Geometry-Consistent Alignment for Time-Lapse Crack Monitoring in Field-Based SHM

Xinxin Sun* (ORCID: 0000-0002-9963-2109) and Peter Chang (ORCID: 0009-0006-5819-9819)

Department of Civil and Environmental Engineering, University of Maryland, College Park, MD 20742, USA

*Corresponding Author

Email: xinxin68@umd.edu

Abstract

Keypoint-based detectors such as SIFT and SURF—built on Gaussian scale-space—are widely used in image-based structural health monitoring (SHM), but frequently fail under field conditions. Their reliance on isotropic smoothing suppresses fine crack edges and yields unstable alignment under occlusion, surface texture, or varying illumination. Binary descriptors like ORB and BRISK, while faster, inherit similar drawbacks due to Gaussian pyramids and weak structural saliency.

This study presents a geometry-consistent alignment framework that addresses these challenges by repurposing anisotropic diffusion–based feature detection and integrating it with RANSAC-driven homography estimation. Compared to the isotropic blurring of Gaussian filtering, anisotropic diffusion constructs a nonlinear scale space that selectively smooths homogeneous regions while preserving crack-relevant discontinuities—enhancing keypoint stability under low contrast and visual clutter. RANSAC further performs robust outlier rejection to fit a reliable geometric transformation under sparse or noisy conditions.

The proposed pipeline is unsupervised, training-free, and calibration-free—designed for crack imagery acquired by UAV or handheld platforms. Validated on over 100 field-acquired image pairs across varied degradation scenarios—including textured masonry, strong perspective distortion, and moving shadows—the method achieved average area errors below 5% and spine length errors under 15% relative to manual ground truth. In comparison, classical methods frequently exhibited 2–4× higher errors; for instance, SURF showed over 25% length error in occluded scenes due to clustered inliers, while SIFT failed to recover alignment under similar conditions.

Fully interpretable and lightweight, the framework supports scalable crack evolution tracking without data annotation or parameter tuning. Its compatibility with mobile sensing and transparency in edge preservation lay the foundation for integration into smart SHM workflows and adaptive control systems. Future directions include stereo or depth-based correction for non-planar surfaces and integration with learned descriptors—extending the interpretable, calibration-free foundation established in this work. Furthermore, unsupervised mapping of crack evolution patterns from aligned imagery may enable cluster-based or topology-informed monitoring strategies.

Keywords: Smart structural health monitoring; Crack evolution tracking; Anisotropic diffusion; UAV-based inspection; Training-free alignment; Vision-based diagnostics; Homography estimation; Unsupervised feature matching

1. Introduction

Structural Health Monitoring (SHM) increasingly relies on image-based diagnostics to assess material degradation under real-world conditions. Surface cracks, as early indicators of fatigue, corrosion, or bond deterioration, play a central role in non-destructive evaluation [1], [2], [3]. Tracking their evolution over time is critical for assessing structural safety and serviceability [4], [5], [6].

Recent structural failures—such as the 2024 collapse of the Francis Scott Key Bridge—underscore the risks associated with undetected crack propagation and have renewed interest in scalable, vision-based SHM frameworks [7]. Compared to dense sensor networks, image-based approaches offer a cost-effective and flexible solution for long-term infrastructure monitoring [8], [9]. However, field-acquired images often suffer from viewpoint variations, shadows, surface textures, and dynamic lighting, which introduce geometric inconsistencies that compromise crack tracking fidelity [10], [11], [12].

Even modest angular deviations can lead to significant geometric errors. In controlled experiments (see Fig. 1.1 and Table 1.1), we observed crack area changes exceeding 5%, length deviations over 15%, and width fluctuations up to 25% under slight perspective offsets—demonstrating the importance of alignment correction in time-lapse SHM workflows.

Fig. 1.1: Sample crack images acquired under slight viewpoint variations.

Table 1.1: Variations in crack measurements due to geometric distortion (in pixels and percentage error)

Image	Crack Area	Area Change	Spine Length	Length Change	Avg. Width	Width Change
Image 1	4968	-	1339	-	4.0	
Image 2	5231	+5.3%	1558	+16.4%	3.9	–2.5%
Image 3	5236	+5.4%	1149	-14.2%	5.0	+25.0%

Note: Full measurement values are reported in Table 1.1, but only percentage variations are discussed in the main text for clarity.

Feature-based image registration methods are commonly employed to address this issue. Classical detectors such as SIFT, SURF, ORB, and BRISK are widely used in SHM and vision tasks [13], [14], yet their reliance on Gaussian scale-space often suppresses fine crack boundaries—particularly in shadowed or textured scenes. Binary descriptors like ORB, while computationally efficient, degrade in noisy or uneven lighting [15], [16]. In a shadow-interference test case (Table 3.9), SIFT and ORB achieved fewer than 30% correct matches, resulting in alignment drift over 20 pixels.

Additionally, most current image alignment approaches are tuned for single-frame or laboratory-controlled conditions and fail to support consistent, long-term monitoring [17]. These limitations motivate the need for a geometry-consistent alignment workflow suited for SHM deployments.

This study constructs a lightweight, calibration-free pipeline for aligning crack imagery across field-acquired sequences. Rather than proposing a new algorithm, we integrate anisotropic diffusion–based scale space construction with RANSAC-driven homography estimation to preserve crack-relevant features under real-world visual distortions. Compared to DoG- and binary-descriptor methods, our framework retains greater edge fidelity and stabilizes image geometry in time-lapse sequences without requiring training data, calibration, or manual tuning.

In a series of field-simulated tests, over 100 crack images from concrete and masonry structures were evaluated across four degradation scenarios: oblique views, partial occlusion, high surface texture, and dynamic shadows. Under these conditions, classical detectors showed frequent failure, while the proposed workflow maintained alignment stability and spatial coherence.

The contributions of this study are as follows:

1. We present a geometry-consistent alignment workflow tailored for SHM monitoring of crack evolution under field-induced distortions;
2. We demonstrate that anisotropic diffusion preserves crack-localizing edge features more reliably than traditional Gaussian detectors;
3. We validate the system on four representative degradation scenarios, achieving high alignment consistency without training, calibration, or supervised tuning.

2. Methodology

This study addresses a core challenge in vision-based SHM: maintaining geometric consistency across time-lapse crack images captured under variable field conditions. UAV or handheld imagery is particularly prone to viewpoint variation, perspective distortion, and acquisition noise. When combined with crack propagation caused by aging, loading, or environmental exposure, these effects distort spatial correspondence and compromise the reliability of evolution tracking.

To mitigate such effects, a robust alignment workflow is required—one that corrects field-induced distortion while preserving fine structural features critical to consistent monitoring. Feature-based image registration offers a practical strategy, typically comprising two stages: (1) keypoint detection and (2) geometric transformation estimation. Once repeatable features are identified across image pairs, a global transformation (e.g., homography) is computed to align them.

In current SHM practice, detectors such as SIFT, SURF, ORB, and BRISK are frequently used due to their speed and multi-scale design [18]. These algorithms rely on Gaussian-filtered pyramids and saliency-based keypoint selection. However, field imagery often contains narrow, low-contrast cracks obscured by shadow, noise, or material texture. Under such conditions, these detectors show reduced localization accuracy and inlier consistency—prompting the need for edge-preserving alternatives.

Empirical results in this study confirm that traditional detectors frequently fail to identify stable, crack-relevant features under field acquisition. This limitation stems from the smoothing nature of Gaussian filtering, which suppresses high-frequency components and weakens crack-continuity cues. These observations motivate the integration of feature detectors better suited to the geometric and visual properties of crack imagery.

The following subsections summarize the mathematical formulation of four classical detectors—SIFT, SURF, ORB, and BRISK—and analyze their performance limitations in SHM tasks.

2.1 Traditional Feature Detection Methods

Feature-based alignment pipelines are widely used to correct geometric distortion in SHM imagery. This section reviews four classical approaches and highlights the specific limitations they exhibit when applied to crack monitoring under field conditions.

2.1.1 SIFT

SIFT constructs a Gaussian scale space:

$$L(x, y, \sigma) = G(x, y, \sigma) \cdot I(x, y) \tag{1}$$

and detects keypoints via Difference of Gaussians:

$$DoG(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}$$

where k controls scale progression [19].

2.1.2 SURF

SURF approximates Hessian determinants using box filters:

$$Det(H) = L_{xx}L_{yy} - (0.9 \cdot L_{xy})^2 \tag{3}$$

where L_{xx}, L_{xy}, and L_{yy} as second-order Gaussian derivatives [20].

2.1.3 ORB

ORB applies FAST for keypoint detection:

$$|I(p) - I(p_i)| > T \tag{4}$$

and employs BRIEF for binary descriptors [21].

2.1.4 BRISK

BRISK forms binary descriptors by comparing intensity pairs:

$$Decriptor = \sum_{i,j} sgn(I(p_i) - I(p_j)) \tag{5}$$

[22].

2.1.5 Limitations of Gaussian-Based Approaches

All four methods use Gaussian pyramids for scale invariance. While effective in well-textured scenes, they perform poorly under field conditions involving shadows, blur, or fine-scale cracks. Gaussian filtering suppresses high-frequency edges as shown by:

$$L(x, y, \sigma_{n+1}) = \frac{1}{4}\sum_{m,n} G[m, n] \cdot L(2x + m, 2y + n, \sigma_n) \tag{6}$$

This effect causes unstable or missing keypoints along crack trajectories. ORB and BRISK, though not explicitly DoG-based, inherit similar weaknesses due to their Gaussian-pyramid foundations. These limitations reduce alignment robustness and compromise the continuity of crack evolution tracking.

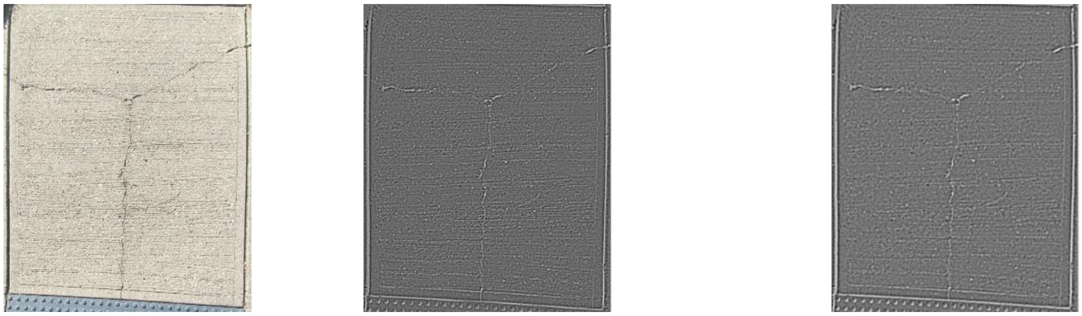
2.2 Failure Modes of Gaussian-Based Detectors in SHM Crack Alignment

2.2.1 Crack Detail Loss from Gaussian Scale-Space

In Gaussian-based detectors such as SIFT, the scale parameter σ is incremented geometrically:

$$\sigma_{i+1} = \sigma_i \cdot \sqrt{2} \tag{7}$$

While this facilitates scale invariance, it also causes progressive blurring that degrades fine crack structures. As shown in Fig. 2.1, increased σ reduces edge clarity in the Difference of Gaussian (DoG) images, diminishing keypoint saliency and matchability.



(a) Original image

(b) DoG at $\Delta\sigma = (1.6 \cdot \sqrt{2}) - 1.6$

(c) DoG at $\Delta\sigma = (1.6 \cdot (\sqrt{2})^2) - (1.6 \cdot \sqrt{2})$

Fig. 2.1: DoG Images at Various Scales

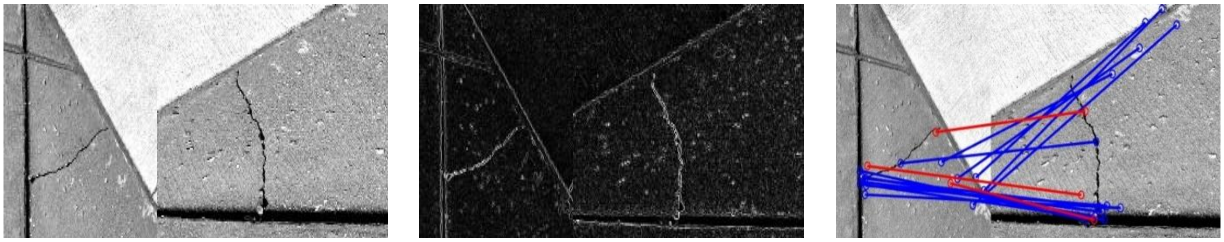
In SHM applications, this blurring leads to inconsistent crack localization across image sequences, making evolution tracking unreliable—particularly for fine cracks under field noise.

