

몬테카를로 트리 탐색을 이용한 다양한 음악 감정 햅틱패턴 생성

박태화[○], 유원상[○], 전현창[○], 김경중^{*}
광주과학기술원 융합기술원 융합기술학제학부
광주과학기술원 AI 대학원

taehwa-p@gm.gist.ac.kr, u.wonsang0514@gm.gist.ac.kr, kevinjeon119@gm.gist.ac.kr, kjkim@gist.ac.kr

Generating Diverse Emotional Haptic Patterns for Music using Monte Carlo Tree Search

Tae-Hwa Park[○], Won-Sang Yoo[○], Hyeon-Chang Jeon[○], Kyung-Joong Kim^{*}
School of Integrated Technology, Gwangju Institute of Science and Technology
AI Graduate School, Gwangju Institute of Science and Technology

요약

음악 감상 중 청각이 아닌 다른 자극을 동시에 주게 되면 사용자는 음악만 들을 때보다 풍부한 경험을 얻는다. 그러나, 이를 위해 생성된 주파수 기반 햅틱패턴은 사용자에게 박자 정보를 전달하는데 충분하지 않, 음악의 감정 등 다른 요소를 전달하는 데 한계가 있다. 본 논문에서는 이러한 한계점에 기반하여 감정을 잘 전달 할 수 있는 햅틱을 생성하는 모델을 제안한다. 제안하는 모델은 감정 햅틱패턴 데이터베이스와 몬테카를로 트리 탐색을 이용하여 주어진 감정과 음악에 맞는 감정 햅틱패턴을 생성하며, 실험을 통하여 제안하는 모델이 음악의 비트에 맞추어 햅틱패턴을 생성하는 것을 보이고, 주파수 기반 햅틱패턴과 비교하여 더 나은 다양성을 보임을 확인했다.

1. 서론

사람과 컴퓨터가 상호작용을 하는 분야에서, 햅틱은 촉각에 대한 감각 정보를 제공하는 수단 중 하나이다. 햅틱은 다양한 분야에서 사람에게 풍부한 경험을 제공하도록 도움을 준다. 음악 감상에서의 햅틱 또한 촉각을 통해 사용자에게 더 풍부한 경험을 제공한다. 이러한 햅틱 자극을 이용하여 음악 감상에 도움을 준 예로는 청각 장애인들에게 햅틱 자극을 통해 음악의 비트를 느끼게 도와주는 연구가 있다[1]. 청각장애인뿐만 아니라 일반적인 사용자도 음악 감상에 있어 햅틱 자극이 도움이 된 연구도 있다[2].

그러나, 이러한 햅틱 자극이 항상 도움을 주는 것은 아니다. 대표적으로 뮤지컬 햅틱 웨어러블 장치(Musical Haptic Wearable for Audiences)(MHW)는 햅틱 자극과 들리는 음악의 관계를 고려하지 않고 음악의 비트나 분위기에 맞지 않는 진동을 사용자에게 주어, 오히려 햅틱을 이용한 음악 경험을 불쾌하게 느끼는 사례를 발견할 수 있었다[2].

본 논문에서는 이러한 문제점에 기반하여 일반인과 청각장애인의 음악 감상 경험을 높이기 위한 감정 햅틱패턴을 생성하는 모델을 제안한다. 제안하는 모델은, 감정-햅틱패턴을 나누어 감정 햅틱패턴 데이터베이스를 설계했으며, 생성된 데이터베이스를 이용하여 몬테카를로 트리 탐색을 이용하여 주어진 감정과 음악에 맞는 패턴을 생성하는 모델이다.

성능 확인을 위해 5초 길이의 짧은 음악을 입력으로 그에 맞는 햅틱패턴을 생성하는 실험을 진행했으며, 원본 음악과 비트 수 차이와 비트가 찍힌 시간대를 비교함으로써 성능을 비교했다. 그 결과, 제안하는 모델은 무작위 모델 대비 비트 수 차이와 비트가 찍힌 시간대가 더 정확한 모습을 보여 음악과 잘 맞는 햅틱패턴을 생성했으며, 시간 동적 와핑(Dynamic Time

Warping)(DTW) 알고리즘을 사용하여 생성한 햅틱패턴의 차이를 측정하는 거리가 주파수 기반 햅틱패턴보다 더 떨어진 것을 통해 다양한 햅틱패턴을 생성한 것을 확인했다.

