EUROGRAPHICS 2022 / R. Chaine and M. H. Kim
(Guest Editors)

Volume 41 (2022), Number 2

Supplementary Materials
Learning Spectral Unions of Partial Deformable 3D Shapes

Luca Moschella[1], Simone Melzi[1,2], Luca Cosmo[3], Filippo Maggioli[1], Or Litany[4], Maks Ovsjanikov[5], Leonidas Guibas[6], Emanuele Rodolà[1]

[1] Sapienza University of Rome [2] University of Milano-Bicocca [3] Ca' Foscari University of Venice [4] NVIDIA 5 LIX, École Polytechnique, CNRS [6] Stanford University

We report additional results that were not included in the main manuscript. Further, we describe with more details the architectures involved in the proposed method.

1. Additional results

In this section, we collect additional results for the experiments and applications described in the main manuscript.

1.0.0.1. Additional results on geometry reconstruction In Fig. 1, we report additional examples of shape-from-spectrum recovery. These qualitative results confirm that we outperform [MRC*19] when applied to partial shapes.

1.0.0.2. Associativity with more than 3 partial shapes In Fig. 2 we show an example of iterative spectral union, with four different partial shapes. Deeper iterative unions are more difficult since the prediction error in each step is amplified by the subsequent steps.

1.0.0.3. Additional results on the *horse* class We report additional results on the horse class. In these experiments, the spectral union operator is pre-trained on humans and fine-tuned to horses from the TOSCA dataset. Then, a region localization model is trained specifically for horses, slightly modified to account for the different number of vertices in the template.

In Figure 3 we report qualitative results of region localization on horses; in Figure 4 we show associativity examples; in Figure 5 we present qualitative results on horses with different triangulation, vertex density and style with respect to the horses used in the training phase; in Figure 6 we show that our method is able to generalize to non-isometric but similar enough deformations.

1.0.0.4. Additional results on *Point clouds* To show the flexibility of our approach we consider a further class of shapes composed by airplanes from [CFG*15] represented as point clouds. We report in Figure 7 some qualitative results on the region localization task with this class. We also report additional results on headphones [MZC*19] in Figure 8. These results show that our model generalizes to different source geometries, as long as the class shape does not change.

2. Architecture

In this section, we describe in detail the proposed neural architecture. Note that since surface area directly affects the magnitude of the eigenvalues, at test time the shapes are normalized to have the same area of the shapes seen at training time.

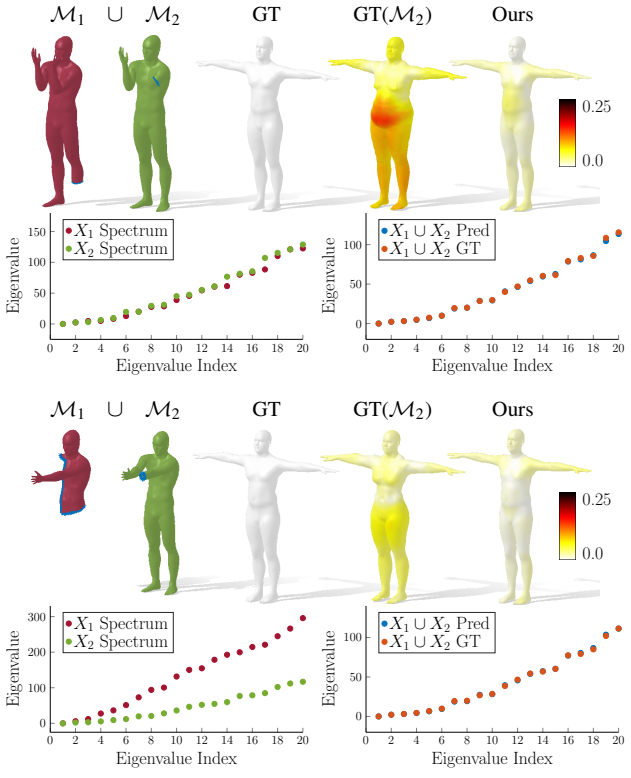

Figure 1: Comparison of the reconstruction obtained by running the state-of-the-art method of [MRC*19] on the green shape, yielding the fourth shape, and the reconstruction obtained from our predicted full spectrum, yielding the last shape.

