Adaptive Fusion-based 3D Keypoint Detection for RGB Point Clouds

Muhammad Zafar Iqbal (mzafar.Iqbal@tum.de),

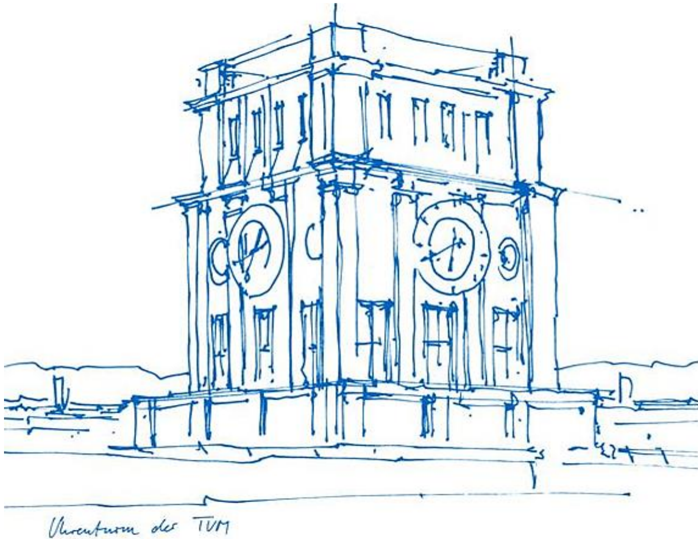
Dmytro Bobkov (dmytro.bobkov@tum.de),

Eckehard Steinbach (eckehard.steinbach@tum.de)

Chair of Media Technology,
Technical University of Munich,
Germany

IEEE International Conference on Image Processing (ICIP), 2019

25th September 2019



Overview

- Introduction
- Contributions
  - Geometric salient points detection
  - RGB salient points detection
  - Fusion of salient points
- Experimental Evaluation
- Conclusion

Why 3D Keypoint Detection?

- 3D object recognition, retrieval and matching
  - Global descriptors
    - Require pre-segmented objects
    - Performance degrades with occlusion and clutter
    - Computationally efficient
  - **Local descriptors**
    - No pre-segmentation is required
    - Robust against occlusion and clutter
    - Computationally expensive
    - Require ***keypoint detection***

3D Keypoint Detectors

- Fixed scale
  - Local Surface Patches (LSP) [1]
  - Heat Kernel Signature (HKS) [2]
  - Intrinsic Shape Signature (ISS) [3]
  - Keypoint Quality (KPQ) [4]
  - Harris 3D [5]
  - Harris 6D [6]
  - Histogram of Normal Orientation (HoNO) [7]
- **Adaptive scale**
  - Laplace-Beltrami Scale Space (LBSS) [8]
  - Salient Point (SP) [9]
  - Mesh Difference of Gaussian (Mesh-DoG) [10]
  - Keypoint Quality Adaptive Scale (KPQ-AS) [4]

Adaptive scale keypoint detector are more robust and scale invariant

The detectors shown in black use the geometrical structure for keypoint detection while the keypoint detectors shown in blue use either RGB/geometry or use both RGB and geometry information

Existing Algorithms Drawbacks

- Existing 3D keypoint detectors have limited performance
  - Low repeatability
  - Low distinctiveness
- Most of the existing 3D detectors either use
  - Geometric information
  - Photometric appearance
- More efficient detectors are required in terms of repeatability and distinctiveness, and use both available information

Proposed Technique

RGB point cloud

$$P = \{p_i \in \mathbb{R}^6 : i = 1, .., N\}$$

Mario of the SHOT-SpaceTime dataset [1]

Saliency measure from geometric structure motivated by [3]

$$P^{cor} = [P_x, P_y, P_z]$$

$$P^{rgb} = [P_r, P_g, P_b]$$

Saliency measure from photometric appearance motivated by [10]

Fusion of salient points

Saliency Measure from Geometric Structure

- Fixed scale EVD of covariance-based detector is proposed by Zhong [3]

