Resource Allocation and Optimization
Volume 1 · Number 1 · 2026

Nine Disparity Anchors, Two Local Models: Geometry and Appearance Beyond the Fitting Points

Adrian Barnes[1] · Jeffrey Webb[2] · Benjamin Lewis[3]

[1] Norway · [2] Belgium · [3] Austria

Original publication: October 31, 2026 | Article 609

Abstract

Nine known disparity values can support either a constant surface or a more flexible local fit. Across 48 eligible tiles in two Middlebury scenes, an affine fit reduced held disparity MAE from 2.8826 to 2.0986 pixels and three-view mean RGB error from 20.9466 to 19.0452 channel levels. The average gains concealed exceptions: four tiles worsened geometrically, six worsened photometrically, and eight had opposed change directions across the two outcomes. Better fitting of sparse geometry therefore needs separate evaluation beyond the anchors and on the appearance measurement it is intended to support.

Executed sparse-geometry comparison on two static Middlebury scenes. Source reference disparity supplies privileged anchors; no learned depth or camera-motion transfer method is evaluated.

Nine Disparity Anchors, Two Local Models: Geometry and Appearance Beyond the Fitting Points

> Executed sparse-geometry comparison on two static Middlebury scenes. Source reference disparity supplies privileged anchors; no learned depth or camera-motion transfer method is evaluated.

Abstract

Nine known disparity values can support either a constant surface or a more flexible local fit. Across 48 eligible tiles in two Middlebury scenes, an affine fit reduced held disparity MAE from 2.8826 to 2.0986 pixels and three-view mean RGB error from 20.9466 to 19.0452 channel levels. The average gains concealed exceptions: four tiles worsened geometrically, six worsened photometrically, and eight had opposed change directions across the two outcomes. Better fitting of sparse geometry therefore needs separate evaluation beyond the anchors and on the appearance measurement it is intended to support.

1. Spending observations and spending parameters

Sparse geometric observations can be used in several ways. A simple representation may average them into one local disparity value, while a more flexible representation can fit a spatial trend. The number of supplied observations can remain identical even though model complexity changes. A resource comparison should therefore distinguish the observation budget from the number of fitted coefficients and from the later cost of evaluating the representation.

We ask whether an affine disparity fit improves held geometry and multi-view photometric agreement over a constant fit when both receive the same nine disparity anchors per image tile. The question has two endpoints because a geometrically better approximation need not produce lower encoded-image error in every tile. Texture, occlusion, interpolation, and the distribution of geometric residuals can affect the relationship between those measurements.

Meng, 2026 addresses scale-calibrated camera-motion transfer through a gauge-invariant parallax number. That focus motivates examining how a limited local geometry representation behaves when used across view separations. Our experiment is narrower: two elementary disparity models receive known source anchors and are evaluated on static released views. We do not calculate the target parallax number, transfer a camera trajectory, or run its model.

The study separates a guaranteed training property from the actual held findings. An affine least-squares model contains the constant model as a special case, so its anchor squared error cannot be larger when fitted successfully. That property does not guarantee lower absolute error at unobserved pixels. Nor does it guarantee that a warp based on the fit yields a more faithful image measurement. Both downstream comparisons must be calculated.

2. Local geometry as a measurement device

The public scenes and disparities come from Middlebury's structured-light stereo work (Scharstein and Szeliski, 2003), within the broader stereo correspondence evaluation framework of Scharstein and Szeliski, 2002. These resources provide a controlled setting with real image texture and known reference geometry. They do not represent a random sample of driving scenes, moving subjects, or the conditions of a generated-video evaluation.

Multi-view geometry explains why correspondence and depth assumptions matter to image alignment (Hartley and Zisserman, 2004). Reconstruction pipelines such as COLMAP organize substantially richer

geometric estimation (Schonberger and Frahm, 2016). Here we avoid fitting a complete reconstruction system. The affine surface is a linear function of reference-image coordinates in disparity units; calling it an affine disparity fit does not assert that a physical three-dimensional plane has been recovered.

