T4DT: Tensorizing Time for Learning Temporal 3D Visual Data

BMVC 2022 · Skoltech · ie · ETH zürich

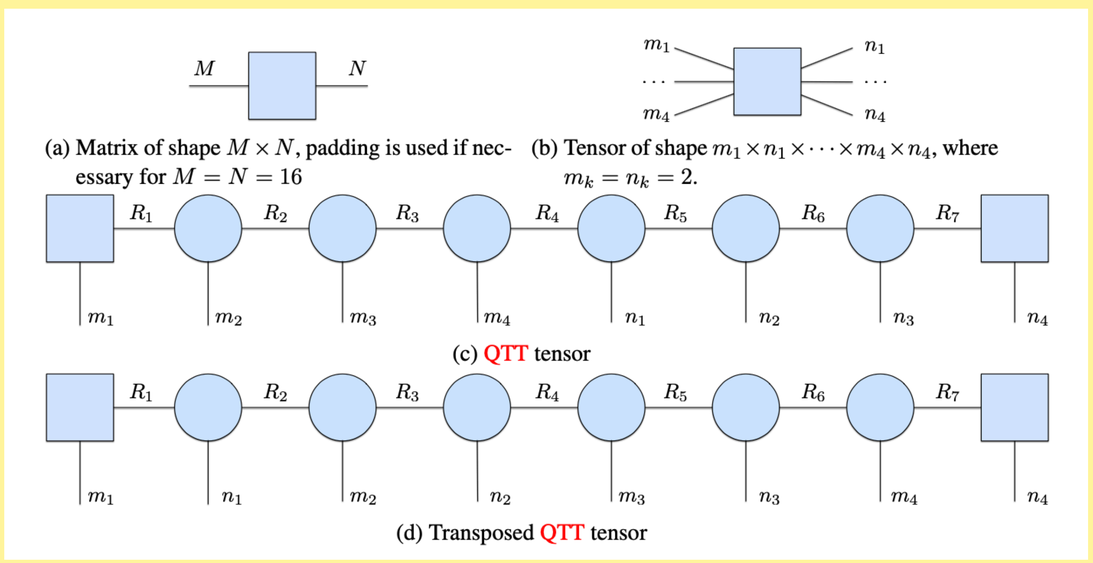
Usvyatsov M., Ballester-Rippol R., Bashaeva L.,
Schindler K., Ferrer G., Oseledets I.

Method

SDF and TSDF

$$\mathrm{SDF}(p) = \begin{cases} \mathrm{dist}(p, \partial\Omega) & \text{if } p \in \Omega, \\ -\mathrm{dist}(p, \partial\Omega) & \text{otherwise,} \end{cases} \qquad \mathrm{TSDF}(p) = \begin{cases} -\tau & \text{if } \mathrm{SDF}(p) \le -\tau, \\ \tau & \text{if } \mathrm{SDF}(p) \ge \tau, \\ \mathrm{SDF}(p) & \text{otherwise.} \end{cases}$$

Quantics Tensor Train



(a) Matrix of shape $M \times N$, padding is used if necessary for $M = N = 16$

(b) Tensor of shape $m_1 \times n_1 \times \cdots \times m_4 \times n_4$, where $m_k = n_k = 2$.

