Fast 3D keypoints detector and descriptor for view-based objects recognition

Ayet SHAIEK and Fabien MOUTARDE

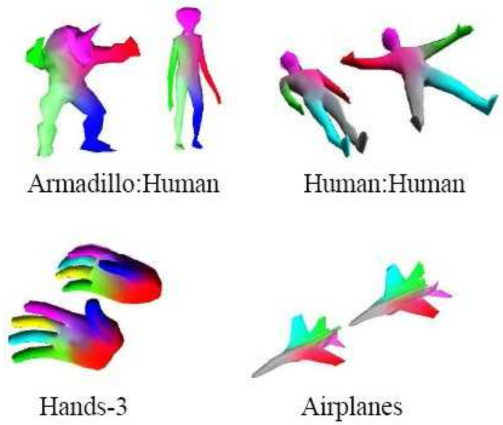
PHD Student

Robotics laboratory (CAOR)
Mines ParisTech- France

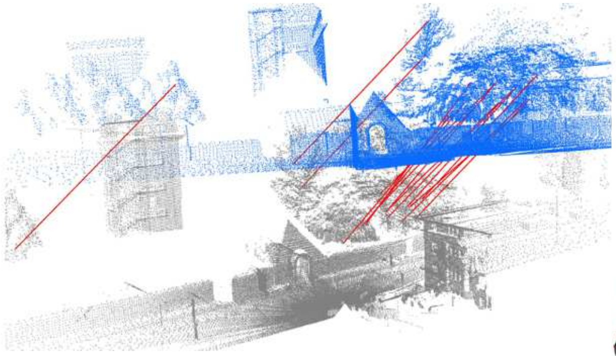
11/11/2012

Overview

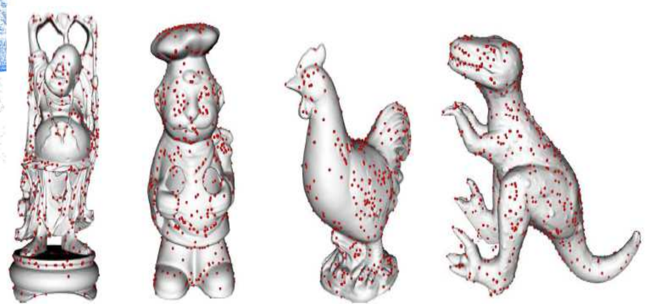
1. Introduction
2. Keypoint Detection
3. Keypoint Description
4. Keypoint Matching
5. Conclusion



3D shape Modeling



Surface registration



Object Recognition

- **Technological advances have made possible the production of reliable and accurate 3D data**
- **For a 3D capture of an object: Physical shape + robustness**
  -> Growing interest for 3D data
- **Processing 3D data presents some issues:**
  - Big amount of data
  - Capture conditions (sensor or scene)

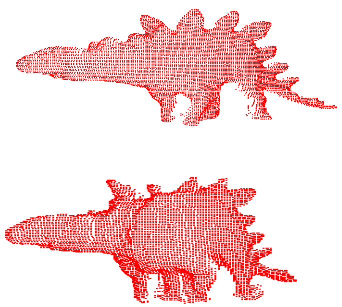
➢ Need compact, invariant and robust representation
-> Assures effectiveness for tasks:

Objectif

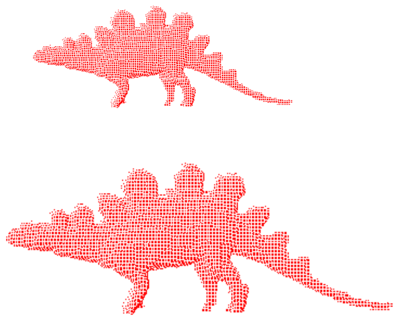
How to characterize the 3D shape for 3D object recognition

Problems related to the real world and sensors:

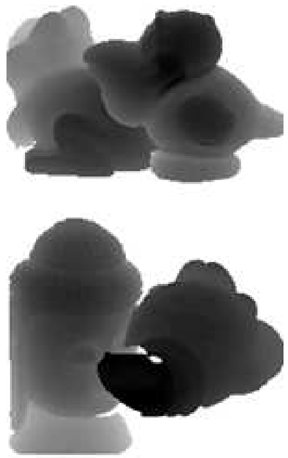
- Invariance to scale, sampling and geometric transformations
- Robustness to noise and occlusions