Bundle adjustment and sparse optimization address global geometric estimation questions beyond this local approximation (Triggs et al., 2000; Lourakis and Argyros, 2009). Camera-pose estimation also requires its own measurements and model, as illustrated by EPnP (Lepetit et al., 2009). We estimate none of those quantities. The view-separation multiplier is fixed by the captured rectified-view arrangement, leaving the comparison focused on how the same nine disparity values are represented spatially.

Anchor placement is another distinct design question. Sensor selection and placement can be optimized for an information objective (Joshi and Boyd, 2009; Krause et al., 2008). Our anchors are fixed on a regular grid rather than selected adaptively from full disparity or image error. The experiment consequently compares model forms at one observation layout; it does not establish that the nine locations are optimal or that acquiring them would have equal physical cost.

Learned monocular depth methods such as MiDaS and self-supervised depth estimation concern how geometry is inferred from observations (Ranftl et al., 2019; Godard et al., 2018). In this analysis, the nine anchors are taken from provided reference disparity. Their information is deliberately privileged and their acquisition uncertainty is unmeasured. Results should not be interpreted as the performance of a predicted-depth pipeline simply because both outputs use disparity units.

3. A common anchor grid and an explicit retained cohort

The source contains two static nine-view scenes, cones and teddy. We use im2 as the reference and divide each 375-by-450 image into a five-by-six grid of nonoverlapping 75-by-75 tiles, producing sixty candidate tiles. Source disparity is decoded by dividing its stored values by four. Positive reference values define valid geometry for the present measurement.

Within every tile, the nine anchor locations form a regular three-by-three grid with local row and column indices 12, 37, and 62. All nine must have valid source disparity. No anchor is moved to a favorable nearby location when it is invalid. This requirement excludes ten candidate tiles. Every remaining tile must also provide at least 1,000 valid nonanchor geometry pixels for held evaluation.

Photometric comparisons use three fixed target views, indices zero, four, and eight. A tile is retained only when each target comparison contains at least 1,000 common nonanchor pixels whose constant and affine sampling coordinates are both in bounds. Two further tiles fail this joint view requirement. The final cohort has 48 tiles: 21 from cones and 27 from teddy. It is therefore a selected cohort with unequal scene contributions, not the complete image area or an equal-scene average.

The in-bounds requirement is part of what the image comparison measures. In the least-covered retained tile-view cell, common pixels are about 24.46% of that tile's valid nonanchor geometry pixels. We do not infer full-tile appearance quality from such a cell. Saved pixel counts make the coverage boundary visible alongside the conditional error difference.

For the constant model, all tile coordinates receive the mean of the nine anchor disparities. The affine model fits an intercept, a horizontal coefficient, and a vertical coefficient by ordinary least squares on the same nine observations. Image coordinates are centered at 37 and divided by 25 for numerical scaling. Those transformations do not add geometry observations or turn the comparison into an equal-parameter-budget experiment.

The linear algebra uses the local NumPy and SciPy libraries (Harris et al., 2020; Virtanen et al., 2020). Full rank is checked for the three-column fitting design, and fitted anchor squared error is checked against the nested constant fit. Numerical linear algebra and floating-point considerations remain relevant to these checks (Higham, 2002; Goldberg, 1991). No coefficients are quantized or rounded before prediction, and no runtime benchmark is performed.

The withholding operation is analytical rather than physical. We possess the released disparity array but expose only the nine declared values to each fit. Remaining positive values become evaluation targets. A real sparse sensor could have noisy, delayed, or strategically located observations instead, and the fixed grid might intersect different surfaces. Those effects are not estimated by the current calculation. Similarly, the supplied disparity values have their source encoding and resolution; treating them as the reference for a paired comparison does not establish error-free physical depth. The fitted policies and uncertainty summaries remain conditional on that captured reference resource.

