UNIVERSITÀ DEGLI STUDI FIRENZE

Structural Change Detection
by
Direct 3D Model Comparison

Marco Fanfani
Carlo Colombo

Computational Vision Group – CVG
Dept. of Information Engineering – DINFO
University of Florence, Italy

VISAPP 2019 – Prague, February 26th 2019

Introduction

- Detecting **structural changes in a 3D scene** over time is required in several application scenarios
  - Archaeological campaign documentation
  - Hydro-geological monitoring
  - Building construction management
  - Tracking urban architectural changes

Introduction

- Detecting **structural changes in a 3D scene** over time is required in several application scenarios
  - Archaeological campaign documentation
  - Hydro-geological monitoring
  - Building construction management
  - Tracking urban architectural changes

- **Change detection methods**
  - Human visual inspection
  - Comparison of **two 2D images** or sequences (Radke et al. 2005) – requires fixed viewpoint and illumination compensation
  - Compatibility check between **an image sequence and a 3D model** (Taneja et al. 2011) – needs a subsequent classification step into structural (e.g. buildings) vs non structural (e.g. pedestrians, cars) changes

Introduction

- Detecting **structural changes in a 3D scene** over time is required in several application scenarios
  - Archaeological campaign documentation
  - Hydro-geological monitoring
  - Building construction management
  - Tracking urban architectural changes

- **Change detection methods**
  - Human visual inspection
  - Comparison of **two 2D images** or sequences (Radke et al. 2005) – requires fixed viewpoint and illumination compensation
  - Compatibility check between **an image sequence and a 3D model** (Taneja et al. 2011) – needs a subsequent classification step into structural (e.g. buildings) vs non structural (e.g. pedestrians, cars) changes

- **Our solution**
  - Fully automatic
  - Based on the direct **comparison of 3D models** of a time-changing environment obtained by image sequences acquired at two different times
  - Relatively insensitive with respect to illumination and viewpoint changes
  - Detects only changes related to structural (fixed) elements in the scene (SfM)

Method overview

T_0

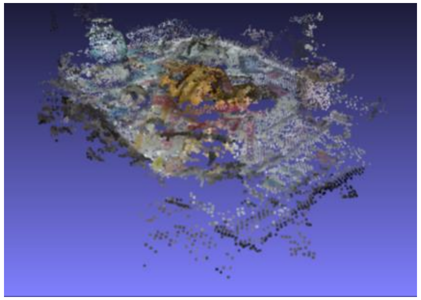
- Input: **image sequences** of the same scene acquired at two different times

T_1

- Pre-processing: construction of **sparse 3D models** with off-the-shelf software (VisualSfM, COLMAP)
- Step 1: **photometric rigid registration** to align the 3D reconstructions
- Step 2: **structural change detection** by comparison of the 3D models

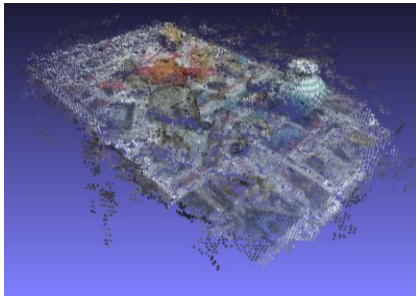
Method overview

T_0



- Input: **image sequences** of the same scene acquired at two different times

T_1

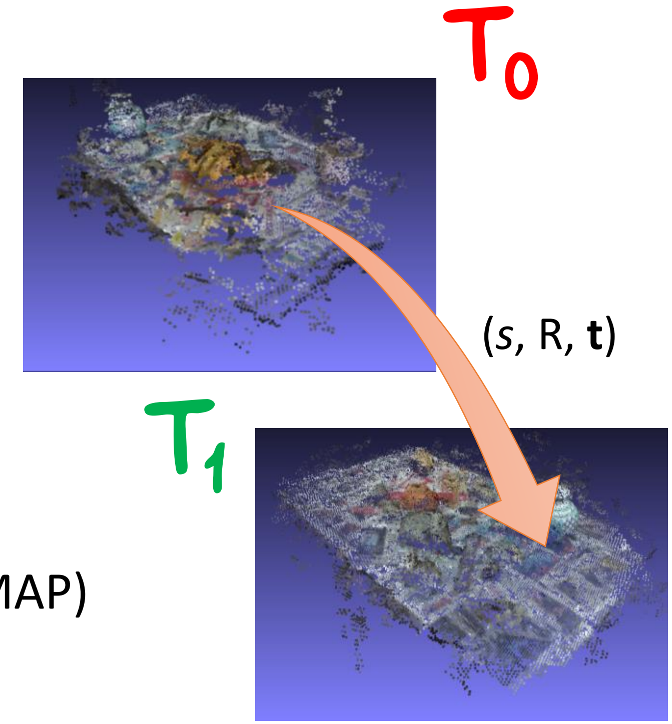


- Pre-processing: reconstruction of **3D models** with off-the-shelf software (VisualSfM, COLMAP)