Different viewpoints



Different scales



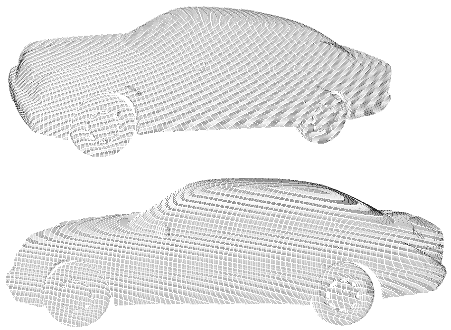
Occlusions

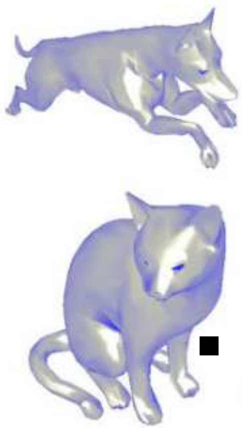
Different input data **formats** in recognition task :

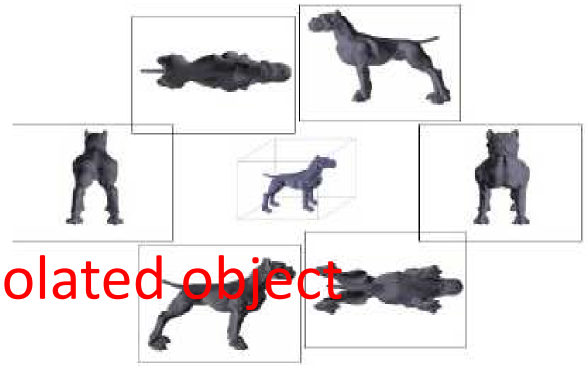
– Point cloud
– 3D Mesh
– **Range image (2.5D)**

❖ Partial views
❖ Whole 3D object

▪ Situated in a whole scene (need segmentation)

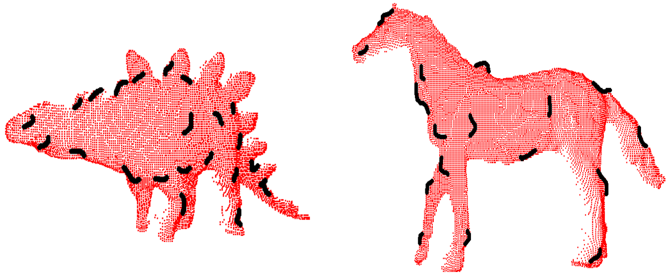






▪ Isolated object

Keypoint detection



- Salient point
  - Repeatable and robust
  - Well localized
  - Fast detection
  - Existing:
    - **3D Surf** detector [Jan et al., 10]
    - **3D Harris** detector [Ivan et al., 10]
    - **Gaussian filters** [Castellani et al.,08]

Largest shape variation: How much do the dominant directions of the surface change locally?

- **Differential geometry: curvatures and surface normals.**
- **Curvature based detectors**

Knopp, Jan and Prasad, Mukta and Willems, Geert and Timofte, Radu and Van Gool, Luc, "Hough transform and 3D SURF for robust three dimensional classification", 589—602,ECCV'10.
SIPIRAN I., BUSTOS B.: A robust 3D interest points detector based on Harris operator. In Proc. Eurographics Workshop on 3D Object Retrieval (2010), Eurographics Association, pp. 7–14.
U. Castellani, M. Cristani, S. Fantoni and V. Murino, « Sparse points matching by combining 3D mesh saliency with statistical descriptors"; EUROGRAPHICS 2008; Vo l u m e 2 7(2008), Number 2.

MINES PARIS ParisTech

Shape index
[Chen et al., 07]

$$SI_p = \frac{1}{2} - \frac{1}{\pi} arctg\left(\frac{k_p^1 + k_p^2}{k_p^1 - k_p^2}\right)$$

where k^1_p and k^2_p are maximum and minimum principal curvatures

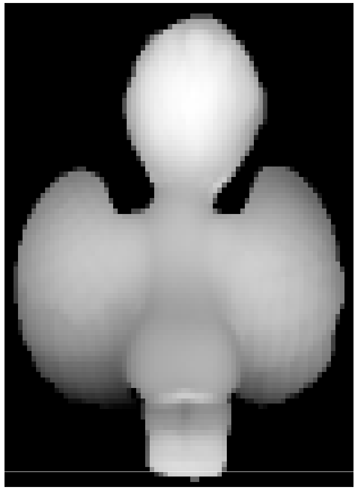
- SI_p = **max** of shape indexes and $SI_p \geq (1+\alpha) * \mu$; (convex surfaces)
- or SI_p = **min** of shape indexes and $SI_p \leq (1- \beta) * \mu$; (concave surfaces)

where μ is the mean shape index over neighbours' values and **$0 \leq \alpha , \beta <= 1$**

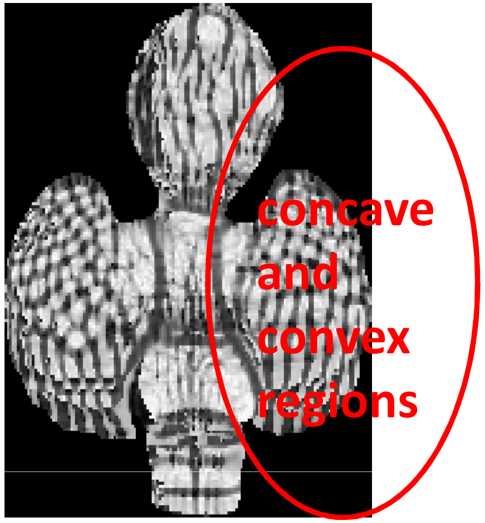
» **Invariance to orientation and scale.**

Chen, H. and Bhanu, B. "3D free-form object recognition in range images using local surface patches," Pattern Recognition Letters, 28(10), 1252-1262 (2007).