(c) QTT tensor

(d) Transposed QTT tensor

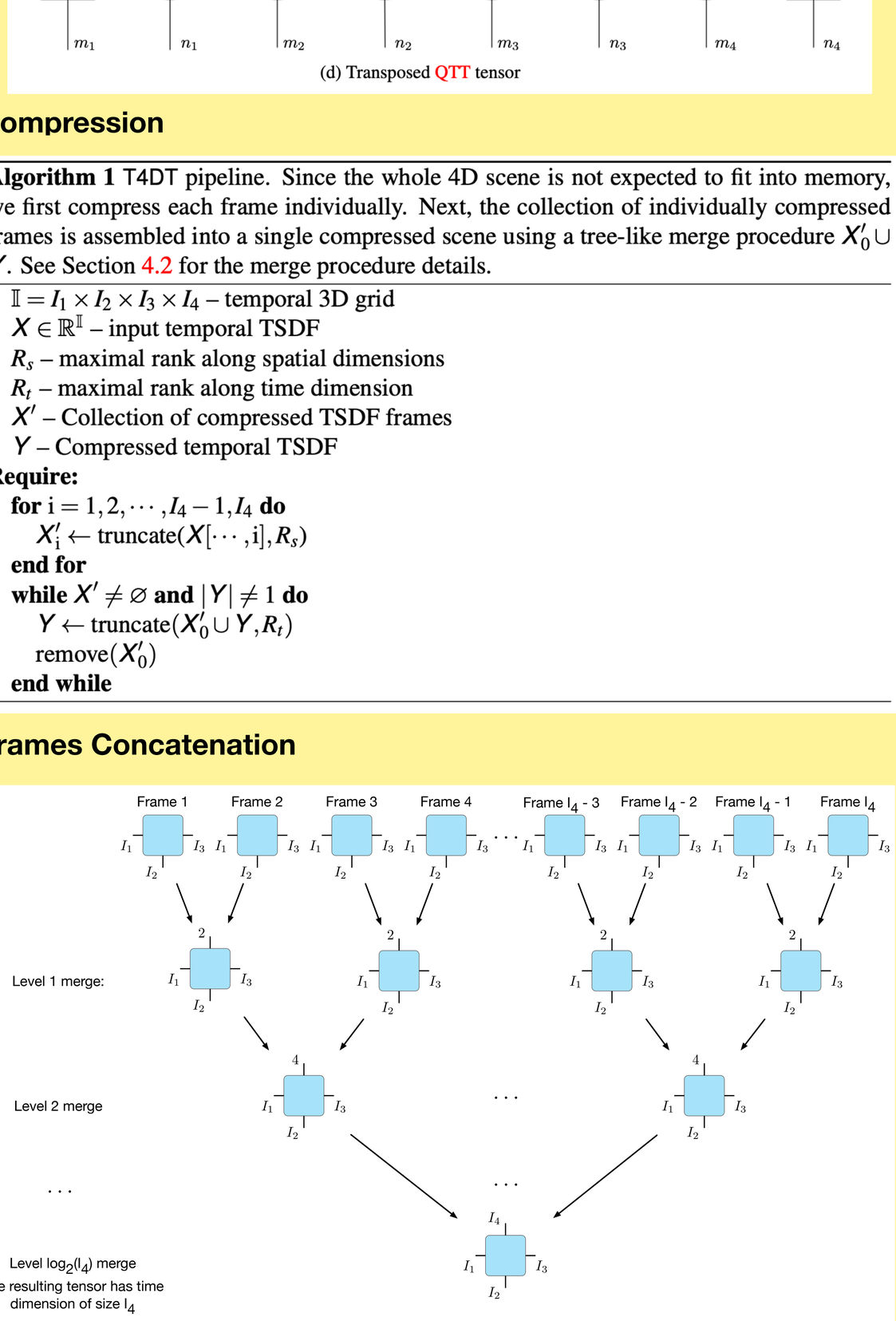
Compression

Algorithm 1 T4DT pipeline. Since the whole 4D scene is not expected to fit into memory, we first compress each frame individually. Next, the collection of individually compressed frames is assembled into a single compressed scene using a tree-like merge procedure $X'_0 \cup Y$. See Section 4.2 for the merge procedure details.

$\mathbb{I} = I_1 \times I_2 \times I_3 \times I_4$ – temporal 3D grid
$X \in \mathbb{R}^{\mathbb{I}}$ – input temporal TSDF
R_s – maximal rank along spatial dimensions
R_t – maximal rank along time dimension
X' – Collection of compressed TSDF frames
Y – Compressed temporal TSDF

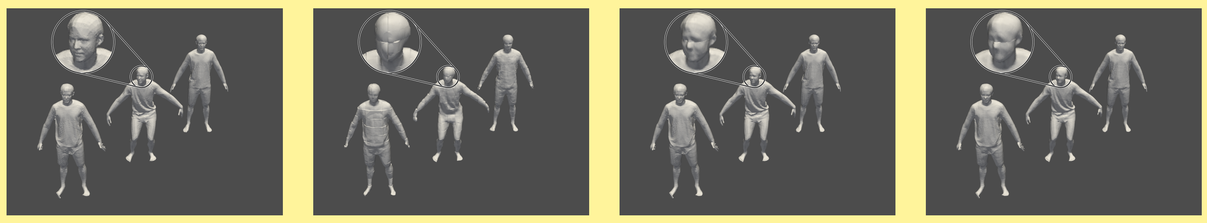
Require:
  for i $= 1, 2, \cdots, I_4 - 1, I_4$ **do**
    $X'_{\mathrm{i}} \leftarrow \mathrm{truncate}(X[\cdots, \mathrm{i}], R_s)$
  end for
  while $X' \neq \varnothing$ **and** $|Y| \neq 1$ **do**
    $Y \leftarrow \mathrm{truncate}(X'_0 \cup Y, R_t)$
    $\mathrm{remove}(X'_0)$
  end while

