

# 대화형 유전자 알고리즘을 이용한 감성적 음악검색

이준승<sup>0</sup> 조성배

연세대학교 컴퓨터과학과

odradeck@candy.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

## Emotional Musical Expression Retrieval Using Interactive Genetic Algorithm

Jun-Seung Lee<sup>0</sup> Sung-Bae Cho

Department of Computer Science, Yonsei University

### 요 약

본 논문에서는 대화형 유전자 알고리즘을 이용한 음악검색 시스템을 개발한다. 기존의 음악검색 시스템은 찾고자 하는 음악에 대한 정보를 필요로 하기 때문에 사용자가 시스템이 요구하는 정보를 가지고 있지 않는 경우 검색이 힘들다. 하지만 대화형 유전자 알고리즘을 통한 질의어 생성방식을 이용하면 사용자의 주관적 감정에 의한 음악검색을 할 수 있다. 먼저 사용자가 초기 유전자형에 의해 선택된 음악을 듣고 주관적인 평가를 내리면 이 평가값으로 유전자 알고리즘을 이용하여 질의어를 생성, 가장 가까운 음악을 검색하여 들려준다. 사용자는 이러한 과정을 반복하여 자신의 주관적 평가에 따라 진화되는 유전자형에 의해 원하는 음악을 점진적으로 검색한다. 이를 통해 사용자는 전문적인 음악적 지식이나 찾고자 하는 음악에 대한 특별한 정보없이 개인의 주관에 의한 검색을 할 수 있을 것이다.

### 1. 서 론

디지털화된 정보가 점점 많아지면서 그것들을 효과적으로 찾아낼 수 있는 방법에 대한 관심이 커지고 있다. 또한 검색의 대상 역시 텍스트기반의 정보에서 음악이나 그림 등의 매체에 까지 확대되고 있다. 하지만 이런 감성적 매체는 텍스트기반 정보와는 다르게 사용자의 주관적 감정에 따라 그 정보의 가치 역시 변하게 된다. 같은 음악을 들더라도 어떤 사용자는 슬프게 느낄 수 있고 어떤 사용자는 우울하게 느낄 수 있음은 물론, 때에 따라 한 사용자가 동일한 음악에 대해 다르게 느낄 수 있다. 기존의 음악 검색 시스템에서는 이러한 음악의 감성적 특징보다는 빠르고 효율적인 음악패턴추출, 사용자의 허밍(Humming)음의 변환 등에 중점을 두었다. 하지만 이러한 검색 방법은 사용자가 자신의 감성을 시스템이 요구하는 질의조건으로 표현할 수 있어야 한다. 즉 음악의 일부를 부를 수 있거나, 멜로디의 계이름을 알아야 검색이 가능하다.

본 논문에서는 주관적 감정에 의한 음악검색을 위해 대화형 진화 알고리즘을 이용한 질의어의 생성 방법을 제안한다. 사용자가 유전자형에 의해 선택된 음악을 듣고 평가한 값으로 유전자 알고리즘을 통해 질의어를 생성한다. 이 질의어는 다음 음악을 검색하고, 사용자는 다시 음악을 듣고 평가를 내리는 과정을 반복한다. 이러한 과정을 통해 사용자 감정에 의해 진화된 질의어에 의한 검색이 이루어지는 것이다. 이 연구를 통해 미적인 매체에 대한 감성적 검색의 가능성을 알아본다.

### 2. 배경

#### 2.1. 기존의 음악검색시스템

이제까지의 음악 검색 시스템에 대한 연구는 크게 두 가지 분야에 관심을 두고있다. 즉, 음악검색에 사용될 질의어 생성과 음악에서 패턴이나 주제의 추출이다. 전처리를 통해 빠르고 효율적인 검색을 위한 방법이 연구되는 것이다[1].

특히 '허밍에 의한 질의어(Query by Humming)'에 대한 관심이 높아지고 있는데, 이는 사용자가 마이크로폰을 이용하여 녹음한 허밍음을 통한 검색으로, 허밍음을 기호화된 표현으로 변환시키는 방법에 중점을 두고 있다. 대부분 검색대상이 되는

음악정보는 기호화되어 저장되고 검색에서도 이 기호화형식을 이용한 인덱스를 사용하기 때문이다. 실제로 MelDex[2]라는 웹기반의 음악 검색 시스템은 사용자가 음악을 녹음하여 전송하면 이를 변환하여 비슷한 음악을 찾아준다.

