# A Appendix

## A.1 Metrics

Throughout this work, we convert temporal 3D data from mesh format to tensor format for compression using sampled truncated SDF and from tensor format back to temporal 3D data in the form of meshes for quantitative tests using the marching cubes algorithm.

We selected a tensor-based l2 reconstruction metric for the comparison in tensorial form. We use the metrics below to evaluate the quality of the obtained mesh.

### A.1.1 Intersection over Union (IoU)

IoU is a commonly used metric for comparing 3D shapes [20]. For two given 3D volumes $A$ and $B$:

$$\mathrm{IoU}(A,B) = \frac{\|A \cap B\|}{\|A \cup B\|}. \tag{5}$$

3D mesh is converted to an occupancy grid, and the operation is easily performed with implementation from [4].

### A.1.2 Hausdorff distance

One-sided Hausdorff distance is computed as:

$$d_{\mathrm{h}}(A,B) = \max\{\min\{||a-b|| : a \in A\} : b \in B\}. \tag{6}$$

Note, that $d_{\mathrm{h}}$ is not symmetric. The symmetric version is defined as :

$$d_{\mathrm{H}}(A,B) = \max\{d_{\mathrm{h}}(A,B), d_{\mathrm{h}}(B,A)\}. \tag{7}$$

We use the implementation from [7].

### A.1.3 Chamfer distance

Symmetric Chamfer distance is defined as:

$$d_{\mathrm{CD}}(A,B) = \sum_{a \in A} \min_{b \in B} ||x-y||^2 + \sum_{b \in B} \min_{a \in A} ||x-y||^2. \tag{8}$$

Following [2], we sample 30'000 points from each mesh and compute sampled symmetric Chamfer distance. Finally, we use the implementation from [4].

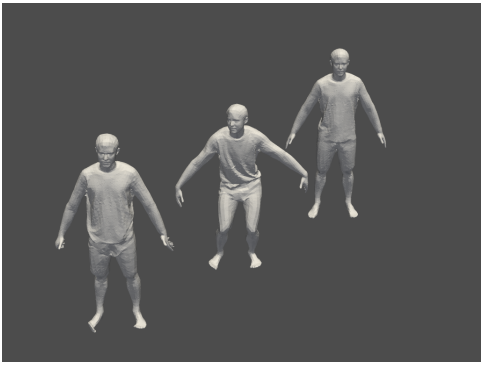
### A.1.4 Mesh Structural Distortion Measure (MSDM2)

Introduced in [11] MSDM2 metric is used to estimate the correlation with human visual perception. The metric assumes one mesh to be original and the second distorted. See Algorithm 3 for an algorithmic view of MSDM2 computation. Refer to [11, 12] for the more details on metric formulation. We use original implementation from [11] and self written pybind11 [8] interface to Python.

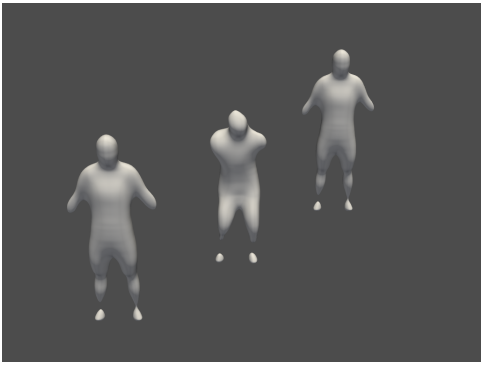
**Algorithm 3** MSDM2 high-level computation scheme

$M_o$ – original mesh, $M_d$ – distorted mesh, $\{v_i \in M\}$ – set of mesh vertices
**Require:**
  **for** $v_i \in M_o$ **do**
    Find $v_j = \min_{v_k \in M_d} \mathrm{dist}(v_i, v_k)$
  **end for**
  **for** $v_i \in M_o$ **do**
    calculate curvature of $v_i$
  **end for**
  **for** $v_i \in M_d$ **do**
    calculate curvature of $v_i$
  **end for**
  interpolate curvature per face for $M_o$ and $M_d$
  **for** $v_i \in M_d$ **do**
    calculate MSDM2 for $v_i$ based on curvature features and nearest neighbours $v_j$ from $M_o$
  **end for**
  Integrate global asymmetric MSDM2