4. Two held outcomes on their own scales

The geometry outcome is absolute disparity error averaged over all valid nonanchor pixels in a retained tile. Anchor pixels are excluded from this score, preventing a lower fitting residual from being reported as held performance. The full source disparity is used only as the held target outside those nine locations; it does not refit either model or choose its coefficients.

For a target view k, the horizontal sampling coordinate subtracts predicted disparity multiplied by (k-2)/4. Vertical coordinates remain the reference rows. Bilinear interpolation samples each encoded RGB channel, and its absolute difference from the reference image is averaged over the common valid mask. We perform no color calibration, occlusion certification, or conversion to a perceptually uniform color space.

The three target-view errors are averaged equally within a tile, and tiles then receive equal weight. The resulting endpoint differs from a pooled error over all valid pixels or views. It describes three-view mean photometric error for the retained tiles, not a guarantee about the largest error at any pixel. No predicted disparity is clipped; none of the retained affine held predictions happens to be negative in this execution.

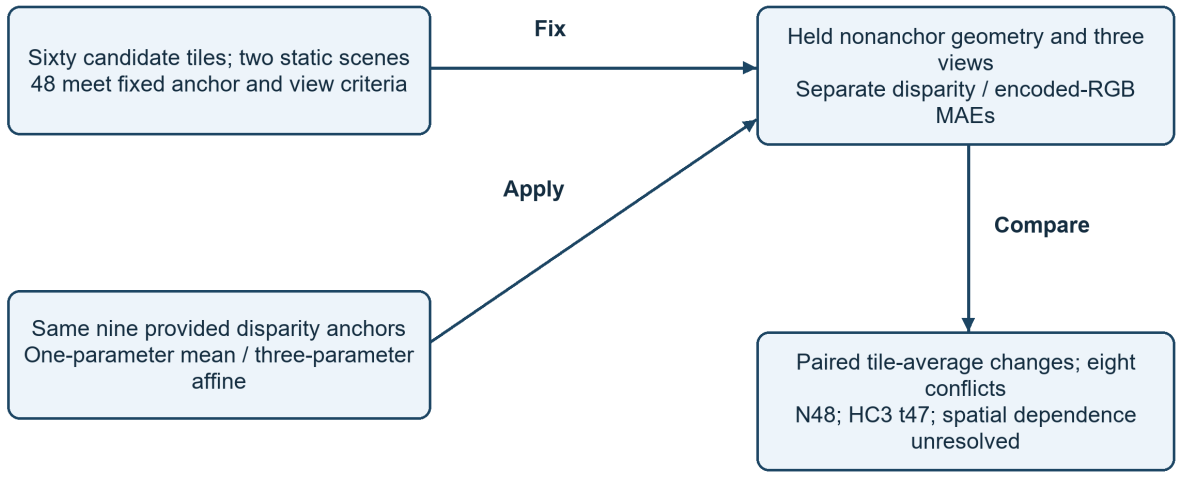


Figure 1. *A fixed nine-point geometry budget feeds nested constant and affine fits. Nonanchor disparity error and paired common-pixel image error are evaluated separately before tile-level contrasts are summarized.*

Two intercept-only regressions estimate affine-minus-constant differences across the 48 tiles. Model 1 concerns disparity MAE in pixels; Model 2 concerns the tile's three-view mean RGB error in encoded channel levels. Both use HC3 standard errors and t47 intervals, with no controls or fixed effects. Intercept-only R-squared is zero and should not be read as a diagnostic of the local surface fits.

Robust variance estimation addresses heterogeneous row variation (White, 1980), while dependence requires its own assumptions (Cameron and Miller, 2015). Adjacent tiles share scene content, acquisition conditions, and boundaries. HC3 does not resolve that spatial dependence, and two scenes provide insufficient replication for broad across-scene inference. The reported intervals are conditional working summaries for this selected tile cohort, rather than certified population coverage or uncertainty for future camera-motion transfer.

