IRIS-VIS: a New Dataset for Visibility Estimation in an Industrial Environment — Supplementary Material —

Flavien Armangeon[1,2] Thibaud Ehret[1] Enric Meinhardt-Llopis[1]
Rafael Grompone von Gioi[1] Guillaume Thibault[2] Marc Petit[2] Gabriele Facciolo[1]
[1] Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France
[2] EDF R&D (Electricité de France), France
https://centreborelli.github.io/iris-vis

1. Implementation details

As discussed in the main paper, the quantitative experiments are averaged over nine viewpoints located in three subparts of the TEST scene, illustrated in Figure 1. The missing parts are redundant as the structure of the objects is similar to those in the selected areas.

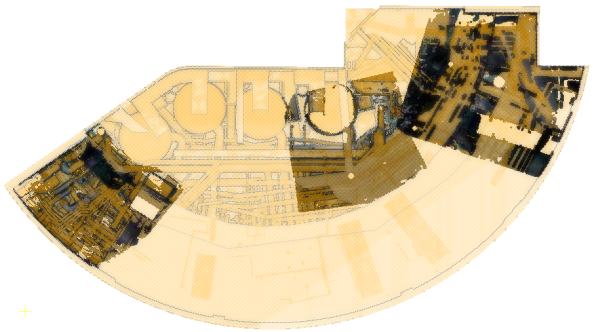

Figure 1. The three point clouds used for the quantitative experiments. They cover 30% of the whole TEST scene ($530m^2$).

Experiments are performed on a Intel Xeon W-2235 CPU for traditional methods (DVPS [4], VEVD [1],VEVD-I and on a RTX 3090 GPU for deep learning based methods (Vis2Mesh [5] and NKSR [2]).

We used our own code for input and ground truth construction, as well as for methods [4] and [1]. This is because we did not find any previously available implementation for DVPS including the three *inversions* mentioned in our paper.

In their paper, Biasutti *et al*. [1] suggest the use of a 3D-to-2D perspective projection before the KNN search step. In order to estimate the visibility in all directions around the viewpoint, we perform instead a spherical projection. The projected point coordinates are kept in 3D to avoid the discontinuity in the limits of the spherical angles range. It implies that the KNN search step is performed in 3D.

Vis2Mesh [5] takes as input a set of camera views in addition to the point cloud. We arbitrarily chose 50 cameras in each subscene located in reasonable distance from the objects (a few tens of centimeters) and placed close to different complex equipment.

In order to choose the most appropriate parameters, we evaluated DVPS [4] and VEVD [1] with different parameters. The results for these experiments are shown in Tables 4 and 5. We also observed during our experiments that a higher amount of neighbors in VEVD improves slightly the results. However, with more than 75 neighbors, the KNN search step becomes computationally expensive for the whole scene while leading to marginal improvements. We consider that the improvements are not worth the computational cost and thus chose to only keep 75 neighbors. Only one parameters configuration is given for NKSR because of the low variability in the results. In the end, the chosen parameters for the different methods are:

- DVPS [4]: An exponential inversion with parameter -10^{-4},
- VEVD [1]: 75 neighbors and a visibility threshold of 0.7,
- VEVD-I: 75 neighbors, a visibility threshold 0.55 and the median as the threshold function,
- Vis2Mesh [5]: Default parameters,
- NKSR [2]: A chunk size of 50.

2. Quantitative evaluation

Table 1 gives the results on the synthetic clouds, uniformly sampled from the CAD model. These clouds contain the same number of points and are located in the exact same areas as the LiDAR point clouds (Figure 1). As the synthetic clouds are noiseless by construction, they are useful to study the impact of the acquisition noise by comparing the metrics

with the LiDAR clouds provided in the main paper. First, it is important to remark that the results are significantly different between the two types of clouds. For all the methods, except VEVD, the recall, accuracy and f1-score are higher in the synthetic experiments for both densities, especially on the complex visibility estimation metrics. However, the precision decreases most of the time. These observations tend to demonstrate the impact of the noise for the visibility estimation task and the advantage of providing a real point cloud in the IRIS-VIS dataset.

On scene Quali-1 (Table 2), the results are better because of the simplicity of its environment compared to the other scenes. The results on scene Quali-2 (Table 3) follow the same trend as in the quantitative scenes (main paper). For every scene, the computation times for DVPS [4], Vis2Mesh [5] and NKSR [2] are mostly in the same order of magnitude and lower than other methods. Note however that Vis2Mesh and NKSR were run on GPU.