2. 관련 연구

2.1. 햅틱과 음악

기존에는 음악이 고려된 햅틱패턴을 생성하는 연구들이 있었다. 유전 알고리즘을 사용한 기존연구는 모델이 개인에 맞는 경험을 전달하기 위해 기존 진동의 세기를 조절해 사용자가 인식한 것에 맞게 바뀌는 역할을 했다. Macaron 햅틱 인터페이스에서는 사용자가 원하는 햅틱패턴을 생성하여 직접 사용할 수 있게 만들었다[3]. 다른 연구로는 연구진이 생성한 햅틱패턴에 설문문을 통해 얻은 감정 레이블을 이용하여 VibViz라는 데이터셋을 구축하여 햅틱패턴을 분석, 시각화한 연구가 있다[4].

이러한 기존 연구는 햅틱만 따로 생성하여 감정을 측정하거나 직접 패턴을 디자인하여 사용했다. 이러한 데이터들은 음악과 음악의 감정을 고려하여 햅틱을 생성한 것이 아니기 때문에 비록 햅틱패턴의 질은 좋아도 바로 음악을 표현하기 어려운 점이 있다. 이와 반대로 음악의 특정 주파수 대역을 추출하여 햅틱패턴으로 사용하면, 박자 정보를 전달하기 좋지만, 이외의 정보를 전달하기 어려워진다. 제안하는 모델은 잘 디자인된 햅틱패턴을 음악의 박자 정보에 맞추어 배치함으로써 음악에 맞는 차별화된 햅틱패턴을 생성한다는 차별점을 가진다.

2.2. 몬테카를로 트리 탐색

몬테카를로 트리 탐색(Monte Carlo Tree Search)(MCTS)은 현재 상태를 뿌리 노드로 설정 후 행동 가능한 선택지를 자식 노드로 하여 Selection, Expansion, Simulation, Backpropagation 네

단계를 반복하며 충분한 데이터를 쌓는다. 이후, 현재 상태에서 가장 가치가 높다고 평가되는 노드를 선택하는 탐색 알고리즘이다. 각 단계의 세부사항은 아래와 같다[5].

- Selection : 트리 정책(Tree Policy)에 맞추어 뿌리 노드부터 다음에 시뮬레이션을 진행할 노드를 선택한다.
- Expansion : 선택된 노드가 자식 노드를 생성할 수 있다면, 자식 노드를 생성하고 생성한 노드를 선택한다.
- Simulation : 선택된 자식 노드의 상태에서 기본 정책(Default Policy)에 따라 시뮬레이션을 진행한다.
- Backpropagation : 시뮬레이션의 결과를 해당 노드에서 루트 노드까지 관련 있는 모든 부모 노드를 업데이트한다.

MCTS는 현재 상태에서 다음 액션을 선택하는데 특화된 알고리즘으로 다양한 분야에 적용하기 쉽고, 신뢰성이 높은 강점이 있으며, 주로 AlphaGo와 같이 Game AI 분야에서 게임을 하는 인공지능을 만들기 위하여 사용되었다. 본 논문에서는 현재까지 생성된 햅틱패턴에 가장 잘 어울리는 햅틱패턴을 선택하고 처음 선택한 햅틱패턴에 따라 다양한 햅틱패턴이 생성되는 것을 기대하여 MCTS를 사용했다.