Point cloud coordinates

$$P^{cor} = [P_x, P_y, P_z]$$

$$Cov_i^a = \frac{1}{M}\sum_{j=1}^{M}(p_j^{cor} - p_i^{cor})(p_j^{cor} - p_i^{cor})^T$$

$$NN_i^a = \{p_j^{cor} : |p_i^{cor} - p_j^{cor}| < r^a\}$$

$$R = \{r^a \in \mathbb{R}^+ : a = 1, \cdots, A\}$$

"A" represents number of support radii

Compute covariance matrices at predefined support radii

Compute eigenvalue decomposition of covariance matrices

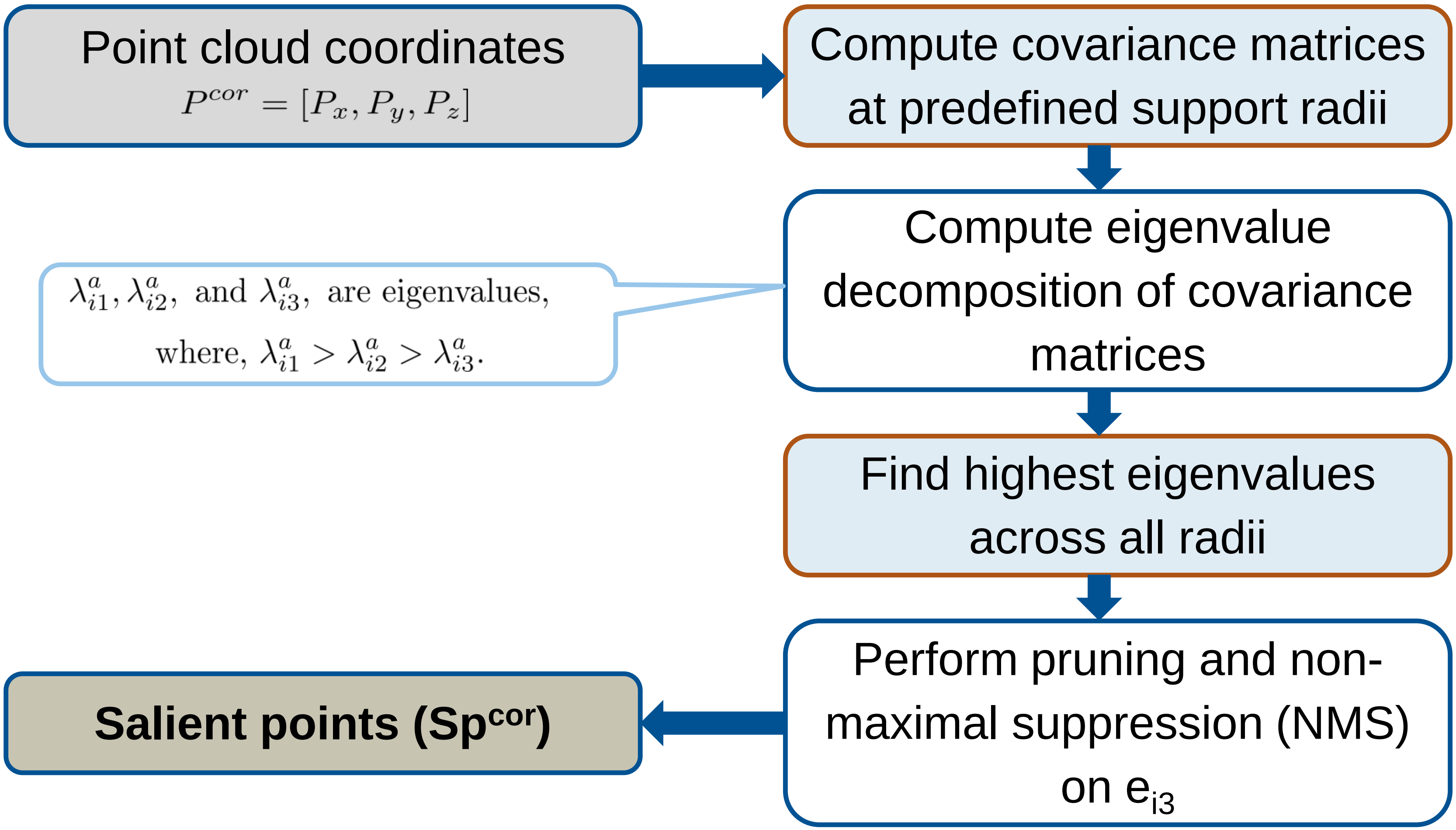
Find highest eigenvalues across all radii

