

사용자 맞춤 음악 제공을 위한 음악 편집 모델의 새로운 Condition 제안

오주원, 손봉기*, 이재호
덕성여자대학교, *서원대학교

2000violet@duksung.ac.kr, *bksohn@seowon.ac.kr, izeho@duksung.ac.kr

Proposal of new condition for Music Editing Model To Provide Customized Music

JooWon Oh, Bong-Ki Son*, Jaeho Lee
Duksung Women's Univ., *Seowon Univ.

요 약

본 논문은 기존 음악 편곡 모델에 컨디션 정보를 제공하여, 보다 효율적으로 사용자 맞춤 음악을 제공해주는 방법을 제안한다. 음악 생성 모델을 통해 생성한 멜로디 라인이 음악 편집 모델과 강화 학습을 거쳐서 사용자 취향 맞춤 음악을 제공한다. 본 논문에서는 음악 편집 모델과 강화 학습을 실행 과정에서 사용되는 음악 편집 모델의 condition 정보의 새로운 형태를 제안한다. 본 논문에서는 본래에 존재하는 InstructME를 활용하여, text condition을 본 논문에서 제안하는 다른 형태의 condition 정보를 제안한다.

I. 서 론

본 논문에서는 사용자 맞춤 음악 제공을 위한 conditional 한 음악 편집 모델을 제안한다. 기존에 개발이 완료된 InstructME를 활용하여, condition 정보를 본 논문에서 제안하는 새로운 방식을 적용하는 방식이다. 사용자 맞춤 음악을 제공하기 위해 음악 생성 모델, 음악 편집 모델과 강화 학습을 동시에 병행한다. 생성 모델을 통해 생성한 멜로디 라인을 사용하여, 음악 편집 모델을 통해 편곡 과정을 거쳐 사용자 맞춤 음악을 제공한다.

본 논문에서는 그 중에서도 음악 편집 모델을 중점적으로 다룬다. 본래의 InstructME에 다른 condition 정보를 주입하여 사용자 취향 맞춤의 음악을 제공하고, 강화 학습과 함께 학습을 병행하기에 좋은 형태를 제안한다.

편곡을 통해 리듬과 화성, 악기 구성요소를 변경해가며, 하나의 멜로디를 활용해 여러가지 분위기의 음악을 만들어낸다. 편곡 과정에서는 강화 학습이 함께 사용되어, 사용자 맞춤 편곡을 진행한다. 사용자 취향 맞춤 음악을 제공하기 위해 음악 편집 모델의 Condition 이자 Input을 사용해야 하는지에 대한 아이디어를 제시한다.

II. 본론

2.1 InstructME

Instruct ME는 음악의 장기적인 일관성을 위해 Chunk Transformer와 Conditional Latent Diffusion 모델을 사용하여 음악 편곡을 진행하는 음악 편곡 모델이다. 음악 특성상 장기적인 일관성과 지속성을 위해서 Chunk Transformer를 사용하여 음악 편곡을 진행한다. 특히나 편곡 작업을 할 때는 원곡을 잘 유지하여, 하모니와 조화를 이루는 것이 더욱 중요하다.

InstructME를 통해서 여러 가지 편곡 작업이 가능한데, 그 중에서는 add, remove, replace, extract, remix 등 하나의 음악에 대해 약 5 가지의 기능을 수행할 수 있다.

add의 경우에는 Source 음악에 대해서 “add a piano”, “remove a guitar”, “replace guitar with drum”, “extract a bass” 등 과같은 명령어로 해당 기능을 수행한다.