3. 제안하는 방법

3.1. 감정-햅틱 데이터베이스 생성

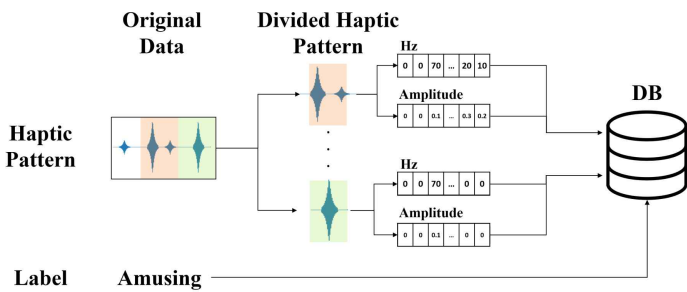


그림 1. 햅틱 데이터베이스 생성과정

위 그림은 데이터베이스 생성 방법에 대해 간략하게 표현한 것이다. 데이터베이스는 VibViz[4]에서 제공하는 햅틱패턴 데이터와 감정 레이블값 데이터를 이용하여 생성했다. 하나의 햅틱패턴을 여러 개로 나누기 위하여 비트를 추출하여 이를 특징점으로 이용했다. 원본 햅틱패턴 데이터는 특징점 기준으로 이전, 이후에 소리가 없는 부분을 추출하여 하나의 원본 패턴 데이터당 최소 1개에서 최대 4개의 패턴으로 나누었다. 잘라낸 햅틱 데이터는 주파수(Hz)와 세기(Amplitude)로 값을 변환하여 저장했으며, 해당 햅틱패턴에 맞는 감정을 레이블 값으로 같이 저장했다. 총 120개의 햅틱패턴 데이터를 이용하여 'amusing', 'annoying', 'anxious', 'beautiful', 'calm', 'energizing', 'indignant', 'joyful', 'sad' 9개의 감정 레이블에 대해 293개의 햅틱패턴 데이터를 생성했다.

3.2. MCTS 기반 햅틱패턴 생성

제안하는 모델은 생성된 햅틱패턴 데이터베이스와 MCTS를 이용하여 그림 2와 같이 동작한다. 음악의 비트 배열을 추출하여 감정값과 함께 모델에 입력으로 들어간다. MCTS는 입력받은 비트 배열 데이터를 기반으로 트리의 깊이와 비트의 수를 정한다. 이후 MCTS가 탐색을 진행하여 현재 비트 위치에 배치할 햅틱패턴 부분 1개를 선택하게 된다. 마지막으로 선택된 햅틱패턴 부분을 비트 배열 데이터에 삽입한 뒤 업데이트된 비트 배열 데이터를 입력으로 다시 MCTS를 통해 햅틱패턴 부분을 선택한다. 이 과정을 반복하여 입력 비트 수만큼 햅틱패턴 부분을 선택했다면, 생성된 전체 햅틱패턴을 출력한다. 이를 위해

MCTS에서 Tree Policy는 Upper Confidence Bound, version 1(UCB1)을 사용하였고, Default policy는 데이터베이스의 클립을 무작위로 선택하여 붙이는 방법을 선택했다.

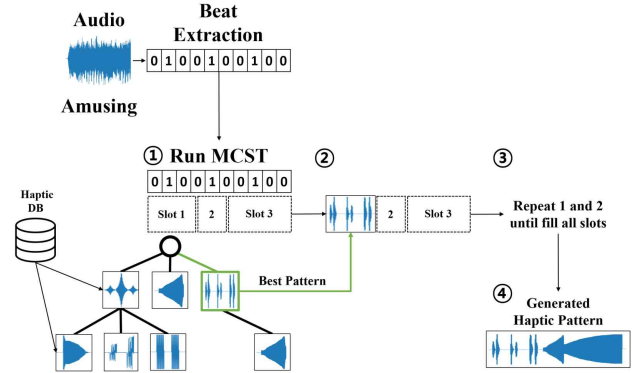


그림 2. 제안하는 모델의 동작 예

시뮬레이션을 통해 생성된 햅틱패턴과 각 노드에 점수를 기록하기 위한 평가함수는 아래 식과 같다.

$$Score = -(1 - BeatCoverage)^2 + BeatHitRate$$

$$BeatCoverage = \frac{b_g}{b_m}$$

$$BeatHitRate = \frac{b_a}{b_g}$$

수식 1. 제안하는 모델이 사용하는 햅틱패턴 평가함수

평가함수는 입력 비트와 생성된 햅틱패턴의 비트가 잘 정렬되는 패턴을 만들기 위하여 위와 같이 설계했다. 비트 커버리지(Beat Coverage)는 입력 비트와 생성된 햅틱 비트의 비율을 계산하는 함수이다. 비율이 1에 가까울수록 생성된 비트 수가 입력된 비트 수와 가까워진다는 것을 의미한다. 비트적중률(Beat Hit Rate)은 입력 비트와 생성된 비트가 같은 시간대에 있는 비율을 계산하는 함수이다. 비율이 1에 가까울수록 생성된 비트와 입력된 비트가 같은 시간대에 있다는 것을 의미한다. 본 평가 함수와 감정 레이블이 있는 데이터베이스를 통해 감정 입력에 따라 다양하며, 음악과 비트가 어긋나지 않는 햅틱패턴을 생성하는 것을 목표로 한다.