2.2.2 Influence of Weak Gradients in Low-Contrast Conditions

Gradient-based detectors rely on strong intensity transitions, defined as:

$$|G(x, y)| = \sqrt{(\frac{\partial I}{\partial x})^2 + (\frac{\partial I}{\partial y})^2} \tag{8}$$

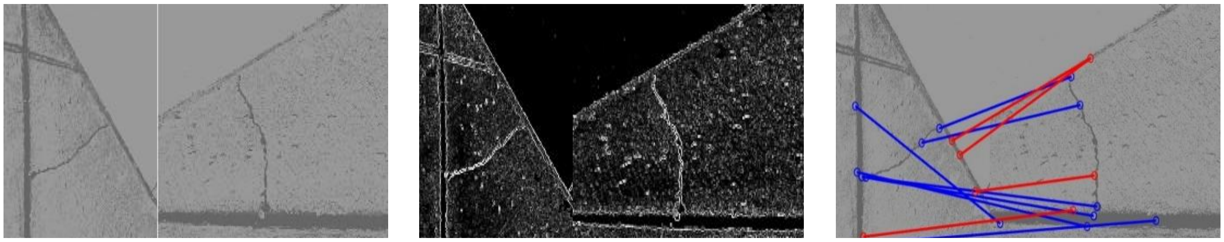
In low-contrast crack imagery, gradients are weak due to minimal brightness differences, leading to unstable keypoints and poor matches.



(a) High-contrast crack images

(b) Gradient magnitude of high-contrast images

(c) Feature matching results for high-contrast images



(d) Low-contrast crack images

(e) Gradient magnitude of low-contrast images

(f) Feature matching results for low-contrast images

Fig. 2.2: Effect of contrast on gradient magnitude and matching accuracy

As illustrated in Fig. 2.2, high-contrast images (a–c) exhibit strong gradient magnitudes and yield a high number of correct feature matches (blue lines). In contrast, low-contrast images (d–f) produce blurred gradients, increased background interference, and a higher rate of mismatches (red lines), resulting in degraded alignment accuracy. These findings underscore the need for alternative detectors that preserve fine edge details and remain stable under visually degraded conditions. KAZE, an edge-preserving detector based on nonlinear diffusion, addresses these challenges directly.

2.3 KAZE Integration for Edge-Preserving Alignment in Crack Monitoring

To address the geometric instability introduced by Gaussian-based methods, this study integrates the KAZE framework—an edge-preserving, scale-aware detector—into a robust alignment workflow tailored for SHM.

$$\frac{\partial L}{\partial t} = div(c(x, y, t) \cdot \nabla L) \tag{9}$$

where $c(x, y, t)$ is a contrast-adaptive conductivity function:

$$c(x, y, t) = \frac{1}{1+(\frac{|\nabla L|}{\kappa})^2} \tag{10}$$

This formulation preserves high-gradient structures (e.g., crack edges) by reducing diffusion near discontinuities, while suppressing noise in homogeneous regions.

Unlike Gaussian smoothing, KAZE maintains edge sharpness across scale levels and adapts to local structure. This makes it especially effective in SHM scenarios featuring shadows, blur, and low contrast [14], [23], [24], [25], [26].

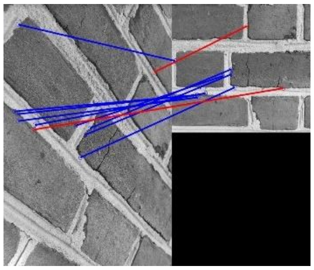
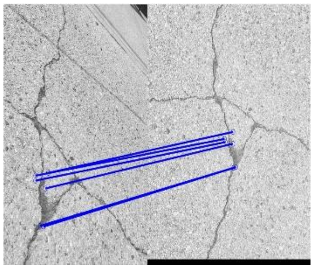
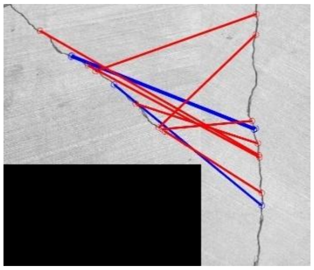
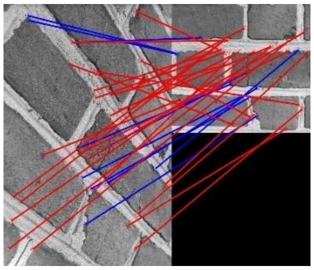
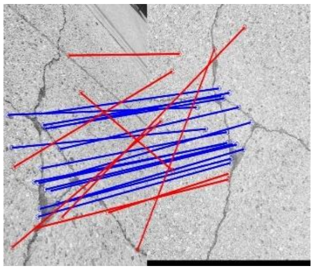
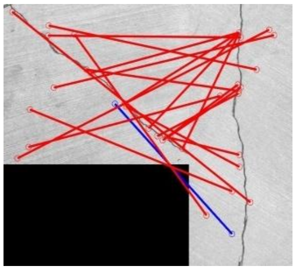
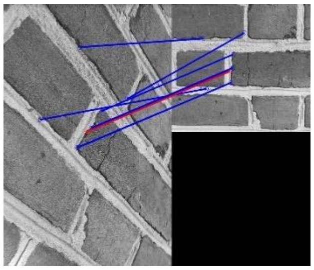
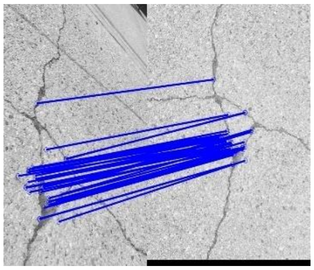
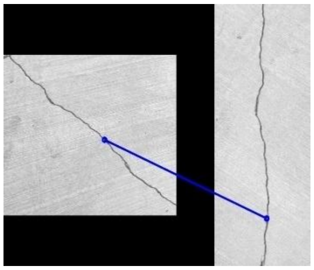
As shown in subsequent sections, KAZE significantly outperforms traditional detectors in geometry-consistent crack alignment across time-lapse sequences— supporting its application in long-term SHM monitoring scenarios, where consistent image alignment is required under uncontrolled conditions.

2.4 Comparative Evaluation of Feature Detectors Under Challenging Conditions

To evaluate the suitability of KAZE for SHM crack alignment, we conducted controlled experiments comparing its performance against four widely used detectors—SIFT, SURF, ORB, and BRISK—under field-simulated conditions. All images were captured using an iPhone 11 at chest height (~1.3–1.4 m) under natural outdoor lighting, without artificial enhancement. The evaluation covered five key visual conditions—texture, contrast, noise, blur, and perspective distortion—designed to simulate the most common challenges encountered in SHM field imagery. Each detector was assessed using both qualitative keypoint correspondence and quantitative alignment stability, measured by RANSAC-derived inlier counts.

Table 2.1. Keypoint Matching Results Across Texture Conditions (High, Medium, Low)

The images illustrate keypoint matches overlaid on crack surfaces under varied texture complexity.

	High Texture	Medium Texture	Low Texture
Images			
SIFT			
SURF			
ORB			

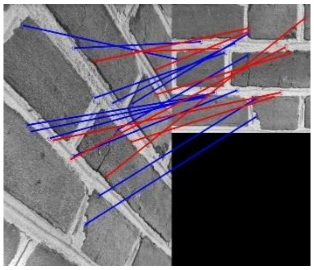
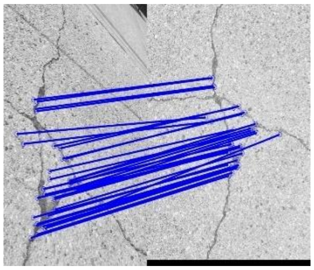
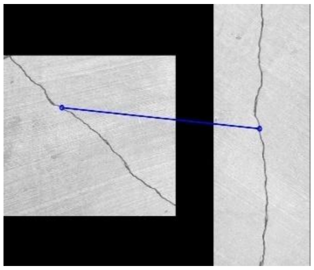
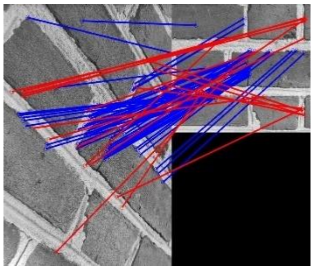
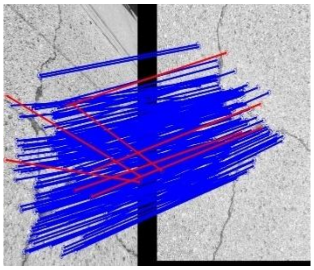
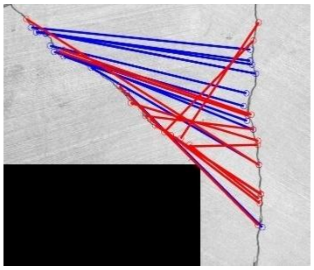
BRISK			
KAZE			

Table 2.2. Mean Inlier Counts by Detector Across Texture Levels

Inlier counts reflect the number of geometrically consistent matches retained by RANSAC-based homography estimation.

Texture \ Mean Inliers	SIFT	SURF	ORB	BRISK	KAZE
High	9	10	5	5	53
Medium	6	17	43	35	180
Low	3	1	1	1	14

Across all simulated SHM scenarios, KAZE consistently outperformed SIFT, SURF, ORB, and BRISK in both match density and geometric alignment stability. Under low-texture, low-contrast, high-noise, and blurred conditions—common in real-world monitoring—traditional detectors showed significant performance degradation, often producing few or no valid matches. In high-texture and high-contrast settings, all methods achieved moderate success, though KAZE still demonstrated denser and more spatially consistent matches. However, under low-texture, low-contrast, high-noise, and blurred conditions—common in real-world SHM—traditional detectors exhibited severe performance degradation, often producing zero or near-zero valid matches. In contrast, KAZE retained critical crack features and maintained stable correspondence, with inlier counts exceeding baseline methods by 2–10× depending on the condition. These results demonstrate that KAZE's nonlinear scale-space enables robust retention of crack-relevant edge features, providing stable alignment under diverse field-induced visual disturbances—supporting its integration in long-term SHM workflows. Supplementary comparisons under additional contrast, noise, blur, and distortion conditions are reported in Appendix A (Tables A.1–A.8), supporting the evaluation of KAZE's robustness and generalizability across diverse SHM scenarios.

2.5 Outlier Elimination and Homography Estimation

Although KAZE improves feature stability, real-world SHM imagery—affected by shadows, occlusions, or surface clutter—can still lead to mismatched correspondences. To ensure reliable alignment, we integrate RANSAC [27] to filter out outliers and retain only geometrically consistent matches [28], [29], [30], [31]. We estimate the homography H relating two image planes using a Direct Linear Transform (DLT) model, constructed from keypoint correspondences. A homogeneous system A·h = 0 is solved via Singular Value Decomposition (SVD), yielding the least-squares estimate of the 3×3 projective matrix H.

$$\mathrm{X}_2 = H \cdot \mathrm{X}_1 \tag{11}$$

Where $\mathrm{X}_1 = [x_1 \quad y_1 \quad 1]^T$ is a point in the source image, and H is the 3×3 projective transformation matrix:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \tag{12}$$

To estimate *H*, a linear system $A \cdot h = 0$ is constructed from *n* correspondences, where h is the 9-element vectorized form of *H*. This homogeneous system is solved via Singular Value Decomposition (SVD):

$$A = U\Sigma V^T \tag{13}$$

The solution h is the last column of *V*, minimizing the algebraic error [32]:

RANSAC proceeds by randomly sampling 10 correspondences per iteration to compute candidate homographies. For each match, the reprojection error is computed as:

$$error_i = \sqrt{(x_2 - {x_2}^{Projected})^2 + (y_2 - {y_2}^{Projected})^2} \tag{14}$$

Inliers are determined using a dynamically refined threshold based on the root-mean-square reprojection error, updated iteratively after each model estimation.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N} {error_i}^2} \tag{15}$$

The required number of iterations N_{iter} is adjusted based on the estimated outlier ratio *e* and confidence level $p = 0.99$:

$$N_{iter} = \frac{\log{(1-p)}}{\log{(1-(1-e)^k)}} \tag{16}$$

where k =10 is the sample size per iteration. The best homography is selected based on the highest inlier count or lowest residual error and refined using least-squares optimization over all inliers. This process yields stable geometric alignment for time-lapse SHM, ensuring measurement consistency across varying acquisition conditions.

The complete RANSAC-based homography estimation process—including random sampling, error-based inlier filtering, dynamic iteration control, and final model refinement—is summarized in Fig. 2.3. This procedure ensures robust alignment despite viewpoint variation and visual noise, making it well-suited for SHM applications in uncontrolled field environments. This forms the foundation of the alignment pipeline described in the following implementation section.

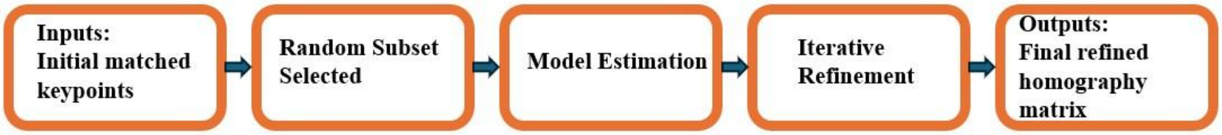


Fig. 2.3. Schematic of RANSAC-based homography estimation, highlighting random sampling, inlier filtering, and model refinement steps for crack image alignment.

