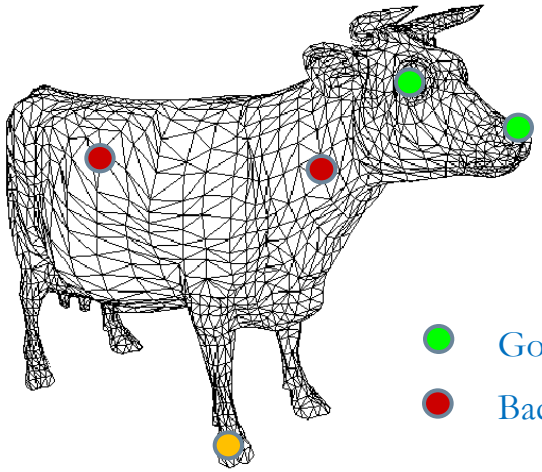
A Performance Evaluation of 3D Keypoint Detectors

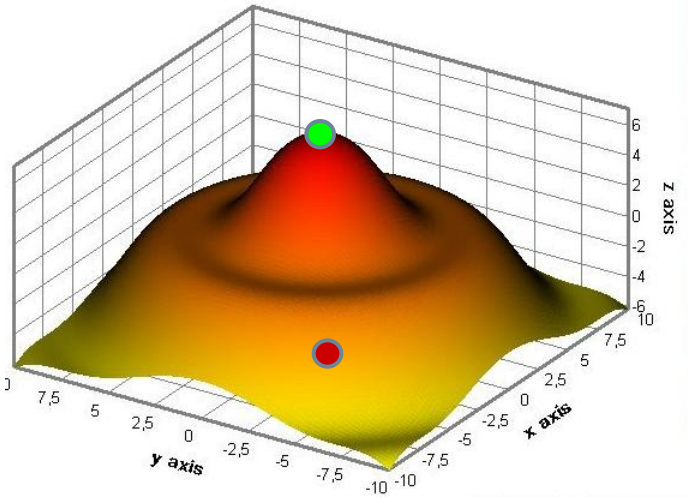
Samuele Salti, Federico Tombari, Luigi Di Stefano

CVLab – University of Bologna, Italy

3D keypoints

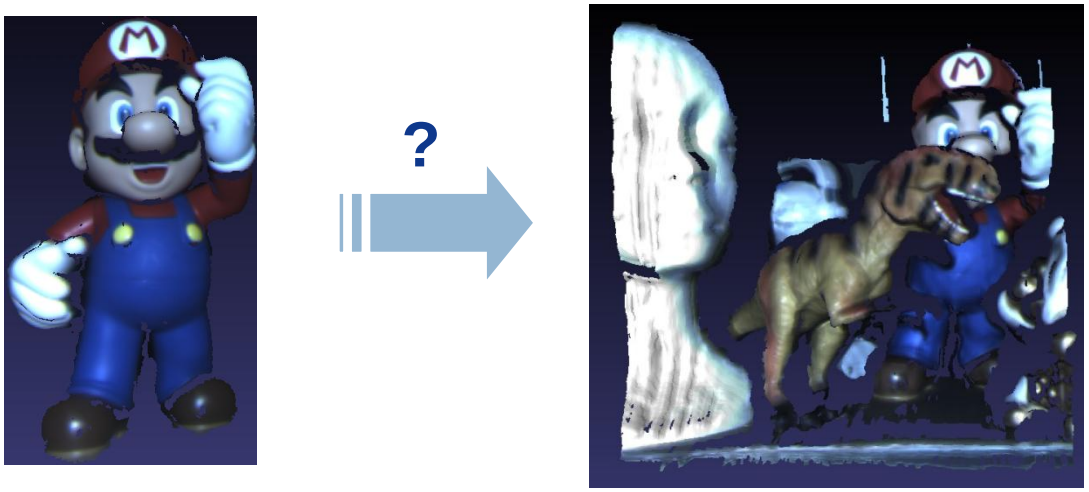
- Lots of 3D applications rely on **3D keypoints** (or features)
  - **Distinctive**, i.e. suitable for effective description and matching
  - **Repeatable** with respect to point-of-view variations





3D keypoints for object recognition

- To be useful in an object recognition scenario 3D keypoints must be also repeatable in presence of o**cclusions, clutter, noise**
- Several techniques proposed recently but we don't know which one performs better



