DTDR: Distributed Transform-Domain Representation

A Persistent Representation for High-Dimensional Semantic Data

1 Overview

Distributed Transform-Domain Representation (DTDR) is a method for representing numerical data, including machine-learning model parameters and vector embeddings, in a structured transform domain that preserves computational functionality while reducing memory footprint and bandwidth demands.

Unlike conventional compression schemes, DTDR produces a representation that is not merely suitable for storage and later reconstruction, but can serve as a *native numerical domain* for inference, similarity search, and approximate nearest-neighbour (ANN) computation.

1.1 Core Idea

DTDR applies one or more structured orthogonal transforms (such as Walsh–Hadamard or discrete cosine transforms) to numerical data, followed by optional quantisation. Crucially, the transformed representation distributes information across coefficients rather than localising it in contiguous regions of memory.

This distribution gives rise to several distinctive properties:

- Robustness to partial corruption or loss
- Graceful degradation of computational functionality
- Compatibility with existing kernels and distance metrics
- Reduced memory bandwidth pressure

These properties hold even when the representation is not reconstructed to full numerical fidelity.

1.2 DTDR as a Computational Domain

Recent experiments demonstrate that DTDR can function as a *unified computational domain*, supporting multiple stages of modern machine-learning pipelines without reverting to full-precision representations.

In particular, DTDR representations have been shown to support:

- Model inference using reconstructed working precision

- Vector similarity computation directly in the transform domain
- Approximate nearest-neighbour search using established methods such as IVF and HNSW
- Binary distance estimation and re-ranking

All of these operations can be performed directly on DTDR vectors.

1.3 Emergent Structure and Multi-Resolution Signals

Because information is distributed across coefficients, DTDR representations exhibit meaningful structure across multiple aggregation scales. By examining similarity persistence under progressive aggregation or dilution, it is possible to extract additional localisation signals that do not exist in the original representation.

This multi-resolution persistence has been shown to improve coarse localisation in ANN pipelines, particularly in low-probe regimes where conventional centroid-based methods discard the most information.

1.4 Scope

DTDR is not proposed as a replacement for existing inference or ANN methods, but as a representation that is compatible with them while exposing additional computational structure. Its primary relevance is therefore to memory-bound and bandwidth-limited systems, where reducing data movement while preserving functionality is critical.