하지만 이런 시스템은 사용자가 찾고자 하는 음악에 대한 시스템이 요구하는 정보를 가지고 있어야 한다. 허밍음에 의한 검색은 사용자가 일부분의 음악을 정확히 부르지 못하면 검색할 수 없다는 단점이 있다. Themefinder[3]라는 테마별 음악검색 시스템은 찾고자 하는 음악에 대한 계이름이나 음악의 진행, 음의 진행간격 등에 대한 정규식 표현을 요구한다. 새롭게 정의한 음악기호형식과 이를 통한 다양한 정보를 이용한 빠른 검색이 이루어지지만, 음악에 대한 많은 전문적 지식이 필요하기 때문에 음악 이론이 없는 사람은 이를 이용하여 검색하기가 어렵다.

#### 2.2. 대화형 유전자 알고리즘

유전자 알고리즘이란 자연계에서 적자생존의 원리에 따라 세대가 지나면서 우량의 형질을 지닌 개체가 생성되는 과정을 모방한 알고리즘으로 최적화 문제나 탐색문제의 해를 구하는데 적절하다. 집단이라는 탐색공간이 초기화되고 집단의 각 개체는 평가함수에 의해 평가된다. 이 평가값에 의하여 다음세대에 선택될 개체의 선택확률이 정해진다. 이 과정에서 돌연변이와 교차의 발생을 통해 더 효율적인 탐색이 가능하도록 한다. 이렇게 세대가 되풀이되며 더 적합한 해를 찾아간다.

대화형 유전자 알고리즘이란 일반적인 유전자 알고리즘에 있는 평가 함수 대신 사용자가 직접 평가값을 부여하는 것이다. 본 논문에서는 음악검색이라는 개별적인 감성적 접근이 필요한 문제의 해결을 위해 대화형 유전자 알고리즘을 사용하였다. 이는 사용자와 긴밀한 상호작용을 통해 개별적인 특성을 반영할 수 있게 해준다. 사용자의 주관에 의해 평가되어 생성된 질의어를 이용하여 음악을 검색하여 들려주고, 다시 평가받는 진화과정을 반복한다. 대화형 유전자 알고리즘은 이미 여러 가지 감성적 매체에 대한 연구에서 사용되고 있다[4]. 비슷한 예로 감성적 이미지 검색[5]에서 실제로 적용되기도 하였다.