Perform pruning and non-maximal suppression (NMS) on e_{i3}

Salient points (Sp^{cor})

Saliency Measure from Geometric Structure

- Fixed scale EVD of covariance-based detector is proposed by Zhong [3]



Saliency Measure from Geometric Structure

- Fixed scale EVD of covariance-based detector is proposed by Zhong [3]

Point cloud coordinates
$P^{cor} = [P_x, P_y, P_z]$

Compute covariance matrices at predefined support radii

Compute eigenvalue decomposition of covariance matrices

Find highest eigenvalues across all radii

$$e_{it} = max(\{\lambda_{it}^{a} : a = 1, \cdots, A\}),$$
$$\text{where } t = 1, 2, 3.$$

Perform pruning and non-maximal suppression (NMS) on e_{i3}

Salient points (Sp^{cor})

Saliency Measure from Geometric Structure

- Fixed scale EVD of covariance-based detector is proposed by Zhong [3]

Point cloud coordinates
$P^{cor} = [P_x, P_y, P_z]$

→ Compute covariance matrices at predefined support radii

↓ Compute eigenvalue decomposition of covariance matrices

↓ Find highest eigenvalues across all radii

↓ Perform pruning and non-maximal suppression (NMS) on e_{i3}

if $\left(\left(\frac{e_{i2}}{e_{i1}} > Th_{12}\right) \wedge \left(\frac{e_{i3}}{e_{i2}} > Th_{23}\right) \wedge (e_{i3} < Th_3)\right)$ **then**
  prune = true
else
  prune = false
end if
Perform NMS at support radius of r^{nms}.

← **Salient points (Sp^{cor})**

Saliency Measure from Photometric Appearance

- Adaptive DoG-based detector is proposed by Zaharescu et al. [10]

Photometric appearance
$$P^{rgb} = [P_r, P_g, P_b].$$

$$I_i = 0.299 * R_i + 0.587 * G_i + 0.114 * B_i$$

Compute intensity from RGB channel

Compute Gaussian filters with different variance value

Convolve the intensity data with Gaussian filter bank

Find difference of Gaussians and find optimum values

Perform pruning and NMS on dd_i^{max} and dd_i^{min}

Salient points (Sp^{rgb})

Saliency Measure from Photometric Appearance

- Adaptive DoG-based detector is proposed by Zaharescu et al. [10]

Photometric appearance

$$P^{rgb} = [P_r, P_g, P_b].$$

Compute intensity from RGB channel

Compute Gaussian filters with different variance value

$$g_{ik}(s^b) = exp(-\frac{|p_k^{cor} - p_i^{cor}|^2}{2(s^b)^2}),\ where$$

$$NM_i^b = \{p_k^{cor} : |p_i^{cor} - p_k^{cor}| < r^b\},$$

$$s^b = s_m * 2^{\frac{b-1}{B}},$$

$$S = \{s^b \in \mathbb{R}^+ : b = 1, \cdots, B\}.$$

"B" represents the number of scales.

