Forward-First Radial Lens Distortion Modeling with Deterministic Near-Exact Undistortion and Domain-Aware Monotonicity Analysis

Ahmet Basaran

◆

Abstract—Radial lens distortion is conventionally modeled with a fixed family of terms—the odd-degree Brown polynomial, or a rational or fisheye variant—fitted from distorted observations by minimizing reprojection error. This carries two limitations. First, real optics include profiles these families cannot represent: the mustache (barrel-to-pincushion) curve, the mid-spatial-frequency zonal ripple of molded aspheres, and the deliberate two-zone compression of foveated and panomorph lenses. Second, a reprojection-error fit is not constrained to stay monotonic, so the recovered mapping can fold and then silently fabricate content when applied. This paper presents a forward-first framework that treats the distortion function as the primary model and represents it as a sum of arbitrary differentiable basis functions—polynomials of unrestricted degree together with localized terms such as Gaussian zonal bumps and sigmoid knees—selected automatically from data. The selection is monotonicity-constrained, so the recovered mapping is invertible by construction in every case and never folds. Undistortion is derived analytically; because the forward model is explicit, the full distort-then-undistort pipeline is validated against a synthetic ground truth at the interpolation-noise floor—a methodology we apply to the baseline families as well. On out-of-family profiles, even an ideal-solver fit of the fixed families leaves pixel-level errors up to 16 px, whereas the open basis reaches the corner-noise floor (0.26 px on a zonal-ripple lens) and recovers a foveated lens's design parameters directly. On a five-board strong-barrel calibration the estimated function matches ground truth to an RMSE of 6.7×10^{-5} (normalized radius), automatically reduced to a parsimonious monotonic model; and under a non-monotonic ground truth a standard pipeline calibrates to sub-0.1 px reprojection error yet silently renders mirrored content, which the proposed hard/soft monotonicity diagnostics detect and bound. Distortion constants come from a direct linear least-squares solve—no iterative optimization. The implementation is available as a commercial toolbox for MATLAB and Python.

Index Terms—lens distortion, undistortion, camera calibration, radial distortion, basis functions, model expressiveness, monotonicity, polynomial model, least squares, forward modeling, corner detection.

1 INTRODUCTION

Optical systems inevitably introduce geometric aberrations. Radial lens distortion—the most common and practically significant form—displaces image points along the radial direction from the optical axis, causing straight lines in the scene to appear curved. Correcting this distortion is a prerequisite in precision metrology, photogrammetry, machine vision, and computational photography.

The established approach to lens distortion modeling, originating with Conrady [11] and Brown [1] and extended through Zhang's calibration method [2] and related photogrammetric and self-calibration techniques [12], [13], [14], proceeds *inverse-first*: the camera is treated as a black box producing distorted images, and calibration estimates a correction mapping that recovers undistorted scene geometry. Distortion coefficients are fitted by minimizing reprojection error through nonlinear optimization, using correspondences between detected image points and their known undistorted positions. This paradigm underpins virtually all widely used camera calibration software.

The inverse-first paradigm, as practiced, leaves two issues unaddressed. First, the distortion is modeled by a *fixed* family of terms—Brown's odd-degree polynomial, or a rational or fisheye variant—chosen before any data is seen. Real optics depart from these families: the mustache curve that runs from barrel at the center to pincushion at the edge, the mid-spatial-frequency zonal ripple of molded aspheres, and the deliberate two-zone profile of foveated and panomorph lenses are all smooth and physically valid yet lie outside the span of any fixed family, so no choice of coefficients fits them. Second, the fit is not constrained to remain monotonic—that is, invertible: nothing in a reprojection-error fit prevents the recovered mapping from folding, most commonly where it is extrapolated beyond the calibrated radial range, after which undistortion silently fabricates mirrored content. Reprojection error, the usual figure of merit, measures the self-consistency of the fit rather than the pixel-level fidelity of the corrected image, so neither issue surfaces in it.

This paper proposes a forward-first framework that addresses both. The distortion function is the primary model and is represented as a *sum of arbitrary differentiable basis functions*—polynomials of any degree together with localized terms (Gaussian zonal bumps, sigmoid knees)—chosen automatically from data, so profiles outside the fixed families are represented directly rather than approximated. The same selection is monotonicity-constrained, so the recovered mapping is invertible by construction and never folds—expressiveness without the overfit risk. Undistortion

---

- *A. Basaran is an Independent Researcher, Ankara, Turkey. E-mail:* `ahmet.basaran@opticaldistortionlab.com`.
- *Manuscript in preparation (draft, not peer-reviewed), June 18, 2026.*

is derived analytically from this explicit forward model, which also lets the corrected image be checked pixel-by-pixel against a synthetic ground truth (distort a known image, undistort, compare). That synthetic round trip is exact by construction and reaches the interpolation-noise floor; it is a validation *methodology*, not a capability unique to our model, and we apply it in Section 9 to benchmark the standard model families on the same footing.

The contributions of this work are:

1) **Forward-first distortion framework**—the distortion function is the primary, precisely defined model; undistortion is a derived near-exact inverse.
2) **Open-basis distortion model (function freedom)**—the distortion function is modeled as a sum of arbitrary differentiable basis functions: polynomials of any degree (odd, even, mixed) *plus* localized terms such as Gaussian zonal bumps and sigmoid knees, all selected automatically from data. This represents real out-of-family profiles—the mustache curve, molded-asphere zonal ripple, foveated/panomorph two-zone compression—that fixed Brown / rational / fisheye families cannot.
3) **Monotonicity-constrained estimation (invertible by construction)**— forward selection admits a term only if the model stays monotonic over the domain, so the recovered mapping is invertible in *every* case and the procedure never returns a non-monotonic (folding) model—unlike an unconstrained reprojection-error fit, which can. This is what makes the open basis safe: expressiveness without overfit folding.
4) **Hard/soft monotonicity loss metrics**—principled separation of physically irreversible information loss (folding) from recoverable attenuation, with quantitative domain-referenced metrics.
5) **Near-exact undistortion with synthetic ground-truth validation**— round-trip pixel error at the interpolation-noise floor; we use the synthetic ground truth to validate our pipeline and, on the same footing, the baseline families (Section 9).
6) **Deterministic precomputed inverse**—the image-space inverse is applied without any per-pixel iterative solver, eliminating convergence risk.
7) **Direct linear calibration**—distortion constants from a single least-squares solve, no iterative optimization, supporting multi-board pooling and both classical and learned corner detection.
8) **FOV-adaptive output geometry**—output image dimensions computed from the distortion model, adapting correctly to both barrel and pincushion profiles.

2 BACKGROUND AND RELATED WORK

2.1 The Classical Radial Distortion Model

The radial distortion model introduced by Brown [1] expresses the distorted radial coordinate r_d as a polynomial correction to the undistorted coordinate r_u:

$$r_d = r_u + k_1 r_u^3 + k_2 r_u^5 + k_3 r_u^7 + \cdots \tag{1}$$

where coefficients are normalized by focal length. The convention of exclusively odd-degree terms arises from the rotational symmetry assumed in the original derivation and has persisted largely unchanged in subsequent literature and tooling. Alternative parametrizations include the single-parameter division model [15], plumb-line formulations exploiting scene lines [16], and wide-angle/omnidirectional models [10], [20]; comparative evaluations of these families are given in [19]. All share the inverse-first orientation discussed below.

2.2 Calibration Methodology

Zhang's method [2] estimates distortion coefficients jointly with camera intrinsics by minimizing reprojection error between detected checkerboard corners and their projected positions under a pinhole-plus-distortion model. The optimization is nonlinear (typically Levenberg–Marquardt [23], [24], often within a bundle adjustment [25]) and requires diverse calibration views, careful initialization, and convergence monitoring. Representative implementations are described in [3], [4], [18]. A parameter-free variant that recovers the radial profile without committing to a fixed polynomial order [26] is the closest inverse-first analogue to the data-driven model structure pursued here.

The primary object of this estimation is effectively the *undistortion* model: coefficients are sought such that their inverse application to image coordinates best explains the observations. The forward direction—how an undistorted scene projects through the lens—is not the target of estimation.

2.3 Structural Constraints of the Inverse-First Paradigm

The inverse-first paradigm imposes structural constraints that follow from its problem formulation rather than from physical or mathematical necessity.

Absence of forward ground truth. When the distortion model is estimated from distorted observations, no independent forward ground truth exists. Reprojection error measures the geometric self-consistency of the fit. It does not directly quantify the pixel-level fidelity of the undistorted output, and it cannot be compared against an absolute reference.

Fixed polynomial structure. The conventional odd-degree polynomial basis—fixed before calibration data is collected—may not be the minimal or most accurate representation for a given lens. No mechanism exists within the inverse-first framework to adapt model structure to data while simultaneously enforcing physical validity.

Iterative inverse mapping. Undistortion in inverse-first systems is applied per pixel through an iterative solver (Newton–Raphson or fixed-point iteration), since the polynomial (1) has no closed-form inverse; approximate analytic inverses have been proposed to mitigate this [17]. The resulting residuals are bounded by convergence tolerance rather than by any fundamental physical limit.

Monotonicity unmonitored. The requirement that the distortion mapping be monotonically increasing—necessary for bijectivity and hence for information-theoretically lossless undistortion—is not routinely checked or reported in existing calibration workflows.

2.4 Prior Work on Monotonicity and Forward Models

Claus and Fitzgibbon [5] examined bijectivity conditions for rational lens models. Brito et al. [6] discussed self-calibration of radial distortion under invertibility assumptions. Quantitative metrics separating irreversible information loss from recoverable attenuation have not, to our knowledge, been proposed in prior work. Forward distortion modeling appears in computer graphics for physically based lens simulation [7] but has not been adopted as a calibration paradigm.

