

허밍 질의 기반 음악 검색 시스템의 유사도 계산 알고리즘

오 동 열*, 오 해 석**

A Similarity Computation Algorithm for Music Retrieval System Based on Query By Humming

Oh Dong Yeol *, Oh Hae Seok **

요 약

사람은 음악에서의 선율을 악보의 기보법과 같이 음표의 높이와 음표의 길이가 조합된 형태로서 기억하는 것이 아니라, 전반적인 음표 간의 높낮이의 흐름과 음표 사이의 상대적인 지속시간으로 구성된 음조 곡선 형태로 기억한다. 이와 같은 선율의 기억 방식으로 인해 기존 음악 검색 시스템과 같이 건반을 이용한 주선율 입력이나 악보에 기보된 형태로 음악 검색의 질의를 이용하는 방법을 그대로 적용하기 어려운 점이 있다. 이에 본 논문에서는 사용자의 허밍을 질의로 사용하는 음악 검색 시스템에서의 고려 대상들과 기존에 연구된 허밍 질의 기반의 음악 검색 시스템을 살펴본다. 또한 사람이 선율을 기억하는 방식인 상대적인 음표 간 높낮이와 음표 지속 시간을 이용하여 음악 내에 특정 정보를 추출하고 이를 기반으로 허밍 질의와 단음과의 유사도 계산 알고리즘을 제안한다. 제안된 유사도 계산 알고리즘은 선율내의 음높이 차만 가지고 선율의 유사도를 비교하는 경우 발생할 수 있는 문제를 연속된 음간에 길이 차이를 이용하여 해결하였다.

Abstract

A user remembers a melody as not the combination of pitch and duration which is written in score but the contour which is composed of the relative pitch and duration. Because of the way of remembering a melody, the previous Music Information Retrieval Systems which uses keyboard playing or score as the main input melody are not easily acceptable in Query By Humming Systems. In this paper, we mention about the considerable checkpoints for Query By Humming System and previous researches. And we propose the feature extraction which is similar with the way of remembering a melody and similarity computation algorithms between melody in humming and melody in music. The proposed similarity computation algorithms solves the problem which can be happened when only uses the relative pitches by using relative durations.

▶ Keyword : 음악 검색 시스템(Music Information Retrieval System), 허밍 질의(Query By Humming), 멀티미디어 검색 시스템(Multimedia Information Retrieval System)

• 제1저자 : 오동열

• 접수일 : 2006.08.31, 심사일 : 2006.09.11, 심사완료일 : 2006.09.20

* 숭실대학교 컴퓨터 공학과 박사 과정 ** 경원대학교 컴퓨터 공학과 교수

I. 서론

웹의 급속한 발전과 더불어 기존 텍스트 기반의 콘텐츠가 멀티미디어 형태로 변화됨에 따라 멀티미디어 정보의 생성 및 처리와 더불어 필요한 정보를 효과적으로 검색할 수 있는 방법이 요구되었다.

멀티미디어 데이터 중 오디오 데이터의 특징 정보를 기반으로 한 오디오 검색 방법에 대해서는 점차적으로 연구가 진행 중이나 사람의 허밍 질의를 이용한 음악 데이터에 대한 연구는 사람이 재생해 내는 허밍 질의의 불확실성에 의해서 기존 오디오 데이터의 검색 방법이 그대로 적용되기 어려운 부분들이 있다.

사람들은 일반적으로 자신이 기억하고 있는 선율과 연관된 곡의 작곡가, 가수, 노래 제목, 가사와 같은 메타 데이터를 정확하게 기억하지 못하는 경우가 많다. 선율 정보와 선율이 포함된 곡의 메타 정보들은 사용자가 관심을 갖지 않는 이상 명확하게 연관 지어 기억해 내기 힘들며, 반대로 음악의 제목을 알더라도 그 선율 정보를 기억하기 어려울 수 있다. 따라서 선율 정보와 선율이 포함된 곡의 메타 데이터를 사용자가 보다 용이하게 검색하기 위해서 사용자의 허밍 정보로 해당 곡의 메타 데이터를 검색할 수 있는 방법이 요구된다. 또한 사용자는 특정 곡의 선율을 기억할 때, 악보의 기보된 형태로 음표의 높이(Pitch)와 지속 시간(Duration)의 조합으로서 선율을 기억하지 않는다. 일반적으로 사용자는 음조의 곡선(Contour)을 본능적으로 기억하기 때문에 이러한 기억을 기반으로 발생되는 사용자의 허밍 정보는 여러 가지 특징들을 내포하고 있다. 따라서 사용자의 허밍을 질의로 사용하는 허밍 기반의 음악 검색 시스템에서는 여러 가지 특징들이 고려되어야 한다. 본 논문의 구성은 다음과 같다. 2장에서는 선율에 대한 기본 개념과 더불어 허밍 질의를 이용한 음악 검색 시스템의 고려 대상에 대해서 기술한다. 이와 더불어 기존에 관련된 연구를 살펴본다. 3장에서는 본 논문에서 제안된 유사도 계산 알고리즘을 제안한다. 4장에서는 제안된 유사도 알고리즘을 적용하여 실험 데이터를 이용 제안된 알고리즘의 타당성을 검토한다. 마지막으로 5장에서 결론 및 향후 연구를 언급한다.