Convolve the intensity data with Gaussian filter bank

Find difference of Gaussians and find optimum values

Perform pruning and NMS on dd_i^{max} and dd_i^{min}

Salient points (Sp^{rgb})

Saliency Measure from Photometric Appearance

- Adaptive DoG-based detector is proposed by Zaharescu et al. [10]

Photometric appearance

$$P^{rgb} = [P_r, P_g, P_b].$$

Compute intensity from RGB channel

Compute Gaussian filters with different variance value

Convolve the intensity data with Gaussian filter bank

$$h_i^b = \frac{\sum_k I_k * g_{ik}(s^b)}{\sum_k g_k(s^b)}$$

Find difference of Gaussians and find optimum values

Perform pruning and NMS on dd_i^{max} and dd_i^{min}

Salient points (Sp^{rgb})

Saliency Measure from Photometric Appearance

- Adaptive DoG-based detector is proposed by Zaharescu et al. [10]

Photometric appearance

$$P^{rgb} = [P_r, P_g, P_b].$$

Compute intensity from RGB channel

Compute Gaussian filters with different variance value

Convolve the intensity data with Gaussian filter bank

Find difference of Gaussians and find optimum values

$$d_i^b = h_i^b - h_i^{b-1}$$

$$dd_i^{min} = min(d_i^b : b = 1, \cdots, B),$$

$$dd_i^{max} = max(d_i^b : b = 1, \cdots, B)$$

Perform pruning and NMS on dd_i^{max} and dd_i^{min}

Salient points (Sp^{rgb})

Saliency Measure from Photometric Appearance

- Adaptive DoG-based detector is proposed by Zaharescu et al. [10]

Photometric appearance
$$P^{rgb} = [P_r, P_g, P_b].$$

Compute intensity from RGB channel

Compute Gaussian filters with different variance value

Convolve the intensity data with Gaussian filter bank

Find difference of Gaussians and find optimum values

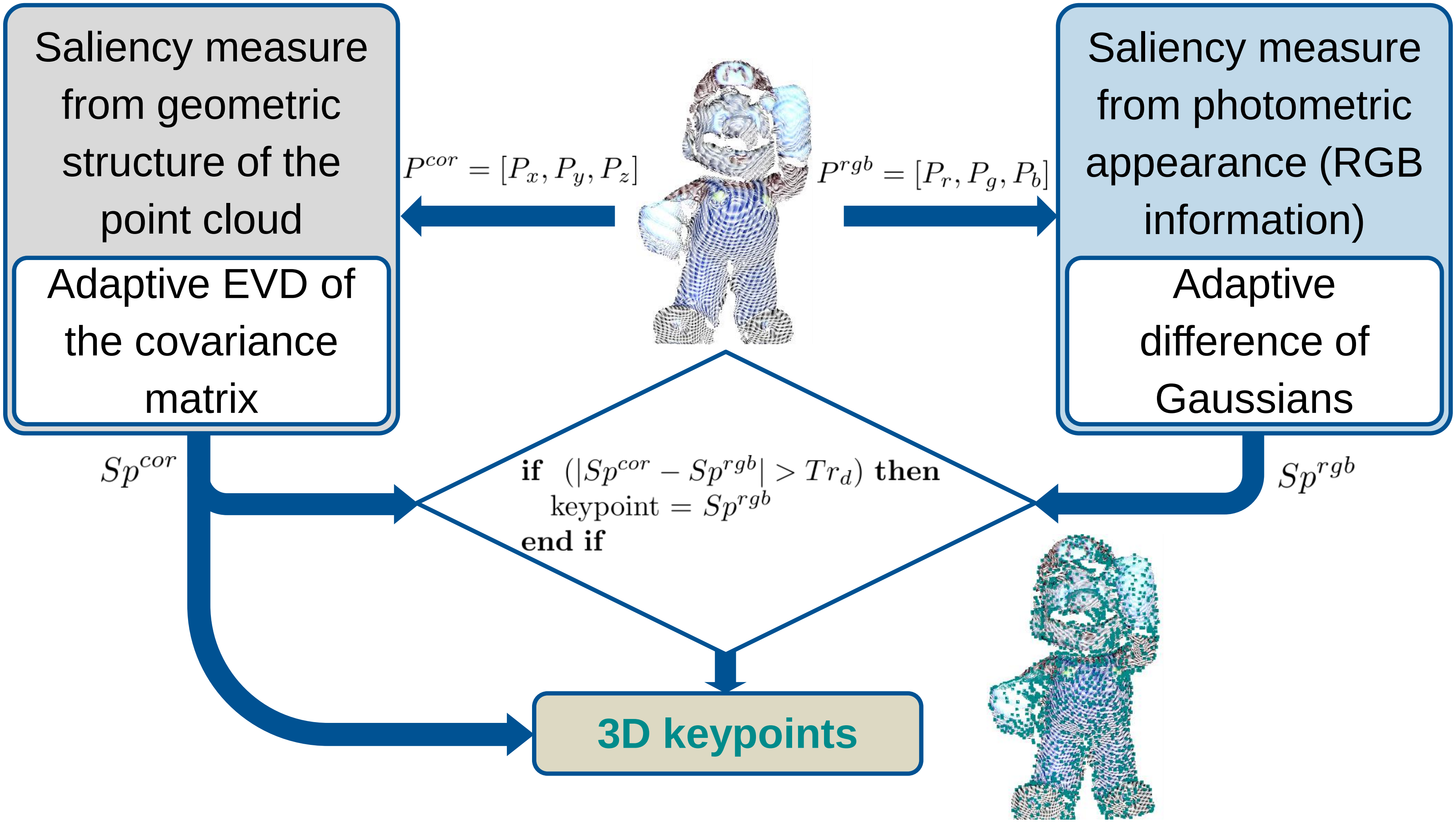
Perform pruning and NMS on dd_i^{max} and dd_i^{min}

