DEGAS: Detailed Expressions on Full-Body Gaussian Avatars

Supplementary Material

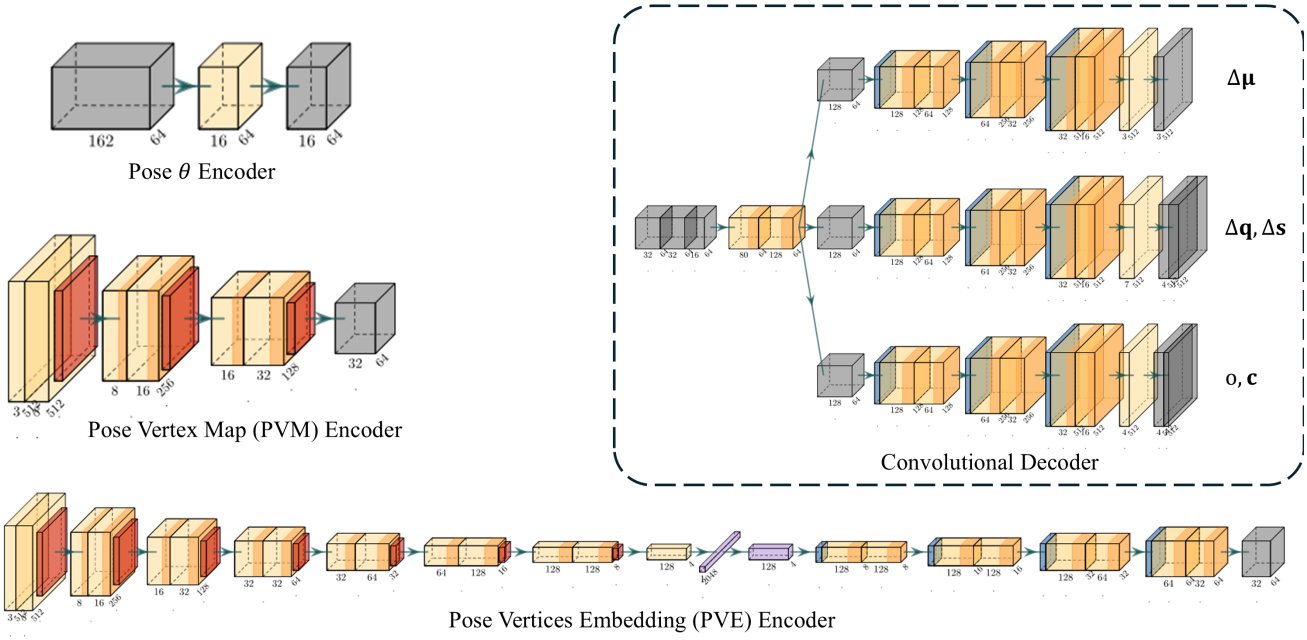

Figure A1. **Network structures.** We show the network structures of the three encoder branches and the convolutional decoder.

Figure A2. **Ablation on sparse views.** Our method can be trained with sparse views, i.e., 3 views, 6 views, and 12 views.

A. Implementation Details

There are three encoder branches and one convolutional decoder defined in this paper. We show the network structures in Figure A1. As can be seen that though both the encoder PVM and PVE work on the position map, PVE is much deeper than PVM. We find all three encoder branches help the modeling. The convolutional decoder has three groups of CNNs, i.e., one for position, one for rotation and scaling, and the last one for opacity and color.

B. Sparse Views

We show in Figure A2 that our method can be trained with sparse views.

C. Ethics

We captured six human subjects in our proposed DREAMS-Avatar dataset. All subjects have given written consent for the captured footage in this dataset. We make the data publicly available for research purposes.

Our method could be extended to synthesize communication media contents of real people. We do not condone using our work to generate fake images or videos of any person with the intent of spreading misinformation or tarnishing their reputation.