Supplementary Material
LDP-FEAT: Image Features with Local Differential Privacy

Francesco Pittaluga
francescopittaluga@nec-labs.com

Bingbing Zhuang
bzhuang@nec-labs.com

NEC Labs America

The supplementary material contains (1) a proof that LDP-FEAT satisfies ϵ-LDP, (2) additional experimental results on Aachen night-time localization and Structure-from-Motion (SfM), and (3) an analysis of the paper's assumptions.

S1. Local Differential Privacy of LDP-FEAT

Here, we prove that LDP-FEAT satisfies ϵ-LDP. For clarity, we first prove that ω-SM satisfy ω-LDP and the proof for LDP-FEAT follows very similarly.

S1.1. Theorem 1 (ω-Subset satisfies ϵ-LDP)

For any inputs v_1 and v_2, and their output $\mathcal{Z}_1$ and $\mathcal{Z}_2$ returned by ω-SM, there are four possible scenarios $\{v_1{\in}\mathcal{Z}_1, v_2{\in}\mathcal{Z}_2\}$, $\{v_1{\notin}\mathcal{Z}_1, v_2{\in}\mathcal{Z}_2\}$, $\{v_1{\in}\mathcal{Z}_1, v_2{\notin}\mathcal{Z}_2\}$, $\{v_1{\notin}\mathcal{Z}_1, v_2{\notin}\mathcal{Z}_2\}$, each with different probability distributions. Below, we show that the probability inequality required by ϵ-LDP, *i.e*. Eq. (3) of the main paper, is satisfied for all the four scenarios.
1) $\{v_1{\in}\mathcal{Z}_1, v_2{\in}\mathcal{Z}_2\}$. In this case,

$$\begin{aligned}\Pr(\mathcal{Z}_1|v_1) &= \frac{\omega e^\epsilon}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\\ \Pr(\mathcal{Z}_2|v_2) &= \frac{\omega e^\epsilon}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\end{aligned} \tag{S1}$$

meaning that $\Pr(\mathcal{Z}_1|v_1){=}\Pr(\mathcal{Z}_2|v_2)$, hence $\Pr(\mathcal{Z}_1|v_1) \leq e^\epsilon \Pr(\mathcal{Z}_2|v_2)$ holds.
2) $\{v_1{\notin}\mathcal{Z}_1, v_2{\in}\mathcal{Z}_2\}$. In this case,

$$\begin{aligned}\Pr(\mathcal{Z}_1|v_1) &= \frac{\omega}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\\ \Pr(\mathcal{Z}_2|v_2) &= \frac{\omega e^\epsilon}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\end{aligned} \tag{S2}$$

meaning that $\Pr(\mathcal{Z}_1|v_1) = e^{-\epsilon} \Pr(\mathcal{Z}_2|v_2)$, hence $\Pr(\mathcal{Z}_1|v_1) \leq e^\epsilon \Pr(\mathcal{Z}_2|v_2)$ holds since $\epsilon > 0$.
3) $\{v_1{\in}\mathcal{Z}_1, v_2{\notin}\mathcal{Z}_2\}$. In this case,

$$\begin{aligned}\Pr(\mathcal{Z}_1|v_1) &= \frac{\omega e^\epsilon}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\\ \Pr(\mathcal{Z}_2|v_2) &= \frac{\omega}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\end{aligned} \tag{S3}$$

meaning that $\Pr(\mathcal{Z}_1|v_1) = e^\epsilon \Pr(\mathcal{Z}_2|v_2)$, hence $\Pr(\mathcal{Z}_1|v_1) \leq e^\epsilon \Pr(\mathcal{Z}_2|v_2)$ holds.
4) $\{v_1{\notin}\mathcal{Z}_1, v_2{\notin}\mathcal{Z}_2\}$. In this cases,

$$\begin{aligned}\Pr(\mathcal{Z}_1|v_1) &= \frac{\omega}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\\ \Pr(\mathcal{Z}_2|v_2) &= \frac{\omega}{\omega e^\epsilon + |\mathcal{K}| - \omega} / \binom{|\mathcal{K}|}{\omega},\end{aligned} \tag{S4}$$

meaning that $\Pr(\mathcal{Z}_1|v_1){=}\Pr(\mathcal{Z}_2|v_2)$, hence $\Pr(\mathcal{Z}_1|v_1) \leq e^\epsilon \Pr(\mathcal{Z}_2|v_2)$ holds.
This concludes our proof.

