

강화학습을 통한 음악 작곡 연구

정택현*, 이승호*, 김기천^o

A Study on Music Composition through Reinforcement Learning

Tack-hyun Jung*, Seung-ho Lee*, Keecheon Kim^o

요약

본 논문은 음악의 작곡(composition)이라는 분야에 이미지 인식(image recognition)과 강화학습(reinforcement learning)을 접목하려는 시도와 음악적 지식과 인간의 직관을 배제하고 음악을 창작하려는 목적의 알고리즘 작곡(algorithmic composition)에 관한 연구결과를 기술한다. 제안하는 강화학습 환경은 소리의 파형이나 형태에 초점을 맞추는 소리 인식 기반의 연구들과는 달리 이미지 인식 관련 기술을 활용하여 악보의 정보를 추출하는 차별화된 새로운 시도를 수행하였으며, 제안하는 구조와 방법론에 대한 명세를 상세히 기술한다. 또한, 각기 다른 두 가지 actor를 활용하여 학습을 수행하는 구조를 고안함으로써, 기존 마르코프 체인(Markov Chain)을 통한 접근법보다 더욱 심도 있는 접근 방향을 제시한다. 이러한 시도는 향후 더욱 고도화를 위한 초석이 되며, 이를 개선하기 위한 후속 연구의 방향성을 제시한다고 생각한다.

Key Words : 강화학습, Music Composition, Image Recognition, Markov Chain

ABSTRACT

In this paper, we propose a new approach to music composition through the image recognition and reinforcement-learning and we conducted a study to creating music, excluding musical knowledge and human intuition, and described the results. The proposed learning environment in this paper, is a new differentiated approach to extract the information of a sheet music using image recognition-related technologies, unlike studies based on Sound Recognition that focus on sound waves. also, we designing an architecture of learning environment that utilizes two different actors, proposing a more in-depth approach than a traditional Markov Chain. thus, we detail the proposed structure and methodology. These experiments will be a cornerstone for research to advancement in the future works and It will be direction of our future research to improve it.

I. Introduction

음악의 작곡(composition)이란 음의 배열을 음악적 이론에 근거하여 체계적인 방식으로 창조하는 행위와 그 결과로서의 작품이다. 일반적으로 작곡은 작곡 이

론, 화성법, 대위법, 음악 형식론 등을 바탕으로 하며, 작곡가의 창의성이 더해지면서 창작물이 만들어진다.

이러한 음악의 창작은 높은 수준의 음악적 지식과 감성이 요구되며, 창작의 주체는 사고가 가능한 ‘인간’으로 전제되어왔다.^[1] 하지만, 최근 복잡한 확률을 통제하는 알고리즘과 기계 학습 등 다양한 기술이 개발

* 본 연구는 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-차세대정보-컴퓨팅기술개발사업의 지원(No. NRF-2017M3C4A7083678, 더 나은 Web 경험을 위한 자율제어 네트워크 애플리케이션 기술)을 받아 수행되었습니다.

♦ First Author : Konkuk University Department of IT Convergence Information Security, tackhyun12@konkuk.ac.kr, 학생회원

° Corresponding Author : Konkuk University Department of Computer Engineering, kckim@konkuk.ac.kr, 중신회원

* Konkuk University Department of Computer Information & Communications Engineering, phg0726@konkuk.ac.kr, 학생회원

논문번호 : 201910-254-D-RN, Received October 28, 2019; Revised December 11, 2019; Accepted December 12, 2019

됨에 따라, 이러한 컴퓨터 과학을 음악에 접목하는 연구가 활발하게 진행되고 있다.^[1-10]

이러한 연구 중 가장 대표적인 연구로는 Google의 AI 프로젝트인 'Magenta'가 있으며^[2], 이는 소리의 파형에 초점을 맞춰 소리 인식을 구현하고 이러한 원리를 통해 학습과 작곡을 수행한다. 이처럼 소리의 특성을 이용하는 해당 접근법은 소리를 해석하는 알고리즘이 반드시 포함되어야 하므로, 음악에 대한 전문적인 지식의 필요성이 제기된다.

본 연구는 알고리즘 작곡(algorithmic composition) 연구의 일환으로 음악적 지식과 사람의 직관(Heuristic)을 배제하는 방향의 학습 환경을 정의 및 구현하고자 한다. 이를 통해 명확한 기준과 정답이 없는 음악이라는 분야에서 알고리즘 작곡이 어떠한 가능성을 가지는지를 살피고자 한다.

