Guiding Local Feature Matching with Surface Curvature
–Supplementary Material–

Shuzhe Wang[1] Juho kannala[1] Marc Pollefeys[2,3] Daniel Barath[2]

[1]Aalto University [2]ETH zurich [3]Microsoft

In this supplementary, we first prove that the curvature similarity is translation, rotation, and scaling invariant. Then, we detail the principal curvature calculation, ellipsoid fitting and radii calculation algorithms. Last, we provide additional discussion and results to our new curvature similarity estimator.

1. Proof of Curvature Similarity Invariance

As mentioned in the main paper, the normal curvature $k_n \in \mathbb{R}$ measures how curved the surface $\mathcal{S} \in \mathbb{C}^2$ is at a surface point p along a direction d. By definition [3], the curvature $k(l) = |\alpha''(l)|$, where $\alpha : \mathbf{I} \to \mathbf{R^3}$ is the curve parameterized by the arc length l, $\alpha'(l)$ is the tangent vector. Note that by a change in the orientation and position, the value of arc length l does not change. Thus, the curvature is invariant to translation and rotation. We refer the readers to [3] for the definition of curvatures and the mathematical form of the proof.

Proof: We define the curvature similarity as

$$S(k_1, k_2) = \frac{\min(|k_1|, |k_2|)}{\max(|k_1|, |k_2|)}, \ \ (0 \leq S(k) \leq 1). \tag{1}$$

(1) Since the normal curvatures are invariant to translation and rotation, the value of curvature similarity $S(k_1, k_2)$ remains the same after these operations.

(2) The normal curvature is not invariant to scaling as the scaling affects the arc length s. Following [6], we consider a point on the surface with arbitrary normal curvature k, and the normal curvature k' after the surface scaling with factor a ($a \neq 0$). The scaling alters the arc length. Thus, we have

$$k' = \frac{d\alpha'}{dal} = \frac{1}{a}\frac{d\alpha'}{dl} = \frac{1}{a}k. \tag{2}$$

As a result, both the minimal and maximum normal curvatures are changed by the same factor, and we have the curvature similarity

$$\begin{aligned} S'(k_1, k_2) &= \frac{\min(\left|\frac{1}{a}k_1\right|, \left|\frac{1}{a}k_2\right|)}{\max(\left|\frac{1}{a}k_1\right|, \left|\frac{1}{a}k_2\right|)} \\ &= \frac{\frac{1}{a}\min(|k_1|, |k_2|)}{\frac{1}{a}\max(|k_1|, |k_2|)} = S(k_1, k_2). \end{aligned} \tag{3}$$

Therefore, after scaling, the value of curvature similarity is unchanged and $S(k_1, k_2)$ is invariant to translation, rotation, and scaling.

2. Principal Curvature Calculation

The principal curvatures are obtained by calculating the partial derivative of k_n w.r.t. λ and making it zero. Where k_n is

$$k_n = \frac{L + 2M\lambda + N\lambda^2}{E + 2F\lambda + G\lambda^2}. \tag{4}$$

from the main paper. Then, we have the following equations:

$$k_1 k_2 = K = \frac{LN - M^2}{EG - F^2}, \tag{5}$$

$$k_1 + k_2 = 2H = \frac{LG - 2MF + NE}{EG - F^2}, \tag{6}$$

where K, H are the Gaussian and mean curvatures. The two real roots are the estimated principal curvatures at $\mathbf{p}_i = (u, v, Z_i(u, v)) \in \mathbf{P}_i$, where

$$k_1, k_2 = H \pm \sqrt{H^2 - K}, \tag{7}$$

where H is the mean curvature and K is the Gaussian curvature in the main paper.

3. Ellipsoid Fitting with Constraints

Similar to the ellipsoid fitting approach in [4], we show that the general function for an ellipsoid with point $\mathbf{p}' =$

