3D scene reconstruction using a Time-Of-Flight camera

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

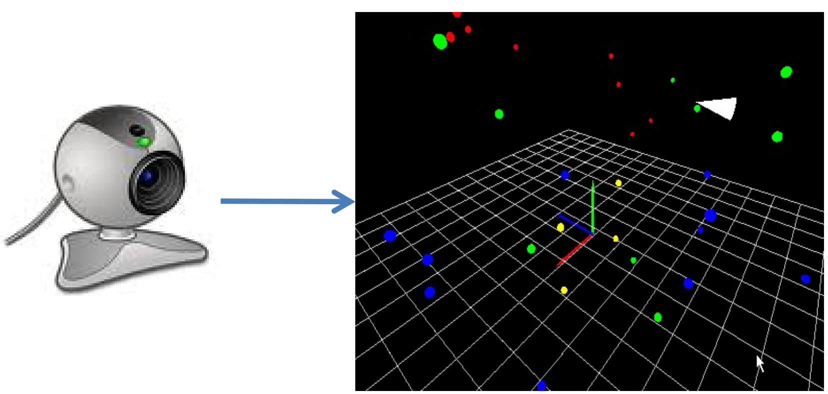
Outline

- Introduction
- Camera Time-Of-Flight
- Data filtering
- Camera motion estimation and 3D map
- Meshes generation from points cloud
- Conclusion & perspectives

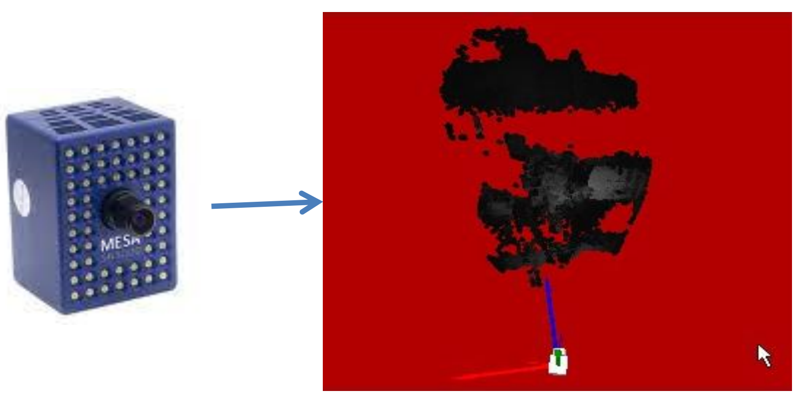
Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

Introduction

- Goal: perform a 3D reconstruction of the surrounding environnement using a Time-of-flight (TOF) camera
- Context: static scene and moving camera (Simultaneous Localization And Mapping problem: SLAM)
- Previous works:
    - SLAM using monocular camera
    - Limitations: sparse reconstructed map
- New method: based on Time-Of-Flight sensor (dense reconstruction)



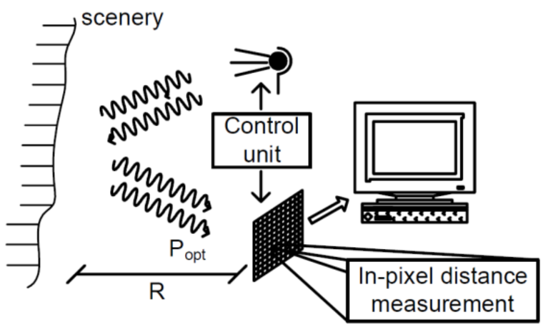
Webcam SLAM

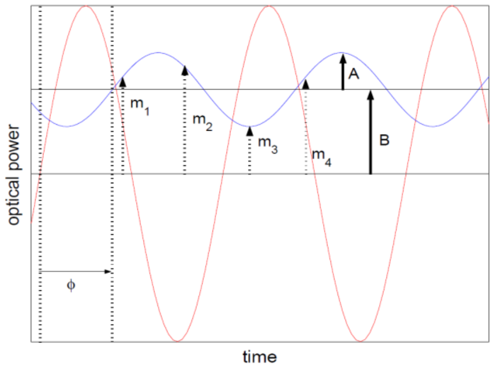


TOF SLAM

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

TOF camera (introduction)





