

이미지 분할 및 처리를 이용한 작곡 기법

윤지원*, 김준**

계명대학교 미디어아트대학 뮤직프로덕션과*

동국대학교 영상대학원 멀티미디어학과**

Musical Composition Based on Image Partitioning and Processing

Jiwon Yoon* and Jun Kim**

Department of Music Production, College of Media Art, Keimyung University*

Department of Multimedia, Graduate School of Digital Image & Contents, Dongguk University**

요 약

음악은 일반적으로 여러 성부가 함께 어우러진 집합체로, 다양한 성격을 가지는 서로 다른 요소의 레이어가 모여 하나의 작품을 이룬다. 음악의 이러한 특징은 다양한 시각적 요소로 구성된 이미지와 매우 큰 유사성을 가진다. 특히 음악에서 사용되는 악보는 다양한 수준의 음악적 '이벤트'를 시각화한 2차원 이미지로 해석될 수 있으며, 이는 이미지 안에 포함된 시각 정보로부터 음악 구성에 필요한 여러 요소를 추출할 수 있다는 전제 하에, 사진이나 그림 등의 일반적인 시각 이미지로 음악 악보를 대체·보완하여 작곡 및 연주에 활용할 수 있는 가능성을 시사한다. 본 연구에서는 단순히 소리가 아닌 음악의 구성에 필요한 요소를 얻기 위하여 이미지로부터 서로 다른 수준의 정보를 얻기 위한 기본적인 개념을 논의하며, 실제 구현기법 및 가능성 등을 살펴보고 향후 지속적인 추가 연구에 대한 방향을 제시한다.

I. 서 론

음악은 문명의 발달과 더불어 자연스럽게 변화하고 발전하여 왔다. 현대의 음악은 매우 복잡한 양상을 보여주고 있으며, 음악 구성의 요소는 과거에 비해 훨씬 더 많아지고 다양화되고 있다.

일반적으로 한 음악의 구조 혹은 텍스처(texture)는 그 안에 사용된 성부(聲部, part/voice) 또는 악기의 수 및 상호 관계에 의하여 구분될 수 있으며, 좀 더 자세하게는 화성 및 리듬의 변화 뿐 아니라 사용된 악기의 음색(timbre)과 같은 세부적인 요인에도 영향을 받는다.

또한 과학의 발전에 힘입은 심리음향학의 학문적 성과에 따라, 완성된 하나의 음악을 분석하고 이해하는데 사용되는 요소는 매우 다양한 것으로 밝혀졌으며, 인간은 이러한 요소들을 빠르게 분석하고 이해하며 그 변화 양상을 즐기는 것으로 나타났다. 이러한 바탕 위에, 현대의 작곡가들은 완성된 악보를 유한한 음악의 완성으로 여겼던 과거로부터 벗어나 여러 요소들을 사용하여 새로운 음악을 구성하기 위한 아이디어를 필요로 하고 있다.

본 연구에서는 이러한 시도의 일환으로 시각적 이미지로부터 음악의 구성 요소를 만들어내는 개념을 제시하고, 그에 필요한 기술적 기법을 살펴본다. 이러한 이미지 기

법의 작곡 기법은 한편으로는 음악의 내용을 추상적인 시각 정보로 표현하는 방식인 악보의 진통과도 연결되는 측면을 가진다. 즉 기존의 음악 제작과 연주의 흐름이 작곡자의 의도를 표현한 시각 정보로부터 왔다면, 본 연구의 방향은 이와 반대로 이미 존재하는 시각 정보로부터 음악 구성에 필요한 요소를 얻어내는 작업으로 생각할 수 있다.

본 연구는 연주자의 창의력을 동원하여 악보로부터 이끌어낸 음악적 결과물 그 이상의 것을 얻기 위하여, '주어진 이미지로부터 어떠한 방식으로 음악 구성에 필요한 데이터를 얻을 것인가' 라는 질문에서 시작되었다.