$[0,0,0]^{\mathrm{T}}$ on the ellipsoid is

$$\begin{aligned} ax^2+by^2+cz^2+2dyz+2exz+ \\ 2fyz+2gx+2hy+2iz+j=0. \end{aligned} \tag{8}$$

and with the constraint

$$4M-N^2=1, \tag{9}$$

where

$$\begin{aligned} N &= a+b+c, \\ M &= ab+bc+ac-d^2-e^2-f^2, \end{aligned}$$

For each point $\mathbf{p_i}=[u_i,v_i,z_i]^{\mathrm{T}}$ in the point set, it should satisfy Eq. (8) with the constraint Eq. (9). Then, we define

$$\mathbf{X}_i=[u_i^2,v_i^2,z_i^2,2v_iz_i,2u_iz_i,2u_iv_i,2u_i,2v_i,2z_i]^{\mathrm{T}}, \tag{10}$$

and formulate a least square fitting problem that minimizes the algebraic distance

$$min||\mathbf{C}\mathbf{v}||^2 \quad s.t. \quad 4M-N^2=1 \tag{11}$$

Where $\mathbf{C}$ is coefficient matrix

$$\mathbf{C}=[\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3,...,\mathbf{X}_n] \tag{12}$$

and

$$\mathbf{v}=[a,b,c,d,e,f,g,h,i]^{\mathrm{T}} \tag{13}$$

For constraint Eq. (9), we write it in a matrix form that

$$\mathbf{v_1}^{\mathrm{T}}\mathbf{K_1}\mathbf{v_1}=1 \tag{14}$$

Where $\mathbf{v_1}=[a,b,c,d,e,f]$, and $\mathbf{K_1}$ is a $6*6$ matrix that

$$\mathbf{K_1}=\begin{bmatrix} -1 & 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 \end{bmatrix} \tag{15}$$

Taking all variables into consideration, we rewrite the condition function as $\mathbf{v}^{\mathrm{T}}\mathbf{K}\mathbf{v}=1$, and define

$$\mathbf{K}=\begin{bmatrix} \mathbf{K}_1 & \mathbf{0}_{6\times 3} \\ \mathbf{0}_{3\times 6} & \mathbf{0}_{3\times 3} \end{bmatrix} \tag{16}$$

Minimizing the Eq. (11) becomes solving the equation with constraint by the Lagrange multiplier approach

$$\mathcal{L}(\mathbf{v},\lambda)=||\mathbf{C}\mathbf{v}||^2-\lambda(\mathbf{v}^{\mathrm{T}}\mathbf{K}\mathbf{v}-1) \tag{17}$$

where λ is the Lagrange multiplier, and we solve the equation by setting $\partial\mathcal{L}/\partial\mathbf{v}=0$, $\partial\mathcal{L}/\partial\lambda=0$. The following equations are obtained

$$\mathbf{C}\mathbf{C}^{\mathrm{T}}\mathbf{v}=\lambda\mathbf{K}\mathbf{v} \tag{18}$$

and

$$\mathbf{v}^{\mathrm{T}}\mathbf{K}\mathbf{v}=1 \tag{19}$$

Eq. (18) is a general eigenvalue system, $\mathbf{v}$ and λ can be solved by the Eigen decomposition. To obtain the semi-axes of the ellipsoid, we need to determine the centre of the ellipsoid first. Recall from the main paper that we have the matrix form of an ellipsoid

$$\mathbf{x}\mathbf{Q}\mathbf{x}^{\mathrm{T}}+2\mathbf{P}\mathbf{x}^{\mathrm{T}}+\mathbf{R}=0 \tag{20}$$

where

$$\mathbf{Q}=\begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}, \quad \mathbf{P}=\begin{bmatrix} g & h & i \end{bmatrix}$$

Assuming that an ellipsoid is centred at the origin and moves to centre $\mathbf{c}=(cx,cy,cz)$, We compute $\mathbf{c}$ by

$$\mathbf{c}=\mathbf{Q}^{-1}\mathbf{P} \tag{21}$$

At the same time, the length of semi-axes in an ellipsoid remains the same. Thus, we have

$$(\mathbf{x}-\mathbf{c})^{\mathrm{T}}\mathbf{Q}(\mathbf{x}-\mathbf{c})-\mathbf{c}^{\mathrm{T}}\mathbf{Q}\mathbf{c}+\mathbf{R}=0 \tag{22}$$

Divide by $\mathbf{c}^{\mathrm{T}}\mathbf{Q}\mathbf{c}-\mathbf{R}$, we have

$$(\mathbf{x}-\mathbf{c})^{\mathrm{T}}\mathbf{U}(\mathbf{x}-\mathbf{c})=1 \tag{23}$$

where

$$\mathbf{U}=\frac{\mathbf{Q}}{\mathbf{c}^{\mathrm{T}}\mathbf{Q}\mathbf{c}-\mathbf{R}} \tag{24}$$

Since Eq. (23) is the standard form of an ellipsoid with the centre $\mathbf{c}$, the semi-axes $R=(\alpha,\beta,\gamma)$ of the ellipsoid are the square root of reciprocals of eigenvalues of $\mathbf{U}$.