Tables 4 and 5 provide the results respectively of DVPS and VEVD, VEVD-I according to different parameters. We selected the best parameters regarding the results in dense configuration, with an emphasis on false positives related metrics. In fact, we consider that the false positives are more disruptive than the false negatives for most of the computer vision applications involving visibility estimation. For all methods and densities, we see that the results vary considerably depending on the parameters. For VEVD [1] and VEVD-I, the choice of the best parameters seems to be robust to the density. This comes from the fact that the minimum and maximum depths among the neighbors do not significantly change with the density.

3. Qualitative evaluation

Figures 2 and 3 show additional qualitative visualizations on scenes Quali-1 and Quali-2 respectively. In both scenes, except for DVPS, the results are similar in dense and sparse conditions as discussed in the main paper. The differences in the results given by DVPS, mostly located behind the pole, are also discussed.

References

[1] Pierre Biasutti, Aurélie Bugeau, Jean-François Aujol, and Mathieu Brédif. Visibility estimation in point clouds with variable density. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, 2019. 1, 2, 3, 5, 6, 7

[2] Jiahui Huang, Zan Gojcic, Matan Atzmon, Or Litany, Sanja Fidler, and Francis Williams. Neural kernel surface reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4369–4379, 2023. 1, 2, 3, 6, 7

[3] Sagi Katz and Ayellet Tal. On the Visibility of Point Clouds. In *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE. 4

[4] Sagi Katz, Ayellet Tal, and Ronen Basri. Direct visibility of point sets. In *ACM SIGGRAPH 2007 papers*, pages 24–es. 2007. 1, 2, 3, 4, 6, 7

[5] Shuang Song, Zhaopeng Cui, and Rongjun Qin. Vis2mesh: Efficient mesh reconstruction from unstructured point clouds of large scenes with learned virtual view visibility. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 6514–6524, 2021. 1, 2, 3, 6, 7

Density	Sparse					Dense				
Method	DVPS	VEVD	VEVD-I	Vis2Mesh	NKSR	DVPS	VEVD	VEVD-I	Vis2Mesh	NKSR
t(s)	10^1	10^2	10^2	10^1	10^1	10^1	10^2	10^2	10^2	10^1
TP	23.43	**24.03**	18.12	21.48	23.46	21.66	**24.98**	24.35	21.25	23.38
FP	**2.43**	19.05	3.76	6.52	06.03	**1.68**	18.23	5.18	3.56	06.09
FN	2.79	**2.19**	8.10	4.74	2.76	4.68	**1.37**	2.00	5.10	2.96
TN	**71.35**	54.74	70.02	67.26	67.76	**71.97**	55.42	68.47	70.10	67.56
Precision	**90.61**	55.79	82.82	76.73	79.57	**92.79**	57.80	82.46	85.65	79.33
Recall	89.36	**91.66**	69.11	81.94	89.48	82.22	**94.80**	92.42	80.65	88.75
Accuracy	**94.78**	78.77	88.14	88.75	91.22	**93.63**	80.40	92.82	91.34	90.94
F1-score	**89.98**	69.36	75.35	79.25	84.23	**87.19**	71.82	87.15	83.08	83.77
TP-c	16.12	21.10	14.30	**23.11**	19.56	17.63	**30.17**	25.35	26.85	24.30
FP-c	**1.54**	18.41	2.94	14.15	7.21	**0.50**	16.50	3.00	11.11	07.02
FN-c	12.71	7.73	14.53	**5.72**	9.27	18.81	**6.27**	11.09	9.58	12.14
TN-c	**69.63**	52.76	68.23	57.02	63.96	**63.06**	47.06	60.57	52.45	56.55
Precision-c	**91.26**	53.41	82.94	62.02	73.07	**97.24**	64.65	89.43	70.73	77.60
Recall-c	55.91	73.20	49.59	**80.15**	67.85	48.38	**82.80**	69.56	73.70	66.68
Accuracy-c	**85.75**	73.87	82.53	80.13	83.52	80.69	77.23	**85.91**	79.30	80.85
F1-score-c	69.34	61.76	62.07	**69.93**	70.37	64.61	72.61	**78.25**	72.18	71.73

Table 1. Quantitative results on the point cloud sampled from the CAD model. Positive predictions are the visible points in the outputs. The "-c" metrics are the complex visibility estimation metrics. Vis2Mesh and NKSR were run on a GPU, the others on CPU. VEVD [1], DVPS [4], Vis2Mesh [5] and NKSR [2] are state-of-the-art methods.

