

# Lokalizacija proizvoljno zakrenutog teksta u slikama

Filip Zelić

Fakultet Elektrotehnike i računarstva

*filip.zelic@fer.hr*

July 8, 2015

# Opis zadatka



# Algoritmi i matematičke metode

Metode korištene za naglašavanje aspekata slike:

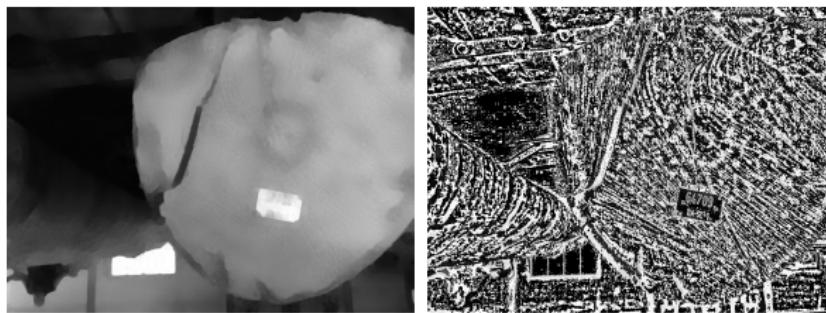
- Strukturno-teksturna dekompozicija
- Upravljivi filter

Metode korištene za daljnju obradu nakon naglašavnja:

- Transformacija širine poteza
- Morfološka obrada binarne slike

# Struktorno-teksturna dekompozicija

- Uklanjanje šuma primjenjujemo na izvornu sliku
- Ideja je dobiti potpuno zamućene dijelove slike na kojima je tekst
- Sljedeći korak je oduzimanje zamućene slike od izvorne slike
- Koristimo metodu Denoise\_TLV1 koja prima dva parametra
- Najbolje rezultate dobivamo za  $brojIteracija = 40$  i  $lambda = 0.65$

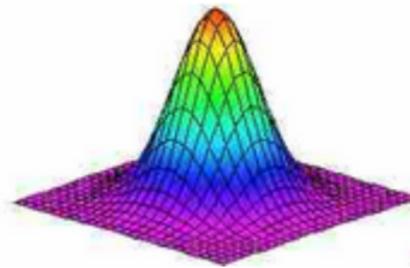


- Metoda rezultira velikim šumom i nije korištena, ali je moguće s njom detektirati markice na slici debla

# Digitalni filter

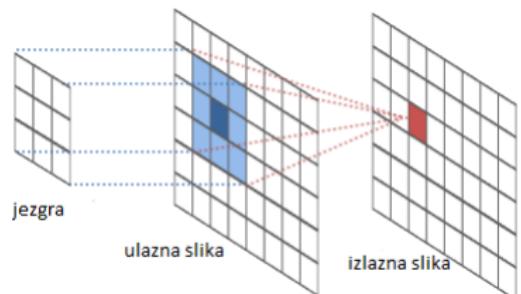
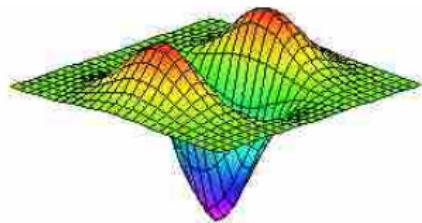
- Digitalno filtriranje je proces provođenja matematičkih operacija nad signalom
- Cilj je potisnuti ili naglasiti neke aspekte slike
- Na slici je prikazana dvodimenzionalna Gaussova funkcija:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (1)$$



# Dvodimenzionalni linearni filter

- Pri konstruiranju filtra parametar  $\sigma$  utječe na veličinu jezgre filtra
- Desna slika prikazuje konvoluciju jezgre filtra s ulaznom slikom
- Upravljivi filter koristi drugu derivaciju Gaussove funkcije
- Komponente koje su nam potrebne za izračunavanje upravljivog filtra:
  - 1 Filtar  $G_{xx}$
  - 2 Filtar  $G_{yy}$
  - 3 Filtar  $G_{xy}$
- Lijeva slika prikazuje usmjerenu drugu derivaciju Gaussove funkcije  $G_{xx}$