이때 원본 이미지 그 자체 (혹은 이미지를 만들거나 선택한 작가의 예술적 의도) 뿐 아니라 이러한 시각적인 정보를 다양한 기법을 통하여 음악으로 재해석하는 작업의 전 과정은 작곡가의 예술적 영감과 의도를 반영하는 창작의 영역이 되는 동시에, 한편으로는 디지털 이미지 처리(digital image processing)를 비롯한 여러 기술적인 도전을 수반하게 된다. 이는 시각예술과 청각예술 사이의 한 단계 높은 교류를 의미하는 동시에, 청각 기반의 음악이 새로운 차원의 예술로 나아가기 위한 의미 있는 시도가 될 것이다.

II. 시청각 데이터의 상호 대응

이미지를 활용한 작곡 기법은 시각과 청각 데이터의 결합이라는 측면에서 수많은 미디어 아티스트들의 관심사가 되어 온, 오디오비주얼 예술(audiovisual art)[1]과 매우 큰 연관성을 가진다. 예술적인 관점에서 본다면 음악으로부터 영감을 얻어 창작된 시각 예술(또는 그 반대의 경우)의 수많은 사례가 포함될 수 있으나, 본 논문에서는 소리와 이미지 사이의 상호 변환을 중심으로 시각 데이터의 대응 관계(mapping)를 살펴본다.

2.1 소리의 시각화

소리의 시각화(sound visualization), 특히 소리를 이미지로 변환하는 작업의 주요 목표는 기본적으로 소리의 음향학적 특성을 시각적으로 보여주고 분석하기 위한 것이다. 예를 들어 소리의 파형(waveform) 이외에 대표적인 시각화 방법인 스펙트로그램(spectrogram)의 경우 시간에 따른 주파수 성분의 변화를 보여준다(그림 1). 한편 이러한 스펙트로그램 기반의 시각화 기법은 예술 창작 분야에도 활용되고 있으며, 주로 역 스펙트로그램(inverse spectrogram) 대응을 이용한 소리 합성을 통하여 다음에서 언급할 이미지의 청각화(image sonification)와 함께 사용된다.

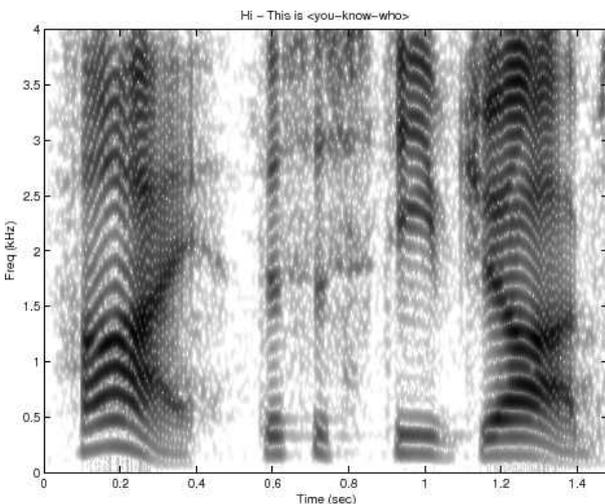


그림 1. 목소리의 스펙트로그램[2]

2.2 이미지의 청각화

컴퓨터 음악 분야에서의 새로운 시도 중 하나인 이미지의 청각화(image sonification)는 시각적인 이미지로부터 음악에 필요한 소리를 만들어내는 기법으로, 본 연구에 매우 근접한 작업이라 할 수 있다. 특히 앞서 살펴본 스펙트로그램의 경우 시간과 주파수 양 축을 가지는 구조로, 이는 시간의 흐름에 따른 음 높이의 변화를 보여주는 전통적인 악보와 유사한 틀을 가진다.

지금까지의 선행연구를 살펴보면 소리를 얻기 위하여 이미지를 효과적·효율적으로 활용하기 위한 방법들을 찾아내는 것을 주된 관심사로 여겨 왔으며, 이를 통해 얻어진 데이터는 기본적으로 곡 전체가 아니라 곡의 일부를 이루는 소리로 청각화 되었다. 또한 이들 선행연구의

경우 소리 생성에 필요한 데이터 검출을 위한 스캐닝(scanning) 방식 및 데이터 대응 자체가 가장 중요한 연구 주제였다. 스펙트로그램과 유사한 형식의 구조를 가지는 이미지 청각화의 사례는 Yellowtail[3], Metasynth[4], 또는 AudioSculpt[5]와 같은 소프트웨어를 들 수 있으며, 좀 더 단순한 형태로는 소리 샘플과 이미지 픽셀이 1:1 대응관계를 이루는 래스터 스캐닝(raster scanning) 기법을 활용한 Rasterpiece[6]와 그리고 Monalisa[7] 등이 있다.

