MSPR 2x: Multisample Phase Reconstruction

A Tile-Local Spatial Upscaling Method for Real-Time Rendering

Technical Report

Abstract

This report presents MSPR 2x, a deterministic phase-reconstruction method for real-time rendering. It converts four pre-resolve MSAA color samples into four output phases per low-resolution pixel. The method is designed for tile-based GPUs: it consumes sample and neighborhood evidence before resolve, writes the 2× result directly, and avoids a persistent full-frame intermediate.

An archived nine-scene matrix compares Bilinear 2×, FSR1 EASU+RCAS, MetalFX Spatial, MSPR-Quality, and MSPR-Lite against Direct 2×. Quality and GPU time use the same scene set and execution contract. MSPR-Quality recorded the highest mean at 30.301 dB with a 3.656 ms full-path GPU mean, or 1.067 ms above Bilinear. MSPR-Lite reached 30.065 dB at 3.218 ms, only 0.629 ms above Bilinear.

Evidence boundary. The nine-scene matrix was captured by one earlier release on one iPhone 13 (iPhone14,5, iOS 18.2.1), with same-frame Direct 2× references. **Historical asset disclosure:** the assets behind the Temple and full-screen-image rows came from Apple's downloadable order-independent-transparency demo [1] and were used only for testing. They, their runtime scenes, and their raw logs are not distributed because this project does not grant rights to redistribute Apple's sample assets. Users may obtain the demo from the linked Apple page under Apple's applicable terms. The two rows and their contributions to the archived quality and full-path performance means remain historical, non-reproducible evidence; current source reproduction covers seven project-authored procedural scenes. The broader 100-frame full-scene set retains its calibrated offline-sample protocol and is not a Metal imageblock hardware dump.

Contents

1 Executive findings

1. **MSPR directly uses pre-resolve subpixel evidence.** Fixed phase coefficients and bounded content-adaptive blending reconstruct four output phases from four MSAA samples and local resolved color, with particular value on geometry edges, transparency boundaries, and highlight contours.
2. **Both MSPR profiles improve quality with small added full-path time.** Their archived nine-scene means both exceeded MetalFX, Bilinear, and FSR1. Against the 2.589 ms Bilinear full-path mean on the same matrix, MSPR-Quality adds 1.067 ms for a 1.021 dB gain, while MSPR-Lite adds 0.629 ms for a 0.785 dB gain.
3. **MSPR-Quality is the quality-first operating point.** Its archived nine-scene mean is 30.301 dB, leading MetalFX, Bilinear, and FSR1 by 0.891, 1.021, and 2.934 dB. It also leads Bicubic by 0.815 dB on the 100-frame full-scene set.
4. **MSPR-Lite is the cost-first operating point.** Its archived nine-scene mean is 30.065 dB, only 0.237 dB below MSPR-Quality. Removing threadgroup staging and the barrier yields a 3.218 ms full-path mean, 0.629 ms above Bilinear under the same workload contract.
5. **The two deployment profiles cover different budgets.** MSPR-Quality uses a 3×3 resolved neighborhood for greater content adaptivity; MSPR-Lite uses four current-pixel samples and two horizontal resolved taps in a simpler fixed six-tap path.
6. **The release includes an end-to-end reproduction entry point.** One command on a physical iPhone regenerates the seven distributed procedural-scene rows, representative raw image sets, and report-style comparison figures, with MetalFX, MSPR, FSR1, Bilinear, and Direct 2× sharing the same frame and color contract.

2 Algorithm and rendering contract

2.1 Why MSPR executes before resolve

FSR1 is a resolved-color spatial upscaler: EASU reconstructs directionally adaptive edges and RCAS sharpens the result [2, 3]. SGSR1 similarly consumes one resolved texture and combines a 12-tap Lanczos-like filter with adaptive sharpening in one mobile-oriented shader [4]. MetalFX Spatial accepts an input color texture and emits a larger output texture through Apple's runtime [5]. All three start after conventional resolve.

MSPR instead reads the four MSAA colors while they are tile-local. For output phase $p \in \{0, 1, 2, 3\}$,

$$L_p = \sum_{s=0}^{3} w_{p,s} C_s, \tag{1}$$

$$B_p = 0.5625R + 0.1875R_x + 0.1875R_y + 0.0625R_{xy}, \tag{2}$$

$$\widetilde{B}_p = \text{clip}(B_p + \gamma D, 0, 1), \tag{3}$$

$$O_p = \text{clip}\left(\widetilde{B}_p + \alpha\,(L_p - \widetilde{B}_p), 0, 1\right), \tag{4}$$

where C_s are pre-resolve samples; R is their resolve; R_x, R_y, and R_{xy} are selected by output phase; and $D = R - 0.25(R_L + R_R + R_T + R_B)$ is local detail. The resolved contrast bounds the detail scale γ, while resolved contrast and sample disagreement control α. Four phases are written to the corresponding 2× pixels.

MSPR-Lite retains the same four-phase output but replaces the adaptive neighborhood with a fixed six-tap reconstruction:

$$O_p^{\text{Lite}} = \text{clip}\left(\sum_{s=0}^{3} a_{p,s} C_s + b_{p,L} R_L + b_{p,R} R_R, 0, 1\right), \tag{5}$$

where R_L and R_R are the left and right resolved colors read directly from the imageblock. A neighbor falls back to the current resolve when it lacks full four-sample coverage. This path does not stage a 3×3 threadgroup neighborhood and requires no barrier. Its fixed coefficients live in `formulaV2MSAATile`, so both Lite and Quality write the final 2× image inside the rendering pass.

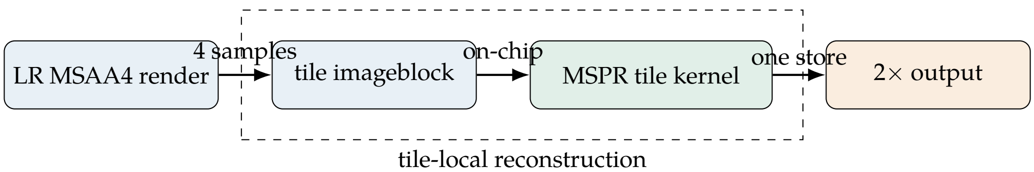

Figure 1: Intended Metal execution. The four-sample data is never materialized as a full-frame external texture.

