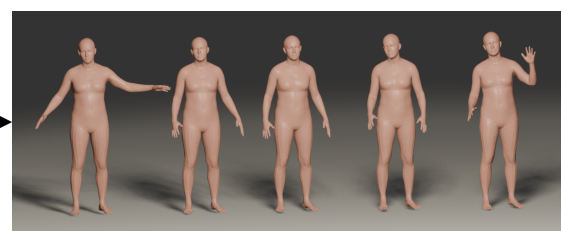
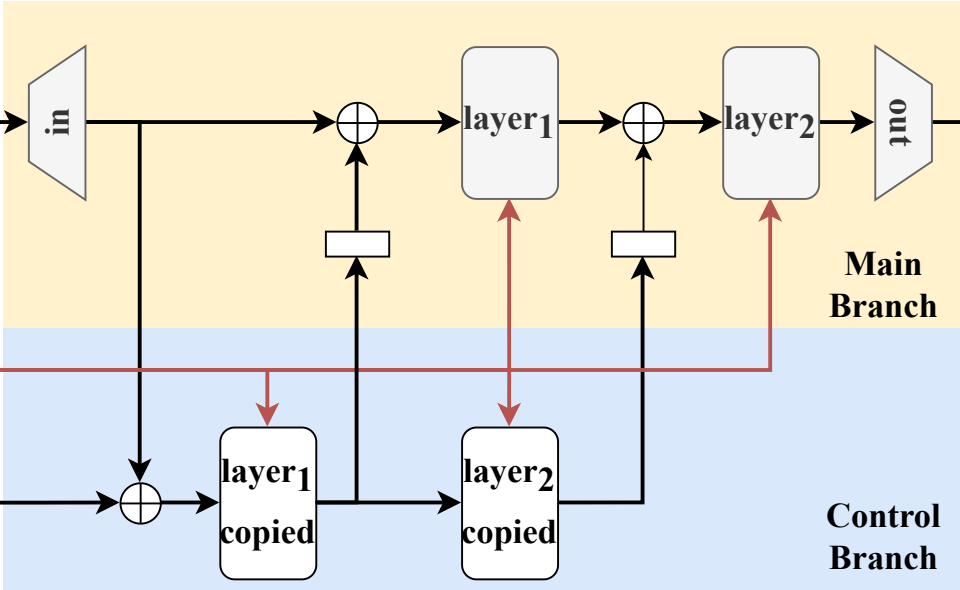


noise



Main
Branch

Control
Branch



bridge module



element-wise addition

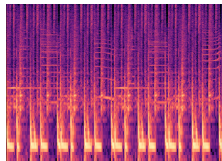


fully connected layer



audio condition

a man is giving a speech
while waving hands



condition
encoders

layer1
copied

layer2
copied