if $(|dd_i^{min}| \vee |dd_i^{max}| > Th_{cn})$ **then**
  if $((\frac{v_{i2}}{v_{i1}} > Tr_{12}) \wedge (\frac{v_{i3}}{v_{i2}} > Tr_{23}) \wedge (v_{i3} > Tr_3))$ **then**
    prune = false
  else
    prune = true
  end if
else
  prune = true
end if

where, v_{i1}, v_{i2}, v_{i3} are eigenvalues (in descending order of magnitude) of covariance matrix of RGB.
Perform NMS at support radius of r^{nms}.

Salient points (Sp^{rgb})

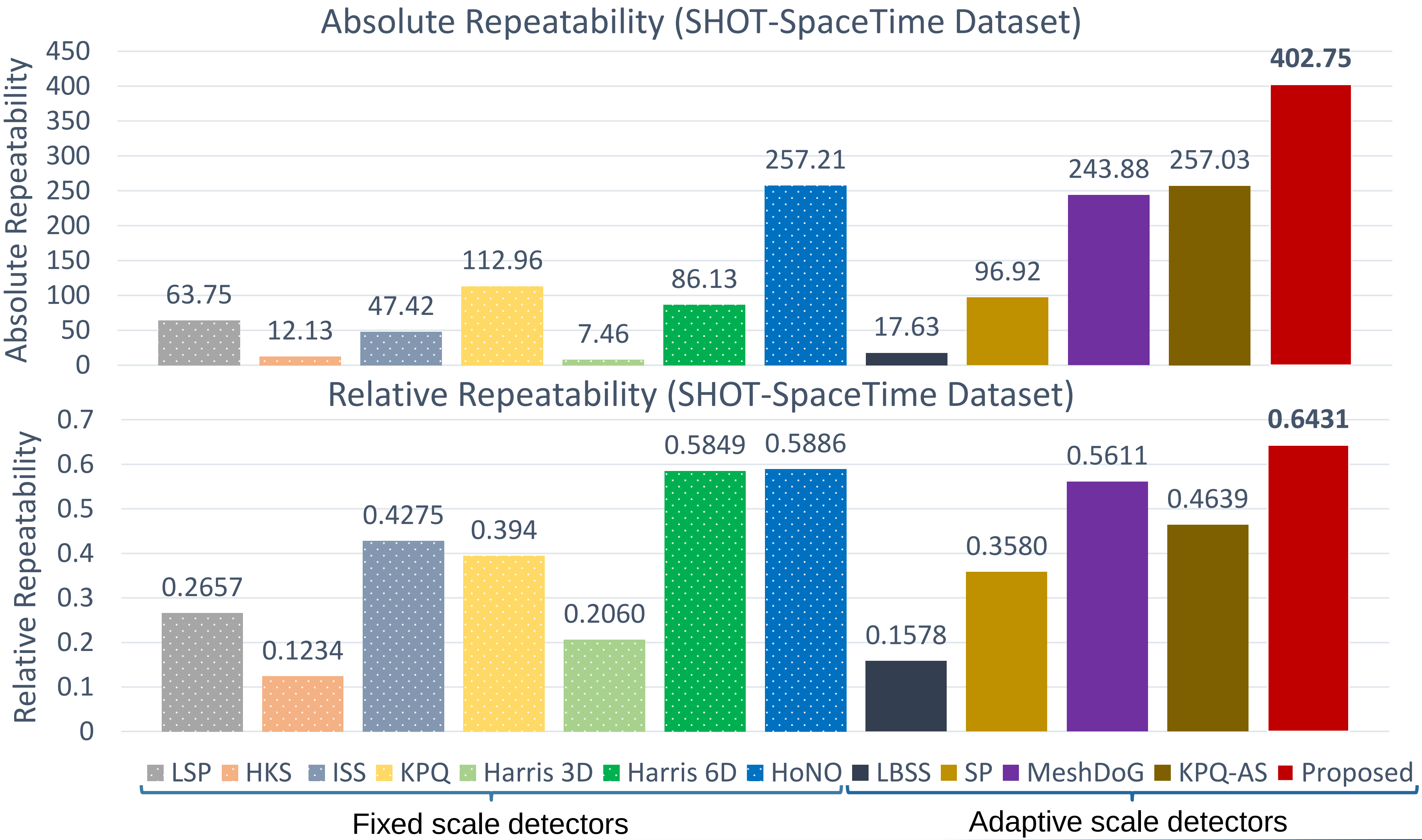
