

# 화성 진행 학습 모델을 적용한 규칙 기반의 4성부 합창 음악 생성

조원익\*, 김정훈\*, 천성준\*, 김남수<sup>o</sup>

## Rule-Based Generation of Four-Part Chorus Applied With Chord Progression Learning Model

Won Ik Cho\*, Jeung Hun Kim\*, Sung Jun Cheon\*, Nam Soo Kim<sup>o</sup>

### 요약

본 논문에서는 규칙 기반의 4성부 합창 음악 생성 과정에 화성 진행 학습 모델을 적용해 보고자 한다. 제안하는 시스템은 32음의 멜로디를 입력으로 받아 다른 세 성부를 화성학의 규칙에 맞게 완성시켜 주며, 그 과정에서 사용하는 화성 진행을 CRBM 모델을 이용하여 예측한다. 학습 데이터는 화성학 교육 자료집에서 다수 발췌하였으며, 화성 진행을 조성에 독립적으로 추출하여 주어진 데이터를 효과적으로 활용할 수 있도록 하였다. 학습 모델을 적용한 결과물이 기존의 규칙 기반 4성부 합창 음악에 비해 보다 자연스러운 진행을 보임이 확인되었다.

**Key Words** : automatic composition, four-part chorus, rule-based approach, machine learning, conditional restricted Boltzmann machine

### ABSTRACT

In this paper, we apply a chord progression learning model to a rule-based generation of a four-part chorus. The proposed system is given a 32-note melody line and completes the four-part chorus based on the rule of harmonics, predicting the chord progression with the CRBM model. The data for the training model was collected from various harmony textbooks, and chord progressions were extracted with key-independent features so as to utilize the given data effectively. It was shown that the output piece obtained with the proposed learning model had a more natural progression than the piece that used only the rule-based approach.

### I. 서론

4성부 합창은 소프라노, 알토, 테너, 베이스(SATB)의 네 성부로 구성되어 있는 음악 형식이다. 그 근간을 중세 교회 음악에 두고 있어 매우 엄격한

규칙을 준수하며, 조성, 음계, 화음, 화성 진행과 종지의 개념이 잘 드러나 있어 화성학 교육 과정에서 많이 사용된다. 이 경우 하나의 성부를 주고 다른 세 성부를 완성하는 형태의 과제가 주로 수행되는데, 이 과정을 인공 지능을 통해 자동화하는 연구가 자동 작곡 연

※ 이 연구는 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2015R1A2A1A15054343).

• First Author : Seoul National University Department of Electrical and Computer Engineering and Institute of New Media and Communications, wicho@hi.snu.ac.kr, 학생회원

o Corresponding Author : Seoul National University Department of Electrical and Computer Engineering and Institute of New Media and Communications, nkim@snu.ac.kr, 종신회원

\* Seoul National University Department of Electrical and Computer Engineering and Institute of New Media and Communications, jhkim@hi.snu.ac.kr, sjcheon@hi.snu.ac.kr, 학생회원

논문번호 : KICS2016-09-250, Received September 9, 2016; Revised October 10, 2016; Accepted November 18, 2016

구 분야의 하나로 진행되어 왔다<sup>11</sup>.

이 분야에서 오랫동안 사용되어 온 방법은 규칙 기반의 생성 방식이다. 규칙 기반의 생성 방식은 계산 속도가 빠르고 화성학의 규칙에 잘 들어맞는다는 장점이 있으나, 실제 작곡가들이 작곡하는 방식을 재현하는 데에 한계가 있다. 사람이 작곡을 할 때는 이전에 학습했거나 주로 사용하였던 화성 진행이나 멜로디 진행을 사용하게 되는데, 규칙 기반의 생성 방식에서는 이러한 경향을 충분히 활용하기 어렵기 때문이다. 이러한 단점을 보완하기 위해 genetic algorithm (유전 알고리즘, GA)을 이용하거나<sup>12</sup> 화성 진행을 점수화시켜 정형화된 진행이 더 잘 선택되도록 설계할 수는 있지만, 이 방법은 모든 경우에 대해 경향성을 설정해 둘 수 없다는 한계가 있다.