2.3. 음높이 변화 분석을 통한 음악적 특징 표현

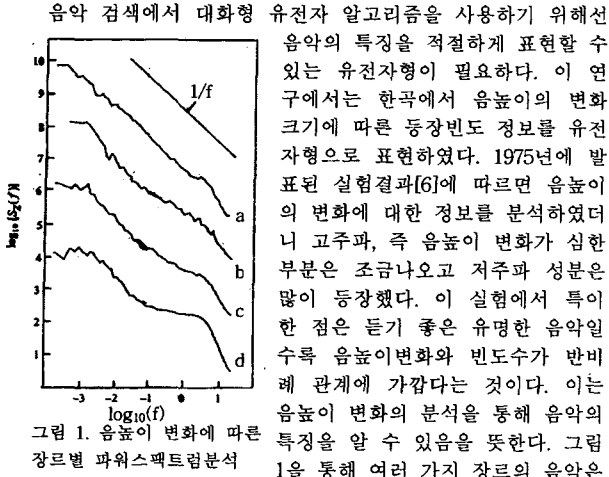


그림 1. 음높이 변화에 따른 장르별 파워스펙트럼분석

음악 검색에서 대화형 유전자 알고리즘을 사용하기 위해선 음악의 특징을 적절하게 표현할 수 있는 유전자형이 필요하다. 이 연구에서는 한국에서 음높이의 변화 크기에 따른 등장빈도 정보를 유전자형으로 표현하였다. 1975년에 발표된 실험결과[6]에 따르면 음높이의 변화에 대한 정보를 분석하였더니 고주파, 즉 음높이 변화가 심한 부분은 조금나오고 저주파 성분은 많이 등장했다. 이 실험에서 특이한 점은 듣기 좋은 유명한 음악일 수록 음높이변화와 빈도수가 반비례 관계에 가깝다는 것이다. 이는 음높이 변화의 분석을 통해 음악의 특징을 알 수 있음을 뜻한다. 그림 1을 통해 여러 가지 장르의 음악은 서로 다른 주파수별 강도를 나타냄을 확인할 수 있다. 가로축은 음높이의 변화의 주파수를 나타내고 세로축은 이를 분석한 파워스펙트럼이다. 4개의 다른 장르 라디오채널에 대한 결과로 a는 클래식, b는 재즈, c는 락, d는 뉴스 채널에 대한 분석결과이다. 이는 음높이변화에 대한 강도가 음악의 특성을 나타낼 수 있음을 보여준다. 본 연구에선 이를 이용하여 음악에 대한 유전자형을 표현하였다.

3. 감성적 음악 검색

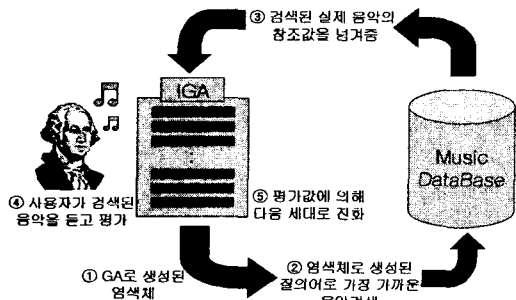


그림 2. 시스템 작동 과정

그림 2는 감성적 음악검색 시스템이 작동하는 과정이다. 그림에서처럼 이 시스템은 크게 두분으로 나뉜다. 대화형 유전자 알고리즘을 이용하여 염색체 형태의 질의어를 생성하는 부분과 이렇게 생성된 질의어를 이용하여 실제 저장된 음악데이터에서 음악을 찾아주는 부분이다.

먼저 난수적으로 초기 염색체 집단이 생성된다. 이렇게 생성된 염색체를 이용하여 각 염색체와 가장 가까운 특징의 음악을 찾게되고 사용자는 이렇게 검색된 실제 음악을 듣고 염색체에 대한 평가를 내리게 된다. 이러한 과정의 반복을 통해 사용자의 주관에 의해 평가되고 진화된 염색체형태의 질의어를 이용한 검색이 이루어지는 것이다.

3.1. 음악의 표현

이 시스템에서 유전자형은 사용자의 주관적 감성에 의해 진화되면서 음악검색에서 질의어로 사용되기 때문에 유전자형에 음악의 특징을 잘 표현해야 효과적인 검색이 가능하다. 이를

위해 음높이 변화의 범위에 따른 빈도수측정을 통해 유전자형을 설계하였다. 실제 신호형식의 음악이라면 음높이 변화에 따른 정보를 따로 분석한 후 이렇게 생성된 신호를 다시 파워스펙트럼 분석을 통하여 음높이 변화의 관계에 대한 값을 얻을 수 있겠지만, 이 시스템은 MIDI형태의 기호화된 음악을 사용하기 때문에 음높이에 대한 정보를 MIDI화일에서 추출한후 각 음높이 변화에 대한 빈도수를 계산하였다.

염색체	11111101110001000111000100111010001
질의어	음높이변화 0~1 2~3 4~5 6~7 8~9 10~11 12~13 14~
빈도수	15 13 12 4 12 5 13 1

<생성된 염색체와 질의어>

가장가까운 음악검색      거리 : 15.88

음높이변화	0~1	2~3	4~5	6~7	8~9	10~11	12~13	14~
빈도수	3	2	2	1	1	2	4	0

<검색된 음악>

그림 3. 음악 인코딩과 검색

먼저 유전자 알고리즘에 의해 32비트의 염색체가 생성된다. 이 염색체는 4비트 마다 음높이 변화에 따른 빈도수를 나타내고 있다. 또 저장된 실제 음악에서도 음높이변화에 따른 빈도수 정보를 추출한다. 음악검색은 이 염색체를 통해 생성된 질의어와 실제 음악에서 추출한 정보를 비교하여 가장 가까운 음악을 추출하는 방법으로 이루어진다. 그림 3은 생성된 염색체의 질의어 형태와 이 질의어와 가장 가까운 형태를 가진 검색된 음악의 모습이다.

