



CoMusion: Towards Consistent Stochastic Human Motion Prediction via Motion Diffusion



Jiarui Sun and Girish Chowdhary



Motivation

- **Goal:** stochastic human motion prediction (HMP)

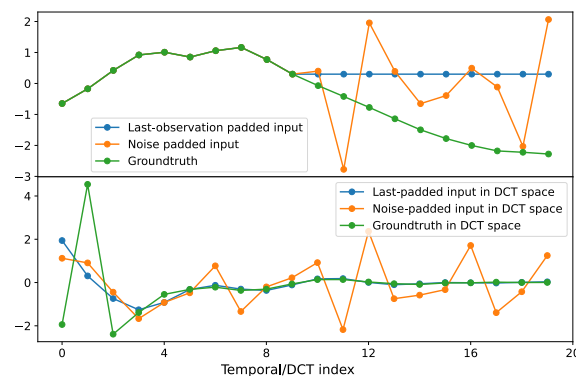


Prior work limitations

- Inconsistent, even unrealistic behavior
- Multi-stage training process

Insight

- GCN + DCT design from deterministic HMP

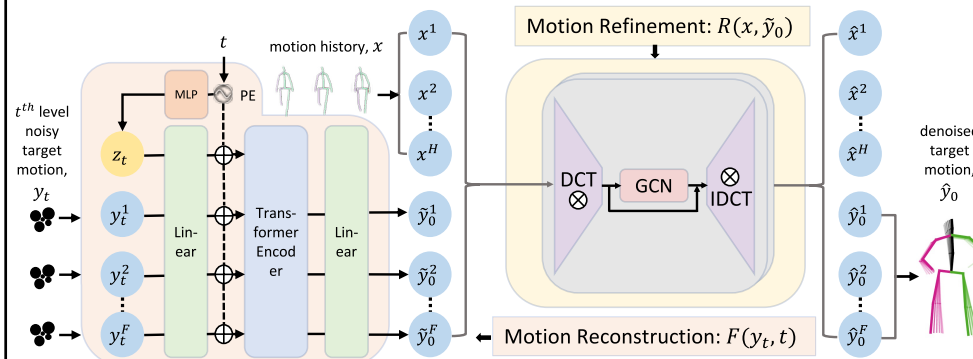


- GCN for spatial learning
- DCT for temporal residual learning

Our approach

- Guess a smooth future first
- Then, GCN+DCT for refinement
- Single-stage diffusion model

Method



A Transformer-based **motion reconstruction module** first generates an initial guess from noised future motion. Then, a GCN-based **motion refinement module** refines the prediction in the DCT space

- **Spatial-temporal** motion matching loss

- Spatial: joint-based re-weighting
- Temporal: history + future reconstruction
- Best-of-many: implicit relaxation

$$\mathcal{L}_{\text{rec}} = \frac{1}{J} \sum_{j=1}^J (\gamma \cdot \|x^j - \hat{x}^j\|_1 + \lambda^j \|y_0^j - \hat{y}_0^j\|_1)$$

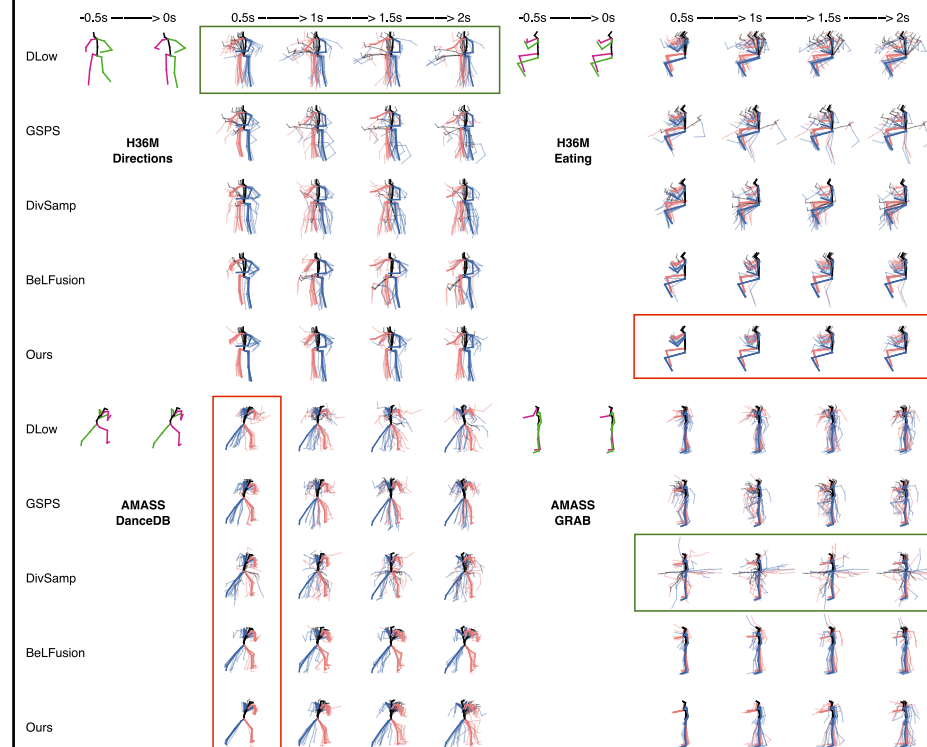
$$\mathcal{L}_{\text{final}} = \min_k (\mathcal{L}_{\text{rec}}^k, y_0, x)$$

Quantitative Results

		Diversity		Accuracy				Consistency	
Type	Method	One-Stage	APD ↑	APDE ↓	ADE ↓	FDE ↓	MMADE ↓	MMFDE ↓	CMD ↓
VAE-based	TPK	✓	9.283	2.265	0.656	0.675	0.658	0.674	17.127
	DLow	✗	13.170	4.243	0.590	0.612	0.618	0.617	15.185
	GSPS	✗	12.465	4.678	0.563	0.613	0.609	0.633	18.404
	DivSamp	✗	24.724	15.837	0.564	0.647	0.623	0.667	50.239
DM-based	HumanMAC	✓	9.321	-	0.511	0.554	0.593	0.591	-
	BeLFusion	✗	9.376	1.977	0.513	0.560	0.569	0.585	16.995
	Ours	✓	10.848	2.328	0.494	0.547	0.469	0.466	9.636

- Properly **diverse** predictions with great **accuracy**
- More importantly, **consistent** in terms of behavior (mode, speed)

Qualitative Results



- **Prior work:** behaviorally inconsistent, physically unrealistic
- **Ours:** diverse predictions adapted to context, gradual change of variation

Conclusions

- **CoMusion:** An end-to-end diffusion-based stochastic human motion prediction framework that benefits from the GCN-DCT design used in deterministic works
- **Website** is available at: <https://jsun57.github.io/CoMusion/>