

Univerza v Mariboru
Fakulteta za elektrotehniko, računalništvo in informatiko



Doktorska disertacija

OCENJEVANJE MODELOV IN POSTOPKOV ZA 3D SLEDENJE GIBANJU ČLOVEŠKEGA TELESA S STEREOKAMERO

Kandidat: mag. Matjaž Divjak

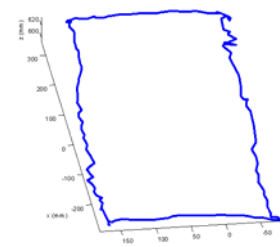
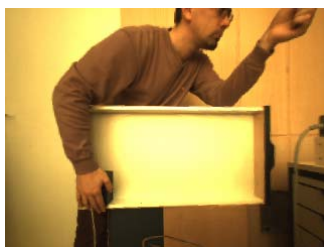
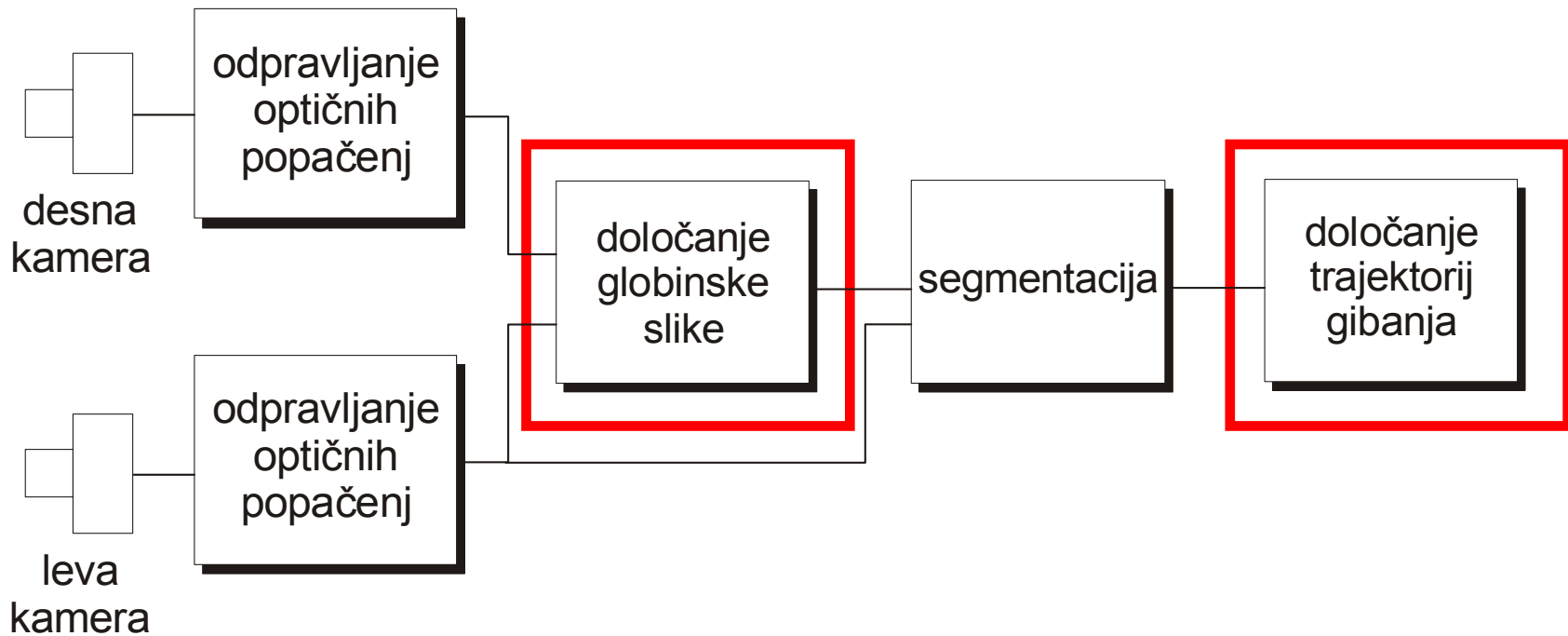
Mentor: red. prof. dr. Damjan Zazula

14. oktober, 2005

Vsebina predstavitve

- Definicija problema
- Pregled področja
- Cilji in omejitve
- Teza disertacije
- Določanje globine s postopkom CKC2D
- Ocenjevanje natančnosti video sledenja
- Rezultati
- Povzetek prispevkov

Problem: sledenje gibanju človeškega telesa



Postopki za določanje globinskih slik - pregled

- Običajno so sestavljeni iz 4 korakov:
 - računanje stroškov ujemanja pikslov
 - združevanje stroškov v regiji
 - iskanje najboljšega ujemanja med regijami
 - izboljševanje disparitet

- Metrike za oceno razlik med objekti:
 - absolutna (MAD) ali kvadratna (MSE) razlika intenzitete
 - normalizirana prečna korelacija
 - primerjava polj gradientov
 - primerjava faznih zamikov v Fourierjevi transformaciji

Postopki za določanje globinskih slik - pregled

- Združevanje stroškov:
 - posamezni piksli
 - 2D okna: konstantna ali dinamična
 - 3D okna
 - poljubne regije
- Iskanje ujemanja:
 - lokalno: *winner-takes-all*
 - globalno: minimizacija energijske funkcije
 - dinamično programiranje
 - simulirano ohlajanje
 - klestenje grafov
 - nevronske mreže

Ocenjevanje natančnosti sledenja - pregled

- Pomanjkanje zanesljivih referenčnih podatkov
 - prosto gibanje objektov – kaj uporabiti za referenco?
- Običajne rešitve:
 - ročno označevanje položaja
 - polavtomatski postopki
 - primerjava z natančnejšo sledilno napravo:
 - elektromagnetni sledilnik
 - laserski prebirknik



Cilji in omejitve

- Osnovni cilj: natančno sledenje gibanju obraza in dlani z navadno stereokamero v realnem okolju
- Omejitve:
 - kalibrirana kamera
 - dve vhodni sliki
 - epipolarna geometrija
 - Lambertske površine
 - brez eksplicitnega modela telesa
 - pred kamero nastopa ena oseba