이러한 특징으로 인해, 최근의 자동 작곡 연구에서는 통계적 방법론이나 심화 신경망 학습 모델이 많이 사용되고 있는 추세이다. hidden markov model (은닉 마르코프 모델, HMM)이나<sup>13</sup> recurrent neural network (재귀 신경망, RNN)을 사용한다<sup>14</sup> 자동 작곡 양식을 확인할 수 있으며, 4성부 합창 자동 작곡 연구에서는 Markov decision process (MDP)을 사용하여 그림 1과 같은 음악을 얻을 수 있다고 보고된다<sup>15</sup>.

본 논문에서는 conditional restricted Boltzmann machine (CRBM)을 이용한 화성 진행 학습 모델을 통해 보다 자연스러운 화성 진행을 가진 4성부 합창 음악을 생성해 보고자 한다. 한 시점의 4성부 음을 묶어 프레임이라 정의하자. CRBM의 학습 과정에서는 현재 프레임보다 이전의 데이터들을 추가적으로 사용하여 다음 프레임의 특징 벡터를 결정하는 과정을 학습한다. 이 때 사용되는 데이터는 조성에 독립적으로 수집된다. 이러한 데이터를 활용한 특징 추출 방법은

각각의 데이터의 크기를 줄여 주며 조성을 모두 반영한 학습 과정보다 적은 양의 데이터로 학습이 가능하게 한다. 이렇게 학습된 모델을 기존의 규칙 기반 4성부 합창 음악 생성<sup>6</sup>에 적용하여, 보다 자연스러운 화성 진행을 구현할 수 있게 하였다.

## II. 규칙 기반의 화음 후보 생성

4성부 합창에서 중요한 것은 주어진 조성에 맞는 화음을 사용하는 것이다. 따라서 우리가 사용할 화음은 주어진 조성에서 나올 수 있는 유한한 크기의 화음 집합에서 선택되어야 한다. 본 논문에서 다루는 화음은 3화음과 7화음, 그리고 그 전위형으로 제한한다. 나폴리 화음이나 증화음의 경우는 다루지 않을 예정이며, 전조와 부화음의 경우는 최소 두 프레임 이상이 쌓을 이를 경우를 생각해야 하는데 이는 아직 구현의 어려움이 있어 논의에서 제외하였다.

### 2.1 화음 공간 생성

화성학에서 다루는 조성은 총 서른 가지이다. 이 가운데 중복되는 것을 제외하면, 장단조 여부와 으뜸음이 무엇이냐에 따라 스물네 가지의 조성을 얻는다. 하나의 조성이 결정되면 그에 따른 음계가 결정되는데, 여기서는 모든 단음계에 대해 화성 단음계를 사용하기로 한다. 결정된 음계는 7개의 음으로 구성되어 있고, 이 중 어떤 것을 근음으로 하나에 따라 몇 도 화음 인지가 결정된다. 하나의 근음에서 나올 수 있는 화음은 3화음 3개 (기본형, 1전위, 2전위)와 7화음 4개 (기본형, 1전위, 2전위, 3전위)로 총 7가지이며 따라서 총 49가지 경우를 생각해 볼 수 있다. 각 화음을 구성하는 음들을 순열로 배열하는 방법을 통해 음을 결정하는데, 이 때 음들을 하나의 옥타브 안에 사상하는 방법은 music information retrieval (음악정보검색, MIR)에서 사용되는 크로마와 같은 개념이다<sup>17</sup>. 물론 4성부 합창에서는 3화음 중 하나의 음을 중복해 넣어야 하지만, 전위형은 근음이 무엇이냐에 따라 결정되므로 위의 세 가지에서 벗어나지 않는다.