Density	Sparse					Dense				
Method	Katz	Biasutti	Biasutti-I	Vis2mesh	NKSR	Katz	Biasutti	Biasutti-I	Vis2mesh	NKSR
t(s)	10^0	10^0	10^0	10^0	10^0	10^0	10^0	10^1	10^0	10^0
TP	43.33	41.52	29.39	**43.62**	42.68	41.68	41.93	41.65	**43.56**	42.56
FP	3.60	4.06	**0.07**	1.61	7.62	**0.84**	3.58	1.26	1.41	7.16
FN	1.34	3.15	15.28	**1.05**	1.98	2.96	2.71	2.98	**1.07**	2.07
TN	51.74	51.28	**55.26**	53.73	47.72	**54.52**	51.78	54.11	53.96	48.21
Precision	92.34	91.10	**99.76**	96.45	84.86	**98.01**	92.13	97.07	96.87	85.60
Recall	97.00	92.96	65.79	**97.65**	95.56	93.37	93.94	93.32	**97.59**	95.36
Accuracy	95.06	92.80	84.65	**97.35**	90.40	96.20	93.71	95.76	**97.52**	90.77
F1-score	94.61	92.02	79.29	**97.05**	89.89	95.64	93.02	95.16	**97.23**	90.22
TP-c	43.64	41.60	7.44	**46.15**	40.62	37.52	46.39	35.45	**51.26**	45.95
FP-c	5.76	7.09	**0.14**	3.48	16.36	1.12	5.14	**1.01**	2.68	14.04
FN-c	6.80	8.85	43.01	**4.30**	9.83	20.17	11.31	22.24	**6.43**	11.74
TN-c	43.79	42.47	**49.42**	46.07	33.19	41.18	37.17	**41.29**	39.63	28.27
Precision-c	88.34	85.44	**98.20**	92.98	71.28	97.09	90.03	**97.22**	95.03	76.60
Recall-c	86.51	82.46	14.75	**91.48**	80.52	65.04	80.40	61.45	**88.86**	79.64
Accuracy-c	87.43	84.06	56.86	**92.22**	73.81	78.70	83.56	76.75	**90.89**	74.22
F1-score-c	87.42	83.92	25.65	**92.22**	75.62	77.90	84.94	75.31	**91.84**	78.09

Table 2. Quantitative results on Scene Quali-1. Positive predictions are the points set as visible in the outputs. Vis2Mesh and NKSR was run on GPU, the others on CPU.

Density	Sparse					Dense				
Method	Katz	Biasutti	Biasutti-I	Vis2mesh	NKSR	Katz	Biasutti	Biasutti-I	Vis2mesh	NKSR
t(s)	10^0	10^1	10^1	10^1	10^1	10^1	10^2	10^2	10^1	10^1
TP	25.73	27.70	**27.87**	27.79	27.59	21.20	27.78	**28.54**	27.74	27.64
FP	**2.69**	18.20	8.72	3.16	3.72	**0.89**	17.51	8.01	2.84	3.42
FN	4.94	2.97	**2.80**	2.88	3.08	9.46	2.88	**2.12**	2.92	3.02
TN	**66.64**	51.13	60.60	66.17	65.60	**68.45**	51.83	61.33	66.50	65.92
Precision	**90.53**	60.35	76.16	89.78	88.10	**95.99**	61.34	78.09	90.71	89.00
Recall	83.90	90.31	**90.86**	90.62	89.95	69.15	90.62	**93.10**	90.49	90.14
Accuracy	92.37	78.83	88.47	**93.96**	93.19	89.66	79.61	89.87	**94.24**	93.56
F1-score	87.09	72.35	82.86	**90.20**	89.02	80.39	73.16	84.93	**90.60**	89.56
TP-c	21.87	26.23	22.11	**27.16**	24.40	15.88	32.63	28.97	**32.51**	28.52
FP-c	4.39	14.01	**4.31**	9.45	9.67	**0.95**	10.59	2.12	7.85	8.84
FN-c	13.99	9.62	13.75	**8.70**	11.46	27.26	**10.51**	14.18	10.64	14.62
TN-c	59.75	50.13	**59.84**	54.70	54.47	**55.90**	46.27	54.73	49.00	48.01
Precision-c	83.28	65.18	**83.69**	74.19	71.61	**94.35**	75.50	93.18	80.55	76.34
Recall-c	60.98	73.16	61.66	**75.74**	68.04	36.81	**75.64**	67.14	75.35	66.11
Accuracy-c	81.62	76.36	**81.95**	81.85	78.87	71.79	78.90	**83.70**	81.51	76.54
F1-score-c	70.41	68.94	71.01	**74.95**	69.78	52.96	75.57	**78.04**	77.86	70.86

Table 3. Quantitative results on Scene Quali-2. Positive predictions are the points set as visible in the outputs. Vis2Mesh and NKSR was run on GPU, the others on CPU.