2.2 MSPR-Quality calibration

The release quality profile uses the fixed phase coefficients, contrast blend, bounded detail, and sample-disagreement gate below. The current release-shader scene matrix was used to choose this profile over experimental alternatives, so that matrix is release-validation evidence rather than an untouched benchmark. The separate 100-frame full-scene set was not used to fit these coefficients.

$$W_{\text{MSPR-Quality}} = \begin{bmatrix} 0.63305 & 0.11351 & 0.27293 & -0.01922 \\ 0.39027 & 0.56770 & 0.02330 & 0.01899 \\ 0.01989 & 0.02368 & 0.56413 & 0.39257 \\ -0.01760 & 0.27210 & 0.11202 & 0.63377 \end{bmatrix}. \tag{6}$$

The contrast interval is 0.0400–0.2000, low/high blend weights are 0.7000/0.3000, bounded detail gain is 0.3000, and sample-disagreement floor is 0.7000. Every coefficient row is constrained to unit DC gain. The machine-readable profile is `Reports/MSPR/parameters/mspr-quality.json`; the shader uses the same constants at half precision.

2.3 Tuning for a production rendering workload

Freeze the input contract before tuning: MSAA sample order and locations, tone map, color space, attachment format, exposure, and the Direct 2× reference must remain fixed during a search. Build a development set stratified by content family, including diagonal geometry, alpha cutouts, transparency, highlights, fine lines, repeated textures, particles, UI, and low-contrast gradients. Keep both static frames and moving-camera sequences for every important family.

1. **Fit the phase matrix.** Fit each output phase from $C_0 \ldots C_3$ with regularized constrained least squares against the Direct 2× target, enforcing a unit row sum. Optimize this matrix alone first so gate changes cannot hide a coefficient error.
2. **Calibrate the contrast gate.** Initialize `contrast_low/high` from luminance-gradient percentiles. More Local weight at high contrast strengthens phase edges but also amplifies wrong sample ordering, texture ringing, and HDR overshoot.
3. **Calibrate sample disagreement.** Set `sample_low/high` from the distribution of the maximum channel difference across the four samples. Raise `sample_floor` for true geometry and transparency boundaries; lower it for noisy textures or unstable highlights.
4. **Tune detail last.** Increase `detail_gain` from zero only to compensate for softness in the spatial base. Monitor Scharr-gradient L1, negative-lobe overshoot, and temporal shimmer together with single-frame PSNR.
5. **Accept by content family.** Optimize mean and P10 PSNR, L1, gradient L1, and the worst family score. Freeze the selected JSON and SHA-256 before a one-shot validation on full rendered sequences excluded from the search.

These parameters change constants, not sample reads, neighborhood reads, barriers, or stores, so parameter variants within the MSPR-Quality topology normally have the same GPU cost. A project with a different sample pattern, HDR pipeline, or transparency order should rerun this calibration instead of copying the reported ranking.

3 Evaluation protocols

3.1 Archived nine-scene method matrix

The archived comparison used nine deterministic scene families at 958×442 input and 1916×884 output, scene time 3.0, and seed 17. The families were transparency, flat shaded geometry, checker detail, fine lines, alpha cutout, specular highlights, gradients, a full-screen image, and Temple geometry. Every scored method was compared with the corresponding Direct 2× frame using RGB PSNR.

Bilinear 2×, the project FSR1 EASU+RCAS port, MetalFX Spatial, MSPR-Quality, MSPR-Lite, and Direct 2× were generated by one earlier release on one iPhone 13 running iOS 18.2.1. All nine scenes shared the same resolution, time, seed, tone map, and RGBA8Unorm quantization contract. Performance used 20 warm-up frames and 120 measured frames per method and scene, with command-buffer GPU duration as the full-path metric. The assets used by the archived Temple and full-screen-image scenes came from the downloadable demo accompanying Apple's "Implementing order-independent transparency with image blocks" documentation [1]. They were test inputs only. Because this project does not grant rights to redistribute Apple's sample assets, neither scene nor its assets are included; users can obtain the original demo from the linked Apple page and use it subject to Apple's applicable terms. The current package can regenerate only the seven project-authored procedural scenes, not the two archived Apple-demo rows.

3.2 Broader full-scene rendering set

This dataset was built and frozen by the project; it is not a public or industry-standard benchmark. Scripted real-time scenes produced 100 frames in ten 10-frame sequences at 5760×3240 native-6, followed by 960×540 LR sample proxies and a 1920×1080 MSAA4 target. The content covers high-contrast thin geometry and diagonals, layered indoor occlusion, curved vehicle highlights, animated character silhouettes, translucent effects, and world-space UI. Camera profiles include orbit, yaw, strafe, sprint, micro-shake, and FOV changes. GI, TAA, FXAA, and post-AA were disabled to isolate spatial-reconstruction error. The set has zero sequence, capture-session, and content-hash overlap with the training and calibration plans.

Table 1: Content distribution of the 100-frame full-scene rendering set.

Family	Frames	Camera profiles	Stress factors
Alias geometry	20	orbit, yaw turn	thin poles, silhouettes, high-contrast diagonals
Indoor	20	sprint, FOV zoom	occlusion, hard boundaries, depth layers
Vehicle	20	strafe, micro-shake	specular curves, wheel and body contours
Character	20	FOV zoom	animated silhouettes, fine attachments
Translucent VFX	10	sprint	alpha blending, reactive detail
World-space UI	10	yaw turn	text-like edges, planar overlays

The four sample positions are fixed D3D-style positions extracted from native-6 captures: $(x, y) = (2, 0), (5, 2), (0, 3), (3, 5)$ in sixths of an LR pixel. Tone-mapped inputs and targets are quantized to RGBA8Unorm-equivalent precision. This is reproducible and preserves subpixel phase, but it is not a hardware Metal sample dump.