II. 관련 연구

본 장에서는 음악에서의 선율에 대한 기본적인 개념과 허밍 질의를 이용한 음악 검색 시스템에서의 고려 사항을 설명하고 기존에 연구된 음악 정보 검색 시스템을 언급한다.

2.1 음악과 선율

음악은 글자 그대로 ‘소리(音)를 즐긴다(樂)’는 의미로서 사전적 의미로는 “박자, 가락, 음성 따위를 갖가지 형식으로 조화하고 결합하여, 목소리나 악기를 통하여 사상 또는 감정을 나타내는 예술”을 일컫는다[1]. 음악을 구성하는 여러 가지 요소 중에서 선율은 여러 개의 음을 일정한 규칙에 맞추어 미적, 시간적으로 리듬을 가지고 연속적으로 배치한 것으로서 하나의 곡을 다른 곡과 분별할 수 있는 중요한 구성 요소이다. 선율은 오선 상에 표기된 관점에서 볼 때 여러 음들의 조합으로 구성되며, 각 음에 대한 구분은 표1과 같은 특징 정보에 의해 가능하다[2].

표 1. 음의 구성 정보
Table 1. Elements in a Sound Note

음의 구성 정보	설 명
음의 발생 시간	최초 음이 발생한 시간. 오선에서는 음의 가로 방향의 의미
음의 높이	1초 동안에 진동수를 의미하며 헤르츠(Hertz)라 표시. 오선에서는 세로 방향의 위치가 높이를 의미
음의 길이	일정한 높이의 음이 지속된 시간. 오선에서는 음표의 형상으로 표현
음의 크기	음압을 의미하며 데시벨(dB)로 표시. 오선에서는 직접적으로 표현되지 않으며 부가 정보에 의해 결정

음을 구성하고 있는 특징 정보 중에서 음의 발생시간, 음의 높이, 음의 길이는 악보에 기보가 되면 기본적으로 모호함이 없고, 정보의 속성으로 보자면 디지털 적이다. 그러나 음의 크기는 객관적인 표현으로 나타나지 않고 표정 기호나 부가 정보로 표현되기 때문에 아날로그적 성격을 가지고 있는 정보이다. 따라서 본 논문에서는 허밍 정보에서 특정 정보를 추출할 때, 음의 발생 시간, 음

의 높이, 음의 길이를 이용하며 음의 크기에 대해서는 고려하지 않는다. 왜냐하면 음의 크기 자체는 상대적인 요소이며 악기의 특성이나 사람의 주관적인 특성에 의존적인 요소이기 때문이다.

2.2 허밍 질의 기반 음악 검색 시 고려 사항

허밍을 질의로 사용하는 음악 검색 시스템에서는 기존에 신디사이저 건반을 이용한 질의나 악보를 기반으로 특수한 형태의 문자나 기호로 선율 정보를 생성하고 이러한 질의를 이용한 시스템과는 차이점을 가지고 있다. 이러한 차이점은 인간의 생리학적 구조로 인해 악보나 음악을 보고 선율을 기억하는 방법과 재생하는 방법이 주관적이기 때문이다. 이러한 예로, Richard Moore에 따르면 악기나 노래로 재생된 특정 음의 절대적인 높이를 알 수 있는 사람은 통계적으로 전체의 0.01%이며 대부분의 음악에 관련된 일에 종사하는 사람들의 경우에도 음간의 상대적인 높이만을 구분하며, 이러한 능력은 단순한 음의 높이에만 국한된 것이 아니라, 음의 길이에서도 비슷하다고 언급하고 있다[3]. 일반적으로 대부분의 사람들은 상대적인 음높이 차와 음길이차를 구분하는 능력은 특별한 교육 없이도 어린 시기에 습득된다. 따라서 허밍을 질의로 사용하는 검색 시스템에서는 표2와 같은 개연성을 고려하여 허밍 질의의 특징을 고려해야 한다.