- Step 1: **photometric rigid registration** to align the 3D reconstructions

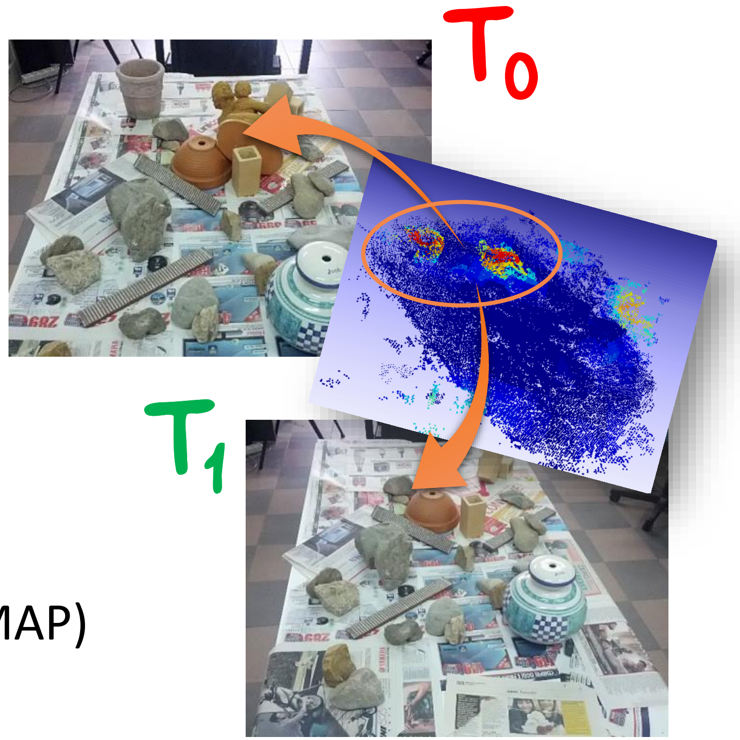
- Step 2: **structural change detection** by comparison of the 3D models

Method overview



- Input: **image sequences** of the same scene acquired at two different times
- Pre-processing: reconstruction of **3D models** with off-the-shelf software (VisualSfM, COLMAP)
- Step 1: **photometric registration** (7 DOFs) to align the 3D reconstructions
- Step 2: **structural change detection** by comparison of the 3D models

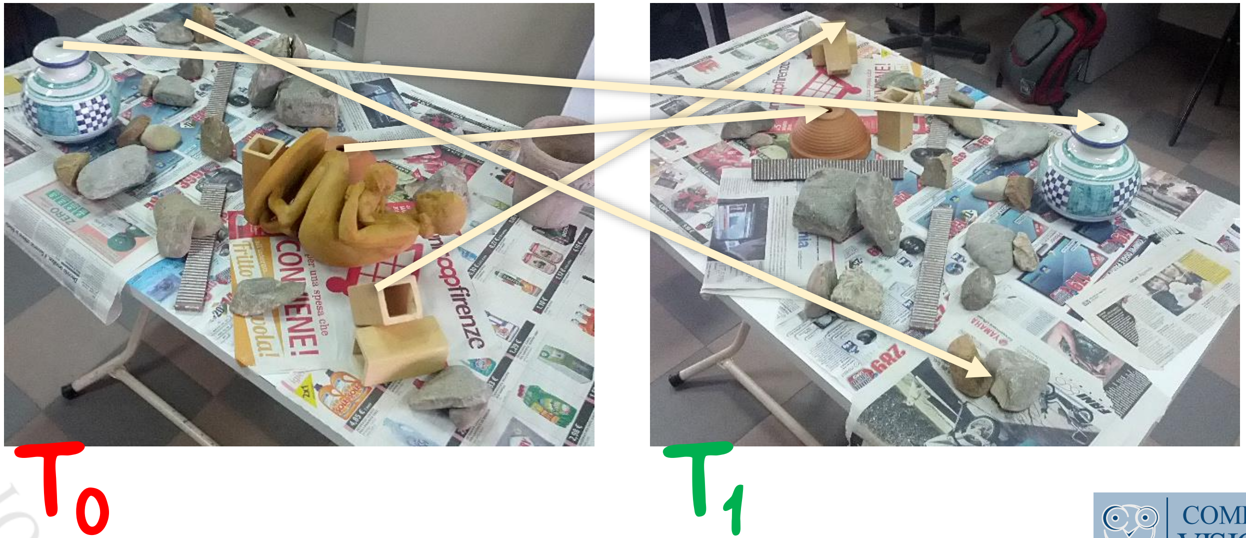
Method overview



- Input: **image sequences** of the same scene acquired at two different times
- Pre-processing: reconstruction of **3D models** with off-the-shelf software (VisualSfM, COLMAP)
- Step 1: **photometric registration** (7 DOFs) to align the 3D reconstructions
- Step 2: **structural change detection ("heat map")** by 3D model comparison

