

Saliency Detection in Large Point Sets – Supplementary 1

Quantitative evaluation

Elizabeth Shtrom
Technion

lizas@tx.technion.ac.il

George Leifman
Technion

gleifman@tx.technion.ac.il

Ayellet Tal
Technion

ayellet@ee.technion.ac.il

We evaluated our algorithm on the *Benchmark for 3D Interest Point Detection Algorithms* [2], whose goal is to provide quantitative evaluation for detecting interest points vs. human-marked points. The benchmark contains 43 models organized in two data sets. Data-set A includes 24 of the models for which 23 human subjects have marked the interest points. Data-set B contains all the 43 models marked by at least 16 human subjects. The benchmark evaluates six techniques: Mesh saliency [4], salient points defined by Castellani et al. [1], 3D-Harris [6], 3D-SIFT [3], scale-dependent corners [5], and Heat Kernel Signature (HKS) [7].

Measurements: Two criteria are being used while constructing the ground truth: the radius of an interest region σ and the number of users n , who marked a point within this interest region. Interest points whose geodesic distances to each other are less than 2σ are grouped together. If the number of points in the set is less than n , this set is discarded. Otherwise, a representative from the set is selected and set as the ground truth interest point.

Three measures are used to evaluate the performance: the False Negative Error (FNE), the False Positive Error (FPE) and the Weighted Miss Error (WME), as explained below.

Let us denote the set of ground truth points for model M as $\mathcal{G}(n, \sigma)$, and the set of interest points detected by an algorithm as \mathcal{A} . For an interest point g in a set \mathcal{G} , a geodesic neighborhood of radius r is defined as $\mathcal{C}_r(g) = \{p \in M \mid d(g, p) \leq r\}$, where $d(g, p)$ corresponds to the geodesic distance between points g and p . The parameter r controls the localization error tolerance. A point g is considered to be correctly detected if there exists a detected point $a \in \mathcal{A}$ in $\mathcal{C}_r(g)$, such that a is not closer to any other point in \mathcal{G} . Denoting the number of correctly detected points in \mathcal{G} as N_C and the number of points in \mathcal{G} as N_G , the false negative error rate at localization error tolerance r is defined as

$$FNE(r) = 1 - \frac{N_C}{N_G}. \quad (1)$$

To calculate the false positive error rate, each correctly detected point $g \in \mathcal{G}$ corresponds to a unique a , the closest point to g among the points in \mathcal{A} . All the points in \mathcal{A} without a correspondence in \mathcal{G} are declared as false positives. The number of false positives is defined as $N_F = N_A - N_C$, where N_A is the number of interest points detected by the algorithm. The False Positive Error (FPE) rate at localization error tolerance r is then defined as:

$$FPE(r) = \frac{N_F}{N_A}. \quad (2)$$

To incorporate the prominence of an interest point into the evaluation, another miss error measure is used, named the Weighted Miss Error (WME). Assuming that within a geodesic neighborhood of radius r around the ground truth point $g_i \in \mathcal{G}$, n_i subjects marked an interest point, the Weighted Miss Error is defined as

$$WME(r) = 1 - \frac{1}{\sum_{i=1}^{N_G} n_i} \sum_{i=1}^{N_G} \delta_i n_i, \quad (3)$$

where $\delta_i = \begin{cases} 1 & \text{if } g_i \text{ is detected by the algorithm} \\ 0 & \text{otherwise.} \end{cases}$

The WME's range is between 0 and 1. An algorithm gets a low WME if the points it detects are frequently voted by the human subjects. This error measures the ability of an algorithm to detect the most prominent interest points. It also makes it possible to take into account less significant interest points, which would have been discarded by hard thresholding with high n .

Results: We ran our algorithm on the vertices of the meshes, without using any connectivity information. To extract key points from our saliency map, we use the same method described in [2]. Specifically, candidate interest points are selected from the local maxima of the saliency map. A vertex is marked as a local maximum if its saliency is higher than that of all its neighboring vertices. Then, the candidate points with saliency higher than a threshold are selected as the final interest points. The threshold used is the average saliency over all the local maxima.

Figures 1 and 2 show the performance of the six methods tested in [2] and ours, with respect to the localization error tolerance r . The lower the graphs, the better the results. With an ideal interest point detection algorithm, the errors are expected to drop very quickly with respect to r . A rapid drop in FNE means the algorithm catches the interest points with a low localization error. A rapid drop in FPE indicates that the algorithm does not return excessive interest points.

As the graphs show, for both datasets and for most of the settings of σ and n , the FNE and WME drop faster with our algorithm, compared to the other methods. As the localization error tolerance increases, our method detects more ground-truth interest points than the competing methods; i.e. achieves lower FNE and WME. The low WME measure demonstrates the ability of our algorithm to identify the semantically-significant interest points. Note that all the algorithms, except the HKS [7] method, mark more interest points than the human users. For the FPE measure, our method is similar to the results of other methods. Only HKS outperforms it significantly, since it detects very few key points, mainly extremities, which usually match the key points marked by humans.