표 2. 허밍 질의의 특징
Table 2. Distinctive Features in Humming Query

허밍 질의의 특징	설명
질의의 무작위성	질의된 허밍은 전체 곡의 임의의 마디에서부터 시작될 수 있다.
음 높이의 불확실성	질의된 허밍은 실제 음높이와는 다른 임의의 음 높이에서부터 시작되므로 조옮김 현상이 발생할 수 있다.
음 길이의 불확실성	질의된 허밍은 실제 음길이와는 다른 임의의 음 길이로 재생될 수 있다.
개수의 불확실성	질의된 허밍의 선율에서 임의로 음이 추가되거나 삭제될 수도 있다.
재생의 합리성	질의된 허밍은 음의 높이나 길이가 원 선율과 다를 수는 있으나 상대적 음의 흐름은 원 선율과 유사하다.

2.3 . 기존 연구

음악 정보를 내용 기반으로 검색하는 시스템에 대한 연구는 다음과 같다.

2.3.1 UDS 기반 음악 검색 시스템

Cornell 대학의 Ghias는 1995년 처음으로 허밍 질의를 이용한 음악 검색 시스템을 제안하였다[4]. Ghias는 허밍 질의 내에 표현된 음 높이가 검색하고자 하는 대상 멜로디의 절대 음 높이와 차이가 있다는 점에 착안하여 음간의 높이 변화를 음악 검색의 특징 정보로 활용하였다. Ghias는 음의 높이 정보를 추출하기 위해서 AutoCorrelation, Maximum likelihood와 Cepstrum 분석 방법을 사용하였다. 각 음간의 높이 차이는 U(Up), D(Down), S(Same)의 세 가지 기호로 표현된다. U는 하나의 음에서 다음 음으로 멜로디가 전개될 때 음의 높이가 높아지는 것을 의미하며 D는 멜로디가 전개될 때 음의 높이가 낮아지는 것을 의미한다. S는 음의 높이가 동일한 경우를 나타낸다. 이와 같은 형태로 허밍 질의 내에 N개의 음이 존재하는 경우, 선율 정보는 N-1개의 기호 집합으로 표현된다. <그림 1>은 동요 개구리송의 특정 부분을 Ghias가 제시한 형태로 표현한 것이다.

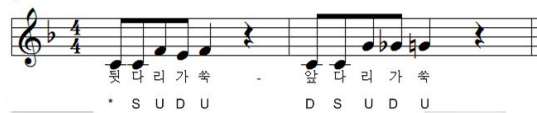


그림 1. UDS를 이용한 선율의 표현
Fig 1. Melody Presentation by Using UDS

이와 같이 음 높이의 변화를 UDS를 이용하여 세 단계로 나누어서 표현하는 경우 검색 과정을 단순화 할 수 있으나, 음악 검색을 위한 데이터의 크기가 커지는 경우 음간의 음높이의 변화를 보다 더 자세하게 표현해야만 하는 단점이 있다.

2.3.2 비교 선분의 기울기 기반 음악 검색 시스템

모/김/구/한은 음악의 선율 정보에 있어서 두 특징인 음을 나타내는 정보(상대적 음높이)와 음들 간의 관계를 나타내는 정보(기울기)를 이용하는 방법을 제시하였다 [5]. 음악에서의 음높이 정보는 Ghias가 음간 차이를 이용하는데 비해서 [5]는 음악의 모든 음표의 음높이 값의 평균에서 해당 절대 음 높이의 차를 이용하였다. 음의 길이의 경우, 각 음의 길이를 그대로 사용하였는데, '임의의 앞 음 A와 뒤 음 B와의 상대적인 음 높이차/앞 음

A의 길이'로 특징 정보를 표현했다. <그림2>는 [5]가 제시한 선율에서의 특징 정보 표현 방식을 나타낸다.

