike razlika, oduzimanje slike pozadine
- Postupci procjene optičkog toka

Analiza dinamičkih scena

- ulaz: slijed 2D slika kroz koji se promatraju događaji u 3D sceni u kojoj se objekti pomiču u odnosu na kameru (postoji relativno gibanje)
- kretanje u sceni sadrži dodatnu informaciju koju biološki vidni sustavi uspješno koriste - korisno je upotrijebiti je i u sustavima umjetnog vida
- promjene u sceni mogu nastati kao posljedica:
  - kretanja kamere
  - kretanja objekta
  - promjene osvjetljenja u sceni (nastojimo izbjeći)
  - promjene strukture, veličine ili oblika objekta u sceni (uobičajena pretpostavka rigidnosti ili kvazi-rigidnosti objekata u sceni)

Zadaci sustava dinamičkog vida

- detekcija promjena u slici (odnosno u sceni)
- segmentacija slike temeljena na kretanju
- rekonstrukcija 3D strukture objekata u sceni
- određivanje značajki kretanja promatrača (vozilo s kamerom) i objekata o sceni
- praćenje objekata u sceni
- prepoznavanje pokretnih objekata
- opis događanja u sceni prirodnim jezikom

Pristupi analizi dinamičkih scena

- izlučivanje dinamičke informacije pronalaženjem razlika između susjednih slika
  - oduzimanje susjednih slika, usporedba dobivenih razlika s pragom
  - modeliranje pozadine, oduzimanje pozadine od ulazne slike
- izlučivanje dinamičke informacije iz slijeda slika računanjem optičkog toka
  - za svaku točku određuje se (prividni) vektor brzine
- analiza svake pojedinačne slike u slijedu, uspostavljanje korespondencije između pronađenih objekata ili značajki u susjednim slikama niza
  - pitanje pronalaženja objekata (lakše – značajki)
  - problem određivanja odgovarajućih točaka u susjednim slikama

Oduzimanje pozadine

Slike razlika

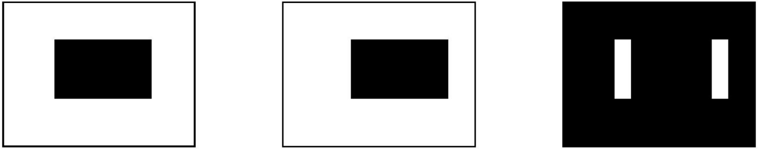
frame differencing

- najjednostavniji pristup izlučivanju dinamičke informacije
- od tekuće slike oduzme se prethodna
  - dobije se slika čije vrijednosti slikovnih elemenata odgovaraju promjeni svjetlosnog intenziteta na odgovarajućoj poziciji
  - usporedba s pragom $\rightarrow$ označavaju se slikovni elementi na kojima je promjena značajna $\rightarrow$ *maska kretanja*



$$I_{izl}(x,y) = \begin{cases} 1 & \text{ako } |I_2(x,y) - I_1(x,y)| > \theta \\ 0 & \text{inače} \end{cases}$$

- u slučaju jednoliko obojenog (neteksturiranog) objekta – slika razlika pokazuje rub objekta (napuštanje odnosno ulazak u područje)
- veća brzina objekta (pomak) rezultira većom površinom na slici razlike; mali pomaci mogu biti neuočljivi !
- ako je objekt teksturiran, slika razlika puno kompaktnije obuhvaća objekt

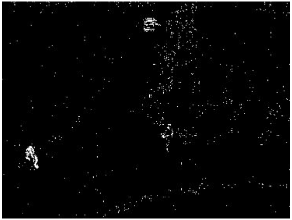


- osjetljivost na izbor praga

ulazna slika

apsolutna razlika

previsok prag



prenizak prag

Oduzimanje pozadine

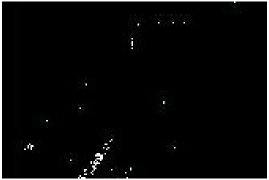
- slike razlika $\rightarrow$ problem “rupa” u detektiranim pokretnim objektima
- mogućnost oduzimanja pozadine
  - treba nam model pozadine, koji ne uključuje pokretne objekte
  - može biti slika s početka sekvence, odnosno slika bez objekata koje namjeravamo detektirati

$$I_{izl}(x,y) = \begin{cases} 1 & \text{ako } |I_{ul}(x,y) - B(x,y)| > \theta \\ 0 & \text{inače} \end{cases}$$

- ograničenje – statična kamera
- problem s promjenom pozadine
  - promjena osvjetljenja u sceni (doba dana, vedro – oblačno)
  - neki pokretni objekt postaje dio pozadine – ako je model pozadine statičan, objekt se stalno detektira (npr. auto koji se uparkira)
- potreba za prilagodbom modela pozadine

ulaz

statička pozadina



prilagodljiva pozadina

- primjer: nekoliko slika iz snimke nogometne utakmice – postupna promjena osvjetljenje, slučajno pomaknuta kamera



