

Branch-Aware Selective Similarity Registration of Multi-Resolution Rock Tomograms

Xingyu Zhu Zhu

King Fahd University of Petroleum and Minerals

Yuzhu Wang

yuzhu.wang@kfupm.edu.sa

King Fahd University of Petroleum and Minerals

Method Article

Keywords: Digital rock, Image registration, Multi-resolution tomography, Similarity transform, Partial overlap, Selective prediction

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11187416/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Branch-Aware Selective Similarity Registration of Multi-Resolution Rock Tomograms

Xingyu Zhu¹ and Yuzhu Wang^{1*}

¹College of Petroleum Engineering and Geosciences, King Fahd University of Petroleum and Minerals, Dhahran, 31261, Saudi Arabia.

*Corresponding author(s). E-mail(s): yuzhu.wang@kfupm.edu.sa;
Contributing authors: xingyu.zhu@kfupm.edu.sa;

Abstract

Multi-resolution rock tomography requires registration across changing image appearance, partial overlap, and repeated pore-grain structures. We propose a training-free framework that couples global similarity estimation with an explicit decision to return a pose or abstain. Its methodological contribution is to connect relative structural geometry, competing pose branches, and image-based evidence within one registration procedure. Scale-space triples combine triangle shape with structure-size cues to generate rotation, translation, and isotropic scale jointly, without an initial alignment. Candidate branches preserve alternative placements for comparison using a constrained appearance model that accommodates blur and intensity response. Absolute evidence and inter-branch separation determine whether a placement is sufficiently supported, rather than treating a high individual score as conclusive. Guarded overlap regions separate appearance fitting, selection and refinement, and a final check without further pose optimization. This design avoids reliance on training pairs or identical intensities and makes unresolved ambiguity an explicit output. Tests on carbonate, sandstone, and shale tomograms support the combined workflow under the evaluated conditions, while component diagnostics delimit its demonstrated benefits.

Keywords: Digital rock, Image registration, Multi-resolution tomography, Similarity transform, Partial overlap, Selective prediction

Acknowledgements. The authors gratefully acknowledge the support of King Fahd University of Petroleum and Minerals under grant number 22011KFUPM.

047 1 Introduction

048 X-ray micro-computed tomography (micro-CT) provides non-destructive three-
049 dimensional observations of rock pores, grains, and mineral arrangement ([Cnudde and](#)
050 [Boone 2013](#)). Its field-of-view/resolution trade-off creates a data-integration problem:
051 high-resolution (HR) scans resolve fine pores and grain contacts within small volumes,
052 whereas low-resolution (LR) scans cover larger regions but suppress fine structure.
053 Digital-rock workflows combine these observations to map local information into a
054 larger domain, support property estimation, and construct paired data for super-
055 resolution ([Botha and Sheppard 2016](#); [Niu et al. 2022](#)). Such uses first require a reliable
056 spatial correspondence between the acquisitions.
057

058 Registration is therefore more than a visualization step. A wrong placement can
059 associate fine-scale structure with the wrong coarse-scale neighbourhood or create an
060 incorrect image pair for subsequent analysis. A useful computational workflow must
061 expose both its recovered transformation and cases in which the images do not support
062 a sufficiently distinct placement. This study addresses that registration and return
063 decision; it does not evaluate downstream permeability prediction or super-resolution
064 training.

065 Registration must accommodate unknown pose, scale, partial overlap, and
066 acquisition-dependent appearance. Coarse sampling also removes or merges visible
067 structures ([Fig. 1](#)); identical intensities or one-to-one pore correspondence cannot be
068 assumed. Repeated pore-grain patterns compound this difficulty by supporting several
069 plausible placements.

070 Material-registration workflows, including multistart microscopy-CT alignment
071 and digital volume correlation, generally target comparable textures, nearby states,
072 or useful initialization ([Latham et al. 2008](#); [Bay et al. 1999](#); [Hall et al. 2010](#); [Tudisco](#)
073 [et al. 2017](#); [Buljac et al. 2018](#)), rather than global HR/LR alignment with unknown
074 overlap.

075 Pyramids, mutual information, and descriptors can accommodate changes in
076 appearance ([Viola and Wells 1997](#); [Pluim et al. 2003](#); [Klein et al. 2010](#); [Heinrich et al.](#)
077 [2012](#); [Ojala et al. 2002](#)), but broad Sim(3) search remains non-convex. Learned regis-
078 tration uses supervision or image-similarity losses, with coarse-to-fine and transformer
079 architectures ([Hu et al. 2018](#); [Balakrishnan et al. 2019](#); [Mok and Chung 2020](#); [Chen](#)
080 [et al. 2022](#)). Transfer requires representative data, scarce for aligned HR/LR rocks,
081 whose repeated textures lack the anatomical organization common in medical images.

082 Geometric methods avoid identical-intensity assumptions: ICP refines nearby
083 poses, FPFH-RANSAC proposes sparse matches, and 4PCS or geometric hashing
084 searches preserved configurations ([Besl and McKay 1992](#); [Rusu et al. 2009](#); [Fischler](#)
085 [and Bolles 1981](#); [Aiger et al. 2008](#); [Mellado et al. 2014](#); [Lamdan and Wolfson 1988](#)).
086 We combine scale-space structure sizes ([Lindeberg 1998](#)) with closed-form transforms
087 from triples ([Umeyama 1991](#)), retaining competing placements rather than assuming
088 geometric agreement identifies a unique pose.

089 Volumetric SIFT and Fourier-Mellin search offer global alternatives ([Rister et al.](#)
090 [2017](#); [Reddy and Chatterji 1996](#)), although cropped support complicates matching.
091 TEASER robustly estimates transforms from putative correspondences ([Yang et al.](#)
092 [2021](#)). Learned sparse matchers include 3DMatch and D3Feat ([Zeng et al. 2017](#); [Bai](#)

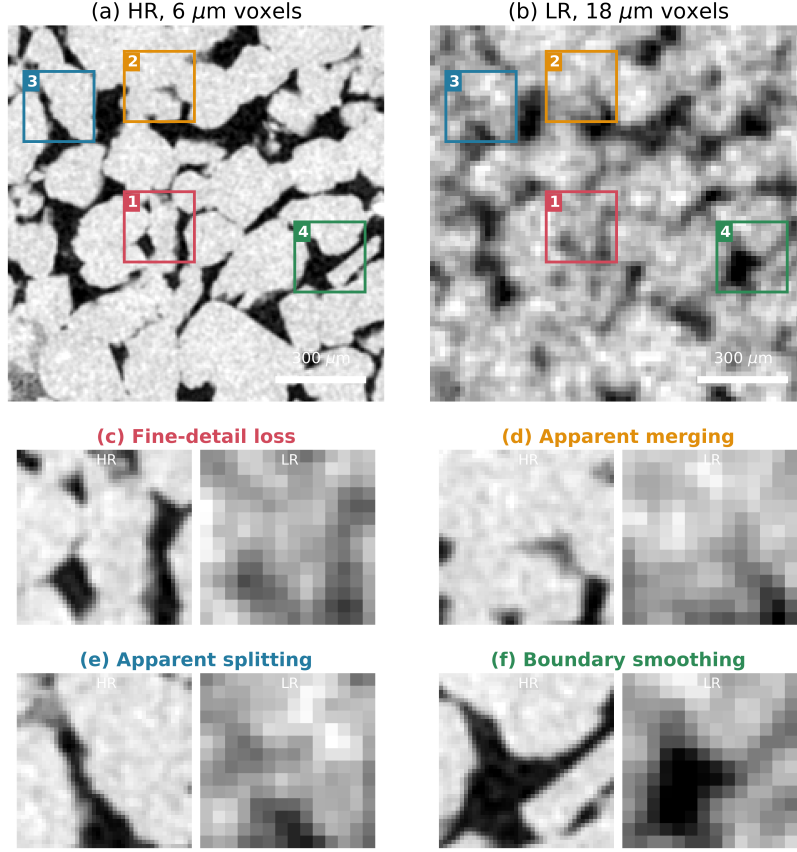


Fig. 1 Resolution-dependent appearance in corresponding central slices from an aligned HR/LR rock pair. The LR slice is displayed on the HR grid by nearest-neighbour replication for visualization only. Numbered regions show (1) fine-detail loss, (2) apparent merging, (3) apparent splitting, and (4) boundary smoothing. Merging and splitting describe appearance in the displayed section, not measured pore-topology changes. Scale bars: 300 μm

et al. 2020); PREDATOR and GeoTransformer add overlap or geometric context (Huang et al. 2021; Qin et al. 2022). Their same-scale surface benchmarks differ from unknown-scale volumetric rock registration; confidence-ranked output also differs from separately checked selective return.