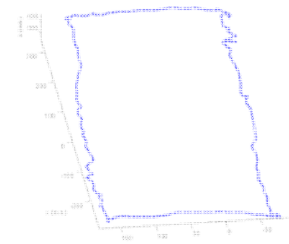
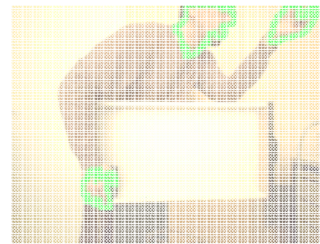
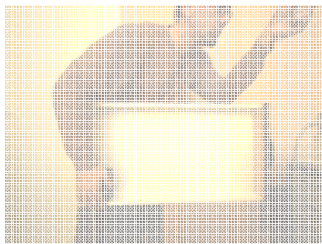
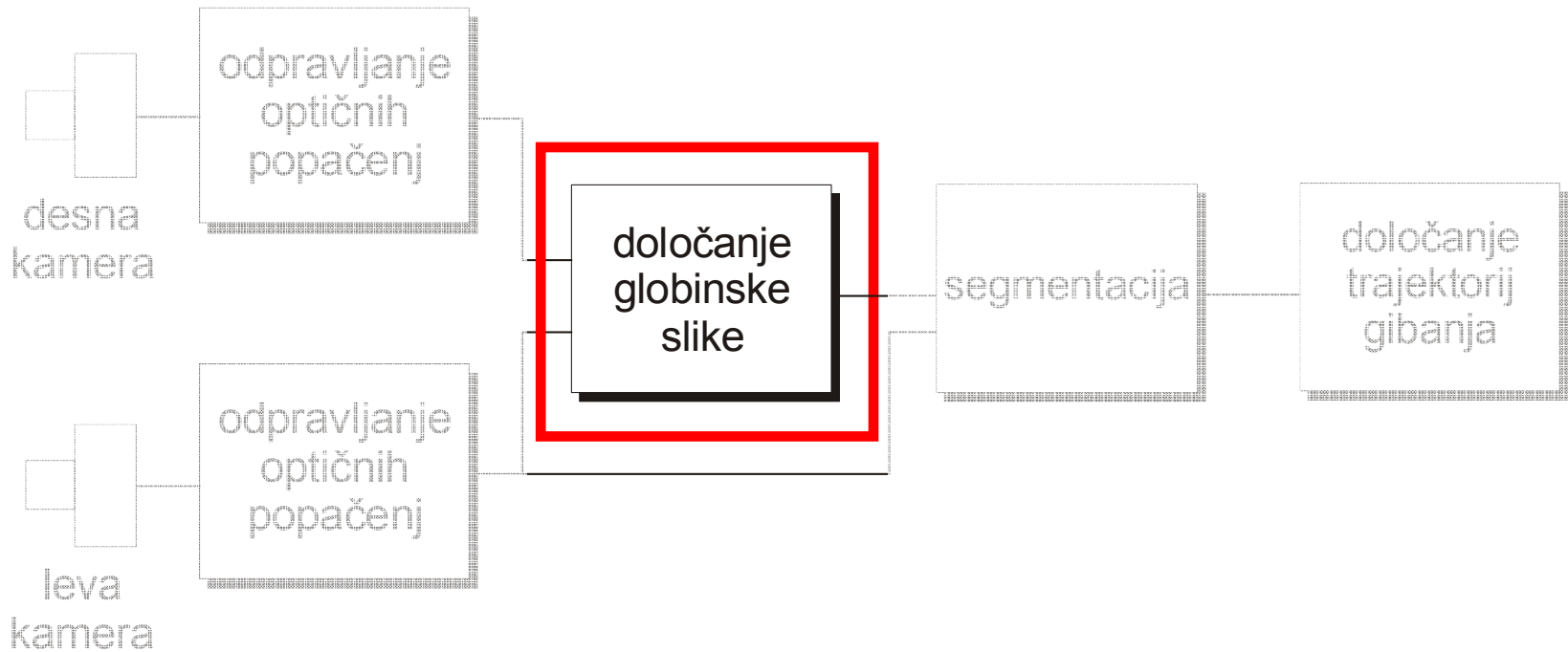
Teza

“Z izboljšavo kvalitete globinskih slik in algoritma za sledenje objektom s predikcijo je mogoče doseči takšno zanesljivost stereo sledilnega sistema, da je uporaben v realnih aplikacijah za navigacijo v navideznih okoljih brez uporabe dodatne opreme.”

Hipoteze:

- Določanje globinske slike s slepo ločitvijo izvorov je bolj robustno kot klasični postopek z iskanjem korelacije.
- Nadgradnja sledilnega postopka s predikcijsko-korekcijsko shemo bistveno poveča robustnost in natančnost sledenja.
- Preprosta stereokamera v kombinaciji z izboljšanim sledilnim algoritmom je primerljiva z obstoječimi rešitvami za navigacijo.
- Razvijemo lahko statistični postopek za analizo natančnosti video sledenja.

Določanje globinske slike



Slepa ločitev izvorov s kompenzacijo konvolucijskega jedra (CKC)

□ 1D CKC:

- iskanje položaja ponavljajočih se komponent
- velika odpornost na šum

□ 2D CKC – stereo slike:

- iskanje skupnih regij na levi (L) in desni (D) sliki

- matematični model: $\mathbf{Y} = \mathbf{H} \mathbf{X}$

- \mathbf{Y} : vektorizirana slika, sestavljena iz ozadja in množice neprekrivajočih se objektov

- \mathbf{H} : intenziteta regij

- \mathbf{X} : položaj regij (binarno)


$$\mathbf{Y} = \text{[black bar] [light gray] [medium gray] [dark gray] [black bar]}$$

$$\mathbf{H} = \text{[light gray] [medium gray] [dark gray]}$$

$$\mathbf{X} = \begin{array}{|c|c|c|c|c|c|} \hline 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \hline \end{array}$$

Določanje globine s postopkom CKC2D

- Istoležni vrstici L in D slike združimo v skupno matriko meritev \mathbf{Y}_{LD} :

$$\mathbf{Y}_{LD} = [\mathbf{Y}_L, \mathbf{Y}_D] =$$


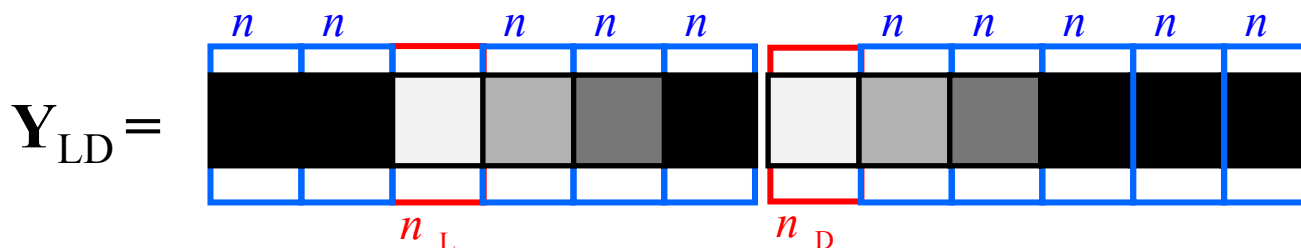
$$\mathbf{X}_{LD} =$$

0	0	1	0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0	1	0	0	0

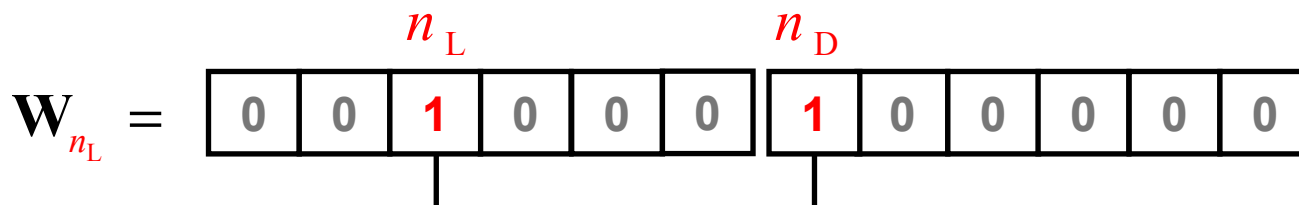
$$\mathbf{Y}_{LD} = \mathbf{H} \mathbf{X}_{LD}$$

Določanje globine s postopkom CKC2D

- Določanje ujemanja med L in D piksli:



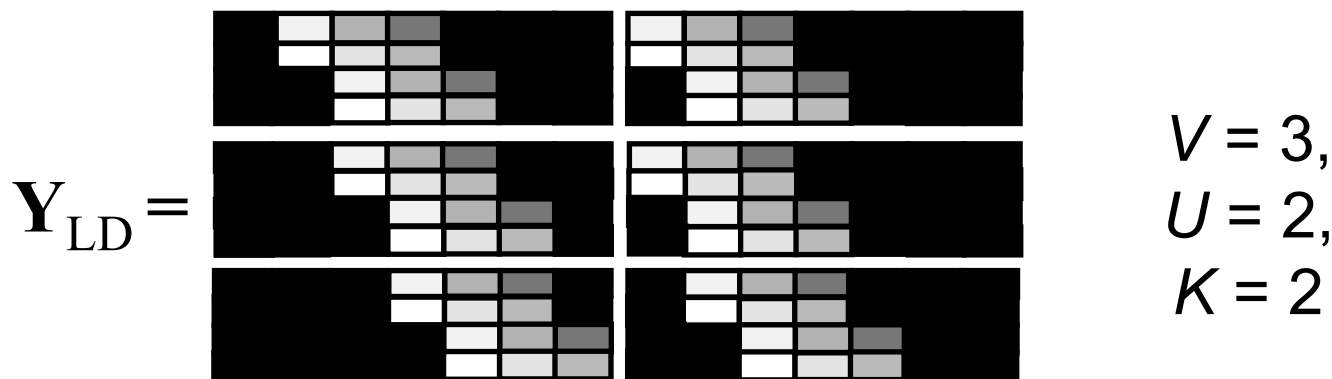
$$\begin{aligned} \mathbf{W}_{n_L}(n) &= \mathbf{Y}_{LD}^T(n_L) \mathbf{R}_{Y_{LD}}^{-1} \mathbf{Y}_{LD}(n) \\ &= \mathbf{X}_{LD}^T(n_L) \mathbf{H}^T \mathbf{H}^{-1} \mathbf{R}_{X_{LD}}^{-1} \mathbf{H}^{-1} \mathbf{H} \mathbf{X}_{LD}(n) \\ &= \mathbf{X}_{LD}^T(n_L) \mathbf{R}_{X_{LD}}^{-1} \mathbf{X}_{LD}(n) \end{aligned}$$



$$\text{dispariteta} = n_L - n_D$$

Določanje globine s postopkom CKC2D

- Težava: premalo vrstic v \mathbf{Y}_{LD} \rightarrow nepopolna kompenzacija \mathbf{H}
- Rešitev: dodajanje sintetičnih vrstic (navideznih meritev):
 - $V-1$ sosednjih slikovnih vrstic
 - $U-1$ lihih potenc
 - $K-1$ zakasnenih meritev