Depth Image



SI image



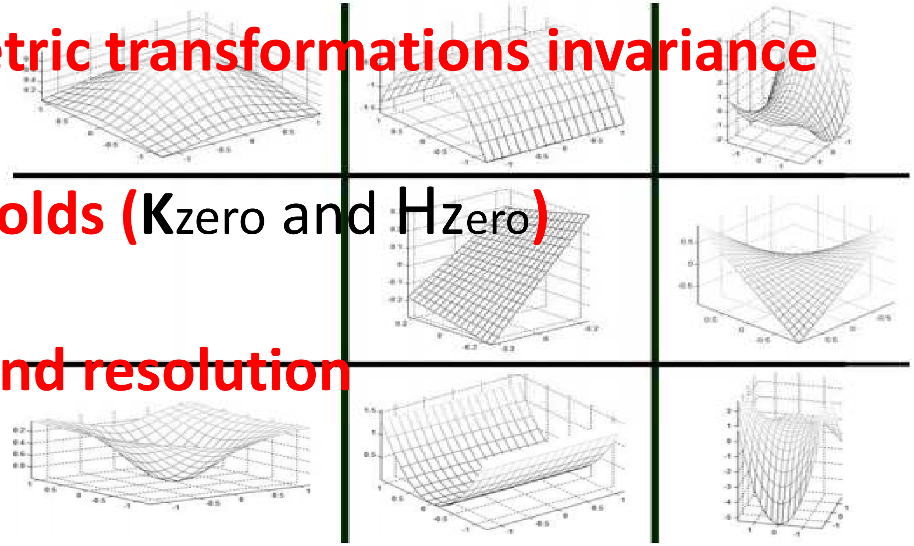
•In SI image: Brighter pixels correspond to the **convex** surfaces (i.e domes and ridge) and darker ones represent **concave** surfaces (rut or cup).

•**Details** are accentuated on SI image: SI **comprises the understanding of the neighbor geometry** whereas range image pixel only indicates its depth ->more **informative**

Shape classification with HK space

MINES PARIS ParisTech

	K>0	K=0	K<0
H<0	Convex [1] (*Elliptic or Umbilic*)	Ridge [2] (Convex *Parabolic*)	Saddle Ridge [3] (*Hyperbolic*)
H=0	(*Not possible*)[4]	Plane [5]	Minimal [6]
H>0	Concave [7] (*Elliptic or Umbilic*)	Valley [8] (*Concave Parabolic*)	Saddle Valley [9] (*Hyperbolic*)



-Orientation … etric transformations invariance

-Depend on zero-thresholds (K_{zero} and H_{zero})

-Not invariant to scale and resolution

$$T_p = 1 + 3(1 + \mathrm{sgn}_{\epsilon_H}(H)) + (1 - \mathrm{sgn}_{\epsilon_K}(K)),$$

$$\mathrm{sgn}_{\epsilon_x}(X) = \begin{cases} +1 & \text{if } X > \epsilon_x, \\ 0 & \text{if } |X| \leqslant \epsilon_x, \\ -1 & \text{if } X < \epsilon_x. \end{cases}$$

Mean curvature H	Gaussian curvature K		
	$K > 0$	$K = 0$	$K < 0$
$H < 0$	Peak $T_p = 1$	Ridge $T_p = 2$	Saddle ridge $T_p = 3$
$H = 0$	None $T_p = 4$	Flat $T_p = 5$	Minimal $T_p = 6$
$H > 0$	Pit $T_p = 7$	Valley $T_p = 8$	Saddle valley $T_p = 9$

pits, peak and saddle surfaces.

P. J. Besl and R. C. Jain (1988). Segmentation Through Variable-Order Surface Fitting, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 2, pp. 167-192.