3.3 Metrics and information contracts

MSPR-Quality receives the four samples. Bicubic and the project's line-equivalent FSR1 EASU+RCAS path receive their arithmetic resolve; FSR1 sharpness is 0.8. Native 2× MSAA4 is the reference rather than a scored method. We report RGB PSNR, SSIM [6], L1, Scharr-gradient L1, edge-weighted L1, binary edge F1, and an unwarped temporal-delta error. The temporal number is diagnostic only because camera motion is not compensated.

The FSR1 numbers characterize the project's Metal-port logic under this report's positive-Reinhard/RGBA8 contract. They do not constitute a quality ranking of every AMD SDK integration; FSR1 documentation recommends an already anti-aliased perceptual-space input, and a production engine may use a different color pipeline or RCAS setting.

4 Quality results

4.1 Archived nine-scene spatial comparison

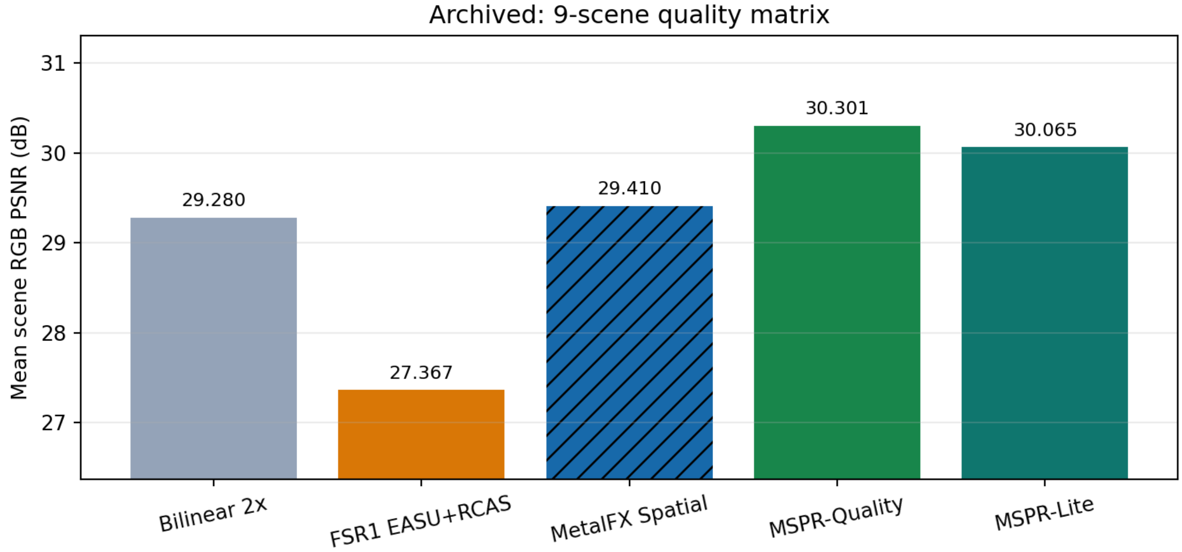

Figure 2: Mean scene PSNR over the archived nine-scene matrix on one iPhone 13. The two Apple-dependent scene rows are retained as non-reproducible historical evidence.

Table 2: Archived nine-scene method matrix summary. Higher PSNR is better.

Method	Mean scene PSNR	Evidence source
Bilinear 2×	29.280 dB	physical iPhone 13
FSR1 EASU+RCAS	27.367 dB	physical iPhone 13
MetalFX Spatial	29.410 dB	physical iPhone 13
MSPR-Quality	**30.301 dB**	physical iPhone 13
MSPR-Lite	30.065 dB	physical iPhone 13
Direct 2×	reference	physical iPhone 13

In the archived matrix, MSPR-Quality led the arithmetic scene mean by 0.237 dB over MSPR-Lite, 0.891 dB over MetalFX, 1.021 dB over Bilinear, and 2.934 dB over FSR1. It recorded the highest PSNR on transparency, specular, and Temple. MSPR-Lite led flat shading and gradients while retaining nearly the same aggregate quality at lower cost. The Temple observation is historical and cannot be rerun from this package.

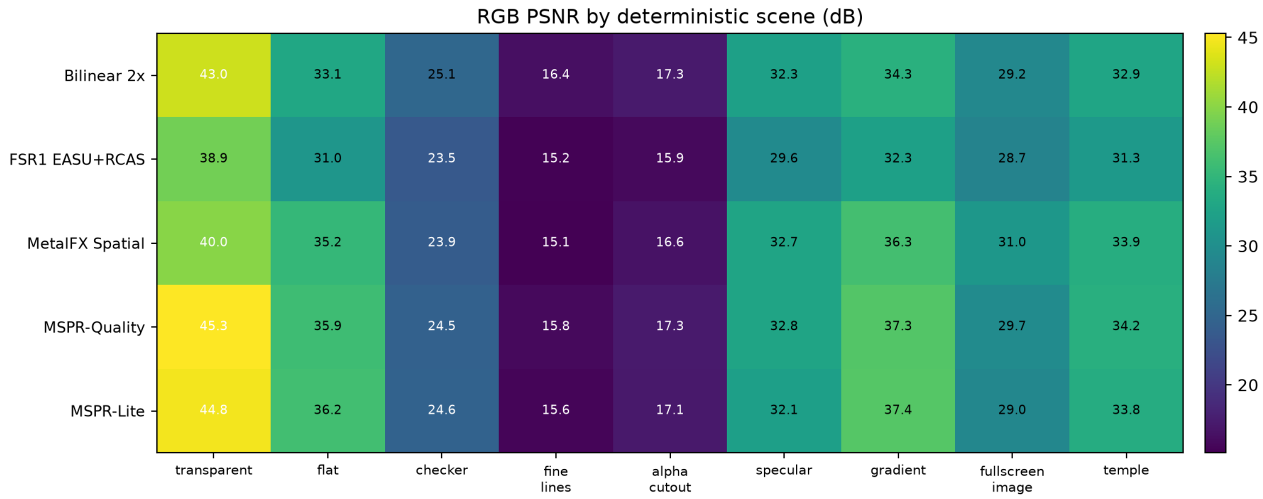

Figure 3: Per-scene RGB PSNR. The matrix exposes the content families hidden by a single average.