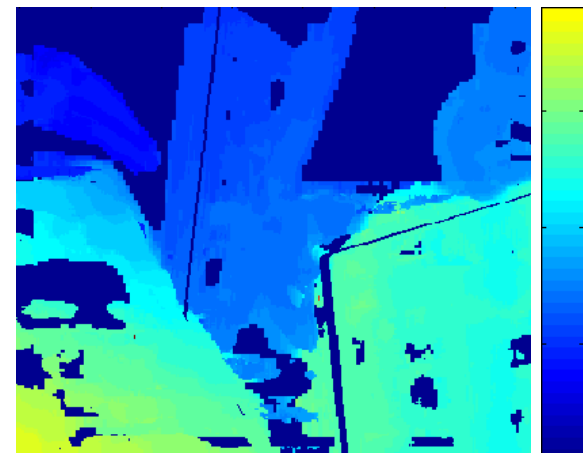
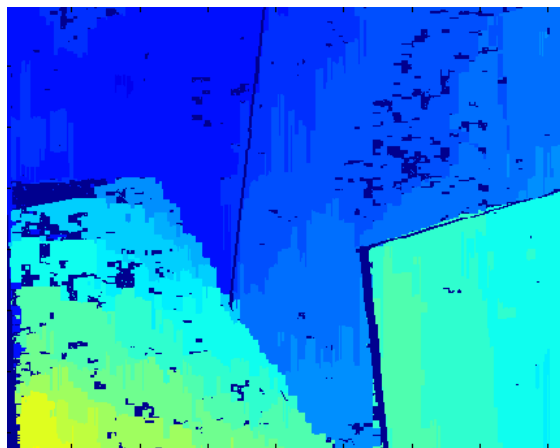
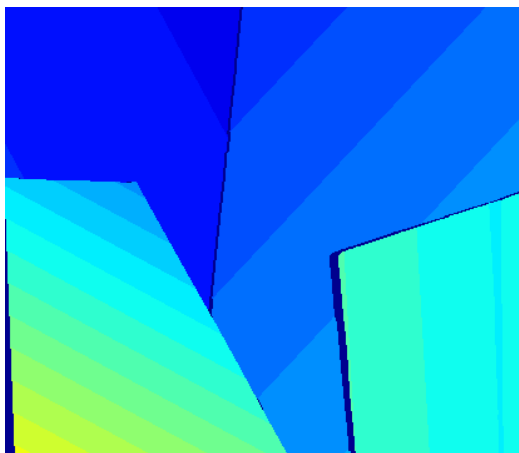
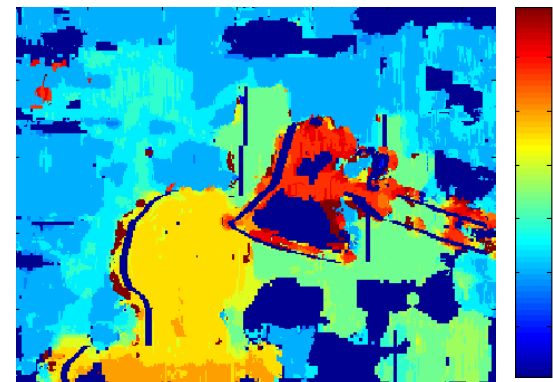
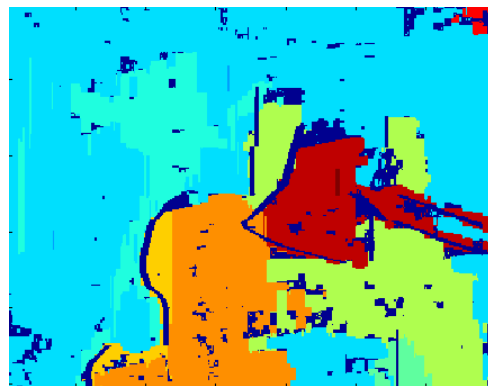


- Pogoji: $V K U \geq \text{rang}(\mathbf{R}_{\mathbf{Y}_{LD}})$

Določanje globine s postopkom CKC2D

- Omejitve pri določanju globine:
 - dovoljen razmik: $disp_{MAX} = 32$
 - preverjanje ujemanja: $L \rightarrow D, D \rightarrow L$
- Hierarhično delovanje:
 - postopek ponovimo z drugačnimi V, K, U
 - disparitete vsake ponovitve shranimo
- Iz shranjenih vrednosti izberemo končne disparitete:
 - zanesljive disparitete: enake na vseh hierarhičnih nivojih
 - preostali piksli: upoštevamo zanesljive sosede

Določanje globine s CKC2D - primerjava z rezultati korelacijskega algoritma



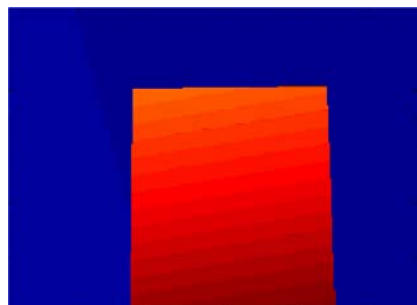
Referenčne globinske slike
iz baze Middlebury [1]

Algoritem CKC2D,
4 hierarhični nivoji

Korelacijski algoritem SVS,
okno 7×7 pikslov

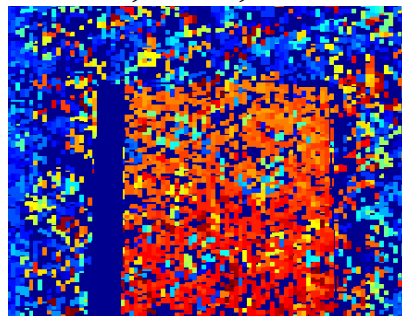
Določanje globine s CKC2D - vpliv šuma

- Globinske slike pri 10 dB šuma:

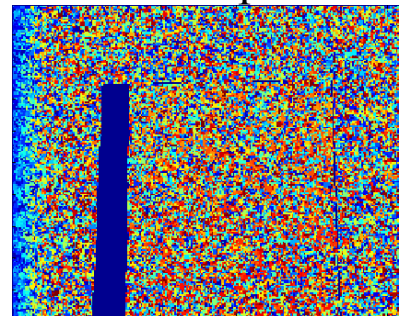


referenčna
globinska slika

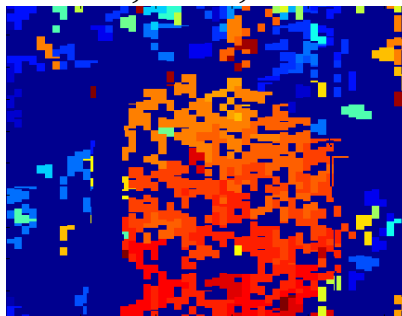
$V=3, K=3, U=11$



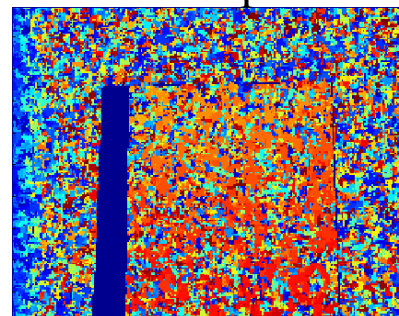
okno 3×3 pikslov



$V=5, K=5, U=11$



okno 5×5 pikslov

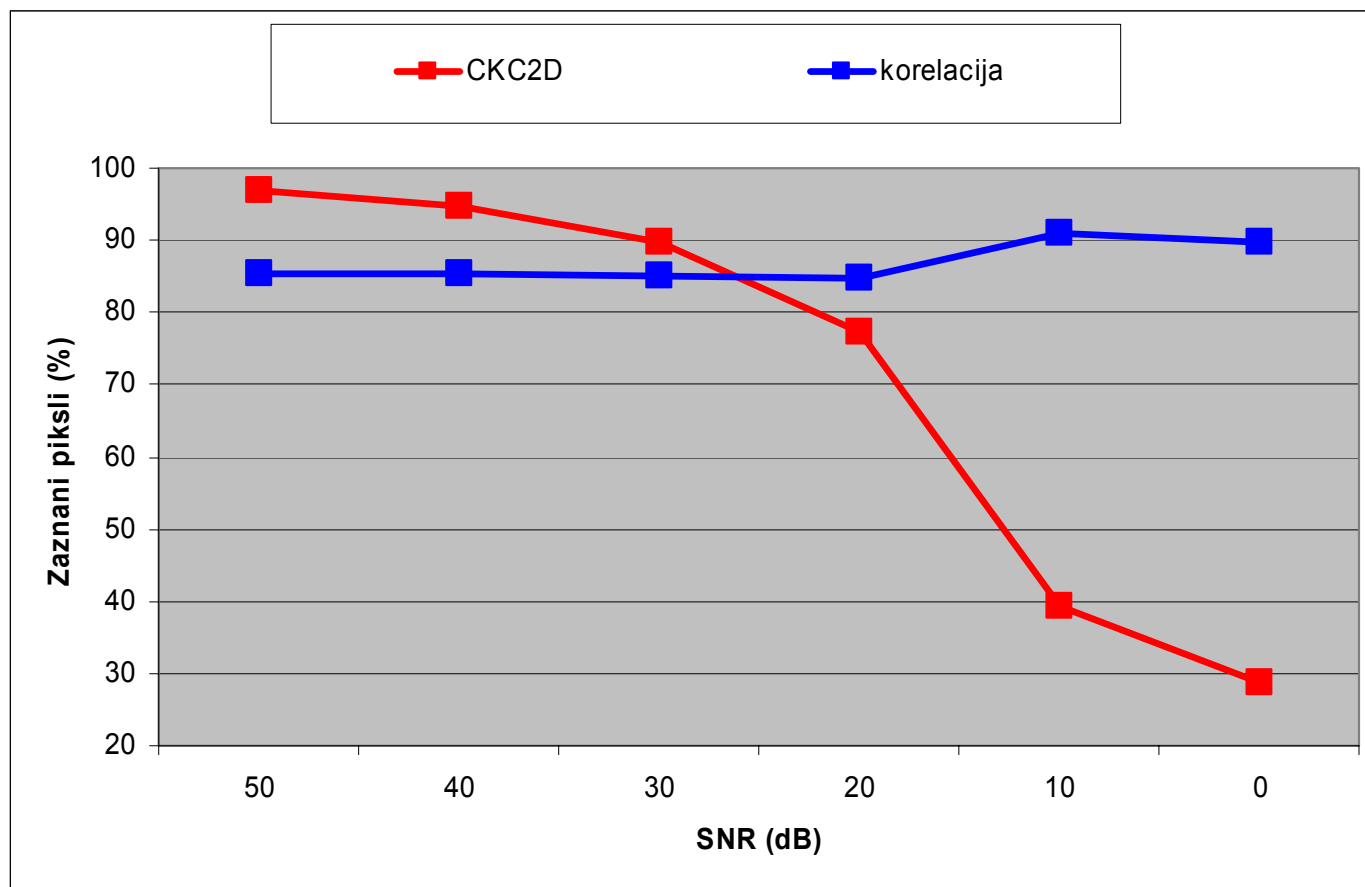


1 hierarhični nivo
CKC2D

lastna implementacija
korelacijskega algoritma

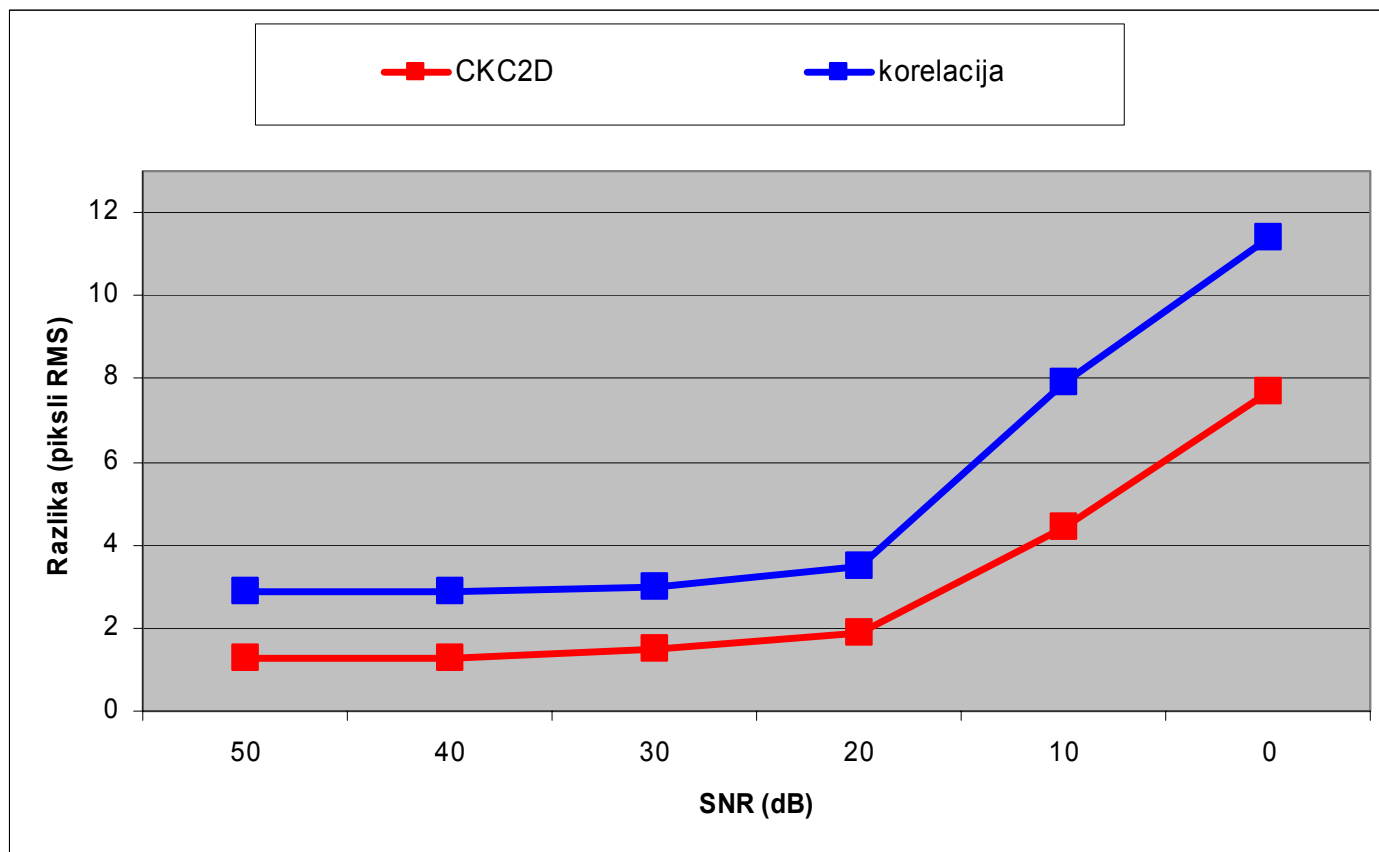