2.1. Spectral union model

In Fig. 9 we show the detailed architecture of the spectral union model.

2.1.0.1. Hyperparameters The dimensionality of each embedding is 32. $\boldsymbol{T_A}$ has 8 heads, 6 layers, the dimensionality of the internal feed-forward layer is 64 and the dropout is 0.1. $\boldsymbol{T_B}$ has 8 heads, 3 layers, the dimensionality of the feed-forward is 32 and the dropout is 0.1. Thus, $\boldsymbol{\rho}$ reduces the embedding dimensionality from 32 to 1.

2.1.0.2. Training The model is trained until convergence. The training randomly augments online each input independently. The batch size is 32. The optimizer used is Adam with learning rate of $2e{-}4$ and weight decay $1e{-}5$. The learning rate changes according to cosine annealing with warm restarts scheduler and it restarts every 10 epochs, doubling the number of epochs between restarts at each restart.

2.2. Region Localization model

In Fig. 10 we show the detailed architecture of the region localization model for humans.

2.2.0.1. Hyperparameters The dense layers increase the dimensionality of the input sequence from 20 to 6890, for humans, i.e. the number of vertices in the fixed template. In particular, the layers apply the following transformations $20 \to 1300 \to 2600 \to 3900 \to 5200 \to 6890$. The dropout is always set to $p = 0.5$.

2.2.0.2. Training The model is trained to localize the region from both the predicted union eigenvalues and all the ground-truth eigenvalues, to which we add random noise. The model is early stopped, monitoring the IoU metric on a validation set. The batch size is 32. The optimizer used is Adam with learning rate of $5e{-}5$ and weight decay $1e{-}6$. The scheduler adopted is again the cosine annealing with warm restarts, with the same hyperparameters.

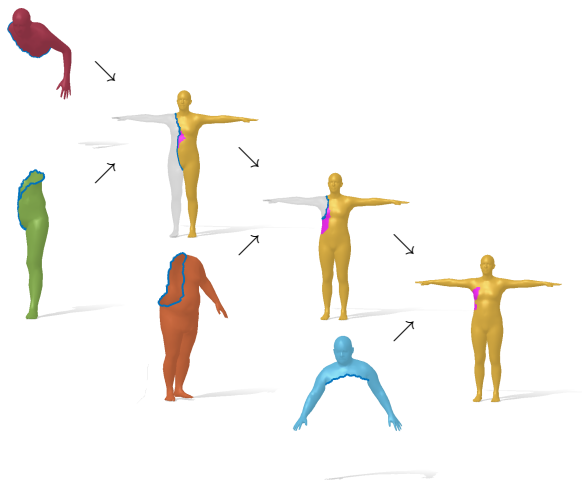

Figure 2: Example of associativity with 4 partial shapes.

2.3. Data processing for point clouds

The aereoplanes from [CFG*15] and headphones from [MZC*19] are point clouds with *semantic* segmentation.

We performed some data processing to: (1) extract shapes with only given segments (e.g., discarding earphones or strange headphones), (2) extract random partialities from each shape and (3) find the segment-level matching between each shape and a fixed template for the region localization task. We did this by defining the graph of the segments for each shape, then searching for subgraphs with determined properties for (2) and solving the graph isomorphism against the template for (3).

We obtained 75 headphones and 964 aereoplanes for the training set, we extract random pairs of partialities from each shape.

References

[CFG*15] CHANG, ANGEL X., FUNKHOUSER, THOMAS, GUIBAS, LEONIDAS, et al. *ShapeNet: An Information-Rich 3D Model Repository*. Tech. rep. arXiv:1512.03012 [cs.GR]. Stanford University — Princeton University — Toyota Technological Institute at Chicago, 2015 1, 2.

[MRC*19] MARIN, RICCARDO, RAMPINI, ARIANNA, CASTELLANI, UMBERTO, et al. "Instant recovery of shape from spectrum via latent space connections". *2019 International Conference on 3D Vision (3DV)*. IEEE. 2019, 37–46 1.

