

Multi-Axis Variance Vector Extraction: Nine Dimensions of Structural Content Identity

by [Nick Clark](#) | Published March 27, 2026

Mechanism

The extractor accepts a content artifact — an image, an audio segment, a video frame sequence, a document rasterization, or any signal admitting a regular sampling grid — and produces an ordered tuple of nine real-valued scalars constrained to a bounded interval. Each scalar corresponds to a distinct structural axis evaluated over one of three analysis bands: a coarse band capturing global structure, an intermediate band capturing regional structure, and a fine band capturing local structure. Within each band the extractor computes three measurements: the variance of the energy distribution across the band, a frequency-compaction figure expressing how concentrated or dispersed the spectral content is within the band, and a structural phase persistence score expressing how stably the dominant structural patterns survive small perturbations. Three bands times three measurements yields the nine dimensions of the output vector.

Energy-distribution variance is computed by partitioning the band into a fixed number of cells, computing the energy contained in each cell, and taking the statistical variance of those per-cell energy values after normalization to unit total energy. The result is bounded below by zero (energy distributed uniformly across cells, yielding identical per-cell values) and above by the variance of a

delta distribution over the cell count (all energy concentrated in a single cell). Because the measurement is computed on relative per-cell proportions rather than absolute magnitudes, it is insensitive to global gain changes and resilient to additive offsets that do not alter the distribution shape.

Frequency-compactness is computed by transforming the band into a frequency-domain representation, ranking the resulting coefficients by magnitude, and recording the proportion of total energy captured by the leading fraction of coefficients. A signal whose energy concentrates in a small number of dominant components yields a high compactness score; a signal whose energy spreads across many components yields a low score. This measurement responds to the underlying structural regularity of the content and remains stable under transformations that preserve regularity even while altering pixel values, sample values, or visual appearance.

Structural phase persistence is computed by applying a family of small, parameterized perturbations to the band — sub-sample shifts, modest rotations, low-amplitude noise injections, and bounded resamplings — recomputing the dominant structural pattern after each perturbation, and measuring the cosine similarity between the perturbed pattern and the original. The persistence score is the mean similarity across the perturbation family. Content with stable internal structure produces a persistence score near unity; content whose dominant patterns are accidental or noise-driven produces a substantially lower score.

Each scalar is normalized into the unit interval before assembly into the output vector. The bounded range serves three purposes: it permits direct vector-similarity comparisons under standard metrics (cosine similarity, Euclidean distance, or weighted L1) without per-axis rescaling; it constrains the search space so that index structures over the vectors remain compact; and it ensures that no single axis can dominate similarity computations through unbounded magnitude.

Operating Parameters

The extractor exposes a small number of operating parameters governing band partition, cell count, perturbation amplitude, and similarity metric. Band partition follows a dyadic decomposition by default — coarse band covering the lowest octave of structural scale, intermediate band covering the middle octaves, and fine band covering the highest octaves resolvable at the input sampling rate — but admits configuration for content classes whose informative structure concentrates at unusual scales. Cell count for variance computation is set so that the expected occupancy per cell remains statistically meaningful even for short signals; the default value is sixty-four cells per band.

Perturbation amplitudes for the persistence measurement are chosen to span the regime in which legitimate transformations (mild compression, modest resampling, small rotations introduced by display geometry) operate without entering the regime in which content has been substantively altered. A perturbation budget of three to five percent of the input signal energy has been observed to discriminate cleanly between identity-preserving transformations and identity-altering edits across the content classes evaluated during reduction to practice.

The similarity metric used for resolution is configurable. Cosine similarity is the default because it disregards differences in vector magnitude that may arise from minor preprocessing variations and emphasizes the geometric orientation of the structural fingerprint. For applications requiring metric-space indexing, weighted Euclidean distance with axis weights derived from the inverse variance of each axis across a calibration corpus produces nearest-neighbor results equivalent to cosine similarity for normalized vectors while admitting standard tree-based indexing. A bounded comparison threshold — for example, a cosine similarity above ninety-two hundredths declares identity — converts the continuous similarity measure into a binary resolution decision suitable for downstream protocol logic.

Alternative Embodiments

Although the canonical embodiment uses three bands, three measurements, and a nine-dimensional output, the architecture admits embodiments at higher and lower dimensionalities. A reduced six-dimensional embodiment omits the persistence axis, producing a vector adequate for applications that tolerate higher false-match rates in exchange for reduced extraction cost. An expanded twelve-dimensional embodiment adds a fourth band — a sub-band capturing structural scales below the coarse band's lower bound — for content classes such as long-form audio in which structural regularities at very low frequencies carry substantial identity information.