Pose generation and selective return are distinct tasks. RANSAC and MLESAC reuse the correspondences that form hypotheses (Fischler and Bolles 1981; Torr and Zisserman 2000); an image score likewise ceases to be a separate check when its voxels fit, select, or refine the pose. Following selective prediction (Chow 1970; Geifman and El-Yaniv 2017), we combine branch-aware global proposal with three-region evidence: $\mathcal{V}_1$ fits appearance, $\mathcal{V}_2$ ranks branches, $\mathcal{V}_1 \cup \mathcal{V}_2$ supports refinement, and reserved $\mathcal{V}_3$ checks the final pose. The main contribution is a computational procedure that retains competing global explanations of cross-resolution rock structure until image evidence supports a selective return. It integrates relative scale-space triples, pose branches,

and spatially separated forward evidence. Triples provide complete pose hypotheses without an initial alignment; branches preserve alternatives that repeated structures make geometrically plausible; image evidence then supports selection and a separate return decision. Geometric proposal uses the full images, so spatial separation does not establish statistical independence. We evaluate the combined procedure and its failure stages; the experiments do not establish a separate accuracy gain for every component.

The remainder of this article is organized as follows. Section 2 presents geometric proposal, branch formation, appearance-based selection, local refinement, and final verification. Section 3 describes the acquisitions, evaluation cohorts, error measures, and comparison implementations. Section 4 analyzes branch decisions, negative controls, limited support, scale recovery, and the trade-off between registration reliability and computational cost. Section 5 summarizes the algorithmic contribution, experimental support, and priorities for further development.

2 Method

2.1 Problem formulation and coordinates

Let V_m and V_f denote moving and fixed three-dimensional grey-scale images. HR is moving and LR is fixed for cross-resolution pairs. Their array sizes, fields of view, voxel spacings, and origins may differ. The desired moving-to-fixed similarity transform is

$$\mathbf{y} = \mathcal{T}(\mathbf{x}) = s\mathbf{R}\mathbf{x} + \mathbf{t}, \quad s > 0, \quad \mathbf{R} \in \text{SO}(3), \quad \mathbf{t} \in \mathbb{R}^3. \quad (1)$$

Fixed-grid resampling uses the inverse map,

$$I_{\mathcal{T}}(\mathbf{y}) = W_{\mathcal{T}}V_m(\mathbf{y}) = V_m(\mathcal{T}^{-1}(\mathbf{y})). \quad (2)$$

When reliable spacings are available, low-pass filtering and resampling place both images on one physical grid before registration. Estimated s then represents residual isotropic physical mismatch, not the known sampling ratio. Without reliable spacings, processing uses native voxel coordinates and s combines unknown sampling ratio with physical size change; it cannot then be interpreted as specimen dilation or shrinkage.

2.2 Global proposal and pose branches

Seven Gaussian levels generate six difference-of-Gaussian (DoG) volumes:

$$\begin{aligned} \sigma_i &= 1.35^i \quad (i = 0, \dots, 6), \\ D_i(\mathbf{p}) &= \sigma_i^2[(G_{\sigma_i} * V)(\mathbf{p}) - (G_{\sigma_{i+1}} * V)(\mathbf{p})], \quad i = 0, \dots, 5. \end{aligned} \quad (3)$$

Bright and dark extrema at the four interior DoG levels receive a quadratic position correction limited to half a voxel; suppression uses a 1.5-voxel radius. At most $N = 128$ points, each with position $\mathbf{p}$ and selected scale σ , are retained. This fixed computational budget balances shared-structure coverage against the $\binom{N}{3}$ triple count and is not an intrinsic rock property.

An extremum is compared with the surrounding $3 \times 3 \times 3$ positions at its own DoG level and the corresponding neighbourhoods at the two adjacent levels. Maxima and minima detect bright and dark structures without one global segmentation threshold. The selected σ is a relative structure-size cue, not an exact pore radius. Blur, reconstruction, and the finite set of levels can change the selected scale even at corresponding locations.

Each unordered triple is indexed by its sorted edge lengths $e_1 \leq e_2 \leq e_3$. The vertex opposite e_i is v_i , which fixes vertex order without testing six permutations. We discard small, nearly collinear, or edge-order-unstable triples using

$$e_1 > 5, \quad e_3 < 0.95(e_1 + e_2), \quad e_2 - e_1 > 0.08e_3, \quad e_3 - e_2 > 0.08e_3. \quad (4)$$

Lengths are measured in working-grid voxels. The relative signature is

$$\phi(Q) = \left(\log \frac{e_2}{e_1}, \log \frac{e_3}{e_1}, \log \frac{\sigma_{v_1}}{e_1}, \log \frac{\sigma_{v_2}}{e_1}, \log \frac{\sigma_{v_3}}{e_1} \right). \quad (5)$$

The first two components are similarity invariants; the final three are approximate size cues affected by sampling and blur. Signatures are compared by

$$d_\phi = \|\mathbf{W}(\phi_f - \phi_m)\|_2, \quad \mathbf{W} = \text{diag}(0.08^{-1}, 0.08^{-1}, 0.15^{-1}, 0.15^{-1}, 0.15^{-1}). \quad (6)$$

The threshold $d_\phi \leq \sqrt{5}$ bounds RMS normalized discrepancy by one, not each component separately. A k -d tree retrieves at most six nearest eligible fixed signatures per moving triple. Ordered vertices define a candidate through

$$\min_{s>0, \mathbf{R} \in \text{SO}(3), \mathbf{t}} \sum_{i=1}^3 \|\mathbf{y}_i - (s\mathbf{R}\mathbf{x}_i + \mathbf{t})\|_2^2, \quad (7)$$

solved in closed form (Umeyama 1991).

Triple enumeration is performed separately in each image, with at most 341,376 triples for 128 detections before pruning. It is not an exhaustive comparison of all moving/fixed triple pairs. Signature retrieval supplies ordered point correspondences, and the closed-form fit converts each match into one complete similarity hypothesis. Under ideal rescaling, both structure size and edge length change by the same factor; cross-resolution sampling need not obey this idealization, motivating the looser tolerances on the three size cues.

Candidates are ranked with all detected structures. A transformed moving point contributes only when it and its nearest fixed point are mutual nearest neighbours. With residual d_i , reciprocity indicator μ_i , and one working-grid voxel h , support is

$$S_{\text{geom}}(\mathcal{T}) = \sum_i \mu_i(\mathcal{T}) \exp[-d_i(\mathcal{T})^2/(2h^2)]. \quad (8)$$

Table 1 Roles of the guarded evidence regions. Separation is spatial and procedural, not a claim of statistical independence.

Region	Permitted use	Output
$\mathcal{V}_1$	Fit the intensity map	Appearance model; fitting residual does not rank poses
$\mathcal{V}_2$	Rank and refine	Selection score ρ_2 and top-two margin
$\mathcal{V}_3$	Evaluate once after refinement	Final score ρ_3 ; no pose change

The first 10,000 candidates are scanned in decreasing support. Candidates $\mathcal{T}_i = (s_i, \mathbf{R}_i, \mathbf{t}_i)$ and $\mathcal{T}_j$ share a branch only if

$$d_R(\mathbf{R}_i, \mathbf{R}_j) < 5^\circ, \quad \|\mathcal{T}_i(\mathbf{x}_m^{\text{ctr}}) - \mathcal{T}_j(\mathbf{x}_m^{\text{ctr}})\|_2 < 2h, \quad |\log(s_i/s_j)| < 0.05. \quad (9)$$

Comparing the transformed moving centre avoids origin-dependent translation vectors. At most 32 branches and four members per branch preserve distinct placements and nearby starting poses.

Reciprocity prevents several moving detections from simultaneously claiming the same fixed structure, but does not make a correspondence geometrically correct. Likewise, branch formation is a search-budget allocation rule, not a posterior distribution over poses. Multiple branches retain distinct placements; multiple members within a branch provide alternative local starting points. The transformed centre in Eq. (9) is used only to compare displacement and does not constrain the rotation axis.

2.3 Spatially separated evidence

Each candidate’s valid fixed-grid overlap $\Lambda(\mathcal{T})$ is split along its longest occupied axis into three regions (Fig. 2) with the roles in Table 1.

Adjacent regions have eight-voxel guards, matching the radius $\lceil 3\tau \rceil$ of the widest Gaussian channel ($\tau = 2.5$). Feasible cuts first maximize the smallest region and then retained volume; boundaries remain fixed in the fixed-image frame. Scoring requires at least 800 overlap voxels, 40% coverage of the smaller common-grid domain, and 400 valid voxels in $\mathcal{V}_1$ and the evaluated region. Otherwise the method abstains.