[MZC*19] MO, KAICHUN, ZHU, SHILIN, CHANG, ANGEL X., et al. "PartNet: A Large-Scale Benchmark for Fine-Grained and Hierarchical Part-Level 3D Object Understanding". *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2019 1, 2.

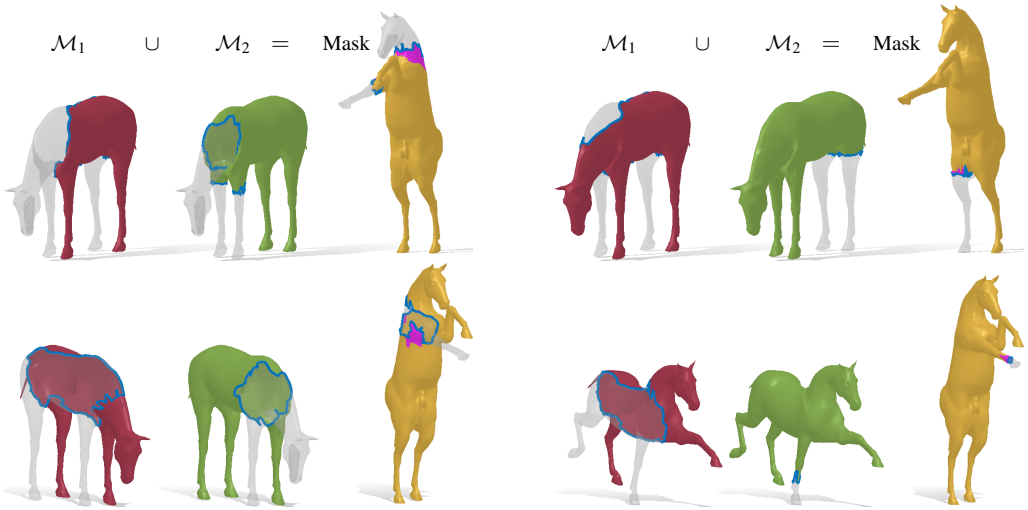

Figure 3: Region localization on horses. We predict the indicator function that describes the union of two partial shapes, given their eigenvalues.

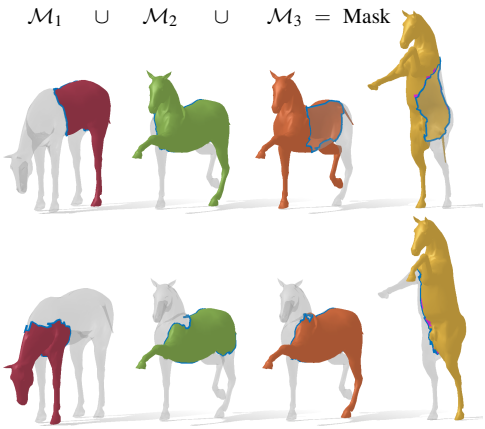

Figure 4: Example of associativity.

Figure 5: Region localization on different horses. The partial shapes have a different triangulation, vertex density and style.

