Dynamic fusion

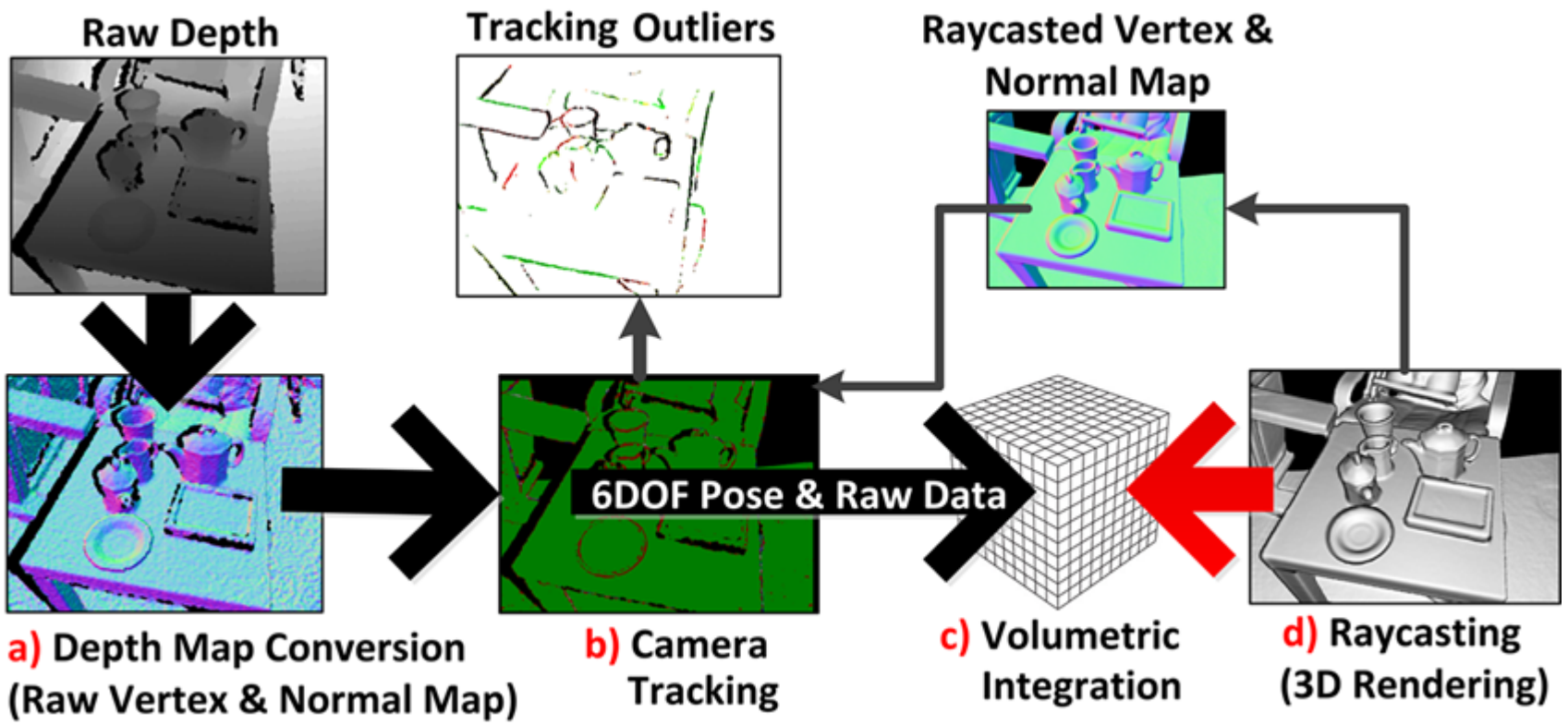
大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics