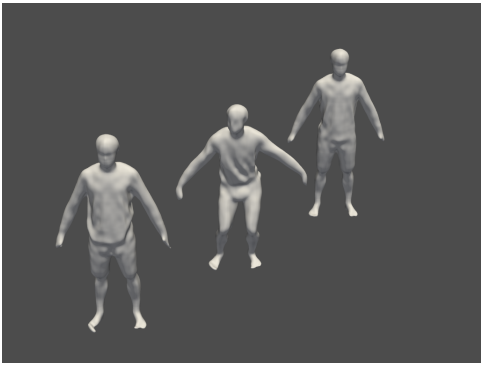
## A.2 Selected Visualizations



(a) Original, 284 GB



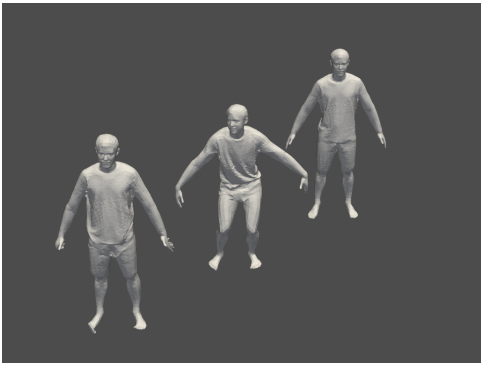
(b) Compression 1:769230, 0.2 MB



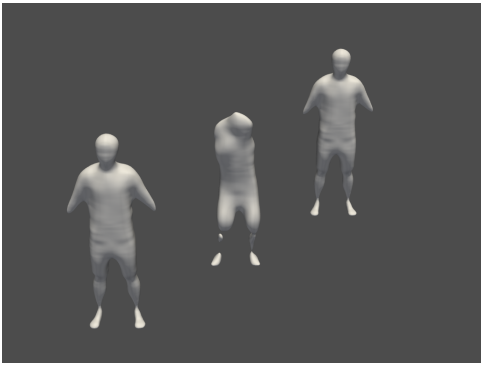
(c) Compression 1:1333, 33 MB

(d) Compression 1:116, 385 MB

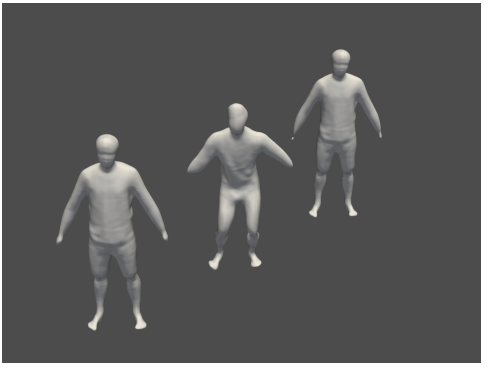
Figure 6: T4DT with different compression levels in TT-Tucker format for a *longshort-flying-eagle* scene of resolution $512^3$ with 284 time frames. Only frames 1, 142, and 284 are depicted.