4. Additional Baselines

We consider baselines that directly encode the geometric embeddings from the monocular depth estimator and curvature estimator to enhance the features for matching. Specifically, we conduct experiments with the QuadTree [9] matcher and the results without finetuning are presented in Tab. 1. For $\oplus F^D\in\mathbb{R}^{2d}$, we concatenate the depth feature extracted from the 2nd-final output layer of the depth estimator with the image features obtained from QuadTree for the matching module. For $+F^D\in\mathbb{R}^{d}$, the two features are summed. For $\oplus F^C\in\mathbb{R}^{d+1}$, we concatenate the one-dimensional curvature value from SCE and image features to compute the matching confidence matrix. For $+\sin(F^C)\in\mathbb{R}^{d}$, the curvature feature is processed with sinusoidal encoding and then summed with the image features. Our method achieves better performance without any fine-tuning.

Method	ScanNet [2]				MegaDepth [5]			
	AUC@5	@10	@20	P(%)	AUC@5	@10	@20	P(%)
$\oplus F^D$	6.7	15.9	28.3	75.5	47.3	64.1	76.6	**98.6**
$+ F^D$	5.3	12.4	22.3	71.2	46.8	63.5	76.3	**98.6**
$\oplus F^C$	24.4	44.3	61.6	**89.6**	52.8	68.5	81.6	98.5
$+ \sin(F^C)$	24.0	43.7	61.0	89.5	52.7	69.1	81.5	98.5
Ours	**25.2**	**45.5**	**62.5**	89.4	**54.0**	**70.3**	**82.1**	97.9

Table 1. **Additional Baselines**. Additional baselines that directly encode the depth features or curvature features to the QuadTree matcher. The best results are bold.

5. Additional Quantitative Results

As mentioned in the main paper, default input sizes differ among the matching approaches and benchmarks. Tab. 2 shows the input image size for the outdoor datasets. Additionally, we also report the AUC of the latest matcher ASpanFormer [1] with CSE on MegaDepth [5] and ScanNet [2]. The longest dimension of image size in MegaDepth are rescaled to 1152 in Tab. 3 by the default setting. Our adds-on curvature similarity estimator still shows performance improved on this SoTA matcher.

Input Size	**YFCC [10]**	**MegaDepth [5]**
SuperGlue [7]-based	1600	1200
LoFTR [8]-based	840	1200
Quadtree [9]-based	832	832

Table 2. **Default image sizes of different matchers.** The numbers are the longest dimension of the rescaled image size.

Methods	**ScanNet [2]**			**MegaDepth [5]**		
	AUC@5°	**@10 °**	**@20 °**	**AUC@5°**	**@10 °**	**@20 °**
ASpanF [8]	25.5	46.0	63.3	55.3	71.5	83.1
ASpanF + CSE	**25.9**	**46.5**	64.5	56.1	**72.0**	**83.4**
ASpanF + CSE (w/ FT)	25.8	**46.5**	**64.6**	**56.4**	**72.0**	83.2

Table 3. **Matching results with ASpanFormer + CSE.** We report the matching results (AUC with different thresholds) on AspanFormer. The best performance is marked in bold.

6. Additional Qualitative Results

We show in Fig. 2 more qualitative matching results. The number of final matches and pose errors are also presented in the figures. In Fig. 3, we visualize the depth maps from different conditions and depth predictors. As described in the main paper, although the ground truth depth provides the most accurate 3D information, missing parts in the depth map may lower the final matching accuracy. Some strategies, such as omitting the invalid depths, might help to improve the performance. However, our experiments show that adopting lightweight monocular depth predictors is sufficient to enhance the matching results. In Fig. 4, we visualize the curvature and depth map of image pairs for both indoor and outdoor scenes on QuadTree [9] + CSE to better explain our curvature similarity-based approach.