2.3 비주얼 악보

Stockhausen[8]이나 Ligeti와 Wehinger[9] 등이 시도한 비주얼 악보(visual score)는 전통적인 악보의 틀을 어느 정도 유지하면서도, 여러 종류의 기호와 색상을 이용하여 다양한 음악적인 정보를 표현하였다(그림 2). 이러한 사례들은 시각 정보의 활용이라는 점에서 매우 중요한 분석 대상이나, 작곡을 의도하지 않은 이미지에서 음악 구성에 필요한 정보를 추출하는 본 연구와는 차이점을 보인다.

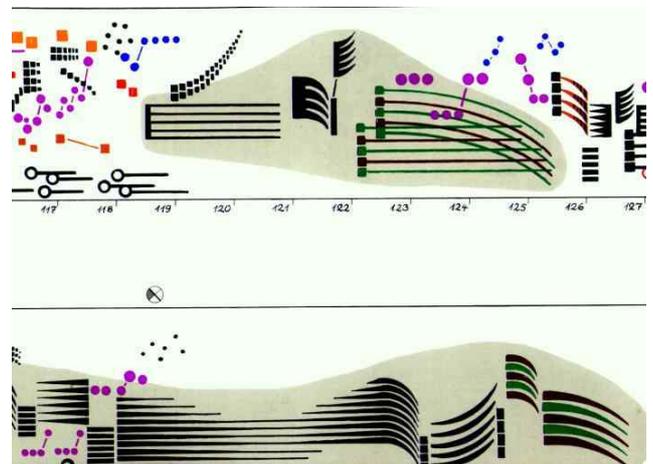


그림 2. G. Ligeti and R. Wehinger, *Artikulation*[9].

III. 음악 정보 구성을 위한 데이터 추출

이상 살펴본 선행연구와 같은 이미지의 청각화에서 한 단계 더 나아가 (또는 비주얼 악보의 역의 개념으로), 본 연구에서는 단순히 소리를 얻는 것에 그치지 않고 하나의 음악 작품을 구성하기 위한 데이터를 얻는 방법들을 제안하고 그 가능성을 확인하는 것을 목표로 한다. 이때 인간에게 있어 시각화된 음악의 가장 일반적이고 친숙한 형태는 역시 악보라 할 수 있으므로, 본 연구에서는 가로 축이 시간에 대응되는 구조를 기반으로 다양한 요소의 대응 관계를 살펴본다.

3.1 영역 분할에 의한 다층 구조

평면적인 이미지로부터 음악 구성에 필요한 다차원의

데이터를 얻어내는 한 가지 방법은 이미지를 가상적으로 분할하여 일정한 영역에서 특정 정보만을 추출하는 것이다. 그림 3은 이러한 분할의 예를 보여주며, 이때 각 영역 안에서의 이미지의 변화 정도는 악곡을 구성하는 개별 성부의 음량 및 음고 변화 등에 대응될 수 있다.

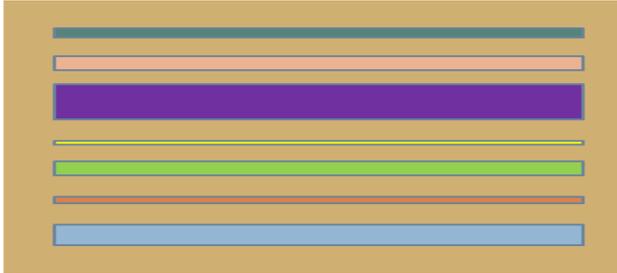


그림 3. 수평 분할.

또 다른 요소는 시간(혹은 음악의 빠르기)에 관한 것이다. 음악은 시간의 예술로, 주어진 시간을 음악적으로 어떻게 구성하느냐 하는 것이 주된 관심사가 되며, 각각의 성부 또는 악기에 따라 서로 다른 계층적 리듬을 부여하는 작업은 수직 방향으로의 분할 간격 또는 스캔 속도를 다르게 설정하여 이루어질 수 있을 것이다(그림 4).