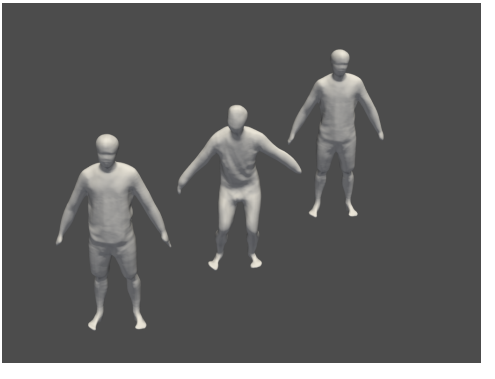
(a) Original, 284 GB



(b) compression 1:476190, 0.61 MB

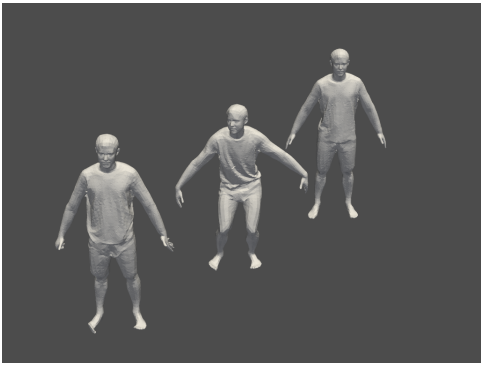


(c) compression 1:6250, 16 MB

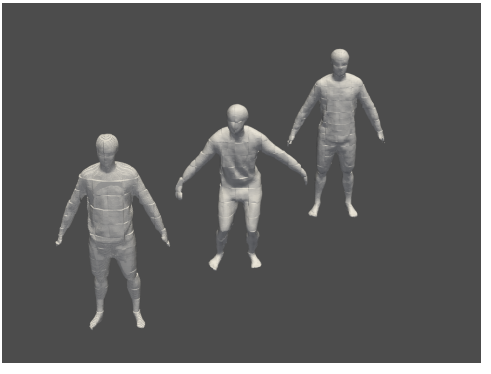


(d) compression 1:2000, 48.5 MB

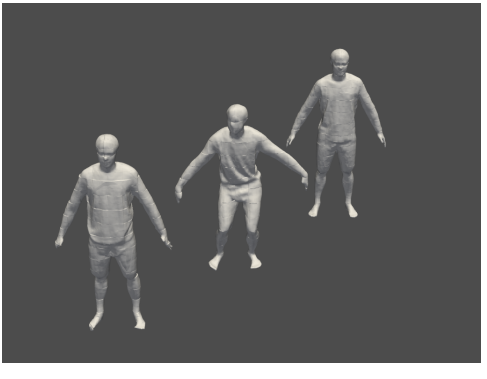
Figure 7: T4DT with different compression levels in TT format for a *longshort-flying-eagle* scene of resolution $512^3$ with 284 time frames. Only frames 1, 142, and 284 are depicted. The compression ratio is different from the actual memory consumption due to the padding of the time dimension to 512.