3.2. 검색 및 평가

사용자는 유전자형과 가장 가까운 형태의 실제 음악을 듣고 평가를 내리게 된다. 이미 저장되어 있는 음악은 전처리리를 통하여 각 음높이변화에 따른 빈도수에 대한 정보를 가지고 있다. 음표의 개수에 영향을 받지 않기 위해 전체음표개수를 이용하여 각 음높이변화별로 비율을 구하여 질의어와 실제음악을 비교한다.  $Q(f)$ 를 질의어의 음높이변화별 빈도비율,  $M(f)$ 를 실제 음악의 음높이변화별 빈도 비율이라 할때, 거리  $D(Q, M)$ 은 다음과 같이 계산한다.

$$D(Q, M) = \sqrt{(Q(A) - M(A))^2 + \dots + (Q(N) - M(N))^2}$$

전처리리를 통해 저장된 실제음악의 음높이변화별 비율과 질의어에서 음높이변화별 비율들 사이의 단순한 거리계산으로 거리를 구한다. 질의어와 거리가 가장 가까운 음악이 선택되면 사용자는 그림 4와 같은 인터페이스를 통해 직접 노래를 듣고 악보로 확인하며 각 음악에 대해 평가를 내리게 된다. 이 평가값은 유전자형에 대한 평가값으로 사용된다.

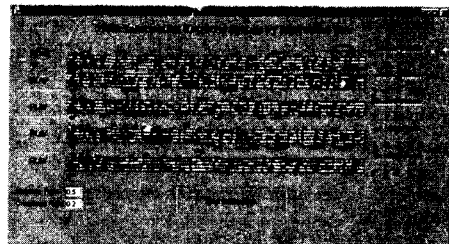


그림 4. 시스템의 인터페이스

3.3. 검색 질의어의 진화

사용자에 의해 평가된 검색체는 유전자 알고리즘을 통해 다음 세대로 진화하게 된다. 평가값에 의한 선택확률에 따라 룰렛 휠 방법으로 다음 세대에 나타날 검색체를 선택한다. 이렇게 선택된 검색체들은 사용자의 입력에 따라 결정된 교차확률과 돌연변이 확률에 따라 변화된 다음 세대가 된다. 교차는 교배가 일어나는 검색체를 임의로 선택한 후 임의의 두 지점을 선택하여 비트스트링을 교환하고, 돌연변이는 돌연변이가 일어나는 검색체에 임의로 선택된 수만큼의 비트를 변환시킨다. 그림 5는 검색체의 교차연산에 의한 질의어와 검색된 음악의 변화를 보여준다. 음영부분은 검색체에서 실제 교체가 일어난 부분이고 그에 따른 1번 검색체의 질의어 변화모습이다.

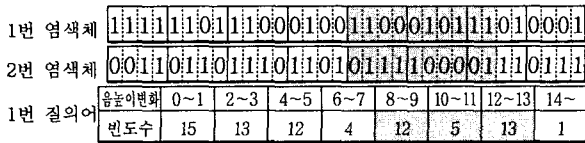


그림 5. 교차연산에 의한 질의어변화와 검색된 음악

4. 실험 및 결과

실험적으로 짧은 소절의 MIDI형식의 음악화일 200개를 가지고 실험하였다. 실험자는 200개의 저장된 음악중 임의로 하나의 목표음악을 선택한 후 그 음악을 반복해서 듣는다. 여러 번 들어서 이 음악에 대한 느낌을 바탕으로 교차확률과 돌연변이 확률 값에 변화를 주면서 실험을 실시 하였다. 초기생성된 집단에서부터 한 세대씩 진화할 때마다 목표 음악과 진화를 통해 만들어진 질의어사이의 거리를 기록하여 음악검색의 과정을 보여주었다. 그림 6은 질의어 진화에 따른 목표음악과의 거리를 돌연변이와 교배확률을 바꿔가며 테스트한 결과로 한 세대의 집단에서 거리가 가장 가까운 값을 뽑아 도식화한 그래프이다. 그래프에서 m은 돌연변이 확률, c는 교차확률을 나타낸다.