### 2.2 4성부 음 생성

다음에 올 4성부 음 생성을 위해서는 2.1에서 생성된 화음 후보들의 배열이 필요하다. 이 때 현재 프레임의 음 정보가 반영되는데, 한 옥타브 이상의 음 도약을 피하는 4성부 합창의 특징을 반영하여 한 성부의 음이 특정 옥타브에 배치될 경우의 수는 세 가지로 제한한다. 위와 같은 방법으로 중복을 고려하지 않고

### Ode to Joy



그림 1. MDP를 사용한 4성부 합창의 예시  
Fig. 1. An example of four-part chorus generation using MDP

셀 때, 하나의 화음 후보 당 81가지의 4성부 음 후보가 나오게 된다.

### 2.3 화성학의 규칙에 따른 후보 제거

화성학의 규칙은 크게 두 가지 측면에서 바라볼 수 있다. 4성부의 음을 한 프레임에서 수직적으로 보는 방법과 진행의 관점에서 수평적으로 보는 방법이다. 본 논문에서 적용한 화성학의 규칙들은 다음과 같다 [6,8].

- 1) 각 성부가 서로의 영역을 침범해서는 안 된다.
- 2) 장조의 경우 3음의 중복을 피한다. 다만 소프라노와 알토, 혹은 소프라노와 테너 간에는 허용한다.
- 3) 인접한 두 성부 간 한 옥타브보다 큰 차이가 나서는 안 된다. 다만 베이스와 테너 간에는 허용한다.
- 4) 시작 화음과 마지막 화음은 1도로 같게 하며, 중지에서 생략이 필요하다면 5음을 생략한다.
- 5) 반중지는 5도 화음으로 한다.
- 6) 모든 성부 간의 병행 1도, 5도, 8도를 금한다.

규칙 1) ~ 3)의 내용은 4성부의 음을 수직적으로 보는 방법, 4) ~ 6)의 내용은 수평적으로 보는 방법으로 해석할 수 있다. 2.2에서 얻은 4성부 음 후보군을 위의 규칙들을 반영하여 수정한다. 일반적인 4부 합창은 베이스가 c2보다 낮거나 소프라노가 c6보다 높은 경우는 취급하지 않기 때문에, 이에 해당하는 경우는 추가적으로 제외하도록 하였다.

## III. 특징의 추출 및 학습

2.3에서 얻은 후보군 중에서 적합한 진행을 선택하는 과정에서는 4성부 합창 데이터를 이용해 학습한 CRBM의 매개변수 집합이 이용된다.

### 3.1 CRBM에 사용되는 특징의 추출

2층 구조인 restricted Boltzmann machine (RBM) 이 한 프레임의 visible layer를 통해 hidden layer의 latent variable을 추정한다면, CRBM에서는 과거의 visible layer들을 추가적으로 이용하여 이후에 나올 정보를 추정한다<sup>9)</sup>. 화음 진행을 결정하는 데에 있어 RBM보다 CRBM이 더 효과적인 이유는, 바로 앞의 화음뿐 아니라 그 앞의 화음들이 어떤 방식으로 진행되어 왔는지도 화성 진행에 영향을 미치기 때문이다. 예컨대 2도 화음 뒤에 어떤 화음이 적절할지 결정할

때 단순히 이전 데이터에서 확률적으로 어떤 화음이 가장 많이 나왔나 살펴보면 5도 화음이 되겠지만<sup>10)</sup> (4도-2도-5도 진행) 만약 그 이전에 1도 화음이 있었다면 뒤에도 1도 화음이 나올 가능성이 높다고 (1도-2도-1도 중지) 예측해볼 수 있다.

