

Normálvektor becslés affin transzformációkkal

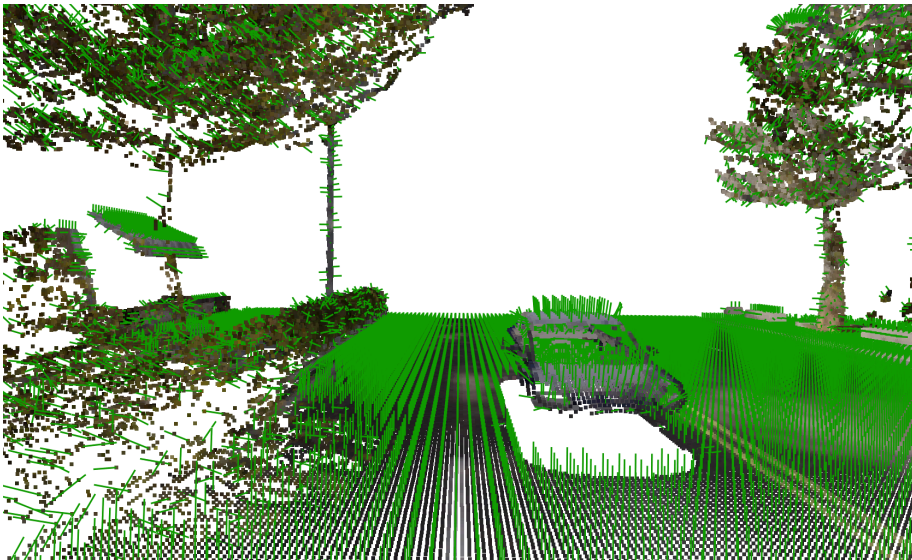
Karikó Csongor

Bolyai Kollégium

2024.03.06

Bevezető

Normálvektorok → Irányított pontfelhő



Felhasználások

- Felületrekonstrukció
- Lokalizáció
- Szegmentálás
- Pontfelhő finomítás

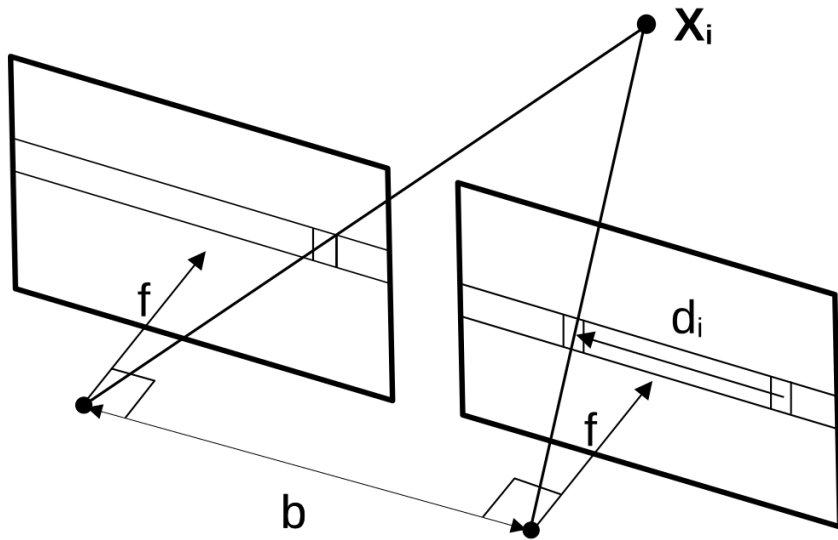
Bevezető

Lehetséges inputok

- Pontfelhő (például Lidar)
- Monó kamera kép (MI)
- **Sztereó kamera kép**

Bevezető

Sztereó kamera kép



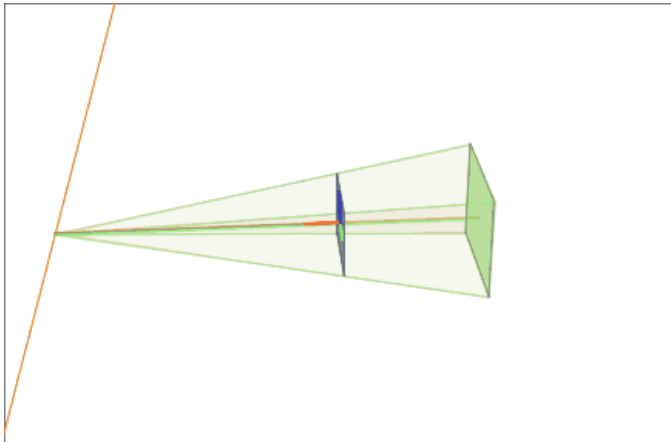
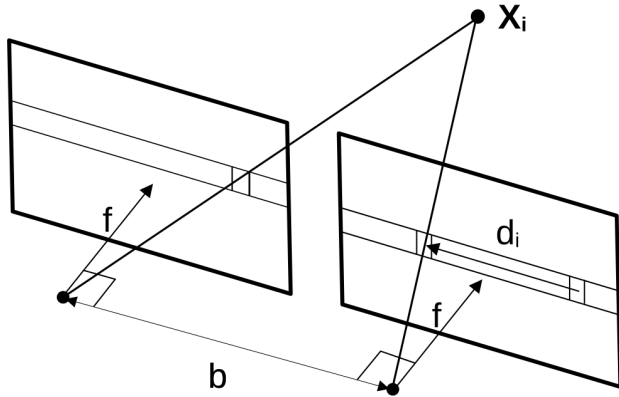
Standard sztereó kamera állás



Bevezető

Standard sztereó kamera állás

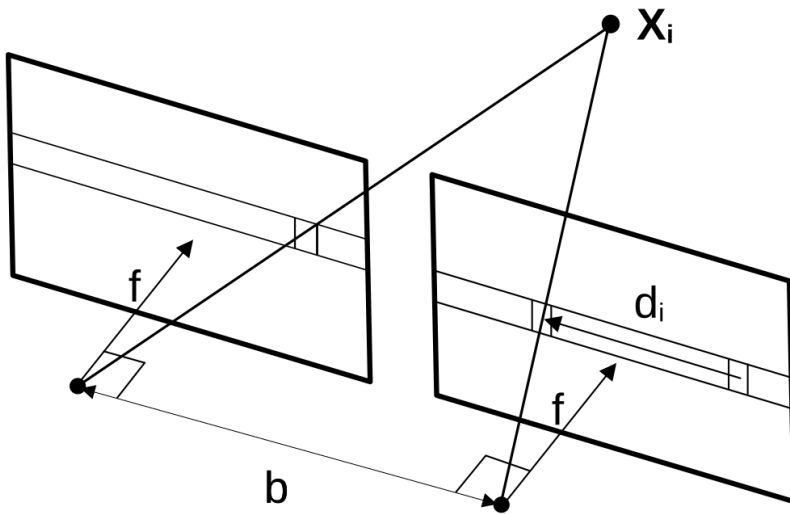
- Párhuzamos szenzor síkok
- Megfelelő képpontok azonos sorban
- Sokszor megkapható rektifikálással általánosabb állásból
- Speciális affin transzformációs mátrix
 - Kevesebb számítás



