

# Mathematical methods for matrix comparison in fiber optic defect detection

Comprehensive mathematical research has revealed exhaustive approaches for comparing matrices derived from fiber optic images, specifically targeting defect detection without machine learning. The following methods provide rigorous mathematical foundations for identifying scratches, digs, blobs, and other anomalies in fiber optic systems.

## Classical mathematical approaches for matrix differences

**Matrix subtraction and norm-based measures** form the foundation of comparison techniques. The Frobenius norm  $\|A - B\|_F = \sqrt{(\sum_{ij} |a_{ij} - b_{ij}|^2)}$  provides global similarity assessment and equals  $\sqrt{(\sum \sigma_i^2)}$  where  $\sigma_i$  are singular values. [\(Wikipedia +3\)](#) The L1 norm  $\|A - B\|_1 = \sum_{ij} |a_{ij} - b_{ij}|$  offers robustness to outliers, while the L $\infty$  norm  $\|A - B\|_\infty = \max |a_{ij} - b_{ij}|$  identifies maximum local variations [\(Wikipedia\)](#) ideal for concentrated defects. [\(Stack Exchange +2\)](#)

**Hadamard product methods** enable element-wise comparisons through  $(A \circ B)[i,j] = A[i,j] \times B[i,j]$ . [\(Wikipedia\)](#) This supports normalized correlation  $\rho = (A \circ B)^T 1 / (\|A\|_F \|B\|_F)$  and element-wise relative error  $E = |A - B| \circ (1 ./ |A|)$ , useful for point-wise intensity analysis.

**Rank-based comparison** leverages singular value decomposition  $A = U \Sigma V^T$  to detect structural changes. The effective rank formula  $R_e = \exp(-\sum_i p_i \log(p_i))$  where  $p_i = \sigma_i^2 / \sum_j \sigma_j^2$  quantifies information content. Trace-based similarity  $\text{tr}(A^T B) / \sqrt{(\text{tr}(A^T A) \text{tr}(B^T B))}$  provides global intensity comparison. [\(Gatech +2\)](#)

## Statistical methods for anomaly detection

**Hotelling's T-squared test** extends t-tests to multivariate matrix data:  $T^2 = (n_1 n_2) / (n_1 + n_2) \times (\bar{X}_1 - \bar{X}_2)^T S^{-1} (\bar{X}_1 - \bar{X}_2)$ , where S is pooled covariance. This enables systematic difference detection across fiber cross-sections. [\(Statistics How To\)](#)

**Mahalanobis distance**  $D_m^2(x) = (x - \mu)^T \Sigma^{-1} (x - \mu)$  provides multivariate anomaly detection. [\(Machinelearningplus\)](#) [\(Built In\)](#) Under normality,  $D_m^2 \sim \chi^2(p)$ , enabling threshold determination at significance level  $\alpha$ . [\(Charles Holbert\)](#) Robust variants using Minimum Covariance Determinant handle contaminated data. [\(SpringerLink\)](#)

**Change point detection algorithms** include matrix-adapted CUSUM:  $S_n = \sum_{i=1}^n \text{tr}((X_i - \mu_0)^T \Sigma^{-1} (X_i - \mu_0))$  and PELT with cost function  $C(y_{1:n}) = \min [\sum_j [\text{cost}(y_{t_{j-1}+1:t_j}) + \beta]]$ . These detect degradation onset and discrete failure events. [\(Arcgis\)](#)

**Local Outlier Factor** adapted for matrices uses local reachability density  $\text{lrd\_k}(A) = 1 / (\sum_{\{B \in N\_k(A)\}} \text{reach\_dist\_k}(A, B) / |N\_k(A)|)$  and  $\text{LOF\_k}(A) = (\sum_{\{B \in N\_k(A)\}} \text{lrd\_k}(B) / |N\_k(A)|) / \text{lrd\_k}(A)$  for anomaly scoring.