- Camera Time-of-flight (TOF) is an active range finder:
  - Emits light (LEDs) and Measures TOF (CCD sensor)

- Reconstructed signal (4 measurements each period) characterized by
  - ➢ Offset B: Mean of the total light incident on the sensor: background plus modulated signal
  - ➢ Amplitude: Object reflectivity and object distance
  - ➢ Phase shift: Distance

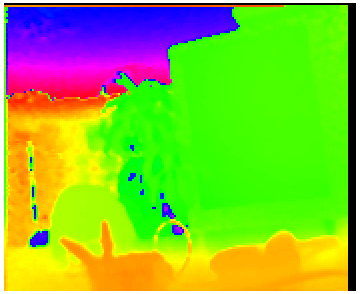
- The distance accuracy of the measurement depends on the acquired amplitude A and offset B

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

TOF camera (pros and cons)

- Swissranger SR3000
  - Number of LEDs: 55
  - Resolution: 176x144
  - Output: amplitude + distance
  - Frame rate: 30 fps

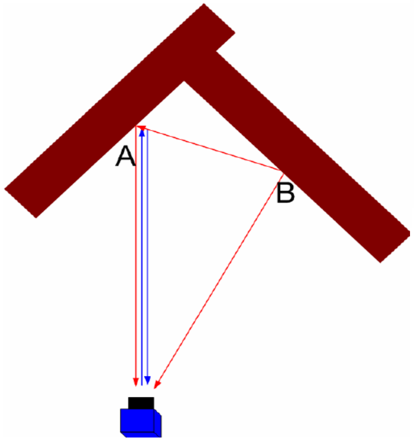
Amplitude



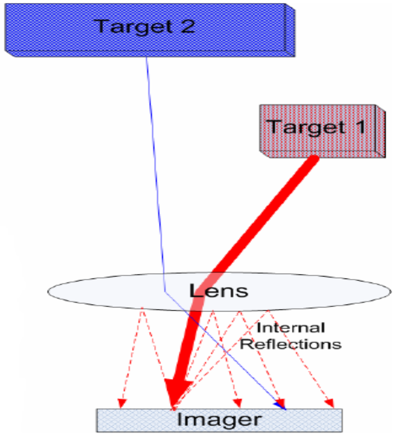
Depth

- Advantages:
  - ➢ Easy to extract range information
  - ➢ High frame rate
  - ➢ Not suffer from missing texture in the scene or bad lighting conditions
  - ➢ Compactness of the sensor

- Disadvantages:
  - ➢ Multiple reflections
  - ➢ Light scattering
  - ➢ Low resolution (176x144)
  - ➢ Camera network