Enforcing monotonicity during fitting connects to the broader literature on shape-constrained regression: isotonic regression [27] and monotone regression splines [28] fit functions subject to an order constraint. Those methods constrain a nonparametric curve directly; the present work instead embeds the monotonicity test inside a parametric forward-selection loop (Section VI-B), so the constraint governs which basis terms enter the model rather than smoothing a fixed-form estimate.

On the detection side, learned methods have been applied to distortion correction directly from single images [21], to joint intrinsic and distortion estimation from a single view [30], [31], and to robust checkerboard corner localization under distortion [22]; the learned detector of Section VI-E belongs to the last line.

3 THE FORWARD DISTORTION MODEL

3.1 Model Definition

Radial quantities are normalized by focal length f and pixel pitch p:

$$r_{\text{norm}} = r_{\text{pixels}} \cdot \frac{p}{f} \tag{2}$$

This normalization makes the model dimensionless and independent of sensor resolution. The use of focal length here is solely for coordinate scaling—defining the numerical domain over which the polynomial operates—and is categorically different from its role in projection-based calibration methods, where it appears as an intrinsic parameter of the 3-D-to-2-D mapping. Alternative normalization conventions that eliminate the focal length requirement are straightforward and do not alter the mathematical structure of the model.

The forward distortion function maps an undistorted normalized radius r_{in} to a distorted normalized radius r_{out}:

$$r_{\text{out}} = f(r_{\text{in}}) = r_{\text{in}} + \sum_i k_i \cdot r_{\text{in}}^{d_i} \tag{3}$$

where $\{k_i\}$ are distortion constants and $\{d_i\}$ are their associated polynomial degrees. No restriction is placed on the degrees: they may be odd, even, or any combination. The identity corresponds to all $k_i = 0$. The classical model of (1) is the special case $d_i \in \{3, 5, 7, 9, 11\}$.

Monomials are only one choice of basis. This is the framework's defining generality: both the linear estimation of Section VI-A and the monotonicity-constrained selection of Section VI-B require only the evaluation of each basis function and its derivative, so the basis may be *any* set of differentiable functions $\varphi_i(r)$,

$$f(r_{\text{in}}) = r_{\text{in}} + \sum_i k_i \cdot \varphi_i(r_{\text{in}}), \tag{4}$$

without changing either procedure. Section 9.8 exploits this with localized bases—Gaussian zonal terms and integrated-sigmoid knees—to represent physically realistic profiles whose local structure no fixed global family captures.

3.2 Coordinate System

Given an $H \times W$ image with optical center at $((W-1)/2, (H-1)/2)$, normalized Cartesian coordinates are assigned to each pixel. Polar coordinates (θ, r) follow from these, with $r_{\max} = \max(r)$ as the domain reference for monotonicity analysis.

3.3 Why Forward-First?

The forward definition of (3) reflects the physical process directly: the lens acts on light from the undistorted scene, displacing it radially. The image-to-scene mapping is derived. The inverse-first convention arose from the practical demands of calibration—observed positions are distorted, desired positions are undistorted—but this workflow orientation imposed polynomial structure and precluded ground-truth validation as side effects. Defining the forward model precisely recovers both.

4 THE DISTORTION AND UNDISTORTION PIPELINES

4.1 Forward Distortion

Given image I and parameters $\{k_i, d_i\}$, the distorted output I_d is produced in three stages.

Stage 1—FOV-adaptive output grid. Evaluating $f(\cdot)$ over the source domain determines the spatial extent of the distorted output. A sampling grid is constructed at the user-specified pitch. Barrel distortion ($f(r) < r$) expands the output field; pincushion ($f(r) > r$) contracts it. Output dimensions adapt to the distortion geometry rather than being fixed to input dimensions.

Stage 2—Deterministic inverse. For each output grid point, the corresponding source position is obtained from a deterministic, precomputed inverse of $f(\cdot)$. The inverse is constructed once and reused for every pixel; no per-pixel iterative solver is involved, so the mapping carries no convergence error.

Stage 3—Resampling. Source coordinates from Stage 2 sample I by image interpolation, producing I_d.

4.2 Undistortion

Given distorted image I_d and the same parameters, undistortion proceeds symmetrically. The undistorted output extent is determined by inverting $f(\cdot)$ at the corner radius. Each output grid point's source position is found by *direct forward evaluation* of $f(\cdot)$—no inversion at this stage. Image interpolation produces I_u.

The two pipelines are structurally complementary and form a round trip:

$$I \xrightarrow{\text{distort}} I_d \xrightarrow{\text{undistort}} I_u \approx I \tag{5}$$

Fig. 1. Non-monotonic distortion under backward mapping ($f(r) = r - 0.5r^3 + 0.2r^5 + 0.5r^7 - 2r^9$, fold at $r = 0.686$). The undistorted interior ($r_{\text{in}} \leq r_{\text{fold}}$) is compressed into the disk of radius $f(r_{\text{fold}}) = 0.523$, the maximum attainable distorted radius (dashed ring). Content at $r_{\text{in}} > r_{\text{fold}}$ is non-invertible ($\mathrm{d}f/\mathrm{d}r < 0$) and is left black in place rather than folded back over the scene; this destroyed content corresponds to the source corners, which extend past the fold radius while the image edges do not. The black strips near the top and bottom edges are not loss — they are simply where the source image, compressed into the barrel disk, does not reach. Frame dimensions and normalization are preserved (pixels are not cropped or resized). Test image: Canterbury Cathedral, Wikimedia Commons, CC BY-SA 2.0.

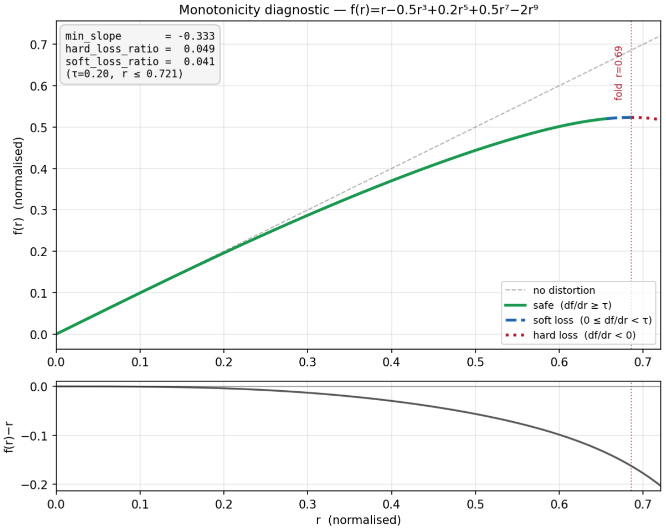


Fig. 2. Monotonicity diagnostic for the non-monotonic model, $f(r) = r - 0.5r^3 + 0.2r^5 + 0.5r^7 - 2r^9$. *Top:* the forward function color-coded by slope regime—safe (green, $\mathrm{d}f/\mathrm{d}r \geq \tau$), soft loss (blue-dashed, $0 \leq \mathrm{d}f/\mathrm{d}r < \tau$), hard loss (red-dotted, $\mathrm{d}f/\mathrm{d}r < 0$)—peaking at the fold $r = 0.686$ and descending beyond it. *Bottom:* deviation from identity $f(r) - r$. With $\tau = 0.20$ over the image domain ($r \leq 0.721$): $\min$ slope $= -0.333$, hard_loss_ratio$= 0.049$, soft_loss_ratio$= 0.041$. The non-monotonic band ($r > 0.686$) is thin—only 4.9% of the domain.

4.3 Non-Monotonic Distortion

The pipeline operates regardless of monotonicity. When $f(\cdot)$ is non-monotonic (hard_loss_ratio > 0, Section 5), the inverse retains only the monotonic branch up to the first fold at r_{fold}. The distorted output is then bounded to the disk of radius $f(r_{\text{fold}})$: undistorted content at $r_{\text{in}} \leq r_{\text{fold}}$ is mapped into this disk exactly, while content at $r_{\text{in}} > r_{\text{fold}}$ is non-invertible and is not rendered—the corresponding output region is left empty (black). Pixels are blacked out in place; image dimensions, and hence the radial normalization, are unchanged (the image is not cropped or resized).

This is a deliberate design choice. Rather than fabricating content in the non-invertible band—for example by folding outer scene points back over the monotonic interior, which produces mirrored pixels that can be mistaken for valid data—the pipeline marks the irrecoverable region explicitly as missing. Fig. 1 demonstrates this on a deliberately non-monotonic model whose fold radius lies inside the image-corner radius, so the corners fall in the non-recoverable region and appear black. The monotonicity diagnostic (Section 5) identifies and quantifies exactly this affected domain.

5 MONOTONICITY ANALYSIS AND INFORMATION LOSS

5.1 Physical Significance

For $f(r_{\text{in}})$ to constitute a physically realizable bijective lens mapping, it must be strictly increasing: each undistorted radius maps to a unique distorted radius. Where $\mathrm{d}f/\mathrm{d}r < 0$, the mapping *folds*: multiple undistorted positions map to the same distorted position, destroying the original information. Where $\mathrm{d}f/\mathrm{d}r$ is small but positive, local detail is compressed—recoverable given the exact model, but noise-sensitive in practice.

5.2 Hard Loss and Soft Loss Metrics

Hard loss ratio—fraction of the undistorted domain with irreversible information loss:

$$\text{hard\_loss\_ratio} = \frac{\left|\{r_{\text{in}} \in [0, r_{\max}] : \; \mathrm{d}f/\mathrm{d}r < 0\}\right|}{r_{\max}} \tag{6}$$

For any physically valid lens, hard_loss_ratio $= 0$. A positive value indicates a non-invertible region that no algorithm can recover.