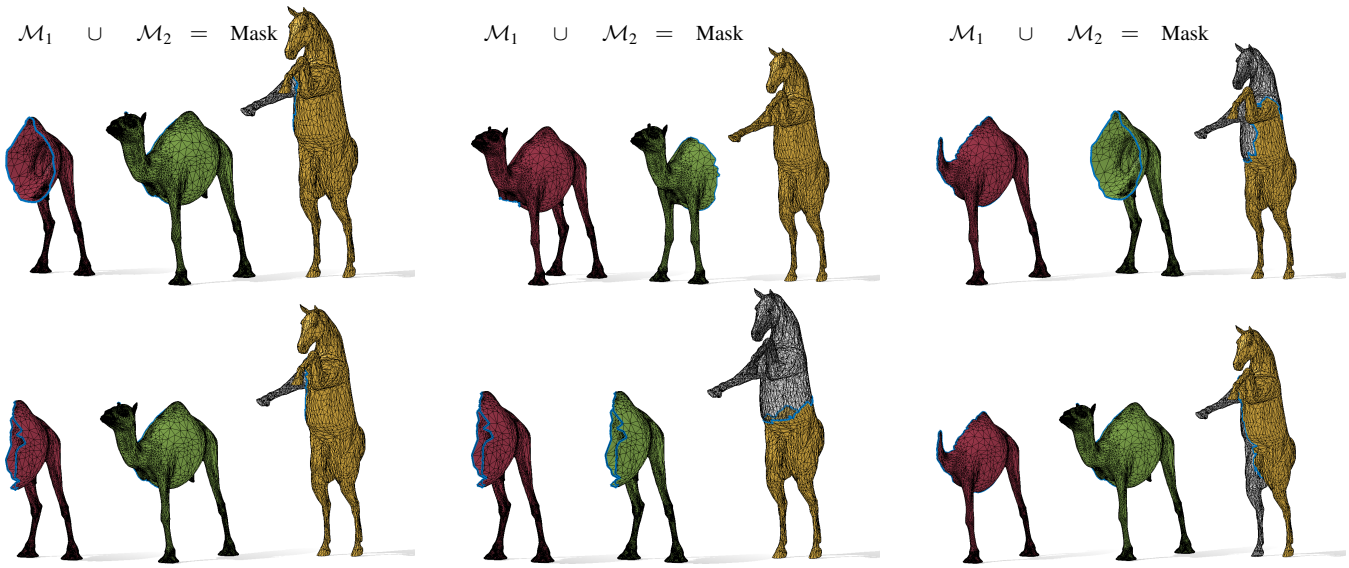

Figure 6: Region localization on a camel, the model is trained on horses.