Shape classification with SC shape

Shape index (**S**) + curvedness (**C**)[Koenderink 1992].

$$S = \frac{2}{\pi} \cdot \arctan\left(\frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2}\right) \quad (\kappa_1 > \kappa_2)$$

- **defines the shape**

$$C = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}}$$

- **square-root of the deviation from flatness**

Convex (Elliptic) (1)	$S \in [+5/8, 1] \cap C > C_{zero}$
Convex (Parabolic) (2)	$S \in [+3/8, +5/8] \cap C > C_{zero}$
Saddle Ridge (3)	$S \in [+3/16, +3/8] \cap C > C_{zero}$
Planar	$C < C_{zero}$
Hyperbola (6)	$S \in [-3/16, +3/16] \cap C > C_{zero}$
Concave (Elliptic) (7)	$S \in [-1, -5/8] \cap C > C_{zero}$
Concave (Parabolic) (8)	$S \in [-5/8, -3/8] \cap C > C_{zero}$
Saddle Valley (6)	$S \in [-3/16, +3/16] \cap C > C_{zero}$

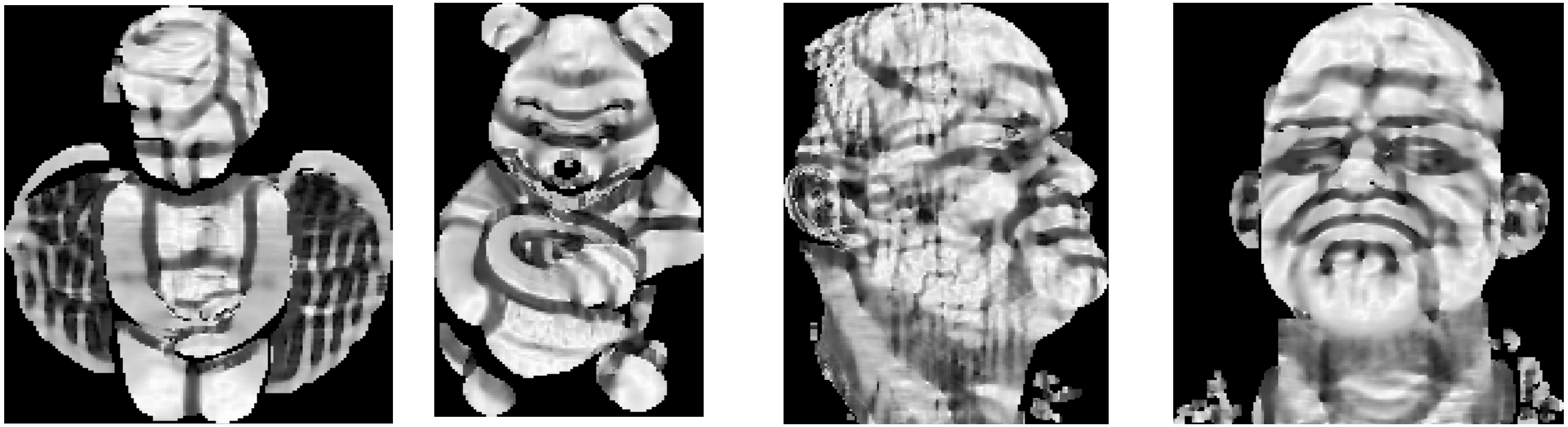
Characteristics of this representation:

- **S:** Orientation, geometric transformations , scale and resolution invariant
- **C** : Orientation and geometric transformations invariant but depends on C_{zero}

Dome, cup and saddle Surfaces

J. Koenderink and A. J. Doorn (1992). Surface shape and curvature scale, Image Vis. Comput., vol. 10, no. 8, pp. 557–565.