2.6 Integrated Implementation and Workflow

To ensure deployment compatibility and reproducibility, all experiments were conducted using MATLAB's built-in `detectKAZEFeatures` function without parameter tuning. This method constructs a nonlinear scale space via additive operator splitting (AOS) to solve the anisotropic diffusion PDE, preserving edge continuity and enhancing descriptor stability on low-texture SHM surfaces. No manual parameter tuning was applied to ensure reproducibility and deployment consistency.

For geometric alignment, a custom RANSAC framework was implemented using Direct Linear Transform (DLT) and Singular Value Decomposition (SVD). Key configuration details include:

- Sampling strategy: 10 matched keypoints per iteration, balancing stability and computational efficiency.
- Inlier classification: Reprojection errors below $\sqrt{5.99} \cdot \sigma$ were accepted, with σ iteratively refined from inlier RMS error.
- Outlier ratio: $e = 0.5$; Confidence level: $p = 0.99$;
- Iteration count: Adaptively updated per Equation (18) using the evolving inlier ratio.
- Sigma refinement: Updated after each iteration based on current homography error.

This adaptive, interpretable implementation proved robust across varied field scenarios involving distortion, noise, and sparse keypoints—especially relevant to UAV-based SHM.

Figure 2.4 summarizes the complete alignment workflow, from keypoint extraction to homography estimation. This modular structure supports reproducible deployment and robust crack tracking under real-world acquisition conditions. The full alignment process completes in approximately 50–60 seconds per image pair on a Dell XPS 9530 laptop (Intel Core i7, 16GB RAM, no GPU acceleration), confirming its practicality for field-level deployment without specialized hardware, especially for periodic inspection or UAV-driven crack imaging.

Fig. 2.4: Full alignment pipeline from keypoint extraction to geometric correction

3. Experiments and Analysis

This section presents four scenario-driven experiments designed to evaluate the geometric alignment robustness of the proposed KAZE-RANSAC framework under representative field conditions commonly encountered in SHM imagery. The validation is based on a dataset of approximately 100 images collected from diverse concrete and masonry surfaces exhibiting varying geometric and visual complexities.

Each experiment evaluates a reference–target image pair that simulates a specific SHM challenge, including perspective distortion, occlusion, high-texture backgrounds, and dynamic lighting. All images were captured using the rear wide camera of an Apple iPhone 11 (12 MP, 1/2.55″ sensor, f/1.8 aperture, 26 mm equivalent focal length), held at chest height (~1.3–1.4 m) under natural outdoor lighting, without artificial enhancement. This configuration represents typical acquisition conditions for UAV- or handheld-based SHM inspections, ensuring both practical relevance and reproducibility.

All experiments involved five detectors—SIFT, SURF, ORB, BRISK, and KAZE—paired with RANSAC-based homography estimation. For SIFT through BRISK, the number of sampled keypoints per iteration varied based on available detections. In contrast, KAZE used a fixed 10-keypoint sampling strategy across all cases to isolate the impact of keypoint quality. For SIFT, SURF, ORB, and BRISK, the number of keypoints per iteration was dynamically adjusted based on detection output. In contrast, KAZE-RANSAC employed a consistent 10-keypoint sampling strategy across all scenarios, isolating the effect of keypoint quality and enabling fair comparison under controlled iterations.

Table 3.1 summarizes the experimental scenarios and associated visual challenges. The assigned challenge level in Table 3.1 reflects the severity of geometric and visual interference present in each scenario and is independent of any algorithmic performance outcome. Among the four scenarios, the "Cropped Crack Images" case posed the most severe challenge, with extreme perspective distortion and partial occlusion. Only the proposed KAZE-RANSAC framework successfully recovered alignment in this setting, justifying its "Very High" difficulty rating.

Table 3.1: Experimental Scenarios and Field-Induced Visual Distortion Types

Experiment	Distortion Type	Interference Type	Challenge Level
Ideal perspective	Moderate geometry	Low visual interference	Baseline
Cropped crack images	Severe geometry	Occlusion, truncation	Very High
Textured brick background	Severe geometry	High-frequency texture	High
Moving shadow	Moderate geometry	Dynamic lighting	High

3.1: Ideal Perspective Correction with Clear Four-Point Correspondences

This baseline experiment assesses alignment performance under moderate geometric distortion using a flat concrete slab with clear four-point corner references. A fine-width shrinkage crack was imaged from two viewpoints: vertical (baseline) and oblique. The scene contained minimal occlusion, low background noise, and no texture interference.

All five detectors—SIFT, SURF, ORB, BRISK, and KAZE—were tested using RANSAC-based homography estimation. Due to the oblique viewing angle, the crack in Image 2 exhibited a 45% reduction in area, 41% reduction in length, and 9% reduction in average width prior to correction (Table 3.2).

Table 3.2 summarizes the uncorrected dimensional differences due to geometric distortion.

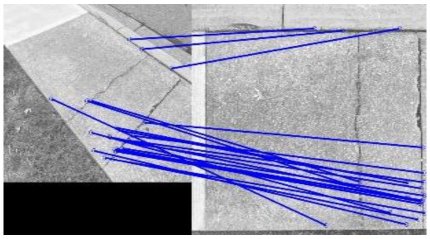
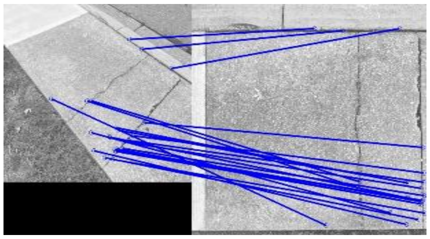
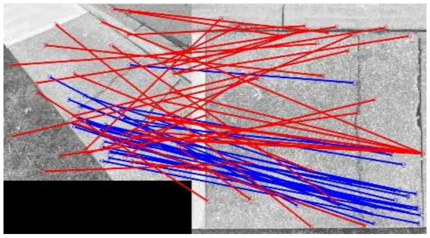
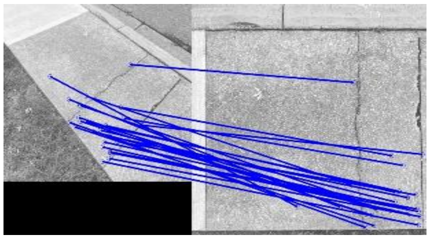
Table 3.2: Measured crack dimensions before geometric correction (in pixels)

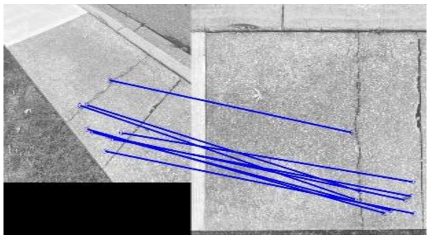
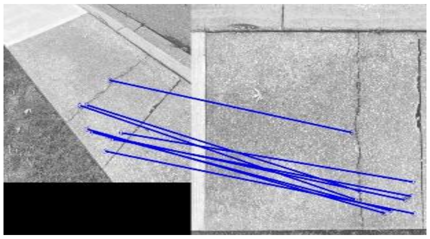
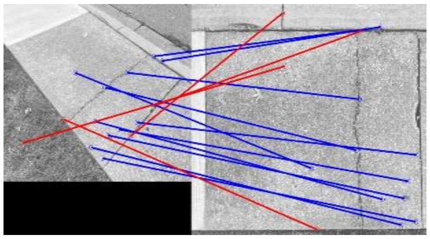
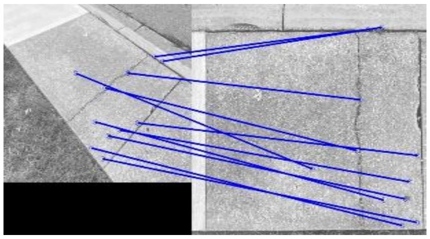
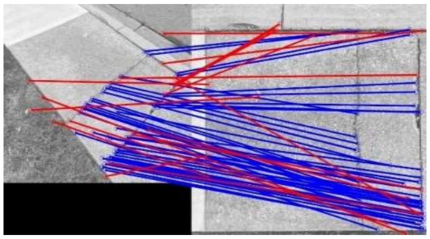
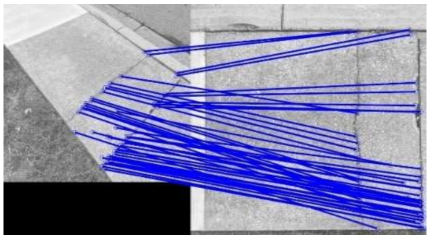
View angle	Baseline	Oblique
Image		
Crack area	2422	1335
Spine length	768	451
Crack average width	3.2	2.9

After alignment, Table 3.3 presents four sub-analyses: (a) match counts before and after RANSAC filtering, (b) visual alignment accuracy using ground anchor overlap, (c) dimensional recovery compared to the baseline, and (d) spatial distribution of inliers with qualitative match interpretation.

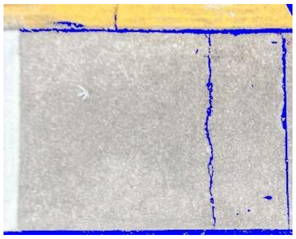
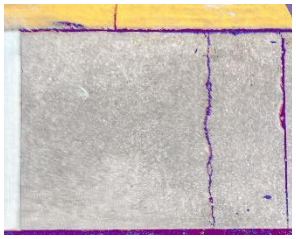
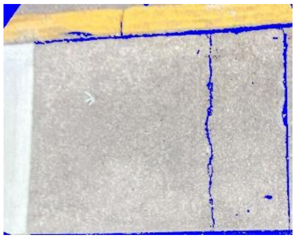
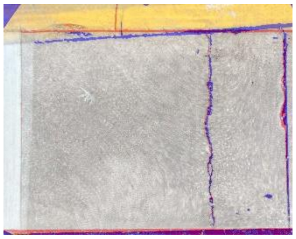
Table 3.3: Keypoint matching results and geometric alignment analysis

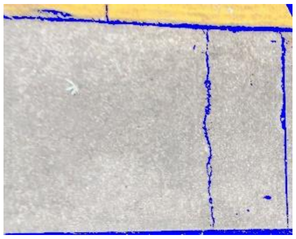
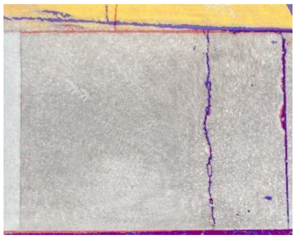
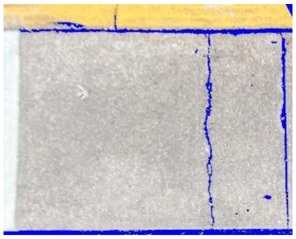
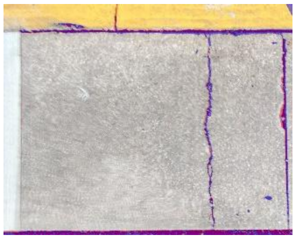
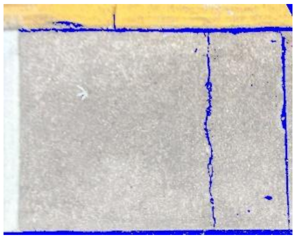
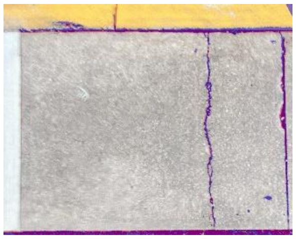
(a) Match counts before/after RANSAC

	Matches Before RANSAC	Matches After RANSAC
SIFT		
SURF		

ORB		
BRISK		
KAZE		

(b) Crack overlap assessment

Method	Corrected Image 2 (Blue)	Baseline Image 1 (Red)	Overlap Accuracy
SIFT			
SURF			

ORB			
BRISK			
KAZE			

(c) Dimensional recovery vs. baseline (pixels and % error)

	Baseline	SIFT	SURF	ORB	BRISK	KAZE
Crack Area	2422	2532	2410	2713	2463	2367
Area Error (%)	–	5	0.5	12	2	2
Spine Length	768	671	622	657	620	784
Length Error (%)	–	13	19	14	19	2
Average Crack Width	3.2	3.6	3.6	3.9	3.7	3.2
Width Error (%)	–	13	13	22	16	0

(d) Spatial inlier breakdown and match interpretation

Detector	Total Inliers	Crack-Aligned	Background	Structural Coverage	Reliability Summary
SIFT	22	2	7	Weak	Overconcentrated on concrete edges; low crack sensitivity

SURF	23	3	18	Very weak	Match density high but spatially misaligned; crack underrepresented
ORB	9	3	6	Moderate	Clean matches, but insufficient crack anchoring
BRISK	18	2	7	Weak	Border-biased matching with minor crack focus
KAZE	59	13	6	Strong	Focused on cracks and corner anchors; structurally meaningful

Among the five methods, only KAZE-RANSAC achieved consistent geometric correction across all metrics, with ≤2% error in area, length, and width. Its 59 total inliers included 13 crack-centered matches and several anchor-aligned points, yielding a well-constrained homography. The resulting overlay showed seamless alignment across the full crack span.

SIFT- and BRISK-RANSAC also yielded accurate visual results. Although each retained only two crack-aligned inliers, their matches were spatially balanced, enabling complete crack overlap with <5% dimensional error.