2.3.1 Forward evidence and selection

The transformed moving image is represented by one unblurred and three Gaussian-smoothed channels,

$$X_k(\mathbf{y}; \mathcal{T}) = (G_{\tau_k} * I_{\mathcal{T}})(\mathbf{y}), \quad \tau_k \in \{0, 0.75, 1.5, 2.5\}. \quad (10)$$

A nonnegative linear combination predicts fixed intensity:

$$\hat{V}_f(\mathbf{y}) = b + \sum_{k=0}^3 a_k X_k(\mathbf{y}; \mathcal{T}), \quad a_k \geq 0. \quad (11)$$

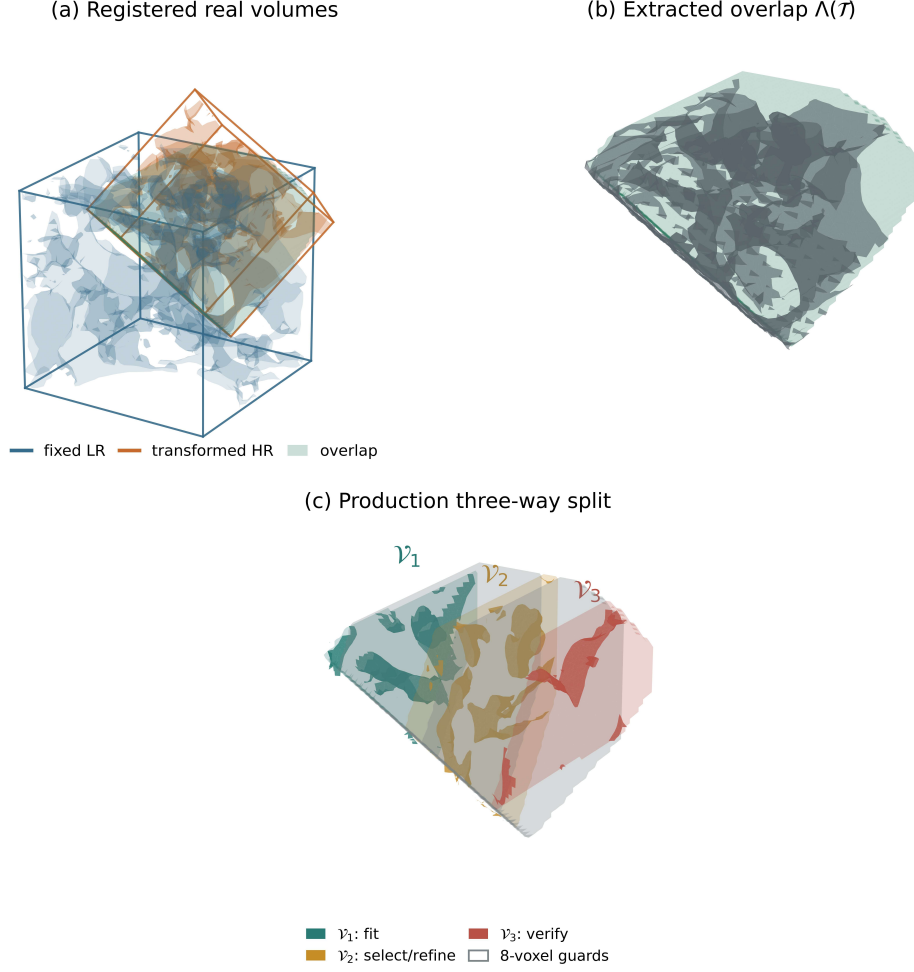


Fig. 2 Candidate-specific overlap and three-region split for an MRCCM-MEC case with target overlap 0.70. The grey slabs are eight-voxel guards. $\mathcal{V}_1$ fits the appearance model, $\mathcal{V}_2$ supports selection and refinement, and $\mathcal{V}_3$ is reserved for verification

Coefficients are fitted only on $\mathcal{V}_1$ by

$$(\hat{b}_1, \hat{\mathbf{a}}_1) = \arg \min_{b, \mathbf{a} \geq 0} \sum_{\mathbf{y} \in \mathcal{V}_1} \left[V_f(\mathbf{y}) - b - \sum_{k=0}^3 a_k X_k(\mathbf{y}; \mathcal{T}) \right]^2 + \lambda \|\mathbf{a}\|_2^2. \quad (12)$$

Here λ is 10^{-4} times the mean non-intercept diagonal of the active feature Gram matrix; the intercept is unpenalized. Predictive evidence on $\mathcal{V}_j$ is

$$\rho_j(\mathcal{T}) = 1 - \frac{\text{Var}_{\mathbf{y} \in \mathcal{V}_j} [V_f(\mathbf{y}) - \hat{V}_f(\mathbf{y}; \hat{\mathbf{a}}_1, \hat{b}_1)]}{\text{Var}_{\mathbf{y} \in \mathcal{V}_j} [V_f(\mathbf{y})]}. \quad (13)$$

This score measures residual-variance reduction, ignores mean residual bias, and may be negative; it is not a probability. The 0.30 threshold requires a 30% reduction relative to fixed-image variance. Its suitability was not calibrated for low-SNR or artifact-heavy acquisitions, and the score has no corresponding upper bound under an incorrect pose. The largest- ρ_2 member represents each branch. The leading representative is rejected below 0.30 and is not selected when no finite score exists or, with at least two finite representatives, the top-two margin is below 0.05. These are developmental operating thresholds rather than correctness bounds.

When exactly one finite branch score remains above threshold, selection proceeds without a margin test. Rejection denotes a finite score below its threshold; abstention denotes missing support or unresolved competition. Both outcomes return no transformation. The appearance weights are refitted for each candidate on $\mathcal{V}_1$, not shared across candidates; its fitting residual never determines their order.

2.3.2 Refinement and reserved verification

For each image and radius $r \in \{1, 2, 4\}$, local intensities are standardized:

$$z_r(\mathbf{y}) = \frac{V(\mathbf{y}) - \mu_r(\mathbf{y})}{\sqrt{\max\{\text{Var}_r[V](\mathbf{y}), 10^{-6}\}}}. \quad (14)$$

The vector (z_1, z_2, z_4) is clipped to $[-5, 5]$ and ℓ_2 -normalized voxelwise. Valid fixed and transformed-moving dot products are averaged over 5^3 neighbourhoods to obtain $\bar{c}$, excluding neighbourhoods with less than half valid support. If $\Gamma(\mathcal{T})$ is the valid part of $\mathcal{V}_1 \cup \mathcal{V}_2$ and q its valid fraction, refinement maximizes

$$J(\mathcal{T}) = q(\mathcal{T})^{1/2} \frac{1}{|\Gamma(\mathcal{T})|} \sum_{\mathbf{y} \in \Gamma(\mathcal{T})} \bar{c}(\mathbf{y}; \mathcal{T}). \quad (15)$$

Adam updates rotation vector, translation, and log scale for 200 steps with learning rates 0.01, 0.1, and 0.002. The best pose is retained. Spatial partition boundaries are fixed, and $\mathcal{V}_3$ never affects parameters or stopping.

Gradients are computed by automatic differentiation through trilinear sampling of the fixed feature field at transformed moving-grid locations. Translation updates use working-grid voxels, and the log-scale parameter maintains positive scale. The validity masks are held fixed within each objective evaluation. The valid-fraction factor discourages improving the score merely by reducing the region over which agreement is measured.

At the refined pose, channels are regenerated and coefficients are refitted on the valid part of the original $\mathcal{V}_1$. The model is evaluated once on the original $\mathcal{V}_3$; the pose is accepted when

$$\rho_3(\hat{\mathcal{T}}) \geq 0.30. \quad (16)$$

Insufficient data causes abstention and a lower finite score causes rejection. Reserved verification excludes appearance fitting, forward-score selection, and refinement from its evaluation region. Detection and geometric proposal still use the full images, so neither statistical independence nor geometric correctness is guaranteed.

Table 2 Frozen primary settings and algorithmic roles.

Setting	Primary value	Role
DoG levels / point budget	7 / $N = 128$	Detection / proposal capacity
Candidate scan / branches / members	10,000 / 32 / 4	Branch-retention budgets
Forward blur channels	0, 0.75, 1.5, 2.5 voxels	Appearance family
Guard / minimum region	8 / 400 voxels	Prevent leakage / require evidence
Minimum overlap fraction	0.40	Candidate support
ρ_2 / top-two margin	0.30 / 0.05	Selection / ambiguity control
Refinement steps	200	Local optimization
ρ_3	0.30	Reserved verification

2.4 Complete procedure

Algorithm 1 summarizes the workflow; Table 2 lists settings held fixed across all primary cases.