S1.2. Theorem 2 (LDP-FEAT satisfies ϵ-LDP)

For any input descriptor d, the output set $\mathcal{Z}$ are obtained by: first map d to an element (let us denote it as $\bar{d}$) in the database $\mathcal{K}$, and then $\bar{d}$ is mapped to the random set $\mathcal{Z}$. Hence,

$$\begin{aligned}\Pr(\mathcal{Z}|d) &= \sum_{\bar{d}\in\mathcal{K}} \Pr(\mathcal{Z}, \bar{d}|d)\\ &= \sum_{\bar{d}\in\mathcal{K}} \Pr(\mathcal{Z}|\bar{d}) \Pr(\bar{d}|d).\end{aligned} \tag{S5}$$

Since the mapping from d to $\bar{d}$ is deterministic – it is mapped to the nearest neighbor d' in the database, we have

$$\Pr(\bar{d}|d) = \begin{cases} 1, & \text{if } \bar{d} = d,\\ 0, & \text{if } \bar{d} \neq d.\end{cases} \tag{S6}$$

Plugging Eq. (S6) into Eq. (S5) yields

$$\Pr(\mathcal{Z}|d) = \Pr(\mathcal{Z}|d') \tag{S7}$$

	DB Size	**Privacy**	**# Desc.**	**Day**			**Night**		
	$\lvert\mathcal{K}\rvert$	ϵ	m	0.25m, 2°	0.5m, 5°	5.0m, 10°	0.25m, 2°	0.5m, 5°	5.0m, 10°
Accuracy Upper Bound	128k	∞	1	73.2	82.8	88.1	24.6	28.8	33.5
	256k	∞	1	76.8	86.3	91.6	28.8	34.6	42.4
	512k	∞	1	78.0	87.4	93.3	33.5	40.8	51.3
	1024k	∞	1	79.7	89.9	94.9	36.1	42.4	51.8
	2^{1024}	∞	1	**84.1**	**91.7**	**96.4**	**50.3**	**61.8**	**73.8**
Impact of Database Size	128k	10	2	39.4	45.4	50.2	1.60	2.10	3.70
	256k	10	2	**42.1**	**49.3**	**54.4**	**3.70**	**5.20**	**6.30**
	512k	10	2	37.4	43.2	48.1	2.10	3.70	4.70
	1024k	10	2	33.3	37.1	42.1	2.10	3.70	3.70
	2^{1024}	10	2	0.00	0.00	0.00	0.00	0.00	0.00
Privacy Guarantee	256k	10	2	42.1	49.3	54.4	3.70	5.20	6.30
	256k	12	2	69.5	78.4	84.1	16.8	19.4	23.6
	256k	14	2	75.1	84.7	89.4	**23.0**	**27.7**	**33.0**
	256k	16	2	**75.4**	**85.3**	**90.2**	**23.0**	27.2	31.9
	512k	10	4	42.1	49.6	55.0	5.20	7.90	9.90
	512k	12	4	69.4	77.9	84.6	19.4	22.5	26.2
	512k	14	4	73.9	84.5	90.2	23.6	29.8	34.6
	512k	16	4	**76.1**	**85.0**	**90.4**	**24.6**	**29.8**	**34.0**
Impact of Subset Size	256k	10	1	34.6	39.7	44.7	2.60	3.70	4.70
	256k	10	2	**42.1**	**49.3**	54.4	**3.70**	**5.20**	**6.30**
	256k	10	4	**42.1**	49.0	**55.1**	**3.70**	4.70	5.20
	256k	10	8	39.1	46.4	51.8	4.70	5.80	6.30
	256k	10	16	32.8	38.0	44.3	2.60	2.60	3.70

Table S1: Aachen Day-Night Localization Challenge.

For any input descriptor d_1 and d_2, their nearest neighbor d'_1 and d'_2, and their output $\mathcal{Z}_1$ and $\mathcal{Z}_2$, there are four possible scenarios $\{d'_1 \in \mathcal{Z}_1, d'_2 \in \mathcal{Z}_2\}$, $\{d'_1 \notin \mathcal{Z}_1, d'_2 \in \mathcal{Z}_2\}$, $\{d'_1 \in \mathcal{Z}_1, d'_2 \notin \mathcal{Z}_2\}$, $\{d'_1 \notin \mathcal{Z}_1, d'_2 \notin \mathcal{Z}_2\}$, each with different probability distributions. Since $\mathcal{Z}_1$ and $\mathcal{Z}_2$ are sampled using the ω-SM, we have shown above that $Pr(\mathcal{Z}_1|d'_1) \leq e^{\epsilon} Pr(\mathcal{Z}_2|d'_2)$ holds for all the four scenarios, and given Eq. (S7), we have $Pr(\mathcal{Z}_1|d_1) \leq e^{\epsilon} Pr(\mathcal{Z}_2|d_2)$. This means that LDP-FEAT satisfies ϵ-LDP.