Internship Week 4-5
Kinect Fusion and Segmentation
24 March 2017

Last meeting

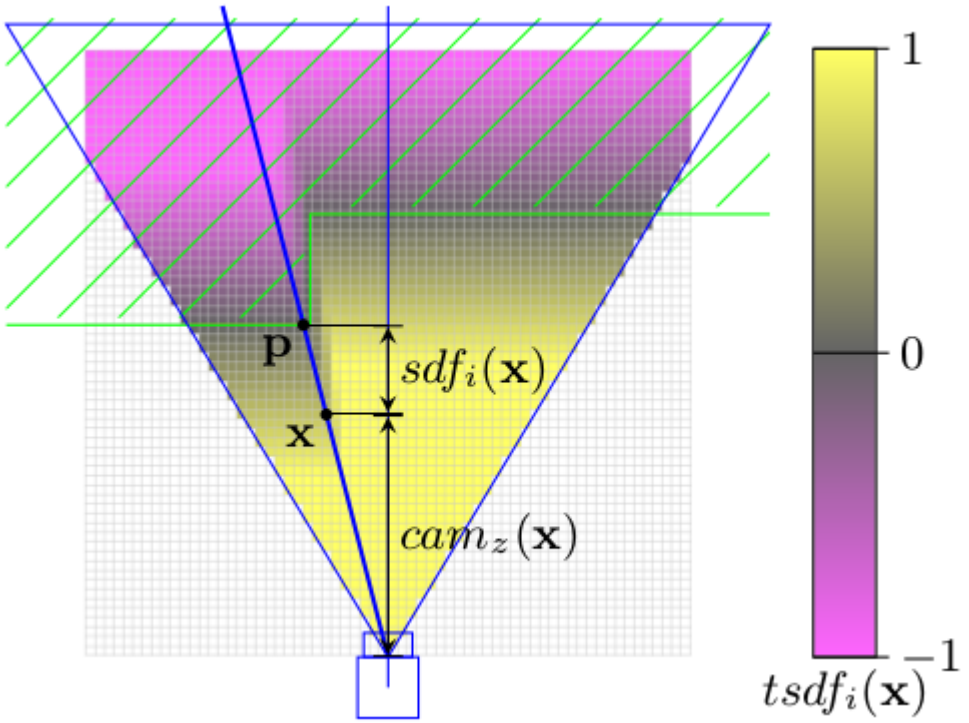
- Previously
  - Segmentation code translated but still partly from Nguyen's data.
  - Tracking
  - Kinect Fusion explanation
- Plan for the week:
  - Finish polygon optimize method
  - Presentation of Kinect Fusion
  - Segmentation from Kinect device data.
  - (Bounding boxes, coordinates changes for each segmented part of the body)

Kinect Fusion



Kinect Fusion

- TSDF



Equations

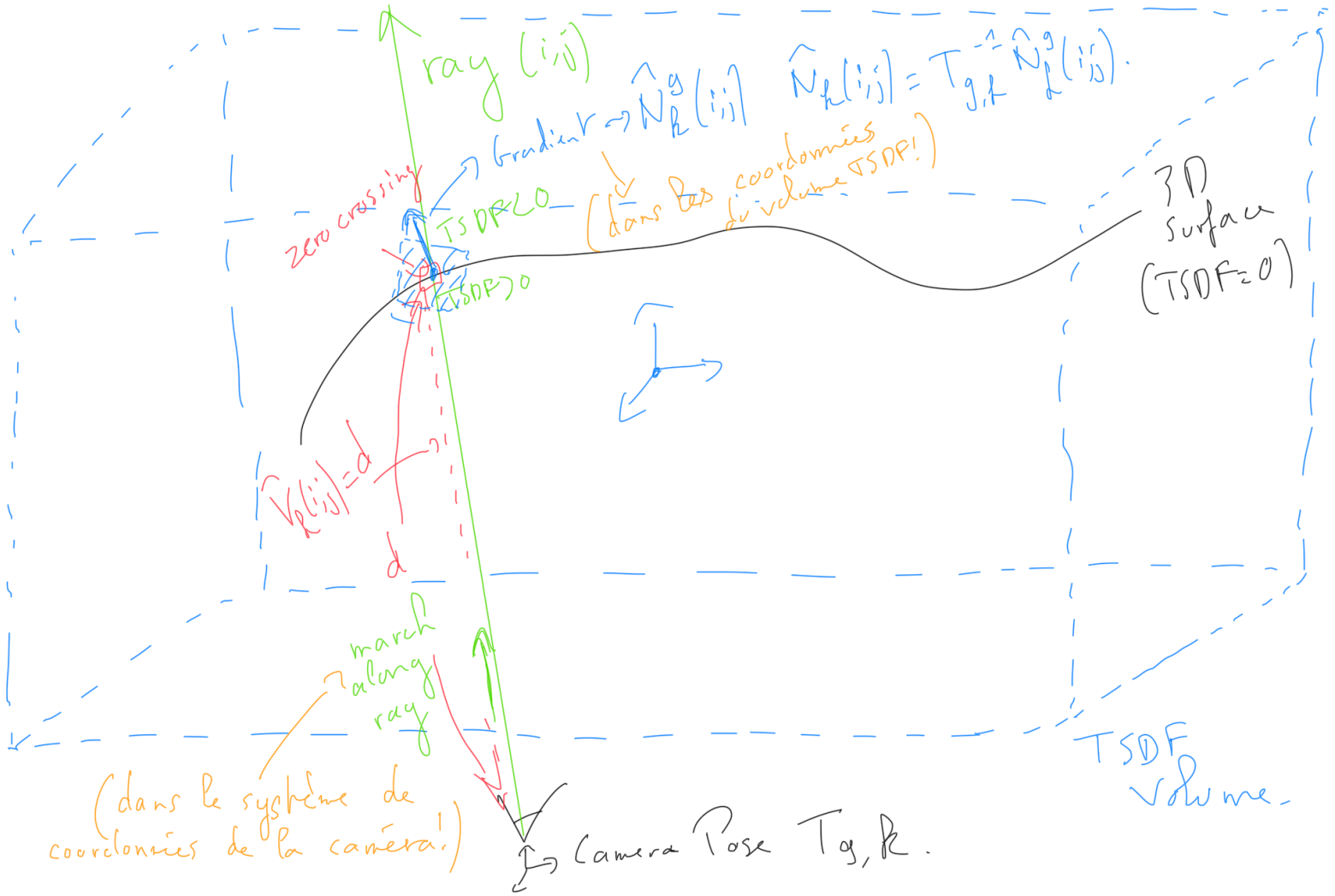
$$F_{R_k}(\mathbf{p}) = \Psi\left(\lambda^{-1}\|(\mathbf{t}_{g,k} - \mathbf{p}\|_2 - R_k(\mathbf{x})\right), \quad (6)$$