Algorithm 1 Branch-aware selective Sim(3) registration

Require: moving volume V_m , fixed volume V_f , optional voxel spacings

Ensure: accepted pose $\hat{T}$, rejection, or abstention

- 1: normalize reliable spacings; detect and match scale-space triples
 - 2: estimate candidates, rank geometric support, and form pose branches
 - 3: **for all** retained candidates T **do**
 - 4: construct guarded $(\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3)$
 - 5: if valid, fit on $\mathcal{V}_1$ and score ρ_2 on $\mathcal{V}_2$
 - 6: **end for**
 - 7: retain the largest- ρ_2 member per branch and rank representatives
 - 8: **if** no finite ρ_2 , $\rho_2^{(1)} < 0.30$, or finite top-two margin < 0.05 **then**
 - 9: **return** abstention or rejection as applicable
 - 10: **end if**
 - 11: freeze the partition and refine on $\mathcal{V}_1 \cup \mathcal{V}_2$
 - 12: refit on $\mathcal{V}_1$ and evaluate ρ_3 once on $\mathcal{V}_3$
 - 13: **if** $\rho_3 \geq 0.30$ **then**
 - 14: **return** accepted pose $\hat{T}$
 - 15: **else**
 - 16: **return** abstention or rejection as applicable
 - 17: **end if**
-

3 Evaluation design

3.1 Data and cohorts

Tomograms span carbonate (MRCCM-MEC/ILS), sandstone (Berea and Bentheimer), and shale (Green River GRS01) (Alqahtani et al. 2021; Goldfarb et al. 2021; Renard and McBeck 2018; Jackson et al. 2021). MRCCM and Bentheimer have calibrated spacings; Berea’s three acquisition conditions use uncalibrated native grids and are excluded from physical-error pooling. GRS01 contains sequential states.

The 383 primary trials use 11 acquisition pairs (Table 3), each with eight non-overlapping fixed source blocks. Moving and fixed crops come from distinct

Table 3 Primary cohort. Spacing is moving/fixed in μm per voxel; “native” denotes equal uncalibrated grid units. Size is moving/fixed in voxels. The final column is returned/correct.

Pair	Material or condition	Spacing	Size	Cases	Returned/correct
MRCCM-MEC	carbonate	2.68/10.72	$192^3/64^3$	41	38/38
MRCCM-ILS	Indiana limestone	2.68/10.72	$192^3/64^3$	41	40/40
Green River GRS01	shale, states 01/02-LR4	5.98/23.92	$192^3/64^3$	41	41/41
Core 1, subvolume 1	Bentheimer sandstone	2/6	$288^3/112^3$	25	25/25
Core 1, subvolume 2	Bentheimer sandstone	2/6	$288^3/112^3$	25	24/24
Core 2, subvolume 1	Bentheimer sandstone	2/6	$288^3/112^3$	25	25/25
Core 2, subvolume 2	Bentheimer sandstone	2/6	$288^3/112^3$	25	25/25
Berea left-500 projections	Berea sandstone	native/native	$96^3/128^3$	40	40/40
Berea right-500 projections	Berea sandstone	native/native	$96^3/128^3$	40	38/38
Berea left-0.75-mm Cu	Berea sandstone	native/native	$96^3/128^3$	40	40/40
Berea right-0.75-mm Cu	Berea sandstone	native/native	$96^3/128^3$	40	39/39
Total				383	375/375

acquisitions, not one resampled image. Seeded rotations, crops, and target overlaps 0.50/0.70/0.90 were hidden from registration; unreachable overlaps reduced case counts. Unequal-spacing trials comprise 182 native pairs and 41 GRS01 state-01/state-02 trials with state 02 coarsened fourfold; 160 Berea trials use equal-grid acquisitions. Cases share acquisitions and rotations and are not independent specimens.

Injected transforms are exact only conditional on inherited acquisition alignment, whose uncertainty is unavailable for MRCCM and Bentheimer. GRS01 may retain local deformation outside Sim(3). Errors therefore do not measure absolute material-point accuracy. Sixteen reused primary rotations span $66.4\text{--}178.0^\circ$, with none below 60° .

Primary and comparison transforms have $s_\star = 1$ and sampling ratios $1/3/4$; the proposed method still estimates scale. Known spacing ratios, detector cues σ_{v_i}/e_1 , and residual physical scale are distinct. A controlled diagnostic renders one $6\mu\text{m}$ CT at five scales $s_\star \in \{0.85, 0.925, 1.00, 1.075, 1.15\}$ for three rotations. Within each rotation, crops and full overlap are fixed; only scale changes.

Post-hoc studies test stage responsibility (18 cases), cross-specimen rejection (100 cases per cohort), and domain size (24 cases). Protocols accompany the results. A controlled branch-retention study uses a fixed, outcome-independent 22-case subset of the primary cohort (Section 4.2). Separate low-overlap tuning/evaluation sets each have 44 cases from separate blocks of the same 11 pairs. These studies are diagnostic, not confirmatory; any negative-control return is false.

3.2 Outcomes and implementations

Except for threshold tuning, truth is used only after registration. Error e_c is RMS moving-corner displacement between estimated and true poses in fixed-grid voxels, combining translation, rotation, and scale. Correctness requires $e_c \leq 2$, equivalent to $12\text{--}47.84\mu\text{m}$ or six/eight moving voxels when calibrated, and two native units for Berea. Secondary errors are rotation e_R and relative scale $e_s = s/s_\star - 1$, reported as $|e_s|$. Coverage is returns/cases, selective accuracy is correct/returned, and unconditional success is correct/cases. Correlation precludes population-level interpretation of case-level binomial intervals.

All six implementations share 50 unit-scale cases at target overlap 0.70 from seven supported pairs; comparison success requires both $e_c \leq 2$ and $e_R \leq 2^\circ$, counting non-returns as failures. Comparators are Mattes mutual information (Mattes et al. 2001), FPFH-RANSAC-ICP (Rusu et al. 2009; Fischler and Bolles 1981; Besl and McKay 1992), OpenGR Super4PCS-ICP (Mellado et al. 2014), GeoTransformer (Qin et al. 2022), and an internal dense search-strength control. Table 10 specifies implementations. All received declared spacings without case-specific retuning. Runtime covers loading through return, excluding interpreter startup, with CUDA synchronization for the proposed method. On one Windows 11/RTX 5090 workstation, it and GeoTransformer used the GPU; Mattes, Open3D, and OpenGR used CPU execution. These are implementation times, not hardware-independent complexity. Resource profiling used one MRCCM-MEC warm-up and three isolated-process runs.

4 Results and discussion

4.1 Registration accuracy and branch selection

The primary result is selective reliability, not universal recovery: the method returned 375 of 383 poses, all correct, but eight cases remained unresolved. Coverage and unconditional success were 97.91%, versus 100% selective accuracy; pair-level coverage ranged from 92.68% to 100% (Table 3). Returned-pose e_c had median and 95th-percentile values of 0.760 and 1.428 fixed-grid voxels; the 218 calibrated returns gave 8.2 and 27.0 μm , respectively. These errors measure alignment recovery, not whether every HR-resolved pore has an LR counterpart.

The decision sequence explains the non-returns (Fig. 3). Within-tolerance representatives existed in 379 cases and led in 375. The margin rule withheld two correct leaders and all four ranking reversals. Of four cases without a correct representative, refinement recovered two and low ρ_2 rejected two. Thus early ambiguity control protected against incorrect leaders but also withheld recoverable poses. No selected pose failed $\mathcal{V}_3$; this strongly filtered subset cannot establish its incremental benefit.

These outcomes explain why geometry and image evidence have different roles. Relative triples supply global starting poses without requiring equal grey levels, while branch retention keeps distinct placements available for later comparison. The appearance model tests those placements using surrounding volume content rather than only the three matched points. Its practical advantage is to accommodate blur and intensity response without demanding one-to-one correspondence between every resolved pore. The high frequency of retained correct representatives shows that this proposal-to-selection path was viable in the primary cohort; it does not isolate the benefit of each component. The controlled branch-retention comparison below tests one such component on a fixed subset.

Absolute evidence cannot replace branch comparison. The highest outside-tolerance scores had median/maximum 0.613/0.962; 324, 77, and 14 of 383 exceeded 0.30, 0.80, and 0.90. Most were nearby alternatives: 310/324 above 0.30 were within 10° , five centre-displacement voxels, and $|\log(s/s_*)| \leq 0.05$; only one exceeded the rotation or centre bound. A continuous appearance score therefore overlaps the hard

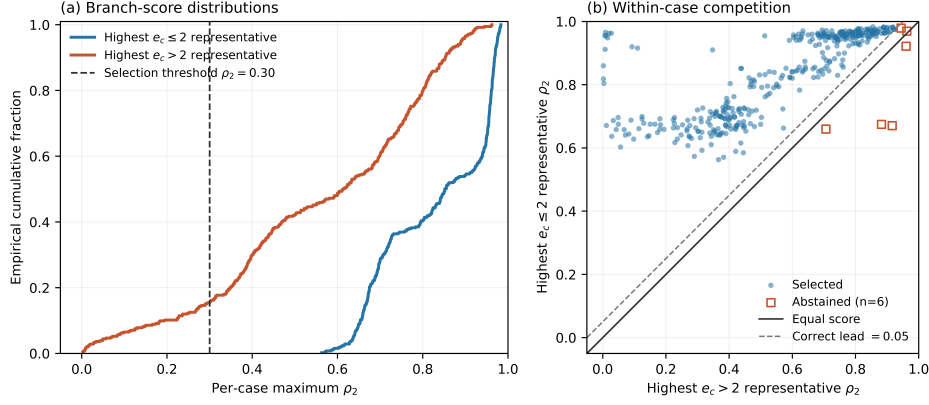


Fig. 3 Pre-refinement evidence among 32 branches per case. Each branch is represented by its highest- ρ_2 member and labelled by whether $e_c \leq 2$. (a) Per-case maxima; four cases lacked a within-tolerance representative. (b) The 379 paired cases; the within-tolerance score led in 375. Squares mark six abstentions. Lines show equal scores and a 0.05 lead

geometric boundary; high evidence need not imply either correctness or a distant structural alias.