## Information-theoretic and spectral approaches

**Singular Value Decomposition** provides low-rank approximation  $A_k = U_k \Sigma_k V_k^T$  with reconstruction error  $\|A - A_k\|_F$  for defect detection. Research shows SVD-based PCA achieves 98.46% sensitivity versus 42.19% for traditional methods. ([ScienceDirect](#)) ([Science](#))

**Fourier transform analysis**  $F(u,v) = \sum \sum f(x,y) e^{-j2\pi(ux/M + vy/N)}$  enables frequency domain defect identification. Power spectrum  $|F(u,v)|^2$  reveals periodic patterns with 80% accuracy for structural anomalies.

**Wavelet transforms**  $W(a,b) = (1/\sqrt{a}) \int f(t) \psi^*((t-b)/a) dt$  provide multi-scale analysis. Discrete Wavelet Transform enables real-time processing, while wavelet packet decomposition enhances frequency resolution for fine defect classification. ([SpringerLink +2](#))

**Information entropy**  $H(X) = -\sum p(x) \log_2 p(x)$  and mutual information  $I(X;Y) = \sum p(x,y) \log(p(x,y)/(p(x)p(y)))$  quantify statistical dependencies. These identify most informative features for defect classification.

## Methods for different-dimension matrices

**Padding techniques** include zero padding and symmetric padding  $A_{\text{padded}} = [A[n,:]; A A[1,:]; A A A[1,:]; A A[n,:]]$  to match dimensions. **Interpolation methods** use bilinear  $f(x,y) = f(0,0)(1-x)(1-y) + f(1,0)x(1-y) + f(0,1)(1-x)y + f(1,1)xy$  or bicubic approaches.

**Projection techniques** employ orthogonal projection  $P = A(A^T A)^{-1} A^T$  ([Stack Exchange](#)) or principal component projection via SVD:  $A_{\text{proj}} = U_k \Sigma_k V_k^T$ . **Registration methods** solve  $\min \|B - T(A)\|_F^2$  where  $T$  represents transformation parameters.

## Morphological and topological defect detection

**Mathematical morphology operations** use erosion  $A \ominus B = \{z \mid (B)_z \subseteq A\}$  and dilation  $A \oplus B = \{z \mid (B^s)_z \cap A \neq \emptyset\}$ . ([Mathworks](#)) ([Wikipedia](#)) Opening  $A \circ B = (A \ominus B) \oplus B$  removes small objects while closing  $A \bullet B = (A \oplus B) \ominus B$  fills gaps. ([Wikipedia](#)) ([Auckland](#))

**Morphological gradient**  $\nabla f = (f \oplus B) - (f \ominus B)$  enhances edges. ([Mathworks](#)) ([Wikipedia](#)) **Top-hat transforms**  $WTH(f) = f - (f \circ B)$  extract bright defects, while  $BTH(f) = (f \bullet B) - f$  extract dark defects. Multi-scale approaches use  $\text{Enhanced\_Image} = \sum_i \alpha_i \times WTH(f, B_i)$ . ([ScienceDirect](#))

**Persistent homology** analyzes filtrations  $\emptyset = X_{-1} \subseteq X_0 \subseteq \dots \subseteq X_n = X$  with persistence = death\_time - birth\_time. ([SpringerOpen](#)) ([Nature](#)) Zero-dimensional persistence detects individual defects, while 1-dimensional persistence identifies loops and holes. ([Wikipedia](#)) ([NCBI](#))

**Watershed segmentation** uses topographic interpretation with watershed lines at  $\{x \mid \forall \epsilon > 0, \exists y, z \in B(x, \epsilon) : CB(y) \neq CB(z)\}$ . ([Wikipedia](#)) ([GeeksforGeeks](#)) Marker-controlled variants separate touching defects using distance transforms. ([Wikipedia](#)) ([Opencv](#))

## Distance metrics and similarity measures

**Matrix norms** include Frobenius  $\|A\|_F = \sqrt{\text{trace}(AA^T)}$ , spectral  $\|A\|_2 = \sigma_{\max}(A)$ , and nuclear  $\|A\|_* = \sum \sigma_i(A)$ . (Wikipedia +3) Each captures different defect characteristics: Frobenius for global changes, spectral for directional defects, nuclear for low-rank patterns. (Wikipedia)