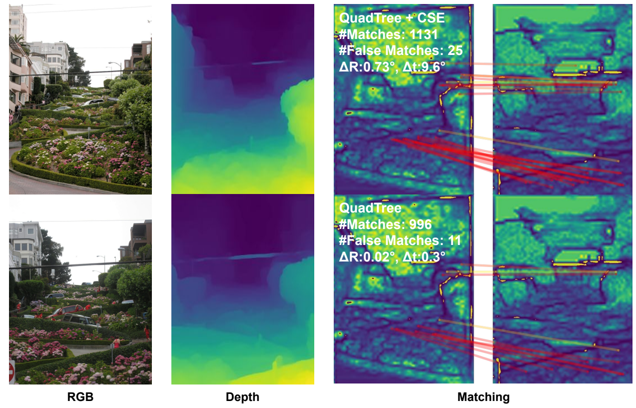

Figure 1. **Failure case** with wrong matches marked with red. Correct matches are not drawn.

7. Failure Case

As we discussed in the Limitation section of the main paper, the matching performance may be downgraded if the depth is inaccurate or no clear surface is observed. We illustrate a specific failure case in Fig 1, where the images show mostly plants without clear surfaces for curvature extraction. Adding the SCE in this scenario increases the number of incorrect matches and results in poorer pose estimation. It is crucial to note, however, that such scenarios are relatively rare. Despite the isolated failure case, our proposed approach still yields substantial improvements in terms of accuracy across various matchers and datasets, as demonstrated in Tables 1 and 2 of the main paper.

References

[1] Hongkai Chen, Zixin Luo, Lei Zhou, Yurun Tian, Mingmin Zhen, Tian Fang, David Mckinnon, Yanghai Tsin, and Long Quan. Aspanformer: Detector-free image matching with adaptive span transformer. In *Proceedings of the European Conference on Computer Vision*, 2022. 3

[2] Angela Dai, Angel X Chang, Manolis Savva, Maciej Halber, Thomas Funkhouser, and Matthias Nießner. Scannet: Richly-annotated 3d reconstructions of indoor scenes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5828–5839, 2017. 3

[3] M.P. do Carmo. *Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition*. Dover Books on Mathematics. Dover Publications, 2016. 1