4. 실험 및 결과

제안하는 모델의 성능을 확인하기 위하여 제안하는 모델과 같은 데이터베이스를 사용하여 무작위로 햅틱패턴을 생성하는 무작위 모델과 원본 음악의 1hz에서 250hz 대역의 소리를 추출하여 햅틱패턴으로 변경하는 주파수 기반 모델을 비교했다. 입력으로 사용한 데이터는 5초 길이의 짧은 음악 클립 데이터[6]를 사용했으며, 36개의 음악 클립을 9개의 감정으로 생성하여, 총 324개의 햅틱패턴으로 비교를 진행했다. 비교를 위한 지표로는 생성된 햅틱패턴이 원본 음악과 잘 어울리는지 확인하기 위하여 비트를 비교한 Coverage와 생성된 햅틱패턴이 원본과 비슷한 수의 비트를 생성했는지 확인하는 Alignment이다.

$$Alignment = \frac{1}{m} \sum_{i=1}^m \exp\left(-\frac{\min_{t_j^y \in B^y} \|t_i^x - t_j^y\|^2}{2\sigma^2}\right)$$

수식 2. 비트 정렬도 계산을 위한 Alignment 수식

Alignment는 입력 음악의 비트와 감정을 고려한 햅틱패턴의 비트의 거리가 가장 작은 값을 구한 뒤, 설계된 Alignment 함

수를 통해 점수를 뽑는다. 세부사항은 식 2와 같다. Coverage와 Alignment 모두 1에 가까울수록 원본음악에 가까운 비트를 생성했다고 볼 수 있다. 마지막으로 감정에 따라 생성된 햅틱패턴의 다양성을 확인하기 위하여 DTW를 이용하여 같은 곡에서 생성된 다른 햅틱패턴 사이의 거리를 측정했다.

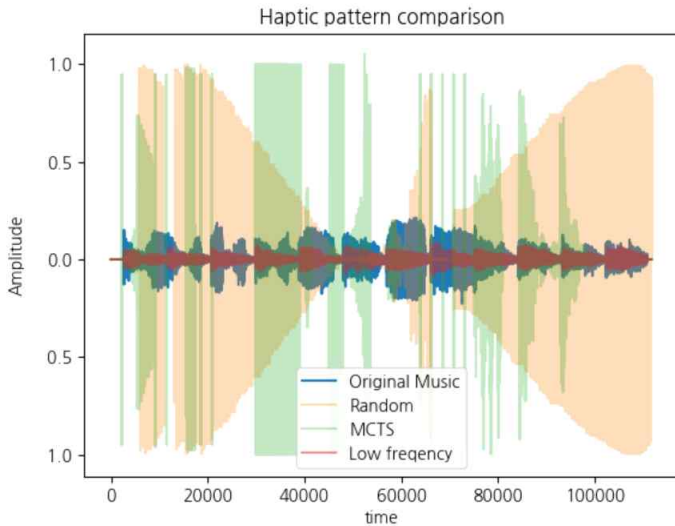


그림 3. 같은 음악에서 각 모델의 햅틱패턴 출력 예

된 햅틱패턴 중엔 무작위 모델이 사용한 오래 지속되는 햅틱패턴도 있지만 제안하는 모델이 사용하는 짧은 햅틱패턴이 섞여 있다. 이에 따라, 긴 패턴을 사용하면 이후 비트를 맞추지 못하여 평가함수의 점수가 낮아지기에 제안하는 모델은 비트를 잘 맞추기 위하여 길이가 짧은 패턴을 이어붙여 햅틱패턴을 생성하는 경향을 보였다. 다양성이 무작위 모델보다 낮아진 것도 긴 패턴을 사용하지 않는 경향 때문으로 보인다.

5. 결론 및 향후 연구

본 연구에서는 기존 주파수 대역을 기준으로 생성하던 햅틱패턴을 감정 데이터를 활용하여 이전보다 더 풍부한 햅틱패턴을 생성하는 연구를 진행했다. 다양한 패턴 생성을 위하여 MCTS를 이용하여 감정이 반영된 햅틱패턴을 생성했다. 감정이 반영된 패턴의 일부는 기존 패턴만큼의 비트 정확도를 갖고 있었으며, DTW를 통해 측정된 거리로 다양한 음악이 생성되었음을 발견했다.

향후 연구에서는 실제 사람들을 대상으로 실험을 하여 생성된 햅틱패턴이 기존 패턴보다 더 감정이 많이 느껴지는지, 기존 패턴보다 음악을 잘 표현하는지 검증할 예정이다. 또한, 본 논문에서 사용한 MCTS는 짧은 햅틱패턴을 잘 생성하기에 1분 이상의 긴 햅틱패턴 생성을 위해 유전 알고리즘(Genetic Algorithm), 탐욕적 탐색(Greedy Search), 임의 탐색(Random Search) 등의 다양한 방법들을 통해 생성된 패턴을 비교해 최적의 방법을 검증할 예정이다.