본 논문에서 도달하고자 하는 최종 목표는 사용자에게 맞춤 취향의 노래를 제공하는 것이다. 그러나, 생성 모델을 통해 이를 수행하게 되면 사용자 취향의 노래가 나올 때까지 랜덤하게 사용자에게 생성 음악을 제공해야 한다는 점에서 굉장히 비효율적이다. 그러나 음악 편곡은 생성 모델에서 하나의 멜로디만 생성하게 되면, 사용자가 취향의 편곡 방향을 지속적으로 제공함으로써 사용자에게 만족스러운 음악을 제공할 수 있다.[1]

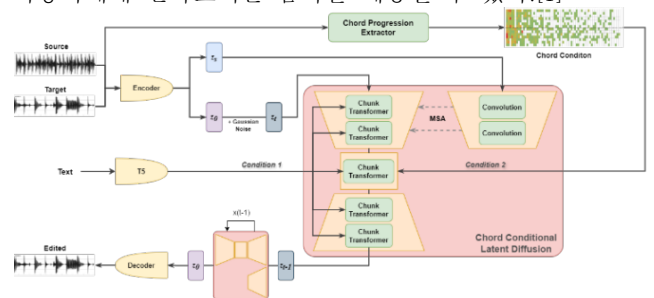


그림 1. InstructME 모델 구조

2.2 강화 학습에 대한 Action 값

사용자 맞춤 음악을 제공하는 과정에서 강화학습을 함께 사용한다. 강화학습을 진행할 때 액션 값이 필요로 하는데 액션 값으로 편곡이 완료된 곡이 입력되게 된다. 그리고 해당 곡을 진행된 강화학습에 대한 output 값이 Condition 정보로 사용된다.

강화 학습은 사용자 반응형으로 진행하게 되며, 음악의 재생 시간, 사용자의 감정 인지 등을 통해 사용자가 제공한 음악이 사용자의 취향인지에 대한 파악을 진행한다. 이에 대한 Input 값은 편곡이 완료된 음악이 될 것이고, 이 과정을 반복해가며 사용자 취향을 강화

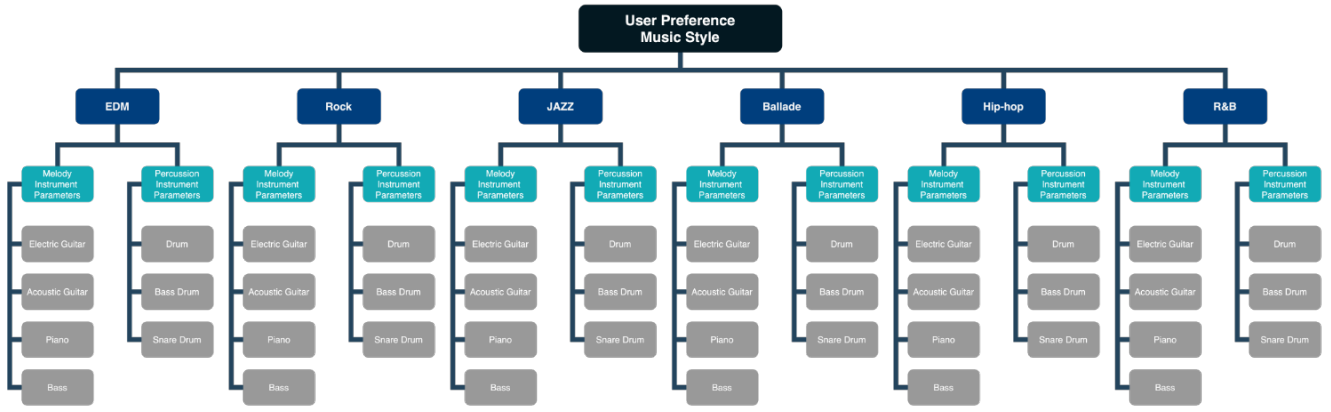
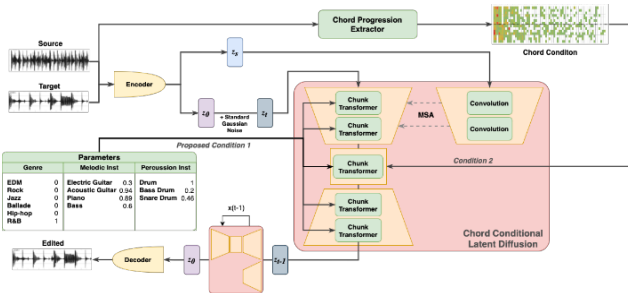


그림 3. Music Editing Condition 정보에 대한 Parameter 구조

학습을 통해 학습한다. 학습이 완료되면, 사용자의 취향에 맞는 다양한 노래를 실패없이 제공할 수 있다.