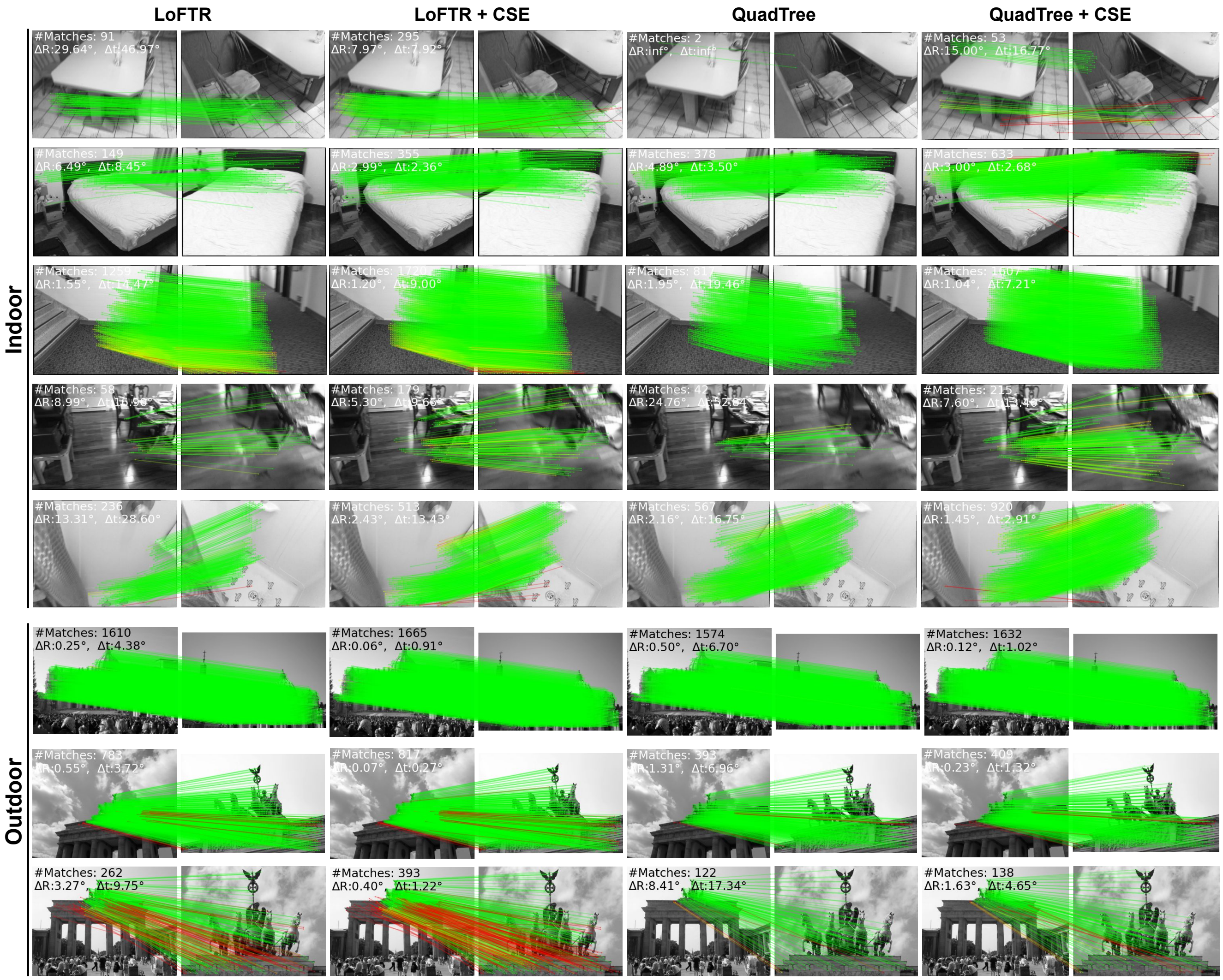

Figure 2. **Additional qualitative matching results**. We provide more matching visualization results on indoor and outdoor datasets with our CSE plug-into LoFTR [8] and QuadTree [9]. Our method consistently obtains more matches and better accuracy independently of the matcher it is combined with.

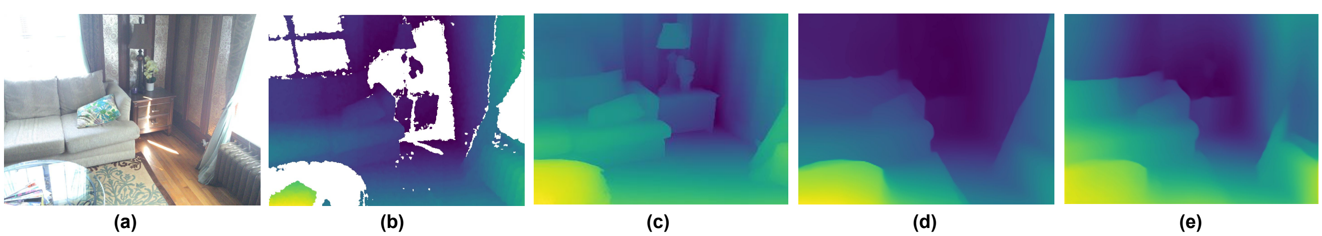

Figure 3. **Depth map with different conditions**. We visualize the depth maps with different predictors. (a) RGB image; (b) Ground truth depth; (c) Depth map from MiDAS + ViT; (d) Depth map from MiDAS + ResNet101; (e) Depth map from MiDAS (real-time).

[4] Qingde Li and J.G. Griffiths. Least squares ellipsoid specific fitting. In *Geometric Modeling and Processing, 2004. Proceedings*, pages 335–340, 2004. 1

[5] Zhengqi Li and Noah Snavely. Megadepth: Learning single-view depth prediction from internet photos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2041–2050, 2018. 3

[6] John Rugis and Reinhard Klette. A scale invariant surface curvature estimator. In *Pacific-Rim Symposium on Image and Video Technology*, pages 138–147. Springer, 2006. 1

[7] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superglue: Learning feature