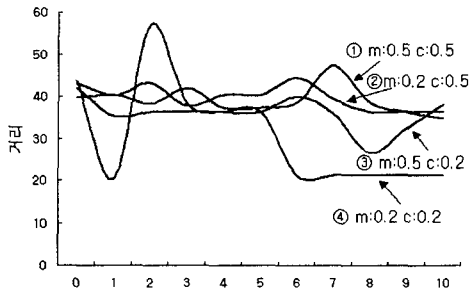


그림 6. 세대별 진화에 따른 목표음악과의 거리변화  
교차확률과 교배확률에 따라 다르지만 초기에 랜덤하게 생성된 질의어에 비하여 진화후의 질의어를 통한 검색이 목표음악에 가까운 거리를 나타내는 것을 볼 수 있다. 그림 6에서 실

제로 목표음악과 진화되가는 질의어의 모습을 비교해 보았다.

음높이변화	0~1	2~3	4~5	6~7	8~9	10~11	12~13	14~
빈도수	5	6	2	0	0	0	0	0

음높이변화	0~1	2~3	4~5	6~7	8~9	10~11	12~13	14~	거리
초기 형태	0	8	9	9	0	6	9	43.68	
1세대	14	12	3	8	3	2	0	20.37	
2세대	0	1	6	8	5	6	6	56.99	
3세대	11	2	1	11	2	1	3	38.33	
4세대	13	8	6	6	6	1	1	36.39	
5세대	13	8	6	6	6	1	1	36.39	
6세대	6	5	1	4	7	2	40.30		
7세대	13	11	10	13	10	14	6	35.91	
8세대	12	12	0	0	12	0	1	26.53	
9세대	11	1	11	0	4	2	1	32.45	
10세대	9	7	8	11	14	5	11	38.13	

그림 6. 목표음악과 진화에 따른 질의어 변화

빈도수가 가장 높은 음높이 변화를 음영의 진화기를 통해 나타내었다. 초기 값은 뒷부분에 나타나지만 진화를 거쳐 목표 음악과 비슷하게 앞으로 움직이는 모습을 관찰할 수 있었다.

5. 결론 및 향후방향

본 논문에서는 사용자와 검색시스템의 상호작용을 통한 질의어의 진화적 생성이라는 방법을 통하여 개인의 주관적 검색이 가능한 시스템을 제안하였다. 이는 이제까지 분류를 통한 인덱스 검색이나, 패턴추출에 의한 비교검색 등과는 다른 검색 환경을 제공한다. 이러한 검색 인터페이스는 음악의 제목이나 멜로디를 정확히 모르더라도, 혹은 음악 자체에 대한 정보가 전혀 없더라도 사용자는 이 시스템을 통해 자신이 원하는 음악을 찾아낼 수 있다. 이처럼 음악이라는 감성적인 요소가 중심이 되는 매체는 그것에 맞게 개인의 감성적 요소를 중심으로 검색되어야 더 효과적인 결과를 얻을 수 있다.

하지만 실험에서도 알 수 있듯이 가까운 음악은 찾을 수 있었지만 정확한 목표음악을 찾기는 어려움이 있었다. 이는 대화형 유전자 알고리즘에서 사용자의 평가방법을 개선하여 더욱 많은 세대수와 집단수를 이용한다면 정확한 해를 찾을 수 있을 것이다. 또한 음악에서 음높이 변화에 따른 분석으로 어느 정도 음악의 특성을 들어낼 수 있었지만 단음이 아닌 화음의 음악이나, 박자, 음색 등에 대한 특징은 반영되지 않았다. 멜로디의 변화뿐만 아니라 음색이나 박자, 음량까지 나타낼 수 있는 표현을 사용하면 더욱 효과적인 검색이 될 것이다.

6. 참고문헌

- [1] P.-Y. Rolland, "Music information retrieval : A brief overview of current and forthcoming research," *Workshop Engineering and Music*, 2002.
- [2] D. Bainbridge, R.-J. McNab and L.-A. Smith, "Melody based tune retrieval over the world wide web," Working Paper , no. 97-17, 1998.
- [3] D. Huron, W. Hewlett, et al. "How themefinder works," <http://www.themefinder.org>, 2001.
- [4] H. Takagi, "Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation," *Proceedings of the IEEE*, Vol. 89, Issue 9, pp. 1275-1296, 2001.
- [5] J.-Y. Lee and S.-B. Cho, "Application of interactive genetic to image retrieval based on emotion," *Korea Information Science Society(B)*, vol. 26, no. 3, pp. 422-430, 1999.
- [6] R.-F. Voss and J. Clarke, "1/f noise in music and speech," *Nature*, vol. 258, pp. 317-318, 1975.