Table 1. *Paired tile-level mean changes, N=48, HC3 standard errors and t47 confidence intervals. Models contain an intercept only. Geometry and image outcomes use different units and different masks; neither regression estimates a causal effect on physical scene quality.*

Model	Term	Estimate (SE)	95% CI	N
M1	Mean outcome/contrast	-0.7840 (0.1063)	[-0.9979, -0.5702]	48
M1: fit	R-squared = 0.0000	-	-	48
M2	Mean outcome/contrast	-1.9013 (0.3657)	[-2.6371, -1.1655]	48
M2: fit	R-squared = 0.0000	-	-	48

Model labels: M1 = Affine minus constant held disparity MAE; M2 = Affine minus constant three-view mean encoded-RGB MAE.

5. Average improvement without local domination

Mean anchor squared error falls from 15.907504 under the constant fit to 7.427132 under the affine fit. This decrease follows the nested least-squares objective and describes the provided fitting points. The held geometry comparison is separate: mean nonanchor disparity MAE falls from 2.882589 to 2.098567 pixels. Model 1 estimates -0.784022 pixels, with standard error 0.106297 and interval [-0.997865, -0.570180].

Mean three-view RGB error falls from 20.946577 to 19.045241 encoded channel levels. The Model 2 difference is -1.901336, with standard error 0.365746 and interval [-2.637121, -1.165550]. Those magnitudes show the benefit in the declared image-error unit. They are not percentages of perceptual quality, a video-generation score, or a measure of driving safety.

The affine fit improves held geometry on 44 tiles and worsens it on four. It improves the photometric endpoint on 42 and worsens it on six. Eight tiles have opposite signs of change across geometry and appearance. Thus the aggregate improvements do not make the endpoints interchangeable, and the richer fit does not dominate at every retained location.

This distinction fits the broader measurement concern that a proxy and its intended construct need separate validation (Carroll et al., 2006). The scale of an outcome also constrains its interpretation (Stevens, 1946). Disparity pixels and encoded RGB levels cannot be combined into an unspecified improvement index without choosing a tradeoff rule. We retain two tables of paired outcomes in the analysis records rather than selecting whichever favors a particular tile.

Method-comparison reasoning emphasizes inspecting differences on their actual scales (Bland and Altman, 1986). The eight opposed signs are one concise manifestation of disagreement, but they do not identify its mechanism. Occlusion, texture gradients, mixed-depth structure, interpolation, and view-dependent image changes are plausible contributors. None is separately measured well enough here to attribute the local failures to a single cause.

6. Appearance agreement is not a perceptual verdict

Structural similarity evaluates image structure in a way different from raw channel error (Wang et al., 2004). Learned perceptual similarity provides another representation of image differences (Zhang et al., 2018), and DISTS combines structural and textural considerations (Ding et al., 2020). FSIM uses feature similarity for image-quality measurement (Lin Zhang et al., 2011). These alternatives clarify what the present endpoint omits; none is computed in this experiment or used to choose a favorable reporting metric.

The image sampling also compares captured views of static scenes rather than newly rendered views. A method could lower the measured reference-to-target RGB error while still producing visible artifacts in a generated image. Conversely, a geometric improvement at an occluded point need not make the sampled target color agree with the reference. These limitations prevent translating a lower photometric number into a claim of perceptually superior synthesis.

One source of generalization optimism would be choosing anchor layouts, view subsets, or tile exclusions after inspecting the comparative losses. Validation structure should reflect the intended transfer question (Roberts et al., 2017), and model-selection bias can arise when evaluation outcomes guide such choices (Varma and Simon, 2006). Here the grid, three views, and minimum pixel counts define the calculation; exclusions are based on available geometry and paired sampling support, not on whether the affine error is favorable.