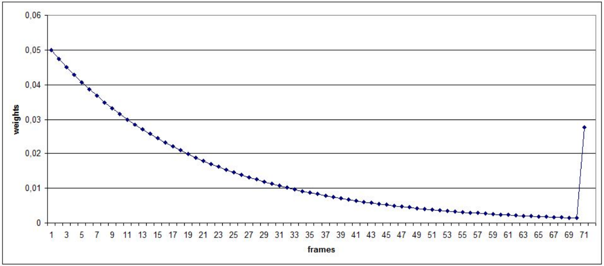

Prilagodljiva pozadina

- ideja: model pozadine se dobija usrednjavanjem prethodnih n slika
  - modelira se svaki slikovni element
  - očekujemo da se pokretni objekt kratko zadržava na poziciji pojedinog piksela
  - usrednjavanje filtrira te kratkotrajne promjene vrijednosti slikovnog elementa
- umjesto srednje vrijednosti može se koristiti i median (robusnija procjena)
- osjetljivost na izbor praga – prag globalan – jedinstven za sve slikovne elemente
- nedostatak: memorijski zahtjevi za pohranu n slika

- uobičajeni pristup – inkrementalnim usrednjavanjem (*exponential moving average*):

$$B_{i+1}(x,y) = \alpha \cdot I_i(x,y) + (1-\alpha) \cdot B_i(x,y)$$

- α je faktor učenja (potrebno ga je pažljivo podesiti)
- procjena uzima u obzir sve prethodne vrijednosti piksela – težinski prosjek uz opadanje težine prema starosti slike



- ovisno o veličini faktora učenja i brzini objekata, u modelu pozadine se neko vrijeme zadržavaju tragovi objekata
- to se pri detekciji objekta manifestira kao trag iza objekta



- primjer pojave “duhova” u modelu pozadine

Selektivno održavanje pozadine

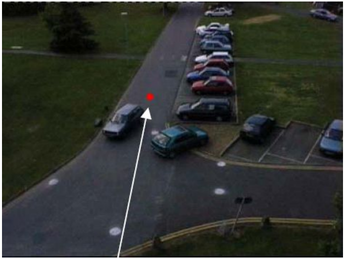
- u prethodnom modelu ažuriranja pozadine područja koja odgovaraju objektu (*foreground*) jednako utječu na pozadinu kao i područja u kojima se objekti ne detektiraju (*background*)
- maska kretanja dobivena nakon oduzimanja pozadine od trenutne slike može se iskoristiti za selektivno ažuriranje
  - ažuriraju se samo oni slikovni elementi pozadine koji nisu klasificirani kao objekt

$$M_i(x, y) = \begin{cases} 1 & \text{ako } |I_i(x, y) - B_i(x, y)| > \theta \\ 0 & \text{inače} \end{cases}$$

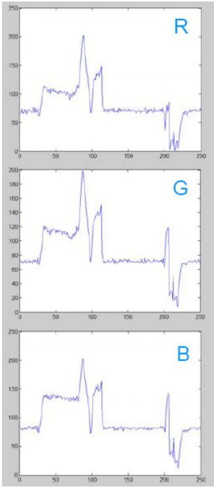
$$B_{i+1}(x, y) = \begin{cases} \alpha \cdot I_i(x, y) + (1 - \alpha) \cdot B_i(x, y) & \text{ako } M_i(x, y) = 0 \\ B_i(x, y) & \text{inače} \end{cases}$$

Distribucija vrijednosti slikovnog elementa

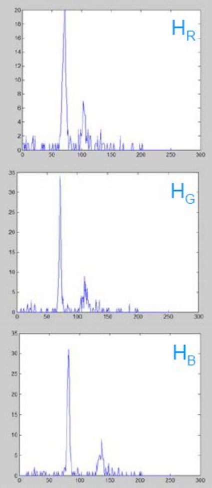
- modelira se ponašanje pojedinog slikovnog elementa, ne uzima se u obzir okolina



položaj
slikovnog elementa



vremenski slijed
vrijednosti



histogrami

Modeliranje piksela normalnim distribucijama

- prilikom oduzimanja pozadine zapravo treba ustanoviti uklapa li se trenutna vrijednost pojedinog slikovnog elementa u njegov model
- modeliranje svakog slikovnog elementa (unimodalnom) Gaussovom razdiobom $\mathcal{N}(\mu_i(x,y), \sigma_i^2(x,y))$
  - svaki piksel ima svoju varijancu na temelju koje se određuje njegov individualni prag segmentacije

$$\mu_{i+1} = \alpha \cdot I_i + (1-\alpha) \cdot \mu_i$$

$$\sigma_{i+1}^2 = \alpha \cdot (I_i - \mu_i)^2 + (1-\alpha) \cdot \sigma_i^2$$

- individualni prag segmentacije $\theta_i(x,y) = k \cdot \sigma_i(x,y)$

$$M_i(x,y) = \begin{cases} 1 & \text{ako } \ |I_i(x,y) - \mu_i(x,y)| > \theta_i(x,y) \\ 0 & \text{inače} \end{cases}$$

- bolje pokriva individualnu varijabilnost slikovnih elemenata, ali ne i multimodalnost

Titranje pozadine

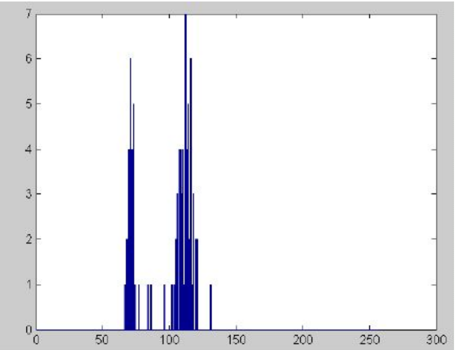
- u nekim situacijama imamo gibanje inherentno pozadini
- tipičan primjer – njihanje grana drveta
- model pozadine bi trebao moći opisati takve dugotrajnije promjene vrijednosti nekih slikovnih elemenata u slici

ulaz

željeni rezultat

Multimodalno modeliranje pozadine

- karakteristično titranje nekih područja u slici (njihanje grana, titranje lišća) $\rightarrow$ multimodalna razdioba