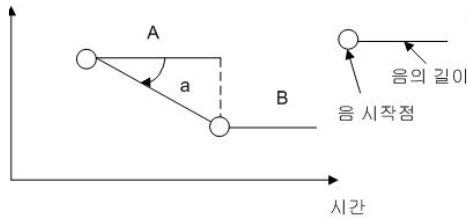


그림 2. 음 A와 음 B간의 관계를 나타내는 음의 각 a
Fig 2. Relative Music Notes Expression By Angle a

음 A와 음 B간의 관계를 평균 음에서의 높이 차로 표현한 방식은 상대적인 조바꿈이나 음악의 높이 변화에 관계없이 동일한 음악으로 간주될 수 있다. 그러나 음의 길이를 사용한 것은 전체적인 음의 빠르기 변화에 민감하게 대처할 수 없다는 단점이 있다.

2.3.3 선율 흐름/음의 높이/음의 길이 기반 음악 검색 시스템

마이크로 소프트 연구소의 Lie Lu는 Ghias의 UDS 시퀀스를 확장하였다[6]. UDS(Up/Down/Same) 시퀀스의 경우, 음간 높이와 반복의 특징 정보를 사용하는데 비해서 Lie Lu는 반복되는 음의 길이는 별도로 구분하지 않았다. 이는 허밍 질의에서 음간 차에 대한 구분에 따른 길이 추출은 용이하나 반복되는 음간 길이 추출은 상대적으로 어렵다는 점에서 착안하였다. 기본적으로 UD(Up/Down) 두 가지의 선율 정보만을 사용하였으며 각 음에 대한 음간 차이 및 음의 지속 시간 세 가지를 특징 정보로 사용하였다.



그림 3. 악보의 일부 표현
Fig 3. Melody Expression

Lie Lu가 제시한 특징 정보 추출에 따르면 <그림 3>은 다음과 같은 세 개의 특징 요소로 구성된 일련의 시퀀스 데이터로 표현 할 수 있다.

$$(*, *, *) (U, 64.1, 1) (U, 71.9, 2) (U, 124.7, 2) (U, 96.0, 3) (D, -96.0, 1) \dots \dots \dots (2.1)$$

처음에 시작된 노트는 참조할 특징 요소가 없기 때문에 '*'로만 구성된 형태이며 검색 과정에서는 사용되지 않는다. 제시된 특징 추출 정보는 선율 흐름에 음의 높이와 음의 길이 정보를 추가함으로써 검색에서의 정확성을 높일 수는 있으나, 절대적인 음의 길이를 특징 정보로 사용함으로써 빠르기 변화에 민감하게 대처하지 못하는 근본적인 문제를 해결하기에 부족한 면이 있다.

2.3.4 형태학적 특징을 이용한 음악 검색 시스템

최근에 들어 [7]에서는 [8]에서 형태 심리학적으로 제시한 음간 위상 관계의 특징 정보를 기본으로 하여 두 선율 간 유사도를 비교하는 방식을 제시하였다. [7]에서는 세 개의 음을 하나의 집합으로 구분하여 음간 높낮이의 차이에 따라서 이를 I/R(Implication/ Realization)이라는 모델 형태로 표현하였다.

<그림 4>는 [7]에서 제시한 형태학적 특징을 이용한 비교 예를 나타낸다.

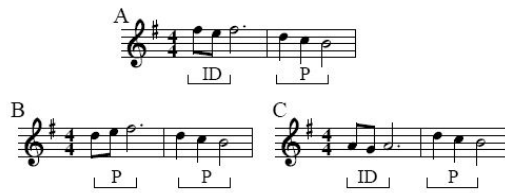


그림 4. I/R 모델을 이용한 선율간 비교
Fig 4. Example of Melody Differency by I/R Model

[7]에서는 음높이의 위상 관계를 이용하여 선율을 비교하는 방식으로 음길이에 대해서는 고려하지 않는다.

III. 허밍 질의 내 특징 정보 추출

본 장에서는 입력된 허밍 질의를 변환하는 과정과 특징을 추출하는 방법을 언급한다.

3.1 허밍 질의에서의 음별 높이/길이 추출

일반적으로 세그먼트 단위에서 주파수를 분석하여 음 높이 정보를 추출하는 경우, 비정상적인 주파수가 검출

될 수 있기 때문에 에너지 정보와 영교차율을 사용하여 허밍 질의 내의 음이 존재하는 영역과 존재하지 않는 영역을 구분한다. 에너지 정보는 기본적으로 허밍 질의 프레임 단위로 분리하기 위해서 사용한다. 추출된 에너지가 매우 낮은 레벨로 떨어졌다가 급속히 다시 올라가거나 혹은 그 이후로 일정 기간 동안 에너지의 변화가 없는 경우 새로운 음이 시작되었다고 판단한다.