Influential tiles and selected scene composition still matter (Cook, 1977). A tile near a depth discontinuity can behave differently from a tile dominated by a smooth surface. The current study preserves the failures but does not fit and validate a classifier for predicting them. Building a reliable gate that decides when to use the richer representation would require another development and held evaluation design.

7. What the equal-anchor comparison establishes

Dataset records retain the sixty candidates, twelve exclusion reasons, nine observed anchors per retained tile, fitted coefficients, view-level errors, and paired regression inputs. Documentation supports understanding those roles and source restrictions (Gebru et al., 2021); reusable provenance supports locating and interpreting the saved calculations (Wilkinson et al., 2016). The richer representation uses more coefficients, so its benefit should not be described as free computation or an equal-storage gain.

At this fixed sparse observation budget, an affine disparity fit improves both held geometry and average image alignment in the selected two-scene cohort. Four geometric failures, six photometric failures, and eight cross-outcome direction conflicts qualify that result. The substantive answer is therefore conditional: added spatial flexibility helps on average here, while training fit alone cannot establish where it helps or whether its image-error gain corresponds to the broader camera-motion and perceptual tasks that motivated the comparison.

References

1. Harris, Charles R.; Millman, K. Jarrod; van der Walt, Stéfan J.; Gommers, Ralf; Virtanen, Pauli; Cournapeau, David; Wieser, Eric; Taylor, Julian; Berg, Sebastian; Smith, Nathaniel J.; Kern, Robert; Picus, Matti; Hoyer, Stephan; van Kerkwijk, Marten H.; Brett, Matthew; Haldane, Allan; del Río, Jaime Fernández; Wiebe, Mark; Peterson, Pearu; Gérard-Marchant, Pierre; Sheppard, Kevin; Reddy, Tyler; Weckesser, Warren; Abbasi, Hameer; Gohlke, Christoph; Oliphant, Travis E.. (2020). Array programming with NumPy. Nature. https://doi.org/10.1038/s41586-020-2649-2.

2. Lin Zhang, ; Lei Zhang, ; Xuanqin Mou, ; Zhang, D.. (2011). FSIM: A Feature Similarity Index for Image Quality Assessment. IEEE Transactions on Image Processing. https://doi.org/10.1109/tip.2011.2109730.

3. Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, IJsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-Willem; da Silva Santos, Luiz Bonino; Bourne, Philip E.; Bouwman, Jildau; Brookes, Anthony J.; Clark, Tim; Crosas, Mercè; Dillo, Ingrid; Dumon, Olivier; Edmunds, Scott; Evelo, Chris T.; Finkers, Richard; Gonzalez-Beltran, Alejandra; Gray, Alasdair J.G.; Groth, Paul; Goble, Carole; Grethe, Jeffrey S.; Heringa, Jaap; 't Hoen, Peter A.C; Hooft, Rob; Kuhn, Tobias; Kok, Ruben; Kok, Joost; Lusher, Scott J.; Martone, Maryann E.; Mons, Albert; Packer, Abel L.; Persson, Bengt; Rocca-Serra, Philippe; Roos, Marco; van Schaik, Rene; Sansone, Susanna-Assunta; Schultes, Erik; Sengstag, Thierry; Slater, Ted; Strawn, George; Swertz, Morris A.; Thompson, Mark; van der Lei, Johan; van Mulligen, Erik; Velterop, Jan; Waagmeester, Andra; Wittenburg, Peter; Wolstencroft, Katherine; Zhao, Jun; Mons, Barend. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. https://doi.org/10.1038/sdata.2016.18.

4. Carroll, Raymond J., David Ruppert, Leonard A. Stefanski, and Ciprian M. Crainiceanu. "Measurement Error in Nonlinear Models." 2006. https://doi.org/10.1201/9781420010138.