Rektifikálás

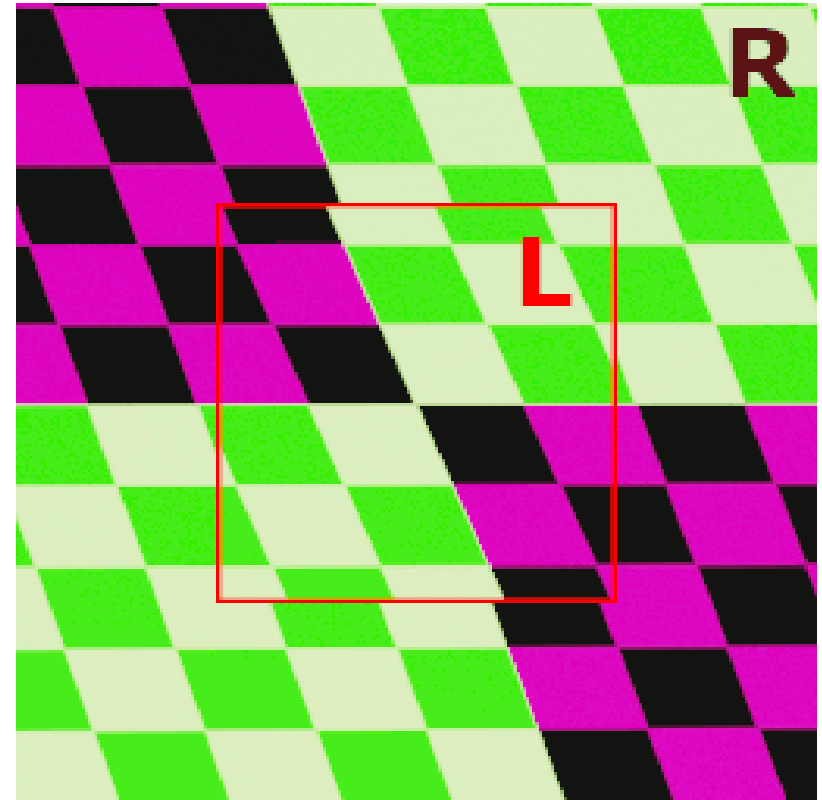
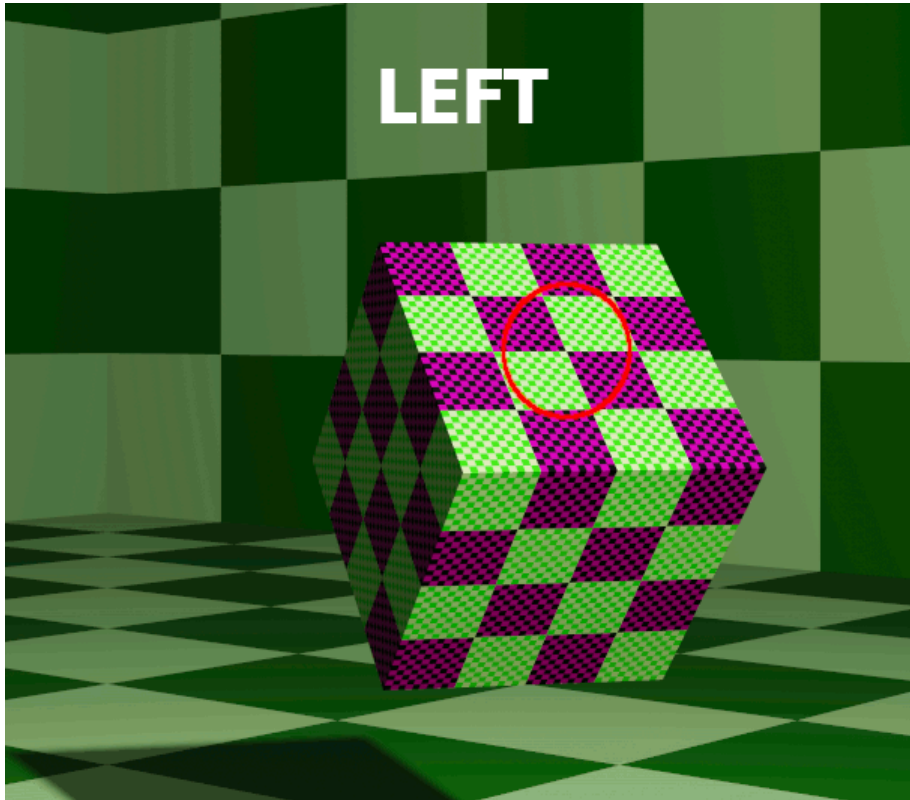
Bevezető

Kalibrációs paraméterek



- Bázis távolság (b)
- Fókusz­táv (f)
- Pixelsűrűség

Elméleti háttér



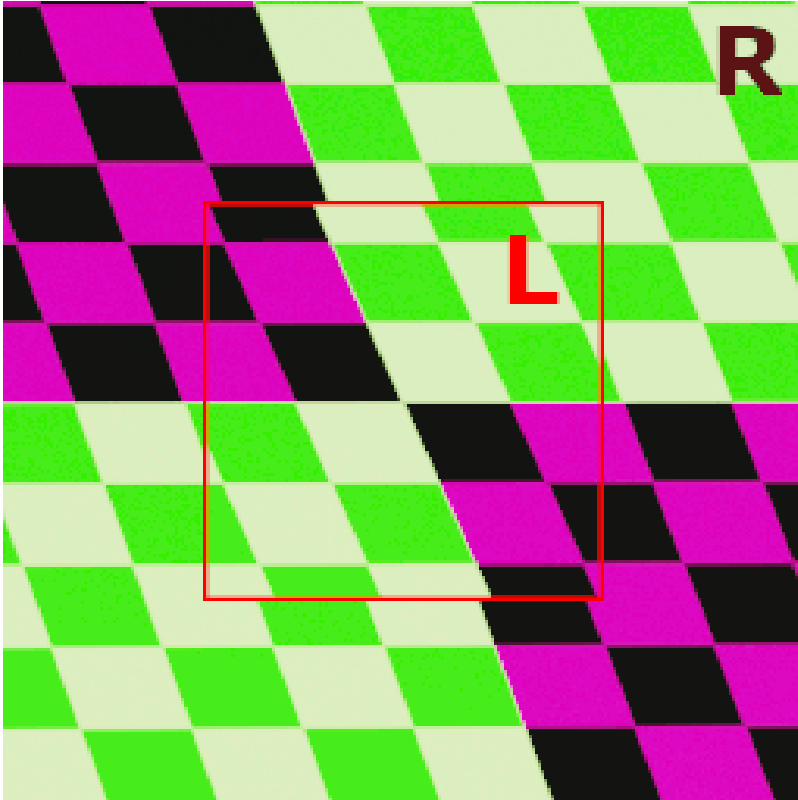
Affin transzformáció megfelelő pontpár környezete között

Elméleti háttér

Összefüggés normálvektorral

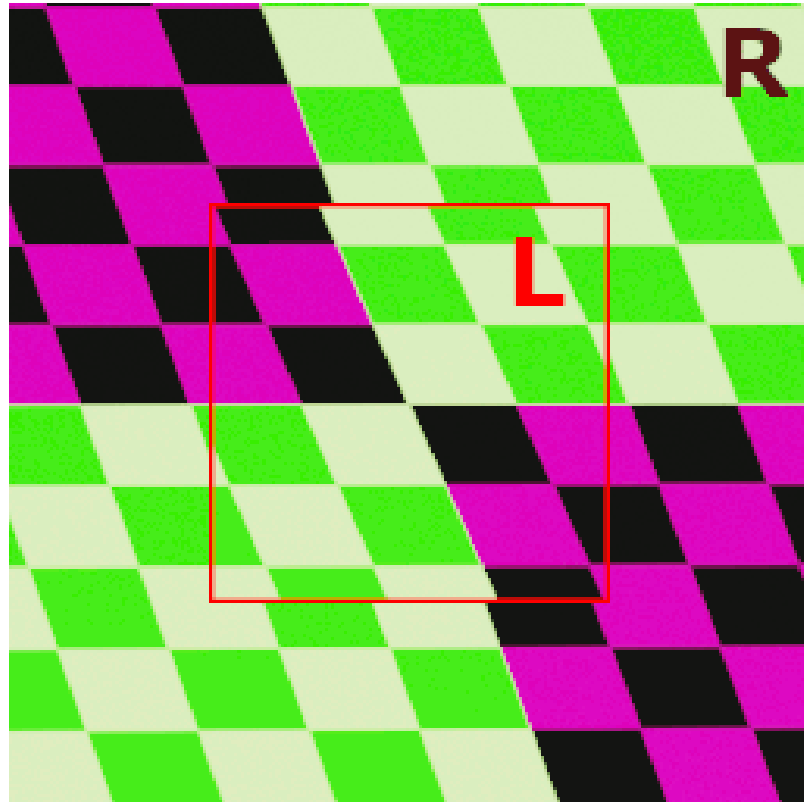
$$A = \begin{bmatrix} a_1 & a_2 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} \frac{n_x(x-b)+n_y y+n_z z}{n_x x+n_y y+n_z z} & \frac{-b f_x n_y}{f_y(n_x x+n_y y+n_z z)} \\ 0 & 1 \end{bmatrix}$$