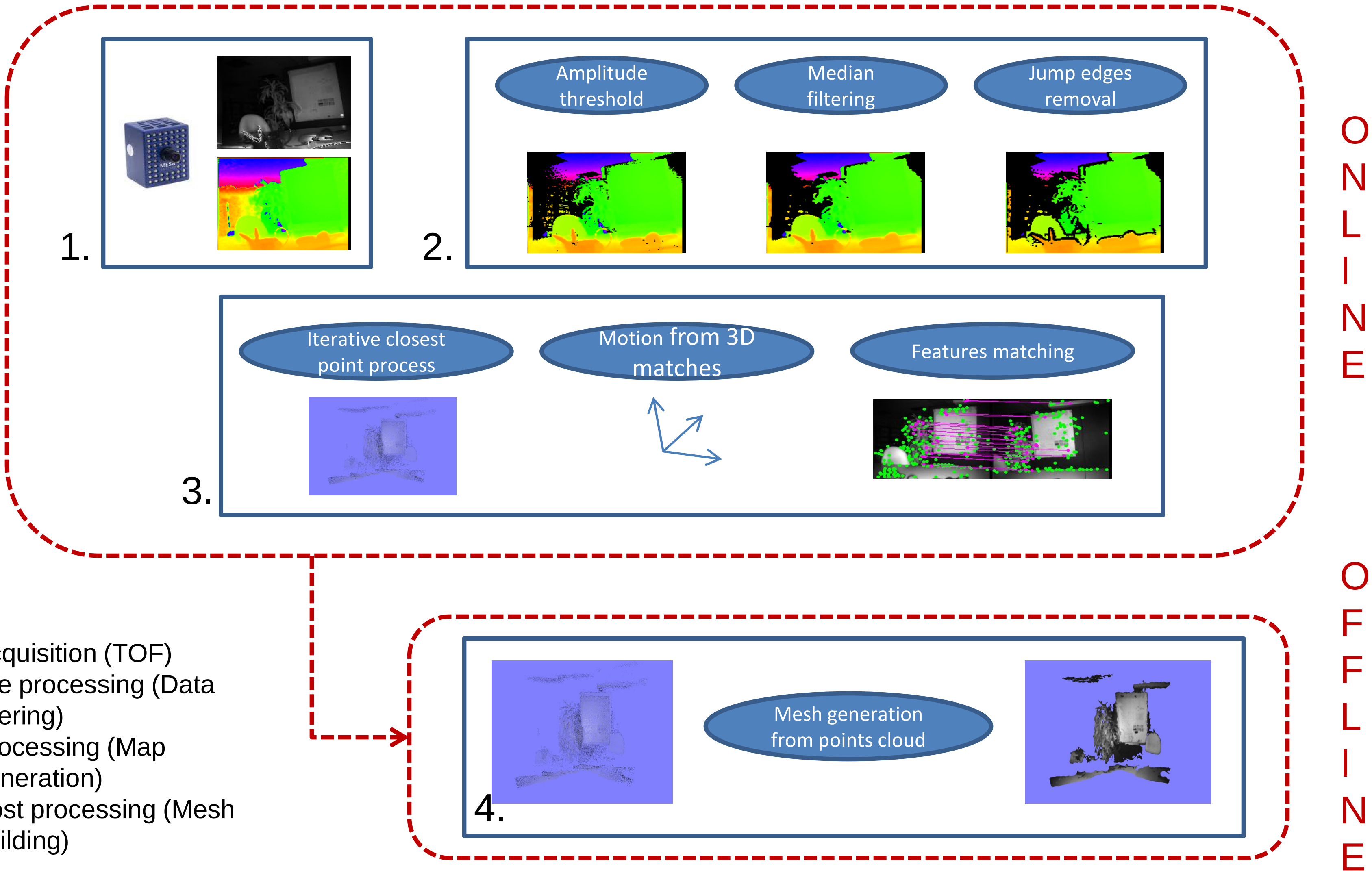


Multiple reflections



Light scattering

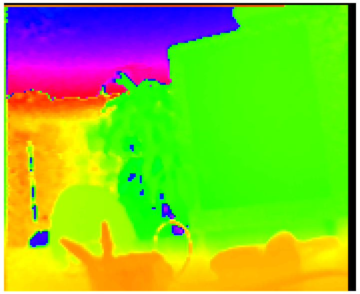
Whole process overview



1. Acquisition (TOF)
2. Pre processing (Data filtering)
3. Processing (Map generation)
4. Post processing (Mesh building)

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

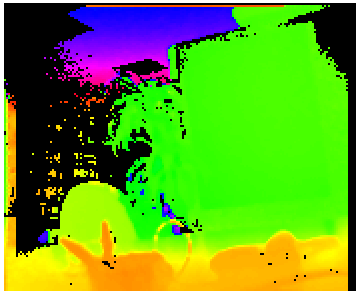
Preprocessing: filtering



Raw distance

Raw amplitude

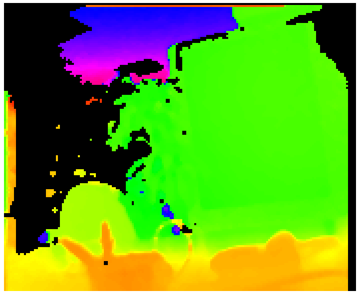
- Raw data provided by TOF camera are very noisy
- Necessary to clean it