S2. Additional Results

S2.1. Aachen Night Localization

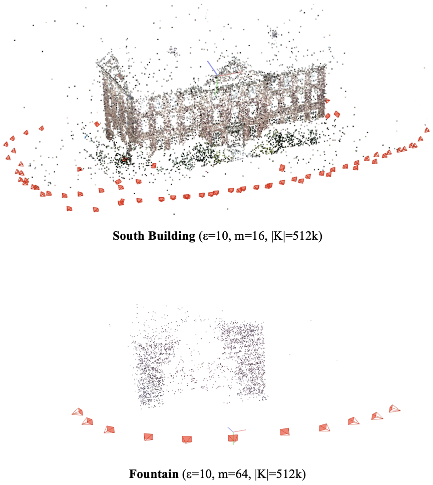
Similarly to the Tab.2 of the main paper, we report in Tab. S1 the localization accuracy for night-time queries in the Aachen Day-Night localization challenge. Overall, we observe a degradation of performance compared to the daytime queries. This is mainly because our database $\mathcal{K}$ was built from the Aachen reference images which contain daytime images only. As aforementioned, this causes a large quantization error Δd in LDP-FEAT, which certainly enhances privacy protection but compromises the utility. We leave the pursuit of a better privacy-utility trade-off for night-time localization as a future work.

S2.2. Structure-from-Motion

We further demonstrate the utility of LDP-FEAT on Structure-from-Motion, as shown in Fig. S1. We adopt COLMAP [2] for SfM by customizing its feature extraction and matching using LDP-FEAT. As an indicator for SfM performance, we report the number of registered images, the number of reconstructed sparse 3D points, the average keypoint track length, and the average reprojection error.

We report results on the "South Building" and "Fountain" scene from the 3D reconstruction benchmark [3]. We first report the results for ($|\mathcal{K}| = 2^{1024}$, $\epsilon = \infty$, $m = 1$). This corresponds to the oracle setting where only the raw descriptor is sent without any privacy protection, and which serves a performance upper bound. We then use a dictionary with $512k$ descriptors, *i.e.* ($|\mathcal{K}| = 512k$, $\epsilon = \infty$, $m = 1$) where the quantization step, *i.e.* mapping the raw descriptor d to its nearest neighbor d', introduces a degree of privacy

Scene	Dict. Size	Privacy	# Desc.	Reg.	Sparse	Track	Reproj.
	$\lvert\mathcal{K}\rvert$	ϵ	m	Images	Points	Length	Error
South Building	2^{1024}	∞	1	128	110,714	5.66	1.29
	512k	∞	1	128	62,194	4.85	1.12
	512k	10	4	88	8,668	3.56	0.85
	512k	10	8	75	10,554	3.66	1.05
	512k	10	16	123	25,451	3.62	0.89
	512k	10	32	123	26,760	3.52	0.94
	512k	10	64	124	21,596	3.31	1.28
Fountain	2^{1024}	∞	1	11	15,332	4.42	2.82
	512k	∞	1	11	8,612	3.80	2.39
	512k	10	4	11	983	2.86	1.26
	512k	10	8	11	1,827	2.98	1.35
	512k	10	16	11	2,598	3.03	1.50
	512k	10	32	11	3,078	3.02	1.52
	512k	10	64	11	3,242	2.89	1.41



South Building (ε=10, m=16, |K|=512k)

Fountain (ε=10, m=64, |K|=512k)

Figure S1: Local Feature Evaluation Benchmark. Structure-from-Motion results using LDP-FEAT with different configurations.

Dim	**Success Rate (%)**			
m	N=50	N=20	N=10	N=5
4	93.89	95.60	96.61	97.46
16	87.27	90.54	92.73	94.38

Table S2: Intersecting Adversarial Subspaces.

Dim=2	Dim=4	Dim=8	Dim=16
97.42%	94.30%	85.46%	73.03%

Table S3: Clustering Attack Collisions.

protection and thus degrades the performance accordingly. Next, we fix $|\mathcal{K}| = 512k$ and $\epsilon = 10$, while increasing m from 4 to 64. The performance varies, and we observe that m=32 yields the best performance. Overall, one observes that good SfM results are obtained from LDP-FEAT under different settings; in particular, most of the cameras are successfully registered, despite the reconstructed points being sparser. We demonstrate the qualitative reconstruction results in Fig. S1.