The frequency-domain transform underlying the compaction measurement may be selected from a family including the discrete cosine transform, wavelet decompositions of varying basis function, and learned transforms produced by autoencoder training over a representative corpus. Selection is driven by the structural priors of the content class: image content benefits from wavelet decompositions that align with natural image statistics, while audio content benefits from constant-Q transforms that align with logarithmic pitch perception. The output vector remains nine-dimensional and bounded regardless of the underlying transform; only the per-axis interpretation shifts.

For streaming media in which extraction must occur without buffering the entire artifact, an incremental embodiment computes the per-band measurements over sliding windows and aggregates the windowed measurements into the output vector through bounded statistical summaries (median, trimmed mean, robust variance). The streaming embodiment yields vectors that are statistically close to those produced by the canonical embodiment over the same content while bounding extraction memory to a fixed window size independent of content duration.

The extractor may also be embodied as a hardware primitive: a fixed-function block within a content-handling pipeline that produces the nine-dimensional

vector at line rate. Such an embodiment enables anchoring of content at the moment of capture or transcoding, attaching the structural fingerprint to the artifact before any opportunity for metadata stripping arises.

Composition with Adjacent Primitives

The variance vector functions as one element of a broader content anchoring system. Within that system the vector composes with cryptographic commitment of the vector to an append-only ledger, with policy objects that govern which similarity thresholds apply to which content classes, and with provenance records that bind the vector to authorship and licensing assertions. The vector itself is content-derived and therefore reproducible by any party in possession of the artifact; the cryptographic commitment establishes a temporal anchor demonstrating when the vector was first registered.

When composed with vector-similarity indices maintained across federated participants, the variance vector enables content resolution without exposure of the underlying artifact. A participant holding a candidate artifact can compute its vector locally and query the index for near matches; the index returns metadata associated with matched vectors without ever requiring the candidate artifact itself to traverse the network. This composition supports privacy-preserving provenance lookups at scale.

The bounded nature of each axis simplifies composition with locality-sensitive hashing schemes that map the vector to a discrete bucket identifier. The bucket identifier is short enough to embed within transport headers and stable enough that minor perturbations of the underlying content produce identical or adjacent buckets. This composition supports first-pass filtering at network edges before more expensive vector-similarity computations are invoked.

Prior-Art Posture

Prior approaches to content fingerprinting fall into three families. Perceptual hashing schemes — pHash, dHash, and their descendants — produce short binary fingerprints suitable for near-duplicate detection but compress information so aggressively that nine-dimensional structural distinctions collapse into a single Hamming distance. Embedding-based approaches derived from neural feature extractors produce high-dimensional vectors with strong discrimination but unbounded per-axis magnitudes and opaque interpretability, complicating reproducibility and standardization. Watermarking approaches embed identity into the artifact itself, but rely on the watermark surviving downstream transformations and provide no fallback when the watermark is stripped.

The disclosed extractor differs from each family along axes germane to patentability. Against perceptual hashing it asserts a richer, multi-axis structural representation supporting graded similarity rather than binary near-duplication. Against embedding-based approaches it asserts a bounded, interpretable, low-dimensional vector reproducible without access to a trained model. Against watermarking it asserts that the fingerprint is derived from the content's own structure rather than injected, eliminating the survivability problem entirely.

A fourth family — content registry systems indexed by submitted hashes of canonical files — is sometimes characterized as a structural primitive but functions as a metadata-binding scheme. A registered hash binds a file's exact bytes to the registry's record of provenance and rights; any transformation that alters the file's bytes produces a hash with no relation to the registered identifier and severs the binding. The disclosed extractor does not depend on byte-level identity. Two artifacts whose bytes differ but whose underlying structure is preserved produce nearby vectors under the chosen similarity metric, and the resolution decision is taken on the basis of that proximity rather than on byte equality. Ordinary content-handling operations — format conversions, modest compressions, viewport-driven resamplings — therefore do not sever the identity binding established at registration time.

The disclosure further departs from prior art in its treatment of per-axis interpretability. Each of the nine output dimensions corresponds to a named structural measurement with a defined computation; an auditor presented with a vector and an artifact can recompute the vector and verify, axis by axis, that the published value matches the artifact's structure. No comparable per-dimension auditability exists in embedding-based fingerprints, where individual coordinates of a learned vector carry no human-interpretable meaning and cannot be independently verified without access to the model that produced them.

Disclosure Scope

The scope claimed under US 63/808,372 covers the extraction of bounded multi-axis variance vectors comprising at least three structurally distinct measurement classes evaluated across at least two analysis bands, the use of vector-similarity comparison over such vectors for content resolution, and the composition of such vectors with cryptographic commitment and federated indexing infrastructure as described. The disclosure further covers the streaming, hardware, and learned-transform embodiments outlined above, and the operating-parameter ranges within which the extractor has been observed to function. The claimed scope expressly excludes single-scalar perceptual hashes, unbounded neural embeddings, and injected watermark schemes; these are recited as prior art rather than as embodiments.