CRBM 학습에 사용할 입력 데이터는 최대한 조성에 독립적인 성분들로 결정하여 가지고 있는 데이터 베이스를 효과적으로 활용하였다. 하나의 조성이 결정되면 그 안에서의 화성 진행은 모두  $n$ 도에서  $m$ 도 화음으로 진행되는 것( $1 \leq n, m \leq 7$ )으로 볼 수 있기 때문에 4성부의 음 높이라는 큰 데이터를 생략할 수 있다. 이렇게 하여 우리가 활용할 정보를 몇 도 화음인지, 3화음인지 7화음인지, 그리고 어떤 전위형을 갖는지 크게 세 가지 정도로 줄일 수 있으며, 이는 6차원의 binary vector로 나타낼 수 있다. 본 논문에서는 CRBM이 다루는 정보를 과거 세 프레임까지의 visible data로 잡았으므로, 총 세 프레임의 입력 데이터와 한 프레임의 출력 데이터를 가지고 학습하였다.

### 3.2 CRBM의 학습

CRBM의 학습 과정은 RBM의 학습 과정과 유사하다. 기존의 RBM에서 학습하는 매개변수는 weight parameter  $W$ 와 bias parameter  $a, b$  총 세 가지이다. 본 논문에서 사용되는 CRBM에서는 과거의 세 프레임이 concatenate되며, 이를 past data로 하여 그림 2에서처럼 past-hidden을 연결하는 weight와 past-visible을 연결하는 weight  $A, B$ 가 추가되고 bias parameter들도 또한 past data를 반영한 모델로 수정된다. 최종적으로는  $W, A, B, a, b$ 의 매개변수를 구하여 이용하게 된다.

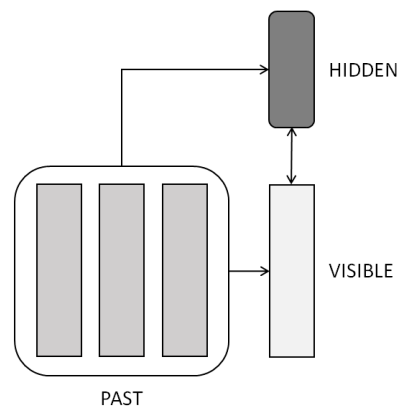


그림 2. 사용되는 CRBM 모델  
Fig. 2. CRBM training model

#### IV. 화음 예측과 최종 결정

우리가 하고자 하는 것은 특정 멜로디를 기반으로 4성부 합창 음악을 완성하는 것이다. II, III에서 다른 내용을 이용해 악보를 생성하는 과정은 그림 3과 같이 이루어진다. 특정 멜로디가 위치할 성부와 첫 4성부 음, 그리고 조성이 주어지고, 첫 마디를 제외한 세 번째 마디까지의 음은 [6]에서 제안한 방법을 통해 결정된다. 이후의 과정은 CRBM 매개변수를 이용해 다음에 나올 화음의 조성-독립적인 특징( $X$ )을 예측하고 이를 2.3에서 얻은 후보군의 특징( $Y_j$ )들과 비교하여 최종 결정을 내리는 단계를 반복한다. 여기서  $j$ 는 후보의 인덱스를 의미한다.

최종 결정을 위한 scoring system은 다음과 같이 정의된다.

- 1) 2.2에서 얻은 후보군 모두에게 5점의 점수를 주고 시작한다. 2.3의 규칙에 어긋나는 후보들은 0점으로 만들고, 0점이 아닌 후보군에서 추출한 특징  $Y_j$ 들을 대상으로 이후의 단계를 시행한다.
- 2)  $X$ 와  $Y_j$ 의 일치도에 따라 점수를 더해 준다. 즉,  $X_i$ 와  $Y_{j,i}$ 가 같으면 1점을 더해 주고 그렇지 않으면 0점을 더해 주는 것이다. 여기서  $1 \leq i \leq 6$ 으로  $i$ 는 특징 벡터의 엔트리를 의미한다.  $1 \leq i \leq 3$ 일 경우는 도수,  $i = 4$ 일 경우는 3/7화음 여부,  $5 \leq i \leq 6$ 일 경우는 전위형을 의미하는 binary variable이다.
- 3) 점수를 더해 주는 과정에서 weighted hamming distance의 개념을 사용하게 된다. 진행을 결정

하는 데에 있어 도수 데이터가 3/7화음 여부나 전위형보다 더 지배적인 역할을 하므로<sup>[8]</sup>  $1 \leq i \leq 3$ 일 경우 0.6의 weight를 곱해 주었고, 그렇지 않은 경우 0.2를 곱해 주었다. 이런 원리로  $X$ 와  $Y_j$ 가 같으면 총 2.4점의 점수가 더해지게 된다.