Določanje globine s CKC2D - vpliv šuma

- Piksli z veljavno disparitetno vrednostjo, $N = 9$



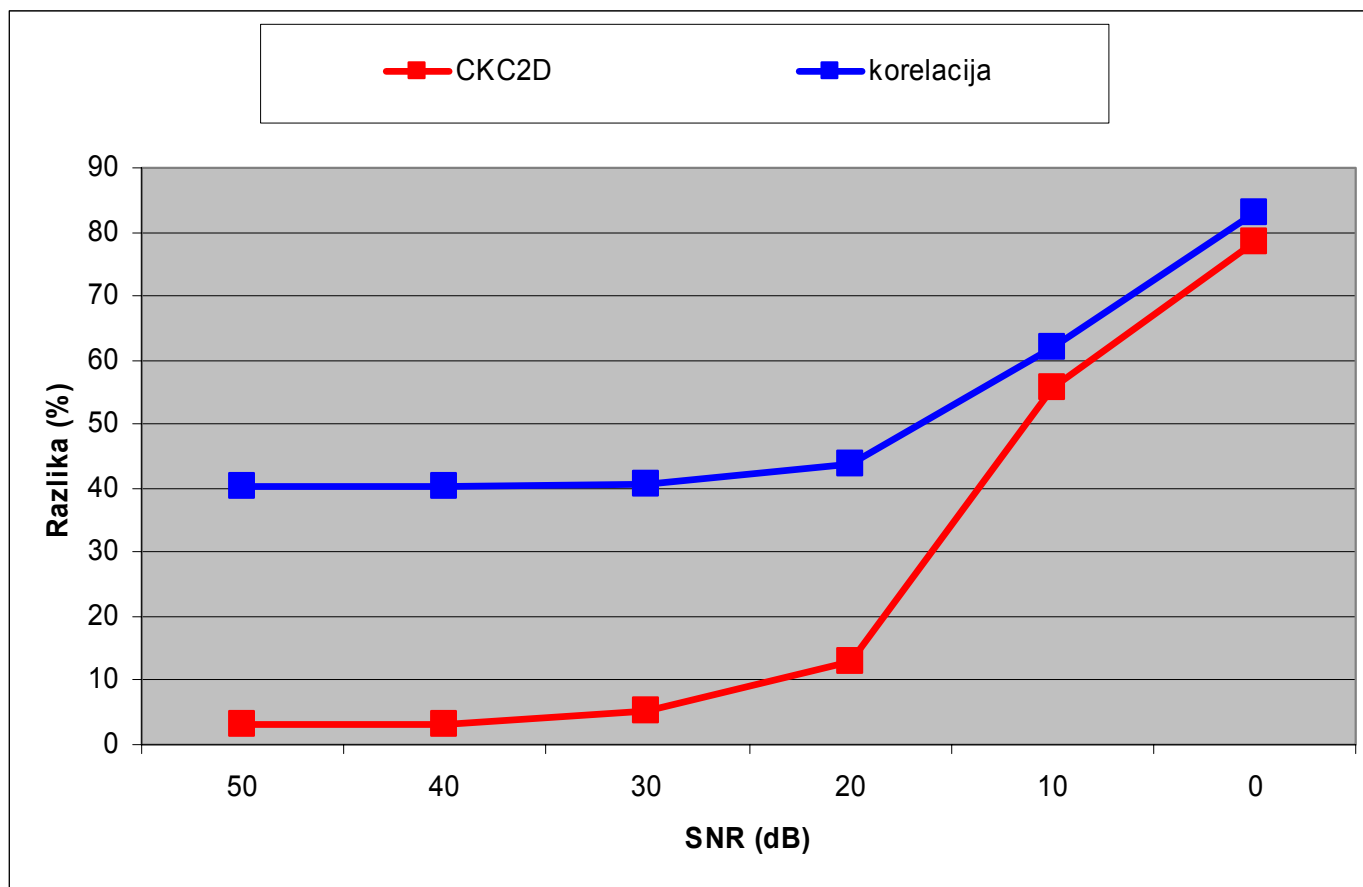
Določanje globine s CKC2D - vpliv šuma

- RMS-razlika glede na referenčne globinske slike, $N = 9$

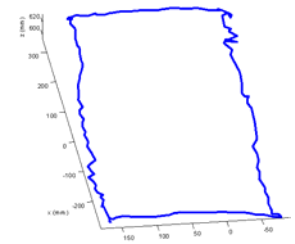
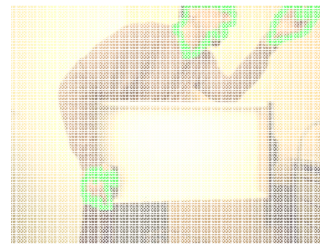
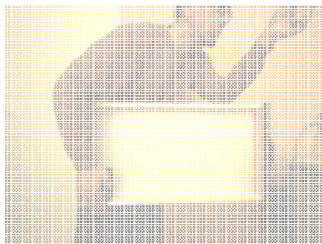
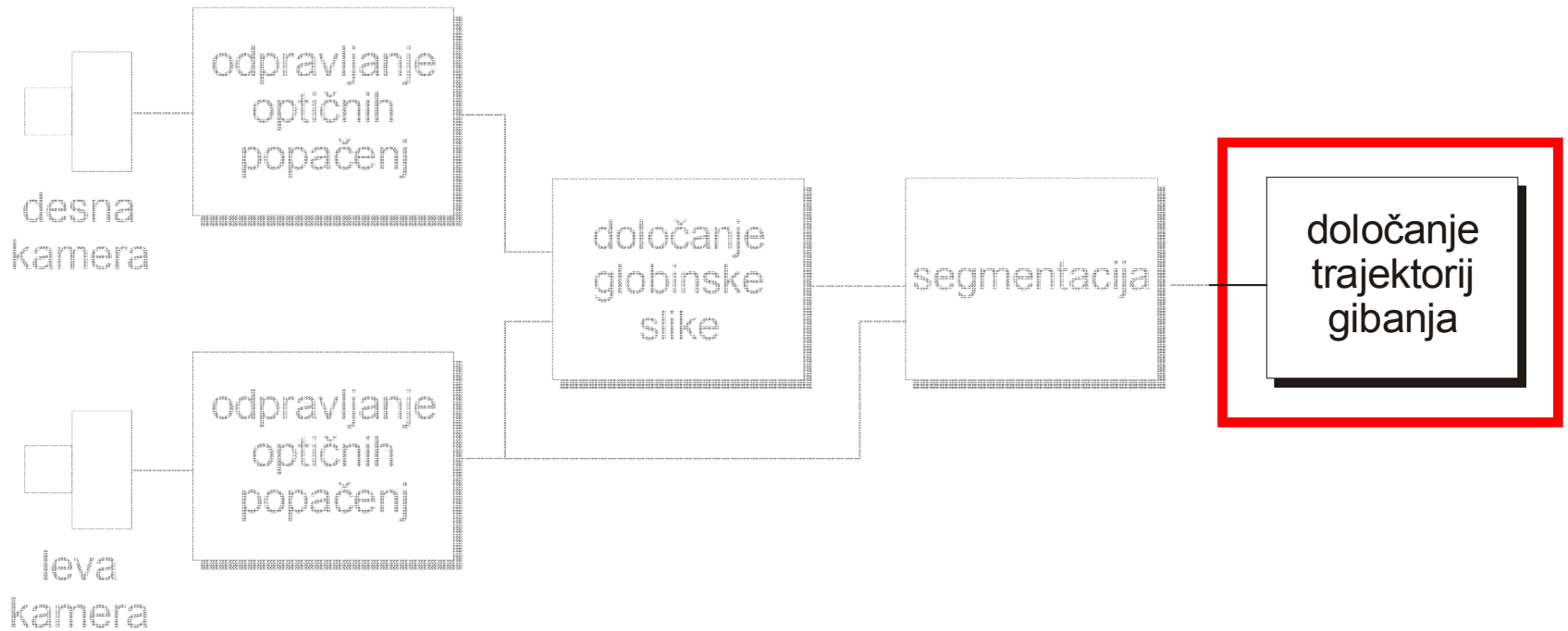