Shape Index Images



Curvedness Images

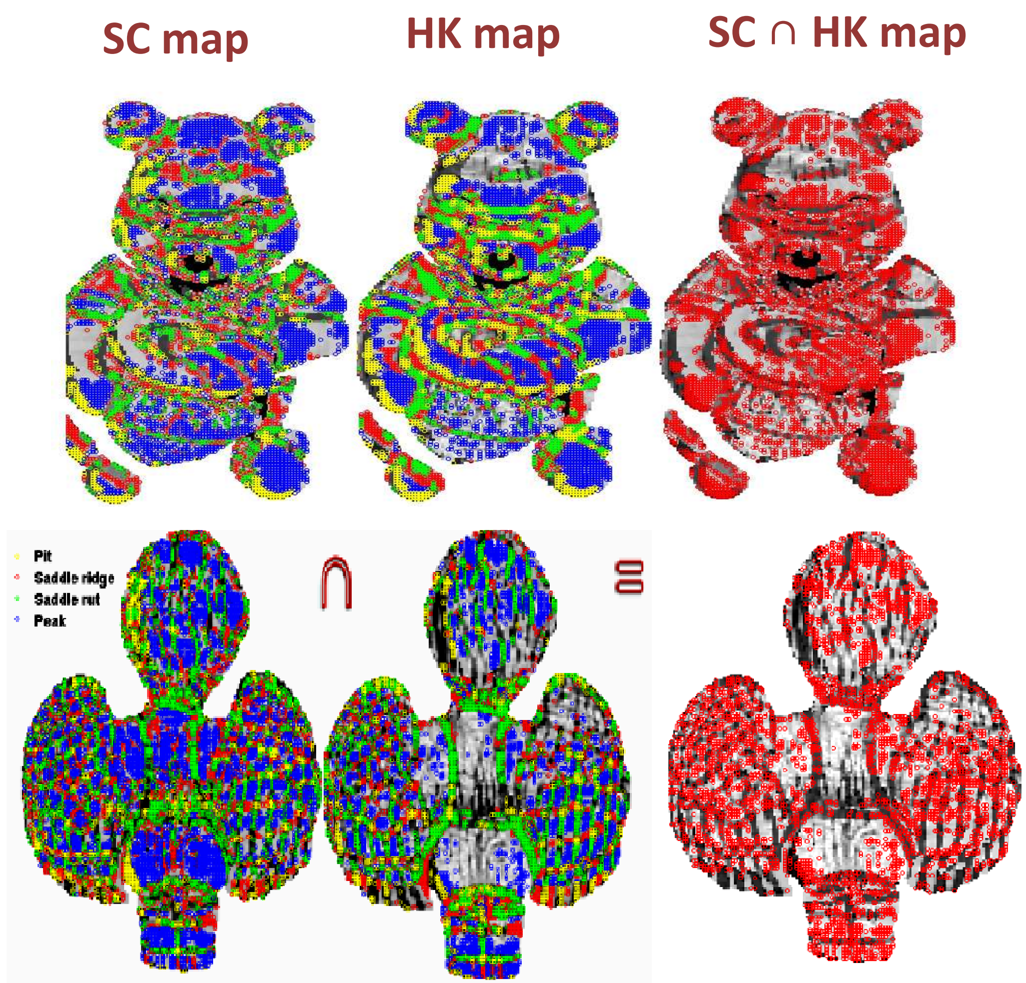
•**Informative regions are visually strengthen and shape details are accentuated**

Principle contribution: Classification on the coupled space HK&SC

Convex (Elliptic) (1)	$H<-\|H_{zero}\| \cap K>+\|K_{zero}\| \cap S \in [+5/8,1] \cap C>C_{zero}$
Convex (Parabolic) (2)	$H<-\|H_{zero}\| \cap K<\|K_{zero}\| \cap S \in [+3/8,+5/8] \cap C>C_{zero}$
Saddle Ridge (3)	$H<-\|H_{zero}\| \cap K<-\|K_{zero}\| \cap S \in [+3/16,+3/8] \cap C>C_{zero}$
Planar (5)	$H<\|H_{zero}\| \cap K<\|K_{zero}\| \cap C<C_{zero}$
Hyperbola (6)	$H<\|H_{zero}\| \cap K<-\|K_{zero}\| \cap S \in [-3/16,+3/16] \cap C>C_{zero}$
Concave (Elliptic) (7)	$H>+\|H_{zero}\| \cap K>+\|K_{zero}\| \cap S \in [-1,-5/8] \cap C>C_{zero}$
Concave (Parabolic) (8)	$H>+\|H_{zero}\| \cap K<\|K_{zero}\| \cap S \in [-5/8,-3/8] \cap C>C_{zero}$
Saddle Valley (6)	$H>+\|H_{zero}\| \cap K<-\|K_{zero}\| \cap S \in [-3/16,+3/16] \cap C>C_{zero}$

Not all regions will be classified: only the common ones

- **HK∩ SC: more reliable result**



Legend

- Saddle ridge
- concave
- Saddle valley
- convex

•The combination process assures **better saliency**.

•**Grouping** extracted keypoints in a connected components and select the most informative ones by **ranking** points according to a confidence value on C value.

- **reduce** the number of selected keypoints

▪Steps of the intersection algorithm SC_HK_Connex:

- For each point, extract **neighborhood** N_p (belong to **spherical support** with radius proportional to the surrounding box of the shape)

- Compute measures of **saliency** (**SC and HK**) and extract Keypoints according to the intersection

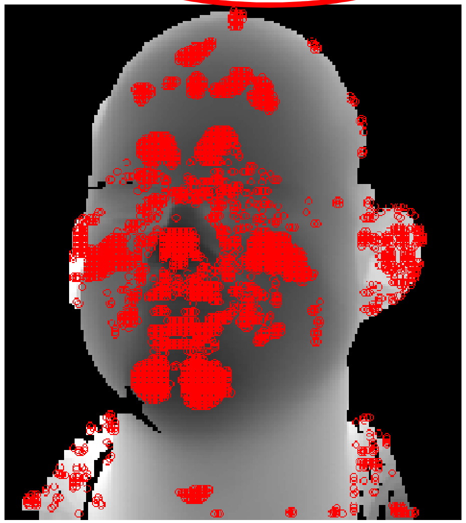
- Compute a **confidence** value for the keypoints (based on **C**)