II. Related Works

2.1 Music composition through Markov chain

마르코프 체인(Markov chain)^[3]을 음악이라는 장르에 접목하는 최초의 시도는 1950년대 Olson, Belar에 의해 수행되었다.^[4] 해당 연구는 11개의 멜로디를 대상으로 음고와 리듬을 2차 전이 확률(transition probability)의 행렬로 구성하는 원리로 작곡을 수행하였다.^[4]

이처럼 상태(state)를 정의하고 전이 확률을 계산하는 마르코프 체인(Markov chain)을 음악에 접목하는 연구는 범용적인 접근법이 되어, 다양한 연구와 시뮬레이션을 통해 점차 발전되어 왔다.^[5] 최근 수행된 대표적인 연구로는 마르코프 체인 모델이 음(note)의 지속시간(duration)을 적절히 반영하지 못한다는 한계점을 개선한 모델을 제시한 선행연구^[6]가 있다.

하지만, 마르코프 체인을 활용한 작곡 방법은 무한대에 수렴하는 확률을 통제해야 하는 고질적인 문제점을 가지고 있는 것으로 알려져 있다.^[7] 이는, 하나의 음(note)을 개별적인 하나의 상태(state)로 나타낼 경우, 고려해야 하는 상태가 기하급수적으로 증가한다는 점이다. 차수(order)를 늘리는 해결법을 고려할 수 있지만, 이는 차수를 늘렸을 때 상태의 개수는 여전히 지수배로 늘어난다는 문제점에 따라 큰 어려움이 있다.

따라서 선행연구^[8]는 계층적 클러스터링(hierarchical clustering)을 통해 비슷한 마디를 같은 상태(state)로 묶는 원리로 확률 문제를 개선하고자 하였고, 이와 유사한 다양한 연구에서도 이러한 문제를

개선하기 위한 이론을 연구하였다.

하지만, 이러한 접근법 중 대다수는 음악적 이론과 지식에 근거한 해결방법을 제시하였다. 즉, 인간의 직관, 감각, 지각, 정서, 감정 등과 화성법, 대위법, 음악 형식론, 악기론 등 복합적인 음악 이론이 해결법에 휴리스틱(Heuristic)으로 작용하면서 문제를 해결하였다.

본 연구는 이러한 문제 해결법들과는 달리, 음악적 지식과 인간의 직관을 거의 배제하면서 음악을 창작하는 알고리즘을 구현하는 것을 목표로 한다. 따라서 다양한 접근방법을 고찰하던 중 복잡한 문제에 대해서 사전 지식의 필요 없이 스스로 환경과 상황을 판단하여 학습을 수행하는 강화학습 이론에 주목하였다.^[8]

강화학습(Reinforcement Learning)은 적용하고자 하는 환경(environment)을 상태(state)와 행동(action) 그리고 행동을 수행함에 따라 기대되는 보상(reward)으로 정의하여 미래에 기대되는 가치를 극대화하는 방향으로 자율적인 의사결정을 하는 알고리즘이다.^[8]

본 연구는 이러한 강화학습의 특성을 통해 무한대에 수렴하는 상태 전이 확률(state transition probability)에서 효율적으로 학습을 수행하고 결과를 도출해 내기 위한 시도와 그 결과를 고찰하고자 한다. 또한, 이러한 일련의 과정에서 완전한 탈 휴리스틱을 위해 악보를 읽어내어 데이터를 추출하는 과정을 이미지 인식과 관련된 기술을 활용하는 구조를 고안하였으며, 이 외의 모든 영역에서도 음악적 지식과 사람의 직관 등을 최대한 배제하는 것을 중점적으로 하였다.

III. Music composition with RL

3.1 제안하는 강화학습 모델의 개념

본 연구에서 제안하는 음악을 만들기 위한 강화학습 모델은 [Fig 1]과 같은 일련의 절차를 갖는다. 먼저, 이미지 형태의 악보로부터 정보를 추출하는 과정이 선행된다. 이러한 과정의 선행은 일차적으로 강화학습을 수행하기 위한 데이터를 획득하기 위함이다.

이처럼 획득된 데이터는 본 연구에서 제안하는 두 개의 Actor를 병렬구조로 갖는 강화학습 모델(RL Model)에서 유의하게 활용되는데, 두 Actor는 서로 다른 목적의 행동과 학습을 수행하는 구조를 갖는다. 이러한 일련의 과정이 모두 수행되면, 결과물로 강화학습을 통해 생성된 악보를 이미지 파일의 형태로 저장한다.