SURF-RANSAC, while producing 23 inliers, concentrated most (18) in irrelevant textured regions. Despite reporting only 0.5% area error, its weak spatial distribution led to visible misalignment—highlighting that numerical proximity does not guarantee perceptual accuracy.

ORB-RANSAC retained 9 inliers, with 3 on the crack. However, the sparse and uneven match layout failed to constrain the transformation, resulting in incomplete crack reconstruction.

These results highlight that spatial relevance and structural anchoring are more critical than raw match count in ensuring accurate geometric alignment for crack evolution tracking. This experiment confirms that KAZE's nonlinear scale-space effectively preserves fine edge structures and delivers measurement-grade alignment fidelity—outperforming classical detectors even under ideal, low-distortion conditions.

3.2: Perspective Correction for Cropped Crack Sections

This experiment evaluates alignment performance under compounded challenges involving both severe geometric distortion and partial crack visibility—conditions frequently encountered in real-world SHM applications such as constrained inspections near walls, joints, or sensor occlusions. The tested crack appeared on a concrete surface with bifurcated geometry and limited edge continuity, and was partially cropped out of frame in both reference and distorted views.

The reference image was captured with minimal distortion, while the second view introduced strong perspective skew and partial occlusion. Cracks were only partially visible and lacked strong geometric anchors. Visual clutter and the cropped field of view impaired consistent keypoint detection and alignment, making this an ideal stress test for evaluating detector robustness.

All five detectors—SIFT, SURF, ORB, BRISK, and KAZE—were tested in conjunction with RANSAC.

Table 3.4 summarizes the geometric deformation prior to correction, including a 24% drop in crack area and a 21% drop in spine length.

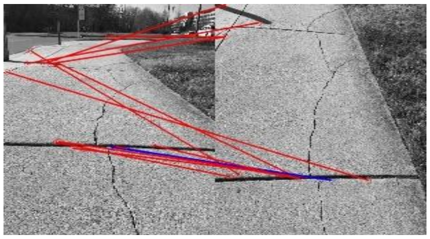
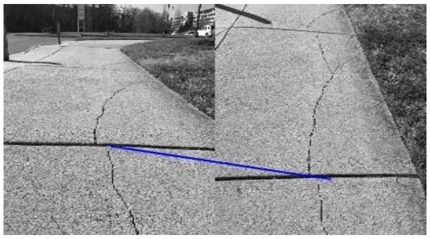
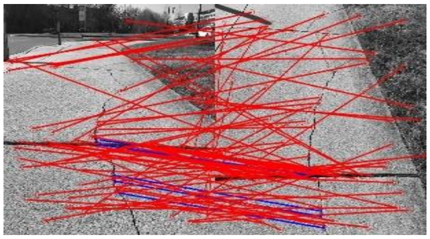
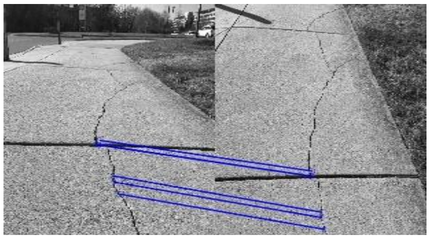
Table 3.4. Crack dimensions before perspective correction (units: pixels)

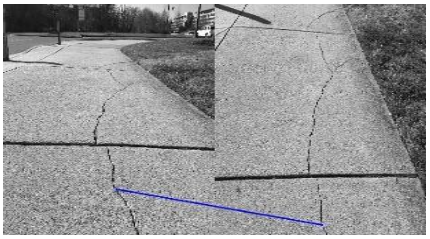
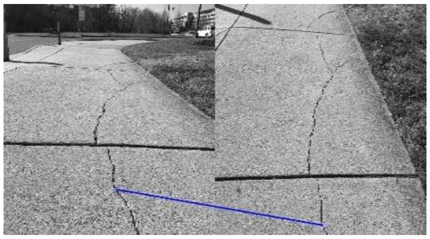
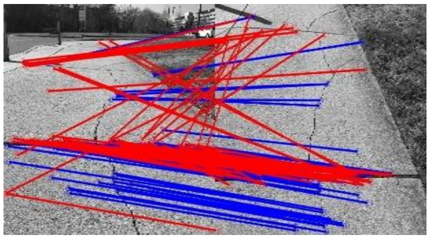
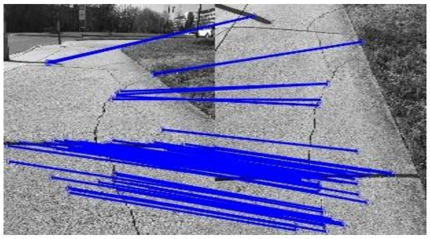
View angle	Baseline	Oblique
Image		
Crack area	1887	1425
Spine length	596	468
Crack average width	3.0	2.9

Table 3.5 presents post-correction results, including match statistics, geometric accuracy, and inlier distribution.

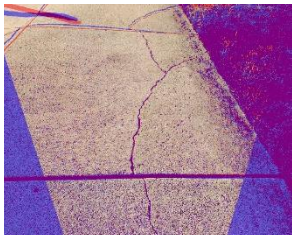
Table 3.5: Keypoint Matching Results and Alignment Analysis Under Occlusion

(a) Match counts before/after RANSAC

	Matches Before RANSAC	Matches After RANSAC
SIFT		
SURF		
ORB	N/A	N/A

BRISK		
KAZE		

(b) Crack overlap assessment

Method	Corrected Image 2 (Blue)	Baseline Image 1 (Red)	Overlap Accuracy
SIFT	N/A	N/A	Failed
SURF			
ORB	N/A	N/A	Failed
BRISK	N/A	N/A	Failed
KAZE			

(c) Dimensional recovery vs. baseline (pixels and % error)

	Baseline	SIFT	SURF	ORB	BRISK	KAZE
Crack Area	1887	–	1695	–	–	1815
Area Error (%)	–	–	10	–	–	4

Spine Length	596	–	439	–	–	509
Length Error (%)	–	–	26	–	–	15
Average Crack Width	3.0	–	3.4	–	–	3.2
Width Error (%)	–	–	13	–	–	7

(d) Spatial inlier breakdown and match interpretation

Detector	Total Inliers	Crack-Aligned	Background	Structural Coverage	Reliability Summary
SIFT	1	0	0	None	Too few matches; insufficient for crack anchoring
SURF	6	5	1	Weak	Mostly crack-aligned, but too sparse for reliable correction
ORB	0	0	0	None	No valid correspondences detected
BRISK	1	1	0	Very weak	Isolated match on crack; no support for global geometry
KAZE	58	23	10	Strong	Focused on cracks and corner anchors; structurally meaningful

Note: “Background” refers to valid RANSAC inliers not located on the crack path, such as edge features or slab texture points. Corner anchors were not individually counted but contributed to total inlier counts.

Despite the compounded challenges of partial occlusion and strong perspective distortion, only KAZE-RANSAC produced both geometric fidelity and structural coherence—maintaining a well-constrained homography and reliable overlap with the baseline image. It was the only method to consistently produce inliers along the crack and structural anchors, resulting in a well-constrained homography and near-perfect visual overlap with the baseline. KAZE also preserved spatial consistency across the cropped region, minimizing projection error even without full contextual visibility.

In contrast, SURF-RANSAC, though able to detect five crack-centered matches, failed to establish a stable homography. Its inliers were spatially clustered, lacking the distribution necessary for global correction, resulting in 26% spine length error and skewed geometry.

SIFT, ORB, and BRISK failed to generate valid homographies under this condition, due to insufficient or spatially clustered inliers. ORB detected no inliers; SIFT and BRISK each retained only one, neither of which provided sufficient spatial distribution. Although BRISK’s match lay directly on the crack, it lacked geometric leverage to constrain the transformation.

These findings reinforce that in real-world SHM scenarios, spatial coverage and structural anchoring are essential for accurate geometric alignment—particularly when crack visibility is limited or contextual cues are missing. KAZE’s nonlinear scale-space preserved edge continuity and resisted visual clutter, enabling reliable correction even when classical detectors failed entirely.

3.3: Perspective Correction on Highly Textured Brick Surfaces

This experiment evaluates alignment accuracy in the presence of high-frequency textured backgrounds—a common challenge in masonry façade inspections where fine cracks coexist with dominant mortar patterns. The tested crack, embedded in a vertical brick wall, was surrounded by repetitive joints and sharp edges that generate dense but semantically irrelevant keypoints.

Two views were acquired: one front-facing as baseline and the other from a strong oblique angle to simulate constrained inspections. The distance to the wall surface was approximately 0.5 meters.

Perspective distortion caused substantial geometric inflation in the oblique image: crack area increased by 176%, spine length by 26%, and width by 38%, highlighting how repetitive textures can amplify distortion effects and mislead feature-based alignment.

Table 3.6: Crack dimensions before perspective correction (units: pixels)

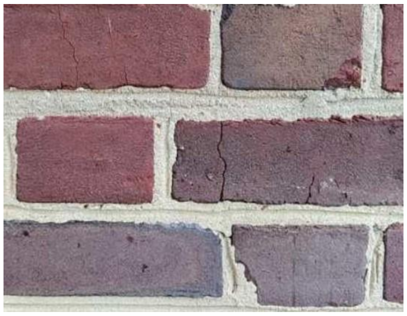
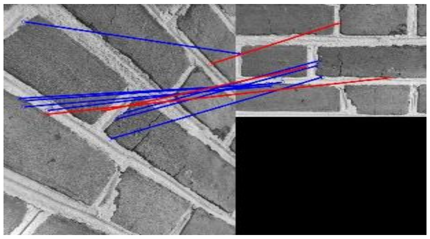
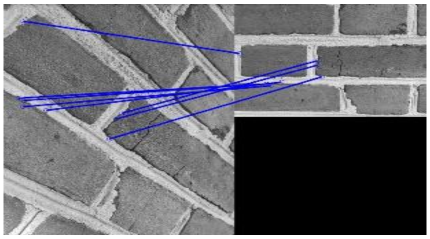
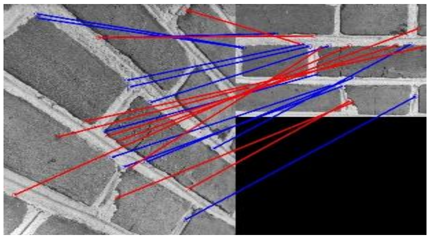
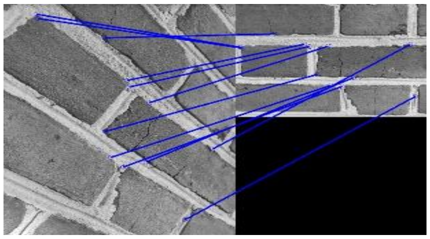
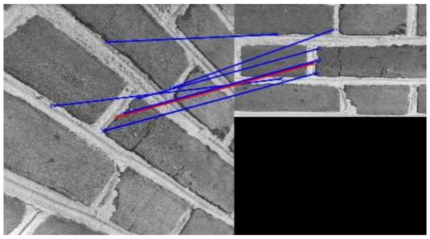
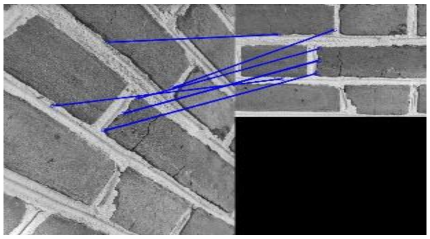
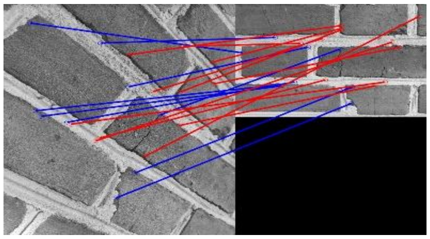
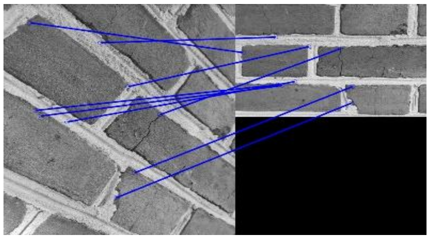
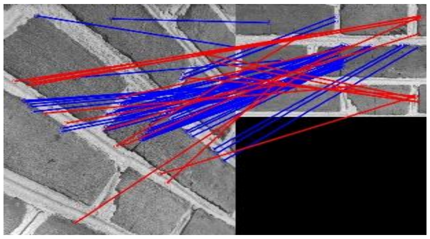
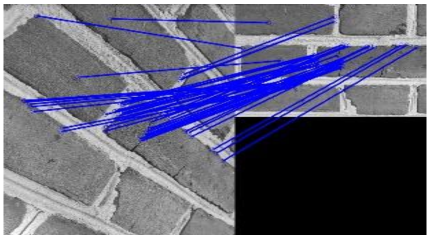
View angle	Baseline	Oblique
Image		
Crack area	278	767
Spine length	103	130
Crack average width	2.6	3.6

Table 3.7 presents post-correction results, including matching performance, alignment accuracy, and inlier distribution.