Density	Inversion	Parameter	t(s)	TP	FP	FN	TN	Precision	Recall	Accuracy	F1-score
Sparse	Lin.	4.0	10^1	36.67	3.93	4.00	55.41	89.54	89.33	92.08	89.37
	Lin.	4.5	10^1	37.88	9.21	2.79	50.13	79.29	92.61	88.00	85.23
	Lin.	5.0	10^1	38.51	19.13	2.15	40.21	65.65	94.34	78.72	77.01
	Exp.	-1e-03	10^1	27.85	**0.48**	12.81	**58.86**	**98.36**	66.20	86.71	78.80
	Exp.	-1e-04	10^1	35.32	1.75	5.34	57.58	95.18	85.57	**92.91**	**90.04**
	Exp.	-1e-05	10^1	38.14	9.67	2.53	49.67	78.97	93.23	87.81	85.38
	Nat Exp.	1e-07	10^1	32.98	1.10	7.68	58.24	96.77	79.44	91.22	87.09
	Nat Exp.	1e-08	10^1	37.18	5.27	3.48	54.07	86.86	90.62	91.25	88.64
	Nat Exp.	1e-09	10^1	**38.81**	23.55	**1.86**	35.78	61.08	**95.10**	74.59	74.03
Dense	Lin.	4.0	10^2	30.87	1.05	9.80	58.28	96.61	74.64	89.15	84.08
	Lin.	4.5	10^2	32.02	1.70	8.65	57.63	94.66	77.78	89.65	85.30
	Lin.	5.0	10^2	32.52	3.16	8.15	56.17	90.40	79.01	88.69	84.22
	Exp.	-1e-03	10^2	19.11	**0.20**	21.56	**59.12**	**98.99**	45.36	78.24	61.91
	Exp.	-1e-04	10^2	29.51	0.70	11.16	58.63	97.72	70.95	88.14	82.03
	Exp.	-1e-05	10^2	32.39	1.70	8.28	57.63	94.83	78.60	**90.02**	**85.86**
	Nat Exp.	1e-07	10^2	26.92	0.51	13.75	58.82	98.18	64.61	85.74	77.73
	Nat Exp.	1e-08	10^2	31.67	1.23	9.00	58.10	96.18	76.74	89.77	85.25
	Nat Exp.	1e-09	10^2	**33.34**	4.25	**7.33**	55.07	88.00	**81.11**	88.41	84.32

Table 4. Quantitative results for DVPS [3, 4], ran on CPU. Positive predictions are the points set as visible in the outputs. Red lines show the chosen parameters. Precision on the dense point cloud is the first criterion for the choice of the best parameters, followed by accuracy and f1-score.

Density	t	α	t(s)	TP	FP	FN	TN	Precision	Recall	Accuracy	F1-score
Sparse	/	0.6	10^2	**37.78**	20.61	**2.89**	38.73	63.56	**92.52**	76.50	74.72
	/	0.7	10^2	36.62	17.37	4.04	41.97	66.53	89.57	78.59	75.79
	/	0.8	10^2	25.47	12.66	15.20	46.67	65.89	62.96	72.14	63.72
	Mean	0.6	10^2	34.58	8.20	6.08	51.14	79.03	83.37	85.72	80.89
	Median	0.55	10^2	33.81	6.56	6.85	52.77	81.91	82.28	**86.59**	**81.92**
	Q1	0.45	10^2	33.27	6.57	7.39	52.77	81.67	80.89	86.04	81.01
	10%	0.4	10^2	27.27	**5.34**	13.39	**54.00**	**82.02**	66.39	81.27	71.77
Dense	/	0.6	10^3	**38.64**	19.05	**2.03**	40.28	65.80	**94.74**	78.92	77.04
	/	0.7	10^3	37.29	16.14	3.38	43.19	68.47	91.33	80.48	77.71
	/	0.8	10^3	24.14	12.10	16.53	47.23	65.88	59.84	71.37	61.98
	Mean	0.6	10^3	37.56	9.09	3.11	50.24	79.31	91.62	**87.80**	**84.88**
	Median	0.55	10^3	37.22	8.96	3.46	50.37	79.73	91.14	87.58	84.79
	Q1	0.45	10^3	36.72	10.24	3.95	49.09	78.14	90.59	85.81	83.24
	10%	0.4	10^3	35.54	**8.36**	5.13	**50.96**	**80.28**	88.37	86.50	83.06

Table 5. Quantitative results for VEVD [1], ran on CPU with 75 nearest neighbors. Positive predictions are the points set as visible in the outputs. Parameter t is the threshold function on the depth values (/ means no thresholding). α is the visibility threshold. Red and blue lines show the chosen parameters for VEVD and VEVD-I respectively. Precision on the dense point cloud is the first criterion for the choice of the best parameters, followed by accuracy and f1-score.