$$\lambda = \|K^{-1}\dot{\mathbf{x}}\|_2, \quad (7)$$

$$\mathbf{x} = \left\lfloor \pi\left(KT_{g,k}^{-1}\mathbf{p}\right) \right\rfloor, \quad (8)$$

$$\Psi(\eta) = \begin{cases} \min\left(1, \frac{\eta}{\mu}\right)\mathrm{sgn}(\eta) & \text{iff } \eta \geq -\mu \\ null & otherwise \end{cases} \quad (9)$$

Kinect Fusion



Raycasting the TSDF

Kinect Fusion

- Tracking

$$\mathbf{E}(\mathrm{T}_{g,k}) = \sum_{\substack{\mathbf{u}\in\mathscr{U} \\ \Omega_k(\mathbf{u})\neq\text{null}}} \left\| \left(\mathrm{T}_{g,k}\dot{\mathbf{V}}_k(\mathbf{u}) - \hat{\mathbf{V}}^g_{k-1}(\hat{\mathbf{u}}) \right)^\top \hat{\mathbf{N}}^g_{k-1}(\hat{\mathbf{u}}) \right\|_2, \quad (16)$$

$$(17)\quad \Omega(\mathbf{u}) \neq \text{null iff} \begin{cases} \mathrm{M}_k(\mathbf{u}) & = & 1, & \text{and} \\ \|\widetilde{\mathrm{T}}^z_{g,k}\dot{\mathbf{V}}_k(\mathbf{u}) - \hat{\mathbf{V}}^g_{k-1}(\hat{\mathbf{u}})\|_2 & \leq & \varepsilon_d, & \text{and} \\ \langle \widetilde{\mathrm{R}}^z_{g,k}\mathbf{N}_k(\mathbf{u}), \hat{\mathbf{N}}^g_{k-1}(\hat{\mathbf{u}})\rangle & \leq & \varepsilon_\theta. \end{cases}$$

$$(24)\quad \sum_{\Omega_k(\mathbf{u})\neq\text{null}} \left(\mathbf{A}^\top\mathbf{A}\right)\mathbf{x} = \sum \mathbf{A}^\top\mathbf{b},$$

$$(25)\quad \mathbf{A}^\top = \mathbf{G}^\top(\mathbf{u})\hat{\mathbf{N}}^g_{k-1}(\hat{\mathbf{u}}),$$

$$(26)\quad \mathbf{b} = \hat{\mathbf{N}}^g_{k-1}(\hat{\mathbf{u}})^\top \left(\hat{\mathbf{V}}^g_{k-1}(\hat{\mathbf{u}}) - \widetilde{\mathbf{V}}^g_k(\mathbf{u})\right)$$