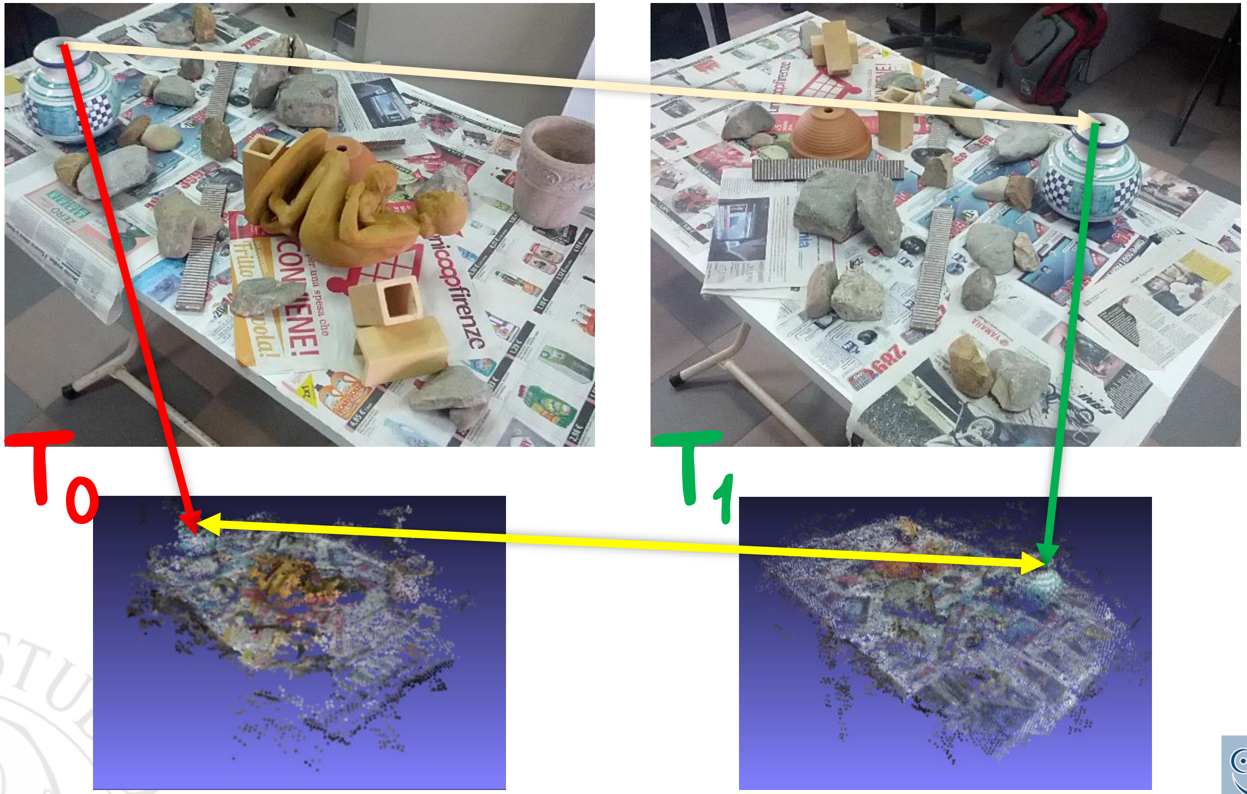
Photometric registration

- After reconstruction we obtain for each sequence:
  - A 3D model, expressed in an arbitrary reference system and scale
  - A set of 2D feature descriptors and inter-frame tracks
  - A list of 2D/3D correspondences
- Registration is obtained by **feature matching** between the images of the two sequences, and then…



Photometric registration

- …2D matches are promoted to 3D, both in T_0 and T_1
- The 3D point map between T_0 and T_1 can finally be estimated from 3D correspondences