The four margin-intercepted reversals had $e_c = 2.13$ or $4.25\text{--}9.76$, rotation error at most 7.54° , centre displacement at most 4.06 voxels, and $|\log(s/s_*)| < 0.033$. The margin detects competition between plausible poses, not their eventual refinement outcome: these candidates were withheld, so their attraction basins remain unknown.

Figure 4 shows the registration result directly: three-dimensional isosurfaces of the fixed LR and moving HR volumes are overlaid before and after registration, using the same view and physical coordinate limits. The returned transform has $e_c = 0.832$ fixed-grid voxels. The rendering shows the recovered spatial placement and partial overlap; resolution-dependent surface differences are not geometric-error measurements.

4.2 Controlled branch-retention comparison

To isolate retention of competing branches, we froze a 22-case subset before running either comparison arm: one case at each target overlap 0.50 and 0.70 from each of the 11 acquisition pairs. Within each pair/overlap stratum, the case with the smallest SHA-256 hash of `esin-branch-retention-20260928:pair:case_id` was selected, using the pair identifiers recorded with the reproducibility files. Outcomes were not used for selection. These are reused acquisitions, not an independent validation cohort.

Both arms use exactly the same geometric proposal pool and full 32-branch clustering, with up to four members per branch. The single-branch arm retains only the first branch in geometric-support order; the other retains all 32. Members are not reassigned when branches are discarded. Appearance scoring, 200-step refinement, support requirements, and the 0.30 selection and final evidence thresholds are identical. Inter-branch margin screening is disabled in both arms because it is undefined for a single branch. This isolates retention under a common decision rule, not the complete primary selective policy. Recomputed scores of shared first-branch members agreed within 10^{-5} in every pair of runs.

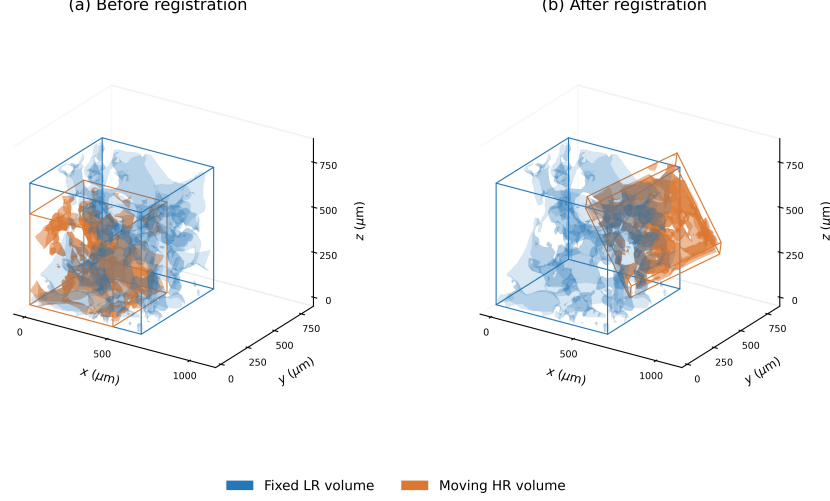


Fig. 4 Three-dimensional registration of MRCCM-MEC case **case-02c7ab44a4aa**: (a) identity placement and (b) the returned similarity transform. Blue denotes the fixed LR volume and orange the moving HR volume; wireframes mark their array extents. Both volumes are represented on the 10.72 μm common grid. Isosurfaces are extracted once at each volume’s Otsu threshold by marching cubes with a two-voxel sampling step; the same moving mesh is transformed for panel (b). Thresholds, opacities (fixed 0.18, moving 0.50), orthographic view, and axis limits are unchanged between panels. Full extents are shown, including non-overlapping regions. These surfaces visualize intensity structure, not validated pore segmentation or pointwise correspondence. Identity placement is a display baseline, not an initialization supplied to the algorithm. This is the first accepted case in ascending ID order; replay matched the stored decision and pose metrics. Coordinates are in μm

Table 4 Paired branch-retention ablation on 22 fixed cases. Correct-candidate retention means at least one within-tolerance member was retained before refinement. Margin screening is disabled in both arms. Times are descriptive single-run measurements with the shared proposal/clustering cost added to the cached single-branch execution, not independent end-to-end benchmarks.

Retained branches	Correct candidate retained	Returned	Correct	Incorrect	Median accounted time (s)
First geometric branch	22/22	22/22	22	0	83.51
All 32 branches	22/22	22/22	22	0	95.23

Both arms returned all 22 cases correctly (Table 4). The first branch already contained a correct member in every case, and both arms selected the same starting candidate for refinement. Restoring the primary 0.05 margin by decision replay changed none of these returns. Thus this subset shows no incremental recall or correct-return benefit from multiple branches. Its measured overhead reflects evaluating alternatives that did not change the selected pose. Timings use one run per arm, always multiple branches first, and include diagnostic work; cache and execution-order effects preclude a general speed comparison. The result does not establish equivalence on ambiguous or low-recall cases, nor does it test the benefit of the ambiguity margin. Broader outcome-independent cohorts are needed to estimate when additional branches justify their cost.

Table 5 Decision-rule replay on 18 matched refined poses, not complete method reruns. “Pre” applies score 0.30 and margin 0.05; post-refinement thresholds are 0.30. The three-way region was not reserved during two-way refinement.

Return rule	Cases	Returned	Returned correct/incorrect	Correct not returned
No score threshold	18	18	16/2	0
Two-way pre-refinement $\mathcal{V}_2$	18	13	13/0	3
Two-way post-refinement $\mathcal{V}_2$	18	16	16/0	0
Both two-way $\mathcal{V}_2$ rules	18	13	13/0	3
Post-refinement $\mathcal{V}_3$ only	18	16	16/0	0

4.3 Decision timing and regional verification

Verification does not itself improve a pose: it changes only whether that pose is returned. Its value must therefore be assessed through decisions on correct and incorrect poses, rather than a rise in median score. The 18-case diagnostic examines decision timing at matched refined poses (Table 5). Early screening withheld three poses that were correct after permissive refinement; either post-refinement check retained all 16 correct poses and rejected two incorrect ones. This supports testing delayed decisions for borderline candidates, but not removing early screening globally: the cohort deliberately overrepresents primary non-returns.

It includes all eight primary non-returns and two correct returns per affected pair. Controls are the first two cases after ascending SHA-256 sorting of `20260916:case_id`. The two-way protocol splits valid overlap at the median coordinate on its longest axis, retaining coordinates $\leq c - 8$ and $> c + 8$. Reversing that axis and recomputing the split exchanges the fitting direction; the two finite forward scores are averaged. Permissive refinement runs for 200 steps without a reserved third region. Three-way rescoring of the resulting poses is therefore diagnostic, not independent verification.

This replay is not a clean ablation of decision timing: two-way partitioning also moved three previously subthreshold margins above 0.05. Changing the partition changes both the available evidence and the decision, so thresholds cannot be transferred between partitions without calibration. At the matched refined poses, all-overlap fitting, $\mathcal{V}_1 \rightarrow \mathcal{V}_2$, $\mathcal{V}_1 \rightarrow \mathcal{V}_3$, and exchanged two-fold fitting made identical decisions at 0.30. Thus this cohort provides no observed accuracy gain from regional separation and does not independently test the primary reserved-region protocol. That protocol limits reuse of optimization evidence, but costs usable support and, below, one correct low-overlap return.

4.4 Cross-specimen rejection performance

Cross-specimen negatives test whether shared rock appearance alone can trigger a return. Both cohorts failed $\mathcal{V}_2$ (Table 6); the matched Bentheimer geometry rules out different array sizes or voxel spacings as the explanation for rejection. However, cross-core non-correspondence need not be as difficult as a nearby wrong placement within one specimen, where overlapping structures can sustain high scores.

Table 6 Cross-specimen controls: (a) primary decisions and (b) score distributions, including forced refinement. Geometry-matched positive counterparts returned 99/100 correct poses.