- Group detected keypoints with **connected components** (label = pair of types)

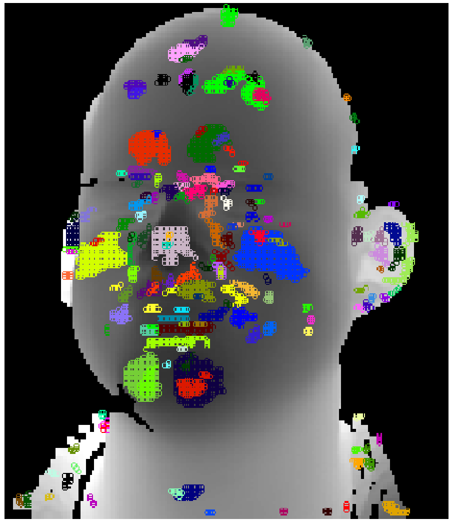
- Extract **largest** connected components and keypoints with the **highest confidence** are taken in each components.

Result of SC_HK_Connex detector

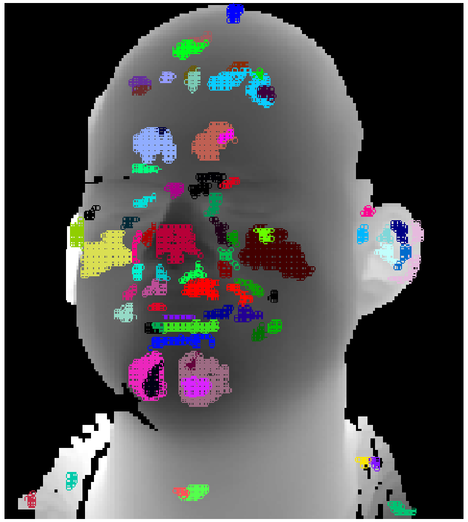
Keypoints without connected components
3118 Kpts

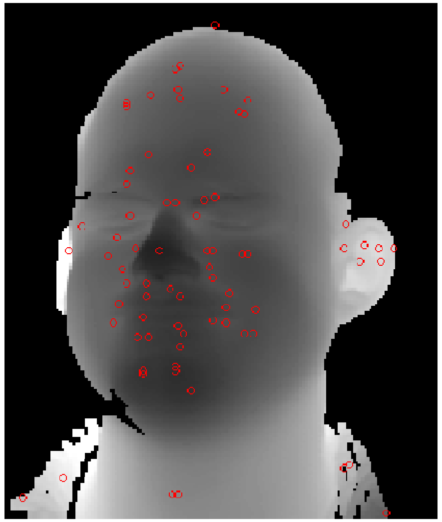


Connected components



Connected components with size > 8 points





Keypoints with connected components

Kpt selection with maximal value of (C)
75 Kpts

Descriptors

- Existing:
  - 3D SURF descriptor [Jan et al., 10]
  - SHOT (Signature of Histograms of OrienTations) and CSHOT descriptor [Tombari and al. ,11]

➢SHOT: normal estimation based on the Eigenvalue Decomposition of a novel scatter matrix defined by a weighted linear combination of neighbor point distances: definition of a Robust Reference Frame (RF)

✓Surface normal information is invariant to sampling density, scale and viewpoints.