3D keypoints for object recognition

- Several recent proposals, two main categories
  - **Fixed scale detectors**
    - Local Surface Patches (LSP) [Chen & Bhanu, PRL07]
    - Intrinsic Shape Signatures (ISS) [Zhong, 3DRR09]
    - KeyPoint Quality (KPQ) [Mian et al., IJCV10]
  - **Scale-invariant detectors**
    - Several proposal working only on range maps (e.g. [Novatnack & Nishino, ECCV08], [Akagunduz & Ulusoy, ICCV07])
    - MeshDoG [Zaharescu et al., CVPR09]
    - KPQ Scale Invariant (KPQ-SI) [Mian et al., IJCV10]
    - Laplace-Beltrami Scale-Space (LBSS) [Unnikrishnan & Hebert, S3D08]
    - Heat Kernel Signature (HKS) [Sun et al., SGP09]
    - 3D-SURF [Knopp et. al, ECCV10]

Methodology

- We focused on **repeatability**
  - **Absolute repeatability**: the number of repeatable keypoints of a model in a scene
  - **Relative repeatability**: the percentage of repeatable keypoints out of the number of model keypoints.
- A keypoint is said to be repeatable if [Mikolajczyk et al., IJCV05]

$$\left\| R_{ms} k_m^i + t_{ms} - k_s^j \right\| < \varepsilon$$

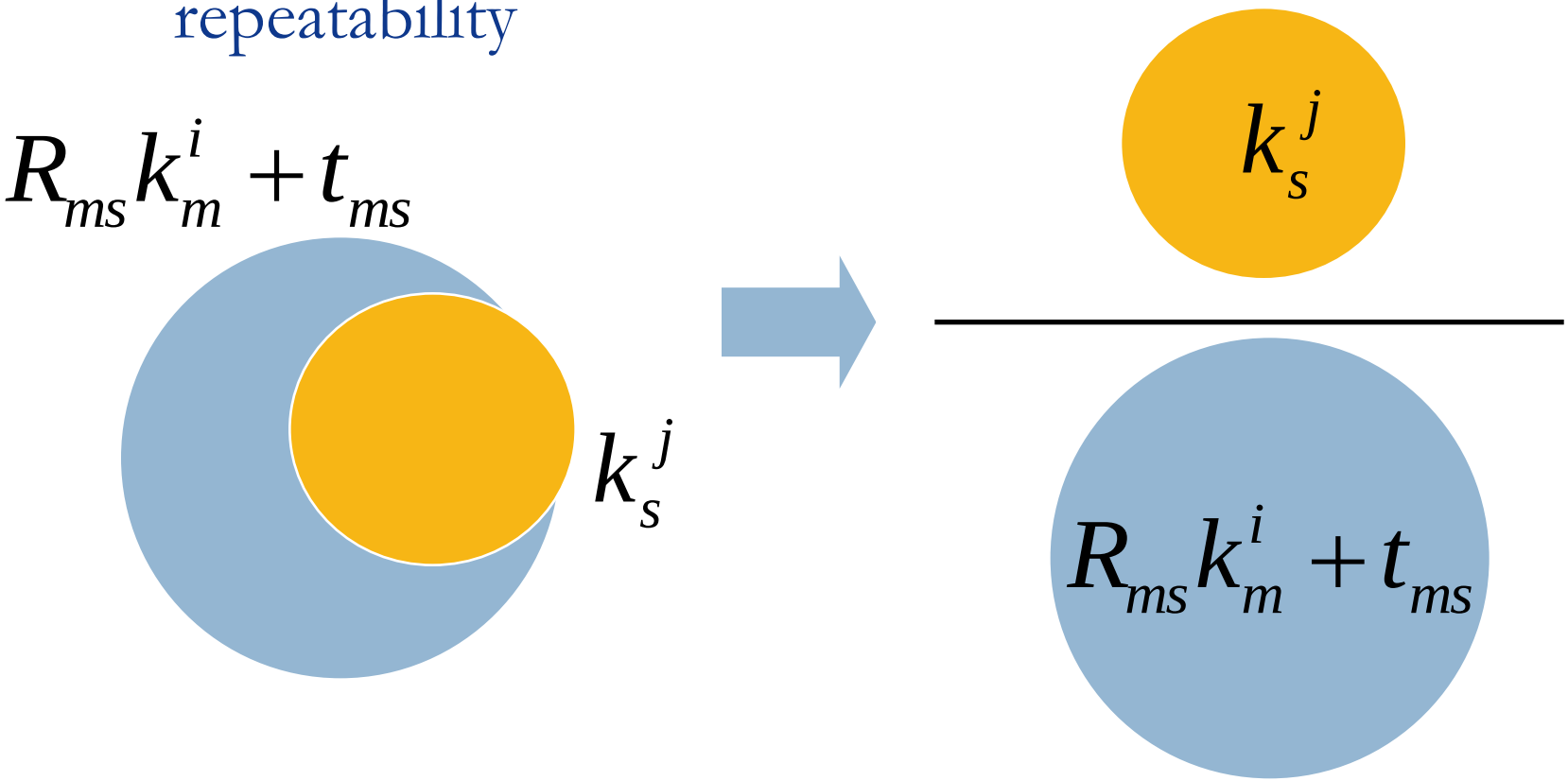
Model keypoint rotated and translated in the scene