Photometric registration

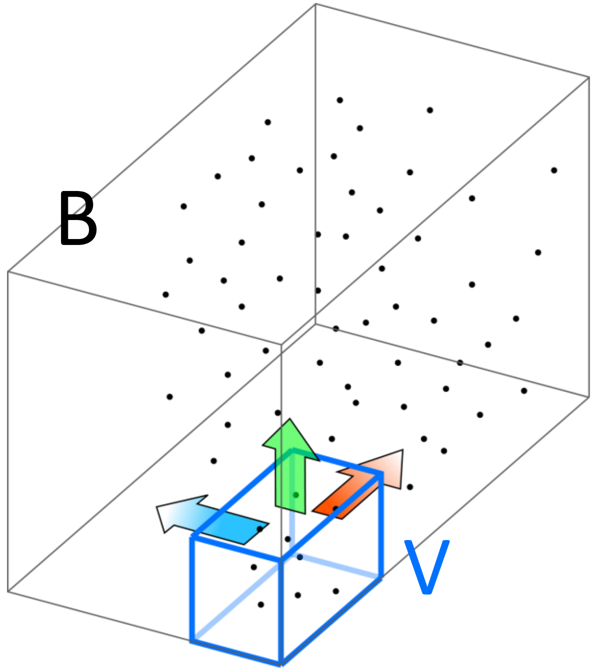
- With initial 3D correspondences given, the transformation τ that maps T_1 onto T_0 is obtained by **minimizing the symmetric re-projection error** in a RANSAC framework:

$$argmin_{\tau} \sum_{\mathbf{x} \in \mathrm{T}_0, \mathbf{X}' \in \mathrm{T}_1} |\mathbf{x} - \tau(\mathbf{X}')|^2 + \sum_{\mathbf{x}' \in \mathrm{T}_1, \mathbf{X} \in \mathrm{T}_0} |\mathbf{x}' - \tau^{-1}(\mathbf{X})|^2$$

- To refine τ, the **Iterative Closest Point** (ICP) algorithm is used:
  - Sparse 3D models are **densified** with a region growing method (Furukawa & Ponce 2010)
  - T_1 is mapped with τ into the reference system of T_0
  - ICP is run to obtain a finer registration

Structural change detection

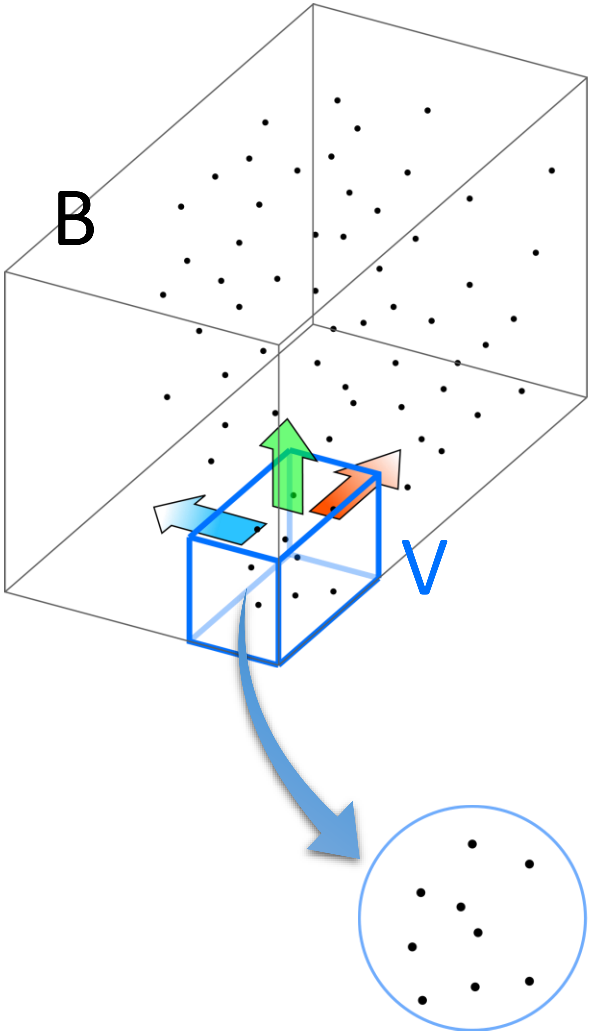
- Structural changes are searched using a **shifting 3D volume V** that slides within a 3D box B including the dense 3D reconstructions
- V is shifted with a volume overlap of 3/4
- Corresponding volumes V on T_0 and T_1 are **compared with three criteria**:
  - Quantity
  - Orientation
  - Occupancy