*Affin transzformáció megfelelő pontpár
környezete között*

Elméleti háttér

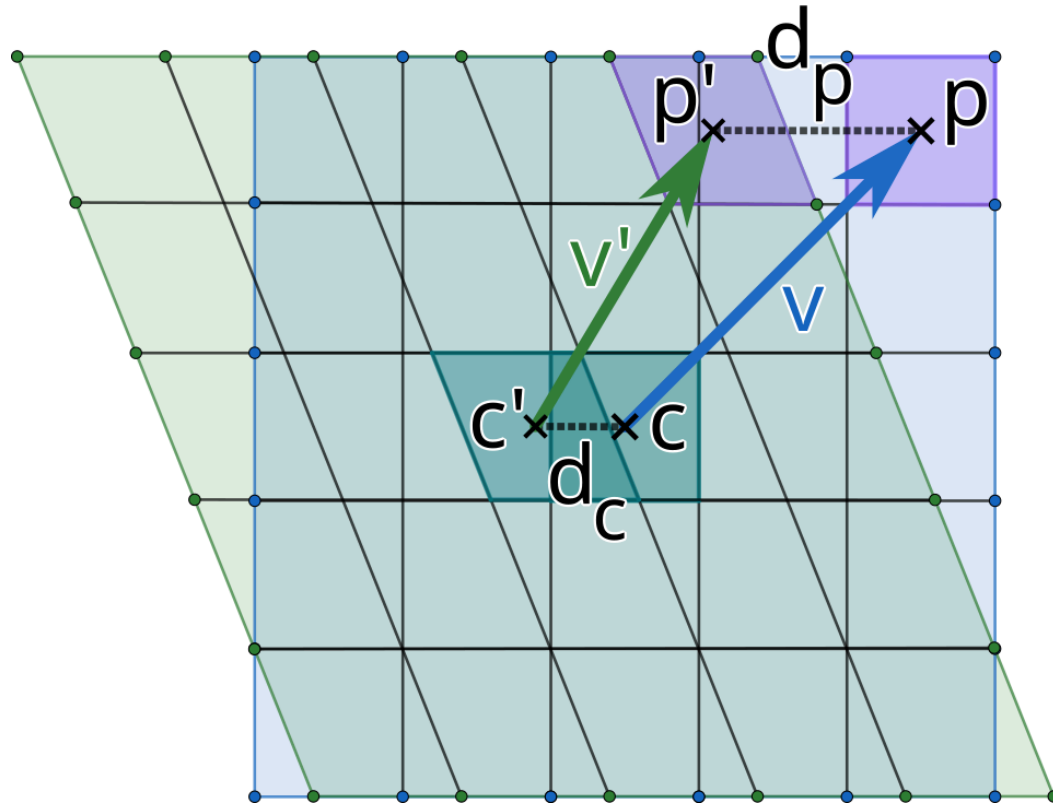


*Affin transzformáció megfelelő pontpár
környezete között*

Összefüggés normálvektorral

$$n = \begin{bmatrix} k_h z (a_1 - 1) \\ a_2 k_v z \\ -a_1 k_h x - a_2 k_v y + b k_h + k_h x \end{bmatrix}$$

Affin paraméter becslés



Egymásnak megfelelő pontpár és környezete a két képen.

$$v + [d_p, 0]^T = [d_c, 0]^T + Av$$

Affin paraméter becslés

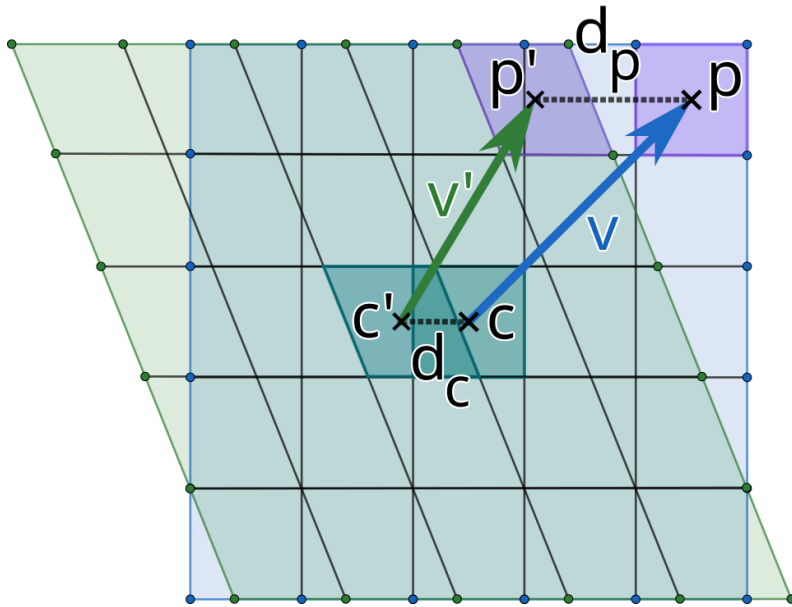
Túlhatározott egyenlet rendszer

$$\underbrace{\begin{bmatrix} v_{1x} & v_{1y} \\ v_{2x} & v_{2y} \\ \vdots & \vdots \\ v_{Nx} & v_{Ny} \end{bmatrix}}_{\mathbf{V}} \underbrace{\begin{bmatrix} a_1 - 1 \\ a_2 \end{bmatrix}}_{\mathbf{a}} = \underbrace{\begin{bmatrix} d_1 - d_c \\ d_2 - d_c \\ \vdots \\ d_N - d_c \end{bmatrix}}_{\mathbf{d}}$$

Ahol $[v_{ix}, v_{iy}]^T$ egy figyelembe vett pont relatív koordinátája, d_i a diszparitása, d_c pedig a középső diszparitás.

Legkisebb négyzetes hibájú megoldás:

$$\mathbf{a} = \underbrace{\left(\mathbf{V}^T \mathbf{V} \right)^{-1} \mathbf{V}^T}_{\mathbf{S}} \mathbf{d}$$



Egymásnak megfelelő pontpár és környezete a két képen.

Affin paraméter becslés

$\mathbf{S} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T$ előre kiszámítható, a képletet kibontva:

$$a_1 - 1 = \sum_{i=1}^N \frac{1}{\underbrace{\alpha\gamma - \beta^2}_{s_{1i}}} (\underbrace{\gamma v_{ix} - \beta v_{iy}}_{s_{1i}}) \underbrace{(d_i - d_c)}_{d_i} \quad (1)$$

$$a_2 = \sum_{i=1}^N \frac{1}{\underbrace{\alpha\gamma - \beta^2}_{s_{2i}}} (\underbrace{-\beta v_{ix} + \alpha v_{iy}}_{s_{2i}}) \underbrace{(d_i - d_c)}_{d_i}. \quad (2)$$

Ahol $\alpha = \sum_{i=1}^N v_{ix}^2$, $\beta = \sum_{i=1}^N v_{ix} v_{iy}$, $\gamma = \sum_{i=1}^N v_{iy}^2$.