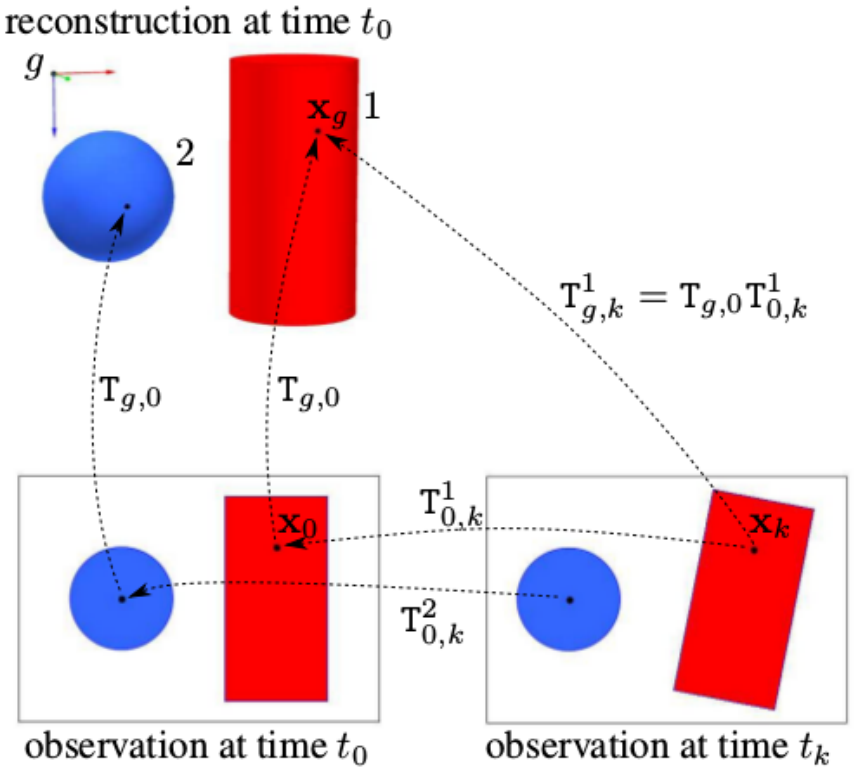
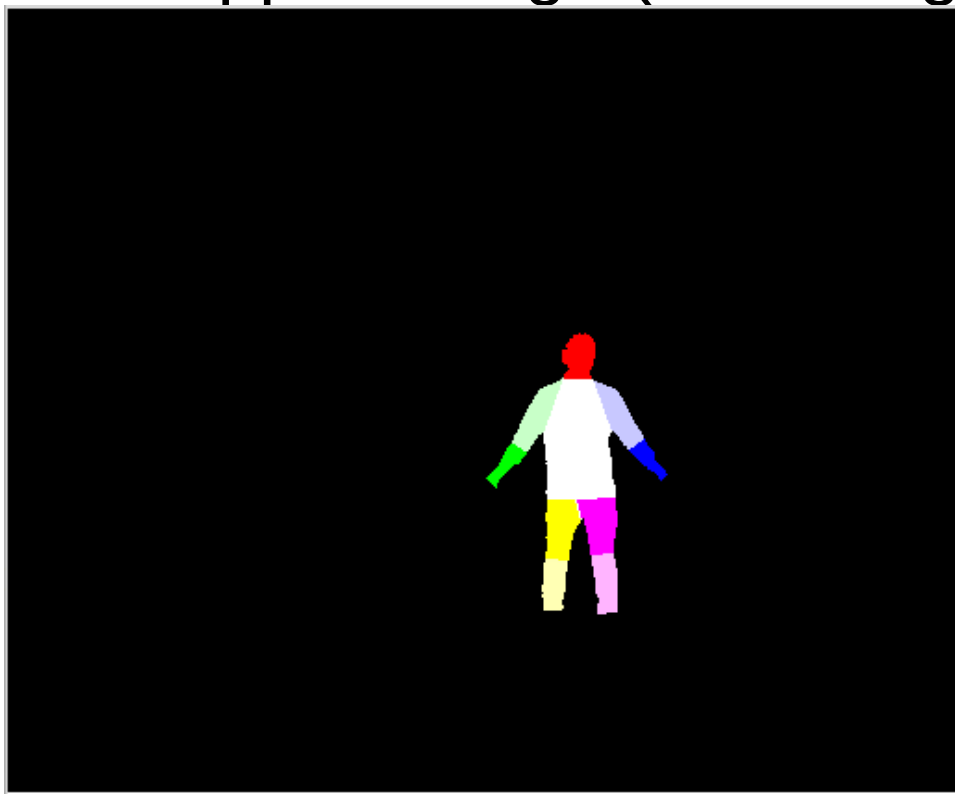