Fusion of Salient Points



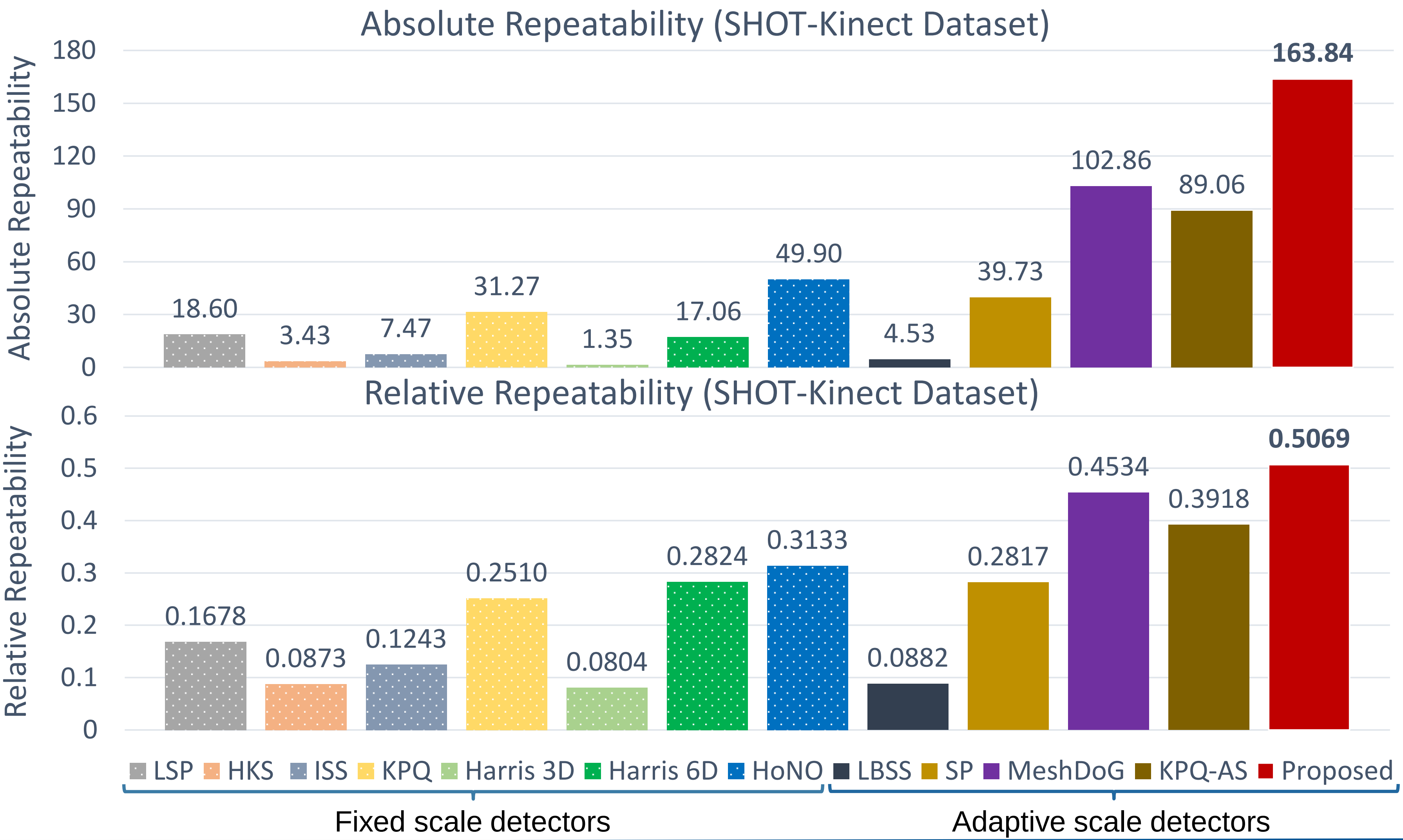
Experimental Evaluation

- Two point cloud datasets with RGB information
  - SHOT-SpaceTime dataset [11,12]
  - SHOT-Kinect dataset [11,12]
- Evaluation metrics
  - Absolute repeatability [11]: number of repeatable keypoints involving model-scene pair
  - Relative repeatability [11]: ratio of absolute repeatability to total number of non-occluded keypoints detected from the model

