



3. Groove MIDI Dataset 리뷰

목표: Groove MIDI Dataset 을 살펴보고 데이터의 구조와 특징을 파악

Groove MIDI Dataset

- 인간이 연주한 템포에 맞춘 표현력 있는 드럼 연주의 **13.6시간에** 해당하는 MIDI 와 (합성된) 오디오로 구성되어 있다.
- 이 데이터셋에는 **1,150개의 MIDI 파일**과 **22,000개 이상의 드럼 연주 마디**가 포함되어 있다.

License

- Google LLC가 Creative Commons Attribution 4.0 International (CC BY 4.0) 라이선스 하에 공개

Dataset

- 각 연주는 드러머가 설정한 특정한 템포에 맞춰 메트로놈과 함께 연주되었다.
- 총 **10명의 드러머**에 의한 연주로 구성, 전체 연주 시간의 **80%이상은 전문 드러머의 연주**로 이루어져 있다.
- 드러머들은 **긴 연주와 짧은 비트 및 fills**를 혼합하여 연주하도록 지시되었다.

- 각 연주는 **장르, 템포 및 익명화된 드러머 ID**로 주석이 달렸다.
- 대부분의 연주는 **4/4박자**로 이루어져 있으며, **다른 박자의 예시도 일부 포함되어**있다.
- 네 명의 드러머에게 같은 10개 비트를 각자의 스타일로 연주하도록 요청했다. 이들은 테스트 세트로 분류되어 있으며, **eval-session/groove1-10**으로 레이블이 지정되었다.
- 실험의 주요 데이터 원본인 MIDI 녹음뿐만 아니라 드럼 세트의 합성 오디오 출력도 캡처하고 해당 MIDI 파일과 2ms 내에 정렬했다.
- 모델 정확도를 다양한 작업에서 쉽게 비교하기 위해 훈련/검증/테스트 분할 구성이 제공된다.

Split	Beats	Fills	Measures (approx.)	Hits	Duration (minutes)
Train	378	519	17752	357618	648.5
Validation	48	76	2269	44044	82.2
Test	77	52	2193	43832	84.3
Total	503	647	22214	445494	815.0

MIDI Data

Format

- 롤랜드 TD-11은 녹음된 데이터를 별도의 트랙으로 분리
 - 메타메시지(템포, 박자, 조표) 트랙, 컨트롤 변경 사항(하이햇 페달 위치) 트랙, 음표 트랙
 - 컨트롤 변경 사항은 채널 0에 설정, 음표는 채널 9에 설정
- 이 데이터를 간편하게 처리하기 위해, 배포하기 전에 원시 MIDI 파일을 두 가지 조정을 가함
 - 모든 메시지를 하나의 트랙으로 합침(meta. control change, and note)
 - 모든 메시지들을 채널9에 설정(10 if 1-indexed)

Drum Mapping

- 롤랜드 TD-11은 MIDI를 사용하여 연주를 녹음할 때, 범용 MIDI(GM)사양과 다른 피치 값을 사용한다.
- 아래에는 롤랜드 매핑이 GM과 비교되는 방식을 보여준다.
- 재생 및 훈련 중에 이러한 불일치에 주의해야 한다.
- 마지막 열은 논문에서 사용된 단순화된 매핑을 보여준다.

Pitch	Roland Mapping	GM Mapping	Paper Mapping	Frequency
36	Kick	Bass Drum 1	Bass (36)	88067
38	Snare (Head)	Acoustic Snare	Snare (38)	102787
40	Snare (Rim)	Electric Snare	Snare (38)	22262
37	Snare X-Stick	Side Stick	Snare (38)	9696
48	Tom 1	Hi-Mid Tom	High Tom (50)	13145
50	Tom 1 (Rim)	High Tom	High Tom (50)	1561
45	Tom 2	Low Tom	Low-Mid Tom (47)	3935
47	Tom 2 (Rim)	Low-Mid Tom	Low-Mid Tom (47)	1322
43	Tom 3 (Head)	High Floor Tom	High Floor Tom (43)	11260
58	Tom 3 (Rim)	Vibraslap	High Floor Tom (43)	1003
46	HH Open (Bow)	Open Hi-Hat	Open Hi-Hat (46)	3905
26	HH Open (Edge)	N/A	Open Hi-Hat (46)	10243
42	HH Closed (Bow)	Closed Hi-Hat	Closed Hi-Hat (42)	31691
22	HH Closed (Edge)	N/A	Closed Hi-Hat (42)	34764

Control Changes

- TD-11은 또한 각 히트마다 하이햇 페달의 위치를 지정하는 컨트롤 변경 사항도 기록합니다. 이 정보는 컨트롤 4로 남겨두었다.

Download

- GMD는 MIDI 파일과 WAV파일, 그리고 CSV 형식의 메타데이터를 포함한 압축 파일로 제공
- [groove- v1.0.0.zip](#)
Size: 4.76GB
SHA256:
`21559feb2f1c96ca53988fd4d7060b1f2afe1d854fb2a8dcea5ff95cf3cce7e9`
- 또한 MIDI만으로 이루어진 데이터셋도 제공된다.
- [groove-v1.0.0-midionly.zip](#)
Size: 3.11MB
SHA256:
`651cbc524ffb891be1a3e46d89dc82a1cecb09a57c748c7b45b844c4841dcc1e`
- 메타데이터 파일(info.csv)에는 각 MIDI/WAV 쌍에 대해 다음과 같은 필드가 있습니다

Field	Description
drummer	An anonymous string ID for the drummer of the performance.
session	A string ID for the recording session (unique per drummer).
id	A unique string ID for the performance.
style	A string style for the performance formatted as "<primary>/<secondary>". The primary style comes from the Genre List below.
bpm	An integer tempo in beats per minute for the performance.
beat_type	Either "beat" or "fill"
time_signature	The time signature for the performance formatted as "<numerator>-<denominator>".
midi_filename	Relative path to the MIDI file.
audio_filename	Relative path to the WAV file (if present).
duration	The float duration in seconds (of the MIDI).
split	The predefined split the performance is a part of. One of "train", "validation", or "test".

Genre List: afrobeat, afrocuban, blues, country, dance, funk, gospel, highlife, hiphop, jazz, latin, middleeastern, neworleans, pop, punk, reggae, rock, soul

TensorFlow Dataset

- TensorFlow Datasets(TFDS)를 사용하여 간단하게 `tf.data.Dataset`으로 로드할 수 있다.
- 다음과 같은 코드 라인만으로 데이터셋을 반복할 수 있다.

```

import tensorflow as tf
import tensorflow_datasets as tfds

# tfds works in both Eager and Graph modes
tf.enable_eager_execution()

# Load the full GMD with MIDI only (no audio) as a tf.data.Dataset
dataset = tfds.load(
    name="groove/full-midionly",
    split=tfds.Split.TRAIN,
    try_gcs=True)

# Build your input pipeline
dataset = dataset.shuffle(1024).batch(32).prefetch(
    tf.data.experimental.AUTOTUNE)
for features in dataset.take(1):
    # Access the features you are interested in
    midi, genre = features["midi"], features["style"] ["primary"]

```

- 데이터를 다양한 방식으로 전처리하기 위한 미리 정의된 구성(Configuration)도 포함시킴
- 예를 들어, 2박자 예제로 훈련하고 16KHz의 오디오를 사용하려면 “groove/2bar-16000hz”를 로드할 수 있다.
- 기존 환경에 반영되지 않는 설정을 사용하려면 자체 GrooveConfig를 생성하고 ‘tfds.load’의 ‘builder_config’ 인수에 전달하면 된다.

References

[Groove MIDI Dataset](#)