Structural change detection

Quantity criterion

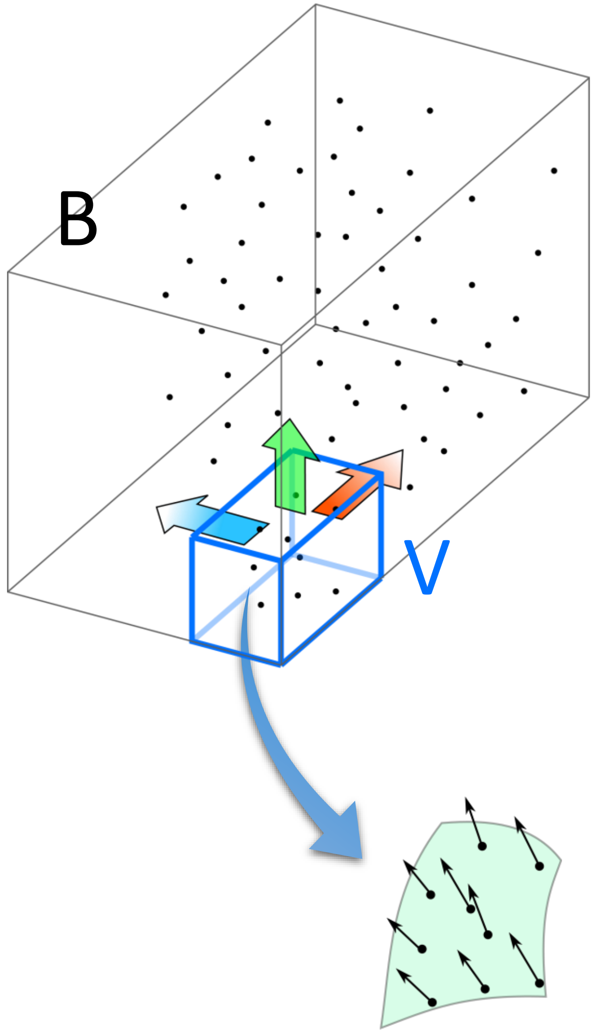
- Counts the number of 3D points included in V both in T_0 and T_1
- If a difference greater than α is found, the criterion is satisfied
- α is set to the average number of points within V, considering both T_0 and T_1



Structural change detection

Orientation criterion

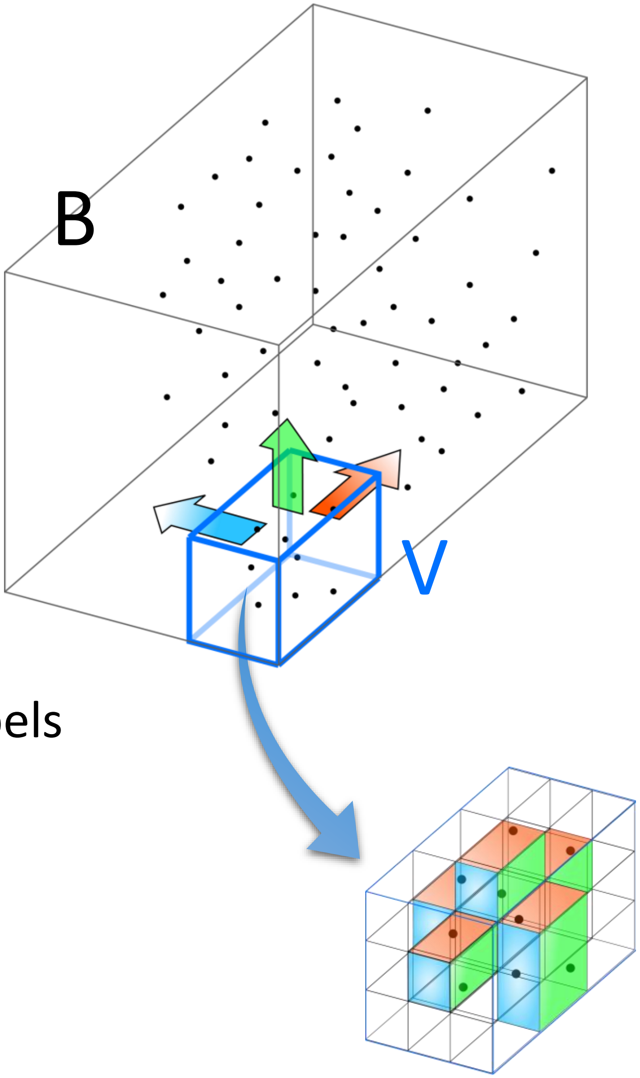
- For each 3D point a **normal vector** is computed and parametrized by its angles
- For each position of V, the **distribution** of the angles of the normal vectors is stored into a 2D histogram
- **Histograms are compared**: if their distance is greater than β, the criterion is satisfied