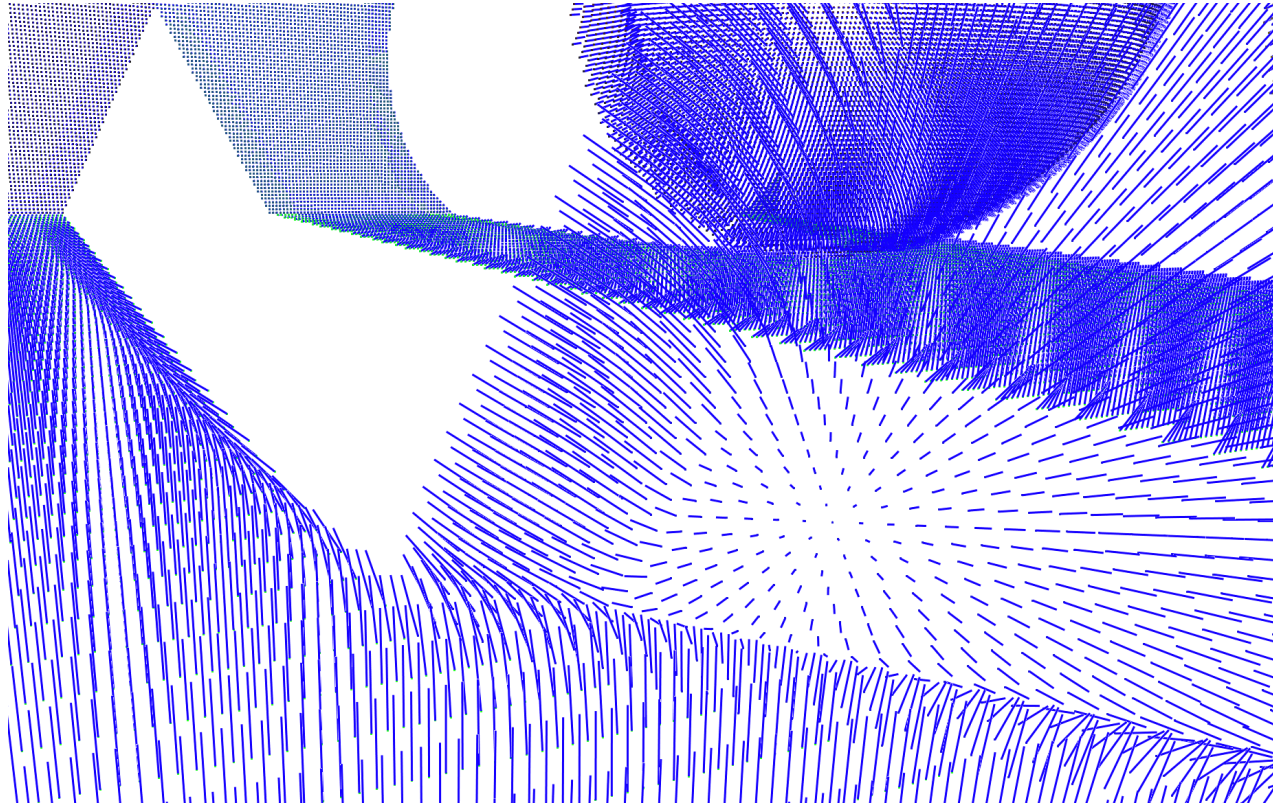
Affin paraméter becslés

$$a_1 = \sum_{i=1}^N s_{1i}(d_i - d_c) = \underbrace{\sum_{i=1}^N s_{1i}d_i}_{1. \text{ Lépés}} - d_c \underbrace{\sum_{i=1}^N s_{1i}}_{\delta_1} + 1$$

$$a_2 = \sum_{i=1}^N s_{2i}(d_i - d_c) = \underbrace{\sum_{i=1}^N s_{2i}d_i}_{1. \text{ Lépés}} - d_c \underbrace{\sum_{i=1}^N s_{2i}}_{\delta_2}$$

Így az első lépés egy 2D konvolúció, a második lépés pedig csak néhány művelet.

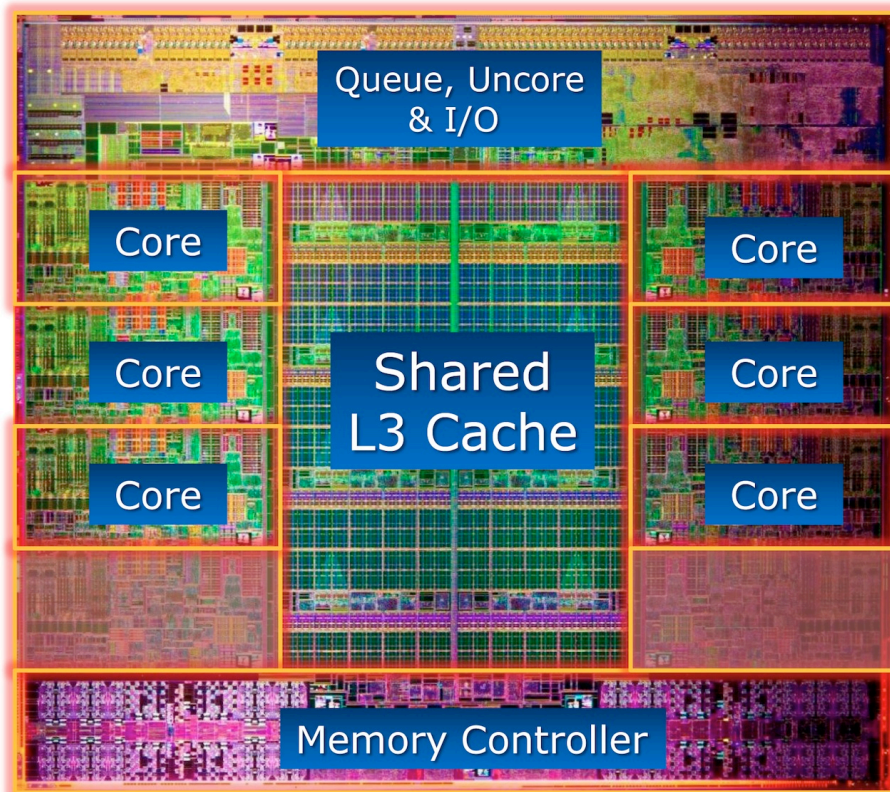
Egy fix méretű és alakú kernel esetén a kapott normálvektorok elmosódnak a felületek találkozásakor:



A figyelembe vett szomszédokat adaptívan kell megválasztani!

Hatékonyági kérdések

CPU



- kevés nagy, gyors mag
- szekvenciális utasításokban hatékony

GPU

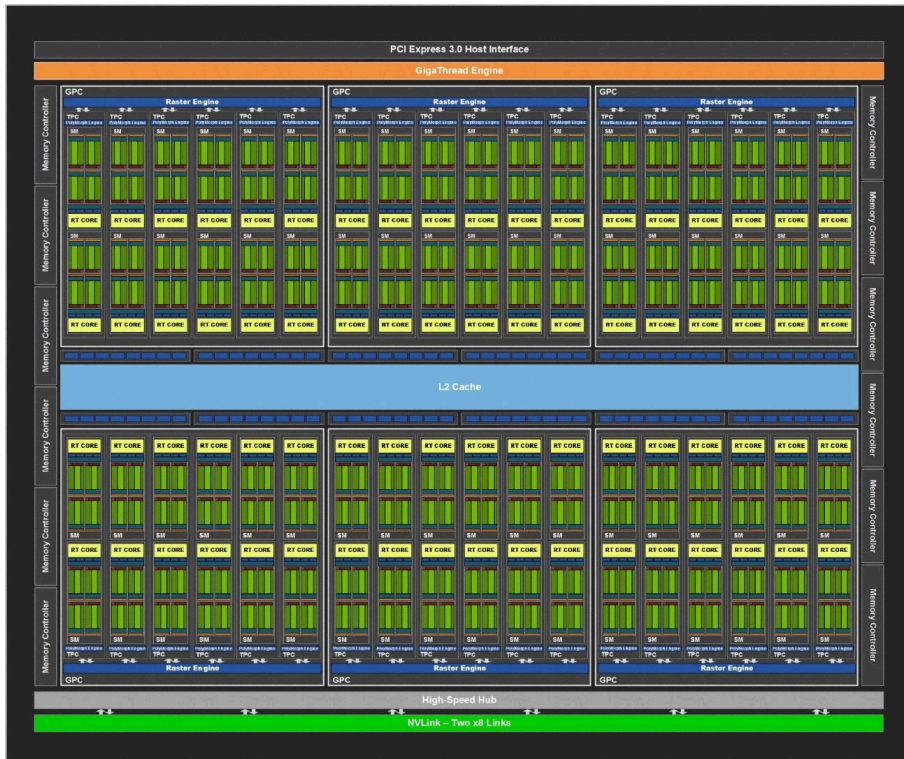


- sok kicsi, egyenként lassabb mag
- párhuzamos utasításokban hatékony

Hatékonyági kérdések

GPU

Single Instruction Multiple Data



- Divergens szálak be kell egymást várják (if mindkét ága mindig végrehajtódik)
- Lokális memória korlátolt, ha sokat használunk egyszerre kevesebb szál futhat
- Létezik szinkronizáció (barrier), de költséges

Cél: minél jobban megőrizni a futásidőt

Adaptív régió kiválasztás

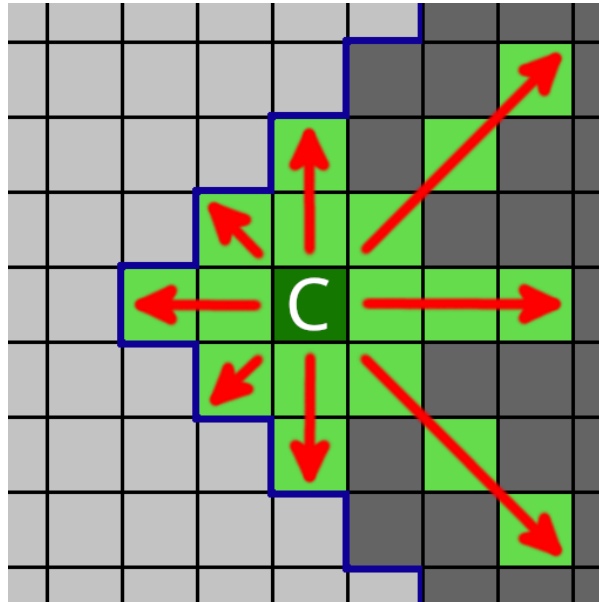
Ötlet:

1. Éldetektálás a mélységtérképen
2. Flood-fill a régiók megkereséséhez
3. Régióként affin paraméter becslés
4. Paramétereiből normálvektor

Problémák:

- Éldetektálás zaj mellett nehéz
- Flood-fill nem GPU hatékony
 - Mely pixelekből induljon?
 - Mindegyikből?
 - Memória?
 - Divergens szálak?