4.2 Representative current-release method cases

The three release-shader examples cover transparency, alpha cutout, and specular content at time 3.0 and seed 17. Every grid shows the LR resolve, Bilinear, FSR1, MetalFX Spatial, both MSPR profiles, Direct 2×, a fixed crop, per-method PSNR, and an amplified MSPR-Lite error panel generated by the same iPhone 13.

The release can regenerate the seven-scene protocol and current assets from source. Install the quality-profile iOS scheme on a physical iPhone, then run from the repository root:

```
uv sync --frozen
uv run python Demo/Tools/reproduce_release_quality.py --device <iPhone-UDID>
```

The script runs the seven distributed scenes, writes `Reproduction/MSPR-Quality/quality-matrix.json`, pulls representative PPM/PNG image sets from App Documents, and creates report-style summaries and case grids under `figures/`. A physical iPhone with MetalFX Spatial support is required. It does not regenerate the two removed historical scene rows.

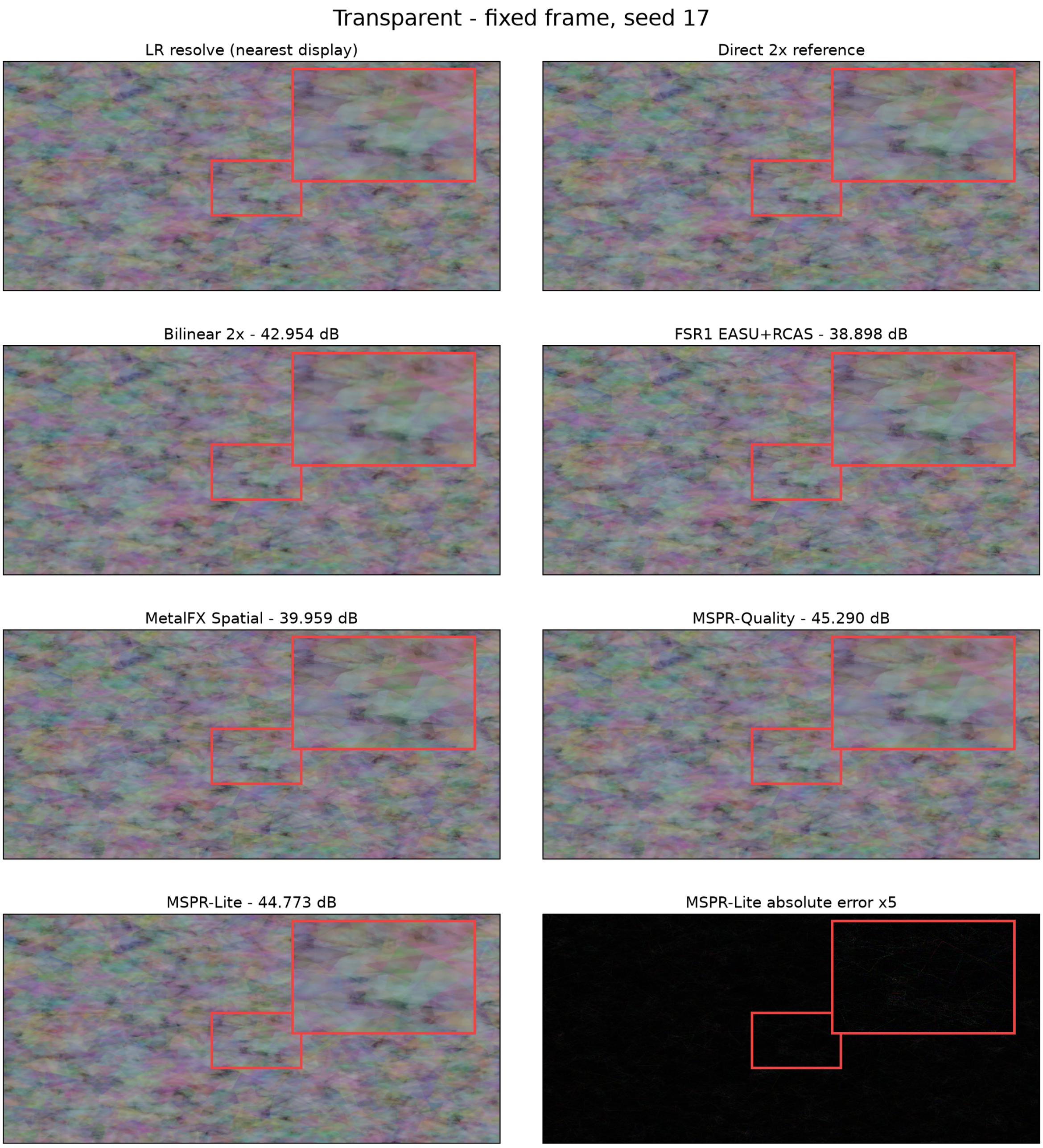

Figure 4: Transparent stress case. MSPR preserves pre-resolve phase evidence and leads this scene numerically.