- modeliranje slikovnih elemenata multimodalnom normalnom razdiobom (*mixture of Gaussians*) – $\mu_i, \sigma_i, \omega_i$ (Stauffer and Grimson, 1999)
- pitanje broja očekivanih modova (3-5)
- inicijalizacija, ažuriranje vrijednosti
- način pronalaženja modova
  - EM algoritam (*expectation maximisation*)
  - Mean-shift algoritam
  - aproksimacija razdiobe jezgrenim funkcijama (Parzenovi prozori)

ulaz

Gaussov model

mješavina Gaussa

- mješavina Gaussa modelira i pozadinu i prednji plan – svaka od K komponenti distribucije opisuje jedan od objekata pozadine ili prednjeg plana
- komponente (modovi) se rangiraju na temelju omjera njihove težine i standardne devijacije – što je viša i kompaktnija distribucija, veće je vjerojatnost da odgovara pozadini
- prvih B modova koji zadovoljavaju

$$\sum_{i=1}^{B} \omega_i > T$$

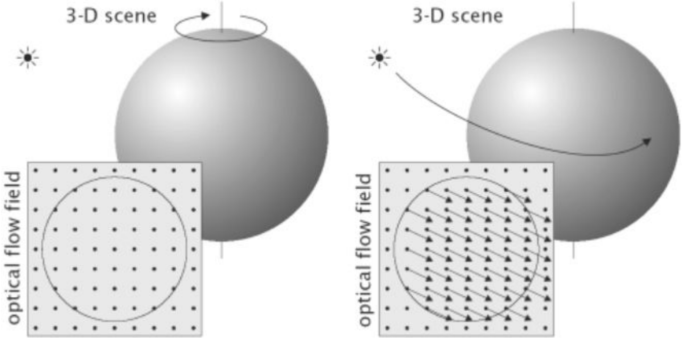
  smatraju se pozadinom
- u svakoj slici, nova vrijednost piksela pridružuje se jednom od modova
- model se ažurira tom novom vrijednošću

- mod se smatra kandidatom za podudaranje ako: $(x_t - \mu_{i,t})/\sigma_{i,t} < 2.5$
- prvi takav kandidat po rangu prihvaća se kao odgovarajući za x_t
- parametri $(\mu_{i,t}, \sigma_{i,t}, \omega_{i,t})$ ažuriraju se samo za podudarajuću distribuciju
- model se ažurira novom vrijednošću piksela korištenjem kumulativnog prosjeka
- ako se ne pronađe odgovarajući mod, najslabije rangirana distribucija se zamjenjuje novom, centriranom u x, s malom težinom i velikom varijancom

Optički tok

Optički i slikovni tok

- **Slikovni tok** (*image flow field, motion field*)
  projekcija 3D brzina u slikovnu ravninu (Singh, 1991)
- **Optički tok** (*optical flow, optic-flow*)
  polje prividnih brzina uslijed varijacije svjetlosnih uzoraka u slici
  - gibanje objekata u sceni
  - promjena osvjetljenja u sceni



Procjena optičkog toka

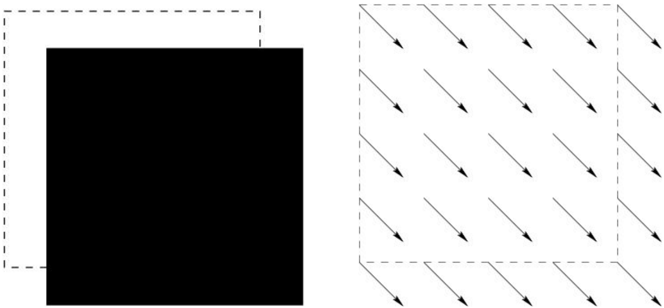
- **Procjena optičkog toka** (*optical flow estimation*)
  na raspolaganju nam je samo prostorno-vremenska razdioba vrijednosti slikovnih elemenata, na temelju koje možemo pokušati procijeniti optički tok

- **Jednakost optičkog i slikovnog toka** (teorem)
  normalne komponente (u smjeru prostornog gradijenta, tj. okomito na rub u sceni) slikovnog i optičkog toka jednake su pod sljedećim uvjetima:
  - osvjetljenje u sceni je prostorno i vremenski nepromjenjivo
  - površina objekata je savršeno difuzna (Lambertova površina)
  - gibanje objekata u sceni je isključivo translatorno

Interpretacija optičkog toka

- za interpretaciju 3D gibanja i strukture objekata u sceni potreban nam je slikovni tok
- na temelju prostorno-vremenske razdiobe vrijednosti slikovnih elemenata nastojimo procijeniti optički tok, a zatim se nadamo da su pretpostavke jednakosti optičkog i slikovnog toka u dovoljnoj mjeri zadovoljene da bismo procijenjeno polje brzina koristili kao slikovni tok