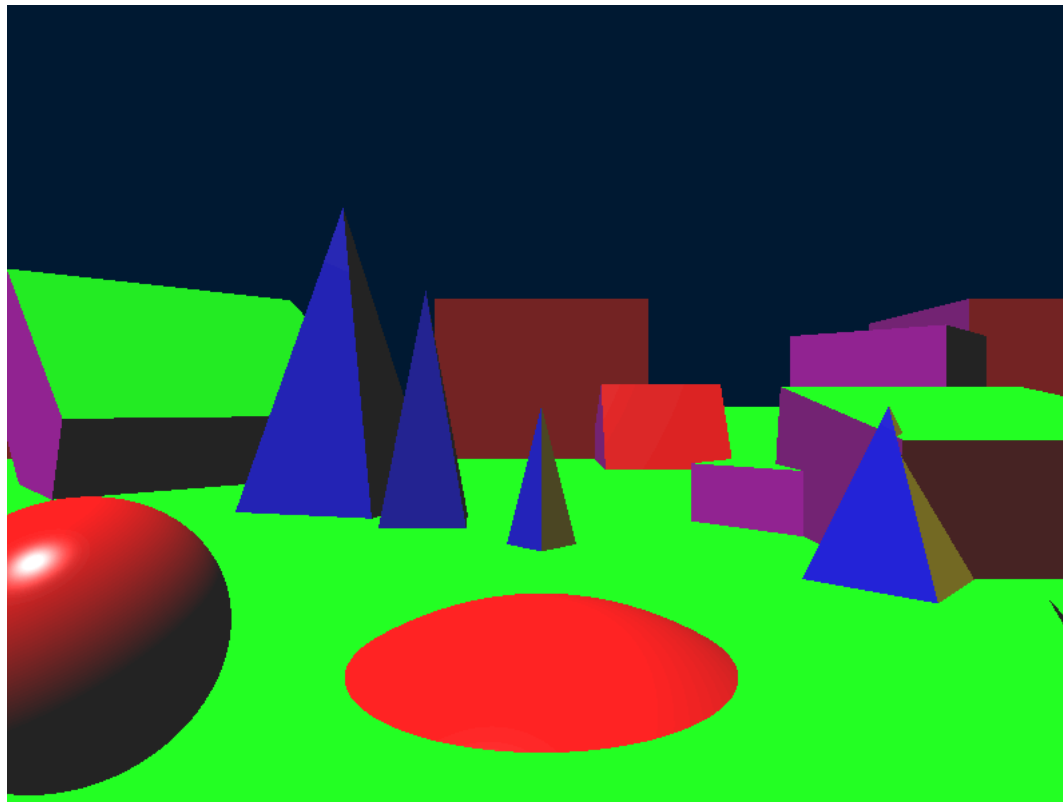
Adaptív régió kiválasztás



- Sugarak mentén bejárás kifelé
- Megállási feltételek:
 - Max lépésszám
 - "Éldetektáló" feltétel

