Music Transformer: Generating Music With Long-Term Structure

- Accepted Conference Name & Year : ex) ICLR2019
- 1st Author Name & Institute : Cheng-Zhi Anna Huang, Google Brain

Keywords

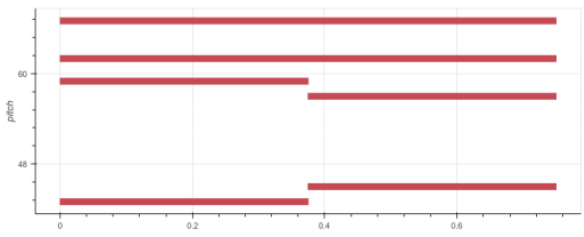
- Related Attention
- Space Complexity
- Memory.

Contribution

- Domain: 기존의 Seq2Seq 작곡모델의 제한된 길이의 메모리 문제를 해결.
- Algorithm: Transformer의 공간복잡도를 O(N^2D)에서 O(ND)로 단축.

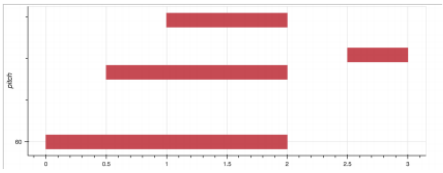
Proposed Architecture

- Pre Processing
  - Multi Track



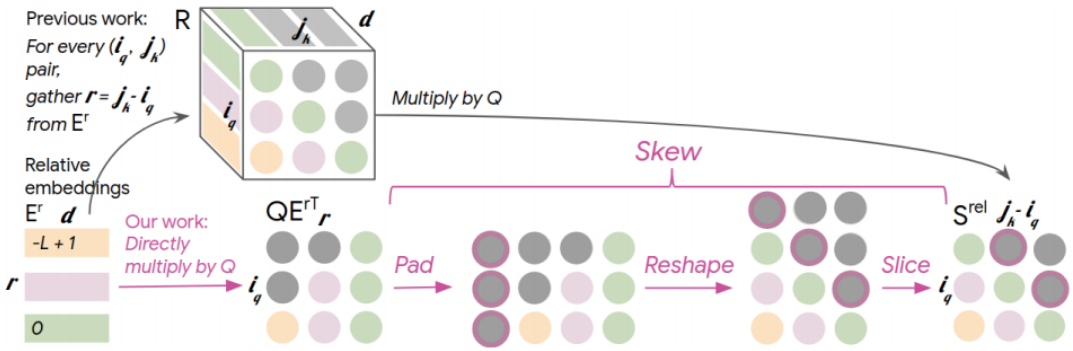

```
S: 67, 67, 67, 67
A: 62, 62, 62, 62
T: 59, 59, 57, 57
B: 43, 43, 45, 45
```

  - Single Track



```
SET_VELOCITY<80>, NOTE_ON<60>
TIME_SHIFT<500>, NOTE_ON<64>
TIME_SHIFT<500>, NOTE_ON<67>
TIME_SHIFT<1000>, NOTE_OFF<60>, NOTE_OFF<64>,
NOTE_OFF<67>
TIME_SHIFT<500>, SET_VELOCITY<100>, NOTE_ON<65>
TIME_SHIFT<500>, NOTE_OFF<65>
```

- Model



$$\text{RelativeAttention} = \text{Softmax}\left(\frac{QK^\top + S^{rel}}{\sqrt{D_h}}\right)V.$$

$$Z^h = \text{Attention}(Q^h, K^h, V^h) = \text{Softmax}\left(\frac{Q^h K^{h\top}}{\sqrt{D_h}}\right)V^h.$$

Dataset

- SERIALIZED INSTRUMENT/TIME GRID (J.S.BACH CHORALES) : pre-process 1
- MIDI-LIKE EVENT-BASED (PIANO-E-COMPETITION) : pre-process2 ←we use this