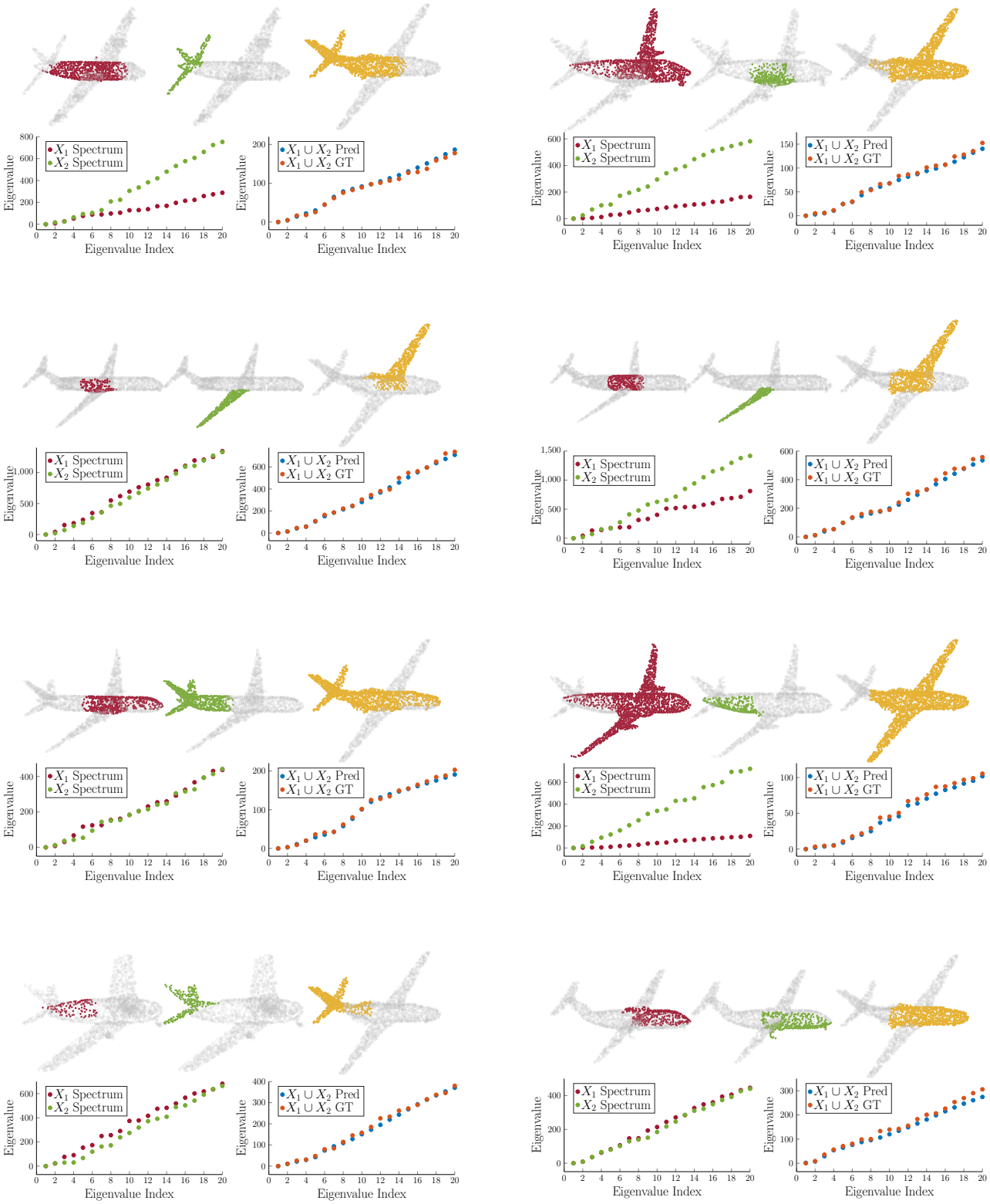

Figure 7: Region localization on aereoplanes. The model is trained and tested on point clouds.