Amplitude threshold

- Depth accuracy depend on amplitude value
  - Lower is the amplitude, lower is the accuracy
    - Far objects
    - Low reflectivity

→ Threshold based on amplitude



Median filtering

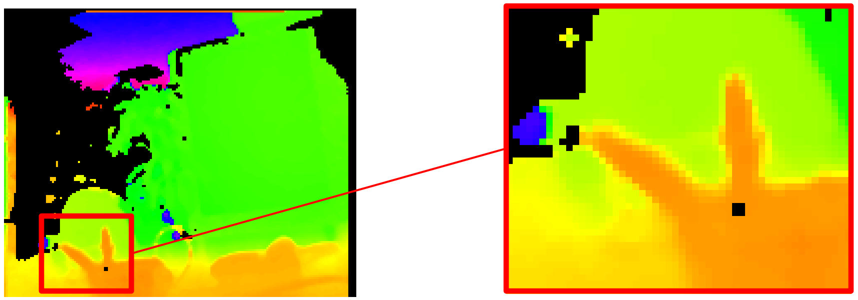
- Pepper and salt noise is very important due to
  - Sensor noise as light scattering
  - Amplitude thresholding

→ Median filtering

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

Jump edge filtering

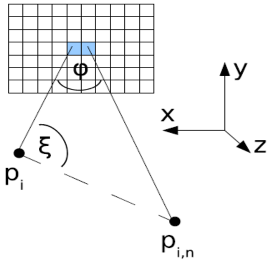
- Problems on abrupt distance transitions
  - TOF camera smooth abrupt depth transitions
  - A small measurement error provide an important error on distance value

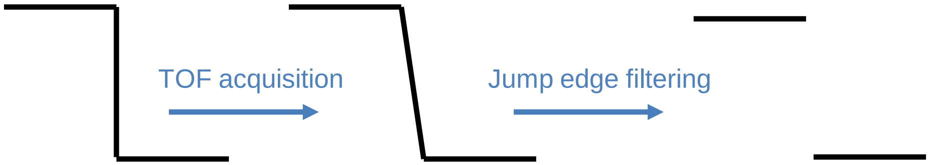


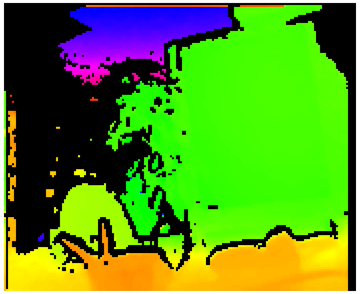
- Green: chair
- Orange: pencils
- Yellow: Noise (smooth transition)

-> JUMP EDGE FILTERING

- Higher is the $\xi_{i,n}$ angle value, higher is the depth transition
- Knowing camera intrinsic parameters, compute $\xi_{i,n}$ (from φ and 3D points)
- Remove pixels with at least 2 neighbors with an angle $\xi_{i,n} > s$

