

ESAA

Model

 $f_{a,n}$

K,V

 $\mu_s + \Delta \mu_i$

 $r_c + \Delta r_i$

 $s_c + \Delta s_i$

 $SH_c + \Delta SH_i$ $\alpha_c + \Delta \alpha_i$

Efficient Gaussian

Deformation Decoder

Splatting

Step 2: Efficient Audio-driven Gaussian Deformation

Audio Feature Extractor

Face Feature

Extractors