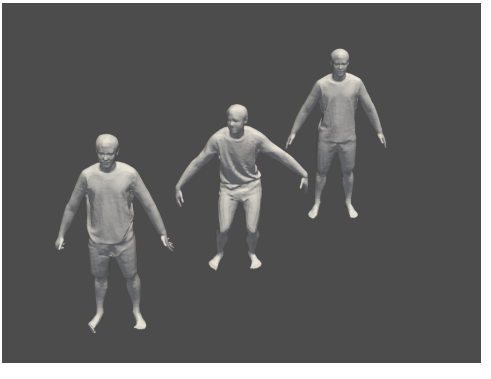
(a) Original, 284 GB



(b) compression 1:6666, 0.08 MB

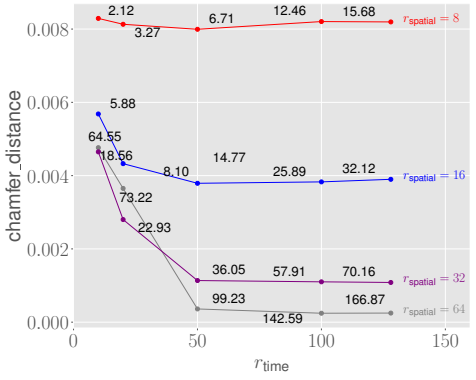


(c) compression 1:1587, 332 MB

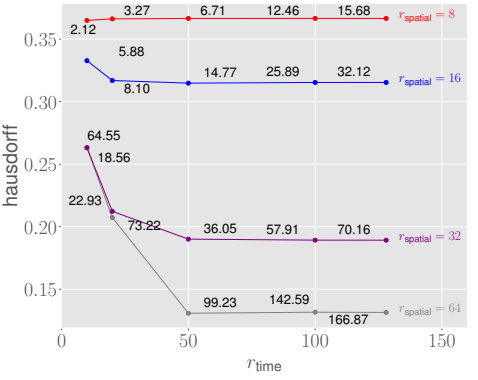


(d) compression 1:244, 1.4 GB

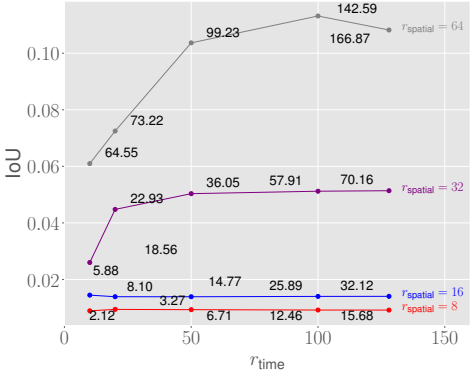
Figure 8: Different levels of compression in QTT format for a *longshort-flying-eagle* scene of resolution $512^3$ with 284 time frames. Only frames 1, 142, and 284 are depicted. The compression ratio is different from the actual memory consumption due to the padding of the time dimension to 512.



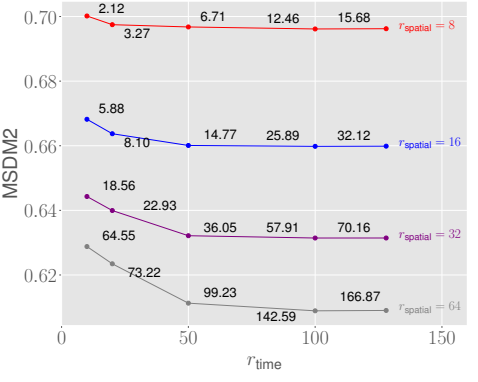

(a) Chamfer distance




(b) Hausdorff distance

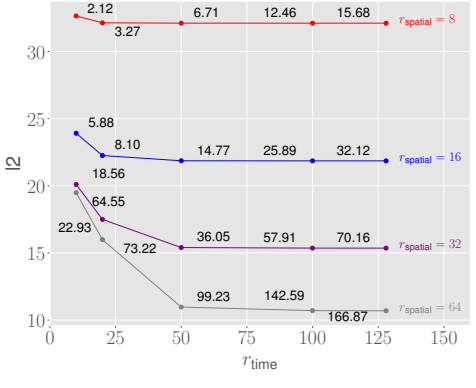



(c) IoU

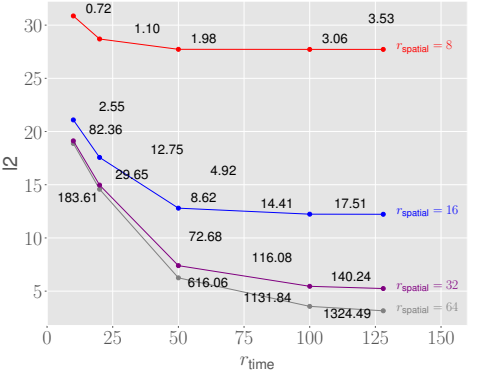



(d) MSDM2

Figure 9: Ranks vs. performance for TT compressed *longshort-flying-eagle* scene of resolution $512^3$ with 284 time frames. Metrics are averaged between frames 1, 142, and 284. Each data point is annotated with the corresponding compression ratio scaled with 1e6.




(a) TT




(b) TT Tucker

Figure 10: Ranks vs. $L_2$ for *longshort-flying-eagle* scene of resolution $512^3$ with 284 time frames. $L_2$ is averaged between frames 1, 142, and 284. Each data point is annotated with the corresponding compression ratio scaled with 1e6.