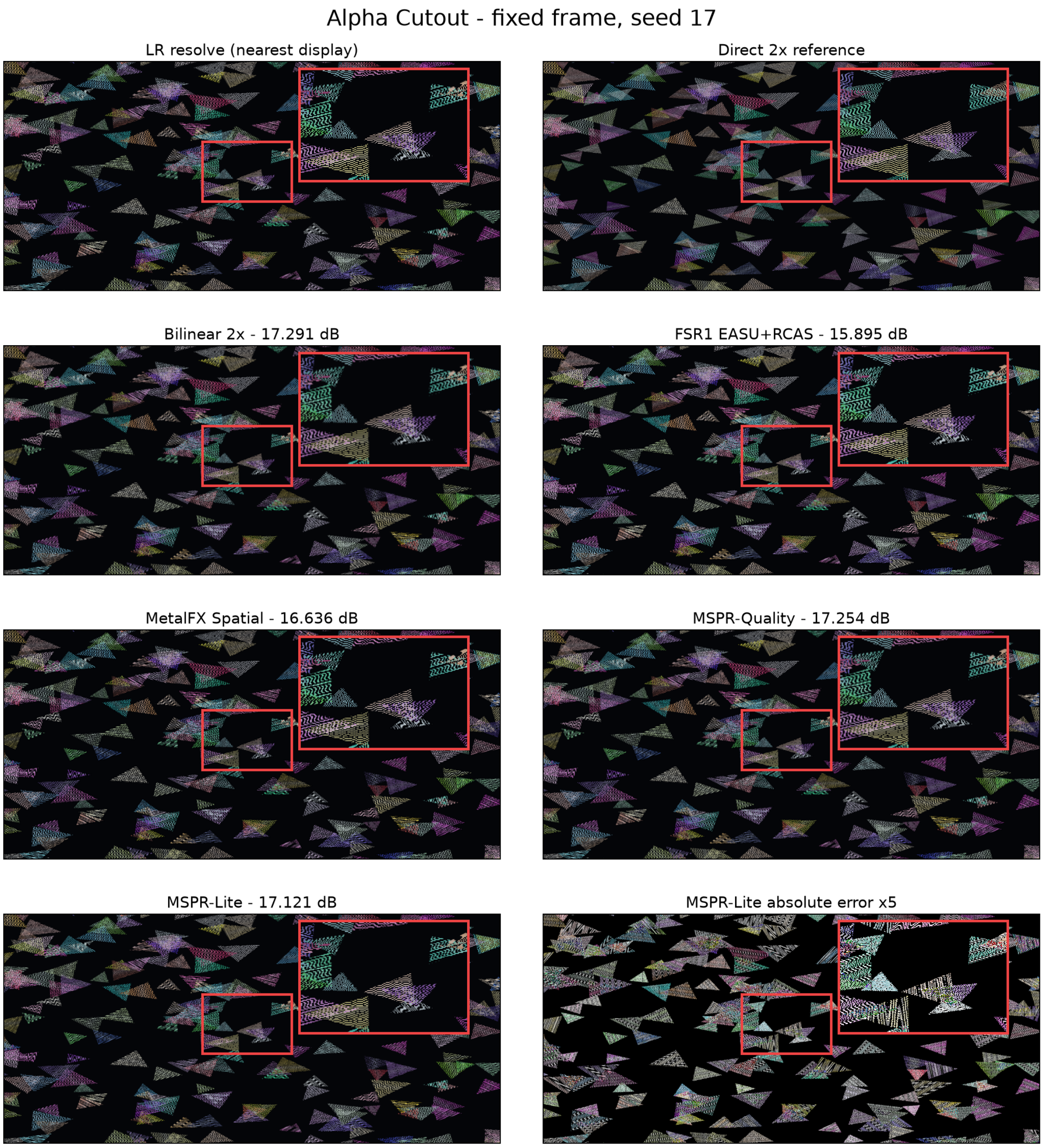

Figure 5: Alpha-cutout case. Bilinear leads MSPR-Quality by only 0.037 dB; the zooms expose their different edge responses.