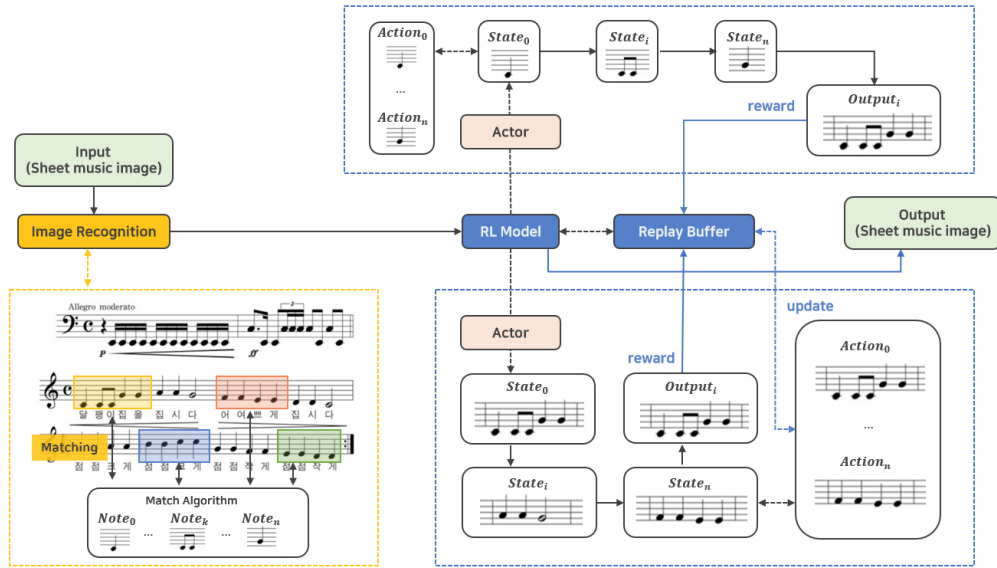


그림 1. 제안하는 음악제작의 강화학습 환경
Fig. 1. A Proposed learning environment for Music composition

3.2 이미지 인식 구조

이미지 인식(Image Recognition, IR)을 통해 악보 이미지로부터 데이터를 추출하는 과정은 템플릿 매칭(template matching)을 기반으로 구현한다. 수식(1)은 이러한 방법론 중 하나로 픽셀의 제공 차를 이용하는 탐색 및 매칭 방법이다.^[9] 탐색 대상인 T 를 탐색영역 I 로부터 점진적으로 이동시켜가며 제공의 합을 계산한다. 이는, 일치하는 영역을 찾는 경우, 0을 반환하기 때문에 이를 통해 T 를 I 의 영역으로부터 탐지하는 원리이다.

$$R(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2 \quad (1)$$

악보 이미지를 대상으로 수행되는 수식(1)의 탐색 메커니즘은 [Algorithm 1]의 방법을 통해 탐색의 효율성을 개선한다. 이는, 수식(1)을 적용하기 이전의 전처리(Pre-Process)로 전체 이미지의 Height(H)에서 탐색영역을 점진적 하강 Shift를 통해 오선보(Staff)가 모두 명확히 나타난 좌표(X, Y)와 크기를 추출하는 개념이며, 이후 [Fig 2]의 형태로 악보를 읽는 원리를 갖는다.

$$\text{if } (R(x, y) \equiv 0) \text{ ORDERBY } I(x + x') \quad (2)$$

수식(2)은 매칭에 성공하였을 때, 음의 순서를 찾는 방법론을 의미한다. $R(x, y) \equiv 0$ 의 의미는 수식(1)에서

나타난 $T(x', y') - I(x + x', y + y')$ 의 결과로써, 탐색 대상(T)에서 탐색영역을 감산하였을 때, 0이 발생하는 경우를 같은 형태를 찾은 것으로 정의된다. 이처럼 탐색에 성공하면, ORDERBY $I(x + x')$ 는 탐색된 x 좌표를 기준으로 하여 반복문을 통해 연속적으로 수행되는 매칭에서 정확한 음의 순서를 찾는다. 즉, $Note_0 \dots Note_k \dots Note_n$ 와 같이 각각의 음이 악보에 존재하는지를 반복문을 통해 모두 대입하여 수식(1)으로 찾고, 수식(2)을 통해 그 순서를 찾는 원리이다.

수식(3)은 수식(1)과 같은 메커니즘이지만, 정규화 계수를 구하여 추가로 나눈 방식이다.^[10] 정규화 계수를 적용한 해당 방법은 잡음(Noise)과 밝기와 빛에 영향에 의한 명암이 있는 이미지에서의 탐지 오류를 크게 줄여준다는 연구결과가 있다.^[10] 따라서, 이러한 결

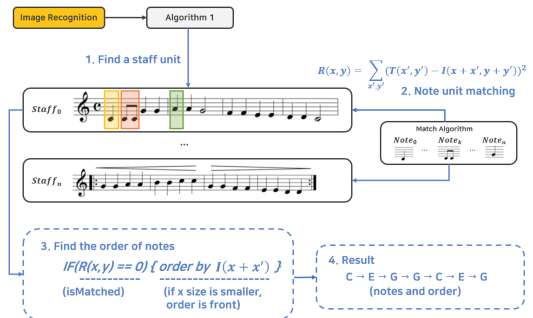


그림 2. 이미지 인식(IR)을 통한 정보 추출의 구조
Fig. 2. The Architecture of Extract data through IR