Slikovni tok



Primjene optičkog toka

- određivanje 3D strukture scene i 3D gibanja objekata u sceni
- segmentacija slike na temelju različitog kretanja objekata (*motion-based segmentation*)
- otkrivanje i izbjegavanje prepreka pri vođenju robota (autonomna navigacija) – npr. procjena vremena do kontakta (*time-to-contact, time-to-collision*)
- analiza medicinskih slika, npr. 2D ultrazvuk srca
- kompenzacija pokreta kamere (*motion-compensated image sequence enhancement*)
- sažimanje video zapisa
- praćenje objekata
- praćenje rasta biljaka

Pristupi računanju optičkog toka

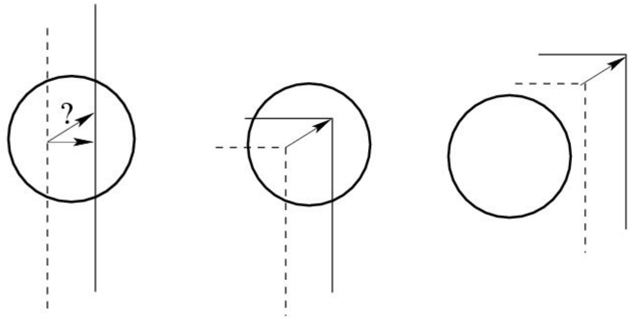
- gradijentni postupci (*gradient-based, differential*)
  - Fennema i Thompson, 1979
  - Horn i Schunck, 1981
  - Lucas i Kanade, 1981

    ...
- korelacijski postupci (*correlation-based, region-based matching*)
  - Anandan, 1989

    ...

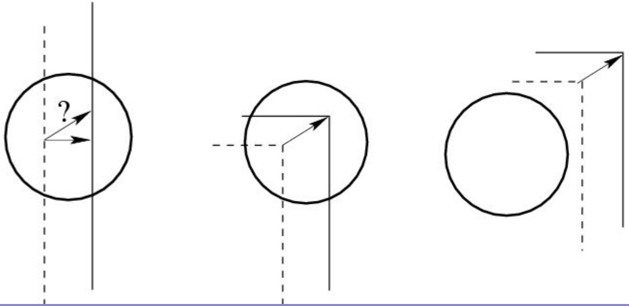
Problem lokalnog pogleda

- problem promatranja kroz rupicu (*aperture problem*)
- problem je inherentan *lokalnom* promatranju gibanja i neovisan o načinu određivanja optičkog toka
- općenito nije moguće odrediti pravu brzinu točke promatranjem njene male okoline



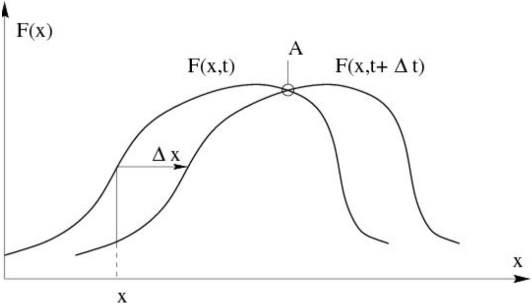
Problem lokalnog pogleda

- za točku s usmjerenim (prostornim) gradijentom intenziteta (rub, 1D diskontinuitet) moguće je dobiti samo normalnu komponentu optičkog toka
- točke s 2D diskontinuitetom intenziteta (područja slike koja imaju izražene varijacije višeg reda - vrhovi, tekstura) sadrže dovoljno informacije za određivanje pune brzine
- problem lokalnog pogleda različito se manifestira kod različitih postupaka određivanja optičkog toka



Gradijentni pristup računanju optičkog toka

- Horn i Schunck, 1981
- temelji se na uočavanju veze prostornog i vremenskog gradijenta s brzinom kretanja slike preko senzora
- ilustracija za 1D primjer
  - $F(x, t)$ – slika objekta u gibanju u trenutku t
  - slika objekta se u vremenu Δt pomakne za Δx
  - funkcija $F(x, t + \Delta t)$ ne mijenja oblik



Gradijentni pristup računanju optičkog toka

$$F(x,t) = F(x+\Delta x, t+\Delta t)$$

$$F_x(x,t) \approx \frac{\Delta F(x,t)}{\Delta x} = \frac{F(x+\Delta x, t) - F(x,t)}{\Delta x}$$

$$F_t(x,t) \approx \frac{\Delta F(x,t)}{\Delta t} = \frac{F(x, t+\Delta t) - F(x,t)}{\Delta t}$$

$$\frac{\Delta F(x,t)}{\Delta x}\frac{\Delta x}{\Delta t} = \frac{\Delta F(x,t)}{\Delta t}$$

$$F_x v = F_t$$

odnosno

$$v = \frac{F_t}{F_x}$$

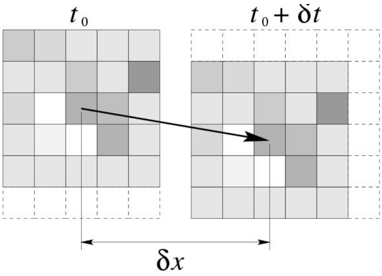
Gradijentni pristup računanju optičkog toka

$$v = \frac{F_t}{F_x}$$

- ako je $F_x \approx 0 \rightarrow$ numerički problem
  - onemogućeno izračunavanje brzine
  - manifestacija problema lokalnog pogleda
- ako je $F_t \approx 0$ (točka A) – procjena brzine je 0 (pogrešno)
  - loša procjena vremenske derivacije, koja se značajno mijenja u vremenu, pa je linearna aproksimacija neprimjerena
  - problem "vremenske zamjene" (*temporal aliasing*) zbog diskretizacije, procjene vrijede za dovoljno male Δx i Δt
  - nužna je derivabilnost promatrane funkcije