Repeatability SHOT-SpaceTime Dataset



Repeatability SHOT-Kinect Dataset



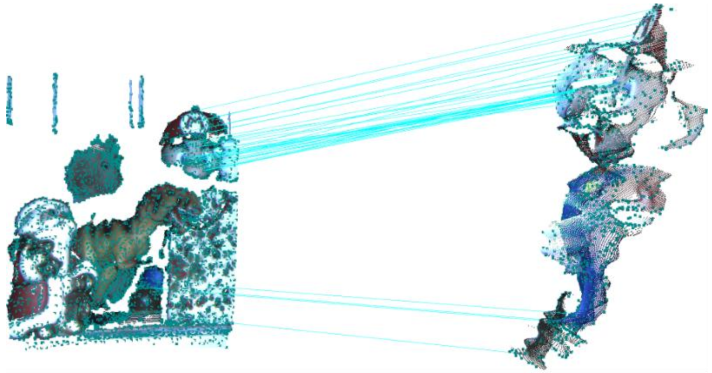
Scale Invariance Evaluation of the RGB-based Keypoint Detectors

Detector	Scale Space	SHOT-Kinect		SHOT-SpaceTime	
		Absolute Repeatability	Relative Repeatability	Absolute Repeatability	Relative Repeatability
Mesh-DoG [10]	3	88.69	0.3953	193.33	0.5113
	5	142.65	0.4650	257.67	0.5548
Proposed DoG	3	134.27	0.4772	337.63	0.6492
	5	136.94	0.4792	333.83	0.6486