(a) Primary outcomes							
Cohort		Cases	Primary outcomes			False accepts	
Unrelated lithology		100	100 $\mathcal{V}_2$ rejections			0	
Same lithology, Core 1/Core 2		100	100 $\mathcal{V}_2$ rejections			0	

(b) Score distributions							
Cohort	Score	Finite	Minimum	Median	95th percentile	Maximum	≥ 0.30
Unrelated lithology	Leading pre-refinement ρ_2	100	–	0.0174	0.0484	0.0752	0
Same lithology	Leading pre-refinement ρ_2	100	0.0010	0.0134	0.0374	0.0453	0
Same lithology	Reused post-refinement ρ_2	100	−0.0113	0.0133	0.0456	0.0627	0
Same lithology	Reserved post-refinement ρ_3	100	−0.1416	−0.0066	0.0110	0.0623	0

Each cohort contains four pairings with 25 crops each. Unrelated pairings cross two calcareous-shale moving volumes with two MRCCM fixed scans. Same-lithology pairings cross two Bentheimer Core 1 with two Core 2 subvolumes, retaining primary $288^3/112^3$ dimensions and $2/6\text{-}\mu\text{m}$ spacings. Both use a $6\text{-}\mu\text{m}$ evidence grid. The forced same-lithology test bypasses score and margin gates, refines each leading branch for 200 steps, and applies threshold 0.30 to both reused and reserved scores.

Bypassing selection exposed every leading same-lithology branch to refinement, yet both reused and reserved scores remained far below 0.30. Rejection is therefore not merely a consequence of stopping early, but neither check uniquely prevented a false acceptance. These controls establish rejection for the tested pairings, not a population error rate: each cohort has 25 correlated crops per pairing, so a binomial bound treating all 100 cases as independent would overstate the information available.

4.5 Rotation, overlap, and proposal recall

Table 7 reports performance across primary rotation and overlap strata. Returns at target overlaps 0.50, 0.70, and 0.90 were 171/176, 143/144, and 61/63, respectively. Coverage does not increase monotonically with overlap, but these strata differ in case composition and are not a controlled overlap sweep. Rotation coverage spans 96.55–98.68%; the small number of unique rotations per bin precludes a reliable angle trend.

The low-overlap study asks whether reduced coverage comes from missing correct proposals or rejecting available ones. Primary angle and overlap strata cannot isolate these mechanisms because rotations are reused and case composition changes. On separate tuning blocks, the margin and overlap support floor were set to 0.02 and 0.25; these settings were frozen for evaluation on new blocks from the same acquisitions.

At target overlap 0.30 (realized 0.275–0.332), the decisive loss occurred before image-based selection: correct-mode recall was 33/44, conditional ranking/refinement success was 33/33, and final acceptance was 32/33 (Table 8). All 11 low- ρ_2 rejections coincided with missing correct modes. The tested threshold relaxations did not recover them; tightening ρ_2 to 0.50 instead lost another correct return. The observed bottleneck is retained proposal recall, not a finely balanced acceptance threshold.

Table 7 Primary rotation and overlap strata under frozen settings. All returns were correct. Rotation matrices are reused across acquisitions and blocks.

Factor	Level	Unique rotations	Returned/cases	Coverage
Rotation	$[60, 90]^\circ$	2	56/58	96.55%
	$[90, 120]^\circ$	7	167/170	98.24%
	$[120, 150]^\circ$	4	77/79	97.47%
	$[150, 180]^\circ$	3	75/76	98.68%
Target overlap	0.50	—	171/176	97.16%
	0.70	—	143/144	99.31%
	0.90	—	61/63	96.83%

Table 8 Low-overlap stages and one-factor sensitivity. Evaluation blocks were excluded from tuning. Unvaried thresholds in (b) remain $(\rho_2, \Delta\rho_2, \rho_3) = (0.30, 0.02, 0.30)$.

(a) Stage outcomes

Subset	Cases	Correct mode recalled	Geometry top correct	Forward/refined top correct	Returned/correct	Low- ρ_2 reject	$\mathcal{V}_3$ reject
Threshold selection	44	37	29	37	37/37	7	0
Evaluation blocks	44	33	30	33	32/32	11	1

(b) One-factor sensitivity on the 44 evaluation blocks

Quantity varied	Values tested	Returned	Correct/incorrect	Correct pose not returned
ρ_2 threshold	0.20, 0.25, 0.30, 0.35, 0.40, 0.45	32	32/0	1
ρ_2 threshold	0.50	31	31/0	2
Top-two margin $\Delta\rho_2$	0, 0.005, 0.01, 0.02, 0.03, 0.05	32	32/0	1
ρ_3 threshold	0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50	32	32/0	1

Low overlap also exposes the value of image evidence beyond geometric support. Geometry ranked a correct mode first in 30/44 evaluation cases, whereas forward selection and refinement yielded correct poses in all 33 cases with a recalled correct mode. The tuning blocks show the same pattern, 29 versus 37. This supports using image content after geometric proposal; because the reported outcome combines selection and refinement, it does not isolate the contribution of ρ_2 ranking alone.

One plausible mechanism is combinatorial loss of shared structure: K common detections permit at most $\binom{K}{3}$ true triples, before geometric pruning and signature matching. Reduced overlap can remove this redundancy while non-overlapping structures still compete for the fixed point and branch budgets. The audit does not distinguish detection loss, failed matching, or discarding correct candidates during truncation. Nor does successful final ranking exclude an upstream effect of partition support on candidate scoring. These are hypotheses requiring stage-specific counts, not established causes.

A small domain-size diagnostic reinforces the distinction between search and score compatibility. One Berea pair used full-overlap native-grid domains $80^3/128^3$, $80^3/160^3$, $160^3/256^3$, and $80^3/256^3$; each used 5° , 10° , 20° , 30° , 40° and 50° rotations about z, y, x, z, y, x , respectively, 150 refinement steps, and otherwise primary settings. The first three conditions returned all six poses correctly. The last retained no correct

Table 9 Sampling strata ($s_\star = 1$, confounded with acquisition pair) and controlled scale (one CT, three rotations, full overlap). Medians use correct returns; e_c is in fixed-grid voxels, uncalibrated for Berea.

Quantity	Level	Cases	Returned/correct	Coverage	Median e_c	Median e_R	Median $ e_s $
Sampling ratio h_f/h_m	1, Berea native grid	160	157/157	98.12%	0.344	0.101°	0.094%
	3, native Bentheimer	100	99/99	99.00%	0.845	0.079°	0.170%
	4, native MRCCM	82	78/78	95.12%	1.072	0.359°	1.569%
	4, coarsened GRS01	41	41/41	100.00%	1.050	0.140°	1.542%
Physical scale $s_\star$	0.850	3	3/3	100%	0.103	0.122°	0.103%
	0.925	3	3/3	100%	0.059	0.074°	0.035%
	1.000	3	3/3	100%	0.088	0.097°	0.034%
	1.075	3	3/3	100%	0.035	0.040°	0.014%
	1.150	3	3/3	100%	0.027	0.028°	0.021%

representative despite true-pose $\rho_2 = 0.725$ – 0.809 . High true-pose evidence cannot rescue a search that never presents that pose. This one-pair diagnostic defines no universal breakdown ratio, but supports prioritizing proposal coverage over threshold relaxation. A paired overlap sweep on held-out acquisitions should fix crop centre and transform while recording detections, valid triples, retained modes, and downstream decisions.

4.6 Sampling ratio and physical-scale recovery

Scale estimation and resolution tolerance are different tests (Table 9). Larger scale errors in both ratio-4 strata are compatible with blur and coarse sampling disturbing the detected σ/e_1 cues even when geometric edge ratios remain invariant. However, acquisition pair and sampling ratio are confounded, and no cue ablation identifies the responsible component.

Conversely, successful recovery of all controlled non-unit scales shows that the implementation is not restricted to $s = 1$. It does not resolve the cross-acquisition problem: those images share one source CT and complete support. Their small errors therefore cannot justify scale accuracy under partial overlap, changing reconstruction, or local deformation. A matched resampling intervention plus signature ablation would separate sampling effects from acquisition differences.

4.7 Reliability and computational cost

The same-case advantage concerns successful capture, not uniformly smaller errors after success (Table 10). The proposed method registered all 50 cases, versus 43 for dense search, yet dense search had the smaller conditional median e_c . Comparing these medians alone would hide its seven failures. The combined pipeline improves reliability here, but this comparison cannot attribute the gain separately to branching, scale freedom, or verification. External results likewise concern bounded-search or zero-shot implementations, not their best attainable performance after domain tuning.