Algorithm 1 : Read to Sheet Music

1. **Input:** Image(X, Y, W, H), Search_Area SW, SH

Loop:

area = (X, Y, SW, SH)

C = Image.crop(area)

x, y, w, h = C(X, Y, W, H)

2. $R(x, y) = \sum_{x', y'} (T(\text{staff}[\equiv]) - C(x + x', y + y'))^2$

IF($R(x, y) \equiv 0$) {

matching($\overset{\text{Notes}}{\equiv} - \overset{\text{Notes}}{\equiv} - \overset{\text{Notes}}{\equiv}$) as like [Fig 3]}

ELSE { $Y += j$ }

과는 명암이 생긴 악보의 스캔본이나 잡음(Noise)이 있는 이미지 환경에서 유용하게 활용될 수 있다.

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}} \quad (3)$$

템플릿 매칭의 정확도(Accuracy)를 보장하기 위해서는 탐색 대상(T)의 크기를 고려하여 이미지의 크기를 조정하는 선 처리를 구현하는 것이 좋다. 또한, 탐색 영역(I)의 크기 또한 환경에 맞게 유의하게 변화시키는 것이 좋다.

3.3. 학습 구조

본 연구는 음악을 학습하고 제작하기 위한 학습 환경을 각기 다른 두 가지의 actor를 활용하는 구조를 통해 고안하였다.

3.3.1 다음 음을 만드는 학습 과정 (actor 1)

첫 번째 actor는 Markov chain을 기반으로 현재 음(note)을 기준으로 다음 음(note)을 학습하고 추론하는 모델을 제안한 선행연구^[11]를 기반으로 구현하였다. 이러한 모델은 [Fig 3]과 같은 상태 전이도(state transition diagram)로 표현되며, 각 상태(state)에서 상태 전이 확률(state transition probability)을 기반으로 다음 상태(state)로 전이하는 것을 본 연구에서는 행동(action)으로 정의한다.

이처럼 상태 전이 확률(state transition probability)을 기반으로 하는 해당 모델은 [Table 1]과 같은 행렬을 갖는다. Action spaces는 12개의 음(C-B)으로 구성되며, 상태(state) 또한 같은 개수를 갖는다. 따라서, 해당 모델은 정해진 action spaces 내에서 전이 확률에 근거하여 다음 음을 찾는 통계/확률적 사실에 입각

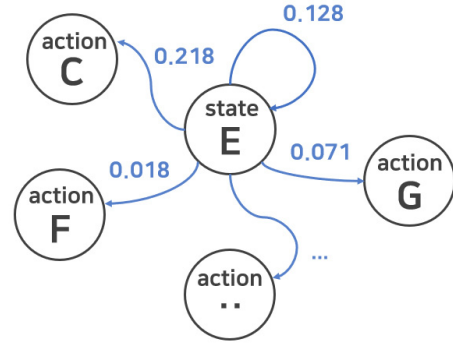


그림 3. 음(note)의 상태 전이도
Fig. 3. The state transition diagram of note

한 음계를 형성하는 결과가 나타난다. 본 연구는 이러한 음계를 4마디 단위로 제한하여 생성토록 하였으며, 이러한 결과물은 두 번째 actor의 action space로써 활용된다.

[Fig 4]는 Actor 1이 상태 전이 확률을 기반으로 4

Table 1. The Example of state transition matrix

state \ next state	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
C	2	1	4	1	6	2	1	3	0	4	2	5
C#	3	0	2	2	1	0	0	0	1	0	0	2
D	7	8	5	4	7	3	0	2	0	6	1	3
D#	0	0	3	1	3	4	1	1	2	0	1	1
E	7	3	7	5	5	4	1	5	1	4	0	2
F	3	2	7	3	6	6	4	4	2	2	1	2
F#	1	0	1	0	2	5	1	7	3	1	2	0
G	5	1	6	1	9	7	6	4	5	1	1	2
G#	0	0	0	1	0	2	3	6	0	4	2	1
A	1	0	2	2	6	8	2	6	6	5	5	4
A#	0	0	0	0	2	1	3	4	2	4	1	6
B	1	0	6	1	2	1	1	8	1	7	5	4