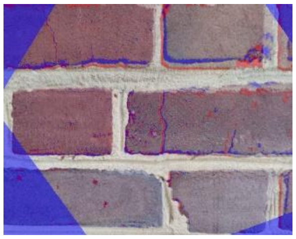
Table 3.7: Keypoint Matching Results and Alignment Accuracy on Textured Surfaces

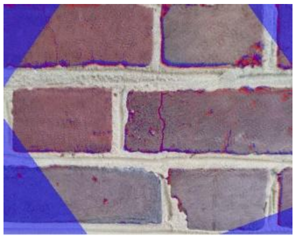
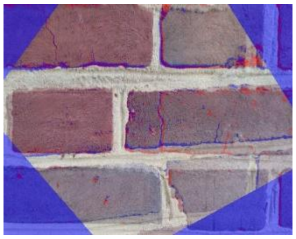
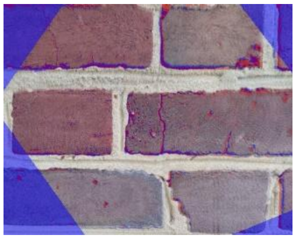
(a) Match counts before/after RANSAC

	Matches Before RANSAC	Matches After RANSAC
SIFT		

SURF		
ORB		
BRISK		
KAZE		

(b) Crack overlap assessment

Method	Corrected Image 2 (Blue)	Baseline Image 1 (Red)	Overlap Accuracy
SIFT			

SURF			
ORB			
BRISK			
KAZE			

(c) Dimensional recovery vs. baseline (pixels and % error)

	Baseline	SIFT	SURF	ORB	BRISK	KAZE
Crack Area	278	291	288	264	303	286
Area Error (%)	–	5	4	5	9	3
Spine Length	103	110	110	101	107	109
Length Error (%)	–	7	7	2	4	6
Average Crack Width	2.6	2.5	2.5	2.5	2.6	2.5
Width Error (%)	–	0.4	0.4	0.4	0	0.4

(d) Spatial inlier breakdown and match interpretation

Detector	Total Inliers	Crack-Aligned	Background	Structural Coverage	Reliability Summary
SIFT	8	0	1	Weak	Concentrated near bricks; no crack-specific focus
SURF	12	0	4	Weak	Background dominant; low semantic relevance
ORB	6	0	1	Very weak	Minimal spatial structure; fragile alignment
BRISK	9	1	2	Moderate	Limited crack sensitivity; partial correction
KAZE	53	10	3	Strong	Edge-preserving; crack and anchor focused

Note: "Background" refers to valid RANSAC inliers not located on the crack path, such as edge features or slab texture points. Corner anchors were not individually counted but contributed to total inlier counts.

All five pipelines produced valid homographies, but crack alignment accuracy varied substantially.

KAZE-RANSAC achieved the most reliable results. Among 53 inliers, 10 were directly aligned with the crack, with others located at corner features. The resulting transformation preserved full crack continuity with only 3% area and 0.4% width error, highlighting resilience to structured background interference.

SURF-RANSAC and BRISK-RANSAC reported low numerical errors (≤9%), but these values were misleading, as most matches concentrated on repetitive textures rather than crack edges.

SIFT-RANSAC, though producing visually acceptable results, misaligned the crack due to mortar-focused keypoints. ORB-RANSAC performed weakest, with insufficient match structure to recover full alignment.

These results indicate that while textured backgrounds may increase keypoint density, only KAZE consistently anchors semantically meaningful features. Its nonlinear diffusion process preserves subtle crack cues while suppressing irrelevant patterns—making it particularly effective for masonry surface diagnostics, where crack signals must be preserved despite high-frequency visual background structure.

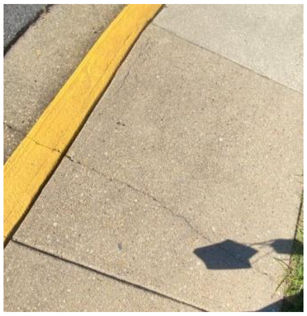
3.4: Perspective Correction for Cracks Under Moving Shadow Conditions

This experiment evaluates alignment robustness under dynamic shadow interference—a common challenge in field-deployed SHM inspections. Moving shadows from personnel or equipment introduce luminance gradients that obscure crack edges and generate spurious keypoints.

Two images were acquired at oblique and frontal angles, with one cast under uniform daylight and the other containing a manually introduced moving shadow. This scenario simulates real-world inspection variability caused by nonuniform lighting.

Before correction, shadow distortion caused crack area to increase by 36% and spine length by 14%, while average width remained unchanged—highlighting the impact of lighting artifacts on geometric accuracy.

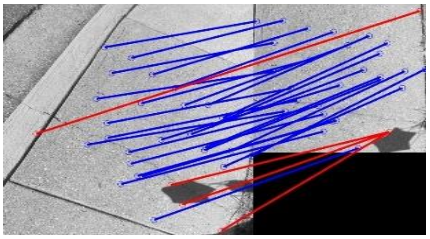
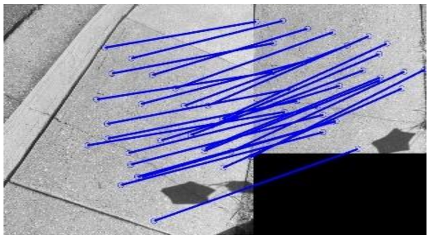
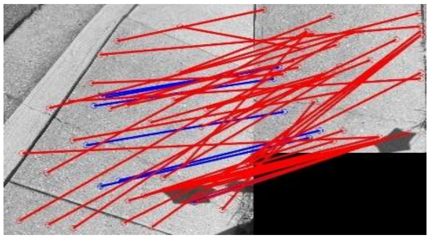
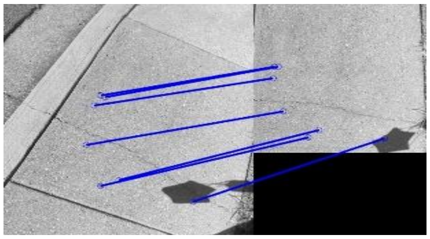
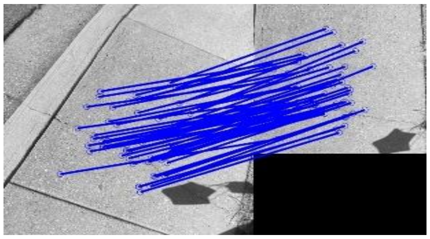
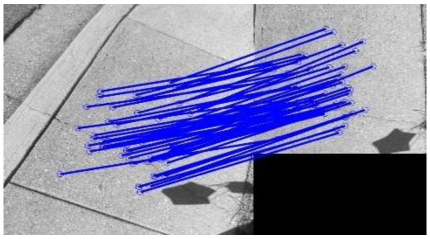
Table 3.8: Crack dimensions before perspective correction (units: pixels)

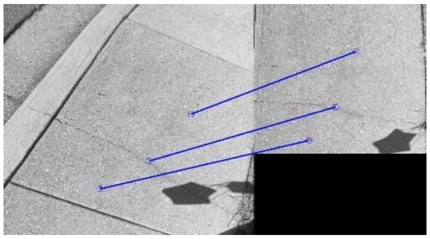
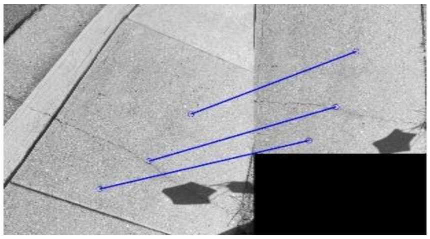
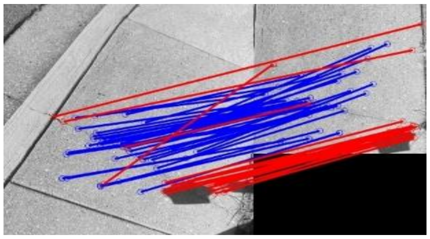
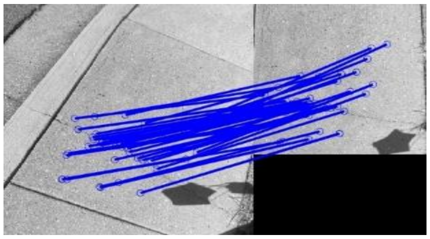
View angle	Baseline	Oblique
Image		
Crack area (pixels)	571	774
Spine length (pixels)	392	445
Crack average width	2.1	2.1

Post-correction results are summarized in Table 3.9, including match statistics, alignment quality, and inlier analysis.

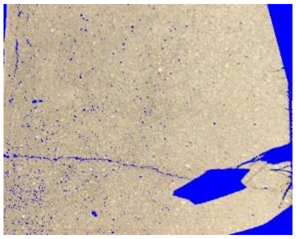
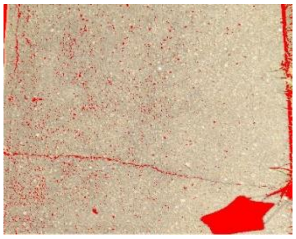
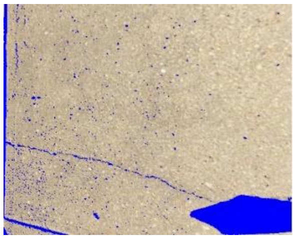
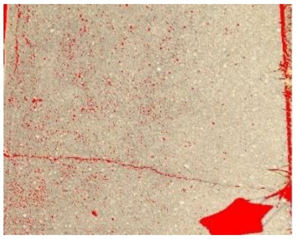
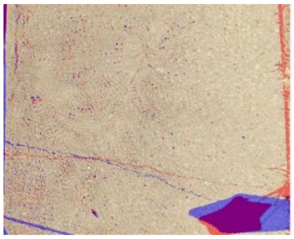
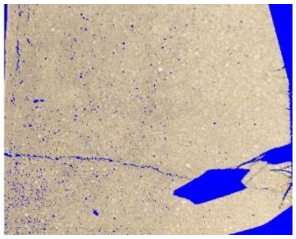
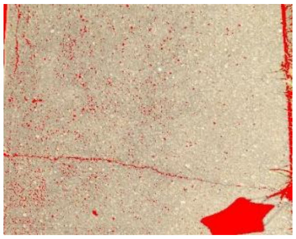
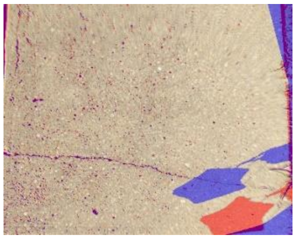
Table 3.9: Matching Results and Alignment Accuracy Under Shadow Interference

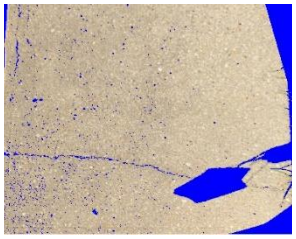
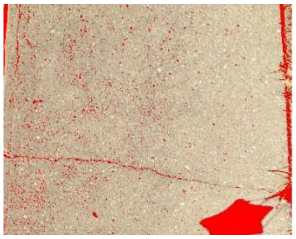
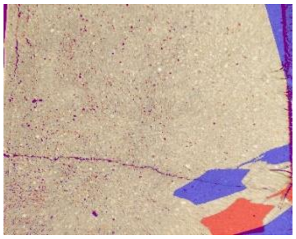
(a) Match counts before/after RANSAC

	Matches Before RANSAC	Matches After RANSAC
SIFT		
SURF		
ORB		

BRISK		
KAZE		

(b) Crack overlap assessment

Method	Corrected Image 2 (Blue)	Baseline Image 1 (Red)	Overlap Accuracy
SIFT			
SURF			
ORB			
BRISK	N/A	N/A	Failed

KAZE			

(c) Dimensional recovery vs. baseline (pixels and % error)

	Baseline	SIFT	SURF	ORB	BRISK	KAZE
Crack Area	571	550	726	603	–	580
Area Error (%)	–	4	27	6	–	2
Spine Length	392	369	358	396	–	385
Length Error (%)	–	6	9	1	–	2
Average Crack Width	2.1	2.0	2.3	2.1	–	2.1
Width Error (%)	–	5	10	0	–	0

(d) Spatial inlier breakdown and match interpretation

Detector	Total Inliers	Crack-Aligned	Background	Structural Coverage	Reliability Summary
SIFT	31	2	27	Moderate	Despite high noise, geometrically valid alignment achieved
SURF	7	0	6	Weak	Shadow-dominated matches; no crack inliers
ORB	59	11	48	Strong	Crack-centered inliers dominate; stable under noise
BRISK	3	1	2	Very Weak	Single crack inlier; insufficient for alignment
KAZE	254	38	213	Strong	High crack match density; robust calibration under shadow

Note: "Background" refers to valid RANSAC inliers not located on the crack path, such as edge features or slab texture points. Corner anchors were not individually counted but contributed to total inlier counts.

Under dynamic shadow interference, KAZE-RANSAC demonstrated the highest robustness and alignment fidelity among all tested methods. It retained 254 inliers, of which 38 were crack-centered and several others anchored along stable corner features. The remaining background inliers—primarily on slab textures and edges—did not interfere with calibration. Instead, their geometric consistency enhanced spatial coverage and reinforced the crack-aligned matches, yielding a well-constrained transformation. The alignment achieved only 2% area error and 0% width error, preserving both geometric accuracy and visual coherence under lighting fluctuations.