5. Virtanen, Pauli; Gommers, Ralf; Oliphant, Travis E.; Haberland, Matt; Reddy, Tyler; Cournapeau, David; Burovski, Evgeni; Peterson, Pearu; Weckesser, Warren; Bright, Jonathan; van der Walt, Stéfan J.; Brett, Matthew; Wilson, Joshua; Millman, K. Jarrod; Mayorov, Nikolay; Nelson, Andrew R. J.; Jones, Eric; Kern, Robert; Larson, Eric; Carey, C J; Polat, İlhan; Feng, Yu; Moore, Eric W.; VanderPlas, Jake; Laxalde, Denis; Perktold, Josef; Cimrman, Robert; Henriksen, Ian; Quintero, E. A.; Harris, Charles R.; Archibald, Anne M.; Ribeiro, Antônio H.; Pedregosa, Fabian; van Mulbregt, Paul; SciPy 1.0 Contributors; Vijaykumar, Aditya; Bardelli, Alessandro Pietro; Rothberg, Alex; Hilboll, Andreas; Kloeckner, Andreas; Scopatz, Anthony; Lee, Antony; Rokem, Ariel; Woods, C. Nathan; Fulton, Chad; Masson, Charles; Häggström, Christian; Fitzgerald, Clark; Nicholson, David A.; Hagen, David R.; Pasechnik, Dmitrii V.; Olivetti, Emanuele; Martin, Eric; Wieser, Eric; Silva, Fabrice; Lenders, Felix; Wilhelm, Florian; Young, G.; Price, Gavin A.; Ingold, Gert-Ludwig; Allen, Gregory E.; Lee, Gregory R.; Audren, Hervé; Probst, Irvin; Dietrich, Jörg P.; Silterra, Jacob; Webber, James T; Slavič, Janko; Nothman, Joel; Buchner, Johannes; Kulick, Johannes; Schönberger, Johannes L.; de Miranda Cardoso, José Vinícius; Reimer, Joscha; Harrington, Joseph; Rodríguez, Juan Luis Cano; Nunez-Iglesias, Juan; Kuczynski, Justin; Tritz, Kevin; Thoma, Martin; Newville, Matthew; Kümmerer, Matthias; Bolingbroke, Maximilian; Tartre, Michael; Pak, Mikhail; Smith, Nathaniel J.; Nowaczyk, Nikolai; Shebanov, Nikolay; Pavlyk, Oleksandr; Brodtkorb, Per A.; Lee, Perry; McGibbon, Robert T.; Feldbauer, Roman; Lewis, Sam; Tygier, Sam; Sievert, Scott; Vigna, Sebastiano; Peterson, Stefan; More, Surhud; Pudlik, Tadeusz; Oshima, Takuya; Pingel, Thomas J.; Robitaille, Thomas P.; Spura, Thomas; Jones, Thouis R.; Cera, Tim; Leslie, Tim; Zito, Tiziano; Krauss, Tom; Upadhyay, Utkarsh; Halchenko, Yaroslav O.; Vázquez-Baeza, Yoshiki. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. https://doi.org/10.1038/s41592-019-0686-2.

6. Triggs, Bill; McLauchlan, Philip F.; Hartley, Richard I.; Fitzgibbon, Andrew W.. (2000). Bundle Adjustment - A Modern Synthesis. Lecture Notes in Computer Science. https://doi.org/10.1007/3-540-44480-7_21.

7. Scharstein, Daniel; Szeliski, Richard. (2002). A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. International Journal of Computer Vision. https://doi.org/10.1023/a:1014573219977.

8. Cook, R. Dennis. (1977). Detection of Influential Observation in Linear Regression. Technometrics. https://doi.org/10.1080/00401706.1977.10489493.

9. Joshi, S.; Boyd, S.. (2009). Sensor Selection via Convex Optimization. IEEE Transactions on Signal Processing. https://doi.org/10.1109/tsp.2008.2007095.

10. Stevens, S. S.. (1946). On the Theory of Scales of Measurement. Science. https://doi.org/10.1126/science.103.2684.677.

