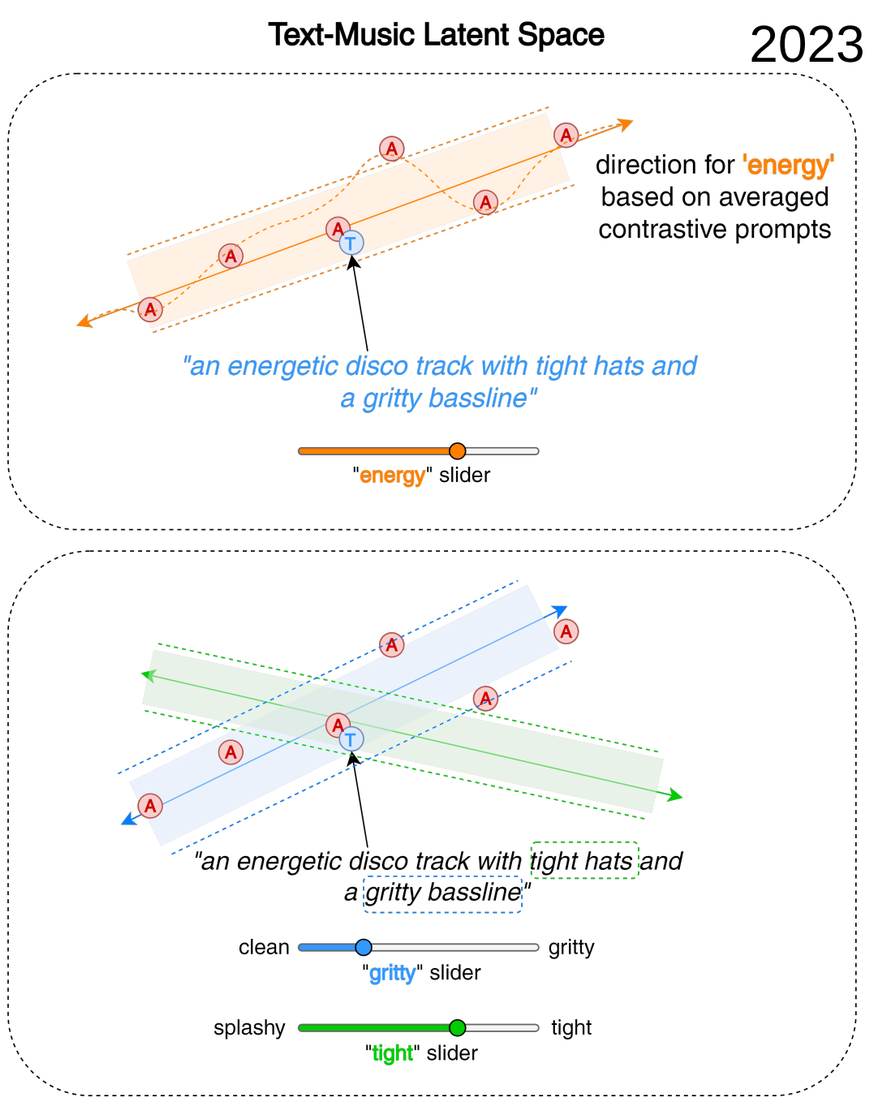
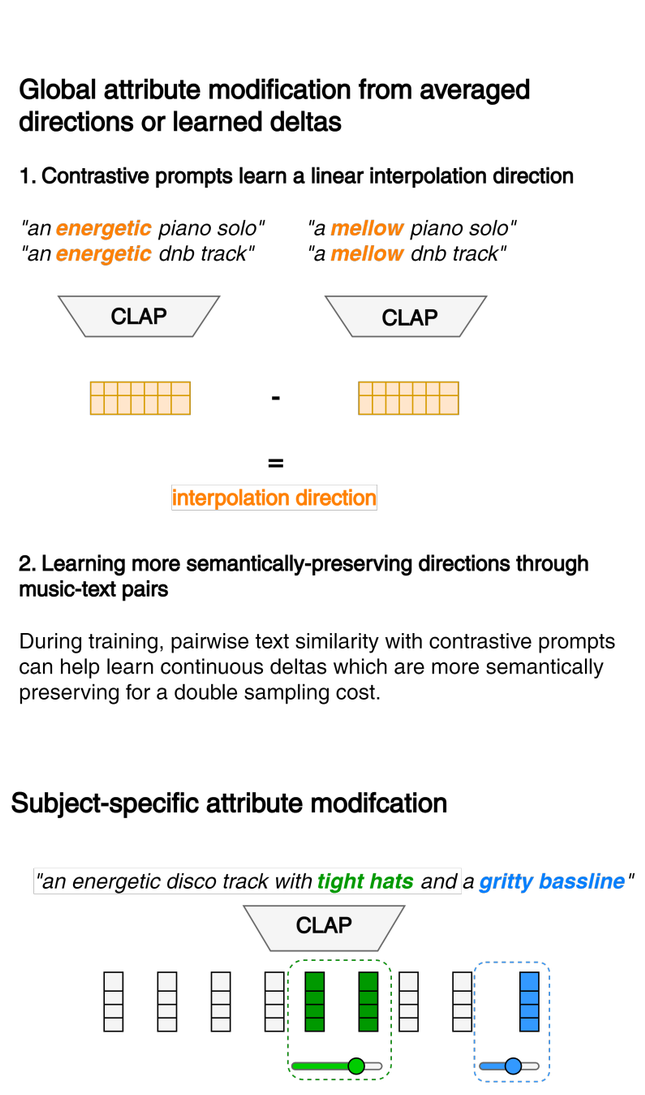
AIM Summer School 2024

Identifying semantic directions in contrastive text-music spaces for continuous and subject-specific attribute control for music retrieval

Julien Guinot





Finding

Some deterministic audio transformations can be easily associated with musical semantic information (time stretching with tempo, pitch shifting with key). We found that it is possible to learn semantic transformations in the latent space by controlling these factors and learning to reconstruct the modified embedding, which benefits retrieval controllability.

Question

Transformations for more nuanced concepts not directly related to any audio transformation are hard to capture in the audio space. We ask whether discovering semantic 'slider' transformations from contrasting prompts in existing text-music latent spaces is a viable method for modeling transformations relating to these concepts.

Supervisor(s): George Fazekas, Elio Quinton