ORB-RANSAC also yielded strong results, with 11 crack-aligned inliers out of 59. The spatial distribution of these matches followed the crack structure closely, enabling a stable transformation with low error across all metrics.

SIFT-RANSAC, despite shadow-induced gradient noise, retained two spatially relevant crack inliers among 31 matches. These proved sufficient for a low-error homography, though high background inlier ratio reduced confidence in structural localization. The resulting transformation achieved 4% area error and preserved average crack width, though minor misalignment was observed due to high background noise ratio.

In contrast, SURF-RANSAC failed to localize any crack-specific inliers. All 7 retained matches were located on shadow edges or irrelevant background structures, producing a poorly constrained transformation with 27% area error and 9% spine length error.

BRISK-RANSAC did not return a valid homography, as only 3 inliers were detected—insufficient for reliable estimation.

These results emphasize that under challenging shadow conditions, the critical determinant of alignment quality is not the absolute number of matches, but the number and spatial distribution of crack-centered and structurally meaningful inliers. While KAZE produced more background matches due to its edge-preserving nature, it simultaneously maintained the highest count of crack-aligned and anchor-supported keypoints, which proved sufficient for accurate calibration. Its nonlinear diffusion filtering suppressed luminance artifacts while preserving subtle crack discontinuities—making it the only method that consistently maintained alignment accuracy under dynamic, real-world lighting variations. This confirms that KAZE offers robust alignment performance for vision-based SHM applications, even under nonuniform or time-varying lighting conditions commonly encountered in field inspections.

3.5 Discussion: Overcoming Limitations of Gaussian-Based Feature Detectors in SHM Alignment

This study reveals that Gaussian-based feature detection—long dominant in image alignment—faces critical limitations in real-world SHM scenarios. Classical detectors such as SIFT, SURF, ORB, and BRISK, though efficient in general vision tasks, often falter under the complex visual and geometric conditions typical of SHM inspections.

1. Gaussian-Based Limitations in SHM

These methods face two critical limitations: (1) suppression of narrow or low-contrast cracks due to Gaussian smoothing, and (2) amplification of false matches on repetitive backgrounds such as masonry textures and shadow edges.

Such limitations were particularly evident in the cropped (3.2), masonry (3.3), and shadowed (3.4) scenarios, where most classical detectors failed to recover accurate alignment—even when yielding high inlier counts.

2. KAZE’s Nonlinear Advantage

In contrast, KAZE constructs a nonlinear scale space via anisotropic diffusion, preserving high-frequency crack edges while suppressing irrelevant gradients. This enables robust detection of crack-centered features under occlusion, noise, and low contrast.

Across all experiments, KAZE consistently produced spatially meaningful inliers along cracks and anchor regions—proving more valuable than raw match count. For example, in 3.4, although KAZE generated over 250 inliers, it was the 38 crack-aligned ones that ensured geometry-consistent registration.

All experiments employed a uniform RANSAC sampling of 10 matched keypoints, empirically determined as the minimum for stable homography estimation across scenes. Higher sample counts often introduced irrelevant background features, while lower counts compromised geometric stability. This fixed strategy balanced geometric reliability and computational efficiency. Moreover, the semantic relevance of crack-aligned features is essential in SHM, where precise localization of degradation guides maintenance decisions and requires crack continuity to be preserved for valid interpretation.

3. Experimental Insights

Each test case exposed specific limitations of classical detectors and demonstrated the robustness of the proposed framework:

- 3.1 (Ideal perspective): Even with minimal interference, classical methods misaligned due to sparse anchors.
- 3.2 (Cropped cracks): Most methods failed entirely; only KAZE-RANSAC achieved valid alignment.
- 3.3 (Masonry wall): Classical methods favored repeated textures; KAZE retained crack geometry.
- 3.4 (Moving shadows): KAZE preserved alignment despite lighting changes, while others degraded or failed.

These results emphasize that the semantic relevance and spatial distribution of features are more critical than quantity—especially in SHM contexts where cracks may be degraded, occluded, or distorted.

4. Practical Considerations and Future Directions

While KAZE-RANSAC showed high robustness, several limitations remain. The framework currently assumes local planarity and was tested on images from a single device (iPhone 11, 12MP), representing a typical UAV/mobile platform. Broader validation is needed. Priorities for future work include:

- Cross-sensor and cross-platform validation across camera types and resolutions
- Adaptive keypoint sampling based on scene characteristics
- Extension to non-planar surfaces via stereo, SfM, or LiDAR
- Integration with deep features for hybrid pipelines using learned descriptors such as SuperPoint or D2-Net [33], [34]

While CNN-based models show potential in semantic segmentation, they often require extensive training and high computational cost, and may lack geometric transparency in uncontrolled settings. By contrast, KAZE-RANSAC is interpretable, unsupervised, and computationally lightweight—making it attractive for field deployment. This work provides a practical alternative for scenarios where training data is scarce and interpretability is essential for field deployment.

5. Contribution in Context

This work presents what we believe to be one of the first detailed assessments of anisotropic scale-space filtering applied to SHM crack alignment under diverse field conditions. While KAZE has been explored in general vision contexts, its robustness for infrastructure imagery—particularly under distortion, occlusion, and shadow—has not been systematically benchmarked.

Unlike conventional vision datasets, SHM imagery often suffers from platform constraints (e.g., UAV capture), sparse textures, and a lack of labeled data. In this context, the KAZE-RANSAC framework offers a transparent, training-free alternative, tailored to SHM deployment needs—enabling reliable crack alignment without calibration or retraining.

While KAZE and RANSAC are established algorithms, their tailored integration and systematic evaluation for crack alignment in SHM settings represent a novel and domain-specific contribution of this study. This work demonstrates their practical value for infrastructure monitoring under real-world conditions and bridging the gap between low-level alignment algorithms and the operational needs of metrological crack tracking in SHM deployments.

3.6 Limitations

Despite its strong robustness under varied SHM conditions, the KAZE-RANSAC framework still presents practical limitations that may affect generalizability and deployment scalability.

1. Device Specificity and Sensor Variability

All experiments were conducted using a single commercial smartphone (Apple iPhone 11, 12 MP, 1/2.55″ sensor), representative of mobile and UAV-based SHM workflows. However, imaging characteristics such as lens distortion, signal-to-noise ratio, and dynamic range may vary significantly across hardware platforms, affecting feature detection and alignment fidelity. Future validation should include professional-grade sensors and low-cost field equipment to ensure generalizability.

2. Homography Assumption and Non-Planar Surfaces

The method assumes local planarity between image pairs. In practice, SHM targets often include curved piers, warped panels, or corrugated facades. These violate planar assumptions, leading to residual misalignments. Extending the framework with stereo vision, multi-view geometry, or LiDAR-assisted depth modeling may enable correction on non-planar scenes.

3. Crack Occlusion and Low Visual Salience

While KAZE is more resilient than traditional methods, it still depends on edge gradients. Under conditions of occlusion, surface contamination, or diffuse clutter, cracks may not yield detectable features. Preprocessing techniques—such as contrast normalization, shadow removal, or thermal-RGB fusion—could enhance performance in degraded environments. For long-term SHM, robust crack presence detection—prior to feature matching—may serve as a critical preprocessing module.

4. Image Quality Sensitivity and Lighting Extremes

All image acquisitions occurred under stable daylight. In practice, UAV-captured SHM images may suffer from motion blur, glare, or low-light artifacts, which compromise keypoint localization consistency and matching reliability. This framework may benefit from integrated image-quality assessment (IQA) or exposure correction modules to support real-time deployment.

5. Short, Hairline, or Textureless Cracks

On smooth surfaces or prefabricated concrete, very short or hairline cracks may produce insufficient gradients for detection. In these cases, dense descriptors, sequential optical flow, or multi-temporal SFM pipelines may be needed to reconstruct fine-grained damage evolution. These cases remain among the most difficult for keypoint-based methods and may require fundamentally different tracking paradigms beyond local descriptor matching.

6. Fixed Keypoint Sampling Strategy

A fixed set of 10 matched keypoints was used for all RANSAC iterations. This number was empirically chosen as the minimum ensuring geometric stability without overfitting to background noise. However, this setting may not generalize across scenes of varying texture or resolution. Adaptive keypoint selection—based on saliency, confidence, or complexity—warrants future exploration, especially for dynamic or large-scale SHM deployments. While this constraint isolates the influence of keypoint quality, it may underutilize spatial diversity—especially in high-resolution or richly textured imagery.


(a) Non-planar surfaces (b) Occlusion or low visual salience (c) Poor image quality (d) Short or textureless cracks

Fig. 3.1: Representative failure scenarios observed in field-aligned SHM conditions

While these challenges exist, the proposed method remains fully unsupervised, interpretable, and training-free—traits that favor transparent and efficient deployment in field environments. To expand applicability and support broader deployment scenarios, the following directions are prioritized:

- Cross-sensor and cross-platform performance validation
- Depth-informed geometric correction using stereo, LiDAR, or SfM
- Scene-aware or saliency-driven keypoint adaptation
- Exploratory hybridization with learned features (e.g., D2-Net, SuperPoint) for comparison in future benchmarks [33], [34]
- On-device QA modules for UAV-based inspections (e.g., match dispersion, real-time feedback)

4. Conclusion

This study demonstrates that classical Gaussian-based feature detectors—such as SIFT and SURF—are insufficient for accurate crack alignment under real-world SHM conditions. Their reliance on isotropic smoothing suppresses fine crack edges and favors irrelevant background textures, particularly under shadows, occlusions, and high-frequency surfaces. While binary descriptors like ORB and BRISK offer improved efficiency, they remain sensitive to unstable or semantically irrelevant keypoints.

In contrast, the proposed KAZE-RANSAC framework integrates nonlinear anisotropic diffusion with robust homography estimation, enabling crack-preserving alignment across four representative SHM challenges. It achieved less than 5% area error and under 15% length error, even in the presence of distortion, shadow, and occlusion.

The key contributions of this work are:

1. A physics-informed, unsupervised alignment pipeline that mitigates Gaussian-related failure modes in SHM;
2. Demonstrated robustness across texture interference, geometric distortion, and dynamic lighting;
3. A lightweight, training-free method deployable on mobile or UAV platforms.

While the current implementation assumes planarity and uniform illumination, it lays the groundwork for interpretable, field-deployable crack monitoring systems in SHM. Future work will explore scene-aware keypoint selection, hybrid integration with learned descriptors [33], [34], and stereo-informed correction for non-planar surfaces. Propagation modeling of alignment uncertainty across image sequences is also a priority to strengthen metrological reliability.

This framework offers a transparent, deployable alternative to data-driven black-box models— supporting explainable, image-based infrastructure inspection across bridges, tunnels, and façades—making it well-suited for low-resource SHM settings where training data or sensor calibration is impractical. Future work may integrate this alignment framework into adaptive SHM pipelines combining vision-based sensing and real-time structural response control, enabling proactive mitigation strategies.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this article, the author utilized ChatGPT to assist with language refinement and editing. Following the use of this tool, the author reviewed and edited the content independently and takes full responsibility for the final version of the work.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Submission Declaration

This article is an original work that has not been published previously and is not under consideration for publication elsewhere.

Acknowledgements

The author would like to thank Professor Peter Chang for his valuable discussions and constructive feedback throughout this work.

Reference

[1] L. Giordano and F. Tondolo, "Crack Width Evolution of R/C Structures Subject to Corrosion and Fatigue," *Key Engineering Materials*, vol. 452–453, pp. 417–420, Nov. 2010, doi: 10.4028/www.scientific.net/KEM.452-453.417.

[2] D. Jian-ming, "Load Fatigue Fracture of Longitudinal Steel Bar in Continuously Reinforced Concrete Pavement," *Transport Standardization*, 2010, Accessed: May 06, 2025. [Online]. Available: https://www.semanticscholar.org/paper/Load-Fatigue-Fracture-of-Longitudinal-Steel-Bar-in-Jian-ming/8b2dc5b6fae4655ff3417e5b2b12f42bb56b81b3

[3] C.-Q. Li and S. Yang, "Prediction of Concrete Crack Width under Combined Reinforcement Corrosion and Applied Load," *Journal of Engineering Mechanics-asce*, vol. 137, pp. 722–731, Nov. 2010, doi: 10.1061/(ASCE)EM.1943-7889.0000289.

[4] M. Modares and N. Waksmanski, "Overview of Structural Health Monitoring for Steel Bridges," *Practice Periodical on Structural Design and Construction*, vol. 18, no. 3, pp. 187–191, Aug. 2013, doi: 10.1061/(ASCE)SC.1943-5576.0000154.

[5] H. S. Munawar, A. W. A. Hammad, A. Haddad, C. A. P. Soares, and S. T. Waller, "Image-Based Crack Detection Methods: A Review," *Infrastructures*, vol. 6, no. 8, Art. no. 8, Aug. 2021, doi: 10.3390/infrastructures6080115.