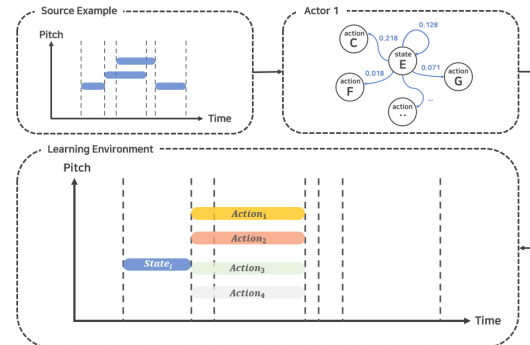


그림 4. 첫 번째 Actor의 강화학습 원리
Fig. 4. The First actor's Principles of RL

마디 단위의 음계가 생성하는 원리를 의미한다. 앞서, 이미지 인식을 통해 악보의 정보를 가로축을 시간 (Time), 세로축을 음높이(Pitch)로 표현하면 Sample Example과 같은 형태를 보인다. 이후, 하나의 음 (Note)을 상태(State)로 정의하고 다음 음을 Markov chain을 통해 확률에 근거하여 선정하는 Action이 4 회 반복적으로 수행된다. 이러한 과정에서 확률론 (Probability theory)에 근거하여 낮은 확률 또한 선정 될 가능성이 있으므로, 학습된 음계를 기반으로 하면서도, 대체로 4마디 음계를 반복적으로 생성할 수 있다.

3.3.2 멜로디 학습 및 생성 과정 (actor 2)

두 번째 actor는 직접적인 결과물을 생성하는 모델로서, 멜로디를 학습하고 생성하는 강화학습의 구조를 정의하였다. Training Features는 [Table 2]에 기재된 내용과 같으며, [Fig 4]와 같은 학습 환경으로 정의한다.

본 모델은 [Fig 5]와 같이 actor 1에서 생성한 값은 action spaces로 활용한다. 따라서, Action은 action spaces에 포함된 4개의 음을 갖는 1마디를 사용하는 것으로 정의한다. State는 현재까지 생성된 멜로디로 정의한다. 이러한 정의에 따라, $state_0$ 는 4개의 음을 갖는 1마디가 되며, $state_1$ 은 8개의 음을 갖는 2마디가 된다. 이러한 원리에 따라, 최종 State는 총 16개의 음을 갖는 4마디를 가지게 되며, 원활한 학습을 위해 문자열 타임을 활용한다. Reward는 두 문자열의 유사도를 판단하는 알고리즘인 Levenshtein Distance¹²⁾를 기반으로 유사도에 근거한 보상 구조와 인간의 직관(Heuristic)에 근거한 기타 보상 구조를 갖는다.

먼저, 유사도 기반 보상 구조는 현재 $state_i$ 의 멜로디를 기존 데이터의 멜로디와 반복적으로 대입하여 유사도를 측정한다. 이때, $state_i$ 의 음의 개수와 비교

대상은 같은 수를 사용하는 것이 중요하다. 이러한 유사도 측정의 결과는 직접적인 보상 구조로 활용된다. 만일, 두 문자열이 완전히 일치한다면 완전한 표절이 되기 때문에 부정적 보상(-1)을 부여한다. 또한, 유사도가 기준치보다 높은 멜로디는 표절이 될 수 있다는 점과 특정 기준치보다 낮은 멜로디는 완전히 새로운 결과물이기 때문에, 해당 두 상황에서는 적은 보상을 부여한다.

본 연구에서는 명확한 정답과 기준이 없는 음악이라는 분야에서의 학습의 목표를 기존 음악과 60~70% 정도의 유사성을 지닌 결과물로 정의하였다. 따라서 해당 범위에 속한 $state_i$ 에게 가장 높은 보상을 부여하는 것으로 학습의 방향성을 유도한다. 나머지는 유사도에 근거하여 보상의 편차를 적절하게 부여하는 것이 중요한데, 반드시 너무 과도하게 큰 값이 부여되지 않도록 조정해야 한다. 인간의 직관(Heuristic)에 근거한 기타 보상 구조로는 중복되는 멜로디가 발생하면, 부정적 보상(-1)을 부여하였다. 만일, 음악적 지식을 해당 모델에 부여하고자 한다면, 이러한 접근방법으로 구현할 수 있다.