Gradijentni pristup računanju optičkog toka

- optički tok – zanima nas vremenska promjena 2D signala
- i ovdje je polazna pretpostavka da se okolina promatrane točke (2D funkcija) pomiče kao cjelina, bez promjene oblika
- translacija krutih objekta u sceni, u ravnini paralelnoj sa slikovnom ravninom, uz nepromjenjive svjetlosne uvjete $\rightarrow$ pretpostavka bi vrijedila za cijelu sliku pojedinog objekta
- za malu okolinu, i uz pretpostavku malih pomaka od slike do slike – pretpostavka je prihvatljiva



Gradijentni pristup računanju optičkog toka

- pretpostavka – okolina promatrane točke se samo pomiče, a njegov lokalni uzorak svjetlosnih intenziteta se ne mijenja (*intensity conservation*)

$$E(x, y, t) = E(x + \delta x, y + \delta y, t + \delta t)$$

- razvoj u Taylorov red

$$E(x, y, t) = E(x, y, t) + \delta x \frac{\partial E(x, y, t)}{\partial x} + \delta y \frac{\partial E(x, y, t)}{\partial y} + \delta t \frac{\partial E(x, y, t)}{\partial t} + R_2$$

- slijedi osnovna jednadžba gradijentnih postupaka računanja optičkog toka:

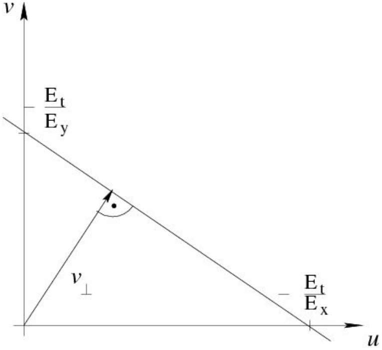
$$\frac{\partial E}{\partial x}\frac{\delta x}{\delta t} + \frac{\partial E}{\partial y}\frac{\delta y}{\delta t} + \frac{\partial E}{\partial t} = 0$$

Gradijentni pristup računanju optičkog toka

- osnovna jednadžba gradijentnih postupaka računanja optičkog toka – lokalno ograničenje optičkog toka (*optical flow constraint equation*)
$$E_x u + E_y v + E_t = 0$$
- E_x, E_y i E_t su parcijalne derivacije intenziteta slikovnih elemenata po prostoru i vremenu – mogu se procijeniti na temelju poznavanja prostorne i vremenske ovisnosti svjetlosnih intenziteta u slijedu slika
- u i v su brzine u x odnosno y smjeru
- lokalno ograničenje nije dovoljno za određivanje brzine – jedna linearna jednadžba s dvije nepoznanice – u i v

Gradijentni pristup računanju optičkog toka

- lokalno ograničenje optičkog toka može se grafički prikazati (*optical flow constraint line*)



$$E_x u + E_y v + E_t = 0$$

Gradijentni pristup računanju optičkog toka

- svaka točka u slici ima svoj lokalni pravac ograničenja, sve brzine koje leže na tom pravcu zadovoljavaju jednadžbu ograničenja
- ako se ograničimo na računanje komponente brzine koja je okomita na pravac ograničenja, možemo je izračunati u svakoj točki

$$\|\mathbf{v}_\perp\| = \frac{|E_t|}{\sqrt{{E_x}^2 + {E_y}^2}}.$$

- ona odgovara projekciji brzine na vektor prostornog gradijenta (sjetimo se problema lokalnog pogleda)
- izraz pokazuje da je određivanje brzine neprikladno u točkama u kojima je prostorni gradijent malen !

Gradijentni postupci s globalnom optimizacijom

- Horn i Schunck, 1981
- postupak se temelji na jednadžbi lokalnog ograničenja
- problem neodređenosti rješava se uvođenjem dodatnog ograničenja koje mora zadovoljavati izračunato polje optičkog toka – glatkoća
- susjedne točke moraju imati slične brzine
- jasno da ovaj uvjet ne vrijedi u točkama diskontinuiteta polja brzina – npr. na granici između dva objekta koji se različito gibaju
- odstupanje od glatkoće u nekoj točki izraženo je veličinom promjene brzine u x i y smjeru

$$\varepsilon_c^2 = \|\nabla u\|^2 + \|\nabla v\|^2 = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2$$

Gradijentni postupci s globalnom optimizacijom

- kombiniranje mjere odstupanja od gradijentne jednadžbe i glatkoće polja optičkog toka $\rightarrow$ funkcional koji je potrebno minimizirati (po cijeloj slici !)

$$\varepsilon^2 = \int\int \left\{ \alpha^2 \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \right] + [E_x u + E_y v + E_t]^2 \right\} dxdy$$

- primjenom varijacijskog računa dobiju se po dvije diferencijalne jednadžbe (za svaku točku slikovne ravnine !)