- 4) 추가적으로, 베이스의 도약을 막기 위해 현재 프레임의 베이스음과 후보군 베이스음의 도수 차이의 역수만큼을 각 후보에 더해 준다.

이렇게 하여 얻은 후보군 중에서 점수가 가장 높은 후보를 골라 업데이트하고, 같은 과정을 반복하여 32 프레임의 악보를 얻게 된다.

#### V. 실험 및 토의

##### 5.1 실험 세팅 및 비교 모델

본 논문에서는 [8]을 포함해 시중에 존재하는 다양한 화성학 교재에서 3화음과 7화음이 섞여 있는 화성 진행들을 발췌하여 CRBM의 학습을 진행하였다. 각 데이터는 4개의 6차원 binary vector로 구성되어 있으며, 총 1516개의 데이터를 이용했다. hidden layer의 크기는  $60^{[11]}$ , 4개의 implicit mixture CRBM을 이용하였으며<sup>[12]</sup> Gibbs sampling은 30회로, iteration은 500회로 설정하였다. 기존에 배포되어 사용되던 gaussian-binary MATLAB CRBM 코드<sup>[13]</sup>는 binary-binary로 수정하여 수행하였다.

각 프레임의 4성부 음을 표현하는 벡터는 [6]과 같이 각 성부의 옥타브(c1~c7)와 크로마(1~12), 도수(1~7), 3/7화음 여부, 그리고 전위형으로 구성하였다. 예컨대 [5]에서도 이용된 베토벤의 ‘Ode to joy’ 초반부의 소프라노 성부를 이용하여 생성한다면, 조성을 C Maj로 설정하고 [3(옥타브) 1(도), 3 8(솔), 4 5(미), 5 3(미), 1(tonic), 3(화음), 0(전위 없음)]의 형식으로 첫 음을 초기화할 수 있다. 그림 4와 5는 각각 [6]의 알고리즘과 본 논문에서 제안한 알고리즘을 통해 얻은 출력 프레임 시퀀스를 편곡 프로그램 FINALE를 통해 기보한 악보의 일부이다.

보다 정량적인 실험을 위하여, 8개의 소프라노 멜로디와 4개의 베이스 멜로디를 각각 [6]의 알고리즘(A)과 본 논문에서 제안한 알고리즘(B)에 적용해 생성한 총 24개의 작품을 이용하여 mean opinion score (MOS) test를 진행하였다. MOS는 원래 통화품질의 측정에서 나온 개념으로, 500ms 이하의 reverberation

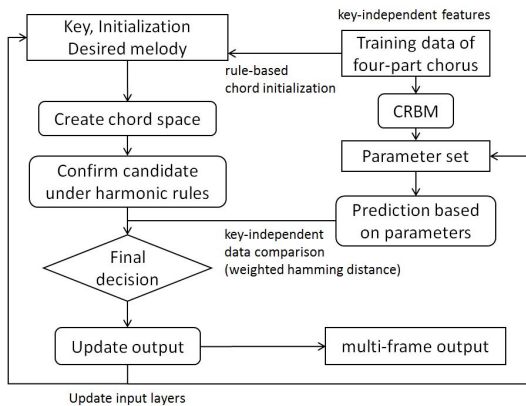


그림 3. 제안하는 생성 알고리즘의 블록도  
Fig. 3. Block diagram for the proposed generation algorithm



그림 4. [6]의 알고리즘으로 생성한 작품의 일부  
Fig. 4. Part of a piece generated by the algorithm of [6]