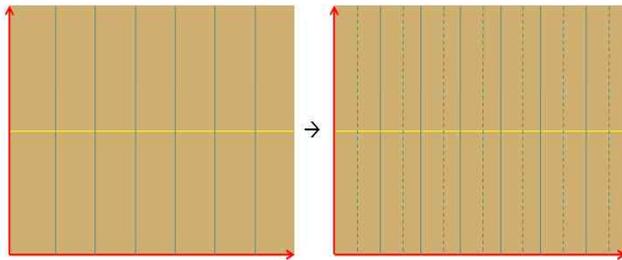


그림 4. 수직(시간) 분할 및 스캐닝

그림 5는 이미지의 구성에 따라 이상의 두 기법을 조합하여 적용한 사례를 보여준다. 작곡가는 자신의 의도에 따라 이미지를 선택하고 영역을 분할할 수 있을 것이다.

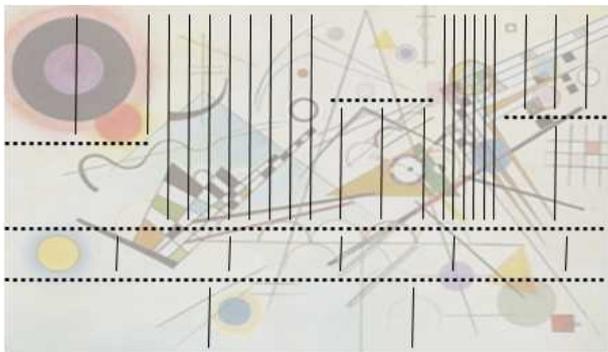
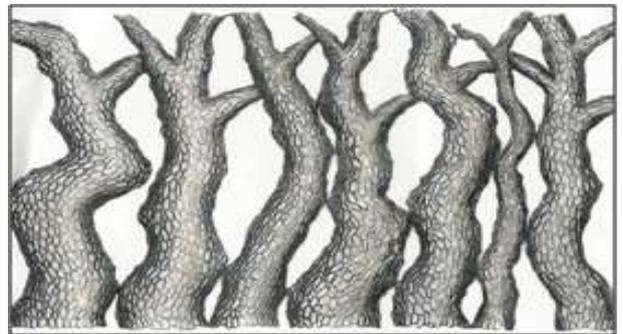


그림 5. 이미지의 구성에 따른 수직·수평 분할의 예 (이미지 원본: Kandinsky, *Composition VIII*, 1923)

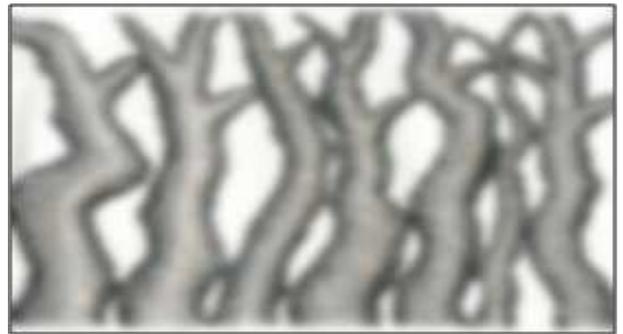
3.2 필터를 이용한 이미지 처리

다차원 데이터를 얻는 또 다른 접근 방법은 하나의 이미지에 서로 다른 여러 이미지 필터(image filter)를 적용한 결과를 활용하는 것이다. 원본 이미지가 같은 경우, 비록 다른 필터를 적용한 경우라도 그 처리 결과들은 어느 정도 상호 유사한 모습을 보이며, 이러한 유사성은 각각의 결과들이 유기적으로 융합되어 하나의 테마를 가진 음악이 될 수 있는 이론적 기반 뿐 아니라 인지적 근거를 제공할 수 있다.