$$E_x^2 u + E_x E_y v = \alpha^2 \nabla^2 u - E_x E_t$$

$$E_x E_y u + E_x^2 v = \alpha^2 \nabla^2 v - E_y E_t$$

Gradijentni postupci s globalnom optimizacijom

- radimo s diskretnim slikama – potrebno je procijeniti derivacije na temelju vrijednosti slikovnih elemenata promatrane okoline, odnosno na temelju vrijednosti slikovnih elemenata susjednih slika
- Horn i Schunck su procjenu temeljili na konačnim razlikama susjednih slikovnih elemenata
- izveli su iterativni postupak rješavanja navedenih diferencijalnih jednadžbi:

$$u^{n+1} = \overline{u}^n - E_x(E_x\overline{u}^n + E_y\overline{v}^n + E_t)/(\alpha^2 + E_x^2 + E_y^2)$$

$$v^{n+1} = \overline{v}^n - E_y(E_x\overline{u}^n + E_y\overline{v}^n + E_t)/(\alpha^2 + E_x^2 + E_y^2)$$

- $\overline{u}^n$ i $\overline{v}^n$ – prosječne vrijednosti brzine u promatranoj okolini, dobivene u prethodnoj iteraciji

Gradijentni postupci s globalnom optimizacijom

- glađenje polja optičkog toka (u svim smjerovima, neovisno o rubovima u slici !) rezultira zamućenjem na granici između objekata $\rightarrow$ nepovoljno
- manifestacija problema lokalnog pogleda – isčezava drugi član u iterativnim izrazima

$$u^{n+1} = \overline{u}^n - E_x(E_x\overline{u}^n + E_y\overline{v}^n + E_t)/(\alpha^2 + E_x^2 + E_y^2)$$

  - nova procijenjena vrijednost pojedine komponente brzine je prosjek procijenjenih brzina okoline iz prethodnog koraka
  - procjene brzine dobivene na rubovima uniformnih područja (postupno) propagiraju u njihovu unutrašnjost

Gradijentni postupci s globalnom optimizacijom

Gradijentni postupci s lokalnom optimizacijom

- drugi pristup korištenju gradijentne jednadžbe ograničenja optičkog toka
  - lokalno rješavanje nedovoljne određenosti te jednadžbe
  - ako se promatra neko područje u slici, za svaku točku tog područja može se postaviti jedna jednadžba ograničenja optičkog toka
  - uz pretpostavku da sve točke promatrane okoline imaju jednaku brzinu, dobivamo sustav jednadžbi s dvije nepoznanice
  - rješavanje npr. metodom najmanjeg kvadratnog odstupanja
  - može se koristiti i neki drugi model promjene brzine unutar područja (npr. linearna promjena)
- Lucas i Kanade, 1981 – u eksperimentalnim ocjenama pokazao se kao jedan od najboljih algoritama – često korišten

Gradijentni postupci s lokalnom optimizacijom

- Lucas i Kanade
- polazna pretpostavka je jednakost brzina svih točaka promatrane okoline
  - traži se brzina (u, v) koja minimizira težinsku sumu kvadrata odstupanja od jednadžbe lokalnog ograničenja optičkog toka po okolini promatrane točke

$$\varepsilon = \sum_{(x,y)\in R} w(x, y)(E_x u + E_y v + E_t)^2$$

  - težinska funkcija $w(x, y)$ obično daje veći utjecaj sredini okoline
  - zbog jednostavnosti, možemo pretpostaviti $w(x, y) = 1$
- primjer rješavanja problema najmanjeg kvadratnog odstupanja

Gradijentni postupci s lokalnom optimizacijom

- promatra se neka slikovna okolina R, traže se takvi parametri u i v da suma kvadratnih odstupanja bude čim manja, tj. želimo minimizirati pogrešku ε

$$\varepsilon = \sum_{(x,y)\in R} (E_x u + E_y v + E_t)^2$$

- derivirat ćemo funkciju po u odnosno v $\rightarrow$ dobivamo sustav s dvije jednadžbe (koliko parametara, toliko derivacija i toliko jednadžbi):

$$\frac{\partial \varepsilon}{\partial u} = \sum 2(E_x u + E_y v + E_t)E_x = 0$$

$$\frac{\partial \varepsilon}{\partial v} = \sum 2(E_x u + E_y v + E_t)E_y = 0$$

Gradijentni postupci s lokalnom optimizacijom

- nakon dijeljenja s 2 i malo sređivanja dobiva se sustav:

$$\sum E_x^2 u + \sum E_x E_y v + \sum E_x E_t = 0$$

$$\sum E_x E_y u + \sum E_y^2 v + \sum E_y E_t = 0$$

- isti sustav može se zapisati i u matričnom obliku:

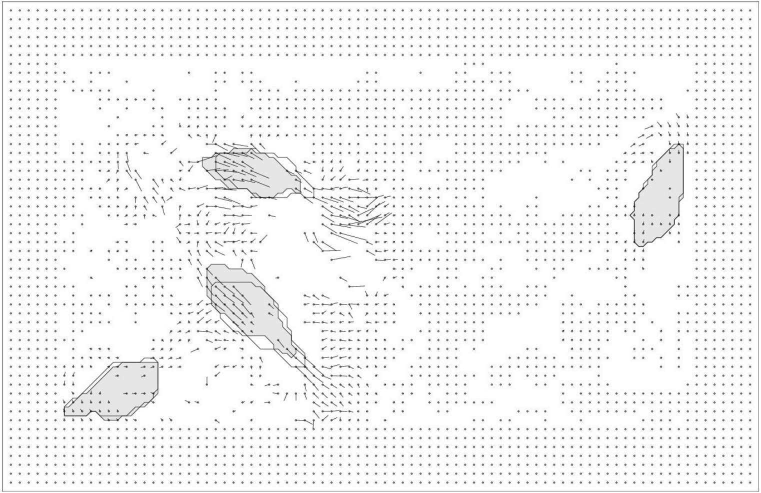
$$\begin{bmatrix} \sum E_x^2 & \sum E_x E_y \\ \sum E_x E_y & \sum E_y^2 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum E_x E_t \\ \sum E_y E_t \end{bmatrix}$$