Scene keypoint

Repeatability threshold

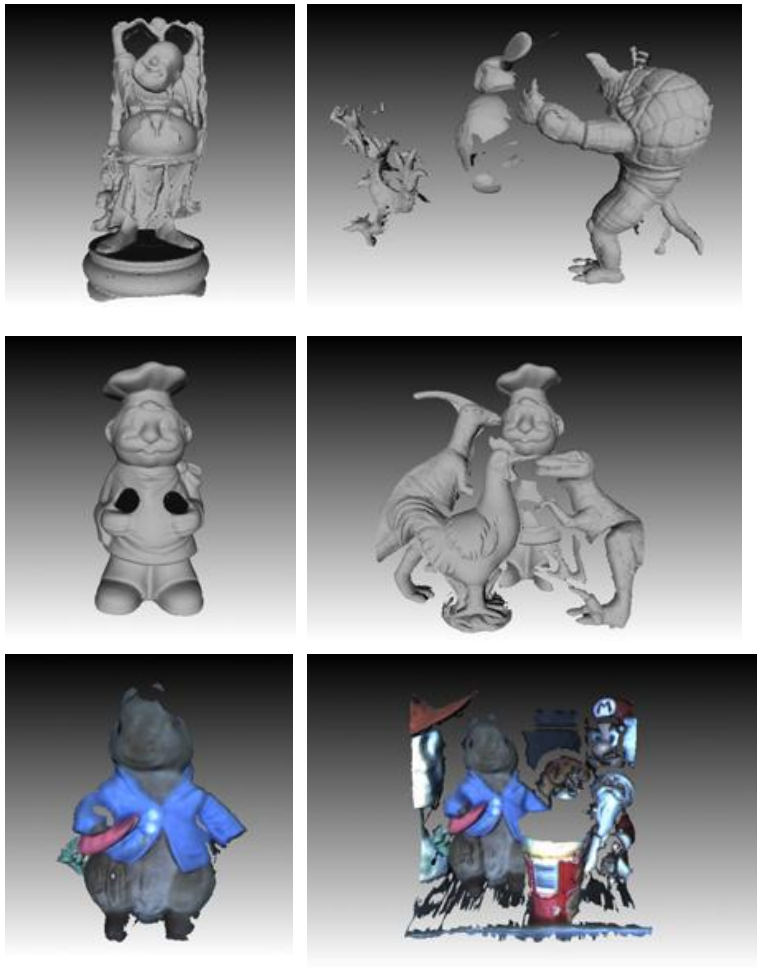
Scale repeatability

- For scale invariant detectors we also consider scale repeatability



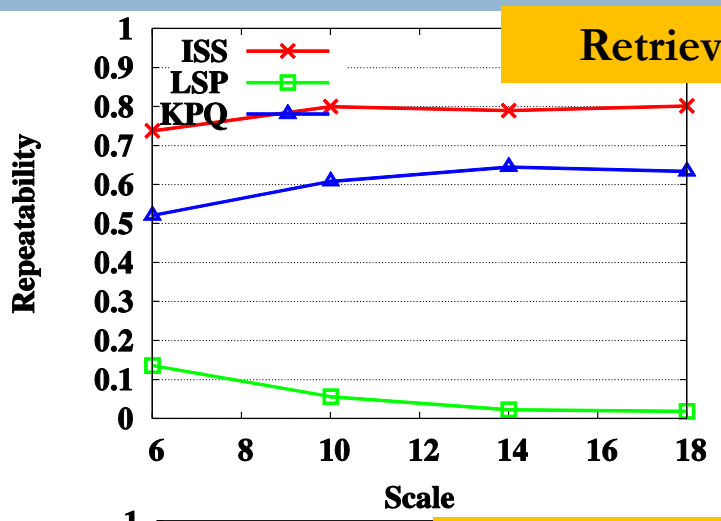
Datasets

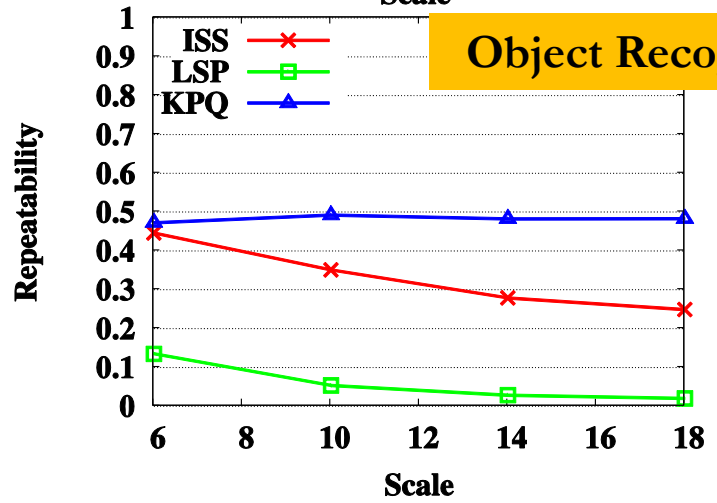
- 2 “Synthetic” datasets created with models from the Stanford Repository
  - 1 for Retrieval
  - 1 for Object Recognition
  - Synthetic Gaussian noise
- the laser scanner dataset [Mian et al., IJCV10]
  - Detailed shapes
  - Full 3D models, 2.5D scenes
  - Point density variations