Structural change detection

Occupancy criterion

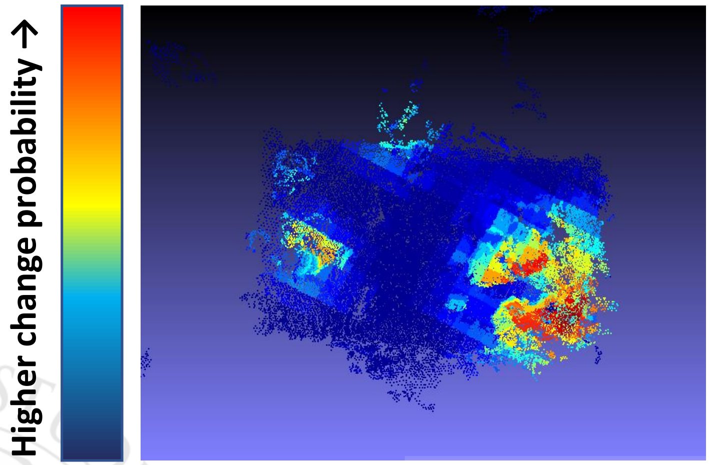
- The volume V is split into 27 sub-volumes v_i
- Each v_i is labeled as *active* if it includes at least one 3D point, otherwise it is labeled as *inactive*
- If more than γ sub-volumes have different labels between T0 and T1, the criterion is satisfied



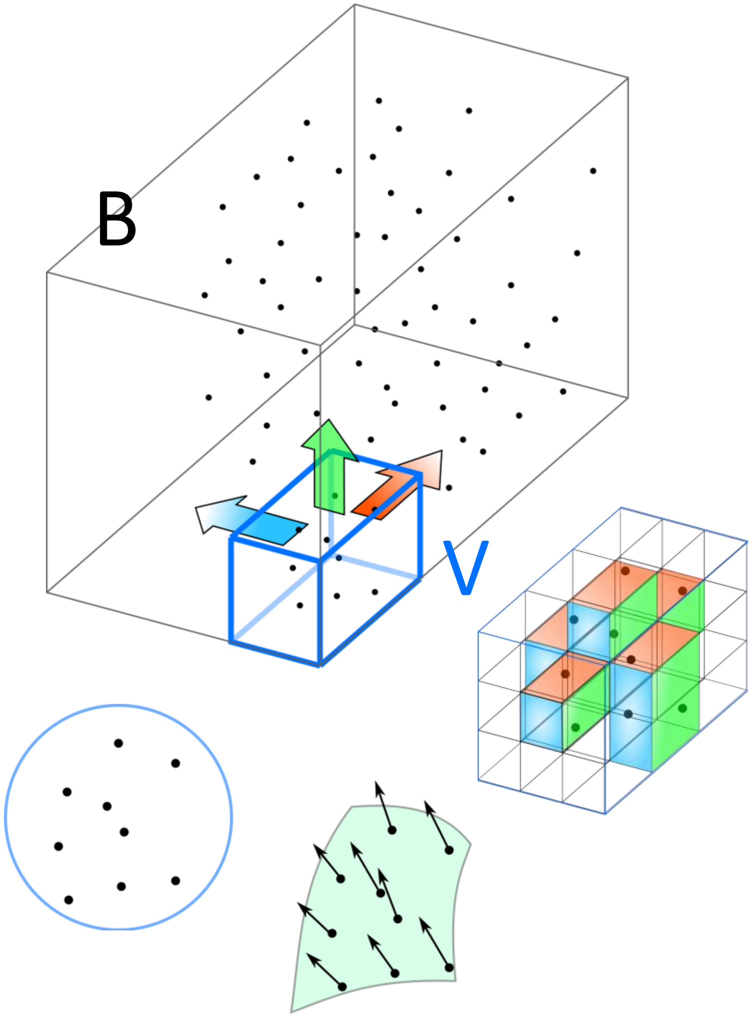
Structural change detection

Majority vote

- If two or more criteria are satisfied, a **token** is given to all the 3D points in V
- Once V has covered the whole B, the number of tokens received by a 3D point can be considered as a **change score**



3D heat-map of change



Experimental results

- Change detection was tested using **two pairs of image sequences** acquired in our lab:
  - **Test 1** – Sequence A, 22 images. Sequence B, 30 images
  - **Test 2** – Sequence C, 22 images. Sequence D, 18 images

- Images were obtained using a (non calibrated) smartphone, with a resolution of 640x480.