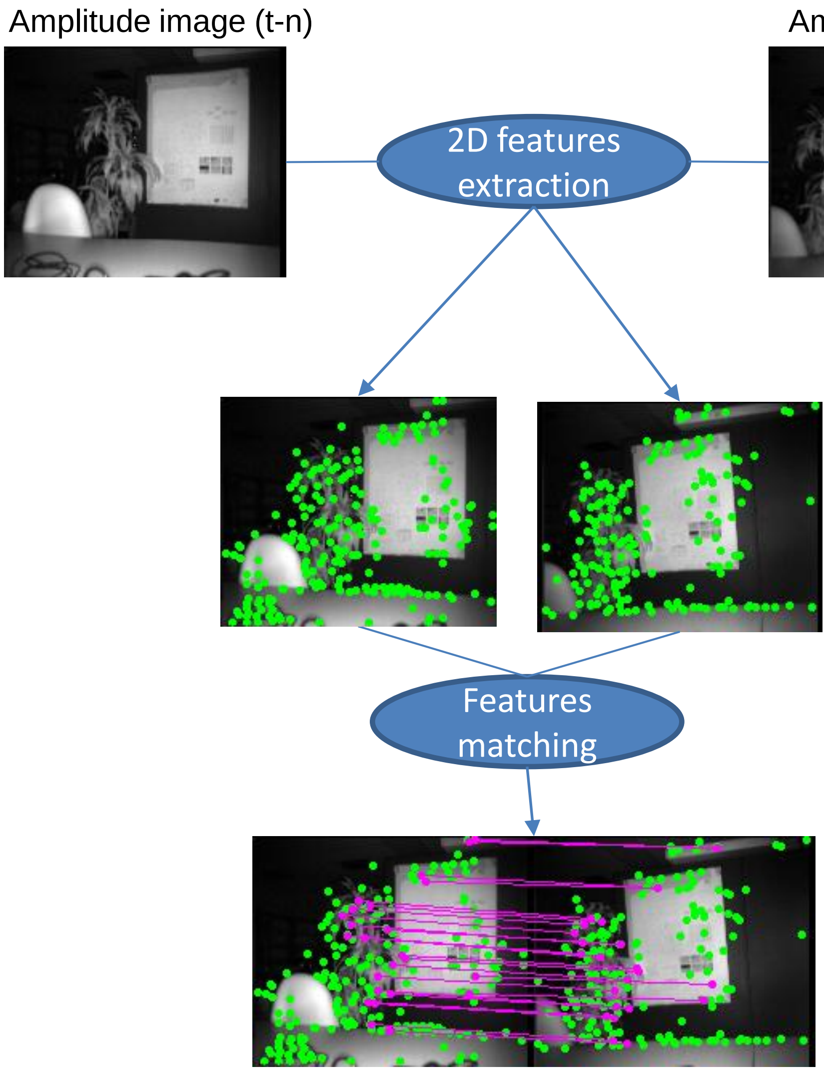






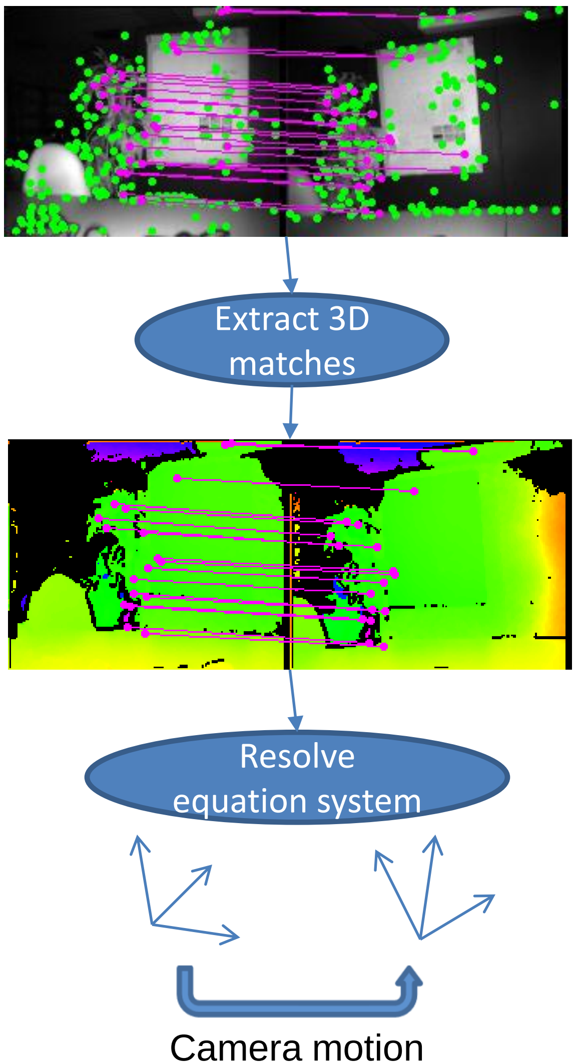
Jump edges filtering

Processing: 2D matching



- Use Amplitude image (2D information) to make matches between 2 images
  - ➢ Detect SIFT features into the current frame
  - ➢ Match features using SIFT descriptors between the last reference frame and the current one