- the SpaceTime stereo dataset [Tombari et al., ECCV10]
  - Smoother, noisier data
  - 2.5D models and scens
  - Same point density



Retrieval vs Object Recognition 1

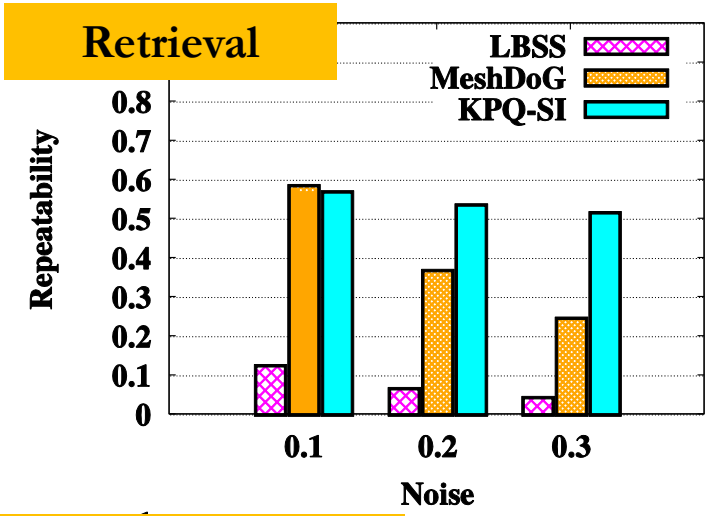
- LSP performs poorly in presence of noise (based on curvature)
- EVD based fixed-scale detectors (i.e. ISS and KPQ) suffer 2.5D views and occlusions
- Presence of clutter favors smaller supports
- KPQ is largely unaffected by noise (smooth surface fitting step)
- Best performer
  - Retrieval: ISS
  - ObjRec: KPQ