그림 5. 제안 알고리즘으로 생성한 작품의 일부  
Fig. 5. Part of a piece generated by the proposed algorithm

time을 가진  $30m^3$  이상의 방에서 발화한 음원에 대해 평가하는 것을 원칙으로 한다<sup>[14]</sup>. 하지만 실험 진행자의 주관적인 평가보다는 훨씬 객관적으로 생성된 음악을 평가할 수 있을 것이라 생각하여 적용하였다. 기반이 되는 멜로디는 장조 8개와 단조 4개로 모두 화성학 교재에서 발췌하였으며, 생성된 음악이 듣기 좋고 자연스럽게 느껴질 경우 5점, 불쾌하고 기계적으로 느껴질 경우 1점을 부여하도록 하였다. 실험은 12명의 음향 분야 연구원을 대상으로 진행하였으며, 구성원 중 6명은 음악 활동 경험이 있고 2명은 합창 경험이 있어 본 실험에 적합할 것으로 판단하였다. 가이드라인이 다소 모호할 수 있어 실제 존재하는 4성부 작품을 5점의 예시로 주고, A나 B로 생성된 음악 중 어색함이 크다고 느낀 음악을 1점의 예시로 주었다.

## 5.2 실험 결과 및 평가

실험 결과는 표 1과 같다. 전체적으로 B를 통해 생성된 4성부 합창 음악이 A를 통해 생성된 경우에 비해 좋은 평가를 받았다.

세부적으로 보면, 장조 음악과 단조 음악에서 모두 B가 A보다 좋은 평가를 받았다. 또한 A와 B 모두에

서 장조 음악이 단조 음악보다 더 대중적으로 선호되는 소리를 들려주었음을 확인할 수 있다.

소프라노 멜로디에 적용한 음악의 경우 B가 A보다 좋은 평가를 받았으나, 베이스 멜로디에 적용한 음악의 경우는 그렇지 않았다. 화성 진행에 사용되는 특징들이 전적으로 베이스음에 의해 결정되기 때문에 베이스음이 정해지면 그만큼 화성 진행 학습 모델이 사용될 여지가 줄어들기 때문이다. 또한 전체적으로 베이스 멜로디에 적용한 경우가 낮은 평가를 받은 것은, 소프라노 멜로디의 자유도를 제한하지 않아 고음의 도약이 어색하게 들리는 데에 그 이유가 있는 것으로 보인다.

## VI. 결 론

본 논문에서는 CRBM을 이용한 학습 모델을 화성 진행에 적용한 규칙 기반의 4성부 합창 음악 생성 알고리즘을 제안하였다. 같은 입력 멜로디에 대하여 기존의 규칙 기반 알고리즘에 비해 좀 더 자연스러운 진행을 보임을 확인할 수 있었다.

본 연구의 발전 방향은 크게 세 가지로 정리해 볼 수 있다.

### 1) DB 품질 개선

학습 모델에서 DB의 품질은 빼놓을 수 없는 요소이다. 보통 화성학 교재의 초반은 3화음만을 사용한 단순한 진행들로 연습을 시작하기 마련인데 그 부분을 제외하고 3화음과 7화음이 복합적으로 사용된 작품들만을 발췌한 것은 다양한 화성 진행의 사용으로 데이터가 골고루 분포하게 하기 위해서이다. 실제 작곡가들의 작품을 분석하여 DB의 품질을 개선할 필요가 있다.

### 2) 3화음, 7화음을 넘어선 다양한 화음의 구현

본 연구에서 3화음과 7화음을 대상으로 실험한 이유에는 많이 사용된다는 것 외에 구현의 용이함도 있다. 나폴리 화음이나 증화음은 부가적으로 들어가면 좋지만 화성 진행에 필수적이지 않을뿐더러 비화성음을 포함하고 있기도 하다. 그렇지만 전조나 부화음은 곡의 분위기를 전반적으로 바꿔 줄 수 있다는 점에서 사용자가 그 필요를 느낄 만하며, 3화음과 7화음만으로 낼 수 없는 다양한 소리들을 시도해 볼 수 있게 한다. 전조는 사용자가 원하는 위치와 조성을 지정하여 그 지점에서 새로 초기화를 하는 방식으로 구현 가능하며, 부화음의 경우는 이후 나올 화음에 대한 고려가