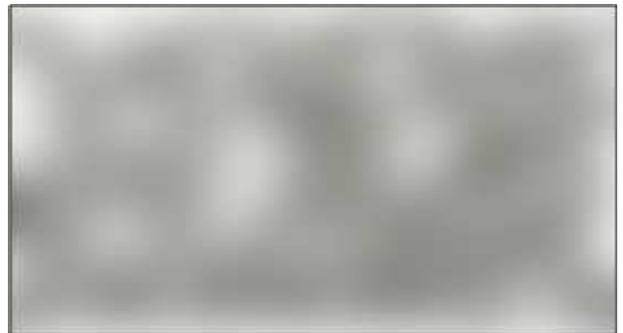
그림 6은 원본 이미지 (a)와 그에 각각 다른 정도의 블러(blur) 필터를 적용한 결과 (b), (c)의 비교를 보여준다. 동일한 높이에서 수평 방향으로 스캔을 하여 얻은 밝기 값을 비교한다면 (a)에 비하여 (b)와 (c)는 상대적으로 수직 변동이 작은 결과를 보여주며, 이는 음악적으로 더 부드럽고 느린 움직임의 선율에 대응될 수 있을 것이다.



(a) 원본 이미지



(b) 블러(blur) 필터 적용 (1)



(c) 블러 필터 적용 (2)

그림 6. 이미지의 구성에 따른 수직·수평 분할의 예 (이미지 원본: 이길래(李吉來), <응집-나무>)

IV. 결론

본 연구는 전체 이미지 또는 이미지의 특정 영역을 선택적으로 스캔하고 그 결과를 음악의 정보로 사용하기 위한 것이다. 스캔을 위한 이미지 분할 및 데이터 산출을 위한 기술적인 면은 Max/MSP와 Jitter를 통해 처리될 것이며, 여기서 추출된 데이터는 음악의 요소로 사용될 수 있도록 미디 이벤트 정보에 대응된다. 또한 하나의 이미지로부터 다차원의 데이터를 얻기 위하여 이미지 처리 프로그램에 사용되는 필터들을 이미지에 적용하며, 이에 따른 결과물을 비교·분석하는 한편, 도출된 결과를 체계적으로 분류하는 작업으로 진행될 것이다.

향후 예상되는 주요 연구 주제는 다음과 같다.

- 동영상 처리: 단순히 정지된 이미지 뿐 아니라 무비 또는 애니메이션 등의 동영상에서 음악 정보를 추출하는 기능은 영화의 사운드트랙을 작곡하는 작업 등에 활용될 수 있는 새로운 형태의 멀티미디어 저작 도구로서의 가능성을 열게 될 것이다.
- 필터의 공감각적 분석: 특정 이미지 필터에 대응되는 소리 필터를 단순히 수학적 연관성 뿐 아니라 인지적인 면에서 접근하여 그 상관관계를 조사하고, 이와 연관될 수 있는 음악적인 요소(구성, 또는 제스처)에 대한 이론적인 기반을 정리하여 새로운 방식의 '공감각적' 작곡법을 제시할 수 있다.
- 자동화 시스템: 이상의 연구 성과를 패턴 인식(pattern recognition) 기법 등과 결합하여, 향후 음악의 장르에 따라 적합한 형태의 이미지를 자동으로 추천할 수 있는 시스템을 구축할 수 있을 것이다.

VI. 참고문헌

[1] G. Levin, "Painterly Interfaces for Audiovisual Performance," M.S. Thesis, MIT, 2000.
 [2] J. Smith, "Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, Second Edition," <http://ccrma.stanford.edu/~jos/mdft/>, 2007, online book, accessed October 2008.
 [3] G. Levin and Collaborators, "Yellowtail," <http://www.flong.com/projects/yellowtail/>, 2000, online page, accessed September 2008.
 [4] U&I Software, "Metasynth," <http://uisoftware.com/MetaSynth/>, 2007, online page, accessed September 2008.
 [5] IRCAM, "Audiosculpt," <http://forumnet.ircam.fr/691.html?&L=1>, online page, accessed September 2008.

[6] W. Yeo, "Rasterpiece: a cross-modal framework for real-time image sonification, sound synthesis, and multimedia art," Proc. of the International Computer Music Conference (ICMC), Copenhagen, Denmark, 2007.
 [7] K. Jo and N. Nagano, "Monalisa: see the sound, hear the image," Proc. of the International Conference on New Interfaces for Musical Expression (NIME), Genova, Italy, 2008.
 [8] K. Stockhausen, "Texte," vol. 2, Cologne, 1964.
 [9] G. Ligeti and R. Wehinger, "Artikulation: Elektronische Musik," Schott, 1970.