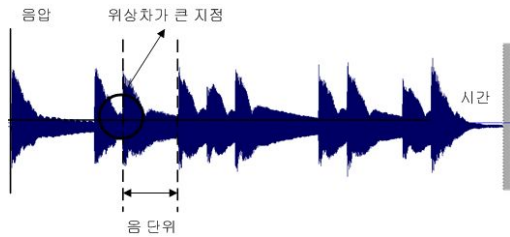


그림 5. 도식적 허밍 질의 표현과 음 단위의 구분

Fig 5. Visual Expression of Humming Query and Division of Notes

위 <그림 5>는 입력된 허밍 질의를 시간 축에 따른 음압의 변화로 표현한 것이다. 에너지 정보의 추출을 위해서 서로 겹치지 않는 10ms 단위의 프레임으로 허밍 질의를 나눈다. 제안된 알고리즘은 (3.1)과 같으며 E_k 는 k번째 프레임의 에너지 정보를 의미한다.

$$E_k = \sum_{n=1}^N x(n)^2 \dots\dots\dots (3.1)$$

영교차율은 음의 파형과 시간축이 만나는 지점으로서 정적 지점을 의미하며 계산된 에너지 정보와 함께 허밍 질의 내에 음을 추출하기 위해서 사용된다. 에너지 정보와 영교차율을 이용하여 추출된 선율 정보는 다음과 같이 음의 높이를 나타내는 주파수 정보(Hz)와 음의 지속 시간을 나타내는 정보(ms)로 표현된다.

입력된 허밍 질의에서 추출된 음의 개수가 n개라 가정하면 (3.2)와 같이 선율은 각 음의 높이(P)와 각 음의 길이(D)를 쌍으로 하는 음의 집합으로 표현한다.

$$\text{선율} = \{(f_1, d_1), (f_2, d_2), \dots, (f_n, d_n)\} \dots\dots\dots (3.2)$$

3.2 허밍 질의에서의 음간 특징 정보 추출

일반 오디오 검색 내 구간 질의는 음의 높이나 음의 재생 길이가 변화하지 않으나 허밍 질의 경우, 사용자에 재생 능력에 따라서 키의 변화나 빠르기의 변화가 발생하기 때문에 제안된 알고리즘은 음간 높이 비율과 음간 재생 길이 비율을 이용한다. 일반적으로 음역의 비율을 나타내기 위해서 순정율과 평균율이 사용된다[9]. 순정율이란 서양의 12음계를 기본으로 두 음간의 진동비가 2:3일때 이 두음을 완전 5도의 음정이라고 한다. 순정율의 경우 그 올림은 사람의 청각에 아름답게 들리지 않음과 리적 비율의 관계로 인해서 조율이 반응이 낮아져 조바꿈을 하면 결국에는 어울리지 않는 화음이 발생하게 된다. 이와는 달리 평균율은 서양에서의 음의 높낮이에 따라서 각자 동일한 비율을 배정하는 방식으로서 자연 배열음에서 음정 비는 한 옥타브에서 1:2의 음정의 비율을 가지고 전체에서의 반응 비율은 $\sqrt[12]{2}$ 로 표현된다. 허밍 질의는 사용자에게 의해서 조바꿈 현상이 발생할 수 있기 때문에 추출된 음의 주파수간 차이를 반응 간격으로 구분하기 위하여 평균율을 사용한다. 추출된 주파수는 계산의 편의를 위해서 상용 로그(log)를 취하여 소수 둘째 자리에서 반올림하여 사용한다. 평균율을 기본으로 반응 단위의 음높이 차이를 계산하는 알고리즘은 (3.3)과 같다.

$$PC = \frac{\log(f_n) - \log(f_{n-1})}{\log(TC)}$$

PC(Pitch Change) : 음높이 차

TC(Temperament_Constant) : 평균율

에서의 반응 비율

f_n : n번째 음 주파수

f_{n-1} : n-1번째 음 주파수

if $f_n = f_{n-1}$, then PC= 0 (3.3)

음 높이 추출 알고리즘은 반응 단위의 시간에 따른 음 높이의 변화를 나타내기 때문에 현재 음이 기존 음보다 높이가 낮은 경우에는 -값을 취하며 현재 음이 기존 음보다 높은 경우에 + 값을 취한다. 만약 연속된 두 음의 높이가 동일하다면 음높이 차는 0으로 설정한다.