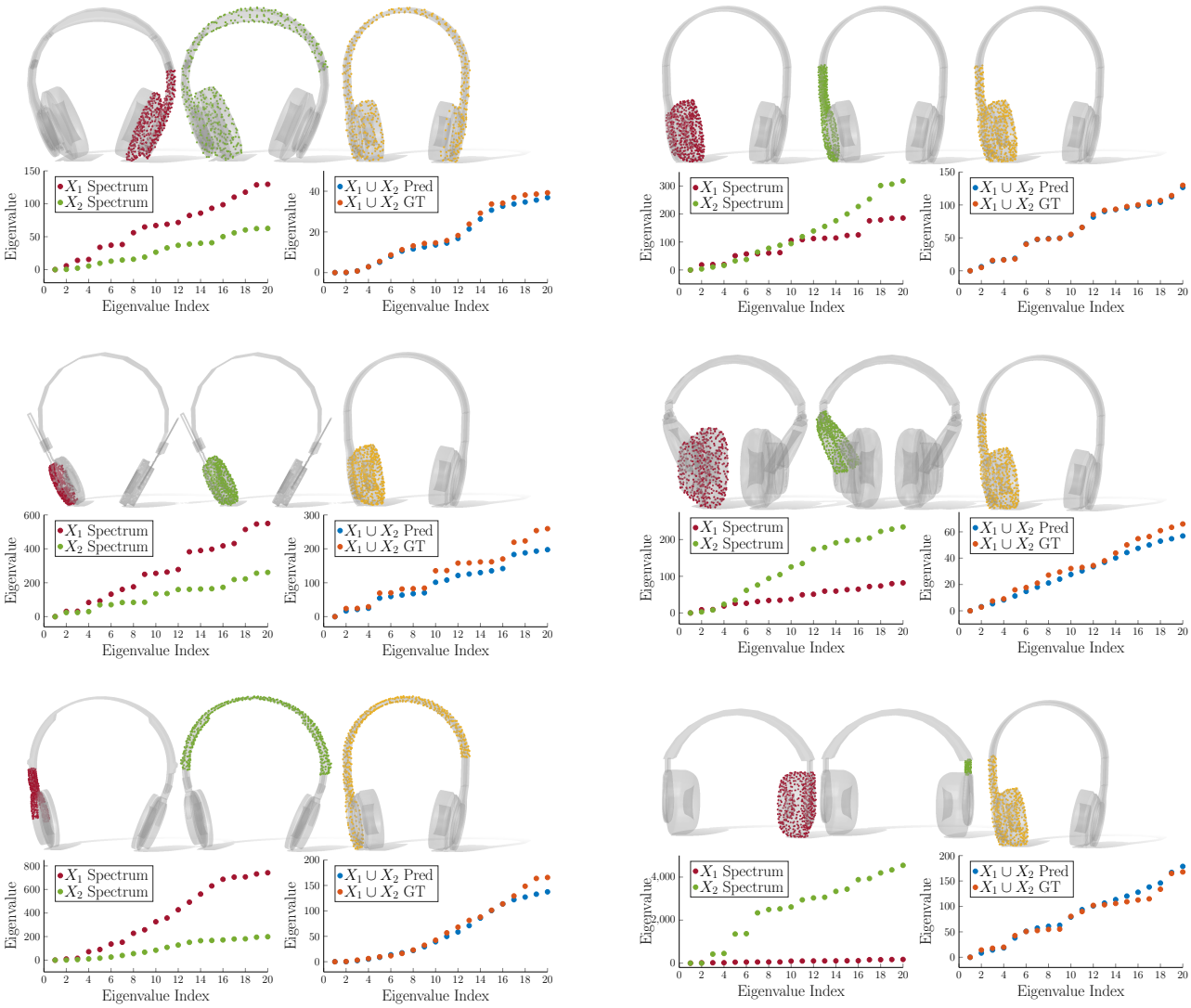

Figure 8: Region localization on headphones, trained and tested on point clouds. In the examples on the right column, despite significant changes in the geometry of the partialities, the model localizes the same correct region.