Processing: 3D matching



- For each match $p1 = Rp2 + t$

- Rotation matrix is represented as a quaternion
  - Reduce number of parameters to estimate
  - Maintain constraints (matrix normalization)

$$\mathbf{q} = (w, x, y, z) \mapsto \mathbf{R} = \begin{bmatrix} w^2 + x^2 - y^2 - z^2 & 2xy - 2wz & 2xz + 2wy \\ 2xy + 2wz & w^2 - x^2 + y^2 - z^2 & 2yz - 2wx \\ 2xz - 2wy & 2yz + 2wx & w^2 - x^2 - y^2 + z^2 \end{bmatrix}$$

- Estimate q from 3D matches (rotation)
  - Linear equation system (SVD)

- Apply rotation found to the first point set
- Compute centroids of each point set
- Extract t from point centroids (translation)

Processing: Outliers

- Two kind of outliers
  - Bad matches between 2 SIFT points
  - Bad distance measurement by TOF camera

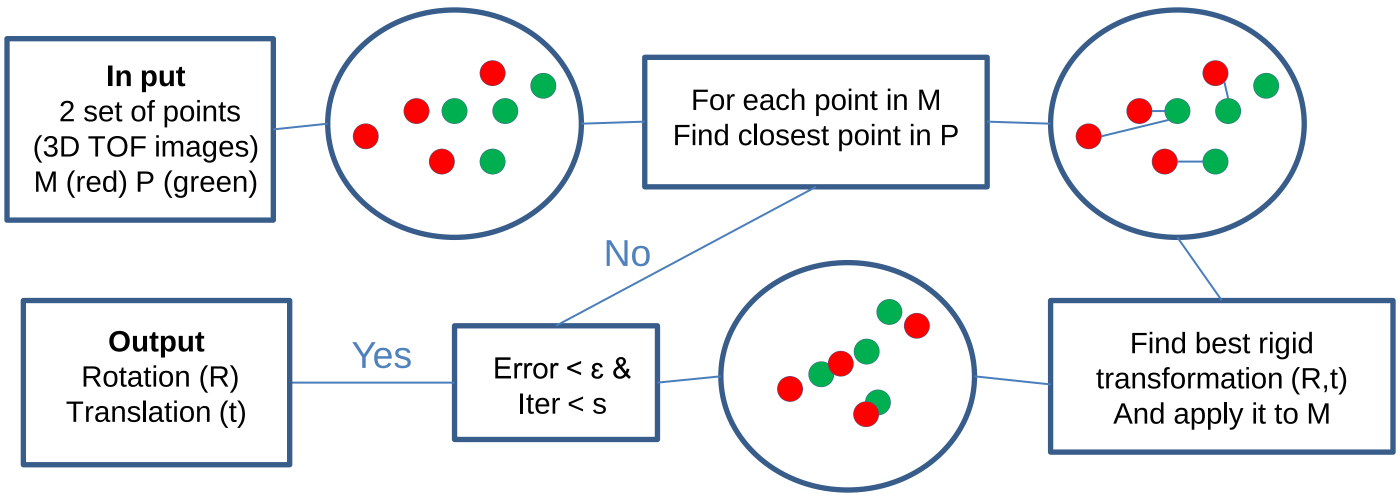
→ RANSAC

- while iterations < **k**
  - Select randomly 3 points into the two 3D point sets S1 and S2
  - Estimate q and t
  - Apply (q,t) to not selected points in S1 (S'1)
  - Compute error between S'1 and corresponding points in S2
  - Extract list of P points with an error < **e** (S''1)
  - If (P>**d**)
    - evaluate (q,t) with all points into (S''1)
    - If (q,t) is the best model, update the best model (q,t)
- Return the best model (q,t)