선율 정보로부터 추출된 음의 길이는 동일한 음높이가 지속된 시간을 의미하며 ms 단위로 표현된다. 서로 다

른 상대적인 음길이의 차이는 재생 시간의 상대적 비를 이용하여 (3. 3)과 같이 계산한다.

$$DC = \frac{d_n}{d_{n-1}}$$

$DC(Duration_Change)$: 음길이 차
 d_n : n 번째 음 길이(ms)
 d_{n-1} : $n-1$ 번째 음 길이(ms) (3.4)

상대 음 길이 추출 알고리즘에서는 길이의 차가 클수록 큰 값을 갖는다. 두 음의 길이가 같은 경우에 해당 값은 1이 된다.

표 3. 음간 특징 정보 추출 예

Table 3. Examples of Distinctive Features between Notes

음의 순번	1	2	3	4
음높이(fn)	164.81	146.52	130.43	146.12
음길이(dn)	311	250	630	320
높이차	-	-2.02	-2.01	-1.61
길이차	-	0.80	2.52	0.50

표 3은 (3.3)과 (3.4)를 이용하여 음간 특징 정보를 추출한 예를 나타낸다.

IV. 두 선율 간 유사도 계산

입력된 허밍 절의는 음길리와 음높이별로 추출되어 음간 높이차와 길이차를 가지고 다른 선율과의 유사도를 비교하게 된다. 본 장에서는 추출된 특징 정보를 이용하여 두 선율간의 유사도를 계산한다.

4.1 제안 알고리즘의 순서도

본 논문에서 제시한 알고리즘의 전체적 흐름은 다음 <그림 6>과 같다. 사용자가 허밍 절의를 입력하는 경우, 허밍 절의 내 음표의 높이 및 길이 정보가 추출된다. 추출된 길이 정보를 기반으로 UDS 기반의 선율 정보를 추출하여 유사한 선율 리스트를 생성한다. 유사 선율 리스트 내의 각 선율별로 음높이 차와 음길이 차가 최소/최대가 되는 행렬을 추출하고 허밍 절의와 유사도를 비

교하여 유사도가 높은 순으로 리스트를 재 출력한다.

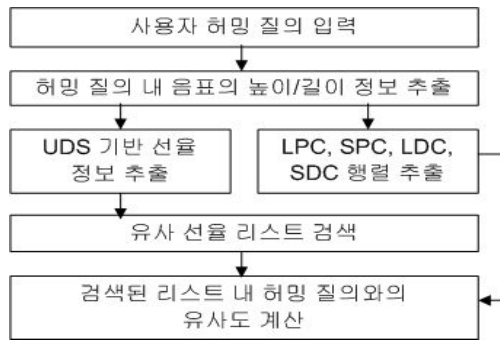


그림 6. 유사도 계산 알고리즘 흐름도
 Fig 6. Flow Chart of Similarity Computation Algorithms

4.2 선율 내 특징 행렬 추출

음높이 차와 음길리의 차의 최대/최소값을 특징 항목으로 사용하게 되는 경우, 실제 선율의 흐름이 유사하지 않더라도 유사한 선율로 선택될 수 있는 불확실성을 내포하고 있다. 따라서, 두 선율간의 유사도를 측정하는 경우, 선별된 네 개의 특징 값을 중앙값으로 갖는 일정 구간을 선택하여 비교 대상 선율 내의 동일 구간의 특징 정보들과의 각 차이를 유사도 측정을 위한 평가 항목으로 사용한다. 특징 값을 중앙값으로 갖는 선율내 구간은 $2 * (2k+1)$ 형태의 2차원 행렬로 표현되며 선별된 특징 값을 중앙값으로 취한다. 행렬의 중앙값으로 선택된 특징 값에 따라서 첫째 중앙값이 가장 큰 음높이 차를 나타내는 경우 LPC(Longest Pitch Change), 가장 작은 음높이 차를 나타내는 경우 SPC(Shortest Pitch Change) 구간이라 칭하며 음길이 차의 최대/최소 값에 따라서 LDC(Longest Duration Change), SDC(Shortest Duration Change)라 칭한다.

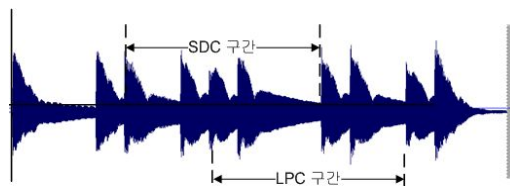


그림 7. K=1일때, SDC 구간과 LPC 구간
 Fig 7. SDC region and LPC region when K=1

<그림 7>은 k의 값이 1인 경우, 음 높이 차가 가장

큰 LPC 행렬과 음 길이차가 가장 작은 SDC 행렬의 예를 나타낸다.