표 1. 제안 알고리즘과 [6]의 알고리즘으로 생성한 음악의 MOS 비교  
 Table 1. MOS test of music generated by proposed algorithm and algorithm of [6]

MOS	Algorithm of [6]	Proposed
Total	3.17	3.35
Major	3.30	3.50
minor	3.00	3.14
soprano	3.26	3.60
bass	3.00	2.85

필요하기에 두 개 이상의 프레임을 묶어 생각하는 방식으로 작품을 구현할 수 있도록 다른 알고리즘을 도입하여야 한다. 이 부분은 후속 연구로 남겨 두도록 한다.

3) 보다 객관적인 평가 방법 도입

우리가 작곡가의 음악에 대해 점수를 매기기 어렵듯, automatic composition의 performance evaluation도 쉽지 않고 그런 객관적인 척도도 마련되어 있지 않다. [2]에서는 GA의 fitness graph를 통해 성능을 측정하지만 이것은 GA의 성능 그래프이지 작곡이 잘 되었는지를 판단하기 어려우며, 4성부의 규칙에 잘 들어맞는지 여부는 모든 알고리즘이 잘 지키고 있으니 좋은 척도가 될 수 없다. MIR에서는 음악의 brightness나 key strength, inharmonicity 등 여러 요소들을 추출하여 특징으로 사용하지<sup>[15]</sup> 이들을 보고 듣기 좋은 음악인지 판단하기는 어렵다. 서로 다른 알고리즘으로 작곡된 음악을 평가할 수 있는 보다 객관적인 방법이 개발된다면, MOS test에서도 나타날 수 있는 주관성을 배제할 수 있을 것이다. 이를 위한 연구는 자동 작곡 연구와 밀접하게 연계하여 진행될 수 있다.

이상으로 연구의 세 가지 발전 방향을 살펴보았다. 마지막으로 연구의 활용 방안에 대해 얘기해 보고자 한다.

본 연구가 진행된 배경에는 4성부 합창 음악 생성을 좀 더 쉽게 할 수 있는 방법을 찾는 과정이 있었다. 이는 향후 성능이 개선되고 접근이 용이한 UI가 개발된다면 이러한 작곡 기법이 화성학 교육의 보조 자료로도 사용될 수 있음을 시사한다. 또 음원 분리<sup>[16]</sup>나 음질 향상<sup>[17]</sup> 기법과 연계하여 음원을 취급하는 방법에 새로운 지표를 제시하거나, 구글의 Deep dream 그림을 통해 그러하듯 다양한 스타일의 음악을 만들

어 대중의 흥미를 불러일으키고 새로운 작품을 창조하는 등의 역할도 할 수 있을 것이다.

References

[1] G. Nierhaus, *Algorithmic composition : Paradigms of automated music generation*, Springer Wien New York, 2010.

[2] D. Patrick and J. Sheppard, "Evolving four-part harmony using genetic algorithms," *Eur. Conf. Appl. Evolutionary Computation*, Springer Berlin Heidelberg, 2011.

[3] M. Allan and C. KI Williams, "Harmonising chorales by probabilistic inference," *Advances in Neural Inf. Process. Syst.*, vol. 17, pp. 25-32, 2005.

[4] D. Johnson, *Hexadria : Composing music with recurrent neural network*(2015), Retrieved Sept., 08, 2016, from <http://www.hexahedria.com/2015/08/03/composing-music-with-recurrent-neural-networks/>

[5] L. Yi and J. Goldsmith, "Automatic generation of four-part harmony," *BMA* 268, 2007.

[6] W. I. Cho, C. M. Lee, H. Y. Kim, S. Y. Jang, and N. S. Kim, "Rule-based generation of 4-part chorus," in *Proc. KICS Int. Conf. Commun.*, pp. 1492-1494, 2016.