Določanje globine s CKC2D - vpliv šuma

- Piksli, ki se od ref. disparitet razlikujejo za > 1 , $N = 9$



Ocena natančnosti zaznanega gibanja



Ocenjevanje natančnosti video sledenja s primerjavo z magnetno sledilno napravo

- Gibanju objekta sledimo z obema napravama hkrati
- Poravnava koordinatnih sistemov (KS): $\mathbf{p}^M = \mathbf{A} \mathbf{p}^K$
- Pretvorbena matrika: odvisna od položaja naprav

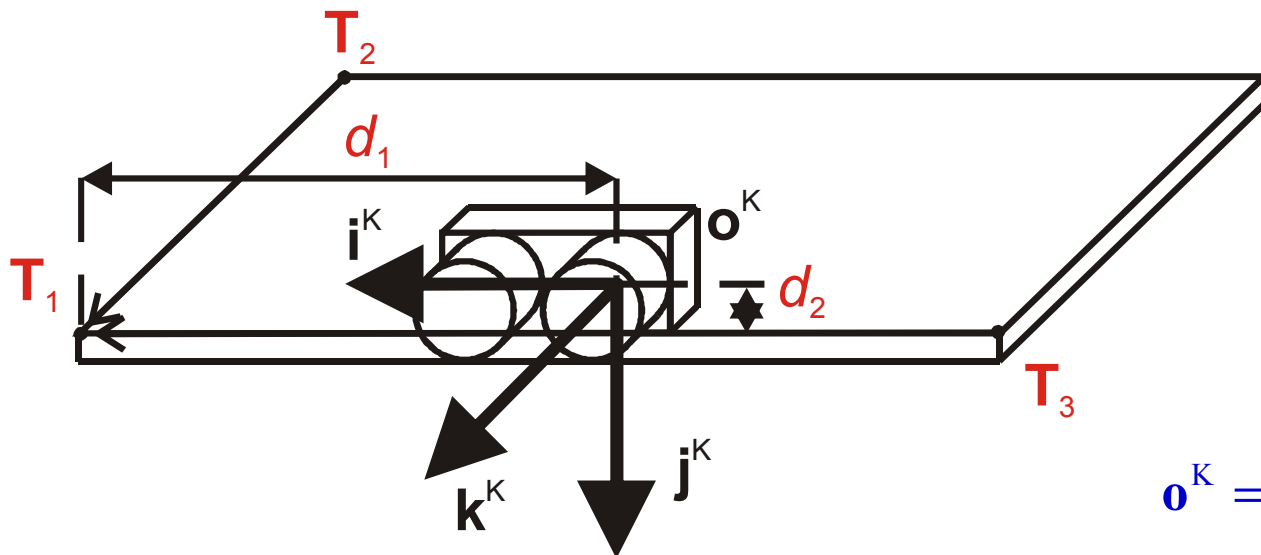
$$\mathbf{A} = \begin{bmatrix} i_1 & j_1 & k_1 & o_1 \\ i_2 & j_2 & k_2 & o_2 \\ i_3 & j_3 & k_3 & o_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{aligned} \mathbf{i}^K &= (i_1, i_2, i_3)^T, \\ \mathbf{j}^K &= (j_1, j_2, j_3)^T, \\ \mathbf{k}^K &= (k_1, k_2, k_3)^T, \\ \mathbf{o}^K &= (o_1, o_2, o_3)^T \end{aligned}$$

- Položaj kamere moramo izraziti z magnetnim sledilnikom
- Težava: bližina kovinske kamere zelo popači meritve
- Rešitev: položaj in orientacijo določimo posredno

Pretvorbeni model A

□ Parametri:

- koordinate točk \mathbf{T}_1 , \mathbf{T}_2 , \mathbf{T}_3
- razdalji d_1 in d_2



$$\mathbf{i}^K = \frac{\overline{\mathbf{T}_3 \mathbf{T}_1}}{\|\overline{\mathbf{T}_3 \mathbf{T}_1}\|}$$

$$\mathbf{k}^K = \frac{\overline{\mathbf{T}_2 \mathbf{T}_1}}{\|\overline{\mathbf{T}_2 \mathbf{T}_1}\|}$$

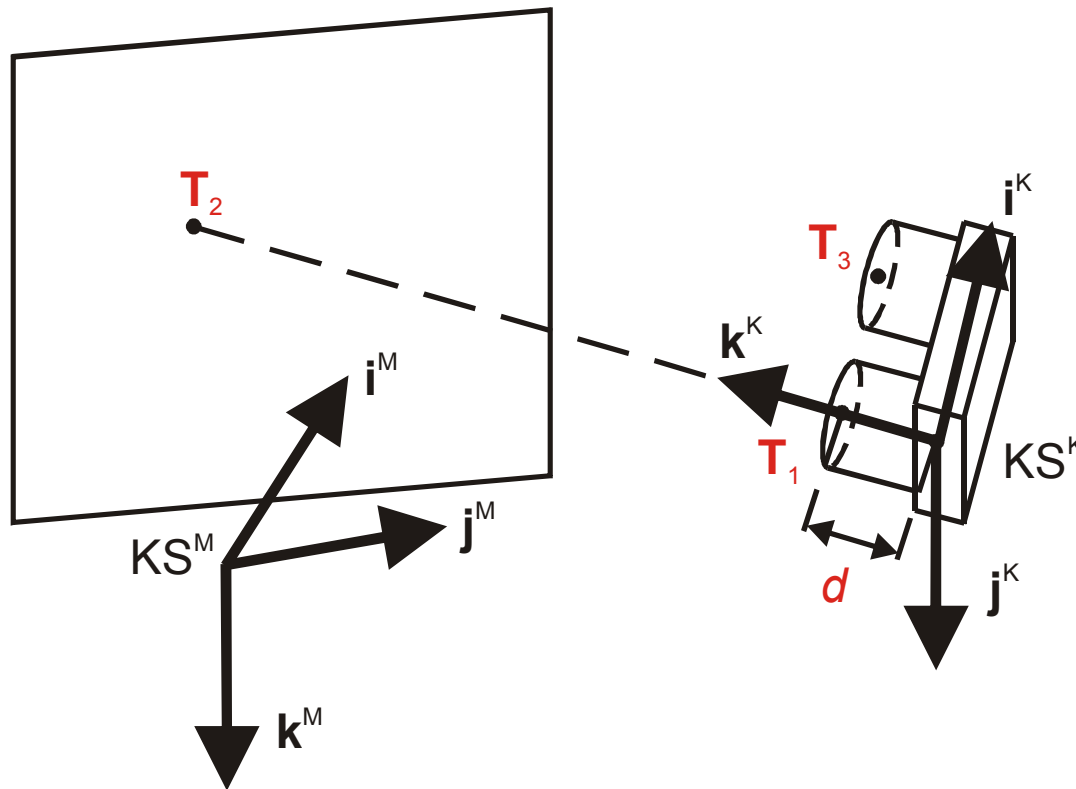
$$\mathbf{j}^K = \mathbf{k}^K \times \mathbf{i}^K$$

$$\mathbf{o}^K = \mathbf{T}_1 - d_1 \mathbf{i}^K - d_2 \mathbf{j}^K$$

Pretvorbeni model B

□ Parametri:

- koordinate točk \mathbf{T}_1 , \mathbf{T}_2 , \mathbf{T}_3
- razdalja d



$$\mathbf{i}^K = \frac{\overline{\mathbf{T}_1 \mathbf{T}_3}}{\|\overline{\mathbf{T}_1 \mathbf{T}_3}\|}$$

$$\mathbf{k}^K = \frac{\overline{\mathbf{T}_1 \mathbf{T}_2}}{\|\overline{\mathbf{T}_1 \mathbf{T}_2}\|}$$