Soft loss ratio—fraction with positive but sub-threshold slope:

$$\text{soft\_loss\_ratio} = \frac{\left|\{r_{\text{in}} \in [0, r_{\max}] : \; 0 \leq \mathrm{d}f/\mathrm{d}r < \tau\}\right|}{r_{\max}} \tag{7}$$

The threshold τ is user-configurable. A model with soft_loss_ratio > 0 but hard_loss_ratio $= 0$ is exactly invertible; attenuation is a practical concern, not a fundamental one. Together these metrics characterize the information content of the distorted image in a way that no scalar summary of the distortion model captures.

To our knowledge, this distinction between what is physically recoverable and what is not—expressed as computable, domain-referenced ratios—has not appeared in prior calibration literature.

5.3 Visualization

The distortion curve $f(r_{\text{in}})$ is color-coded by slope regime: green for safe (slope $\geq \tau$), blue-dashed for soft loss ($0 \leq$ slope $< \tau$), red-dotted for hard loss (slope < 0), with a vertical marker at the fold.

6 CALIBRATION: DIRECT LINEAR ESTIMATION

6.1 Problem Formulation

For each detected corner m in a distorted checkerboard image, with known grid geometry:

- $r_{\text{out},m}$: measured distorted normalized radius,
- $r_{\text{in},m}$: undistorted normalized radius from the corner's grid index and known spacing.

Substituting into (3) and rearranging:

$$\Delta r_m = r_{\text{out},m} - r_{\text{in},m} = \sum_i k_i \cdot r_{\text{in},m}^{d_i} \tag{8}$$

Stacking over M corners yields the linear system:

$$\mathbf{A}\,\mathbf{k} = \boldsymbol{\Delta r}, \qquad A_{m,i} = r_{\text{in},m}^{d_i} \tag{9}$$

solved by linear least squares:

$$\hat{\mathbf{k}} = \arg\min_{\mathbf{k}} \|\mathbf{A}\mathbf{k} - \boldsymbol{\Delta r}\|^2 \tag{10}$$

The solve is deterministic and non-iterative: a single matrix operation, no initialization, no convergence criterion, unique solution for $M > n$. The residual Δr has direct physical units—normalized radial displacement—interpretable without reference to a projection model.

This contrasts structurally with reprojection-error minimization, where distortion coefficients are entangled with intrinsic parameters inside a nonlinear projection, requiring iterative solvers and making individual coefficient errors difficult to interpret.

6.2 Automatic Polynomial Degree Selection with Monotonicity Enforcement

Selecting the polynomial degree set $\{d_i\}$ requires balancing fit quality against over-parameterization, while respecting the physical constraint that the fitted model must remain monotonic. We address this jointly through a forward selection procedure in which model complexity is increased only as long as two criteria are met.

Physical validity check. Each candidate model is admitted only if its slope stays positive over the radial domain:

$$f'(r) = 1 + \sum_i k_i \cdot d_i \cdot r^{d_i - 1} > 0, \quad \forall r \in [0, r_{\max}] \tag{11}$$

A candidate that violates (11) anywhere on the domain is rejected. The monotonicity constraint is thus embedded in the selection itself—not applied as a post-hoc filter—so the procedure cannot produce a physically invalid model regardless of the calibration data.

Fit quality criterion. Complexity is increased only while it produces a meaningful reduction in the fit residual,

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_m \left(f(r_{\text{in},m}) - r_{\text{out},m}\right)^2} \tag{12}$$

and selection stops once the residual reaches an acceptance tolerance. The result is the minimal monotonic model achieving the required fit quality: parsimonious for simple distortion profiles, higher-order only when data require it.

Both odd and even degrees are candidates. The selection is data-driven, not structurally predetermined. A key consequence, demonstrated in Section 9, is that the recovered model need not share the degree structure of the underlying distortion: the procedure recovers the distortion *function* over the observed domain using the fewest terms, not a prescribed coefficient set. The procedure is also basis-agnostic: candidates may be monomials or any differentiable functions (Section III-A), since admission requires only the slope check of (11) on the candidate model.

6.3 Undistorted Radius Estimation

The undistorted radius $r_{\text{in},m}$ for each corner is computed from its grid index (i, j) relative to the board center and the undistorted inter-corner spacing. Spacing is obtained from physical measurements (square size and camera-to-board distance) when available, or estimated from near-axis inter-corner distances—where distortion is negligible—when physical measurements are not supplied.

6.4 Off-Axis Board Correction

When the checkerboard center is displaced from the optical axis, its distorted and undistorted positions differ. A preliminary estimate is used to correct each board's per-corner radial coordinates for this offset, after which the final linear solve is performed on the corrected coordinates.

6.5 Corner Detection: Classical and Learned

The estimation pipeline accepts corner positions from either of two detectors, both implemented without dependency on external computer-vision frameworks.

Harris detector. A Harris corner response [9] with subpixel refinement via the gradient-orthogonality condition. This is a fully classical, training-free detector.

Learned detector. A convolutional corner detector with an EfficientNet-V2-S encoder [8] and a U-Net [29] decoder that regresses a per-pixel corner-response heatmap, trained to localize checkerboard corners under strong distortion. The learned detector maintains reliable detection at larger radii, where strong barrel distortion compresses the checkerboard pattern and degrades the gradient structure that the Harris response relies on.

Detector choice affects the radial coverage of the calibration and hence the domain over which the estimated model is valid. Section 9 quantifies this: under strong distortion, the learned detector extends usable coverage substantially and reduces the estimation error against ground truth by nearly two orders of magnitude relative to the Harris detector.

6.6 Multi-Board Pooling

Corners from multiple board placements are pooled prior to the linear solve, extending radial coverage beyond what any single placement provides. The off-axis correction is applied per board. Automatic degree selection operates on the pooled set. This multi-board capability is validated in Section 9.

TABLE 1
Round-Trip Error E_{rt} vs. Sampling Pitch (Known Coefficients). The strong-barrel model (degrees [3,5,7,9,11], constants $[-0.75,\ 0.52,\ -0.12,\ -0.01,\ 0]$) is applied to a band-limited smooth field and inverted with the same coefficients through a floating-point intermediate, isolating the interpolation-resampling floor.

Pitch	Grid	PSN	E_{rt} mean (GL)	E_{rt} max
h	1200×800	0.0010	0.269	1.45
$h/2$	2400×1600	0.0005	0.068	0.53

7 NEAR-EXACT UNDISTORTION: GROUND-TRUTH VALIDATION

7.1 The Synthetic Ground Truth

Given any image I and parameters $\{k_i, d_i\}$, the round trip

$$I \to \text{distort} \to I_d \to \text{undistort} \to I_u$$

can be evaluated exactly because the forward model is precisely known. The comparison I_u vs. I is an absolute pixel-level measurement, not a consistency check. We make this round trip the primary validation tool; it is available to any method once a forward model is fixed, and Section 9 applies it to the inverse-first baselines on the same footing.

7.2 Round-Trip Error

The normalized mean absolute round-trip error is:

$$E_{\mathrm{rt}} = \frac{1}{3HW} \sum_{c,i,j} |I_u(i,j,c) - I(i,j,c)| \qquad (13)$$

For a monotonic model, E_{rt} is bounded by the image-interpolation error, which scales with the square of the sampling pitch. The error is not convergence error—it is a property of discrete resampling and decreases with finer sampling.

The mean error falls by a factor of $3.97 \approx 4$ as the pitch is halved, confirming the predicted $O(h^2)$ interpolation scaling: the forward model is geometrically invertible to $\sim 10^{-6}$ px, so round-trip fidelity is limited only by discrete resampling—not by model or solver error. Both values are well below one gray level (GL).

7.3 Near-Exact Accuracy: Why It Is Achievable

In the undistortion stage, the mapping applied to each output point is a direct evaluation of the same $f(\cdot)$ that produced I_d—there is no model estimation uncertainty and no iterative approximation. The only error source is interpolation of a discrete image, introducing $O(h^2)$ error for grid spacing h.

In iterative per-pixel undistortion, two error sources are conflated: model estimation uncertainty (irreducible without a ground truth) and solver convergence error (reducible but non-zero by design). The forward-first pipeline eliminates both conflation and estimation uncertainty simultaneously, leaving only the irreducible interpolation floor.

8 SCOPE COMPARISON

The inverse-first and forward-first paradigms address overlapping but distinct problems. Table 2 summarizes the scope of each.

The inverse-first paradigm, as represented by Zhang's method [2] and established implementations [3], [4], excels at jointly estimating intrinsic parameters and distortion from multiple images captured at arbitrary camera-to-board poses—a capability the present framework does not address. The forward-first framework addresses a complementary regime: precise forward modeling, near-exact undistortion, and rigorous pipeline validation when the forward distortion is known or estimable with known intrinsics.

8.1 Model Expressiveness

The additive odd-degree polynomial of (1) is the special case of (3) with $d_i \in \{3, 5, 7, \ldots\}$. Fixing this structure before calibration implicitly assumes the lens is best described by odd-degree terms—a consequence of rotational symmetry assumptions that may not hold for all optical designs. The framework of (3) makes no such assumption; model structure is determined from data subject only to the physical constraint of monotonicity.

8.2 Validation Modality

Reprojection error, the standard optimization objective, measures whether detected corners project consistently under the estimated model. It does not indicate how close the undistorted image is to the true undistorted scene. The E_{rt} metric of (13) directly measures this—but requires a known forward model to define the ground truth. This is the core asymmetry between the two paradigms: the forward-first framework makes its validation objective concrete and absolute; the inverse-first framework makes its estimation objective tractable but its validation objective indirect.