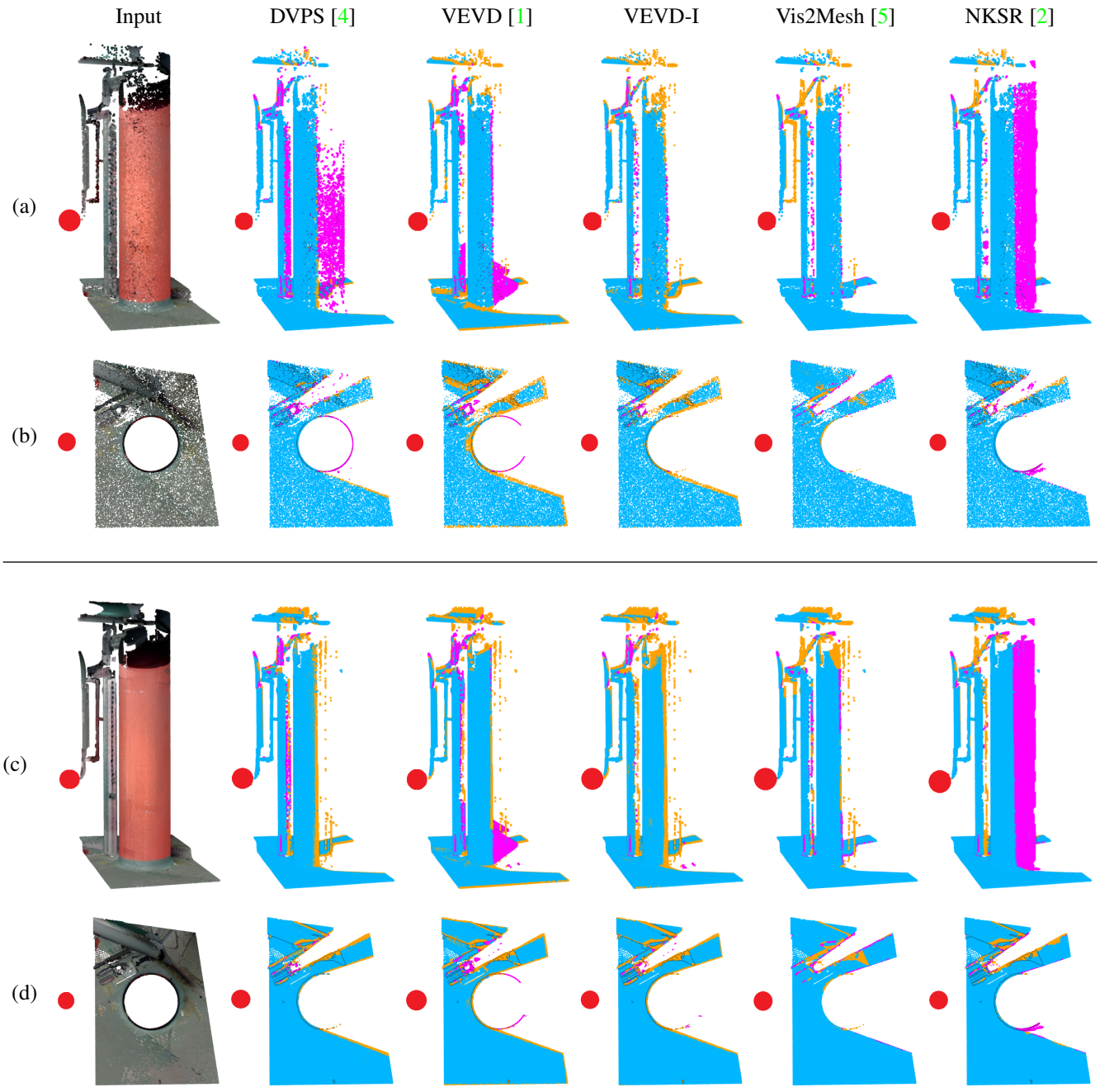

Figure 2. Qualitative results on Scene Quali-1 sparse (a, b) and dense (c, d). TP (blue), FP (purple), FN (orange). Positive predictions are the points set as visible in the outputs.