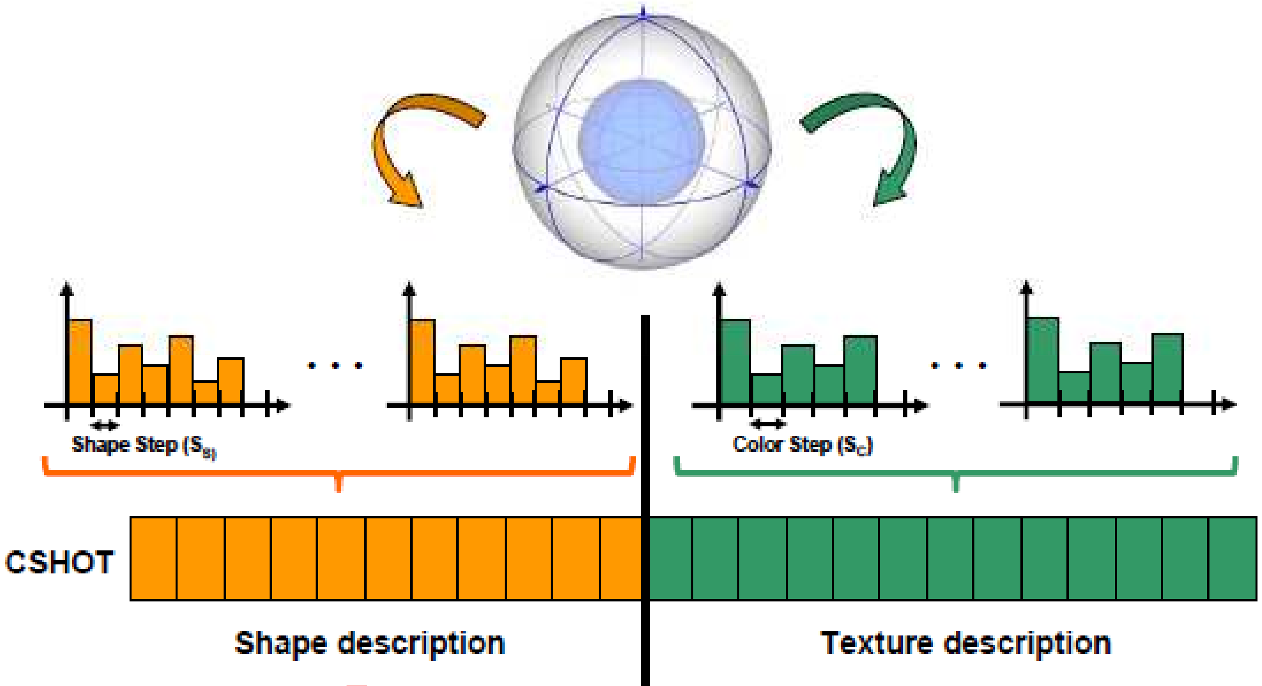
❖CSHOT: add of texture information

✓Succeeds to form more robust and descriptive signature

•F. Tombari, S. Salti, L. Di Stefano ; *"A combined texture-shape descriptor for enhanced 3D feature matching";* IEEE International Conference on Image Processing (**ICIP 201**1), September 11-14, Brussels, Belgium.

•**Proposed descriptor**

•**IndSHOT Descriptor:** joining **shape index histogram** and **histogram of angles** between normals



IndSHOT =

•**histogram of angles between normals**

•**histogram of Index Shape**

Recognition Task

- **Matching:** validating the proposed detector and descriptor using **a view matching approach**

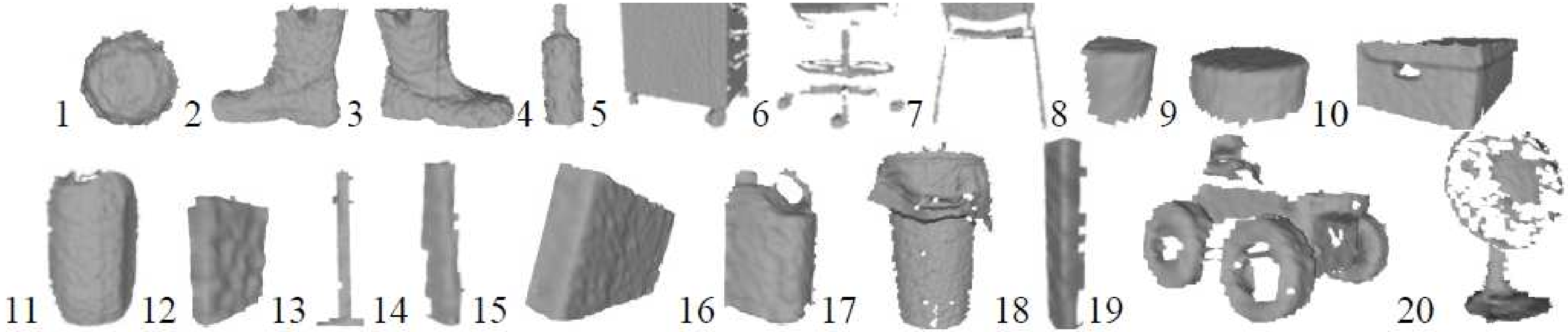
➢**Similarity measure:** Given a test object, we compute a measure of similarity between descriptors extracted on the test view and those of the models in database.

✓For each keypoint , we search for the best nearest neighbor keypoint in the database: Euclidean Distance
✓KDtree to speed up the matching process

➢ **Filtering** the potential corresponding keypoint pairs based on **geometric constraints** :

✓The closest couple of features in term of 3D coordinates distance is the more likely to form a **consistent correspondence.**

Experimental Results



The 20 objects of our lab-Dataset (4 to 15 views/object)