A practitioner of inverse-first tooling can, of course, construct a synthetic round trip externally—distorting an image with a hand-written forward function and inspecting the corrected result. The distinction drawn here is operational: in inverse-first practice the forward model is not an object of the workflow, no validity guard accompanies the fitted model when it is applied, and reprojection error remains the only reported figure of merit. Section 9.9 makes the consequence concrete: a calibration reporting sub-0.1 px reprojection error while its rendered output contains fabricated content.

9 EXPERIMENTAL RESULTS

The central result is the model-family expressiveness comparison of Section 9.8: it isolates what the open basis buys over the fixed families on realistic out-of-family profiles, and is the clearest evidence for the paper's primary claim. The earlier subsections establish the groundwork it rests on—baseline calibration accuracy and the function-not-coefficients property, detector coverage, and undistortion quality—and the later ones probe behaviour under non-monotonic ground truth, where the monotonicity constraint and diagnostics matter most.

TABLE 2
Scope of the Forward-First Framework vs. the Inverse-First Paradigm.

Capability	**Forward-first (this work)**	**Inverse-first paradigm [1], [2], [3], [4]**
Primary model direction	Forward (distortion)	Inverse (undistortion)
Polynomial degree	Unrestricted, data-driven	Fixed, structurally predetermined
Even-degree terms	Supported	Not supported
Local basis functions (zonal, knee)	Supported	Not supported
Automatic term selection	Yes, with monotonicity	No
Inverse mapping	Deterministic, precomputed	Per-pixel iterative solver
Monotonicity monitoring	Hard and soft loss metrics	Not reported
FOV-adaptive output	Yes	No (fixed to input dimensions)
Calibration solve	Direct linear ($\Delta r = \mathbf{A}\mathbf{k}$)	Nonlinear (reprojection error)
Joint intrinsic estimation	Not addressed	Yes
Pixel-level GT validation	Yes (synthetic, integral)	External at best; not practiced
Validity guard at application time	Yes (fold detection, domain report)	None

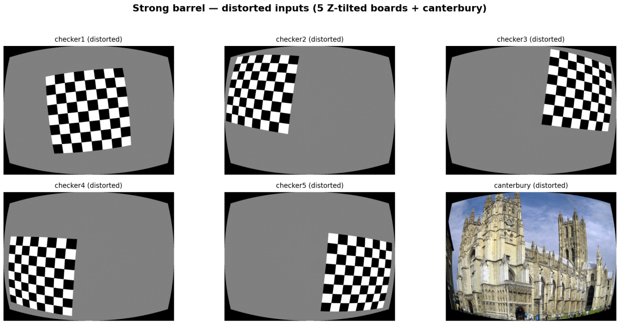


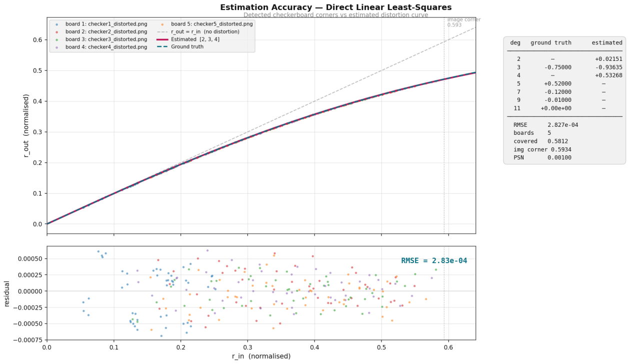
Fig. 3. Calibration inputs under strong barrel distortion: five Z-tilted checkerboards spanning different radial bands, plus the Canterbury test image. In-plane rotation of the boards improves radial coverage without introducing perspective error.

9.1 Test Setup

The calibration experiment uses a deliberately strong barrel distortion as ground truth (degrees [3,5,7,9,11], constants $[-0.75,\ 0.52,\ -0.12,\ -0.01,\ 0]$) applied to synthetic targets at 1200×800 px. Camera parameters: focal length 15 mm, pixel pitch 15 µm (focal length 1000 px), giving normalization $\mathrm{PSN} = p/f = 0.001$. Five 8×8 checkerboards (49 inner corners each) are placed across the field of view with in-plane (Z-axis) rotations between -6.4 and $+7.9$, covering different radial bands (Fig. 3). Each board has a known square size (29 mm) at a known camera-to-board distance (500 mm), giving an undistorted square spacing of 58 px. A Canterbury Cathedral photograph (Wikimedia Commons, CC BY-SA 2.0) is distorted with the same model as a visual test object.

9.2 Calibration Accuracy

The five boards are pooled and the distortion constants estimated by the direct linear solve of (10) with automatic degree selection (Section VI-B). Using the learned corner detector, the procedure selects a three-term model (degrees [2,3,4]) and achieves a fit residual RMSE of 2.8×10^{-4} in normalized radius over 245 pooled corners, with radial coverage to $r = 0.581$—80% of the undistorted image-corner radius of 0.721 (the distorted corner maps to $r \approx 0.593$ under this model). The estimated function matches the ground-truth function to an RMSE of 6.7×10^{-5} over the covered domain (Fig. 4).



deg	ground truth	estimated
2	–	+0.02151
3	-0.75000	-0.93635
4	–	+0.53268
5	+0.52000	–
7	-0.12000	–
9	-0.01000	–
11	+0.00e+00	–

RMSE 2.827e-04
boards 5
covered 0.5812
img corner 0.5934
PSN 0.00100

Fig. 4. Estimation accuracy via direct linear least squares. Detected corners from five pooled boards (colored points, top) and the estimated distortion curve (magenta) overlaid on ground truth (cyan dashed); the two are visually coincident. Residuals (bottom) are unstructured about zero with RMSE 2.8×10^{-4} in normalized radius. The coefficient table reports the automatically selected three-term model against the five-term ground truth.

9.3 Recovering the Function, Not the Coefficients

A notable property of the result is that the estimated model does not share the degree structure of the ground truth. The ground truth uses five odd-degree terms [3,5,7,9,11]; the automatic selection recovers three terms [2,3,4] with entirely different coefficient values. Despite this, the estimated and true *functions* agree to RMSE 6.7×10^{-5} over the observed domain.

This is the expected and correct behavior. Distortion calibration recovers the distortion function over the radial support spanned by the data—not the specific coefficients of whatever polynomial generated it. Over a bounded domain, a low-order polynomial can reproduce a higher-order one to high accuracy, and the monotonicity-constrained selection finds the most parsimonious such representation. The distinction becomes visible only outside the observed domain: beyond the covered radius, the estimated and true functions diverge, since they are genuinely different polynomials that happen to coincide where data constrained them (Fig. 5).

TABLE 3
Detector Comparison on the Five-Board Strong-Barrel Calibration.

Detector	Selected degrees	$r_{\max}$	RMSE vs. GT
Harris (classical)	[2,3,4]	0.335	5.1×10^{-3}
Learned (EfficientNet-V2-S)	[2,3,4]	0.581	6.7×10^{-5}

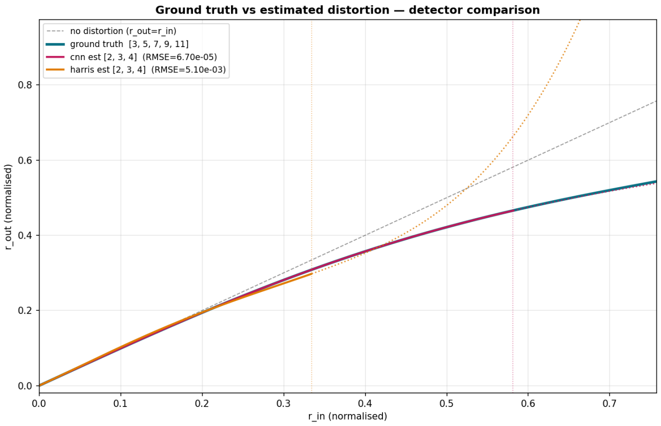


Fig. 5. Ground truth vs. estimated distortion for the two detectors. Within each detector's coverage (solid), the estimate tracks ground truth (cyan); beyond it (dotted) the estimate extrapolates. The Harris estimate (orange) covers only to $r \approx 0.335$ and diverges rapidly outside it; the learned estimate (magenta) covers to $r \approx 0.581$ and remains close to ground truth even into extrapolation. Vertical lines mark each detector's coverage limit.

This reframes what a calibration result means—a faithful reconstruction of the observed mapping, not an identification of latent generating parameters—and motivates reporting the covered radial range alongside any coefficient set.

9.4 Corner Detector Comparison

Detector choice strongly affects the result under heavy distortion. Table 3 and Fig. 5 compare the classical Harris detector with the learned detector on the identical five-board setup.

The learned detector extends usable radial coverage from 0.335 to 0.581 and reduces the error against ground truth by approximately two orders of magnitude. The mechanism is visible in Fig. 5: under strong barrel distortion, the Harris response degrades at large radii where the pattern is compressed, so the classical detector contributes corners only over a limited central band. Its estimated function therefore extrapolates—and diverges sharply—beyond $r \approx 0.335$. The learned detector localizes corners reliably out to the image corner, constraining the fit over nearly the full domain.

9.5 Undistortion with Estimated Coefficients

Applying the estimated model to undistort the five boards and the Canterbury image recovers straight checkerboard edges and rectified architecture across the field (Fig. 6). The end-to-end before/after on the test image is shown in Fig. 7: strong barrel distortion in, rectified image out, using coefficients estimated entirely from the checkerboard data—not the ground-truth coefficients.