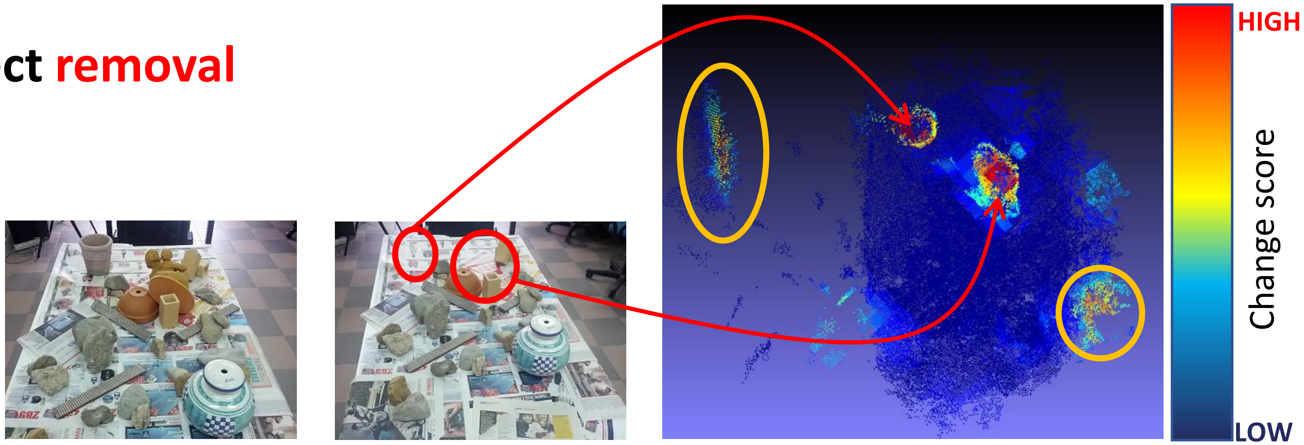
Experimental results

- **Test 1 – Object removal:** the statue and the jar have been removed in the second sequence

- **Test 2: Object insertion / displacement:** insert two cylindrical cans, lay down the statue, change position for the jar, the rocks, and the bricks

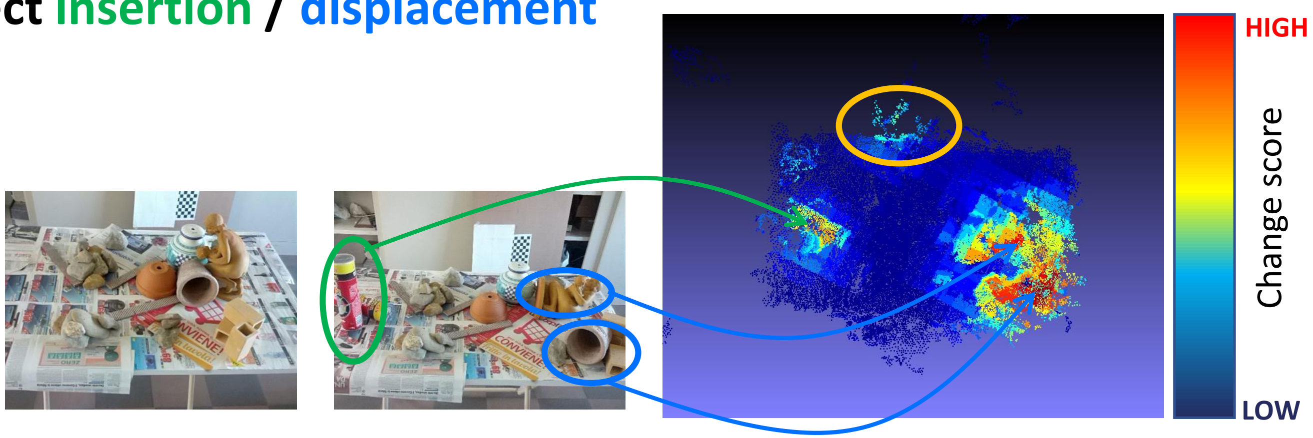
Experimental results: qualitative

- **Test 1: Object removal**



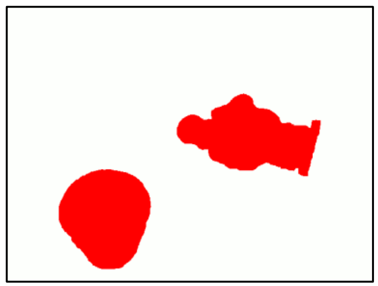
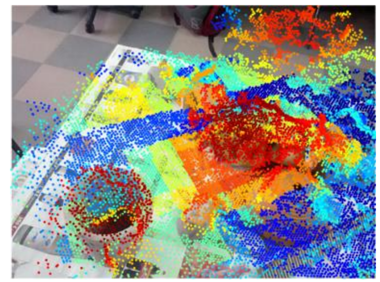
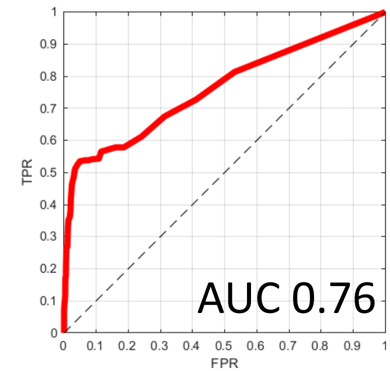
False Positives