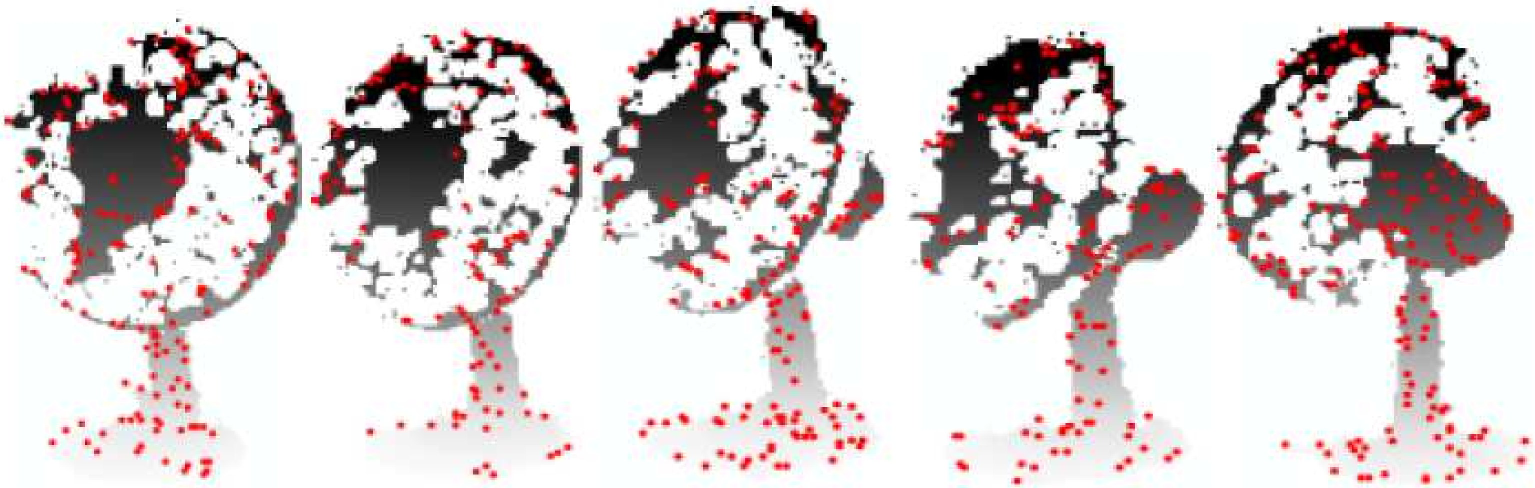
Examples of objects from the RGB-D Dataset *(46 common household objects, 25 views/object)

- Evaluate our detector and descriptor in terms of recognition rate

* http://www.cs.washington.edu/rgbd-dataset/

Keypoint detection

SI detector **SC_HK** **SC_HK_Connex**



Detected keypoint on fan model, with SC_HK_connex, in view angle variation

Recognition rate

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	100																			
2		96																		
3			100																	
4	4			92																
5					96															
6	4					56			12			12						4		
7			4				84								4					
8								100												
9			8			16			56									4		
10										100										
11											96									
12												96								
13													100							
14														100						
15															92					
16																100				
17																8	84			
18				4												16		72		
19																			100	
20																				100

Confusion matrix for the result of SC_HK_connex+IndSHOT methods on Lab-dataset

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
1																																														
2																																											4			
3																																														
4	4			92																				4																						
5																						4																								
6	4					56			12			12						4									4										4					4				
7			4				84								4																												4		4	
8																																														
9			8			16			56									4																			4							12		
10																																														
11																																														4
12																												4																		
13																																														
14																																														
15															92							4	4																							
16																																														
17																8	84																													8
18				4												16		72															4													4
19																																														
20																																														
21					4									4							92																									
22					28															4		68																								
23										4																																				
24																																				4										
25	4	8																							88																					
26																								12		84					4															
27						4						8															84																	4		
28																																				4										
29						4																4							84						4						4					
30														4																																
31																															92					4							4			
32																										8						92														
33																																														
34																																														
35																																														
36	4																							4												88							4			
37																																														
38			4																								4											92								
39																4	4																						76							
40																																														
41																																														
42			4			4																																				92				
43																										8										8							84			
44						4																					8																	88		
45																																					4									
46																	4																													

Confusion matrix for the result of SC_HK_connex+IndSHOTmethod on RGB-D object dataset

Recognition rates

	IndSHOT	SHOT	CSHOT
SC_HK	82.5%	62,5%	67.5%

On Lab-dataset

	IndSHOT	SHOT	CSHOT
SID	89.06%	70,75%	77.77%
SC_HK	91.12%	75,28%	82.14%

RGB-D object dataset

•**The overall recognition rate is quite promising for the SC_HK_Connex**

➢**Computation time** for our recognition process (detection + description + features matching) is quite low (~0.7s) for 100 Keypoints

Conclusion

MINES PARIS ParisTech

- ✓ A new 3D detectors based on **combination** of surface classification **criteria**
- ✓ Combined descriptor IndSHOT: **shape index and angles between normals**

- ✓ **Recognition** rate of 91.12% on public Kinect dataset (*RGB-D object dataset)*

- ➢ Evaluation on other databases
- ➢ Evaluation of scale invariance and noise robustness

Thank you

for your attention