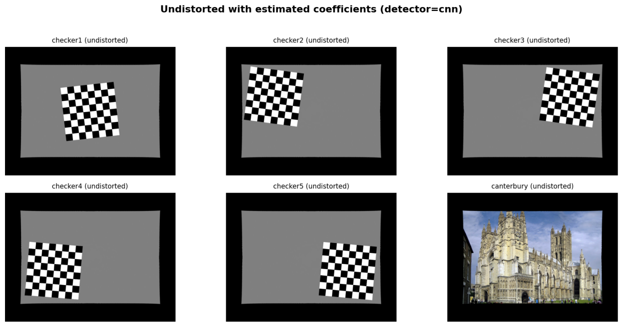


Fig. 6. Undistortion of all five boards and the Canterbury image using coefficients estimated from the boards (learned detector). Checkerboard edges are rectified to straight lines and the cathedral's vertical and horizontal features are restored across the field of view.

Fig. 7. End-to-end pipeline on the test image. *Left:* strong barrel distortion (ground-truth model). *Right:* undistortion using coefficients estimated solely from the calibration boards. The rectification uses no knowledge of the ground-truth coefficients.

9.6 Monotonicity Diagnostics

For the non-monotonic model of Figs. 1 and 2 ($f(r) = r - 0.5r^3 + 0.2r^5 + 0.5r^7 - 2r^9$), the diagnostic over the image domain ($r \leq 0.721$, $\tau = 0.20$) reports $\min \text{slope} = -0.333$, hard_loss_ratio $= 0.049$, and soft_loss_ratio $= 0.041$: the thin band beyond the fold ($r > 0.686$)—just 4.9% of the domain—is correctly flagged as irrecoverable, consistent with the black corners of Figs. 1 and 10. For the strong-barrel calibration model, the same diagnostic returns $\min \text{slope} = +0.408$ with hard_loss_ratio $= 0$ and soft_loss_ratio $= 0$—monotonic over the full image domain, consistent with the successful undistortion of Figs. 6 and 7.

9.7 Calibration and Undistortion Under Non-Monotonic Distortion

A distinguishing test of the framework is its behavior when the ground-truth distortion is itself non-monotonic—a case standard pipelines neither detect nor handle gracefully. We distort the test target with the non-monotonic model of Figs. 1 and 2, $f(r) = r - 0.5r^3 + 0.2r^5 + 0.5r^7 - 2r^9$, whose forward function peaks at the fold radius $r_{\text{fold}} = 0.686$ and decreases beyond it. The model is deliberately strong yet nearly monotonic over most of the frame: the hard-loss band ($r > 0.686$) covers only 4.9% of the image domain (Fig. 2), so the visible content in the distorted image is extensive and well-structured.

Distortion. The forward distortion is bounded: the maximum attainable distorted radius is $f(r_{\text{fold}}) = 0.523$, so undistorted content at $r_{\text{in}} \leq r_{\text{fold}}$ is compressed into the disk of radius 0.523 and no content maps beyond it. Content originally at $r_{\mathrm{in}} > r_{\mathrm{fold}}$ is non-invertible and the pipeline leaves it black in place rather than fabricating it by folding (Fig. 8). The distorted image is therefore a barrel disk—frame dimensions unchanged—with only the thin corner band blacked out.

Calibration. Five checkerboard boards are placed across the monotonic region and distorted with the non-monotonic model. The classical Harris corner detector [9] is used here: within the monotonic region ($r < r_{\mathrm{fold}} = 0.686$) the compression is moderate and the Harris gradient response remains reliable, yielding clean detections across all boards. The calibration data—245 corner pairs covering $r \approx 0$–0.65—are entirely monotonic; the non-monotonicity of the ground truth is never presented to the estimator.

Automatic degree selection (Section VI-B) is run with the monotonicity constraint enforced over the *observed (covered) domain* rather than the full application range (Section X-D). The selection chooses degrees [2,3,4,5,6,7,8] and achieves a fit residual RMSE of 1.43×10^{-4} in normalized radius. The recovered coefficients are numerically large—as expected when many correlated basis functions are fitted over a bounded domain—but the *function* they represent matches the ground truth to an RMSE of 8.75×10^{-4} over the covered domain and is visually coincident with it (Fig. 9). As in Section IX-C, the procedure reconstructs the distortion function, not the specific coefficients: the degree structure [2–8] differs entirely from the ground truth [3,5,7,9].

Accuracy bottleneck. The residual RMSE of 1.43×10^{-4} is governed by corner localization noise—the finite accuracy with which the Harris detector places each corner—rather than by any limitation of the linear estimator or the degree-selection procedure. This was confirmed by supplying the estimator with synthetically generated *exact* (noise-free) corner coordinates: the recovered constants are identical to the ground truth to within machine precision (RMSE $\approx 3.5 \times 10^{-17}$), and the auto-selection correctly assigns zero weight to any degree that does not carry information in exact data. The estimator is exact; the practical accuracy ceiling is set by the detector. Improving corner localization near the fold—where the checkerboard pattern is strongly compressed—is therefore the primary path to lower residuals, and is identified as future work (Section X-E).

Undistortion. Applying the recovered model to the distorted Canterbury image rectifies the interior to an approximately rectangular field (Fig. 10). The corners remain black because the source content there was destroyed in the forward distortion—no algorithm can recover it. This is the geometric signature of the hard-loss region made visible in the output, and matches the prediction of the diagnostic framework exactly.

This experiment closes the loop on the monotonicity framework: the hard_loss_ratio (Section V) predicts the irrecoverable band; the calibration (Section VI), constrained to the monotonic data the boards provide, recovers the invertible sub-function to within detector accuracy; and the undistorted output (Section IV) realizes the maximal valid field, with the black corners as the visible consequence.

Fig. 8. Target distorted with the non-monotonic model $f(r) = r - 0.5r^3 + 0.2r^5 + 0.5r^7 - 2r^9$ (fold at $r = 0.686$). Undistorted content at $r_{\mathrm{in}} \leq r_{\mathrm{fold}}$ is compressed into the disk of radius $f(r_{\mathrm{fold}}) = 0.523$; content beyond the fold is non-invertible and is left black in place. Frame dimensions are preserved. The visible interior is well-structured, confirming that the hard-loss band is thin (4.9% of the image domain).

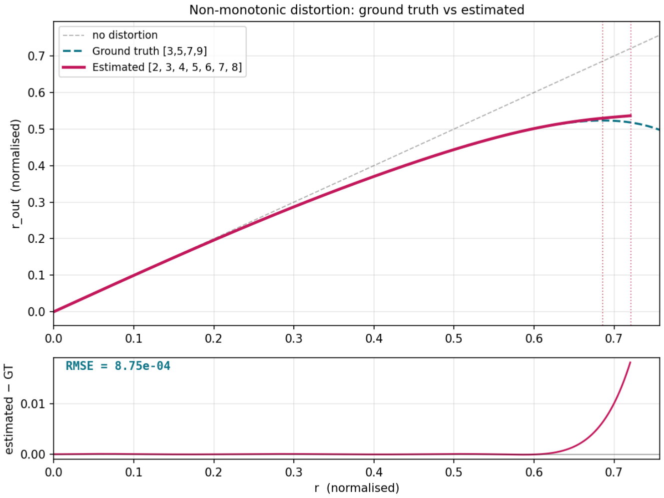


Fig. 9. Ground truth vs. estimated distortion function for the non-monotonic model. The estimated curve (magenta, degrees [2–8]) is visually coincident with the ground truth (teal dashed, degrees [3,5,7,9]) over the covered domain ($r \leq 0.65$), RMSE 8.75×10^{-4}. Beyond the covered range the two curves diverge—a direct illustration of Section IX-C: calibration recovers the function over the observed support, not the underlying coefficients. Vertical lines mark the ground-truth fold (teal) and the limit of the estimated monotonic branch (magenta).

9.8 Model-Family Expressiveness: Controlled Comparison

The preceding experiments use polynomial ground truths that fixed-family models can, in principle, approximate. A sharper question is structural: which realistic distortion profiles can each model family represent at all? To isolate expressiveness from solver and pose effects, all families are fitted to identical $(r_{\mathrm{in}}, r_{\mathrm{out}})$ radial pairs with corner-level localization noise ($\sigma = 0.05\,\mathrm{px}$ at $f = 1000\,\mathrm{px}$), using ideal least-squares solvers—the best case for every family. Five families are compared: the Brown odd polynomial k_1–k_3 [1], the rational model k_1–k_6 [5], the Kannala–Brandt 4-term θ-polynomial [10], the proposed free polynomial

Fig. 10. Undistortion using the recovered model. The interior is rectified to an approximately rectangular field; the corners remain black because the original content there lay beyond the fold radius ($r_{\text{fold}} = 0.686$) and was destroyed in the forward distortion—no undistortion algorithm can recover it. The corner loss is the visible signature of the hard-loss region. Pixels are blacked in place; the frame is not cropped.

TABLE 4
Structural Expressiveness of Model Families: Maximum |Error| in Pixels Against Ground Truth (Ideal Least-Squares Solvers on Identical Radial Pairs, Corner Noise 0.05 px, $f = 1000$ px).

Model family	Smooth barrel (in-family)	Zonal ripple	Foveated knee
Brown odd, 3 terms [1]	0.27	6.04	15.9
Rational k_1–k_6 [5]	0.01	4.43	0.99
Kannala–Brandt 4-term [10]	0.79	4.16	9.34
Ours: free polynomial	0.02	5.66	33.4†
Ours: dictionary (poly + local)	0.32	0.26	< 0.01

†Monotonicity-constrained selection rejects every oscillating candidate and terminates at r^3; the limitation is exposed in the reported residual rather than hidden in an oscillating fit (see text).

with monotonicity-constrained selection, and the proposed *dictionary*—the identical selection procedure over a basis augmented with localized functions (Gaussian zonal terms and integrated-sigmoid knees; Section III-A).

