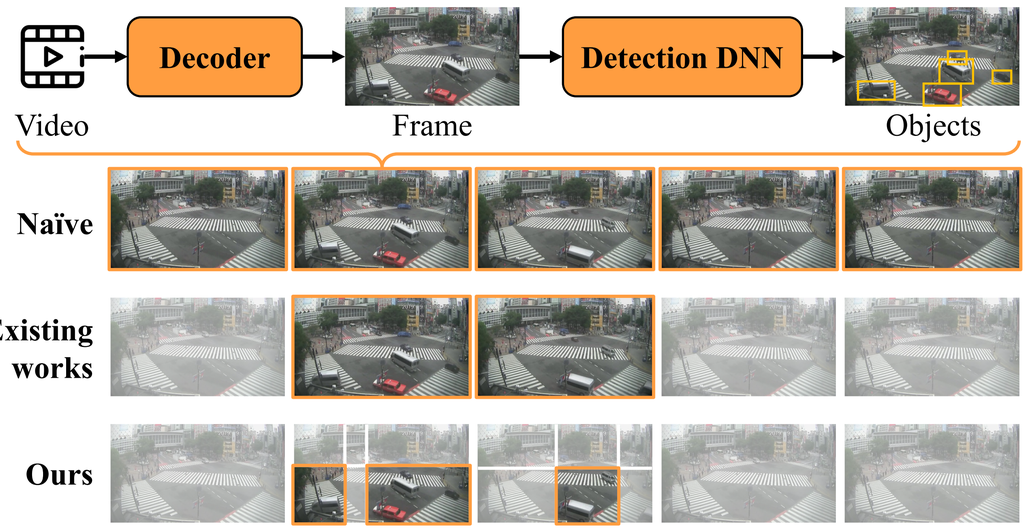
TVM: A Tile-based Video Management Framework

Tianxiong Zhong[1] , Zhiwei Zhang[1] , Guo Lu[2], Ye Yuan[1], Yu-Ping Wang[1], and Guoren Wang[1]

[1]Beijing Institute of Technology [2]Shanghai JiaoTong University

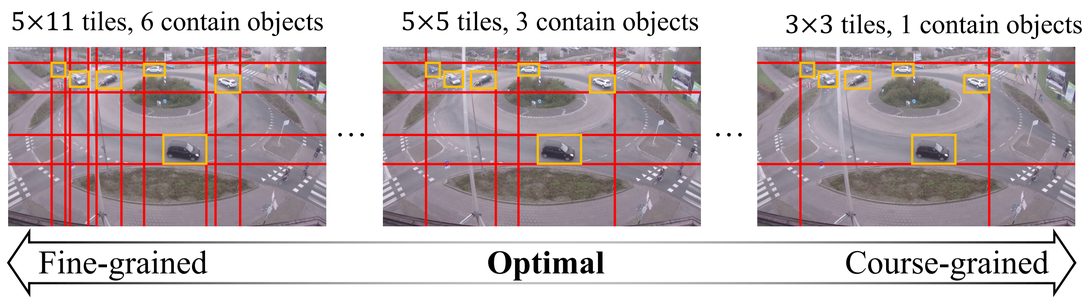
Motivation

Query: Return 99% of the frames in the video that contain cars.



1. Existing work does not consider decoding cost
2. Background pixels are irrelevant to the query, and decoding and processing them is a waste