[6] A. Razak *et al.*, "Structural Crack Detection System Using Internet of Things (IoT) for Structural Health Monitoring (SHM): A Review," *Jurnal Kejuruteraan34*, vol. 6, pp. 983–98, 2022.

[7] M. C. Porcu, M. Buitrago, P. A. Calderón, M. Garau, M. F. Cocco, and J. M. Adam, "Robustness-based assessment and monitoring of steel truss railway bridges to prevent progressive collapse," *Journal of Constructional Steel Research*, vol. 226, p. 109200, Mar. 2025, doi: 10.1016/j.jcsr.2024.109200.

[8] P. Gupta and M. Dixit, "Image-based crack detection approaches: a comprehensive survey," *Multimed Tools Appl*, vol. 81, no. 28, pp. 40181–40229, Nov. 2022, doi: 10.1007/s11042-022-13152-z.

[9] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, vol. 57, no. 2, pp. 787–798, Jun. 2018, doi: 10.1016/j.aej.2017.01.020.

[10] Y.-F. Liu, X. Nie, J.-S. Fan, and X.-G. Liu, "Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 5, pp. 511–529, 2020, doi: 10.1111/mice.12501.

[11] X. Pan *et al.*, "A review of recent advances in data-driven computer vision methods for structural damage evaluation: algorithms, applications, challenges, and future opportunities," *Archives of Computational Methods in Engineering*, pp. 1–33, 2025.

[12] A. S. Rao, T. Nguyen, M. Palaniswami, and T. Ngo, "Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure," *Structural Health Monitoring*, vol. 20, no. 4, pp. 2124–2142, Jul. 2021, doi: 10.1177/1475921720965445.

[13] M. Fatiha and M. Oussama, "Automatic Feature Points Based Registration Technique of Satellite Images," *2024 8th International Conference on Image and Signal Processing and their Applications (ISPA)*, pp. 1–6, Apr. 2024, doi: 10.1109/ISPA59904.2024.10536854.

[14] S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," in *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, Mar. 2018, pp. 1–10. doi: 10.1109/ICOMET.2018.8346440.

[15] M. Ihmeida and H. Wei, "Image Registration Techniques and Applications: Comparative Study on Remote Sensing Imagery," *2021 14th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 142–148, Dec. 2021, doi: 10.1109/DeSE54285.2021.9719538.

[16] R. K. Meghana, S. Apoorva, Mohana, and Y. Chitkara, "Inspection, Identification and Repair Monitoring of Cracked Concrete structure –An application of Image processing," in *2018 3rd International Conference on Communication and Electronics Systems (ICCES)*, 2018, pp. 1151–1154. doi: 10.1109/CESYS.2018.8723898.

[17] V. Belloni, A. Sjölander, R. Ravanelli, M. Crespi, and A. Nascetti, "Crack Monitoring from Motion (CMfM): Crack detection and measurement using cameras with non-fixed positions," *Automation in Construction*, vol. 156, p. 105072, 2023, doi: https://doi.org/10.1016/j.autcon.2023.105072.

[18] D. Mistry and A. Banerjee, "Comparison of feature detection and matching approaches: SIFT and SURF," *GRD Journals-Global Research and Development Journal for Engineering*, vol. 2, no. 4, pp. 7–13, 2017.

[19] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.

[20] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” presented at the Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part I 9, Springer, 2006, pp. 404–417.

[21] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” presented at the 2011 International conference on computer vision, Ieee, 2011, pp. 2564–2571.

[22] S. Leutenegger, M. Chli, and R. Y. Siegwart, “BRISK: Binary Robust invariant scalable keypoints,” in *2011 International Conference on Computer Vision*, 2011, pp. 2548–2555. doi: 10.1109/ICCV.2011.6126542.

[23] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, “KAZE Features,” in *Computer Vision – ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2012, pp. 214–227. doi: 10.1007/978-3-642-33783-3_16.

[24] P. Mukherjee and B. Lall, “Saliency and KAZE features assisted object segmentation,” *Image and Vision Computing*, vol. 61, pp. 82–97, May 2017, doi: 10.1016/j.imavis.2017.02.008.

[25] M. Okawa, “Vector of locally aggregated descriptors with KAZE features for offline signature verification,” in *2016 IEEE 5th Global Conference on Consumer Electronics*, Kyoto, Japan: IEEE, Oct. 2016, pp. 1–5. doi: 10.1109/GCCE.2016.7800470.

[26] Á. Ordóñez, F. Argüello, and D. B. Heras, “Alignment of Hyperspectral Images Using KAZE Features,” *Remote Sensing*, vol. 10, no. 5, Art. no. 5, May 2018, doi: 10.3390/rs10050756.

[27] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981, doi: 10.1145/358669.358692.

[28] S. Choi, T. Kim, and W. Yu, “Performance Evaluation of RANSAC Family,” in *Procedings of the British Machine Vision Conference 2009*, London: British Machine Vision Association, 2009, p. 81.1-81.12. doi: 10.5244/C.23.81.

[29] O. Chum and J. Matas, “Optimal Randomized RANSAC,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 8, pp. 1472–1482, Aug. 2008, doi: 10.1109/TPAMI.2007.70787.

[30] J. M. Martínez-Otzeta, I. Rodríguez-Moreno, I. Mendialdua, and B. Sierra, “RANSAC for Robotic Applications: A Survey,” *Sensors*, vol. 23, no. 1, Art. no. 1, Jan. 2023, doi: 10.3390/s23010327.

[31] F. Zhou, Y. Cui, Y. Wang, L. Liu, and H. Gao, “Accurate and robust estimation of camera parameters using RANSAC,” *Optics and Lasers in Engineering*, vol. 51, no. 3, pp. 197–212, Mar. 2013, doi: 10.1016/j.optlaseng.2012.10.012.

[32] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

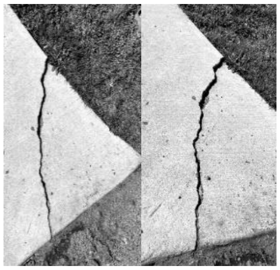
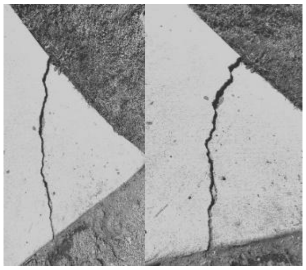
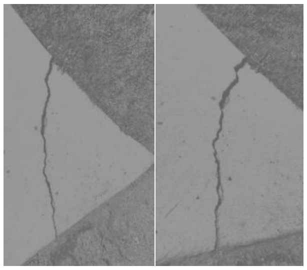
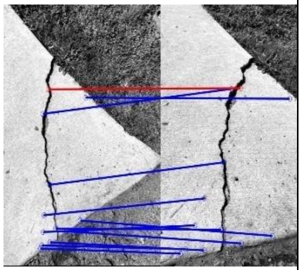
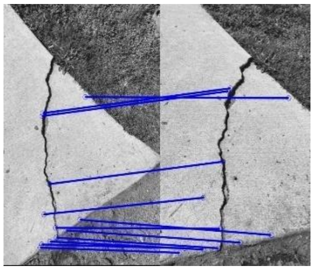
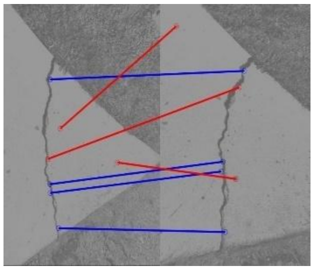
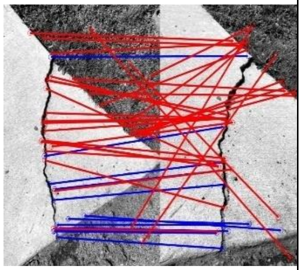
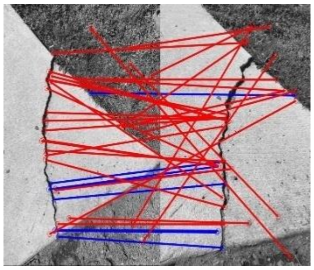
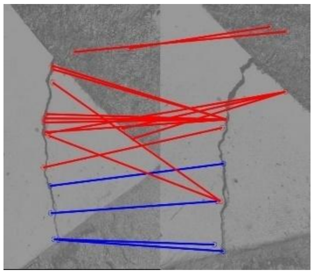
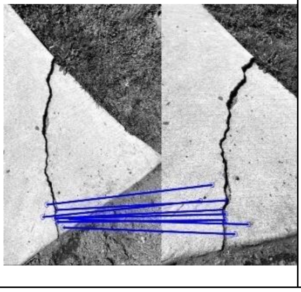
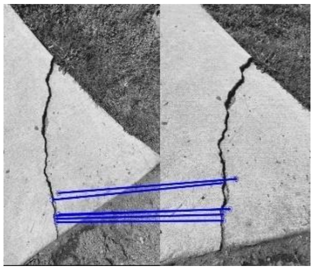
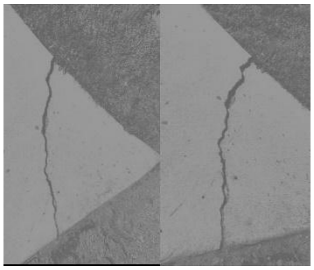
[33] D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperPoint: Self-Supervised Interest Point Detection and Description,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2018, pp. 337–33712. doi: 10.1109/CVPRW.2018.00060.

[34] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, “SuperGlue: Learning Feature Matching With Graph Neural Networks,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA: IEEE, Jun. 2020, pp. 4937–4946. doi: 10.1109/CVPR42600.2020.00499.

Appendix A. Supplementary Evaluation Results for Feature-Based Crack Alignment

This appendix presents additional evaluation results for feature detectors under contrast, noise, and blur variations beyond the four main test scenarios discussed in Section 3. Each table summarizes keypoint match quality, inlier distribution, or dimensional recovery performance under specific degraded imaging conditions.

Table A.1: Keypoint Matching Results Across Contrast Conditions (High, Medium, Low)

	High Contrast	Medium Contrast	Low Contrast
Image			
SIFT			
SURF			
ORB			

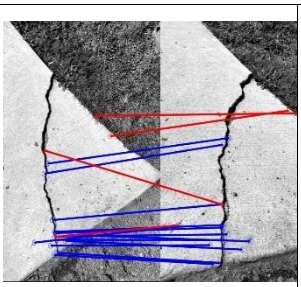
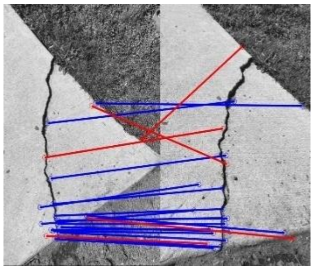
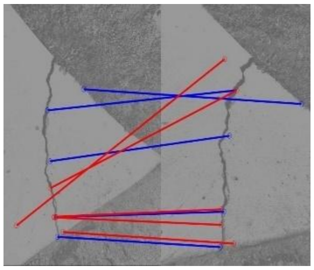
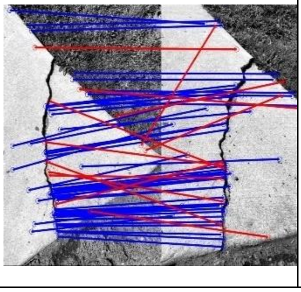
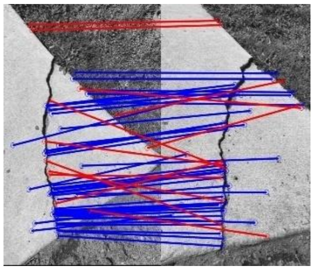
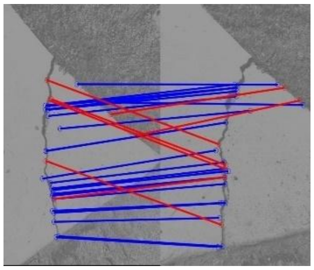
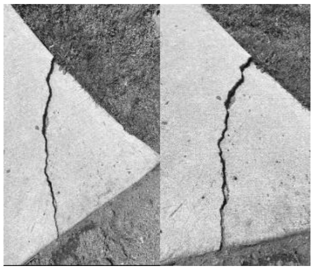
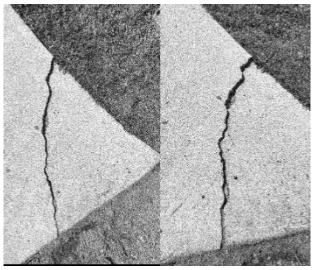
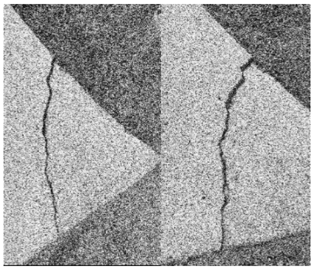
BRISK			
KAZE			

Table A.2: Mean Inlier Counts Under Contrast Variation

Mean Inliers Contrast	SIFT	SURF	ORB	BRISK	KAZE
High	12	12	8	13	48
Medium	13	8	6	17	43
Low	4	4	0	5	24

Table A.3. Crack Alignment Accuracy Under Additive Gaussian Noise (Low, Medium, High)