Three ground-truth profiles probe different regimes. The first is a smooth strong barrel (the model of Section 9-A evaluated to $r = 1.05$, ≈93 diagonal field of view)—an in-family control. The second adds a zonal ripple to a mustache profile,

$$f(r) = r - 0.5r^3 + 1.2r^5 + 0.004\, e^{-((r-0.45)/0.06)^2},$$

smooth, strictly monotonic, and physically motivated—mid-spatial-frequency zonal errors of molded aspheric surfaces produce exactly such localized deviations—but outside the span of any fixed family. The third is a foveated two-zone profile,

$$f(r) = r - 0.6\, w \ln\!\big(1 + e^{(r-0.55)/w}\big), \qquad w = 0.03,$$

whose slope falls smoothly from 1.0 to 0.4 across a narrow transition band: the *designed* behavior of "controlled-distortion" optics (panomorph security lenses, foveated automotive cameras) that allocate resolution to a central zone and compress the periphery. Table 4 reports the maximum pixel-level error of each family against each ground truth; Figs. 11 and 13 show the fitted curves, and Fig. 12 applies the ripple model to a real image.

Three observations follow. First, the in-family control is an honest negative: every family fits the smooth barrel to sub-pixel accuracy. For smooth monotonic profiles, family choice is not the accuracy bottleneck, and the proposed framework claims no advantage there. Second, on the out-of-family profiles the fixed families hit a hard structural limit: even with a perfect solver, the best of them (Kannala–Brandt) leaves 4.16 px and the rest 4.4–6.0 px, while the dictionary stays at the 0.26 px noise floor—a 16× gap. A full `cv2.calibrateCamera` run on the same ripple profile (intrinsics fixed—its best case) cannot absorb the zonal feature at all: its reprojection residual rises to 1.6–2.3 px (the fit visibly fails on its own metric) and the recovered function is off by 47–70 px. The family simply has no degree of freedom shaped like the feature. Third, the free polynomial's large error on the knee profile (33.4 px) is the monotonicity constraint operating as designed: every higher-degree candidate that tracks the knee oscillates into negative slope somewhere on the domain and is rejected, so the selection terminates at r^3—a poor fit, but a physically valid one, with the limitation visible in the residual. The fixed families have no such guard; their best knee fits oscillate around the transition (Fig. 13).

The dictionary restores expressiveness without sacrificing the guard. On the ripple profile it reaches 0.26 px against a 0.05 px noise floor; on the knee profile the selection identifies the single atom $\mathrm{knee}(c{=}0.55, w{=}0.03)$—the design parameters of the lens—and reaches the noise floor (maximum error < 0.01 px). With the knee parameters moved off the dictionary grid ($r_0 = 0.57$, $w = 0.045$, neither value among the dictionary atoms), the selection composes neighbouring atoms to a maximum error of 0.12 px. On the in-family control the enlarged dictionary does no harm (0.32 px). The conclusion is not that polynomials are insufficient—by Weierstrass approximation any continuous profile is polynomial-approximable on a bounded interval—but that under a monotonicity constraint and finite noisy data, physically motivated local bases reach accuracy that global bases cannot, while the constraint continues to exclude invalid models.

9.9 Silent Failure of the Inverse-First Pipeline Under Non-Monotonic Distortion

Section 9-G showed the proposed pipeline calibrating and undistorting under a non-monotonic ground truth. The complementary question is how the standard inverse-first pipeline behaves on the same input—and, critically, whether anything in its output warns the user. The same non-monotonic model (fold at $r_{\text{fold}} = 0.686$) distorts the test image; five 7×7 boards are placed entirely in the monotonic region ($r \le 0.66$), so the calibration data contain no direct evidence of the fold—the realistic situation, since corners cannot be reliably detected in the heavily compressed near-fold band. `cv2.calibrateCamera` with intrinsics fixed (its best case) fits the boards to a reprojection RMSE of 0.093 px (k_1–k_3) and 0.067 px (rational k_1–k_6): by the standard figure of merit, both calibrations are excellent. The

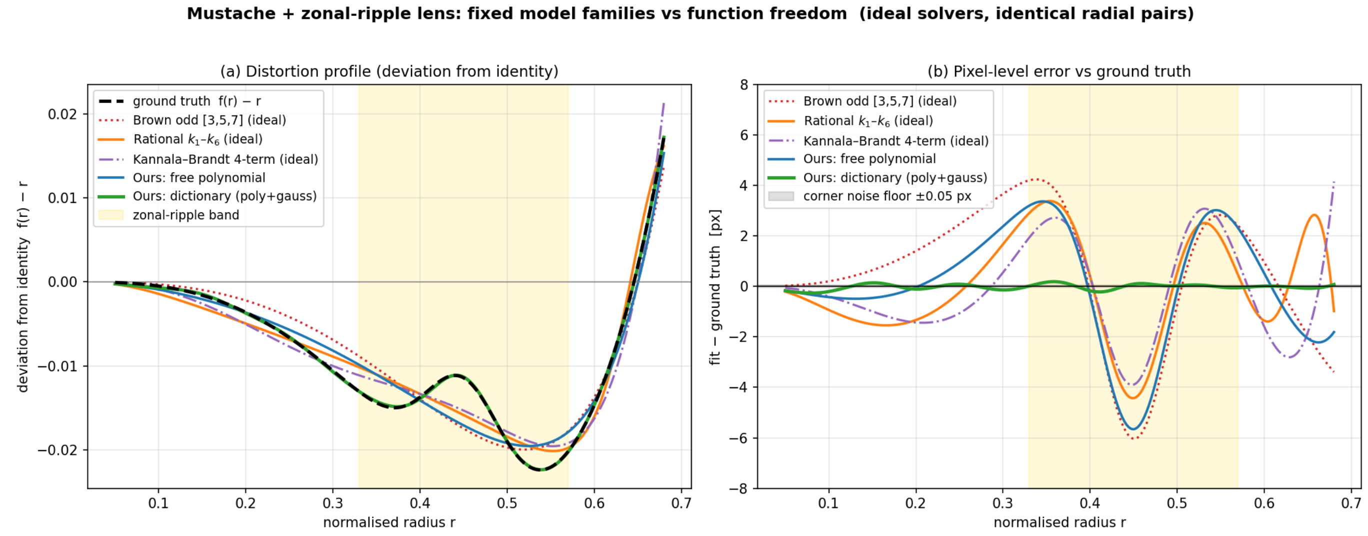


Fig. 11. Mustache + zonal-ripple lens: fixed model families vs. function freedom (ideal least-squares solvers on identical radial pairs, matching Table 4; the two real-`cv2` fits are reported in the text but omitted here for clarity). *(a)* Deviation from identity $f(r) - r$ against ground truth (black dashed); the gold band marks the zonal ripple. *(b)* Pixel-level error of each fit; the gray band is the ± 0.05 px corner-noise floor. The fixed global families (Brown, rational, Kannala–Brandt) and the monotonic free polynomial all swing several px through the ripple band because none has a degree of freedom localized there; only the dictionary fit (green) stays at the noise floor.

proposed estimator on the same corners selects degrees $[3, 2, 9, 6]$ with a fit residual of 4.7×10^{-5} and reports coverage $r \leq 0.622$.

The failure appears at application time. `cv2.initUndistortRectifyMap` evaluates the fitted polynomial over the full output field with no validity guard. Beyond the data support the fitted curves fold—as any polynomial continuation of this profile must—and wherever the fold occurs, destination pixels map back onto *interior* source content: the output contains mirrored, plausible-looking image structure in the affected band, rendered silently (Fig. 14). No error, warning, or flag distinguishes these fabricated pixels from valid ones. The proposed pipeline applies its monotonicity diagnostic before rendering: the fold is detected, the non-invertible band is excluded (blacked), and the hard/soft loss ratios are reported.

The point is not that the OpenCV fit is poor—within its coverage it is accurate—but that the inverse-first workflow has no stage at which invertibility is checked or the valid domain is communicated. Reprojection error is structurally silent about this failure: it is evaluated only at detected corners, which lie inside the monotonic region by construction. The forward-first pipeline makes the valid domain an explicit, reported quantity.

10 DISCUSSION

10.1 The Forward-First Paradigm

The inverse-first convention arose from the practical structure of camera calibration: the target provides ground truth, the camera is the unknown, and estimating the inverse correction is the natural response. This framing is well-suited to joint intrinsic-plus-distortion estimation and will remain the appropriate methodology in that context.

The forward-first paradigm addresses a different starting point: the distortion function is either prescribed (in simulation, synthetic data generation, or optical design) or estimable with known intrinsics. In this regime, defining the forward model precisely unlocks near-exact undistortion and absolute pipeline validation—capabilities with direct value in precision metrology, optical system characterization, and calibration benchmark construction.

10.2 What Calibration Recovers

The detector comparison and the function-vs.-coefficients result together clarify what a distortion calibration delivers. The recovered object is the distortion function over the radial range the data span; coverage, not coefficient structure, governs validity. A model that fits well over a limited band may extrapolate poorly, and two correct calibrations may report entirely different coefficient sets while describing the same lens over their shared domain. Reporting radial coverage alongside coefficients is therefore essential, and extending coverage—here, through a learned detector robust to strong distortion—directly improves the usable accuracy of the result.

10.3 Physical Meaning of the Monotonicity Metrics

The hard_loss_ratio is not a software error metric—it is a statement about the information content of the image. A non-zero value means the distorted image does not contain sufficient information to recover the original in the affected radial band, regardless of the algorithm. Reporting this explicitly, tied to a fraction of the image domain, translates a theoretical invertibility condition into a practical, actionable indicator. The soft_loss_ratio complements it: low-slope regions are recoverable given the exact model but attenuate local detail, increasing noise sensitivity.