2.3 Latent Diffusion Conditional 정보 대체



InstructME에서는 두 가지의 Condition 정보를 요구한다. 사용자가 직접 입력하는 text description 과 Source 음악에서 추출한 코드 진행 정보 두 가지가 Condition 정보로 주입한다.

Condition 정보에서 사용자 입력 텍스트는 “Add a guitar”, “Replace Bass with Electric guitar”와 같이 편곡의 방향성에 대해 제시한다. InstructME는 사용자가 제시한 텍스트를 해석하여, 사용자가 원하는 대로 편곡을 진행한다. Chord Progression의 경우는 Source 음악과의 하모니가 조화를 이루어야 하기 때문에 Chord Progression에 대한 정보를 추출하여 condition 정보로 주입한다.

본 논문은 InstructME에서 text description을 대체할 수 있는 다른 Condition 정보를 주입하는 것을 제안한다. Text(natural language)와 같이 정형화된 형태가 아닌 강화 학습의 결과로 나온 output을 다른 형태의 데이터로 주입한다. 강화 학습의 output이 또 다시 text condition을 대체할 수 있는 다른 condition으로 주입할 새로운 형태를 제안하고자 한다.

본 논문에서는 text condition을 대체할 다른 condition 정보로 파라미터를 사용하는 것을 제안한다. 본 논문에서 제안하는 편곡 방식은 하나의 소스 음악에 대하여 악기를 추가, 제거 및 노래 리믹스를 중점적으로 사용한다. 그렇기 때문에, 각 장르 및 악기를 파라미터 값으로 설정하여 어떤 장르에 어떤 악기가 주요하게 사용될지, 어느 정도의 강도 혹은 빈도로 사용될지에 대한 파라미터 값을 설정해준다.

먼저 파라미터를 크게 장르로 구분한다. 장르에 대해서 1과 0 사이의 실수 값으로 파라미터 값을 설정한다. 그리고 각 장르 하위에는 악기에 대한 파라미터 값이 존재한다. 악기에 대한 파라미터 값 역시 1부터 0 사이의 실수를 사용하여, 1은 가장 강하게, 0은 아예 사용되지 않는 것으로 설정한다.

강화 학습을 통해서 사용자가 각 장르에 대해서 가장

선호하는 악기 스타일을 조정한다. 지속적으로 조정을 진행하다가 장르 별로 사용자 선호에 따른 최적화 가능하다. 그렇게 되면, 장르에 대한 실수 값은 각 장르별 악기의 비중에 대한 일종의 가중치로 사용된다.

같은 장르의 곡이라고 해도 어떤 악기가 중점적으로 사용되는지에 따라 사용자의 선호도가 갈릴 수 있기 때문에 각 장르 별로 악기에 대한 파라미터를 설정하는 방법을 선택하였다.

III. 결론

본 논문에서는 기존 InstructME에 text description을 대체할 파라미터 condition을 강화 학습과 함께 학습을 통해 사용자 취향 맞춤 음악을 어떻게 제시해야 할지에 대한 방법을 제안한다.

아직 아이디어 구현 단계이기 때문에 추후에 개발을 통해 더 보완하고, 구체화 시킬 예정이다. 이는 인공지능으로 생성된 음악으로 사용자에게 취향에 맞는 음악을 제공해줄 수 있을 것이다. 이는 자동차에 삽입되어, 저작권 문제 없이 사용자가 운전 시에 즐거움을 느끼거나, 업무 시간에도 사용자의 무료함을 달래어 줄 수 있을 것이다

ACKNOWLEDGMENT

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원받아 수행된 연구임(과제번호-2022R1A2C1009951).

참고 문헌

- [1] Bing H." InstructME: An Instruction Guided Music Edit And Remix Framework with Latent Diffusion Models," Aug 2023.