Dense failures were not tolerance-edge effects: $e_c = 10.26$ – 97.31 voxels and $e_R = 1.69^\circ$ – 178.93° . Its accepted incorrect output (case-444e8c7bae18, Berea right-0.75-mm Cu) had $\rho_3 = 0.3011$, $e_c = 48.18$, $e_R = 178.93^\circ$, and centre displacement 53.34 voxels. This orientation alias is compatible with local self-similarity, not evidence of a particular pore periodicity. Because dense optimization used the full images, rescoring

Table 10 Comparison on 50 common cases. Error medians use correct returns; time uses all cases. Dense search tests 4096 Halton rotations and exhaustive integer FFT translations of three-channel local contrast, reranks with the overlap-aware objective, and refines for 200 steps at unit scale without rejection. Mattes uses a 4/2/1 pyramid, 24 cube rotations, and refinement of the best four starts. FPFH uses Otsu surfaces, two-grid-voxel downsampling, 100,000 RANSAC trials, and ICP. Super4PCS uses these surfaces, overlap 0.5, two-voxel tolerance, 1000 points, 60 s, and at most 100 point-to-plane ICP iterations. Both are rigid. GeoTransformer uses official 3DMatch weights without retraining; proposed settings are in Table 2. Panel (b) rescores stored dense poses without further optimization; its $\mathcal{V}_3$ voxels were used during dense search.

(a) Registration outcomes						
Method	Returned	Correct	Success	Median e_c	Median e_R	Time (s)
Branch-aware selective Sim(3)	50/50	50/50	100%	0.832	0.138°	84.74
Dense SO(3)–FFT search control	50/50	43/50	86%	0.777	0.168°	25.66
Mattes MI (24-start diagnostic)	50/50	0/50	0%	–	–	27.33
FPFH–RANSAC–ICP	50/50	4/50	8%	0.778	0.649°	0.47
Super4PCS–ICP	50/50	3/50	6%	0.987	0.331°	61.91
GeoTransformer (3DMatch)	50/50	1/50	2%	1.767	0.680°	1.36
(b) Dense-control fixed-pose diagnostic						
Dense output	Cases	Accepted	Rejected	Minimum ρ_3	Median ρ_3	Maximum ρ_3
Correct	43	38	5	0.2364	0.7409	0.9461
Incorrect	7	1	6	0.0001	0.0155	0.3011

is not independent validation. Nevertheless, the overlap between correct and incorrect scores demonstrates why the threshold cannot certify geometry; raising it to remove this one failure would be post-hoc tuning, not transferable calibration.

Reliability costs time: the proposed median is 3.30 times dense search and 37% above Super4PCS, whose 60-s budget was nearly exhausted in 39/50 cases (≥ 59 s). A three-run MRCCM-MEC profile gave median wall/internal times 55.33/50.87 s, peak process-tree RSS 4.94 GiB, and allocated/reserved CUDA peaks 235.5/366.0 MiB. These measurements do not locate the dominant stage or establish large-volume scaling. Profiling proposal enumeration and repeated candidate scoring should precede pruning changes that could worsen the observed recall bottleneck.

5 Conclusions

Registering independently acquired rock tomograms requires both finding a global alignment across resolution changes and deciding whether repeated structures leave that alignment ambiguous. This work contributes a branch-aware formulation of selective global similarity registration: competing structural explanations remain explicit until image evidence supports a placement and a return decision. The algorithmic distinction is that ambiguity is represented through retained pose branches and tested during selection, rather than reduced to the score of a single winning sparse match.

Three connected mechanisms implement this formulation. Relative scale-space triples combine triangle shape and approximate structure sizes to propose rotation, translation, and scale jointly without an initial alignment. Branch retention keeps distinct placements available for comparison by a constrained appearance model that

accommodates blur and intensity response. Absolute evidence and inter-branch separation then determine whether a pose is refined and considered for return. Together, these mechanisms connect cross-resolution global search to an explicit abstention option without training data. The reserved final region separates the last check from pose optimization; its incremental benefit remains unestablished, and its score does not certify geometric correctness.

The experiments support the combined procedure: it returned 375 of 383 primary poses, all correct under the stated criterion, and registered all 50 common comparison cases, versus 43 for dense search and at most four for the external implementations under their stated adaptation limits. Both cross-specimen negative cohorts produced no false returns. The demonstrated advantage is reliable global recovery under the tested conditions, rather than uniformly lower error or runtime. These comparisons do not establish the separate contribution of every component, and the controlled 22-case ablation found no additional correct returns from retaining multiple branches: both arms selected the same initial candidates. The selected stage diagnostic likewise showed no incremental accuracy gain from regional separation. These findings locate the demonstrated contribution in the combined registration procedure, not a proven advantage for every individual mechanism.

Failure analysis identifies two concrete priorities. Limited overlap primarily reduced retained correct-mode recall, motivating better proposal coverage; early screening also withheld recoverable poses, motivating refinement of borderline alternatives before a final decision. Both changes must be tested against false returns and computational cost. Validation on new acquisitions is needed to calibrate thresholds and support requirements, and processing whole-plug volumes near 1000^3 remains untested beyond the current maximum input of 288^3 .

Statements and Declarations

Funding

This work was supported by King Fahd University of Petroleum and Minerals under grant number 22011KFUPM.

Competing interests

The authors declare no relevant financial or non-financial competing interests.

Author contributions

Xingyu Zhu: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, visualization, and writing—original draft. Yuzhu Wang: conceptualization, methodology, supervision, and writing—review and editing.

Data availability

The source tomographic datasets are available from their original repositories: MRCCM ([Alqahtani et al. 2021](#)), variable-condition Berea ([Goldfarb et al. 2021](#)),

875 Green River shale (Renard and McBeck 2018), and multi-resolution Bentheimer
 876 sandstone (Jackson et al. 2021). Persistent dataset identifiers are provided in the refer-
 877 ence list. The public code repository provides a runnable three-dimensional example.
 878 Derived benchmark manifests and machine-readable results are available from the
 879 corresponding author on reasonable request.

880

881 Code availability

882

883 The implementation (Python 3.10 or later; PyTorch, NumPy, SciPy, and tiffle)
 884 and a runnable three-dimensional HR/LR example are available at <https://github.com/XingyuZhuClearfi/rockreg-selective-sim3>. The repository README documents
 885 installation, inputs, outputs, and reproduction commands. Contact: Xingyu Zhu
 886 (xingyu.zhu@kfupm.edu.sa).
 887

888

889 References

890

891 Aiger D, Mitra NJ, Cohen-Or D (2008) 4-points congruent sets for robust pairwise
 892 surface registration. In: ACM SIGGRAPH 2008 Papers, pp 1–10, [https://doi.org/](https://doi.org/10.1145/1399504.1360684)
 893 [10.1145/1399504.1360684](https://doi.org/10.1145/1399504.1360684)

894

895 Alqahtani N, Mostaghimi P, Armstrong R (2021) A multi-resolution complex car-
 896 bonates micro-ct dataset (mrccm). Digital Rocks Portal, [https://doi.org/10.17612/](https://doi.org/10.17612/3T36-Q704)
 897 [3T36-Q704](https://doi.org/10.17612/3T36-Q704), dataset

898

899 Bai X, Luo Z, Zhou L, et al (2020) D3Feat: Joint learning of dense detection and
 900 description of 3D local features. In: Proceedings of the IEEE/CVF Conference on
 901 Computer Vision and Pattern Recognition, pp 6359–6367, [https://doi.org/10.1109/](https://doi.org/10.1109/CVPR42600.2020.00639)
 902 [CVPR42600.2020.00639](https://doi.org/10.1109/CVPR42600.2020.00639)

903

904 Balakrishnan G, Zhao A, Sabuncu MR, et al (2019) VoxelMorph: A learning framework
 905 for deformable medical image registration. IEEE Transactions on Medical Imaging
 906 38(8):1788–1800. <https://doi.org/10.1109/TMI.2019.2897538>

907

908 Bay BK, Smith TS, Fyhrie DP, et al (1999) Digital volume correlation: Three-
 909 dimensional strain mapping using x-ray tomography. Experimental Mechanics
 910 39(3):217–226. <https://doi.org/10.1007/BF02323555>

911

912 Besl PJ, McKay ND (1992) A method for registration of 3-d shapes. IEEE Transactions
 913 on Pattern Analysis and Machine Intelligence 14(2):239–256. [https://doi.org/10.](https://doi.org/10.1109/34.121791)
 914 [1109/34.121791](https://doi.org/10.1109/34.121791)

915

916 Botha PWSK, Sheppard AP (2016) Mapping permeability in low-resolution micro-ct
 917 images: A multiscale statistical approach. Water Resources Research 52(6):4377–
 918 4398. <https://doi.org/10.1002/2015WR018454>

919

919 Buljac A, Jailin C, Mendoza A, et al (2018) Digital volume correlation: Review of
 920 progress and challenges. Experimental Mechanics 58(5):661–708. <https://doi.org/>