Figure 3: Euclidean transformation notation for a scene

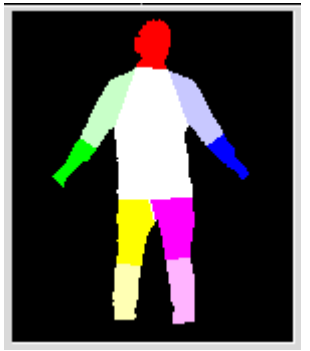
Progress

- Image segmentation from data given by Kinect device but hard-coded values necessary for threshold => cropped image (bounding boxes)



Cropped and then Thresholded image

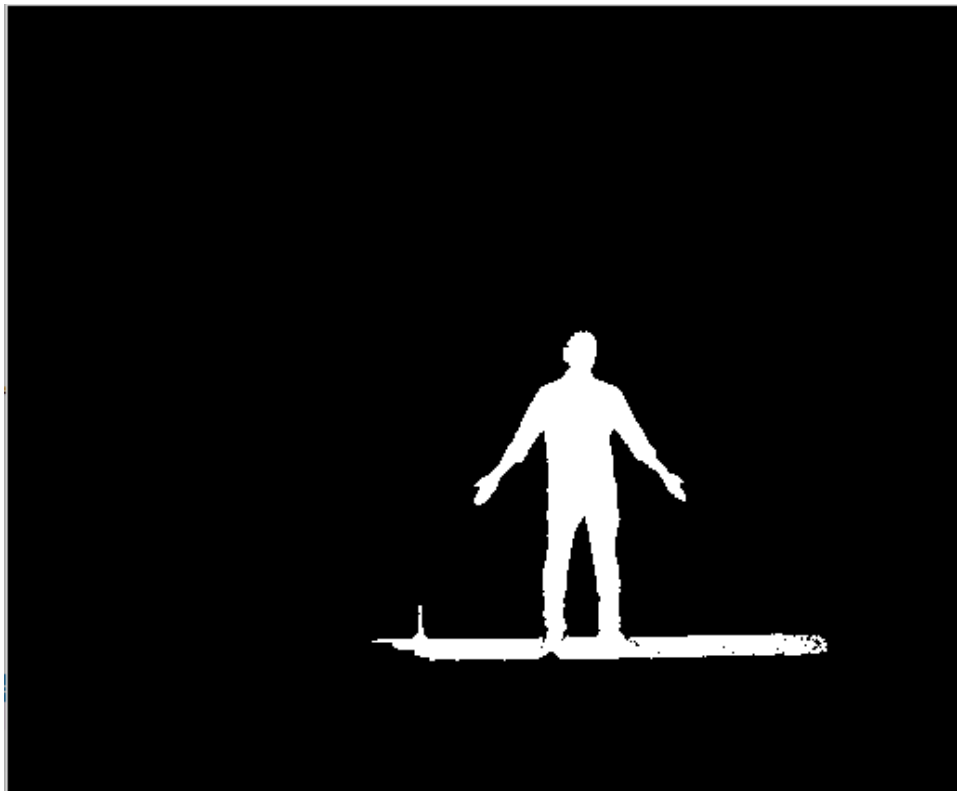
Segmentation through cropped image

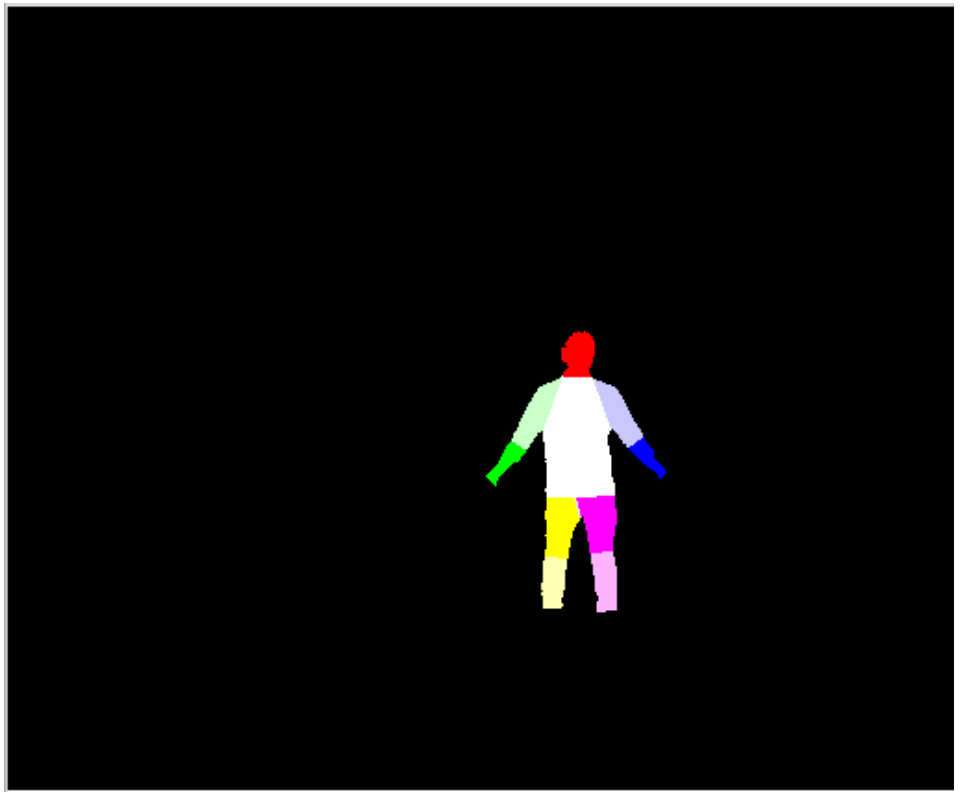