- **Test 2: Object insertion / displacement**

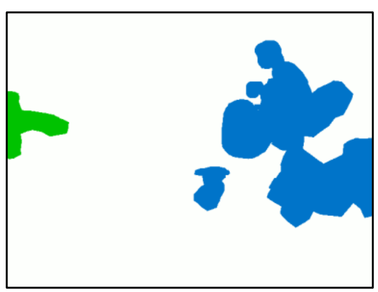
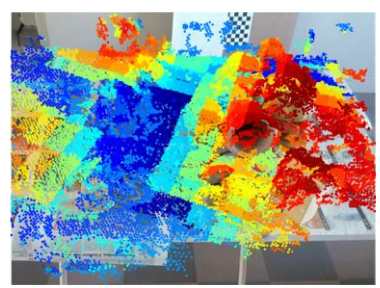
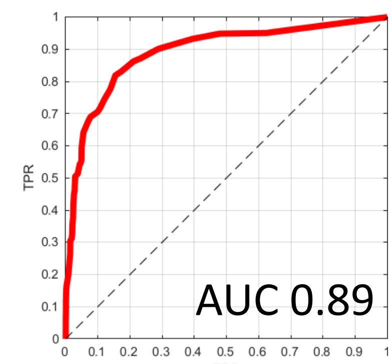


Experimental results: quantitative

- **Test 1: Object removal**

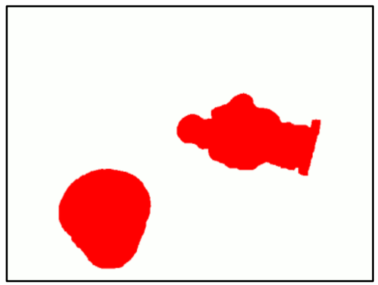
- **Test 2: Object insertion / displacement**

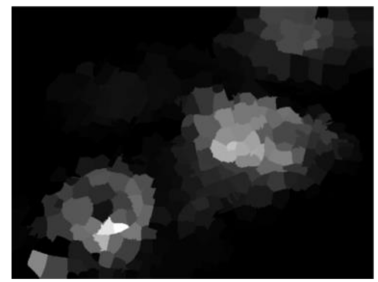





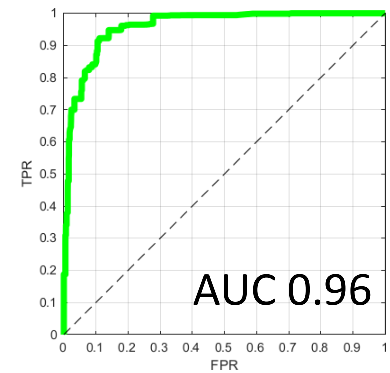
Input | Ground truth | Detection | ROC

Experimental results: quantitative

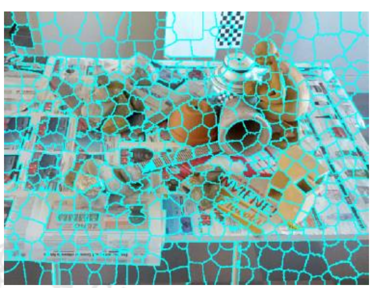
- **Test 1: Object removal**

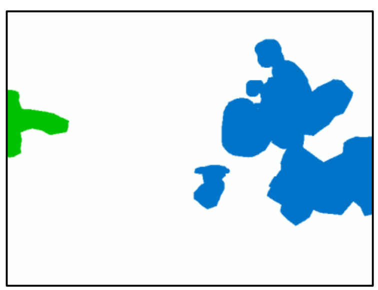


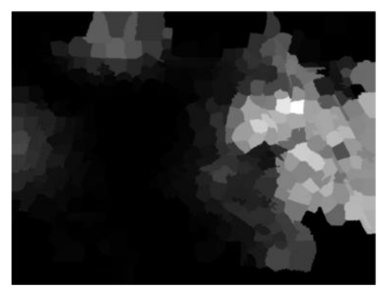


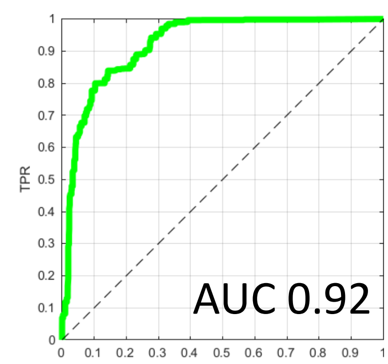


- **Test 2: Object insertion / displacement**









Input with Superpixel Segmentation | Ground truth | Detection | ROC

Conclusions and Future Work

- We proposed a vision based method to **detect structural changes** by comparing 3D reconstructions
- The method is based on 3D model registration, followed by change detection based on three distinct criteria
- Preliminary tests on laboratory scenes show **good results,** with **AUCs greater than 0.90**

- Future work will address:
  - Replacing photometric registration (slow) with direct 3D registration
  - Automatic selection of parameters based on scene scale and image resolution