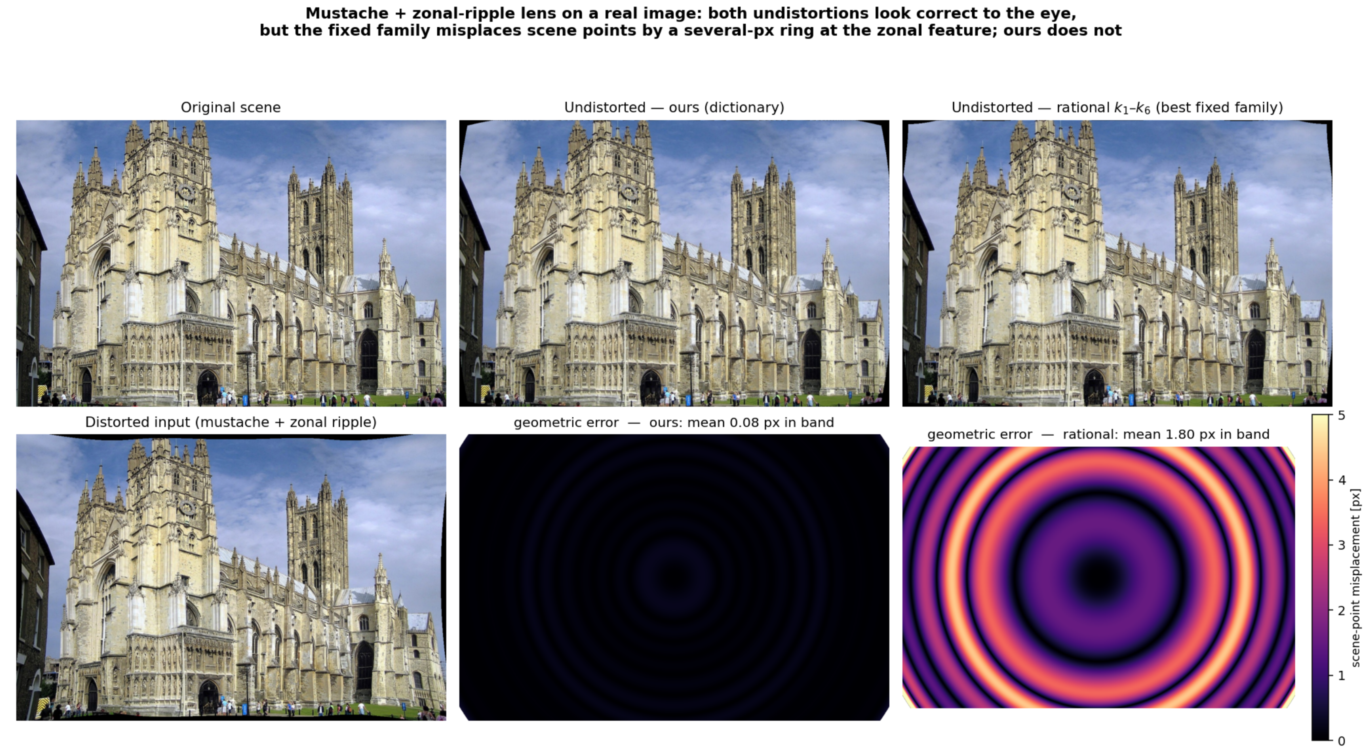


Fig. 12. The same zonal-ripple model applied to a real image and undistorted with each fit. *Top:* original scene; undistortion with our dictionary fit; undistortion with the best fixed family (rational k_1–k_6, ideal solver). All three are perceptually indistinguishable. *Bottom:* the distorted input, then the geometric error map—the per-pixel scene-point misplacement $|g(r) - f_{\mathrm{GT}}(r)|$, which is texture-independent. The rational fit misplaces scene points by a bright concentric ring (mean 1.8 px, peak 11.8 px in the zonal band) because it cannot represent the local feature; the dictionary fit leaves no visible structure (mean 0.08 px). The error is invisible to the eye in the undistorted image yet would corrupt any downstream metrology at the affected radii. Test image: Canterbury Cathedral, Wikimedia Commons, CC BY-SA 2.0.

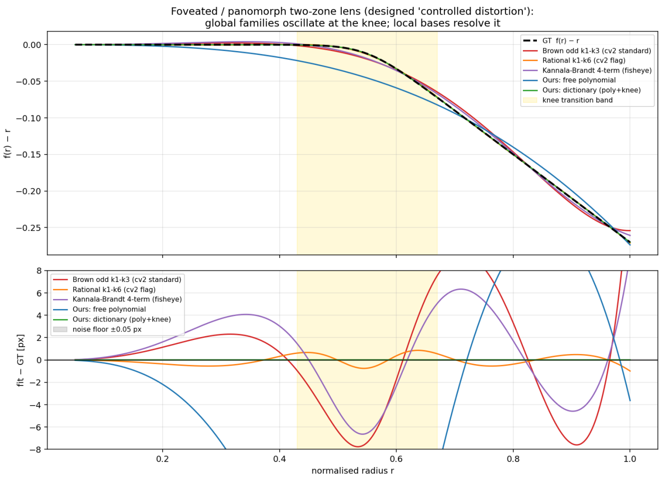


Fig. 13. Foveated/panomorph two-zone lens. The slope of the ground truth (black dashed) falls from 1.0 to 0.4 across a narrow transition band (gold). Global polynomial and rational families oscillate around the knee; the monotonicity-constrained free polynomial refuses every oscillating candidate and reports a large, honest residual; the dictionary fit with knee atoms is visually coincident with the ground truth (max error < 0.01 px) and identifies the design parameters ($c = 0.55$, $w = 0.03$) directly.

10.4 Monotonicity-Constrained Selection vs. Post-Hoc Checking

The degree selection of Section VI-B enforces monotonicity structurally, not as a downstream filter. A degree that improves RMSE but causes folding is excluded from consideration. This guarantees the output is always physically valid regardless of data, threshold, or degree range—a guarantee a post-hoc filter cannot provide without manual iteration. The domain over which monotonicity is enforced is itself a choice: by default it spans the full application range (so undistortion fills the frame without corner folds), but under a non-monotonic ground truth (Section IX-G) it is restricted to the observed domain, where the data lie and the function is monotonic. With that setting the estimation reduces to an ordinary monotonic fit over the invertible sub-function; the boundary of the recoverable field is then fixed by the distorted image itself—the disk of radius $f(r_{\mathrm{fold}})$ beyond which no content exists. The method does not fail or fit spurious oscillations; it recovers the invertible branch and produces the maximal recoverable field.

10.5 Limitations

The calibration module uses focal length and pixel pitch for coordinate normalization. This requirement is not fundamental: unlike projection-based calibration methods where focal length is an intrinsic parameter of the 3-D-to-2-D projection model and cannot be eliminated, the role of focal

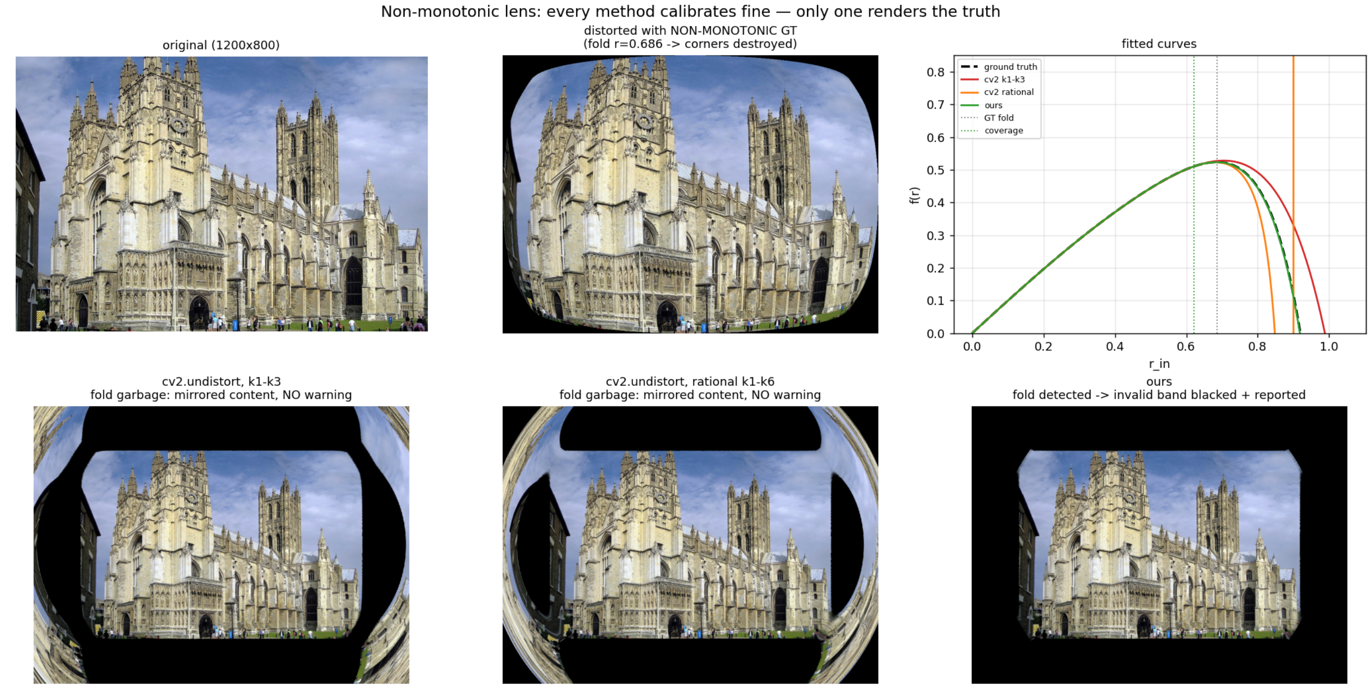


Fig. 14. Non-monotonic distortion: silent failure of the inverse-first pipeline. *Top row:* original image; image distorted with the non-monotonic ground truth (fold at $r = 0.686$, corners destroyed); fitted curves of all methods against ground truth, with the fold and each coverage limit marked. *Bottom row:* full-field undistortion rendered by $\mathtt{cv2}$ with k_1–k_3 and with the rational model—both contain mirrored "fold garbage" with no warning, despite sub-0.1 px reprojection error—and by the proposed pipeline, which detects the fold, blacks the invalid band, and reports hard/soft loss.

length here is solely to define a normalized radial coordinate. Alternative normalization schemes that remove this requirement are straightforward extensions. Joint estimation of intrinsics is not addressed and would require incorporating the projection model, reintroducing nonlinearity. Only radial distortion is modeled; tangential (decentering) components are not included. Board tilt about the X or Y axis introduces unmodeled perspective effects that increase calibration residuals; Z-axis (in-plane) rotation, used throughout the experiments here, is handled correctly.

