

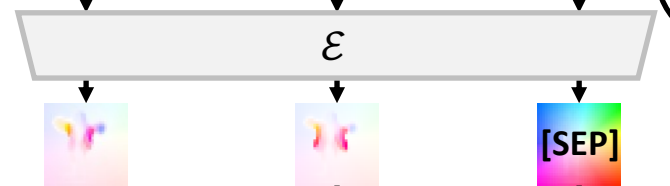
Condition  
Frames ( $x_c$ )



Predicted  
( $x_p$ ) Frames



*Motion Autoencoder  
Compression*



*Motion Autoencoder  
Reconstruction*



*Distribution  
Extrapolation  
Diffusion  
Model*