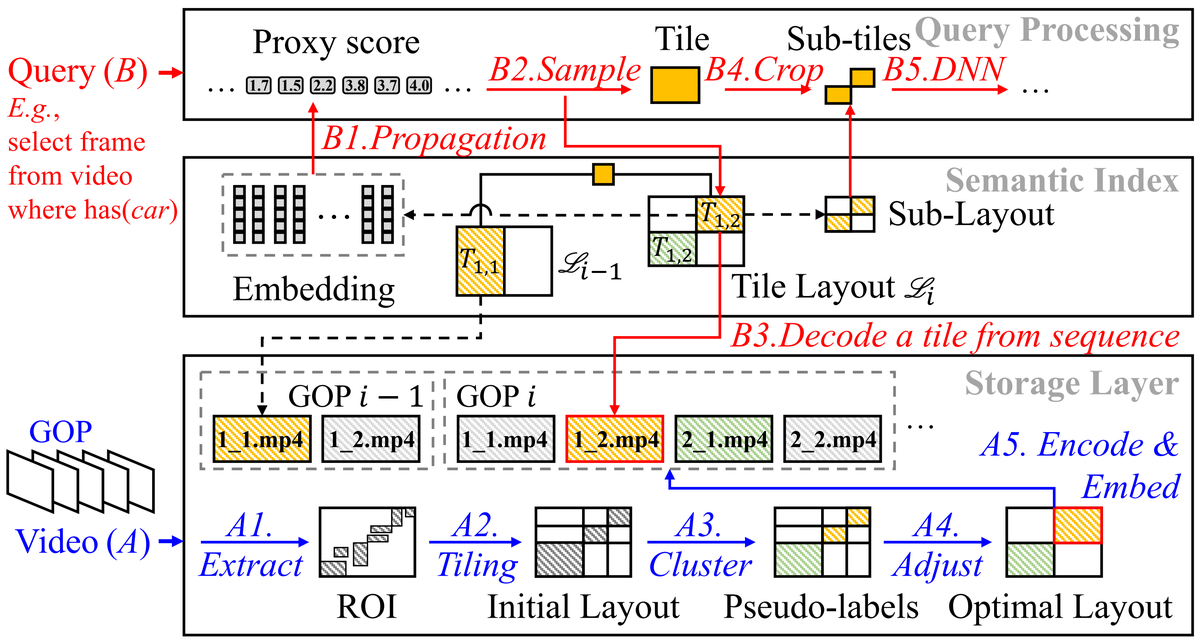
Challenge

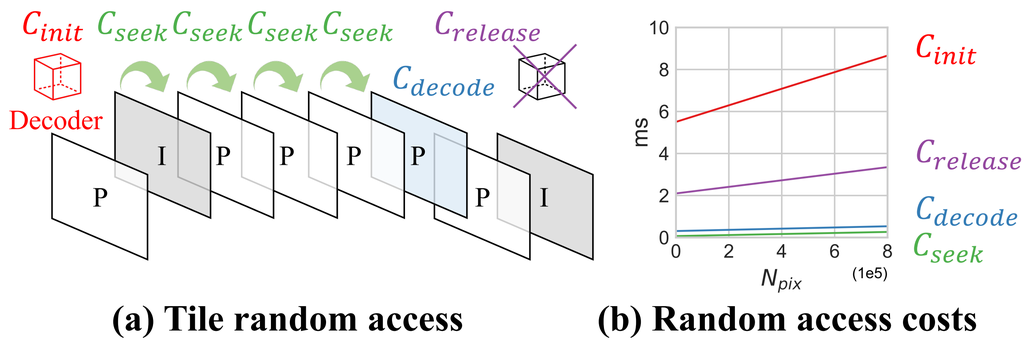


Fine-grained	Optimal	Course-grained
✓ Small tile size	✓ **Suitable tile size**	✗ Huge tile size
✗ Decode many tiles	✓ **Decode less tiles**	✓ Decode only one tile

1. How to efficiently determine the tiling granularity (layout) to minimize the processing cost?
2. How to use tiles to further speed up various exsiting query methods?
3. Query workflows for accessing tiles are different. How to avoid wasting resources?