Adaptív régió kiválasztás

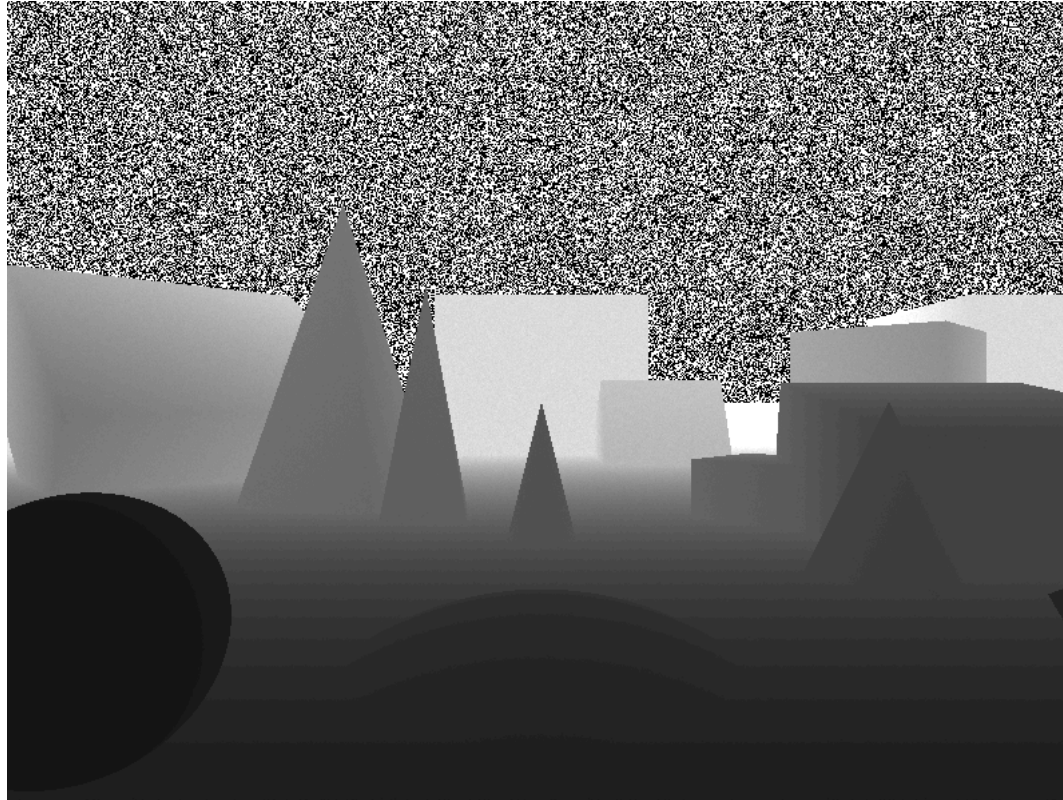
Éldetektálás a mélységtérképen



Példa színtér

Adaptív régió kiválasztás

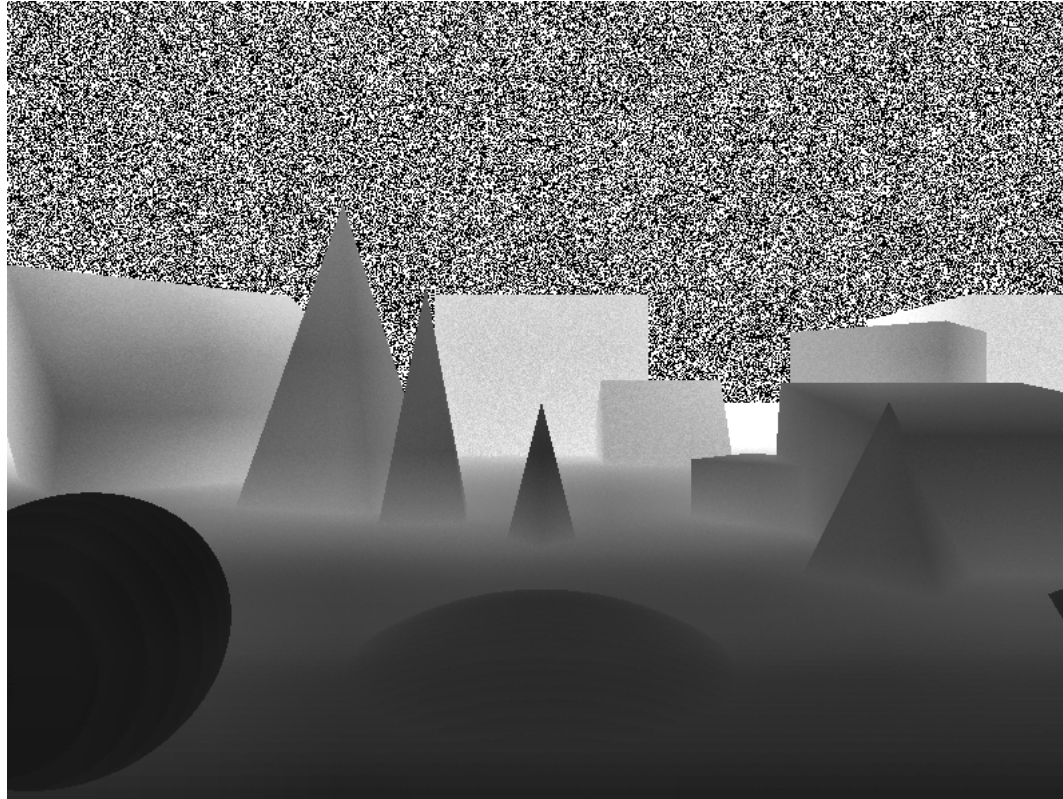
Éldetektálás a mélységtérképen



Mélység $\mathcal{N}(0, 0.2)$ px zajjal

Adaptív régió kiválasztás

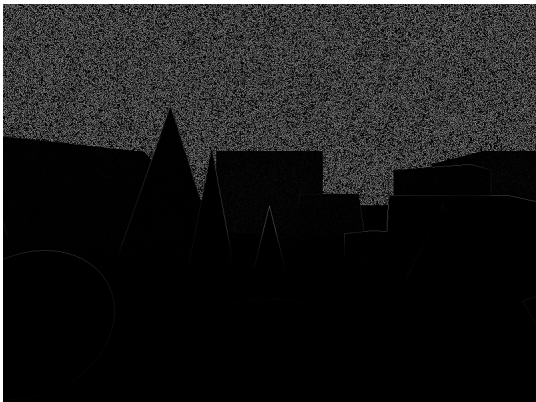
Éldetektálás a mélységtérképen



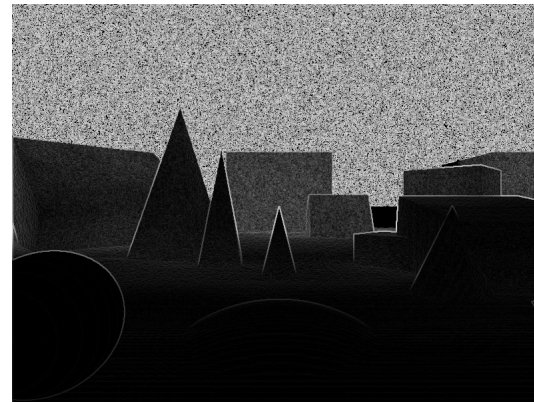
Lokális hisztogram kiegyenlítéssel

Adaptív régió kiválasztás

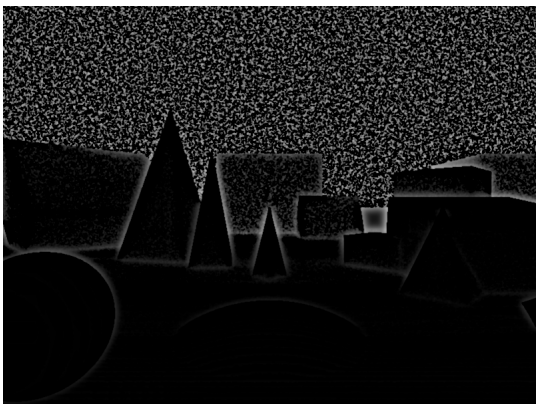
Éldetektálás a mélységtérképen



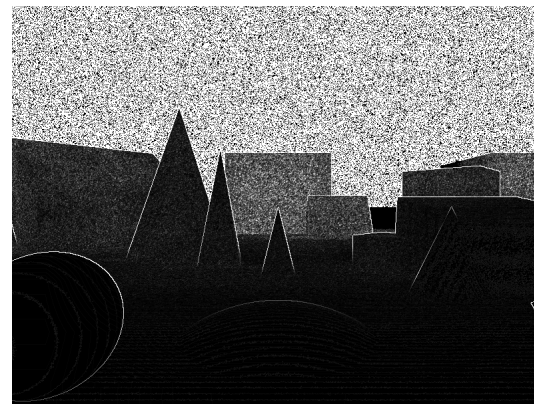
Laplace



Sobel



DoG



Gradient

Adaptív régió kiválasztás

Alternatív éldetektáló megállási feltételek

- Szög szerint:
 - A bejárt utolsó három pixel térdbeli koordinátája által meghatározott szög küszöbölve
- Kumulatív "hiba":
 - Egy éldetektáló szűrést követően a bejárt pixelek értékeinek összege küszöbölve
- **Lefedett mélység:**
 - A bejárás során a mélységek által lefedett tartomány mekkora a kiindulási pixel mélységéhez képest

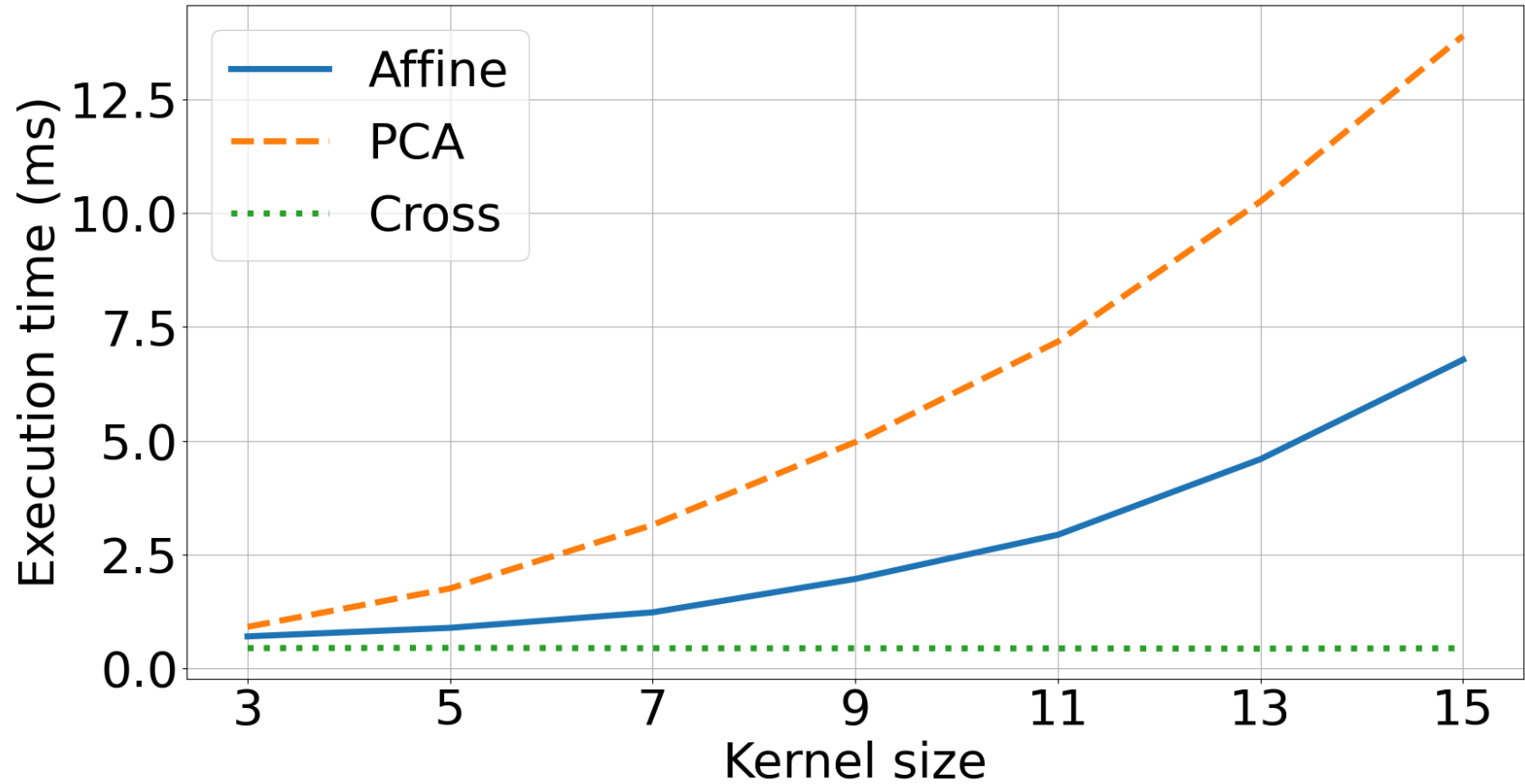
Mérések

Mérések

Összehasonlítás alapja:

- Naív keresztzorzat szomszédokból
- Főkomponens-analízis alapú normál becslés (PCA)
- (SDA-SNE)