4.3 두 선율 간 유사도 비교 알고리즘

특징 정보로 추출된 LPC, SPC, LDC, SDC는 특징 구간별 행렬을 나타내며 실제적인 유사도 계산은 $2 * (2k+1)$ 행렬내의 각 요소간 거리의 차를 더하여 이를 산출한다.

$$SI_{LPC} = \frac{1}{\sum_{n=1}^2 \sum_{m=1}^{2*2k+1} |ContourA_{nm} - ContourB_{nm}|}$$

ContourA_{nm} : 선율 A내의 n*m번째 행렬 원소 값
 ContourB_{nm} : 선율 B의 n*m 번째 행렬 원소 값
 k : 연관 행렬의 크기를 결정하는 가중치
 SI_{LPC} : 최대 음높이차가 중앙값을 이루는 구간 간의 유사도 값 ... (4.1)

식 (4.1)은 임의의 두 선율로부터 음높이차가 최대가 되는 두 행렬의 유사도를 계산하는 방법으로 음높이차가 최소가 되는 행렬의 유사도 계산과 음길이차가 최대/최소가 되는 행렬의 유사도 계산도 동일한 형태로 이루어진다. 특징 정보를 중앙값으로 하는 구간 간의 유사도 값은 행렬을 이루고 있는 각 원소간 거리의 절대 값을 총 누적한 값을 역분포 형태를 취하는데, 이는 결과 값이 작을수록 두 행렬간의 유사성이 높다는 것을 의미한다. 위 알고리즘은 LPC내의 유사성만을 비교한 것이며, 두 선율 간의 음높이 차와 음길이 차의 유사성 비교를 위해서 (4.2)와 같이 LPC, SPC, LDC, SDC의 연관 집합 간에 항목별 연산의 평균 값을 두 선율의 유사도를 평가하는 기준으로 한다.

$$SI_{contour} = \frac{SI_{LPC} + SI_{SPC} + SI_{LDC} + SI_{SDC}}{4}$$

SI_{contour} : 두 선율간 유사도 값
 SI_{LPC} : 음높이차가 제일 큰 값이 포함된 행렬
 SI_{SPC} : 음높이차가 제일 작은 값이 포함된 행렬
 SI_{LDC} : 음길이차가 제일 큰 값이 포함된 행렬
 SI_{SDC} : 음길이차가 제일 작은 값이 포함된 행렬
 (4.2)

V. 실험 및 평가

본 실험에서 비교 대상이 되는 선율 정보는 웹 검색을 통해서 구한 임의의 150개의 미디 파일을 이용한다. 미디 파일에서 주선율이 녹음된 트랙을 분류하여 추출하며, 트랙 분류를 위해서 Cakewalk사의 Sonar라는 미디 시퀀서 프로그램을 사용한다. 허밍 질의는 웨이브 파일 형태로 16bit Mono에 22.1Mhz로 샘플링하여 저장하며 저장된 웨이브 파일은 jAudio를 이용한 주파수 분석을 통해서 허밍 질의내의 음 높이 및 음 길이를 추출한다[10].

실험 평가의 방법은 다음과 같다. 행렬의 크기를 결정하는 가중치 k를 1~3까지 단계별로 설정 하고 매 단계마다 5회에 걸쳐 검색하고자 하는 선율의 허밍 질의를 입력 데이터로 하여 선율을 검색한다. 검색 시 마다 150개의 음악 리스트가 유사도가 높은 순으로 순번에 맞게 출력되며, 사용자 자신이 검색하고자 했던 선율이 포함된 음악의 순위가 전체 리스트에서 몇 번째 순번인지를 확인한다. 표 4는 허밍 질의 시, 찾고자 하는 음악의 매 회당 검색 순위와 평균 값을 나타낸다.

표 4. 허밍 질의 시 해당 음악의 순번
 Table 4. Sequence of the Matched Song

	해당 음악의 순번					AVG
	1st	2st	3st	4st	5st	
k=1	5	7	19	7	6	9.8
k=2	3	4	14	5	4	6
k=3	1	2	9	1	2	3

k의 값이 증가할수록 해당 음악의 순번은 가장 높은 순위로 올라가게 되는데, 이는 특징 행렬의 길이가 길어질 수록 인접한 특징들의 평균값을 취하므로 보다 유사도의 측정의 정도가 보다 정밀해 짐을 알 수 있다. 세 번째 허밍 질의에 경우 상대적으로 다른 허밍 질의에 비해서 낮은 순위를 갖게 되는데 이는 사용자가 자신이 검색하고자 하는 선율을 허밍 질의로 올바르게 표현하지 못하는 경우를 의미한다.