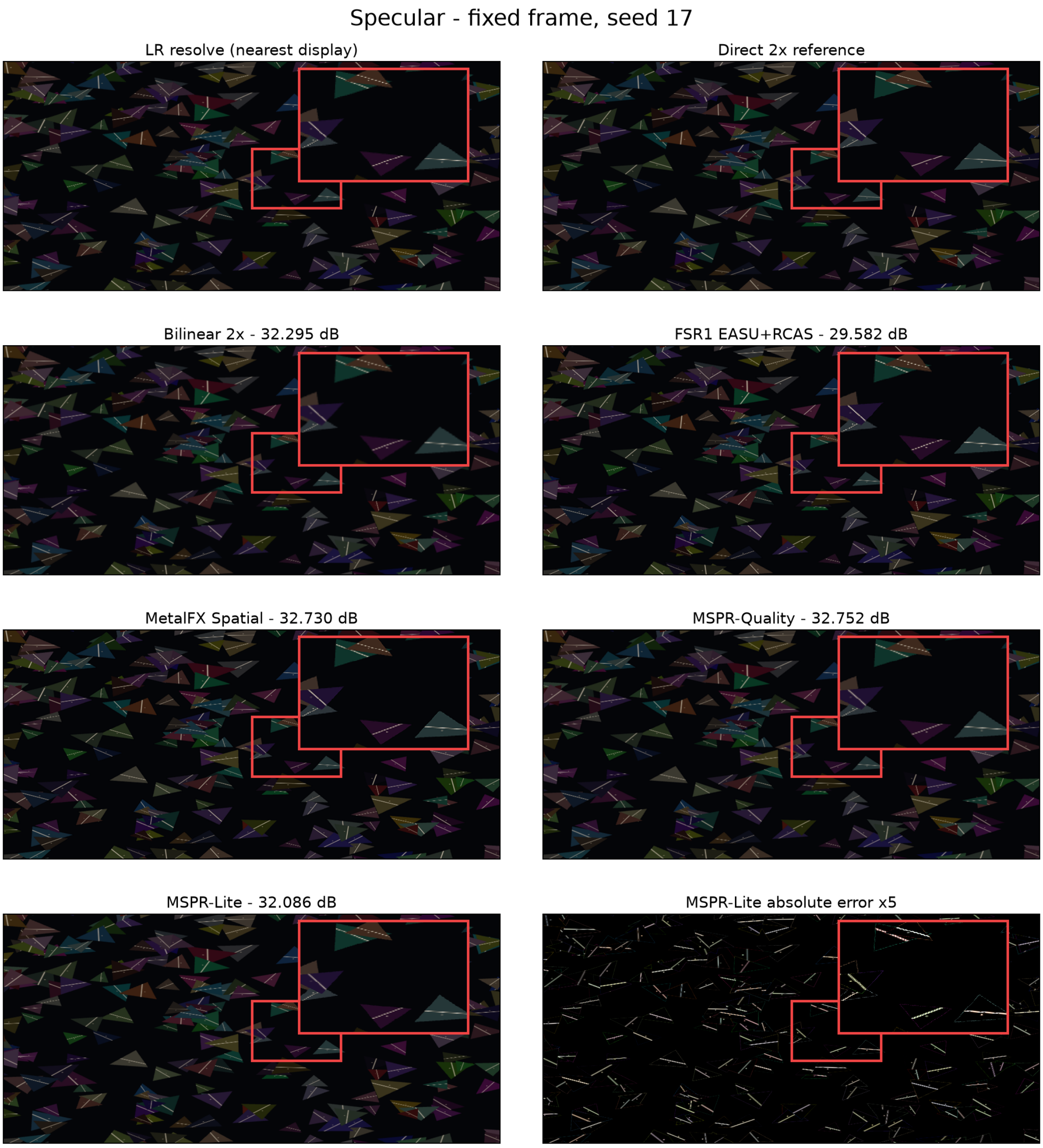

Figure 6: Specular case. MSPR-Quality records the highest PSNR, 0.022 dB above MetalFX Spatial.

4.3 Broader 100-frame full-scene rendering set

The larger offline set cannot run MetalFX and does not contain MSPR-Lite outputs, so it remains secondary evidence for MSPR-Quality, Bicubic, and FSR1 rather than the main method ranking.

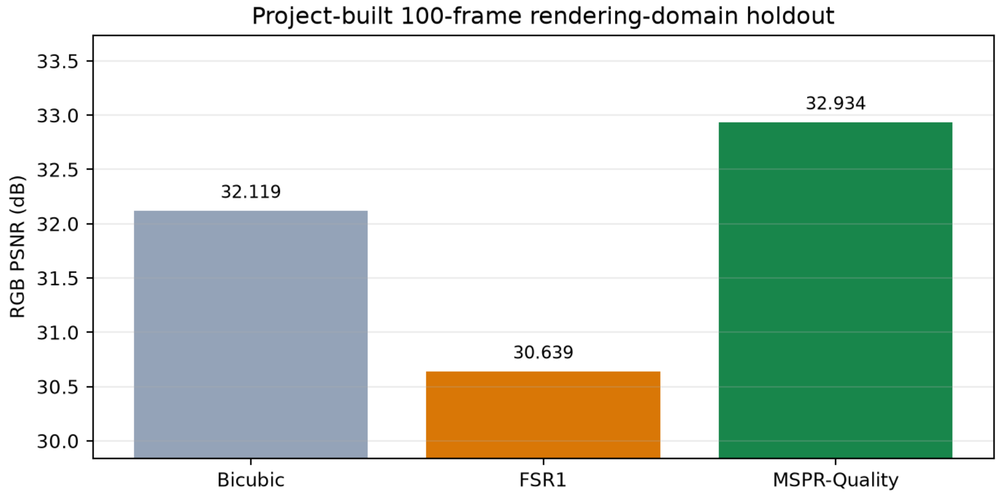

Figure 7: Broader 100-frame full-scene rendering set. Higher PSNR is better.

Table 3: Broader full-scene set averages. Arrows indicate the preferred direction.

Method	PSNR↑	SSIM↑	L1↓	Grad. L1↓	Edge F1↑
Bicubic resolve	32.119	0.97949	0.00444	0.00700	0.8258
FSR1 EASU+RCAS	30.639	0.97474	0.00563	0.00763	0.7858
MSPR-Quality	**32.934**	**0.98504**	**0.00351**	**0.00521**	**0.8770**

MSPR-Quality's 0.815 dB gain over Bicubic is accompanied by lower L1 and gradient error. Its 2.295 dB advantage over FSR1 must be read with the information-contract caveat: FSR1 cannot access phase evidence discarded by resolve.

4.4 Results on full real-time rendering scenes

These examples come from complete rendered sequences rather than isolated procedural patterns. They cover alias-stress geometry, curved vehicle highlights and fine geometry, and translucent VFX under camera-driven real-time rendering, complementing the fixed release-shader matrix above.

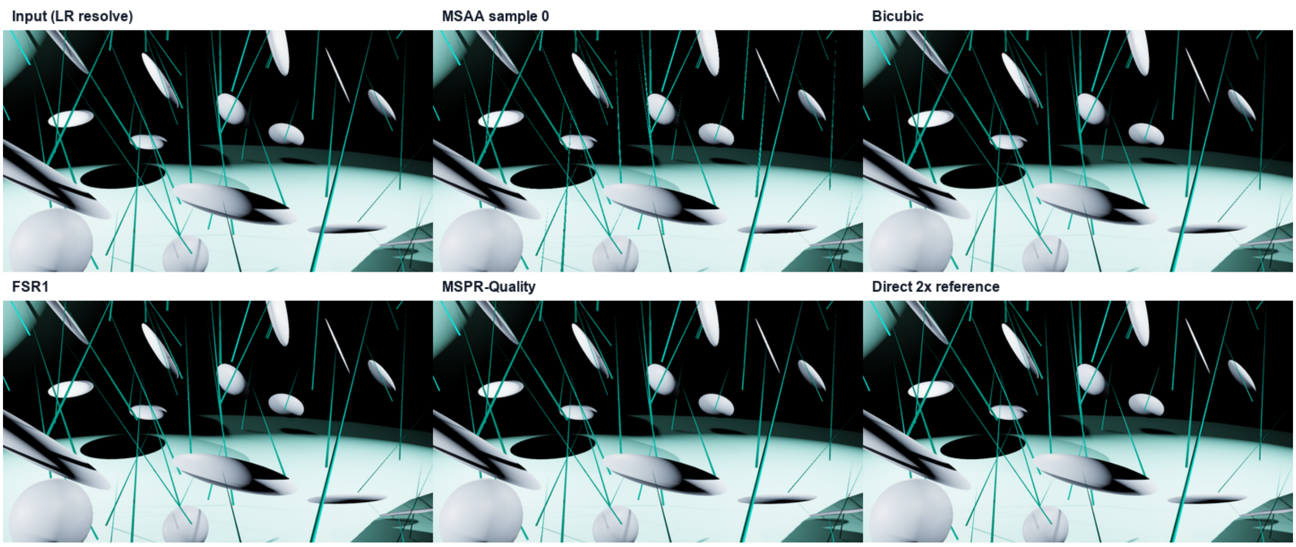

Figure 8: Alias-stress case: LR resolve, one pre-resolve MSAA sample, Bicubic, FSR1, MSPR-Quality, and native 2× reference.