Frames Concatenation



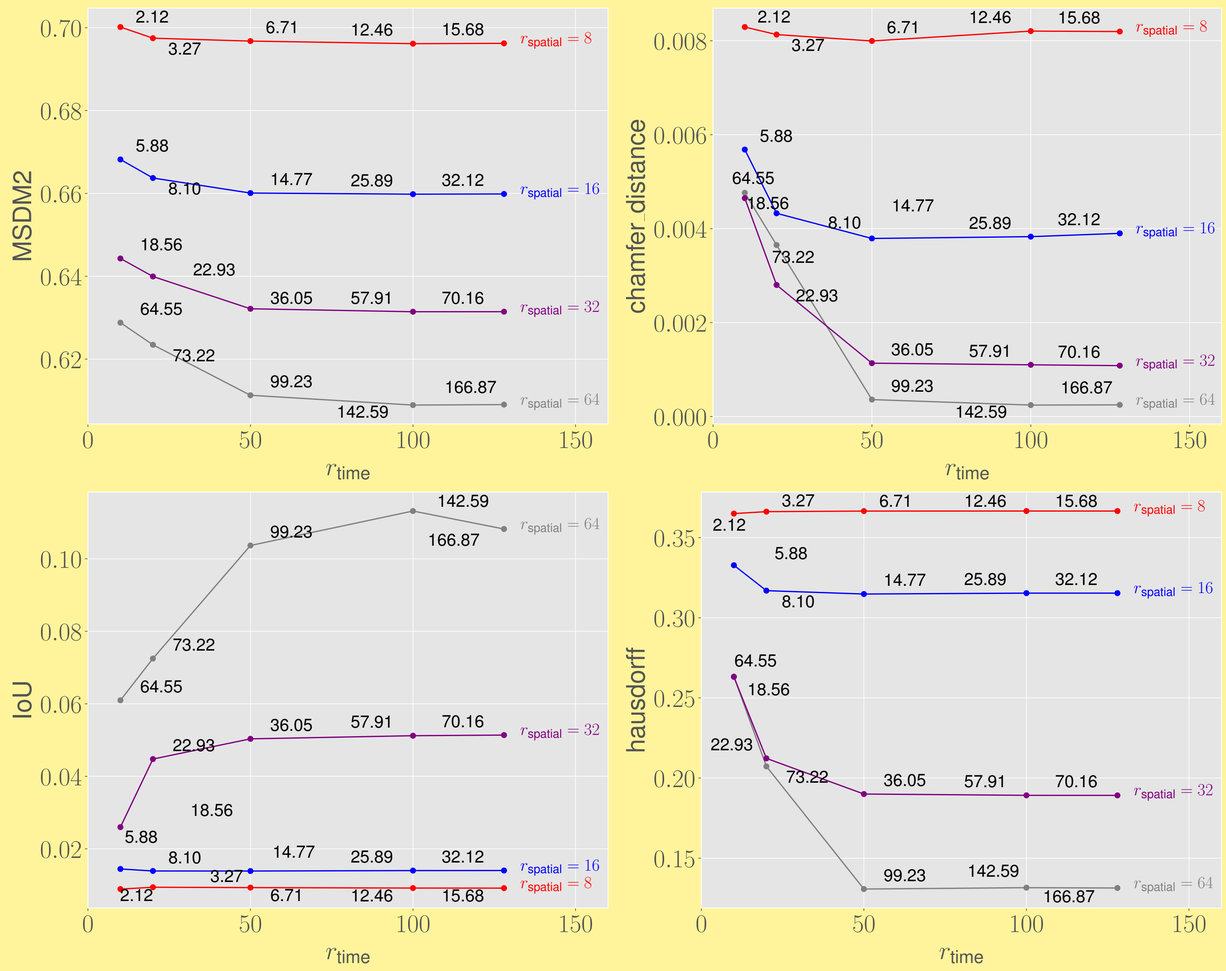
Results

CAPE Dataset

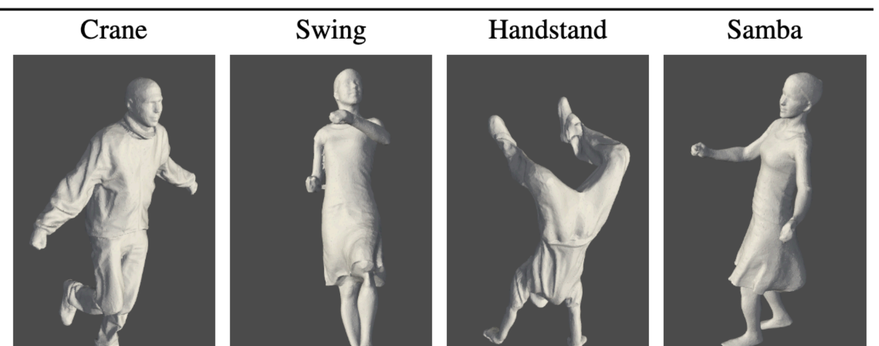


Original, 284GB | 1:12345, 24MB | 1:793, 386 Mb | 1:227, 1.24 GB

Ranks Influence



ANA Dataset



	Crane	Swing	Handstand	Samba
Resolution	$512^3 \times 174$	$512^3 \times 174$	$512^3 \times 174$	$512^3 \times 149$
Metric				
L2 ↓	2.67	2.07	2.45	1.71
Chamfer distance ↓	$5e^{-5}$	$4e^{-5}$	$5e^{-5}$	$4e^{-5}$
Hausdorff distance ↓	0.190	0.015	0.012	0.013
MSDM2 ↓	0.36	0.38	0.35	0.36
IoU ↑	0.41	0.40	0.64	0.54
Compression	1:954	1:1356	1:938	1:935
Size	549 MB	386 MB	558 MB	560 MB