Futásidő



1920x1080, GTX 1660 Ti

SDA-SNE: max 5 iteráció 0.6-0.7ms

Affin 9x9: 0.2-0.25ms

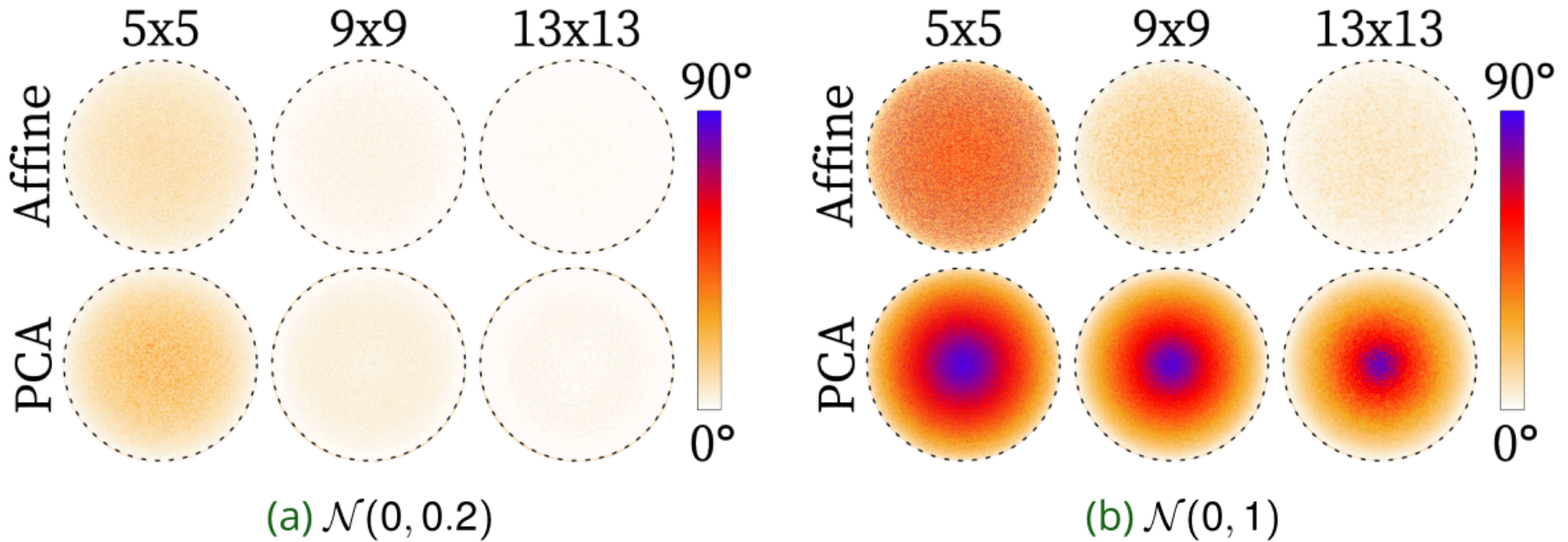
(RTX 4060)

Futásidő

Módszer	Átlag	Min	Max	Med	Std
Affine 9x9	0.65	0.62	0.72	0.65	0.01
PCA 9x9	1.62	1.55	1.73	1.62	0.03
ST[s:10 d:8]	1.80	1.73	2.07	1.79	0.03
ST[s:30 d:16]	8.15	7.73	8.89	8.14	0.21
CD[s:10 d:8]	1.71	1.66	1.86	1.71	0.03
CD[s:30 d:16]	6.85	6.47	8.01	6.84	0.21

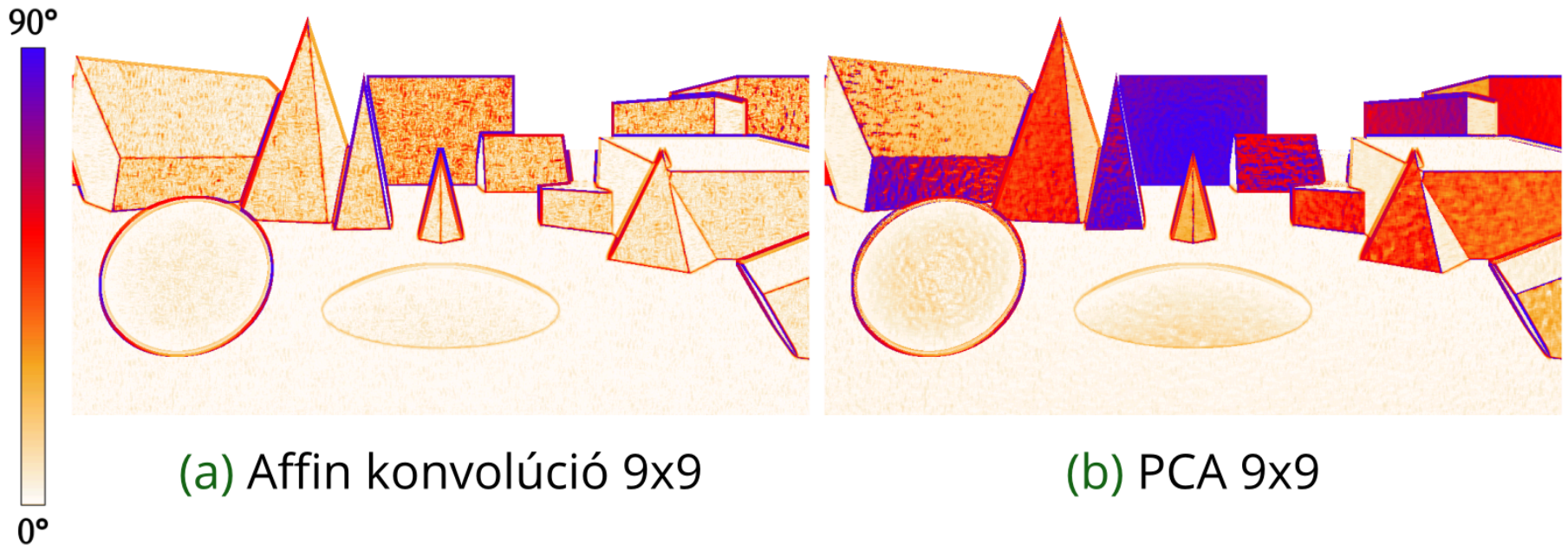
Futásidő statisztika 1000 képkockán (ms), felbontás: 1024x720

Zajszűrés



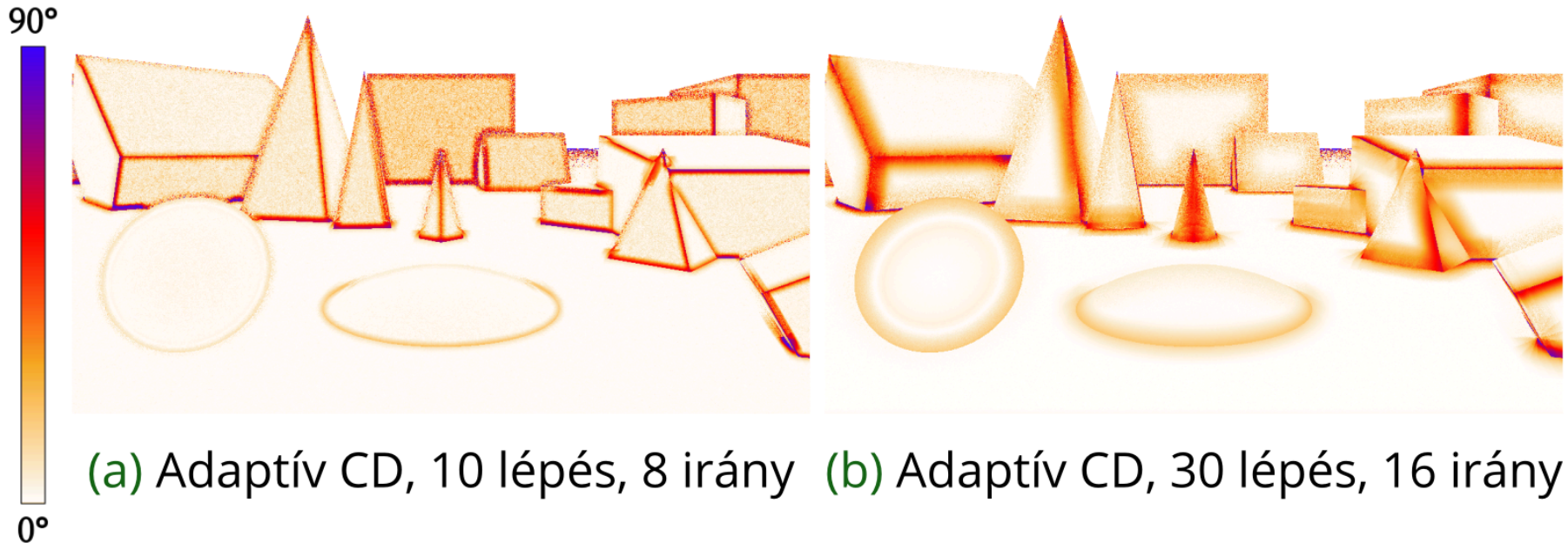
Szögbeli hiba egy gömbön. Sugár: 1.4, távolság: 3 gység, baseline: 0.3, felbontás: 1024x1024

