Graduate School of Science and Technology Master's Thesis Abstract

Laboratory name (Supervisor)	Cybernetics and Reality Engineering (Kiyoshi Kiyokawa (Professor))		
Student ID	2411426	Submission date	2026 / 7 / 22
Name	SIY JOSHUA SAMUEL		
Thesis title	Expressive Human Motion Generation via Kinematic Effort Conditioning		

Abstract

Human motion diffusion models can synthesize action sequences from text, but controlling motion intensity and enabling semantically guided transitions between fixed anchor frames remain challenging. Existing approaches rely on effort-related adverbs for intensity control — which are ambiguous and fail to capture quantitative aspects such as pacing — while motion inbetweening methods either require keyframe supervision or produce geometric transitions without semantic control, resulting in flat and monotonous dynamics.

We propose a unified framework centered on Effort Metric Attention (EMA), a cross-attention module that conditions a latent diffusion model on numerical effort signals inspired by Laban Movement Analysis (LMA). Focusing on the Time and Weight effort factors, we approximate these using two kinematic metrics: peak joint positional change for pacing and collective joint positional change for motion amount. EMA enables fine-grained, region-wise effort control without costly post-hoc optimization.

We further extend EMA to motion inbetweening through inference-time anchor replacement, demonstrating that the model's temporally compressed latent representation naturally produces geometrically consistent transitions without retraining. We introduce two evaluation tasks — metric-to-motion consistency and body-part-level effort modulation — to assess numerical fidelity and localized control. Experiments and a user study show near-monotonic alignment between specified effort levels, generated motion dynamics, and established LMA descriptors. Inbetweening experiments further demonstrate improved boundary smoothness and motion dynamics over hard masking baselines, with architectural analysis confirming that latent diffusion substantially reduces boundary discontinuities compared with frame-space diffusion.

These results indicate that kinematic effort is an effective and interpretable conditioning primitive for both expressive motion generation and semantically controlled inbetweening.