Object Recognition using Keypoints

- SHOT descriptor [13] at keypoints and object recognition protocol defined in [14]

Instance found



Keypoints detected by proposed detector

No instance found

Keypoints detected by Mesh-DoG detector [4]

Conclusion

- Proposed adaptive fusion-based 3D keypoint detector
- Compute saliency using adaptive EVD of the covariance matrices
- Compute saliency by adaptive Difference of Gaussian of the RGB information
- Fusion of the salient points
- The proposed 3D keypoint detector is more repeatable, distinctive and scale-invariant

References

[1] H. Chen and B. Bhanu, “3d free-form object recognition in range images using local surface patches,” *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1252–1262, July 2007.

[2] J. Sun, M. Ovsjanikov, and L. Guibas, “A concise and provably informative multi-scale signature based on heat diffusion,” *Computer Graphics Forum*, vol. 28, pp. 1383–1392, November 2009.

[3] Y. Zhong, “Intrinsic shape signatures: A shape descriptor for 3d object recognition,” in *IEEE International Conference on Computer Vision Workshops*, Kyoto, Japan, pp. 689–696, September 2009.

[4] A. Mian, M. Bennamoun, and R. Owens, “On the repeatability and quality of keypoints for local feature based 3d object retrieval from cluttered scenes,” *International Journal of Computer Vision*, vol. 89, no. 2, pp. 348–361, September 2010.

[5] Ivan Sipiran and Benjamin Bustos, “Harris 3D: a robust extension of the harris operator for interest point detection on 3d meshes,” *The Visual Computer*, vol. 27, no. 11, pp. 963–976, July 2011.

[6] R. B. Rusu and S. Cousins, “3d is here: Point cloud library (pcl),” in *IEEE International Conference on Robotics and Automation,* Shanghai, China, pp. 1–4, May 2011.

[7] S. M. Prakhya, B. Liu, and W. Lin, “Detecting keypoint sets on 3d point clouds via histogram of normal orientations,” *Pattern Recognition Letters*, vol. 83, no. 1, pp. 42–48, November 2016.

[8] R. Unnikrishnan and M. Hebert, “Multi-scale interest regions from unorganized point clouds,” in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Alaska, USA, pp. 1–8, June 2008.

References

[9] U. Castellani, M. Cristani, S. Fantoni, and V. Murino, "Sparse points matching by combining 3d mesh saliency with statistical descriptors," *Computer Graphics Forum*, vol. 27, no. 2, pp. 643–652, April 2008.

[10] A. Zaharescu, E. Boyer, K. Varanasi, and R. Horaud, "Surface feature detection and description with applications to mesh matching," in *IEEE Conference on Computer Vision and Pattern Recognition*, Florida, USA, pp. 373–380, June 2009.

[11] F. Tombari, S. Salti, and L. D. Stefano, "Performance evaluation of 3d keypoint detectors," *International Journal of Computer Vision*, vol. 102, no. 1-3, pp. 198–220, March 2013.

[12] F. Tombari and S. Salti, "3d keypoint detection benchmark," Available: https://vision.deis.unibo.it/keypoints3d/

[13] F. Tombari, S. Salti, and L. D. Stefano, "Unique signatures of histograms for local surface description," in *11th European Conference on Computer Vision*, Heraklion, Crete, Greece, pp. 356–369, September 2010.

[14] F. Tombari and L. D. Stefano, "Object recognition in 3d scenes with occlusions and clutter by hough voting," *in Fourth Pacific-Rim Symposium on Image and Video Technology*, Singapore, pp. 349–355, November 2010.

Thank you!

Questions?

Muhammad Zafar Iqbal (mzafar.Iqbal@tum.de),

Dmytro Bobkov (dmytro.bobkov@tum.de),

Eckehard Steinbach (eckehard.steinbach@tum.de)

Acknowledgement: Higher Education Commission (HEC) Pakistan and DAAD Germany