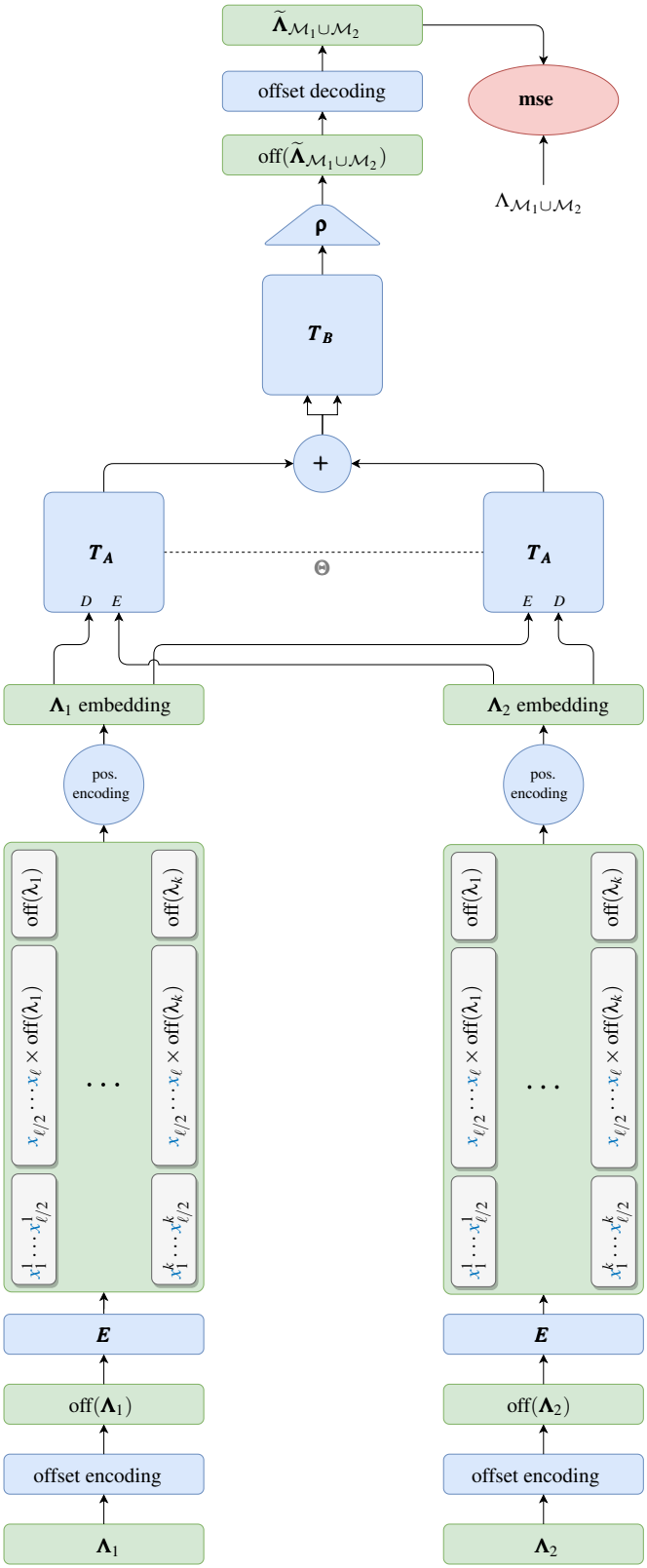

Figure 9: Detailed architecture of the spectral union operator.

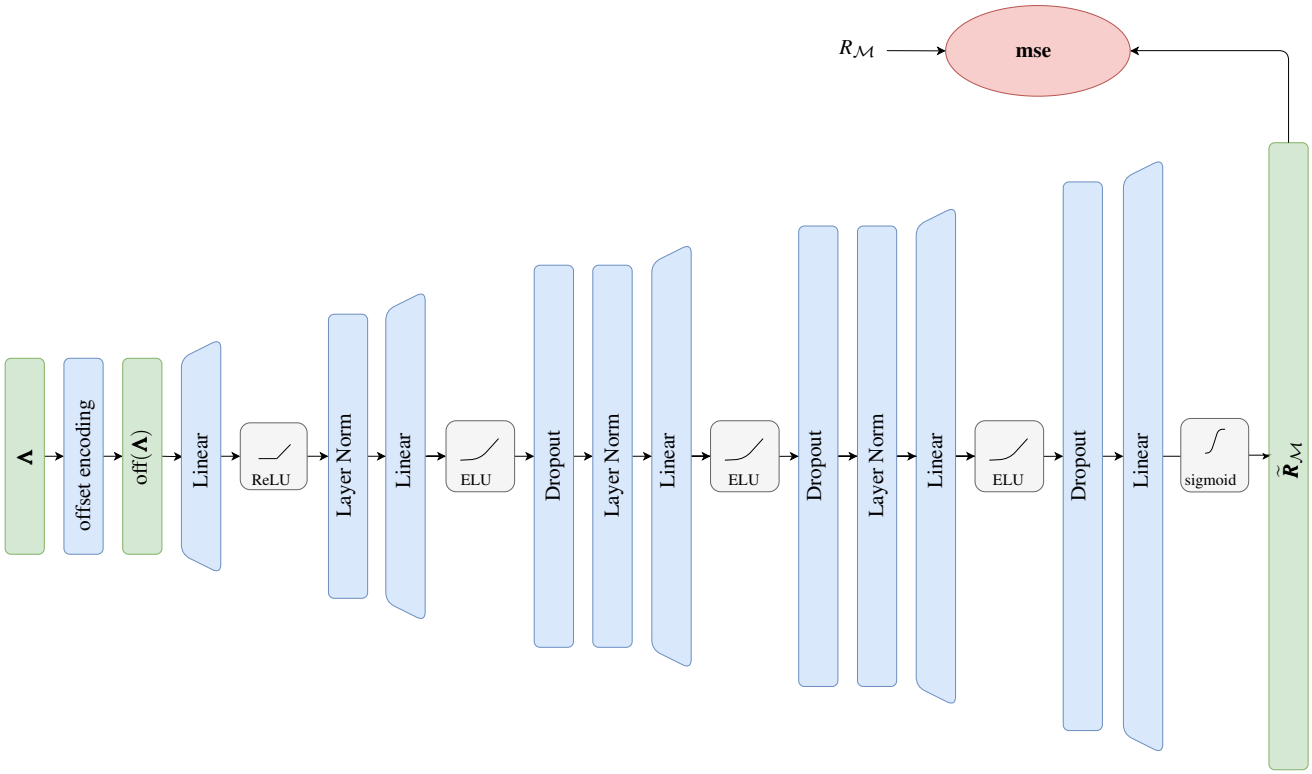

Figure 10: Detailed architecture of the region localization MLP for humans.