마지막으로, 데이터베이스 기반 모델은 수집한 패턴의 길이와 수에 영향을 크게 받는 것을 확인했다. 이를 극복하기 위해 햅틱 데이터베이스의 패턴 수와 길이의 변화에 따른 생성된 음악의 결과를 비교하여 햅틱 생성을 위한 데이터베이스 구축에 대한 방법도 검증할 예정이다.

감사의 글

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 연구개발 지원사업으로 수행되었음(과제번호: R2021040048)

참고문헌

[1] Fletcher, Mark D. "Can haptic stimulation enhance music perception in hearing-impaired listeners?." *Frontiers in Neuroscience* (2021): 1123.
 [2] Nanayakkara, Suranga, et al. "An enhanced musical experience for the deaf: design and evaluation of a music display and a haptic chair." *Proceedings of the sigchi conference on human factors in computing systems*. 2009. pp337-346
 [3] Schneider, Oliver S., and Karon E. MacLean. "Studying design process and example use with Macaron, a web-based vibrotactile effect editor." *2016 IEEE Haptics Symposium (HAPTICS)*. IEEE, 2016.
 [4] Seifi, Hasti, Kailun Zhang, and Karon E. MacLean. "VibViz: Organizing, visualizing and navigating vibration libraries." *2015 IEEE World Haptics Conference (WHC)*. IEEE, 2015.
 [5] Chaslot, Guillaume, et al. "Monte-Carlo Tree Search: A New Framework for Game AI." *AIIDE 8* (2008): 216-217.
 [6] Cowen, A. S., Fang, X., Sauter, D., & Keltner, D. (2020). What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences*, 117(4), 1924-1934.

	Coverage		Alignment		Diversity
	Avg.	Max	Avg.	Max	
무작위	60.14	1.14	0.16	0.94	167.45
저주파	0.98	1	0.97	1	0
MCTS	14.28	0.99	0.36	0.98	112.32

표 1. 모델별 비트 정확도 및 다양성 결과

결과는 위와 같다. 표 1의 값은 생성한 햅틱패턴과 원본 곡을 비교하여 얻은 값이다. 비트 정확도와 비트 정렬도의 경우, 제안하는 모델이 무작위 모델보다 비트에 잘 맞는 햅틱패턴을 생성했지만, 저주파 모델에 미치지 못했다. 이는 그림 3에서도 확인할 수 있는데, 생성된 햅틱패턴(초록색)과 무작위 패턴(주황색)은 입력 음악(파란색)과 큰 차이가 나지만 주파수 기반(빨간색)은 거의 같다.

이는 제안하는 모델의 평가함수가 비트의 수에 크게 영향을 받으면서 동시에 비트 적중률을 높게 설계되어 있기 때문이다. 비트 수가 원곡보다 적은 상태에선 비트의 수를 늘려 수식 1의 BeatCoverage를 1에 가깝게 만들고 동시에 적중률을 높이려고 하지만, 비트의 수가 너무 많을 땐 페널티가 크기 때문에 비트가 적은 패턴을 사용하여 전체적인 비트의 수를 조절하려고 하기 때문이다. 이에 따라, 비트가 많은 햅틱패턴을 배치한 직후 비트가 거의 없는 패턴을 배치하는 모습을 보여주어 비트 수를 정확하게 맞추는 것에 한계가 있다. 다양성의 경우, 저주파 모델은 원본 음악의 주파수 대역이 변하지 않는 이상 항상 같은 출력을 내기에, 다양성이 없는 모습을 보이지만, 다른 두 모델은 다양한 출력이 나올 수 있다.

이를 통해, 제안하는 모델이 한 음악에서 주어진 비트 배열에 따라 다양한 햅틱패턴을 생성하되, 무작위 생성과 비교하여 더 나은 성능을 보임을 확인했다. 다만, 최고 성능과 평균 성능의 차이가 크게 나는데, 이는 평가함수가 비트의 수를 맞추는 것을 중요하게 여기는 점과 구축한 데이터베이스의 데이터가 감정 레이블에 따라 비트를 맞추기 쉬운 패턴과 어려운 패턴이 섞여있는 점 때문이다.

그림 3의 무작위 패턴과 생성된 패턴이 이를 보여준다. 수집