**Procrustes analysis** solves  $\min_R \|Y - XR\|_F^2$  subject to  $R^T R = I$  for optimal alignment, compensating for rotation/scaling/translation in imaging variations. (Scipy +3)

**Earth Mover's Distance**  $\text{EMD}(P, Q) = (\sum \sum f_{ij} d_{ij}) / (\sum \sum f_{ij})$  captures spatial redistribution, (Wikipedia) excellent for displacement-type defects and texture changes. (Wikipedia +3)

**Structural Similarity Index**  $\text{SSIM}(x, y) = ((2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)) / ((\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2))$  aligns with human perception. (Wikipedia +3) Multi-scale MS-SSIM combines multiple resolutions. (Wikipedia)

**Geodesic distances** on matrix manifolds include log-Euclidean  $d(A, B) = \|\log(A) - \log(B)\|_F$  for symmetric positive definite matrices, (Wikipedia) providing natural metrics for covariance-based features.

## Computational procedures for comparative analysis

**Block-wise comparison** partitions matrices into blocks with hierarchical comparison:  $D_{\text{block}} = \sum_{ij} w_{ij} \|A_{ij} - B_{ij}\|_F$ . This enables multi-resolution analysis and computational scalability.

**Weighted difference matrices**  $D_w = W \circ (A - B)$  use importance masks. Gaussian weights  $W[i, j] = \exp(-(i - i_0)^2 + (j - j_0)^2) / (2\sigma^2))$  emphasize critical regions. Adaptive weights  $W[i, j] = 1 / (1 + |A[i, j] - \mu_{\text{local}}| / \sigma_{\text{local}})$  adjust locally.

**Multi-scale analysis framework** combines Gaussian pyramid decomposition with scale-specific comparisons. Automatic scale selection uses  $t^* = \arg\max_t |L_{xx}^{\text{norm}} + L_{yy}^{\text{norm}}|$  where  $L_x^{\text{norm}} = t^{(\gamma/2)} L_x(x, y, t)$ .

## Specific methods for defect types

**Scratch detection** employs Hough transform with polar parameterization  $p = x \cos \theta + y \sin \theta$  and oriented accumulator  $H(p, \theta) = \sum_i w(\theta - \theta_i) \times \delta(p - x_i \cos \theta - y_i \sin \theta)$ . Linear structuring elements in morphological operations preserve scratch-like structures. (ResearchGate)

**Dig detection** uses disk-shaped structuring elements with size determining minimum detectable diameter. Connected component analysis with geometric moments  $m_{pq} = \iint x^p y^q f(x, y) dx dy$  characterizes circular defects.

**Blob detection** combines opening/closing operations with appropriate structuring elements. Shape descriptors include compactness  $4\pi A/P^2$ , aspect ratio, and Hu invariant moments  $\phi_i$  for rotation-invariant classification.

## Integrated computational framework

The optimal approach combines multiple methods hierarchically:

1. **Fast screening:** Euclidean distance, cosine similarity ([Wikipedia](#)) ([datacamp](#)) ( $O(mn)$  complexity)
2. **Detailed analysis:** SSIM, wavelet transforms, morphological operations
3. **Statistical validation:** Mahalanobis distance, hypothesis testing
4. **Verification:** Procrustes alignment, persistent homology

**Performance optimization** leverages:

- BLAS/LAPACK libraries for matrix operations
- Parallel processing for block-wise methods
- Early termination for hierarchical approaches
- Adaptive thresholding based on image statistics

**Quality assessment metrics** include:

- Detection Rate:  $TP/(TP + FN)$
- False Alarm Rate:  $FP/(FP + TN)$
- Dice coefficient:  $2|A \cap B|/(|A| + |B|)$
- Hausdorff distance:  $\max\{\sup_{a \in A} \inf_{b \in B} d(a,b), \sup_{b \in B} \inf_{a \in A} d(a,b)\}$

This comprehensive mathematical framework provides rigorous, non-machine-learning approaches for fiber optic defect detection through matrix comparison, enabling automated quality control with theoretical guarantees and practical applicability across varying imaging conditions and defect types.