- If RANSAC finished correctly and estimate motion is higher than a threshold (rotation or translation)
  - ➢ Update camera motion and 3D map
  - ➢ Current image is the next reference image

Processing: ICP

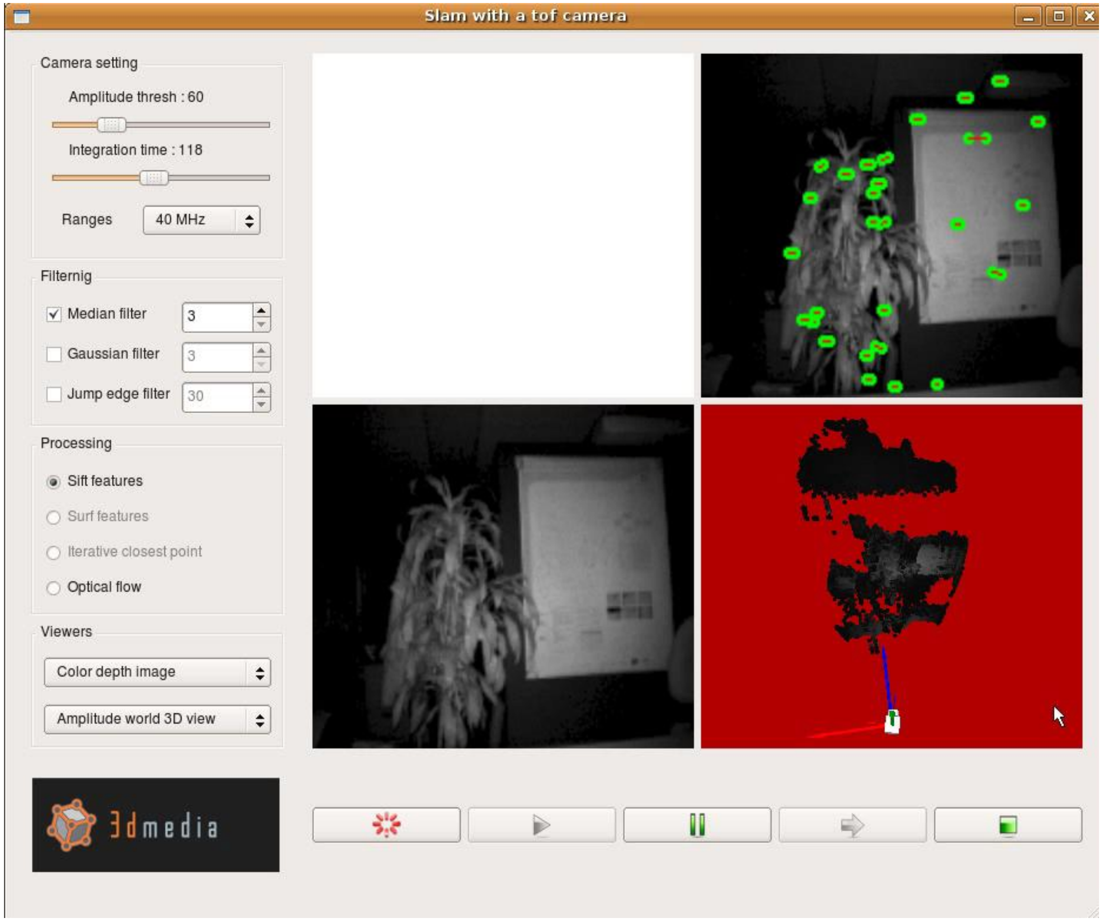
- Estimate the motion using only few matches is not enough
- A global optimization is required to improve preciseness of the motion estimation → Iterative Closest Point (ICP)



- Whole distance maps are used to estimate transformation
- The output motion is validated if error is small (not based on iter) and motion is lower than a threshold

Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

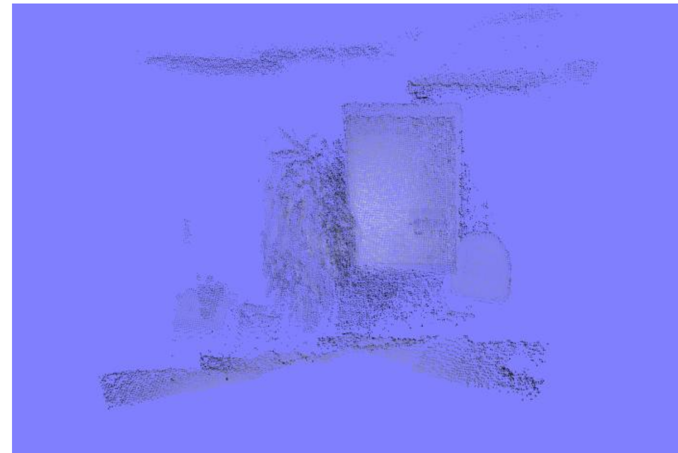
Processing: Demo



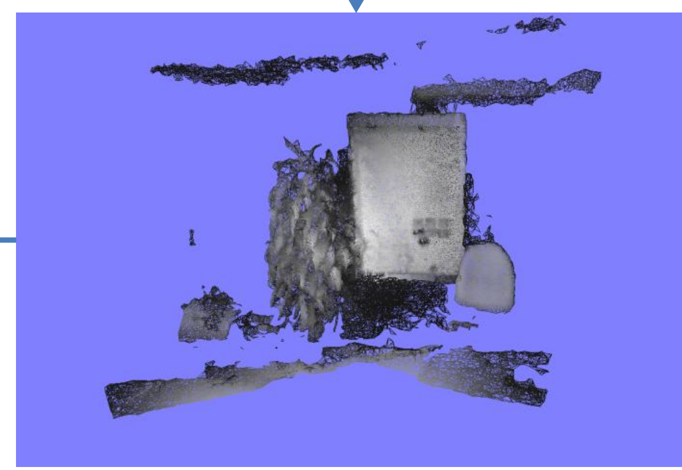
Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

Postprocessing: 3D meshes

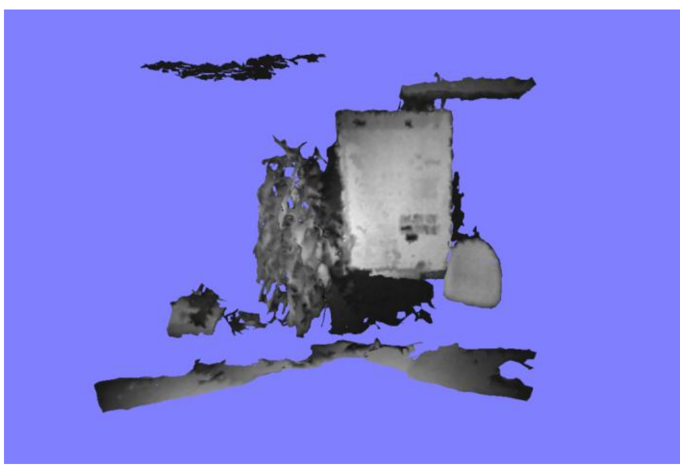
- Generate a mesh from 3D point cloud (meshlab library)
  - ➢ **Ball pivoting**, marching cubes Poisson's reconstruction
- Filter mesh
  - **Based on objects size**
  - Subdivision algorithms
  - **Hole filling**
  - **Laplacian smoothing, …**



Surface reconstruction



Filtering



Le Fonds européen de développement régional et la Région wallonne investissent dans votre avenir

Conclusion & Perspectives

- Conclusion
  - ➢ SLAM using TOF camera
  - ➢ Dense 3D map reconstruction of the scene
  - ➢ Map represented by meshes

- Perspectives
  - ➢ Investigate different versions of ICP algorithm (Trimmed ICP, …)
  - ➢ Calibration of the depth map (relative to amplitude image, …)
  - ➢ Use a temporal motion estimation model (Extended Kalman Filtering, Particle Filtering)

Thank you for your attention

Questions?