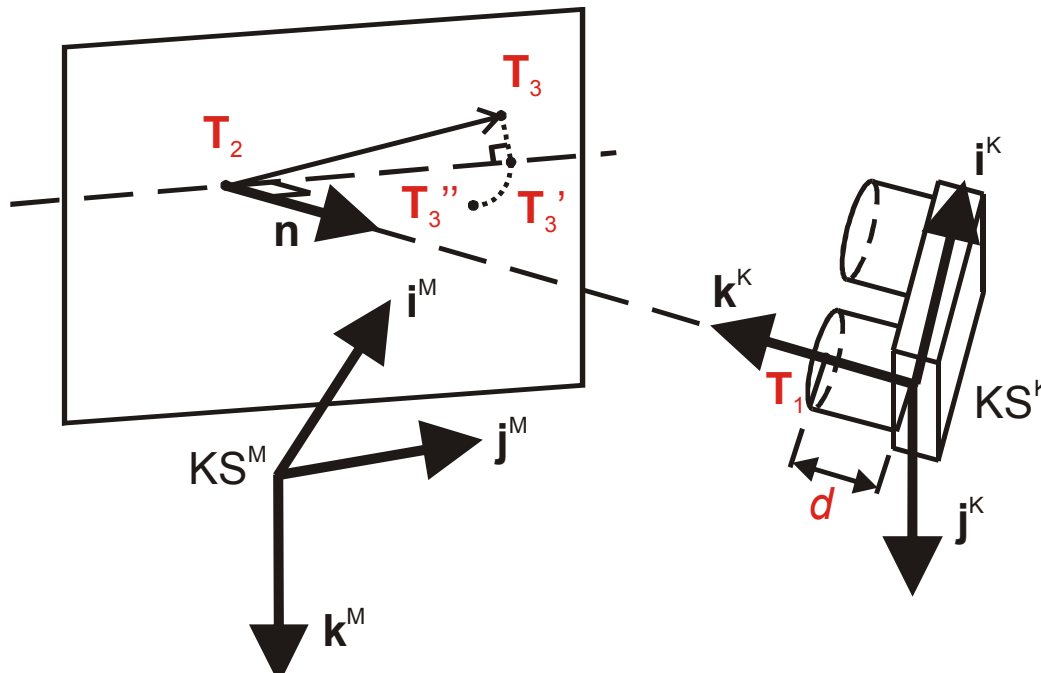
$$\mathbf{j}^K = \mathbf{k}^K \times \mathbf{i}^K$$

$$\mathbf{o}^K = \mathbf{T}_1 - d \mathbf{k}^K$$

Pretvorbeni model C

□ Parametri:

- koordinate točk T_1, T_2, T_3
- razdalja d
- odmik T_3'' od središča slike: m_H, m_V



$$\mathbf{i}^K = \frac{\overline{\mathbf{T}_2 \mathbf{T}_3}}{\|\overline{\mathbf{T}_2 \mathbf{T}_3}\|}$$

$$\mathbf{k}^K = \frac{\overline{\mathbf{T}_1 \mathbf{T}_2}}{\|\overline{\mathbf{T}_1 \mathbf{T}_2}\|}$$

$$\mathbf{j}^K = \mathbf{k}^K \times \mathbf{i}^K$$

$$\mathbf{o}^K = \mathbf{T}_1 - d \mathbf{k}^K$$

Vpliv parametrov pretvorbenega modela

- Občutljivost \mathbf{A} na parameter Θ_l : $\frac{\partial \mathbf{A}}{\partial \Theta_l}$
- Zgornja meja relativne občutljivosti celotnega modela:

$$S^{\text{MAX}} = \left(\left\| \frac{\partial \mathbf{A}}{\partial \Theta_1} \right\| + \left\| \frac{\partial \mathbf{A}}{\partial \Theta_2} \right\| + \dots + \left\| \frac{\partial \mathbf{A}}{\partial \Theta_N} \right\| \right) \cdot \|\mathbf{A}^{-1}\|$$

- Na končno oceno napake poleg občutljivosti vpliva tudi velikost napake parametra $\Delta \Theta_l$:

$$E^{\text{MAX}} = \left(\left\| \frac{\partial \mathbf{A}}{\partial \Theta_1} \right\| \Delta \Theta_1 + \left\| \frac{\partial \mathbf{A}}{\partial \Theta_2} \right\| \Delta \Theta_2 + \dots + \left\| \frac{\partial \mathbf{A}}{\partial \Theta_N} \right\| \Delta \Theta_N \right) \cdot \|\mathbf{A}^{-1}\|$$

➡ Uspešnost modelov je odvisna od dejanske izbire parametrov!

Delitev napake sledenja

□ Možnosti napak:

- skupna napaka sledenja E_{skupni} : ocenimo s primerjavo trajektorij
 - napake pretvorbe: $\Delta\mathbf{A}$
 - napake sledilnega algoritma: $\Delta\mathbf{p}^K$

□ Statistična delitev napake:

- simuliramo vse možne napake $\Delta\mathbf{A}$:
 - napake parametrov Θ_l so enakomerno porazdeljene
 - statistično ocenimo napako pretvorbe
- $\Delta\mathbf{A}$ in $\Delta\mathbf{p}^K$ sta neodvisni, zato velja:
$$\text{RMS } (E_{\text{skupni}})^2 = \text{RMS } (\Delta\mathbf{A})^2 + \text{RMS } (\Delta\mathbf{p}^K)^2$$

Delitev napake sledenja

□ Analitična delitev napake:

- pogoj: 4 točke z enako napako: $\Delta \mathbf{P}^K = \Delta \mathbf{p}^K$
- izračunamo rotacijski del idealne matrike \mathbf{A} :

$$\mathbf{P}^M \left(\mathbf{P}_{\text{izmerjeni}}^K \right)^{-1} = \mathbf{A} \left(\Delta \mathbf{P}^K \right)^{-1} = \begin{bmatrix} i_1 & j_1 & k_1 & -i_1 dp_1 - j_1 dp_2 - k_1 dp_3 + o_1 \\ i_2 & j_2 & k_2 & -i_2 dp_1 - j_2 dp_2 - k_2 dp_3 + o_2 \\ i_3 & j_3 & k_3 & -i_3 dp_1 - j_3 dp_2 - k_3 dp_3 + o_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

\mathbf{A}_{ROT}

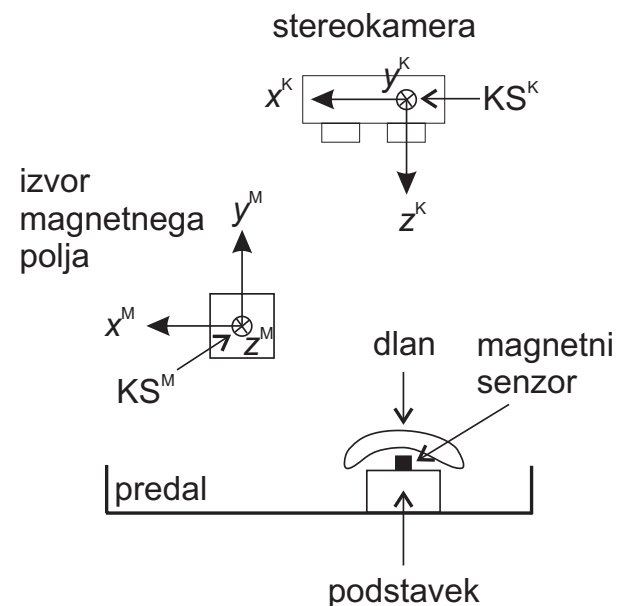
- z \mathbf{A}_{ROT} popravimo $\mathbf{A}_{\text{izmerjeni}} \rightarrow \hat{\mathbf{A}}$, določimo napaki:

$$\square \Delta \mathbf{A} = \left(\hat{\mathbf{A}} \right)^{-1} \cdot \mathbf{A}_{\text{izmerjeni}}$$

$$\square \Delta \mathbf{P}^K = \mathbf{P}_{\text{izmerjeni}}^K \cdot \left(\mathbf{P}^M \right)^{-1} \cdot \mathbf{A}_{\text{izmerjeni}} \cdot \left(\Delta \mathbf{A} \right)^{-1}$$