11. Varma, Sudhir and Richard Simon. "Bias in error estimation when using cross-validation for model selection." *BMC Bioinformatics*, vol. 7, no. 1, 2006. https://doi.org/10.1186/1471-2105-7-91.

12. Wang, Zhou, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. "Image Quality Assessment: From Error Visibility to Structural Similarity." *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004, pp. 600-612. https://doi.org/10.1109/TIP.2003.819861.

13. Gebru, Timnit, et al. "Datasheets for datasets" *Communications of the ACM*, 2021. https://doi.org/10.1145/3458723.

14. Hartley, Richard; Zisserman, Andrew. (2004). Multiple View Geometry in Computer Vision. Cambridge University Press. https://doi.org/10.1017/cbo9780511811685.

15. Ranftl, René; Lasinger, Katrin; Hafner, David; Schindler, Konrad; Koltun, Vladlen. (2019). Towards Robust Monocular Depth Estimation: Mixing Datasets for Zero-shot Cross-dataset Transfer. https://arxiv.org/abs/1907.01341.

16. Roberts, David R., et al. “Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure.” *Ecography*, vol. 40, no. 8, 2017, pp. 913-929. https://doi.org/10.1111/ecog.02881.

17. White, Halbert. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, vol. 48, no. 4, 1980, pp. 817-838. https://doi.org/10.2307/1912934.

18. Bland, J. Martin, and Douglas G. Altman. "Statistical Methods for Assessing Agreement between Two Methods of Clinical Measurement." *The Lancet*, vol. 327, no. 8476, 1986, pp. 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8.

19. Higham, Nicholas J.. (2002). Accuracy and Stability of Numerical Algorithms: Second Edition. Society for Industrial and Applied Mathematics. https://doi.org/10.1137/1.9780898718027.

20. Ding, Keyan, et al. "Image Quality Assessment: Unifying Structure and Texture Similarity." *arXiv preprint arXiv:2004.07728*, 2020. https://arxiv.org/abs/2004.07728.

21. Goldberg, David. (1991). What every computer scientist should know about floating-point arithmetic. ACM Computing Surveys. https://doi.org/10.1145/103162.103163.

22. Meng, Zijie. (2026). ParaScale: Scale-Calibrated Camera-Motion Transfer via a Gauge-Invariant Parallax Number. https://arxiv.org/abs/2606.19805.

23. Scharstein, Daniel; Szeliski, Richard. (2003). High-accuracy stereo depth maps using structured light. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.. https://doi.org/10.1109/cvpr.2003.1211354.

24. Lepetit, Vincent; Moreno-Noguer, Francesc; Fua, Pascal. (2009). EPnP: An Accurate O(n) Solution to the PnP Problem. International Journal of Computer Vision. https://doi.org/10.1007/s11263-008-0152-6.

25. Cameron, A. Colin, and Douglas L. Miller. “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources*, vol. 50, no. 2, 2015, pp. 317-372. https://doi.org/10.3368/jhr.50.2.317.

26. Krause, Andreas; Singh, Ajit; Guestrin, Carlos. (2008). Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies. Journal of Machine Learning Research,9(8),235-284. https://www.jmlr.org/papers/v9/krause08a.html.

27. Schonberger, Johannes L.; Frahm, Jan-Michael. (2016). Structure-from-Motion Revisited. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). https://doi.org/10.1109/cvpr.2016.445.

28. Godard, Clément; Mac Aodha, Oisin; Firman, Michael; Brostow, Gabriel. (2018). Digging Into Self-Supervised Monocular Depth Estimation. https://arxiv.org/abs/1806.01260.

29. Lourakis, Manolis I. A.; Argyros, Antonis A.. (2009). SBA: A software package for generic sparse bundle adjustment. ACM Transactions on Mathematical Software. https://doi.org/10.1145/1486525.1486527.

30. Zhang, Richard, et al. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric" *arXiv preprint arXiv:1801.03924*, 2018. https://arxiv.org/abs/1801.03924.