Éldetektálás zaj mellett



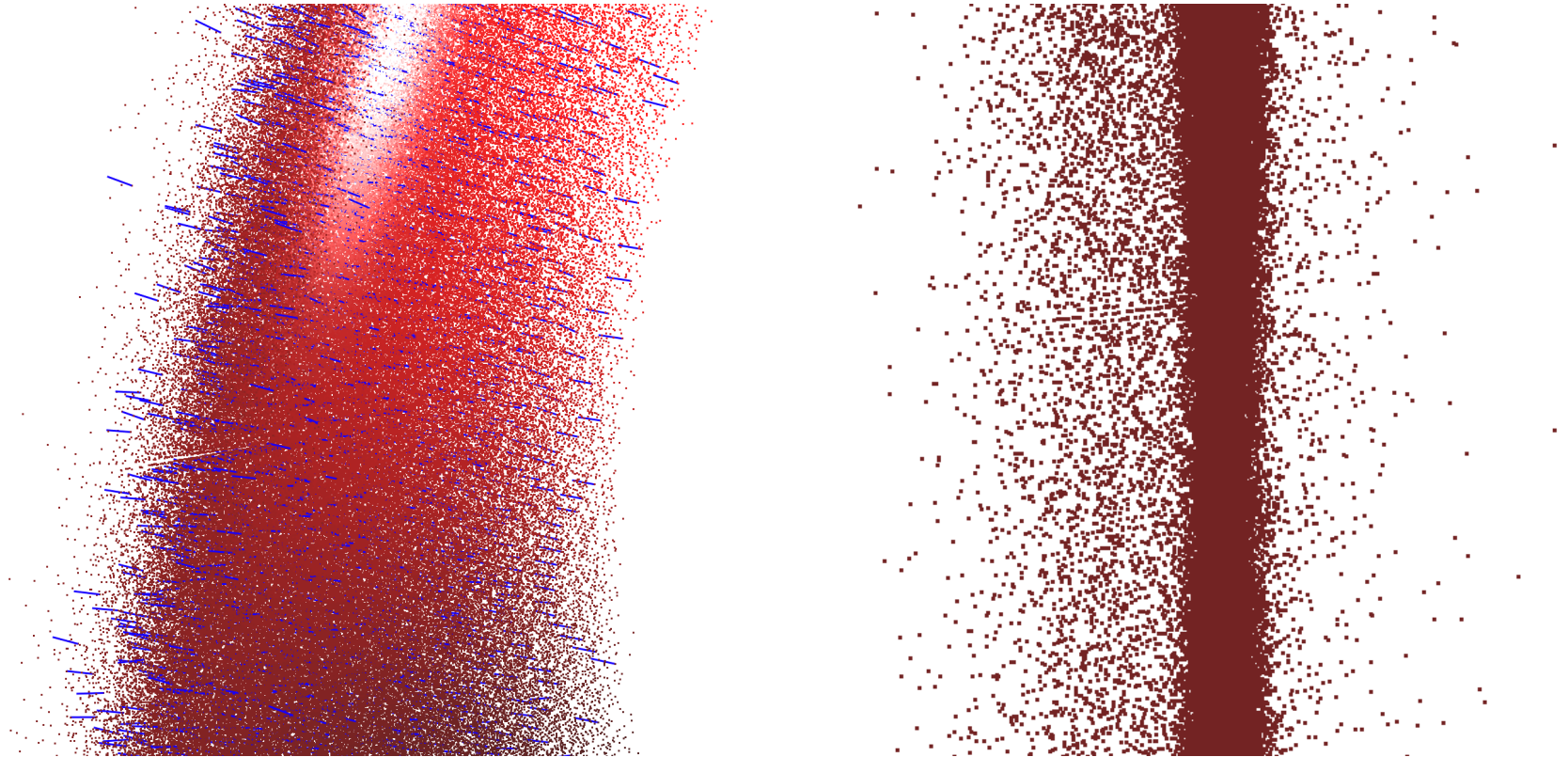
*Szögbeli hiba $\mathcal{N}(0, 0.2)$ px diszparitás zaj mellett.
Középső doboz távolsága: 15 gység, baseline: 0.3,
felbontás: 1024x720*

Éldetektálás zaj mellett



*Szögbeli hiba $\mathcal{N}(0, 0.2)$ px diszparitás zaj mellett.
Középső doboz távolsága: 15 gység, baseline: 0.3,
felbontás: 1024x720*

Kamerára merőleges felületek



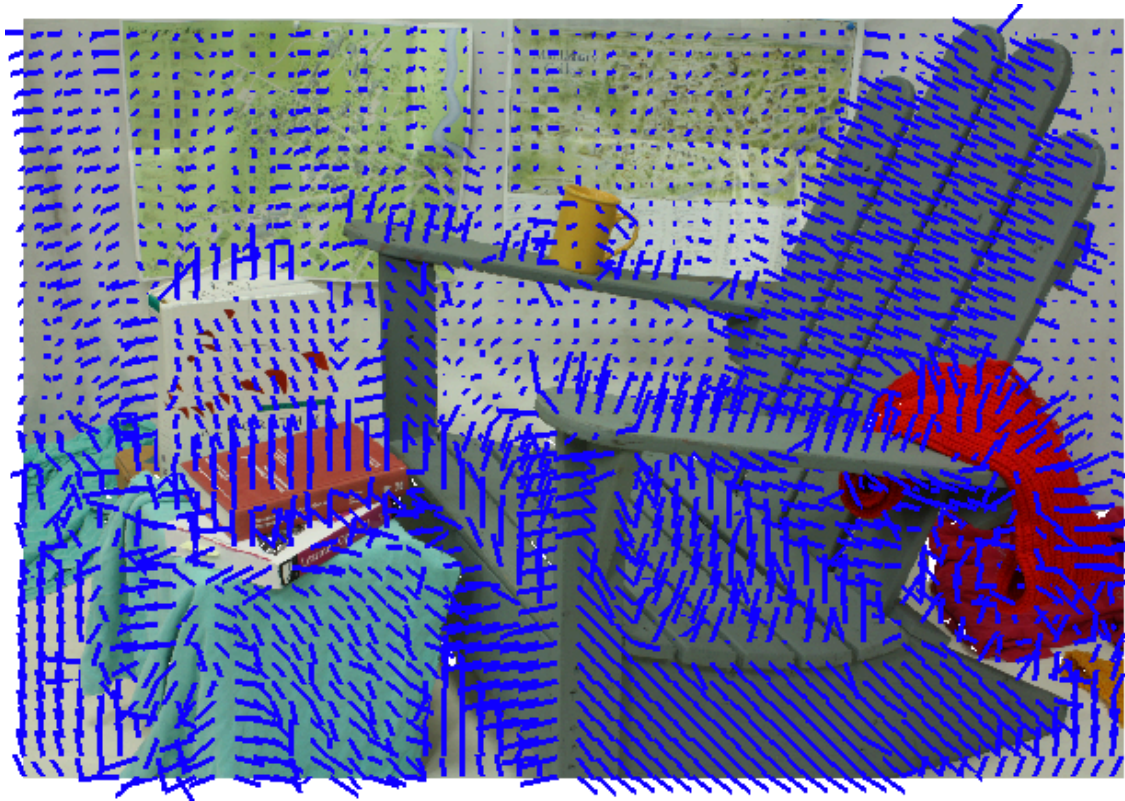
- Ugyanaz a zaj nagyobb hibát okoz

Kamerára merőleges felületek



- Ugyanaz a zaj nagyobb hibát okoz

Valós képeken



Valós képeken



Valós képeken



Valós képeken



Köszönöm a figyelmet