Rezultati eksperimentov na realnih posnetkih

- Sledilni algoritem:
 - segmentacija: barva kože + globina
 - predikcijsko-korekcijski algoritem
 - implementacija v Matlabu
- Testna postavititev:
 - okolje brez kovinskih predmetov
 - fizično omejeni gibi
 - stereokamera **Videre Design MEGA-D**:
 - 3 posnetki s po 120 slikami
 - ločljivost 320×240 pikslov
 - 1 meter oddaljena od testne osebe
 - sledilnik **Polhemus 3Space Fastrak**:
 - napaka $< 0,8$ mm RMS



Rezultati: primerjava pretvorbenih modelov

- Ocene maksimalnih odstopanj parametrov:
 - točke: $\Delta \mathbf{T}_1 = \Delta \mathbf{T}_2 = \Delta \mathbf{T}_3 = 1,6$ mm
 - razdalje: $\Delta d_1 = \Delta d_2 = \Delta d = 1,0$ mm
 - odmiki: $\Delta m_H = \Delta m_V = 1$ piksel
 - poravnava \mathbf{T}_2 v središče slike: 1 piksel
- Primerjava modelov:

	Ocena S^{MAX}	Ocena E^{MAX}
Model A	2027,0	2764,7
Model B	1668,0	2772,4
Model C	1654,2	2788,4

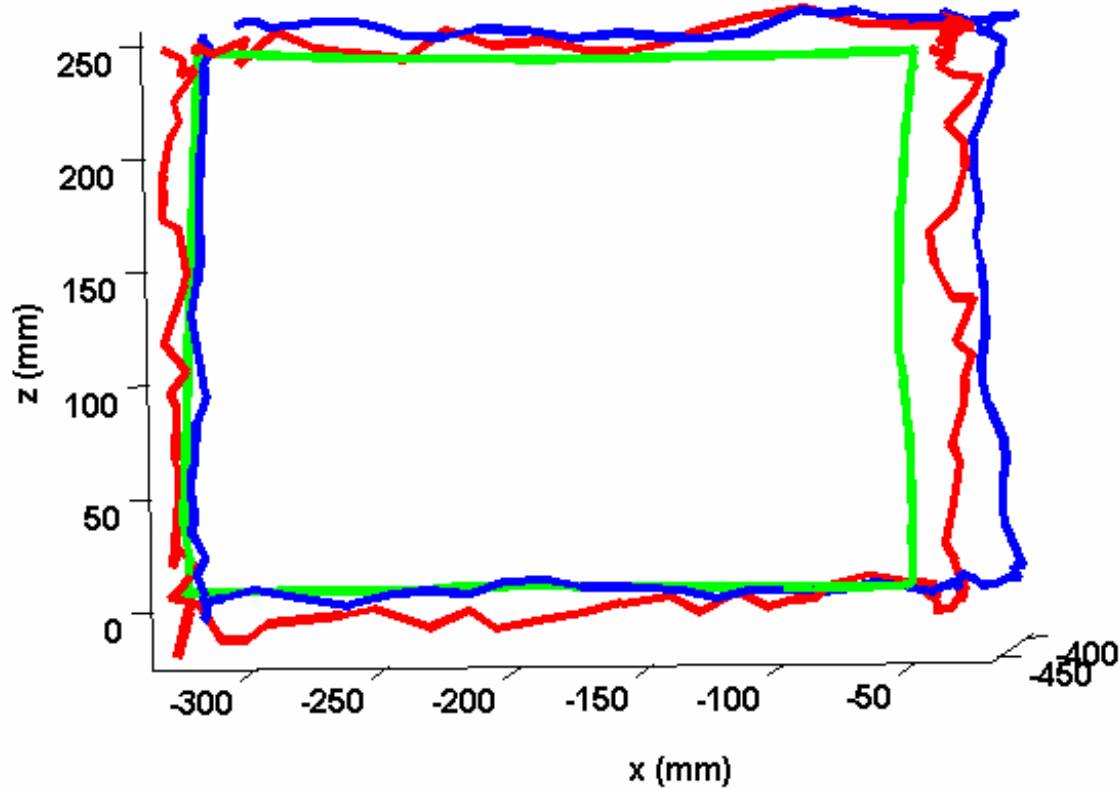
izbran
model

Rezultati: zaznani objekti



Rezultati: trajektorije gibanja

- Trajektorija desne dlani: zelena – referenca, rdeča – algoritmom CKC2D, modra – korelacijski algoritem



Rezultati: primerjava trajektorij

- Povprečna razlika med trajektorijami (mm RMS):

	x	y	z	Skupaj
CKC2D	$11,6 \pm 1,6$	$10,1 \pm 0,4$	$13,6 \pm 3,4$	$20,7 \pm 1,4$
SVS	$16,4 \pm 6,0$	$5,8 \pm 1,3$	$10,9 \pm 1,7$	$20,8 \pm 4,9$

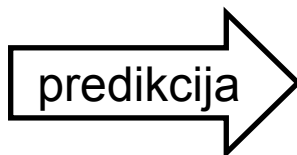
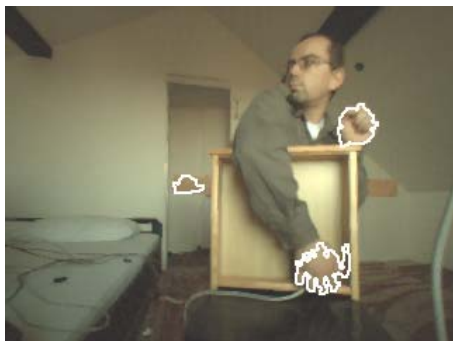
- Statistična ločitev napake sledenja od napake pretvorbe:

	Napaka pretvorbe ΔA (mm RMS)	Napaka sledenja Δp^K (mm RMS)
CKC2D	3,7	20,4
SVS	3,7	20,5

Rezultati: vpliv predikcijsko-korekcijskega algoritma

- prednost: popravljanje segmentacijskih napak
- slabost: večja odstopanja

	Napaka brez predikcije (mm RMS)	Napaka s predikcijo (mm RMS)
CKC2D	20,4	20,7
SVS	19,7	20,8



Rezultati: primerjava s podobnimi sledilnimi sistemi

- Moreno et al.: sledenje glavi s 4 cm napake pri 1 m oddaljenosti
- Jojic et al.: sledenje dlanem s 15 cm napake pri 3 m oddaljenosti



- Naš algoritem: 2 cm napake pri oddaljenosti 1 m od kamere

Znanstveni prispevki

- Nov algoritem za določanje globinskih slik z metodo CKC2D
- Postopek za ocenjevanje natančnosti video sledenja s pomočjo magnetnega sledilnika
- Analiza napak pretvorbe koordinatnih sistemov s tremi različnimi modeli
- Izpeljava analitične in statistične ocene natančnosti sledenja
- Algoritem za sledenje obrazu in rokam na osnovi stereo posnetkov, brez uporabe dodatne opreme
- Analiza učinkovitosti razvitih algoritmov na množici umetnih in realnih posnetkov

Možnosti uporabe

- Enostavna prilagoditev za sledenje poljubnim objektom
- Sistem je dovolj natančen za:
 - navigacijo v VR okoljih
 - krmiljenje in manipulacijo VR objektov
 - nadomestek za 3D miško
 - hibridne sisteme
- Omejitve:
 - aplikacije, kjer je hitrost in natančnost sledenja kritična (medicinske aplikacije)

Teza in hipoteze

- Določanje globinske slike s slepo ločitvijo izvorov je bolj robustno kot klasični postopek z iskanjem korelacije. ✓
- Nadgradnja sledilnega postopka s predikcijsko-korekcijsko shemo bistveno poveča robustnost in ~~natančnost~~ sledenja. ✓
- Preprosta stereokamera v kombinaciji z izboljšanim sledilnim algoritmom je primerljiva z obstoječimi rešitvami za navigacijo. ✓
- Razvijemo lahko statistični postopek za analizo natančnosti video sledenja. ✓

Zaključek: izhodiščno tezo lahko v celoti potrdimo

Zahvala

- Ministrstvu za šolstvo, znanost in šport, ki mi je s pogodbo št. 3311-00-450001/2 omogočilo status asistenta-stažista

- Avtorjem referenčnih globinskih slik:
 - [1] D. Scharstein, R. Szeliski, *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms*, International Journal of Computer Vision, št. 47, str. 7-42, 2002.