10.1007/s11340-018-0390-7	921
	922
Chen J, Frey EC, He Y, et al (2022) TransMorph: Transformer for unsupervised medical image registration. <i>Medical Image Analysis</i> 82:102615. https://doi.org/10.1016/j.media.2022.102615	923
	924
	925
	926
Chow CK (1970) On optimum recognition error and reject tradeoff. <i>IEEE Transactions on Information Theory</i> 16(1):41–46. https://doi.org/10.1109/TIT.1970.1054406	927
	928
	929
Cnudde V, Boone MN (2013) High-resolution x-ray computed tomography in geosciences: A review of the current technology and applications. <i>Earth-Science Reviews</i> 123:1–17. https://doi.org/10.1016/j.earscirev.2013.04.003	930
	931
	932
Fischler MA, Bolles RC (1981) Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. <i>Communications of the ACM</i> 24(6):381–395. https://doi.org/10.1145/358669.358692	933
	934
	935
	936
Geifman Y, El-Yaniv R (2017) Selective classification for deep neural networks. In: <i>Advances in Neural Information Processing Systems</i> , pp 4878–4887	937
	938
	939
Goldfarb E, Ikeda K, Prodanovic M, et al (2021) Effects on digital rock physics models from variable computed tomography scans settings. <i>Digital Rocks Portal</i> , https://doi.org/10.17612/RADZ-GV42 , dataset	940
	941
	942
	943
Hall SA, Bornert M, Desrues J, et al (2010) Discrete and continuum analysis of localised deformation in sand using x-ray micro-ct and volumetric digital image correlation. <i>Géotechnique</i> 60(5):315–322. https://doi.org/10.1680/geot.2010.60.5.315	944
	945
	946
	947
	948
Heinrich MP, Jenkinson M, Bhushan M, et al (2012) MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. <i>Medical Image Analysis</i> 16(7):1423–1435. https://doi.org/10.1016/j.media.2012.05.008	949
	950
	951
	952
Hu Y, Modat M, Gibson E, et al (2018) Weakly-supervised convolutional neural networks for multimodal image registration. <i>Medical Image Analysis</i> 49:1–13. https://doi.org/10.1016/j.media.2018.07.002	953
	954
	955
	956
Huang S, Gojcic Z, Usvyatsov M, et al (2021) PREDATOR: Registration of 3D point clouds with low overlap. In: <i>Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition</i> , pp 4267–4276, https://doi.org/10.1109/CVPR46437.2021.00425	957
	958
	959
	960
Jackson SJ, Niu Y, Manoorkar S, et al (2021) Multi-resolution x-ray micro-ct images of bentheimer sandstones. Zenodo, https://doi.org/10.5281/zenodo.5542624 , version 1, dataset	961
	962
	963
	964
	965
	966

967 Klein S, Staring M, Murphy K, et al (2010) elastix: A toolbox for intensity-based
 968 medical image registration. *IEEE Transactions on Medical Imaging* 29(1):196–205.
 969 <https://doi.org/10.1109/TMI.2009.2035616>
 970
 971 Lamdan Y, Wolfson HJ (1988) Geometric hashing: a general and efficient model-
 972 based recognition scheme. *Proceedings of the Second International Conference on*
 973 *Computer Vision* pp 238–249
 974
 975 Latham SJ, Varslot TK, Sheppard AP (2008) Automated registration for augmenting
 976 micro-ct 3d images. In: *Proceedings of the 14th Biennial Computational Techniques*
 977 *and Applications Conference*, pp C534–C548, [https://doi.org/10.21914/anziamj.](https://doi.org/10.21914/anziamj.v50i0.1389)
 978 [v50i0.1389](https://doi.org/10.21914/anziamj.v50i0.1389)
 979
 980 Lindeberg T (1998) Feature detection with automatic scale selection. *Inter-*
 981 *national Journal of Computer Vision* 30(2):79–116. [https://doi.org/10.1023/A:](https://doi.org/10.1023/A:1008045108935)
 982 [1008045108935](https://doi.org/10.1023/A:1008045108935)
 983
 984 Mattes D, Haynor DR, Vesselle H, et al (2001) Nonrigid multimodality image regis-
 985 tration. In: *Medical Imaging 2001: Image Processing*, vol 4322. SPIE, pp 1609–1620,
 986 <https://doi.org/10.1117/12.431046>
 987
 988 Mellado N, Aiger D, Mitra NJ (2014) Super4PCS: Fast global pointcloud registration
 989 via smart indexing. *Computer Graphics Forum* 33(5):205–215. [https://doi.org/10.](https://doi.org/10.1111/cgf.12446)
 990 [1111/cgf.12446](https://doi.org/10.1111/cgf.12446)
 991
 992 Mok TCW, Chung ACS (2020) Large deformation diffeomorphic image registra-
 993 tion with laplacian pyramid networks. In: *Medical Image Computing and Com-*
 994 *puter Assisted Intervention – MICCAI 2020*, pp 211–221, [https://doi.org/10.1007/](https://doi.org/10.1007/978-3-030-59716-0_21)
 995 [978-3-030-59716-0_21](https://doi.org/10.1007/978-3-030-59716-0_21)
 996
 997 Niu Y, Jackson SJ, Alqahtani N, et al (2022) Paired and unpaired deep learning
 998 methods for physically accurate super-resolution carbonate rock images. *Transport*
 999 *in Porous Media* 144(3):825–847. <https://doi.org/10.1007/s11242-022-01842-z>
 1000
 1001 Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation
 1002 invariant texture classification with local binary patterns. *IEEE Transactions on*
 1003 *Pattern Analysis and Machine Intelligence* 24(7):971–987. [https://doi.org/10.1109/](https://doi.org/10.1109/TPAMI.2002.1017623)
 1004 [TPAMI.2002.1017623](https://doi.org/10.1109/TPAMI.2002.1017623)
 1005
 1006 Pluim JPW, Maintz JBA, Viergever MA (2003) Mutual-information-based regis-
 1007 tration of medical images: A survey. *IEEE Transactions on Medical Imaging*
 1008 22(8):986–1004. <https://doi.org/10.1109/TMI.2003.815867>
 1009
 1010 Qin Z, Yu H, Wang C, et al (2022) Geometric transformer for fast and robust
 1011 point cloud registration. In: *Proceedings of the IEEE/CVF Conference on Com-*
 1012 *puter Vision and Pattern Recognition*, pp 11143–11152, [https://doi.org/10.1109/](https://doi.org/10.1109/CVPR52688.2022.01086)
[CVPR52688.2022.01086](https://doi.org/10.1109/CVPR52688.2022.01086)

Reddy BS, Chatterji BN (1996) An fft-based technique for translation, rotation, and scale-invariant image registration. IEEE Transactions on Image Processing 5(8):1266–1271. https://doi.org/10.1109/83.506761	1013 1014 1015 1016
Renard F, McBeck J (2018) Investigating the onset of strain localization within anisotropic shale using digital volume correlation of time-resolved x-ray microtomography images. NIRD Research Data Archive, https://doi.org/10.11582/2018.00005 , dataset	1017 1018 1019 1020 1021
Rister B, Horowitz MA, Rubin DL (2017) Volumetric image registration from invariant keypoints. IEEE Transactions on Image Processing 26(10):4900–4910. https://doi.org/10.1109/TIP.2017.2722689	1022 1023 1024 1025
Rusu RB, Blodow N, Beetz M (2009) Fast point feature histograms (fpfh) for 3d registration. In: 2009 IEEE International Conference on Robotics and Automation, IEEE, pp 3212–3217, https://doi.org/10.1109/ROBOT.2009.5152473	1026 1027 1028 1029
Torr PHS, Zisserman A (2000) MLESAC: A new robust estimator with application to estimating image geometry. Computer Vision and Image Understanding 78(1):138–156. https://doi.org/10.1006/cviu.1999.0832	1030 1031 1032 1033
Tudisco E, Andò E, Cailletaud R, et al (2017) TomoWarp2: A local digital volume correlation code. SoftwareX 6:267–270. https://doi.org/10.1016/j.softx.2017.10.002	1034 1035 1036
Umeyama S (1991) Least-squares estimation of transformation parameters between two point patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(4):376–380. https://doi.org/10.1109/34.88573	1037 1038 1039 1040
Viola P, Wells WM (1997) Alignment by maximization of mutual information. International Journal of Computer Vision 24(2):137–154. https://doi.org/10.1023/A:1007958904918	1041 1042 1043 1044
Yang H, Shi J, Carlone L (2021) TEASER: Fast and certifiable point cloud registration. IEEE Transactions on Robotics 37(2):314–333. https://doi.org/10.1109/TRO.2020.3033695	1045 1046 1047 1048
Zeng A, Song S, Niessner M, et al (2017) 3DMatch: Learning local geometric descriptors from RGB-D reconstructions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 1802–1811, https://doi.org/10.1109/CVPR.2017.29	1049 1050 1051 1052 1053 1054 1055 1056 1057 1058