Index Structure with Optimal Layout

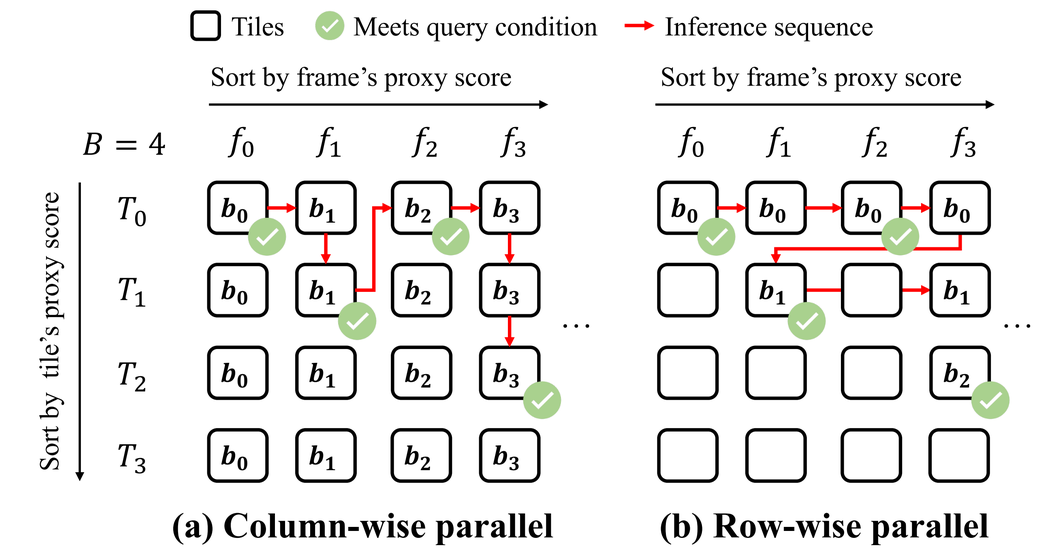




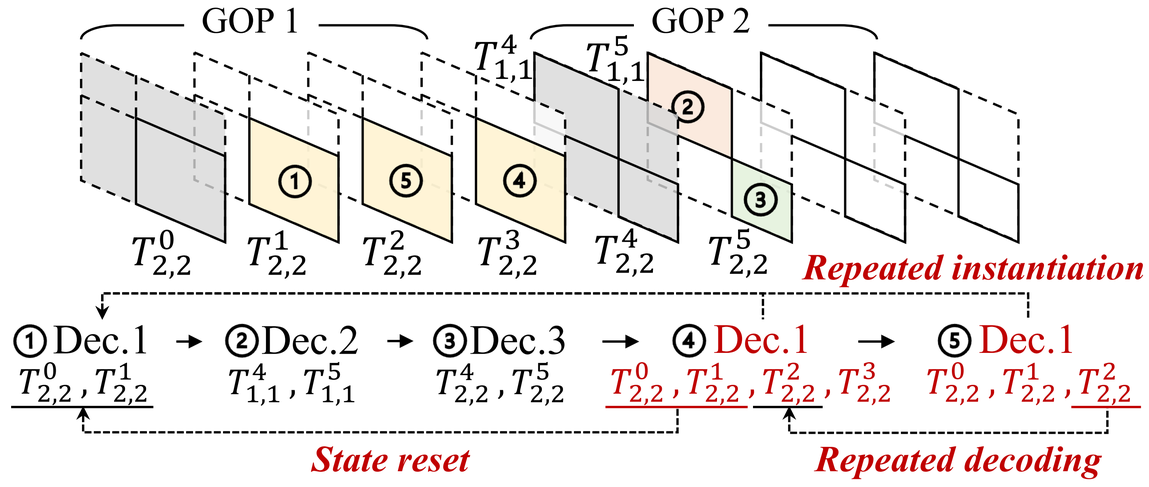
(a) Tile random access (b) Random access costs

- We model the decoding/DNN cost as a function of the number of pixels which can be used to quantify the pros and cons of different layouts.
- To speed up the search for the optimal layout, we also propose DP and greedy algorithms.

Query-Specific Acceleration Algorithms



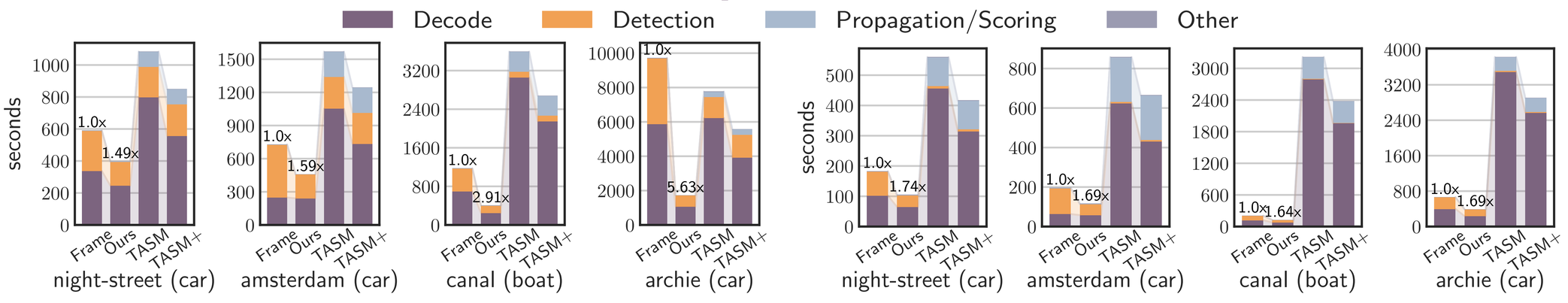
(a) Column-wise parallel (b) Row-wise parallel



Parallel decoding algorithm for limit queries to avoid access tiles that do not need to be processed.

Cache decoders, decoding states, and dependent frames to alleviate the resource waste in track queries.

Experiment



(a) Aggregation Queries (b) Limit Queries

Learn more

VLDB2024 GUANGZHOU