Gradijentni postupci s lokalnom optimizacijom

$$\begin{bmatrix} \sum w_i E_x^2 & \sum w_i E_x E_y \\ \sum w_i E_x E_y & \sum w_i E_y^2 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} \sum w_i E_x E_t \\ \sum w_i E_y E_t \end{bmatrix}$$

- ako je prostorni gradijent u promatranoj točki dovoljno velik, i njegov smjer dovoljno varira u promatranoj okolini, linearni sustav jednadžbi je dobro određen i ima jedinstveno rješenje $\rightarrow$ može se odrediti pouzdana vrijednost optičkog toka u toj točki
- ako je vrijednost gradijenta mala (uniformna područja), ili mu je smjer ujednačen unutar područja (usmjeren gradijent na bridu u slici), ne može se odrediti pouzdana procjena optičkog toka
- problem lokalnog pogleda manifestira se pojavom singularne matrice sustava jednadžbi
- razne modifikacije osnovnog algoritma – primjena robustne statistike, proširenje promatrane okoline vremenskom dimenzijom, primjena različitih filtara za glađenje, načina određivanje derivacija i sl.

Gradijentni postupci s lokalnom optimizacijom



Korelacijski postupci određivanja optičkog toka

- pretpostavka očuvanja lokalne razdiobe svjetlosnih intenziteta u okolini svake promatrane točke
- pomicanjem točke između dvije susjedne slike, cijela njena okolina pomiče kao cjelina
- određivanje optičkog toka u svakoj točki, odnosno njenog pomaka od slike do slike, može se svesti na problem traženja pomaka kojim se postiže najbolje podudaranje okolina promatrane točke u dvije susjedne slike u sekvenci
- *korelacijski prozor* – okolina točke za koju se traži pomak
- *prozor pretraživanja* – određuje raspon pomaka u kojem tražimo najbolje podudaranje
    - veličina prozora pretraživanja određuje najveći pomak (brzinu) koji može biti određen – mora biti usklađen s očekivanom brzinom objekata u slici

Korelacijski postupci određivanja optičkog toka

- mjera podudaranja (ili nepodudaranja)
  - normalizirana križna korelacija
  - suma kvadratnih odstupanja vrijednosti odgovarajućih slikovnih elemenata
- jedinstveno rješenje kojim se minimizira mjera odstupanja može se postići za točke koje imaju prepoznatljivu 2D strukturu
  - za jednolika područja to nije moguće
  - manifestacija problema lokalnog pogleda – više kandidata s podjednakom mjerom podudaranja
  - za točke u kojima postoji jednodimenzionalni diskontinuitet (rub u slici), pojavljuje se neodređenost komponente pomaka u smjeru ruba

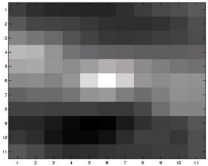
Korelacijski postupci određivanja optičkog toka

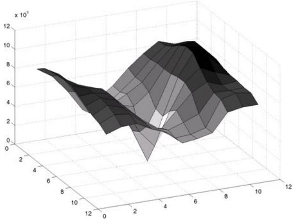
- Anandan, 1989 – najpoznatiji korelacijski postupak
  - hijerarhijska shema u kojoj se oblikuje rezolucijska piramida slika
  - u svakoj razini, počevši od najgrublje, za svaki slikovni element određuje se najpovoljniji pomak između dvije susjedne slike u sekvenci
  - mjera podudaranja – suma kvadratnih odstupanja po elementima okoline
  - računanjem sume kvadratnih odstupanja za svaki potencijalni pomak dobiva se tzv. SSD (*sum of squared differences*) površina:

$$SSD(\delta x, \delta y) = \sum_{i=-n}^{n} \sum_{j=-n}^{n} (I_1(x+i, y+j) - I_2(x+\delta x+i, y+\delta y+j))^2$$

Korelacijski postupci određivanja optičkog toka

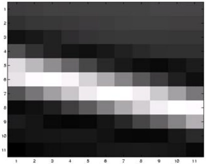
- u SSD matrici pronalazi se najmanja vrijednost – odgovarajući pomak se promatranoj točki pridružuje kao vektor optičkog toka
- interpolacijom površine može se postići i određivanje pomaka s točnošću finijom od slikovnog elementa (*subpixel precision*)

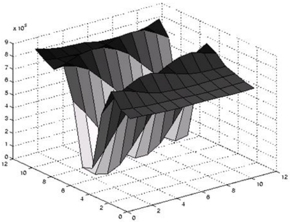




Korelacijski postupci određivanja optičkog toka

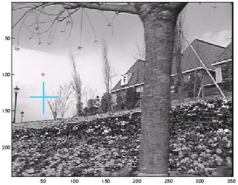
- da bi se smanjile pogreške koje su posljedica pojave problema lokalnog pogleda, uvodi se mjera pouzdanosti, temeljena na zakrivljenosti SSD površine u mjestu pronađenog minimuma

