HAGU: AI-Generated Harmonically Rich Classical Guitar Pieces

Akbudak, Burak Sina (School: Izmir Fen Lisesi)

Current symbolic music generation models and music modeling techniques primarily rely on language models designed for language modeling. Most of these models treat musical features (e.g. pitch, velocity, tempo…) uniformly, hindering the generation of harmonically pleasing musical phrases. To address this issue, we introduce novel contributions, starting with an efficient 2D tokenization technique called RelTok (Relative Tokenization). Through this method, we denote time relative to the preceding note's timing. Moreover, we explore the integration of the future n-gram prediction and the n-stream self-attention mechanism into the music generation task. We train a decoder-only transformer architecture, designed to accommodate n-gram prediction while utilizing 2D tokenization, on a customized dataset comprising classical guitar compositions. Finally, addressing the oversight in assessing the playability of synthesized compositions in music synthesis evaluations, we propose the "playability rate" metric. Tailored for stringed instruments, this technique assesses the composition's playability by examining various aspects of instrument performance. In the conducted experiments comparing our proposed tokenization and architecture with other methods, we evaluate them using a range of metrics, including the novel proposed evaluation metric playability rate, alongside pitch-based, rhythm-based, and subjective metrics. Results showed that while our model outperforms the comparison in certain aspects, key findings reveal the efficiency of our tokenization in terms of training speed and sequence length, the enhanced generation of harmonical structures due to pitch-based metrics, and the successful correlation of our proposed evaluation technique with existing ones.

Awards Won:
Second Award of $2,000