	Low Noise	Medium Noise	High Noise
Image			

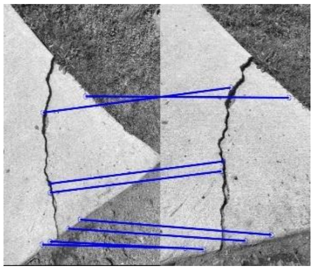
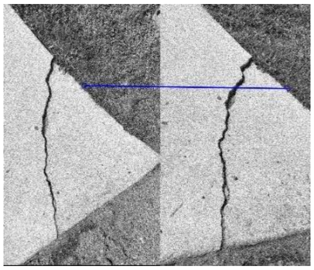
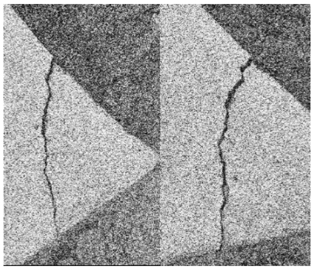
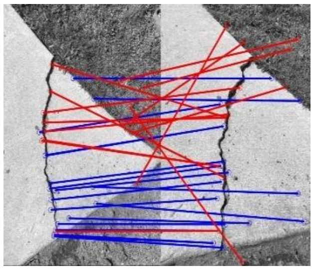
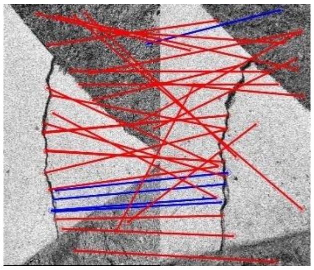
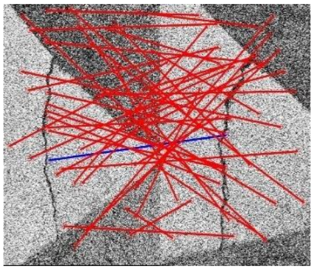
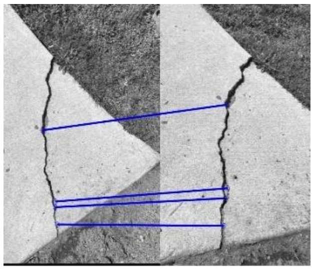
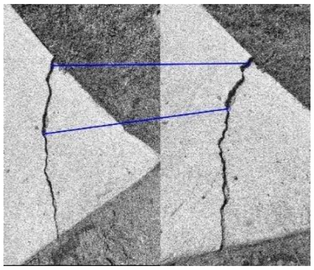
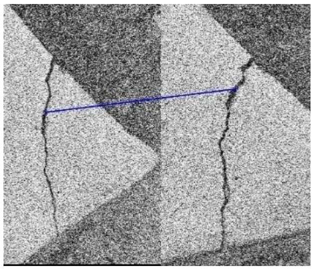
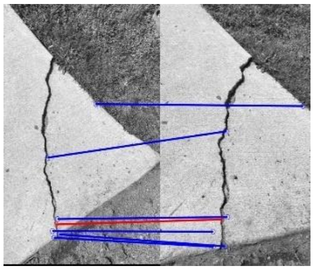
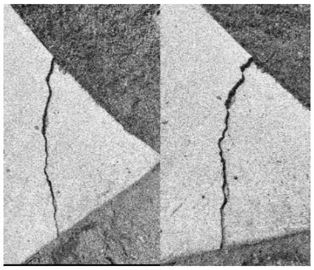
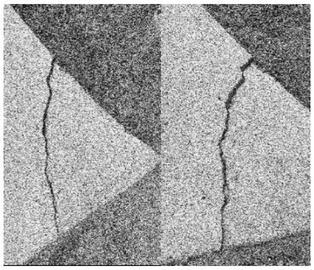
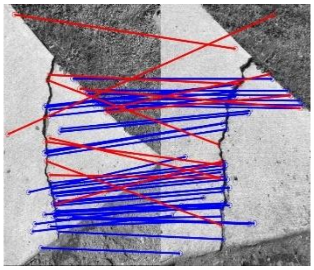
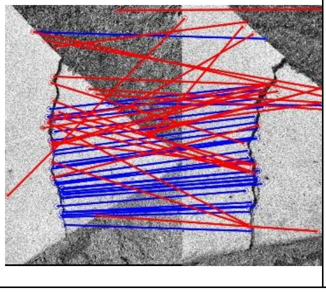
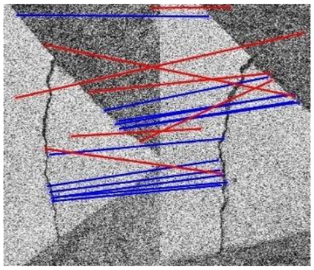
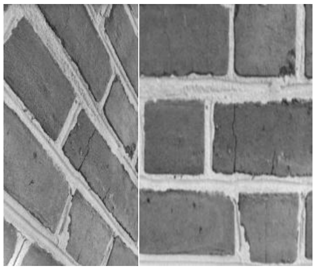
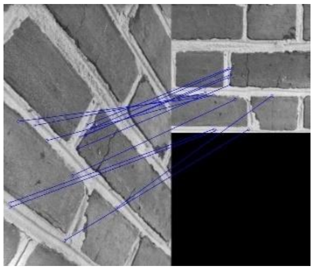
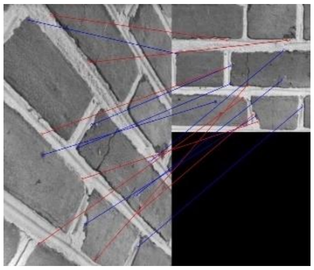
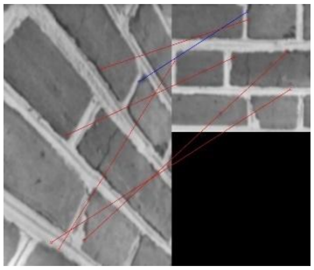
SIFT			
SURF			
ORB			
BRISK			
KAZE			

Table A.4: Keypoint Matching Performance Under Gaussian Noise (Low, Medium, High)

Mean Inliers Noise	SIFT	SURF	ORB	BRISK	KAZE

Low	8	18	4	6	41
Medium	1	5	2	0	31
High	0	1	1	0	12

Table A.5. Dimensional Recovery Under Blur Distortion (Low, Medium, High)

	Low Blur	Medium Blur	High Blur
Image			
SIFT			
SURF			
ORB			

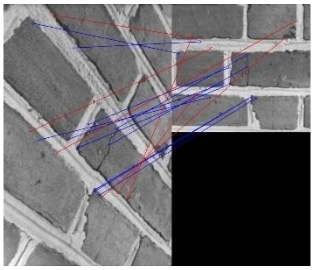
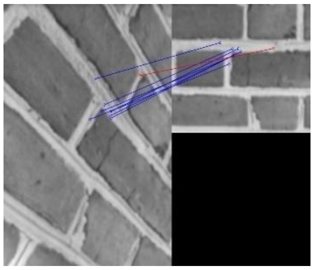
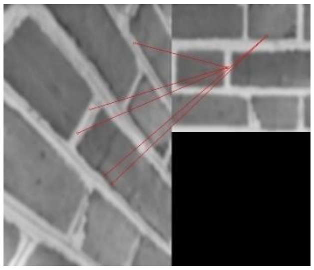
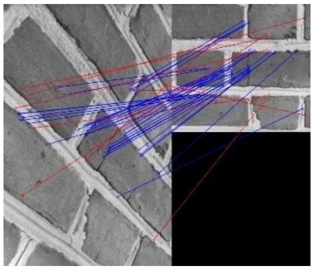
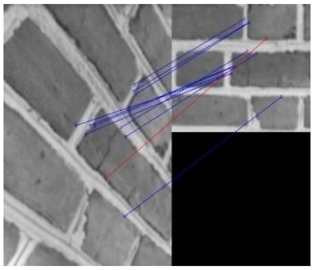
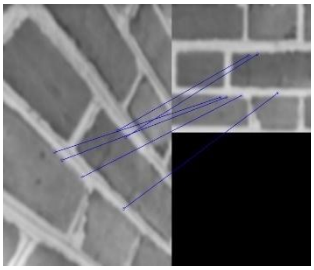
BRISK			
KAZE			

Table A.6: Spatial Inlier Distribution Under Blur Distortion (Low, Medium, High)

Mean Inliers Blurring	SIFT	SURF	ORB	BRISK	KAZE
Low	13	8	0	12	36
Medium	0	1	0	8	10
High	0	0	0	0	6

Table A.7: Keypoint Matching Results Under Perspective Distortion (Mild, Medium, Severe)

Distortion	Mild	Medium	Severe
Image			

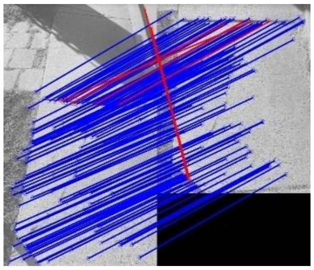
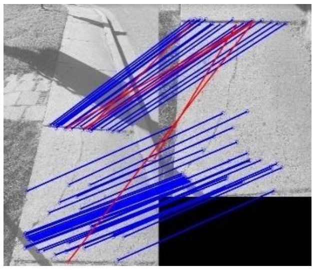
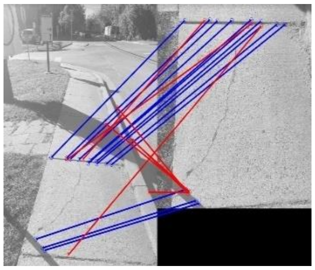
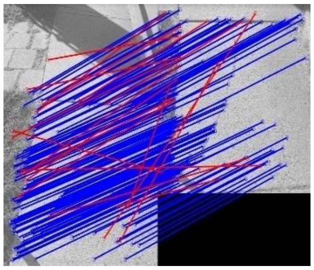
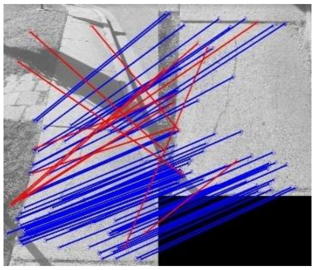
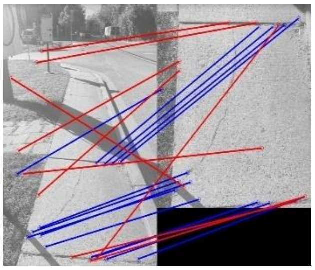
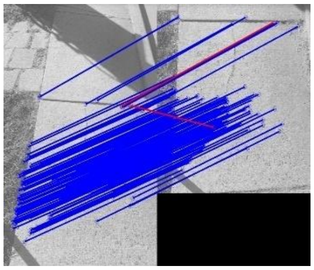
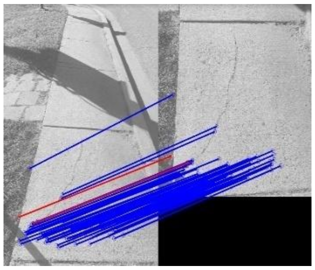
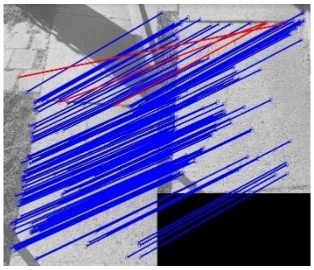
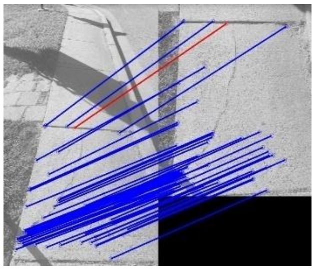
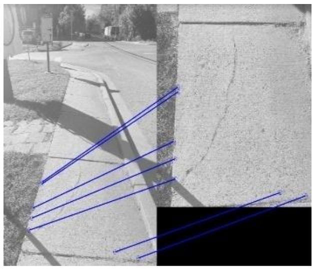
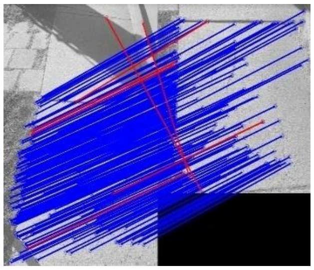
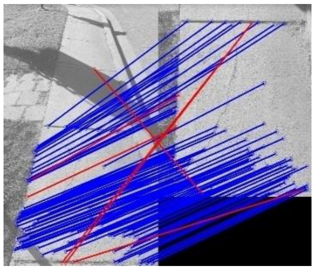
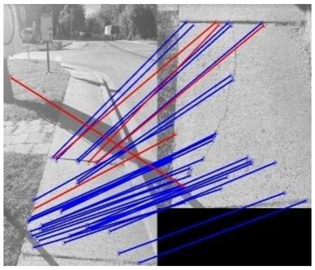
SIFT			
SURF			
ORB			
BRISK			
KAZE			

Table A.8: Mean Inlier Counts Under Perspective Distortion Levels (Mild, Medium, Severe)

Mean Inliers Distortion	SIFT	SURF	ORB	BRISK	KAZE

Severe	14	14	1	7	32
Medium	57	64	62	53	125
Mild	146	152	154	149	263