→ Segmentation done starting with kinect data device

Progress

- Cannot get the same data as Nguyen but segmentation possible



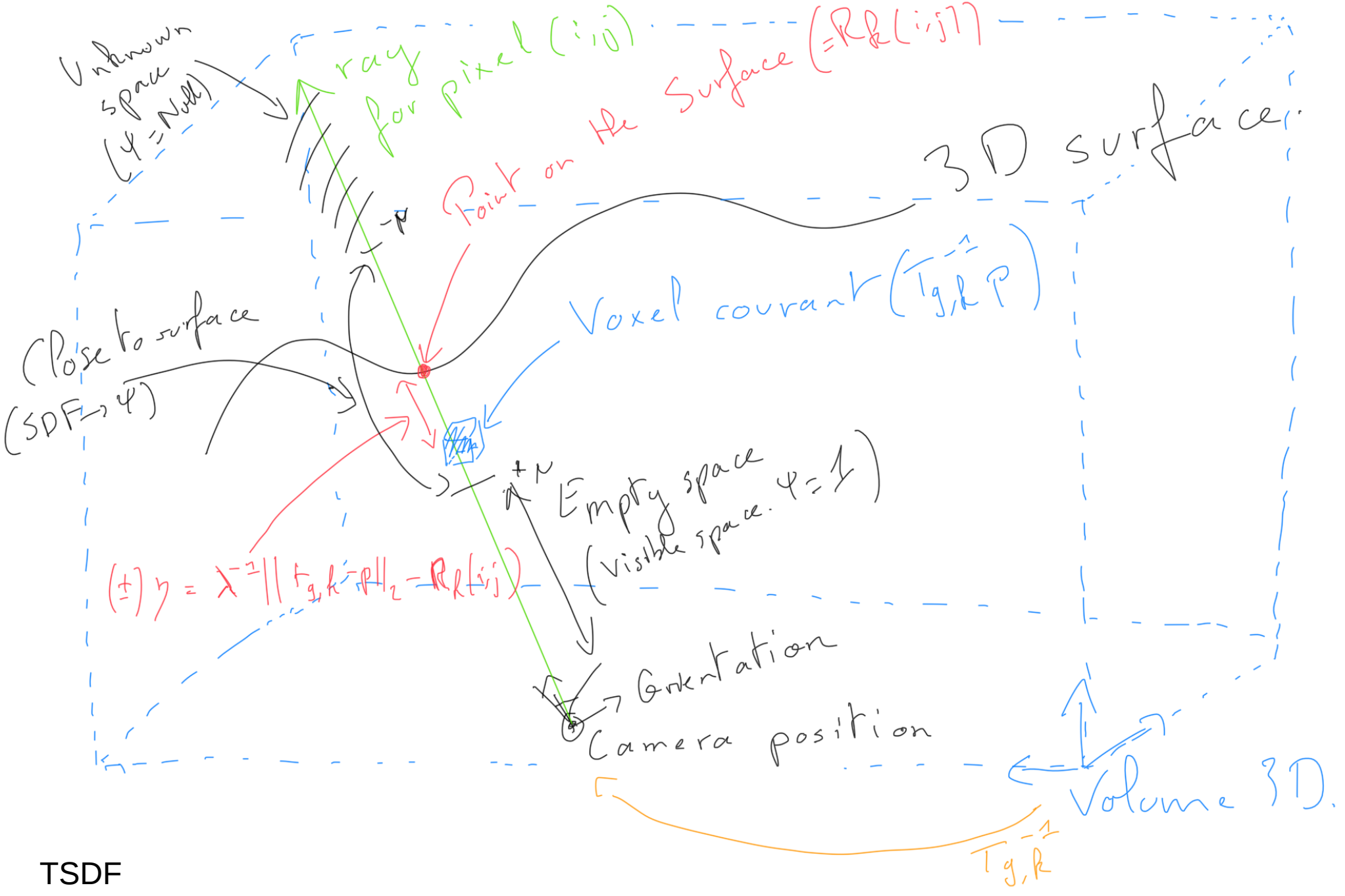


Action plan

- Bounding Boxes + coordinates changes
- Optimize TSDF
- Understand Dynamic Fusion

Q&A

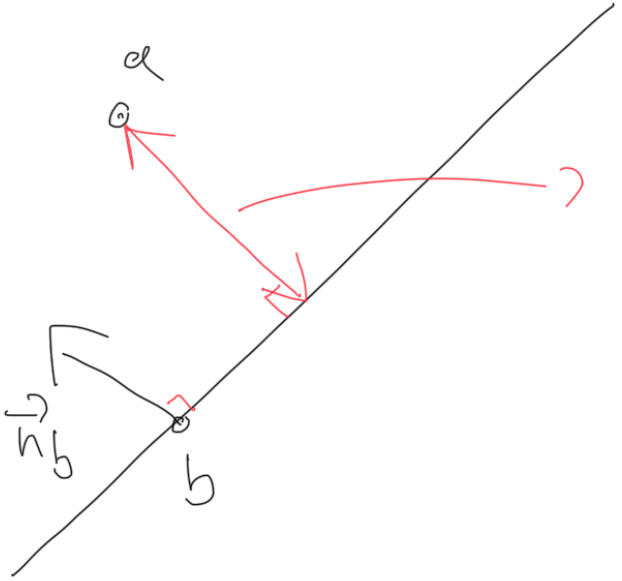
- 4DFusion : http://tinyurl.com/n8cmb9q
- Deadlines?


Unknown space (Ψ = Null)
ray for pixel (i;j)
Point on the Surface (= R_k[i;j])
3D surface
-μ
Voxel coordinate (T_g,k p)
(Pose to surface (SDF → Ψ)
+μ
Empty space (visible space. Ψ = 1)
(±) η = λ^-1 || T_g,k - p ||_2 - R_k[i;j]
Orientation
Camera position
Volume 3D
T_g,k

TSDF

Point-to-plane metric :

$\|\vec{a}-\vec{b}\|_2$ = distance Euclidienne entre 2 points.



$\vec{n}_b \cdot (\vec{a} - \vec{b})$ = Distance du point $\vec{a}$ au plan tangent à $\vec{b}$ (de normale $\vec{n}_b$)