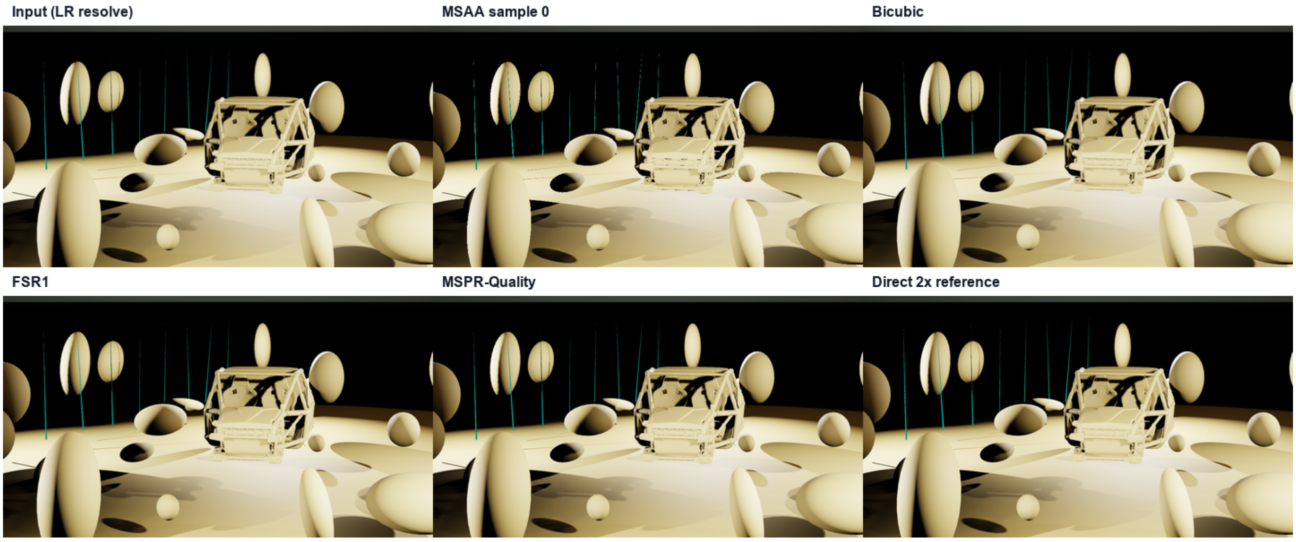

Figure 9: Vehicle case with curved specular boundaries and fine geometry.

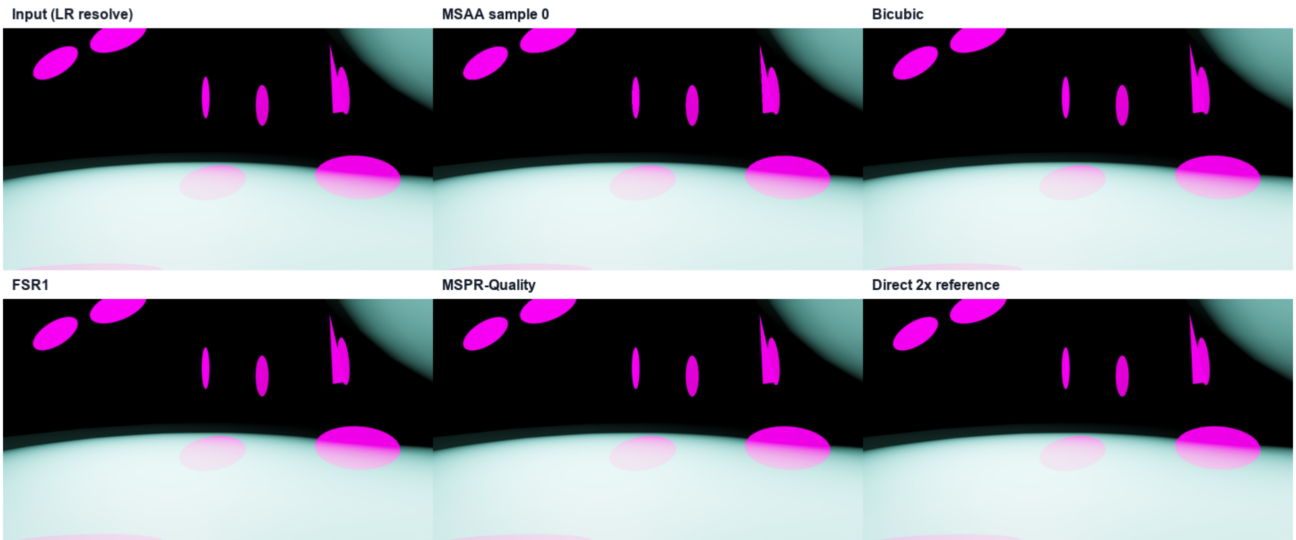

Figure 10: Translucent VFX case. Alpha-blended content remains a deployment risk even when aggregate metrics improve.

5 Performance and bandwidth evidence

5.1 Device and measurement protocol

The archived device run used a iPhone 13 (Apple A15 GPU) on iOS 18.2.1, rendering 958×442 and producing 1916×884. Each of the 9 scenes used scene time 3.0 and seed 17. Every method ran for 20 warm-up frames followed by 120 measured frames in every scene. GPU time is the full command-buffer duration, not CPU encode time or an isolated shader-kernel time. Method means are equal-weight arithmetic means across the same nine scenes used for PSNR.