10.6 Corner Localization Under Strong Compression

The accuracy of the direct linear estimator is bounded by the precision of the corner detector, not by any property of the solver itself: supplying exact corner coordinates yields machine-precision coefficient recovery (RMSE $\approx 3.5 \times 10^{-17}$, Section IX-G). In the non-monotonic experiment the practical ceiling is set by Harris localization noise ($\sim$0.15 px), which translates to a function-level RMSE of 8.75×10^{-4}. Near the fold, strong radial compression minifies the checkerboard pattern below the resolution for which the current learned detector was trained, limiting reliable coverage to $r \approx 0.65$. A contributing factor is the detector's fixed 256-px inference resolution, which further reduces the apparent size of already-compressed squares; at that scale, near-fold squares fall below the 8-px minimum of the training distribution. Extending the learned detector's training distribution to include heavily compressed patterns—specifically, near-fold configurations where square sizes fall below $\sim$10–15 px in the distorted image—together with higher-resolution training and inference (or tiled inference at native resolution), is identified as the primary direction for improving non-monotonic calibration accuracy. The classical Harris detector is sufficient in the monotonic region; the learned detector will be required to push coverage reliably toward the fold.

11 CONCLUSION

We have presented a forward-first framework for radial lens distortion modeling whose distinguishing property is that estimation is monotonicity-constrained: the recovered mapping is invertible by construction in every case, never folding the way an unconstrained reprojection-error fit can. Around this it provides hard/soft monotonicity diagnostics that quantify physically irreversible vs. recoverable information loss, an open basis (polynomial plus local functions) for profiles outside the fixed families, and a deterministic linear calibration with a deterministic inverse whose round trip is validated, at the interpolation-noise floor, against a synthetic ground truth—the same yardstick we hold the baselines to. On a five-board calibration with strong barrel distortion, the estimated distortion function matches ground truth to an RMSE of 6.7×10^{-5} over the observed domain, automatically reduced to a three-term monotonic model, with a learned corner detector extending radial coverage and improving accuracy by two orders of magnitude over a classical detector. Controlled expressiveness comparisons trace the complementary boundary: fixed-family models (Brown, rational, Kannala–Brandt) leave pixel-level errors up to 16 px on realistic out-of-family profiles that the open-basis formulation fits to the noise floor, and a standard pipeline under a non-monotonic ground truth calibrates to sub-0.1 px reprojection error yet silently renders fabricated content that the proposed diagnostics detect and bound.

The generalized polynomial encompasses both odd and even degrees; calibration is a direct linear solve; output geometry adapts to the distortion extent. For monotonic models the recovered image is visually indistinguishable from the original. The implementation is available as a commercial toolbox for MATLAB and Python at opticaldistortionlab.com.

Several extensions are identified as future work. First, the focal length requirement of the current normalization is not fundamental; an intrinsic-free formulation—deriving the normalized domain directly from image geometry—is a natural next step. Second, the calibration currently assumes planar boards with in-plane rotation only; out-of-plane tilt introduces perspective distortion that is not modeled. A correction based on iterative homography estimation, applied as a preprocessing step before the linear solve and leaving the core estimation structure intact, could extend the framework to tilted boards and is under investigation. Third, corner localization near a distortion fold is the current accuracy bottleneck for non-monotonic calibration: the classical Harris detector performs reliably in the monotonic region, but the strong radial compression near the fold reduces checkerboard square sizes below the training distribution of the current learned detector. Extending the training set to include heavily compressed near-fold patterns—where square sizes fall to 10–15 px in the distorted image—is expected to push reliable coverage closer to the fold and reduce the function-level RMSE by approximately one order of magnitude relative to the present Harris baseline; raising the detector's inference resolution above the current 256 px, by retraining at higher resolution or by tiled native-resolution inference, addresses the same bottleneck from the architectural side. Together these extensions move toward a calibration pipeline requiring no prior knowledge of camera intrinsics, tolerant of arbitrary board orientation, and capable of accurately characterizing strongly non-monotonic optical systems.

The core insight—that defining the forward distortion function precisely makes monotonicity enforceable during estimation and the corrected image checkable against a synthetic ground truth—turns physical invertibility from an unmonitored assumption into a guarantee, and may motivate renewed attention to forward modeling in precision optical instrumentation.

APPENDIX

Brown's formulation and equation (3) are algebraically equivalent for odd-degree terms:

$$r_u(1 + K_1 r_u^2 + K_2 r_u^4 + \cdots) = r_u + K_1 r_u^3 + K_2 r_u^5 + \cdots$$

Hence $k_i = K_i$ for the standard terms. Equation (3) generalizes by removing the restriction that degrees be odd.

REFERENCES

[1] D. C. Brown, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, 1966.

[2] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[3] G. Bradski, "The OpenCV library," *Dr. Dobb's Journal of Software Tools*, 2000.

[4] J.-Y. Bouguet, "Camera Calibration Toolbox for Matlab," California Institute of Technology, 2004.

[5] D. Claus and A. W. Fitzgibbon, "A rational function lens distortion model for general cameras," in *Proc. IEEE CVPR*, 2005, pp. 213–219.

[6] J. H. Brito, R. Angst, K. Köser, and M. Pollefeys, "Radial distortion self-calibration," in *Proc. IEEE CVPR*, 2013, pp. 1368–1375.

[7] T. Akenine-Möller, E. Haines, and N. Hoffman, *Real-Time Rendering*, 4th ed. CRC Press, 2018.

[8] M. Tan and Q. V. Le, "EfficientNetV2: Smaller models and faster training," in *Proc. ICML*, 2021, pp. 10096–10106.

[9] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vision Conference*, 1988, pp. 147–151.

[10] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1335–1340, Aug. 2006.

[11] A. E. Conrady, "Decentred lens-systems," *Monthly Notices of the Royal Astronomical Society*, vol. 79, pp. 384–390, 1919.

[12] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Autom.*, vol. RA-3, no. 4, pp. 323–344, Aug. 1987.

[13] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 10, pp. 965–980, Oct. 1992.

[14] J. Heikkilä and O. Silvén, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE CVPR*, 1997, pp. 1106–1112.

[15] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Proc. IEEE CVPR*, 2001, pp. 125–132.

[16] F. Devernay and O. Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, no. 1, pp. 14–24, 2001.

[17] J. Mallon and P. F. Whelan, "Precise radial un-distortion of images," in *Proc. ICPR*, 2004, pp. 18–21.

[18] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[19] C. Ricolfe-Viala and A.-J. Sánchez-Salmerón, "Lens distortion models evaluation," *Applied Optics*, vol. 49, no. 30, pp. 5914–5928, 2010.

[20] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *Proc. IEEE/RSJ IROS*, 2006, pp. 5695–5701.

[21] J. Rong, S. Huang, Z. Shang, and X. Ying, "Radial lens distortion correction using convolutional neural networks trained with synthesized images," in *Proc. ACCV*, 2016, pp. 35–49.

[22] A. Duda and U. Frese, "Accurate detection and localization of checkerboard corners for calibration," in *Proc. BMVC*, 2018.

[23] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[24] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[25] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—a modern synthesis," in *Vision Algorithms: Theory and Practice* (LNCS 1883). Berlin, Germany: Springer, 2000, pp. 298–372.

[26] R. Hartley and S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1309–1321, Aug. 2007.

[27] R. E. Barlow, D. J. Bartholomew, J. M. Bremner, and H. D. Brunk, *Statistical Inference Under Order Restrictions: The Theory and Application of Isotonic Regression*. New York, NY, USA: Wiley, 1972.

[28] J. O. Ramsay, "Monotone regression splines in action," *Statistical Science*, vol. 3, no. 4, pp. 425–441, 1988.

[29] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. MICCAI* (LNCS 9351), 2015, pp. 234–241.

[30] O. Bogdan, V. Eckstein, F. Rameau, and J.-C. Bazin, "DeepCalib: A deep learning approach for automatic intrinsic calibration of wide

field-of-view cameras," in *Proc. 15th ACM SIGGRAPH Eur. Conf. Visual Media Production (CVMP)*, 2018, pp. 1–10.

[31] M. López, R. Marí, P. Gargallo, Y. Kuang, J. Gonzalez-Jimenez, and G. Haro, "Deep single image camera calibration with radial distortion," in *Proc. IEEE/CVF CVPR*, 2019, pp. 11817–11825.