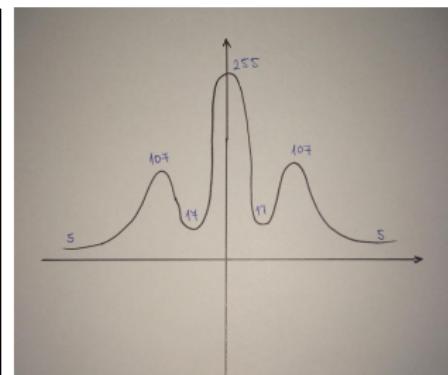


# Upravljivi filter

- Upravljivi filter se sastoji od orientacije i intenziteta
- Filter  $G_\theta$  daje maksimalan ili minimalan odziv za neki kut rotacije.
- Računa se koristeći ranije spomenute komponente filtre  $G_{xx}$ ,  $G_{yy}$ ,  $G_{xy}$  prema formuli:

$$G_\theta(x, y) = G_{xx} \cos^2 \theta + G_{yy} \sin^2 \theta - 2G_{xy} \cos \theta \sin \theta \quad (2)$$

# Upravljivi filter



- Desna slika dolje prikazuje 1.5% nasvjetlijih piksela histograma filtra



# Transformacija širine poteza

- Tekst ima uglavnom konstantnu širinu poteza
- Metoda pretvara vrijednost piksela u najvjerojatniju širinu poteza
- Moguće je detektirati tekst bez obzira na njegovo skaliranje, font, smjer i jezik

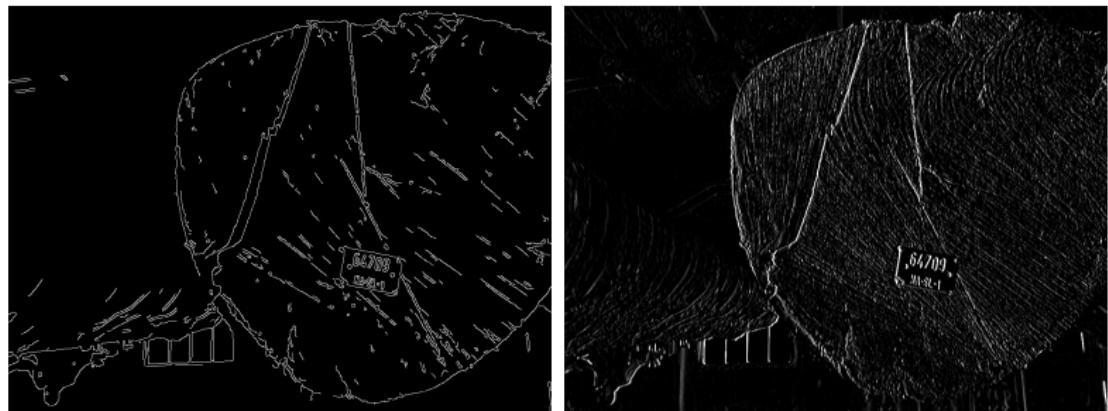


- Pikseli p i q su granični pikseli između kojih se računa udaljenost (širina poteza)

# Transformacija širine poteza

Za provođenje algoritma potrebno je:

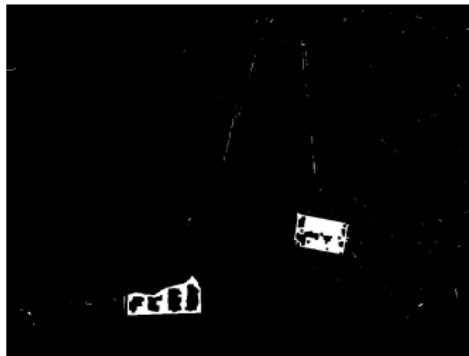
- učitati izvornu sivu sliku i nad njom primijeniti detektor rubova
- izračunati gradijent slike u x i y smjeru
- vrijednosti piksela izlazne slike ove metode postaviti na  $\infty$



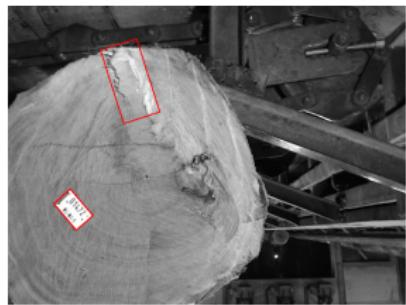
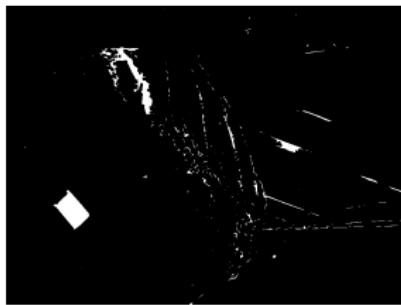
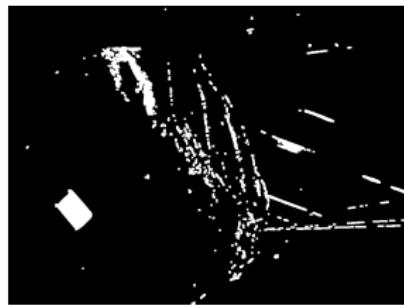
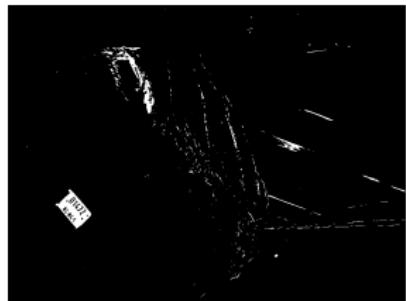
- Metoda nije davala dobre rezultate zbog čega smo koristili morfološku obradu binarne slike