References

- [1] U Castellani, M Cristani, S Fantoni, and V Murino. Sparse points matching by combining 3d mesh saliency with statistical descriptors. *Computer Graphics Forum*, 27(2):643–652, 2008. 1
- [2] H. Dutagaci, C. Cheung, and A. Godil. Evaluation of 3d interest point detection techniques via human-generated ground truth. *The Visual Computer*, 28:901–917, 2012. 1, 2, 3, 4
- [3] A Godil and A I Wagan. Salient local 3d features for 3d shape retrieval. In *IS&T/SPIE Electronic Imaging*, pages 78640S–78640S, 2011. 1
- [4] C.H. Lee, A. Varshney, and D.W. Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3):659–666, 2005. 1
- [5] J. Novatnack and K. Nishino. Scale-dependent 3d geometric features. In *ICCV*, pages 1–8. IEEE, 2007. 1
- [6] I Pratikakis, M Spagnuolo, T Theoharis, and R Veltkamp. A robust 3d interest points detector based on harris operator. In *Eurographics Workshop on 3D Object Retrieval*, 2010. 1
- [7] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5):1383–1392, 2009. 1, 2

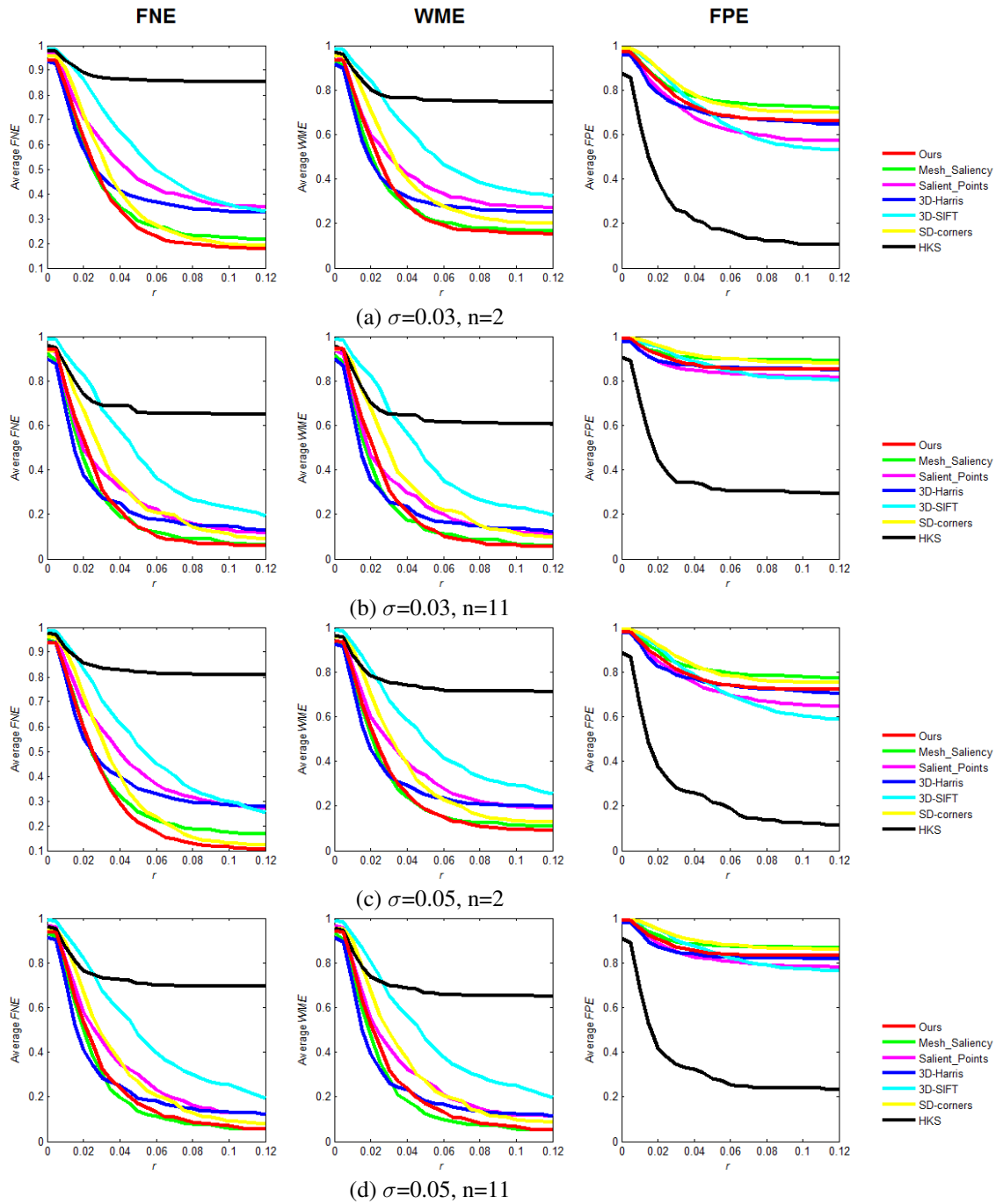


Figure 1. **Evaluation on the Benchmark for 3D Interest Point Detection Algorithms [2]:** Our algorithm’s performance graph is colored with red. The figure shows the performance on Dataset A (24 Models, 23 Subjects). Each column shows a different error: False Negative Error, Weighted Miss Error, and False Positive Error. The settings for obtaining the ground truth are indicated under the plot. The lower the graphs, the better the results.