[7] G. Bickerman, S. Bosley, P. Swire, and R. M. Keller, "Learning to create jazz melodies using deep belief nets," in *Proc. Int. Conf. Computational Creativity*, 2010.

[8] B. D. Baek, *Harmonics*, Sumundang, 2008.

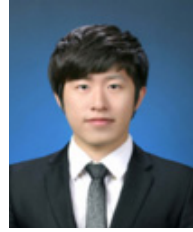
[9] G. W. Taylor, G. E. Hinton, and S. Roweis, "Modeling human motion using binary latent variables," in *Proc. 19th Neural Inf. Process. Syst.*, pp. 1345-1352, 2007.

[10] D. Temperley, *A statistical analysis of tonal harmony*(2009), Retrieved Etp., 12, 2016, from <http://theory.esm.rochester.edu/temperley/kp-stats/index.html>

[11] G. Montufar, N. Ay, and K. Ghazi-Zahedi, "Geometry and expressive power of conditional restricted Boltzmann machines," *J. Machine Learning Res.*, vol. 16, pp. 2405-2436, 2015.

- [12] G. W. Taylor, L. Sigal, D. J. Fleet, and G. E. Hinton, "Dynamical binary latent variable models for 3d human pose tracking," *CVPR*, pp. 631-638, Jun. 2010.
- [13] G. Taylor, *Implicit mixture of conditional restricted boltzmann machine : Code Version 1.000*(2010), Retrieved Apr., 21, 2016, from <http://www.uoguelph.ca/~gwtaylor/publications/cvpr2010/>
- [14] ITU-T, *Recommendation P.800 : Methods for subjective determination of transmission quality*(1996), Retrieved Sep., 09, 2016, from <https://www.itu.int/rec/T-REC-P.800-199608-I/en>
- [15] O. Lartillot, P. Toivainen, and T. Eerola, "A MATLAB toolbox for music information retrieval," in *Data Anal., Machine Learning and Appl.*, Springer Berlin Heidelberg, pp. 261-268, 2008.
- [16] K. S. Kwon, H. Y. Kim, and N. S. Kim, "Speech basis matrix using noise data and NMF-based speech enhancement scheme," *J. KICS*, vol. 40, no. 4, pp. 619-627, 2015.
- [17] C. M. Lee, S. H. Bae, J. H. Kim, and N. S. Kim, "Spectro-temporal filtering based on soft decision for stereophonic acoustic echo suppression," *J. KICS*, vol. 39, no. 12, pp. 1346-1351, 2014.

조 원 익 (Won Ik Cho)



2014년 8월 : 서울대학교 전기정보공학과 학사 졸업  
 2014년 9월~현재 : 서울대학교 전기정보공학부 석박통합과정 박사과정  
 <관심분야> 자동 작곡, 자연어 처리, 음원 분리

김 정 훈 (Jeung Hun Kim)



2013년 2월 : 서울대학교 물리교육과 학사 졸업  
 2013년 3월~현재 : 서울대학교 전기정보공학부 석박통합과정 박사과정  
 <관심분야> 음성 신호 처리, 입체 음향, 음악 정보 검색

천 성 준 (Sung Jun Cheon)



2014년 2월 : 서울대학교 전기정보공학부 학사 졸업  
 2014년 3월~현재 : 서울대학교 전기정보공학부 석박사통합과정 박사과정  
 <관심분야> 기계 학습, 자연어 처리, 자동 작곡

김 남 수 (Nam Soo Kim)



1988년 : 서울대학교 전자공학과 학사 졸업  
 1990년 : 한국과학기술원 전기 및 전자공학과 석사 졸업  
 1994년 : 한국과학기술원 전기 및 전자공학과 박사 졸업  
 1993년~1998년 : 삼성종합기술원 전문연구원

1998년~현재 : 서울대학교 전기정보공학부 교수  
 <관심분야> 음성 신호처리, 음성 인식, 통계적 신호 처리, 패턴 인식, 휴먼 인터페이스