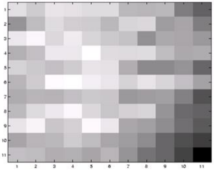


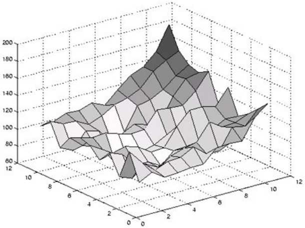


Korelacijski postupci određivanja optičkog toka

- za jednolika područja SSD površina je glatka – ne možemo pouzdano odrediti najbolji pomak





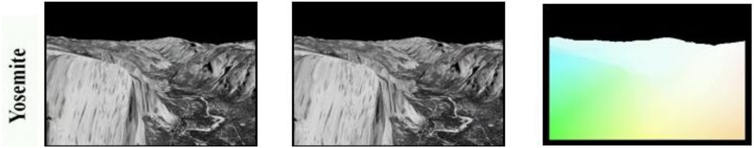


Korelacijski postupci određivanja optičkog toka

- hijerarhijska shema:
  - određuje se optički tok u jednoj razini rezolucijske piramide, određuju se mjere pouzdanosti, obavlja se glađenje
  - dobiveni vektori optičkog toka projiciraju se u odgovarajuće točke u sljedećoj finijoj razini rezolucijske piramide – koriste se za postavljanje prozora pretraživanja u očekivani položaj za svaku točku $\rightarrow$ profinjavanje procjena iz prethodne razine
  - omogućeno određivanje optičkog toka i u slučaju velikih pomaka, uz relativno male dimenzije prozora pretraživanja

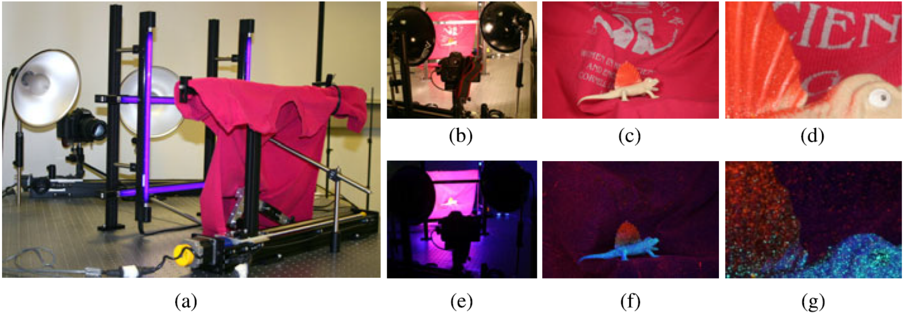
Ocjena uspješnosti

- usporedba različitih postupaka za računanje optičkog toka
- potrebno je poznavati stvarni optički tok za promatrani slijed slika (*ground-truth*) – obično umjetno generirane scene



Ocjena uspješnosti

- primjer: generiranje sekvenci s poznatim referentnim poljem pomaka – skrivena fluorescentna tekstura, osvjetljavanje UV svjetlom (https://vision.middlebury.edu/flow/)



Ocjena uspješnosti

- pitanje mjere podudaranja (odnosno mjere pogreške)
  - kutna pogreška (*angular error*)
  - apsolutna pogreška novog položaja točke (*flow endpoint error*)
- kutna pogreška (Barron, 1994)
  - kutno odstupanje izračunate od referentne brzine u svakoj točki
  - vektori brzine se pritom proširuju dodatnom koordinatom veličine 1
  - ako je $\mathbf{v}_{ref} = (\hat{u}, \hat{v}, 1)^T$ referentna, a $\mathbf{v} = (u, v, 1)^T$ izračunata brzina u nekoj točki, kutna pogreška je definirana kao:

$$\varepsilon_\phi = arccos\left(\frac{\mathbf{v}_{ref}^T \cdot \mathbf{v}}{\|\mathbf{v}_{ref}\|\|\mathbf{v}\|}\right)$$

  - proširenje dodatnom koordinatom rješava problem dijeljenja s 0, a na neki način kažnjava i odstupanje u duljini referentnog i procijenjenog vektora optičkog toka

Ocjena uspješnosti

- neki nedostaci kutne mjere pogreške
  - velika apsolutna pogreška može rezultirati i malom mjerom odstupanja (posebno kod velikih vektora pomaka)
  - nesimetričnost – jednako apsolutno odstupanje od referentne brzine može dati vrlo različite mjere pogreške
  - za male pomake odstupanje se jače kažnjava
- apsolutna pogreška (Otte i Nagel, 1994)
  - kao mjeru pogreške koriste duljinu vektora razlike referentne i izračunate brzine

$$\varepsilon_{\Delta} = \|\mathbf{v}_{ref} - \mathbf{v}\|$$

- prilikom usporedbe algoritama za procjenu optičkog toka najčešće se iskazuju obje pogreške, kao prosjek kroz cijelu sliku, te standardnu devijaciju
- iskazuje se i postotak piksela za koje je pogreška manja od nekog zadanog praga (npr. 3 piksela)

Domaća zadaća

- pročitati članak Viole i Jonesa o detekciji objekata:
  Paul Viola, Michael J. Jones. Robust Real-Time Face Detection. International Journal of Computer Vision 57, 137–154 (2004).
  `https://link.springer.com/article/10.1023/B:VISI.0000013087.49260.fb`
- u terminu sljedećeg predavanja (11. studenoga 2022.) imat ćemo diskusiju o tom članku