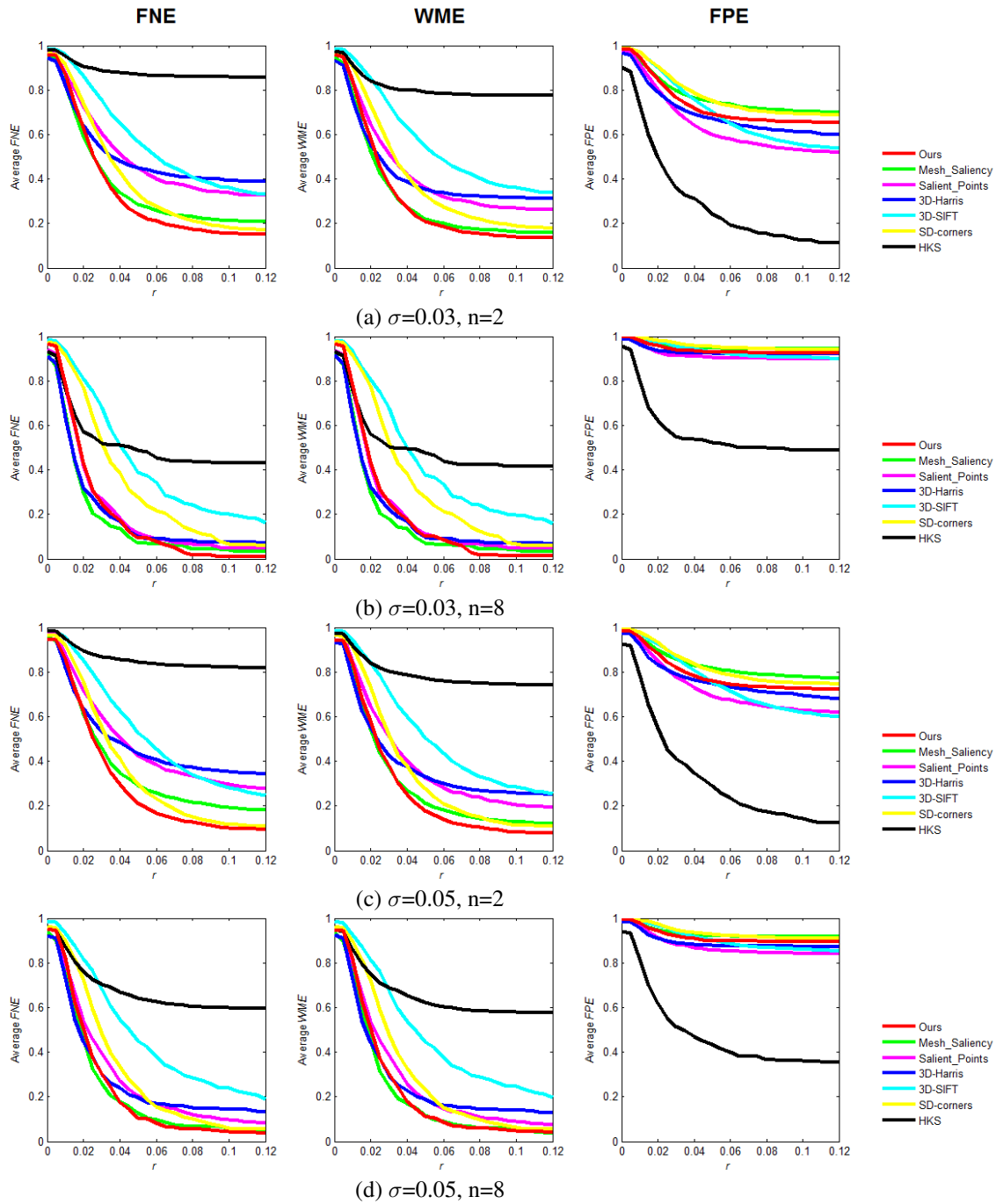


Figure 2. **Evaluation on the Benchmark for 3D Interest Point Detection Algorithms [2]:** Our algorithm’s performance graph is colored with red. The figure shows the performance on Dataset B (43 Models, 16 Subjects). Each column shows a different error: False Negative Error, Weighted Miss Error, and False Positive Error. The settings for obtaining the ground truth are indicated under the plot. The lower the graphs, the better the results.