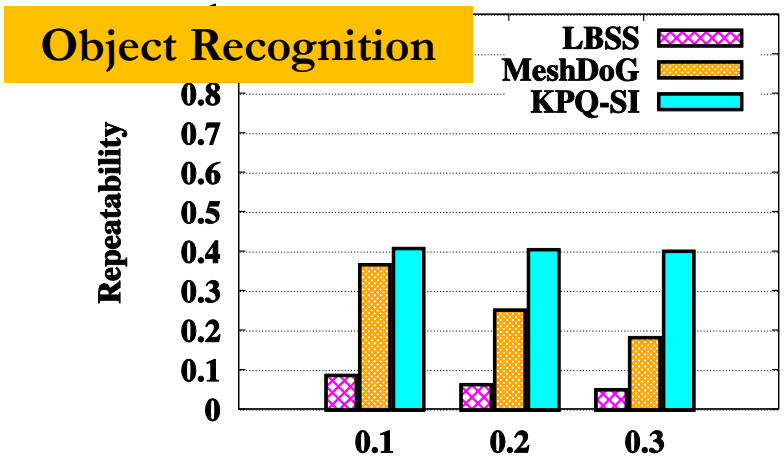




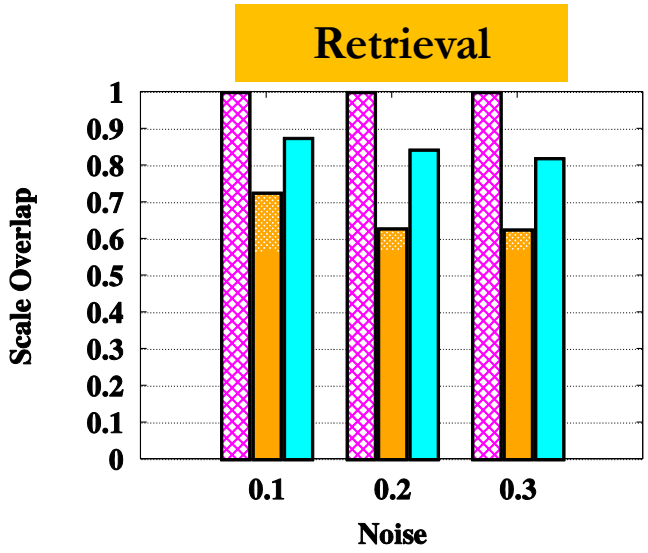
Retrieval vs Object Recognition 2

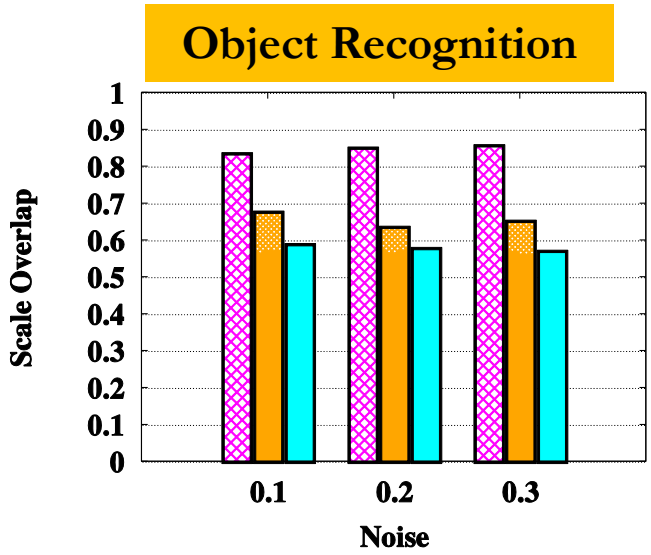
- LBSS exhibits unsatisfactory repeatability
- KPQ-SI is greatly more robust to noise than MeshDoG (smooth surface fitting step vs use of curvatures)
- KPQ and KPQ-SI have similar performance





Scale Repeatability





- LBSS shows superior scale repeatability but poor spatial repeatability.
- MeshDoG suffers point density variations (from Laser Scanner dataset results)

Timing (in seconds)

	Retrieval	Random Views	Laser Scanner	Space Time
LSP	56 ~ 65	31 ~ 100	65 ~ 76	74 ~ 92
ISS	2 ~ 10	2 ~ 7	5 ~ 13	6 ~ 18
KPQ*	266 ~ 493	413 ~ 662	799 ~ 1109	544 ~ 1222
LBSS	1585	461	1148	1397
MeshDoG	198	185	425	469
KPQ-SI*	303	364	634	767

- Average detection time / scene (in seconds)
- ISS is one or more orders of magnitude faster than the other detectors