[Fig 6]은 Actor 2가 강화학습을 통해 16개의 음을 갖는 음계를 생성하는 과정과 원리를 의미한다. 이처럼 각각의 State에서 유사성에 따라 보상을 부여하는 구조는 일종의 테트리스 게임환경에서의 강화학습과 유사한 맥락을 갖는다. actor는 action spaces에서 무작위로 action이 수행될 때마다, 현재까지 누적된 음계의 상태를 기존의 멜로디와 비교한다. 이러한 과정은 표절이 되지 않는 선으로(60~70%) 유사성이 있는 결과물을 얻을 때까지 반복된다. 이때, history queue 혹은 Q-table과 같은 자료구조에 state와 action 그리고 reward를 저장하고 이를 Action 실행 시마다 검증

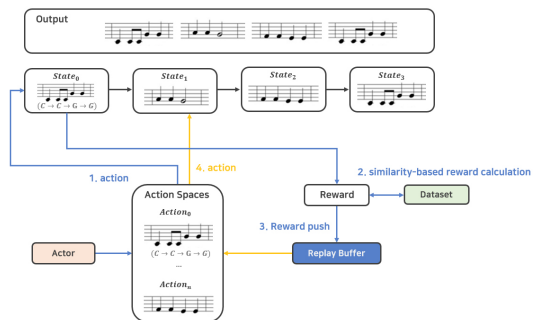


그림 5. 두 번째 actor의 강화학습 환경
Fig. 5. The Learning environment of second actor

Table 2. Proposed training features for Learning

Category	Contents
State	현재까지 생성된 멜로디
Action	4개의 음을 갖는 1마디 eg. C→E→F→C#
Rewards	1) Levenshtein Distance ¹²⁾ 기반 유사도에 근거한 보상 구조 - 생성된 n 마디가 기존 데이터와 일치하면, reward = -1 - 생성된 n 마디가 기존 데이터와 유사하면, reward = s(0.1~0.9) - 생성된 n 마디가 기존 데이터에 존재하지 않으면, reward = (0.1~0.3) 2) 기타 보상 구조 - 마디와 마디 간의 데이터가 중복되면, reward = -1

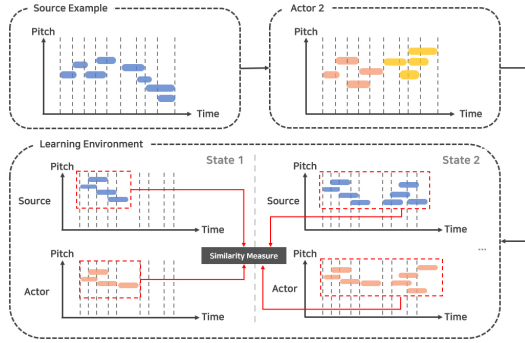


그림 6. 두 번째 Actor의 강화학습 원리
Fig. 6. The Second actor's Principles of RL

한다면, e-greedy 하게 확률을 줄여나가며 학습이 진행된다.

IV. Experiments

본 연구는 제안하는 모델의 PoC(Proof of Concept)를 실현하기 위한 실제 구현을 수행하였다. 이때, 결과물은 업계에서 보편적으로 활용되는 MIDI(Music Instrument Digital Interface)^[13] 형식으로 가공한다.

먼저, 악보 이미지로부터 음을 추출하는 과정에서는 보편적으로 높은 정확도(Accuracy)를 보였지만, 탐색 영역(I)의 크기와 탐색 대상(T)의 크기를 유의하게 조절하는 것이 탐지 성능과 큰 상관성을 보였다. 알고리즘의 특성상 탐색 대상(T)의 회전이나 크기에서 유의한 변화가 있을 때 계산 결과에서 0이 아닌 차이 값이 발생하였고, 이러한 문제는 조건문의 오류를 초래할 수 있다. 또한, 탐색 영역(I)을 탐색 대상(T)과 반복적으로 대조하는 과정에서 속도 문제와 지연이 발생하였다.

본 연구는 이러한 문제를 개선하는 방법으로 [Fig 7]의 프로세스(Section Crop)를 추가로 구현하였다. 이는, 이미지의 세로 크기를 기반으로 수행되는 [Algorithm 1]과 유사한 원리로 이미지의 가로 크기를 기반으로 탐색을 수행하여 하나의 staff로부터 각 마디를 나눈다. 이처럼 분할된 마디를 대상으로 템플릿 매칭을 반복적으로 수행하는 것이 탐색 영역(I)의 크기를 조절하는 것과 속도 문제의 측면에서 훨씬 적합하였다.

[Fig 7]의 개선방안이 적용된 템플릿 매칭의 결과와 기존 템플릿 매칭의 결과는 [Table 3]처럼 요약될 수 있다. 기존 Staff 단위의 매칭보다 Section 단위의 매칭이 탐색영역이 1/4 정도로 적기 때문에, 5회 반복

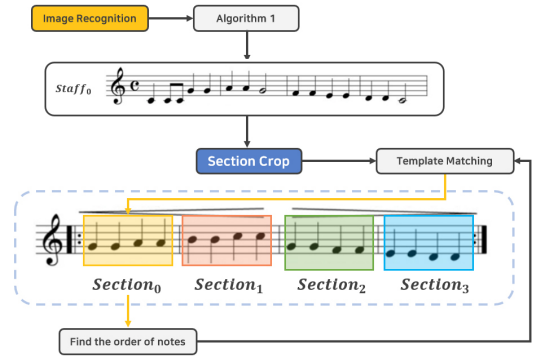


그림 7. 템플릿 매칭을 개선하기 위한 방법
Fig. 7. A Method for improvement of template matching

측정의 평균값에서 일정한 속도 차이를 보였다.