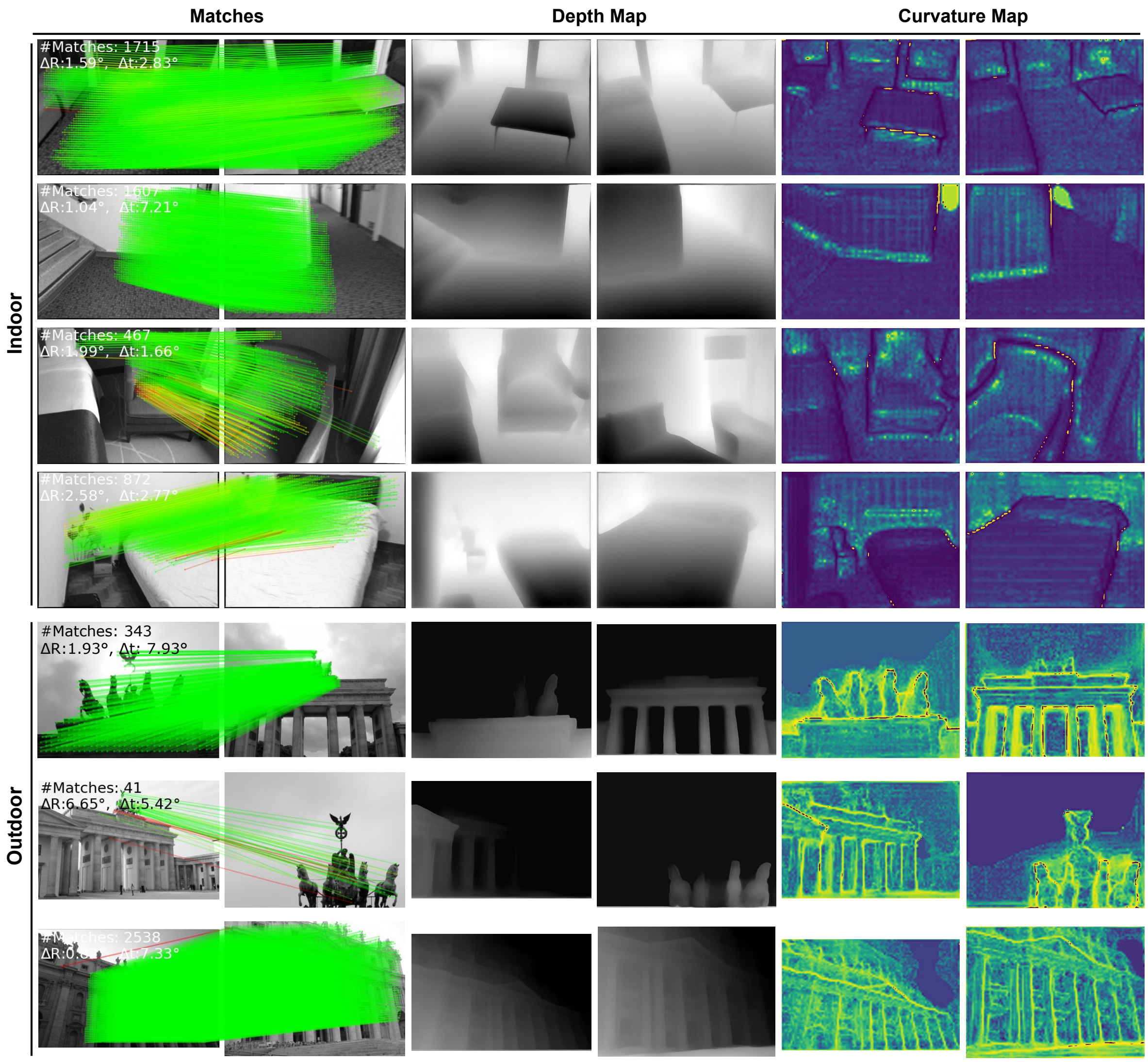

Figure 4. **Curvature and Depth Maps**. We visualize the curvature maps and predicted depth maps of image pairs on indoor/outdoor scenes with QuadTree + CSE. The outdoor curvature maps are presented with a reversed colourmap compared to the indoor ones for better visualization.

matching with graph neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4938–4947, 2020. 3

[8] Jiaming Sun, Zehong Shen, Yuang Wang, Hujun Bao, and Xiaowei Zhou. LoFTR: Detector-free local feature matching with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021. 3, 4

[9] Shitao Tang, Jiahui Zhang, Siyu Zhu, and Ping Tan. Quadtree attention for vision transformers. In *International Conference on Learning Representations*, 2022. 2, 3, 4

[10] Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. Yfcc100m: The new data in multimedia research. *Communications of the ACM*, 59(2):64–73, 2016. 3