5.2 Same-matrix full-path comparison

Table 4: Archived iPhone 13 quality and full-path GPU time on the same nine-scene matrix (lower GPU time is better).

Method	Mean RGB PSNR	GPU mean (ms)	Added time vs Bilinear (ms)
MSPR-Quality	**30.301**	3.656	1.067
MSPR-Lite	30.065	3.218	**0.629**
MetalFX Spatial	29.410	7.606	5.017
Bilinear 2×	29.280	**2.589**	**0.000 (reference)**
FSR1 EASU+RCAS	27.367	8.575	5.986

The final column uses one transparent reference:

$$\Delta_{\text{full}}(m) = \overline{T}_{\text{GPU}}(m) - \overline{T}_{\text{GPU}}(\text{Bilinear}), \tag{7}$$

where each bar is the equal-weight nine-scene mean. This is a complete render-path difference under one workload contract; no separate formula-only baseline is subtracted, and the result must not be described as isolated reconstruction-kernel time. On this evidence, MSPR-Quality improves mean PSNR by 1.021 dB for 1.067 ms of added full-path time; MSPR-Lite improves it by 0.785 dB for 0.629 ms. The same matrix records substantially larger added time for MetalFX Spatial (5.017 ms) and FSR1 (5.986 ms).

MSPR keeps MSAA samples and the resolved neighborhood inside tile memory and persists only the final 2× image. FSR1 normally persists a resolved LR input and a full-resolution EASU intermediate before RCAS. SGSR1 reduces this to one shader pass but still begins from resolved color. MetalFX exposes input/output textures while internal traffic is opaque. Logical pass counts do not equal DRAM bytes because Apple compression, caches, store actions, and tile spill dominate real traffic.

These full-path means combine scene rendering and reconstruction, so they support same-workload deployment comparisons but do not identify standalone kernel cost. The two Apple-demo-dependent scene contributions are historical; a new deployment should repeat the full matrix with redistributable content representative of its own renderer.

6 Position among mobile upscalers

Table 5: Method contracts and the evidence available in this report.

Method	Inputs	Execution	Main strength	Evidence here
MSPR-Quality	four MSAA colors + coverage/neighborhood	tile-local spatial	preserves pre-resolve phase	iPhone 13 quality and full-path time
MSPR-Lite	four MSAA colors + two horizontal resolved taps	barrier-free tile-local spatial	small added full-path time	iPhone 13 quality and full-path time
FSR1	resolved color	EASU + RCAS	portable open spatial baseline	iPhone 13 quality and time
SGSR1	resolved color	one mobile shader	Adreno-oriented bandwidth	design comparison only
MetalFX Spatial	color texture	Apple runtime	platform integration	iPhone 13 quality and time
SGSR2	color, depth, motion, history	temporal	dynamic quality and stability	not an equal-contract baseline
Arm ASR	FSR2-derived temporal inputs	temporal mobile	open cross-GPU temporal path	not an equal-contract baseline
Direct 2×	native scene render	full-resolution raster	reference quality	GT; iPhone time

SGSR2 and Arm ASR can exploit cross-frame evidence and target aliasing/flicker that a stateless MSPR kernel cannot [4, 7, 8]. They also require motion, depth, jitter, history, and disocclusion handling. The appropriate comparison is therefore MSPR versus spatial post-resolve upscalers for the same no-history operating point, with temporal methods reported as a separate quality tier.

A Photo proxy as an out-of-domain scalability probe

Set5, Set14, and Urban100 are photographs, not rendered scenes. They are therefore excluded from parameter selection and from the main shipping conclusion. Their role is limited to testing whether the phase reconstruction can adapt to repeated structures and natural-image statistics under a synthesized four-sample proxy [9, 10, 11].

Table 6: Out-of-domain RGB PSNR (dB). Not a standard single-LR SISR leaderboard.

Method	Set5	Set14	Urban100
Bicubic resolve	31.207	28.144	25.123
FSR1 resolve	28.181	25.946	23.405
MSPR photo profile	37.567	34.948	32.247

The large photo-proxy gains show extensibility, not production validity. The photo profile uses stronger negative lobes and performs worse than MSPR-Quality on rendered internal validation (33.512 versus 36.135 dB). Domain-specific parameters are therefore expected and should be versioned with the renderer.

B Reproducibility and limitations

The release preserves the parameter records, per-frame CSV, aggregate JSON, raw current-release scene logs, and representative full-scene PNG cases. The current scene matrix was used to select the default release profile and is therefore reported as validation evidence, not as an untouched benchmark. The evaluator uses a deterministic positive-Reinhard tone map and RGBA8Unorm-equivalent quantization.

Remaining threats to validity are the spatial sample proxy in the broader scene set, GI-off content, absence of production post-processing, the project-specific FSR1 color contract, lack of motion-compensated temporal metrics, full-path rather than isolated-kernel timing, two historical Apple-demo-dependent rows in the nine-scene means, and no SGSR quality output. These are explicit limitations rather than inferred equivalence.

References

[1] Apple Developer Documentation. Implementing order-independent transparency with image blocks. https://developer.apple.com/documentation/metal/implementing-order-independent-transparency-with-image-blocks. Accessed 2026-08-17.

[2] AMD GPUOpen. Fidelityfx super resolution 1, 2021.

[3] AMD GPUOpen. Fidelityfx-fsr: Fsr1 reference source, 2021.

[4] Snapdragon Game Studios. Snapdragon game super resolution, 2024.

[5] Apple Developer Documentation. Mtlfxspatialscaler.

[6] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[7] Arm. Introducing arm accuracy super resolution, 2024.

[8] Arm. Arm accuracy super resolution, 2025.

[9] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In *British Machine Vision Conference*, 2012.

[10] Roman Zeyde, Michael Elad, and Matan Protter. On single image scale-up using sparse-representations. In *Curves and Surfaces*, pages 711–730. 2010.

[11] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *CVPR*, 2015.