[Fig 8]은 강화학습 모델을 통한 음악창작의 결과를 MIDI 형식으로 생성하여 시각화한 그림이다. 장르와 인간의 개별적 취향을 반영하는 음악의 특성상 결과가 더욱 좋아지고 나뻘짐을 명확히 평가하고 도식화할 근거가 없으므로, 60~70% 유사한 결과물을 학습 종료 시점으로 정의하였다.

학습이 완료된 후, actor 1의 상태 전이 행렬(state transition matrix)을 확인한 결과, 전이 확률이 0%인 행동(action)이 다수 발견되었다. 이러한 결과는 음악이라는 장르에서도 특정 음에서 특정 음으로는 전이될 가능성이 없는 상태가 있음을 시사한다. 이는, 음악적 지식과 이론에 근거하여, 확률이 0%인 상태 전이 확률을 사전에 제외하고 구현한다면 복잡한 확률 문제를 간소화하고 개선할 수 있다는 가능성을 보인다.

본 연구에서 제시한 모델은 음악이라는 장르에 강화학습을 접목하기 위한 초기 단계의 연구로, 음(note)의 지속시간(duration)과 타이밍(timing) 그리고 세기

Table 3. Time Required to template matching

Method	Matching Process	Time Required (5-Times Average)
Original Template Matching	 Note ₀	0.0428788960654774
Improved Template Matching	 Note ₀	0.0070219108

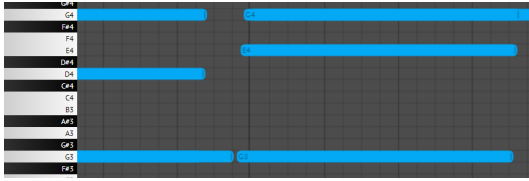


그림 8. 학습 환경에 의해 MIDI 형식으로 생성된 결과물
Fig. 8. The output in MIDI format composited by Learning Environment

(dynamics)는 고려하지 않았다. 음악의 창작은 그 근본이 음의 배열이지만, 음의 지속시간과 타이밍을 통한 박자를 생성하는 것도 매우 중요하다.

따라서, 현재 제시한 모델을 추가로 확장하기 위해 [Fig 9]의 측정방법을 제안한다. 이는 음의 지속시간, 세기, 타이밍을 강화학습에 추가로 접목하여 음악을 더욱 풍부하게 보완하고 실용성을 증가시킨다.

[Fig 9]의 측정방법에서 음의 타이밍과 지속시간은 Actor 1의 상태 전이 행렬을 고도화하는 것으로 구현할 수 있다. [Table 4]는 고도화된 상태 전이 행렬의 예시를 의미한다. 각 상태는 $C(i,n)$ 와 같은 형태로 두 개의 파라미터를 갖도록 개선되며, 하나는 타이밍 정보를 나머지는 지속시간의 정보를 의미한다.

$C(2, 0.5)$ 를 예시로 한다면, 2라는 인덱스의 타이밍에서 C의 음을 0.5초간 지속한다는 의미이며, 해당 상태(state)에서 다음 상태로 전이를 수행할 확률을 구하는 것이다. 파라미터의 값은 악보를 읽어서 음을 찾는 이미지 인식 기능과 같은 Depth에서 별도로 소리 인식(Sound Recognition) 혹은 측정 알고리즘을 구현해야 한다. 이러한 방식으로 개선된 Actor 1은 더욱 개선된 4마디 음절을 만들 수 있으며, 이를 action spaces로 활용하는 Actor 2의 결과물은 더욱 실용적일 것으로 예측된다. 이처럼 개선된 모델은 추후 후속 연구를 통해 원리와 결과를 검증과 고찰을 통해 진행할 예정이다.

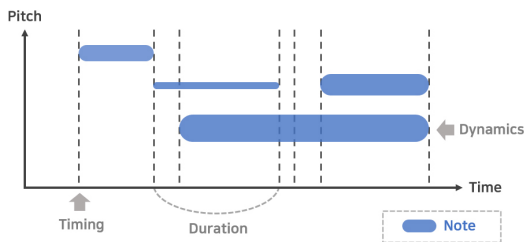


그림 9. 음의 지속시간, 세기 타이밍의 측정방법
Fig. 9. The Measurement method of duration, dynamics and timing in note.