S3. Analysis of Assumptions

S3.1. Intersecting Adversarial Subspaces

As discussed in the paper, our proposed Database and Clustering attacks are based on the following key empirical assumption: a low-dimensional hybrid adversarial affine subspace D likely only intersects the high-dimensional descriptor manifold at $\frac{m}{2}$+1 points corresponding to the original descriptor d and the adversarial descriptors $\{a_1, ..., a_{\frac{m}{2}}\}$ that were sampled from the database. Here, we generate subspaces for 100K descriptors and report how often our assumption holds, i.e., for each subspace, we select the top N database descriptors closest to the subspace, and match them to the $\frac{m}{2}$+1 descriptors forming the subspace. Using the standard ratio test (>0.8) we report the percentage of successful matches in Tab. S2. The high success rates empirically validate our assumption regarding the rareness of intersections beyond the $\frac{m}{2}$+1 forming descriptors. We also note that our assumption is implied by the success of feature matching in [1].

S3.2. Clustering Attack Collisions

For our clustering attack, we assume that the attacker does not have access to the database of real-world descriptors W from which adversarial descriptors $a_{i=1,...,\frac{m}{2}}$ for subspace D are sampled, but does have access to an additional set of adversarial affine subspaces $\mathcal{Q}$ that were lifted with the same database W. We can identify the subset of subspaces $\mathcal{Q}' \in \mathcal{Q}$ that were lifted using one or more of the same adversarial descriptors as D, by noting that each subspace in $\mathcal{Q}'_j \in \mathcal{Q}'$ intersects with D. Assuming that $\mathcal{Q}$ is sufficiently large such that all a_i's were used to lift at least one of the subspaces in $\mathcal{Q}'$, this indicates that the minimal point-to-subspace distance $\min_j \texttt{dist}(a_i, Q'_j) = 0$ for i=$\{1, ..., \frac{m}{2}\}$. On the other hand, since $\mathcal{Q}'$ is selected without any knowledge or specific treatments on d, it is expected that $\min_j \texttt{dist}(d, Q'_j) \gg 0$. In this discrepancy lies the crux of our attack – while $\min_j \texttt{dist}(\hat{a}_i, Q'_j) > 0$ for our estimates of a_i, $\hat{a}_i$, we expect that $\min_j \texttt{dist}(\hat{d}, Q'_j) \gg \min_j \texttt{dist}(\hat{a}_i, Q'_j)$. Hence, we compute the score s_i for

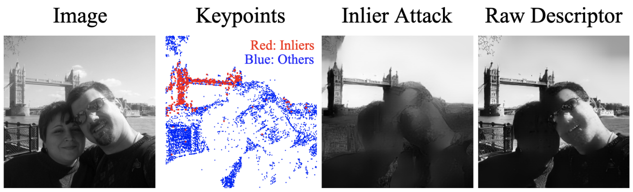


Figure S2: Inlier Attack.

each $\hat{d}_i$ as $s_i = \min_j \texttt{dist}(\hat{d}_i, Q'_j)$ and the largest s_i yields our estimate for d. We note that it is not impossible that $\mathcal{Q}'$ may contain a database descriptor that is close to d too, but the probability of such a collision is low thanks to the high dimension of descriptors. In Tab. S3, we validate this assumption by lifting all the descriptors of our 10 test images to adversarial subspaces and reporting the percentage of them that have no collisions in our attack.

S3.3. Sensitivity of Inlier Content

Since inlier correspondences emerge from RANSAC in the geometric tasks, one natural attack one may think of is leveraging these inlier features to reveal the image content; we term this as inlier attack. We note that this attack is generally applicable to all privacy protocols that are capable of geometric utility tasks where RANSAC returns inliers, *e.g*. ours and [1]. However, RANSAC inliers typically consist of only static background scenes without dynamic foreground (e.g. faces). We clarify that the privacy protection mainly targets at the foreground in both our and [1]'s problem setup, and thus the inlier attack was not a concern. Nonetheless, we perform inlier attack here and show example result in S2. As expected, the attack works only for the background bridge, but not for the foreground faces.

References

[1] Mihai Dusmanu, Johannes L Schonberger, Sudipta N Sinha, and Marc Pollefeys. Privacy-preserving image features via adversarial affine subspace embeddings. In *CVPR*, 2021. 3, 4

[2] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *CVPR*, 2016. 2

[3] Johannes Lutz Schönberger, Hans Hardmeier, Torsten Sattler, and Marc Pollefeys. Comparative Evaluation of Hand-Crafted and Learned Local Features. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. 2