# Morfološka obrada binarne slike

- Za poboljšanje odziva vršimo postupak „zatvaranja“ koji se sastoji od dva postupka – dilatcije i erozije
- Prvo se vrši dilatacija, zatim erozija zadane binarne slike
- Dilatacije je postupno proširenje regija piksela
- Erozija je erodiranje piksela koji se nalaze na granicama
- Nakon morfološke obrade pronalazimo minimalne zarotirane pravokutnike u slici

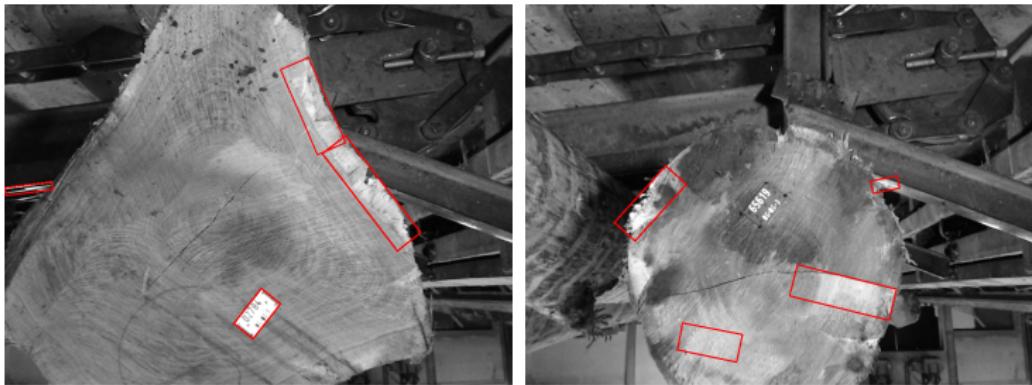


# Cijeli proces obrade slike



# Rezultati

- Broj uspješnih detekcija je najbolji za parametar  $\sigma = 1.2$
- Na slikama se pojavljuju krive detekcije (engl. false positive)
- Problem propuštenih detekcije (engl. false negative)
- Propuštene detekcije se pojavljuju na dijelovima kada osvjetljenje smanji kontrast između pločice, brojki i pozadine



# Rezultati

Podaci	Rezultat
Preciznost	29.3%
Odziv	82.5%
Vrijeme	0.545s

Table: Rezultati evaluacije slike debla s pločicom

- Preciznost niža jer postoji samo jedna pločica
- Osvijetljenje te prozori uzrokuju krive detekcije
- Odziv je najvažniji parametar jer pri integraciji sa sustavom za detekciju, potrebno minimizirati propuštene detekcije

# Rezultati

- Kompleksna pozadina utječe na provođenje evaluacije
- Odabrali smo 20 slika jednostavnije pozadine
- Primjer je lišće koje uzrokuje veliki broj lažnih detekcija zbog zbog velike promjene intenziteta



# Rezultati

Podaci	Rezultat
Preciznost	60.6%
Odziv	75.1%
Vrijeme	0.918s

Table: Rezultati evaluacije slike teksta u urbanoj okolini

- Rezultati za preciznost su bolji, više detekcija tekstualnih regija
- Odziv niži jer postoji i više propuštenih detekcija
- Vrijeme izvođenja dulje jer je potrebno označiti više dijelova slike

# Zaključak

Sljedeći koraci:

- Spajanje lokalizacije teksta sa sustavom za detekciju brojki i slova
- Jednostavno sučelje za korištenje cjelokupnog sustava

Primjena:

- Očitavanje raznih serijskih brojeva proizvoda
- Cjelokupni sustav za prepoznavanje i detektiranje koji bi imao uvid u stanje skladišta

# Pitanja ?