Table 4. The Example of improved state transition matrix

next state \ state	$C(i,n)$	$C(j,m)$	$C(k,o)$
$C(i,n)$	0	2	0
$C(j,m)$	3	0	1
$C(k,o)$	1	3	4

V. Conclusion

본 연구는 이미지 인식(Image Recognition)과 강화학습(Reinforcement Learning)을 통해 음악적 지식과 인간의 직관(Heuristics)을 배제하며, 음악을 창작하는 알고리즘 작곡(algorithmic composition)에 관한 연구를 수행하였다. 제안하는 학습 모델은 음악이라는 장르에 강화학습을 접목하려는 시도의 초기 단계의 연구로서 음의 지속시간이나 타이밍을 고려하지 않았다는 일부 한계점이 있지만, 각기 다른 두 가지의 actor를 정의하고 활용하는 구조를 고안함으로써, 기존 마르코프 체인(Markov chain)을 통한 접근법보다 더욱 심도 있는 접근 방향을 제시하였다. 또한, 소리의 파형이나 형태에 초점을 맞추는 소리 인식 기반 강화학습의 다른 모델들과는 달리 이미지 인식 관련 기술을 활용하여 악보의 정보를 추출하는 차별화된 새로운 시도를 수행하였다. 이러한 시도는 향후 더욱 고도화를 위한 초석이 되며, 이를 개선하기 위한 후속 연구의 방향성을 제시한다고 생각한다.

References

- [1] J. H. Kim, "Composition as reflection of knowledge and science-composers cognitive mind in the human integrative mind," in *Daegu Catholic Univ. Res. Inst. Humanities and Sci.*, no. 23, pp. 143-198, 2014.
- [2] A. Robert, Y. Mann, and J. Engel, "Magenta Studio," Retrieved Oct. 18, 2019, from <https://magenta.tensorflow.org/studio-announce>
- [3] Ka. Ching Chan, "An introduction to Markov chains," *The MAV 49th Annu. Conf.*, 10.13140/2.1.1833.8248, 2012.
- [4] H. F. Olson and H. Belar, "Electronic music synthesizer(1955)," *The J. Acoustical Soc. Am.*, vol. 27, no. 3, Jun. 2005.
- [5] W. J. Lee, "A study on the real-time

performance system based on algorithmic composition(2019),” Retrieved Oct. 22, 2019, from “dspace.ewha.ac.kr”

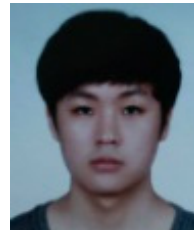
- [6] B. Kim and W. S. Yeo, “Probabilistic prediction of rhythmic characteristics in markov chain-based melodic sequences,” *ICMA*, pp. 429-432, 2013.
- [7] J. Y. Kwon and I. K. Lee, “Music composition using markov chain and hierarchical clustering,” *HCI Soc. Korea*, pp. 744-748, Feb. 2008.
- [9] I. Lee, N. Choi, and I. J. Kim, “Music score recognition system using line tracking and template matching,” *KIISE*, vol. 34, no. 1, pp. 417-421, Jun. 2007.
- [10] K. Briechle and D. Hanebeck, “Template matching using fast normalized cross correlation,” in *Proc. SPIE 4387, Optical Pattern Recognition XII*, 2001.
- [11] J. Pickens and C. S. Iliopoulos, “Markov random fields and maximum entropy modeling for music information retrieval,” *6th Int. Conf. Music Information Retrieval*, London, UK, Jan. 2005.
- [12] L. Yujian and L. Bo, “A normalized levenshtein distance metric,” *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 29, no. 6, pp. 1091-1095, Jun. 2007.
- [13] W. G. Busse, “Toward objective measurement and evaluation of jazz piano performance via midi-based groove quantize templates,” *Interdisciplinary J.*, vol. 19, no. 3, pp. 443-461, Spring, 2002.
- [14] C. Roads and J. Strawn, “*The computer music tutorial*,” MIT Press Cambridge, MA, USA, pp. 879-880, 1996.

정택현 (Tack-hyun Jung)



2018년~현재 : 건국대학교 IT융
합정보보호학과 석사과정
<관심분야> 사이버 보안, 인공
지능, 컴퓨터비전
[ORCID:0000-0002-9172-0817]

이승호 (Seung-ho Lee)



2019년~현재 : 건국대학교 컴퓨
터정보통신공학과 석사과정
<관심분야> 블록체인, 사이버
보안, 인공지능
[ORCID:0000-0002-8142-5136]

김기천 (Keecheon Kim)



1992년 : Northwestern Univ.
공학박사
1998년~현재 : 건국대학교 컴퓨
터공학과 교수
<관심분야> 통신공학, 사이버
보안, 미래인터넷, IoT
[ORCID:0000-0003-3445-3334]