기존 선율 유사도를 위한 알고리즘들이 허밍 질의에서도 사용하기 적합한지 여부를 비교하기 위해서 2.2에서

인급한 허밍 질의의 특징 중에 다음 네 가지 관점에서 상대적으로 비교한다.

- 무작위성 : 허밍 질의에 제한을 두지 않는가? (질의의 길이 및 시작 마디)
- 음 높이의 불확실성 : 허밍 질의에 조옮김이 발생한 경우, 두 선율 간 분별이 가능한가?
- 음 길이의 불확실성 : 허밍 질의 내 서로 다른 음길이 변화가 발생한 경우, 두 선율 간 분별이 가능한가?

UDS 기반의 음악 검색 시스템[4]과 형태학적 특징 기반의 음악 검색 시스템[8]의 경우, 연속된 음간의 높이의 변화차를 특징 정보로 사용하기 때문에 무작위로 질의된 선율에 대해서도 순차적으로 문자열을 비교하는 방식으로 곡 중간에서의 검색이 가능하다. 허밍 질의 내에 조옮김 현상이 발생하는 경우, 선율의 상대적인 높이 변화차를 특징정보로 사용하기 때문에 선율의 음높이 흐름이 동일한 선율의 검색이 가능하다. 그러나 음길이를 선율 간 유사도 비교 시 고려하지 않아 유사한 음높이의 흐름을 가지고 있는 선율들에 대해서 동일한 유사도 값을 부여한다는 문제점이 있다. <그림 8>은 실험을 위한 선율 정보 중에 Comedian가 작곡한 'Veronika, der Lenz ist da'라는 곡과 Mozart가 작곡한 'Don Giovanni I finale'는 곡의 일부분으로서 음길이 차만을 고려하는 경우 서로 다른 선율에 동일한 유사도 값을 부여하는 경우이다.



그림 8. 동일한 유사도 값을 갖는 경우의 예
Fig 8. Example of Same Similarity Score

비교 선분 기반의 유사도 비교 알고리즘[5]과 선율의 음높이/음길이 기반의 유사도 비교 알고리즘(6)의 경우, [4]와 [8]과 마찬가지로 상대적인 음 높이를 사용하기 때문에 허밍 질의 내에 조옮김과는 무관하게 유사한 선율을 검색할 수 있다. [4]의 경우에는 절대적인 음길이

를 선율의 특징 정보로 사용하기 때문에, 사용자의 허밍 질의 내에 모든 음의 길이가 정확하게 일치하지 않는 경우, 동일한 음높이의 흐름을 가지고 있는 선율들내에서 서로 다른 유사도를 부여할 수 없다는 단점이 있다. 제안 유사도 비교 알고리즘의 경우 기본적으로 초기 유사도 검색 시 기존 UDS 기반 유사도 비교 알고리즘을 이용하여 유사한 선율 리스트를 생성 후, 음길이 차와 거리 차의 최대/최소값을 기반으로 생성된 유사 선율 리스트 내에 선율들을 각각 허밍 질의와 비교하게 된다. 따라서 허밍 질의의 무작위성과 음 높이의 불확실성에서 동일한 선율의 검색이 가능하며 상대적인 음길이 차를 이용함으로써 <그림 8>과 같은 경우에도 <그림 9>와 같이 차별된 유사도 계산을 가능케 한다.



그림 9. 음길이 차를 고려한 경우의 예
Fig 9. Example of the Applying Relative Note Durations

VI. 결론 및 향후 연구 과제

본 연구에서는 기존 키보드를 이용한 주선율 입력이나 악보에 기보된 형태의 정보를 질의로 사용하는 음악 검색 시스템이 허밍 질의를 이용한 음악 검색 시스템에 적용될 수 없는 불확실성과 무작위성에 대하여 언급하였다. 이러한 이유는 사람의 생리적인 구조로 인해 야기되는 것으로 사람이 선율을 기억하는 방식이 음의 절대 높이와 길이가 아닌 상대적인 음의 흐름을 기억하는데 있다. 본 연구에서는 사람이 선율을 기억할 때 음간의 높이차와 길이차를 이용한다는 점에 착안하여 두 선율 간에 유사도 계산 알고리즘을 제안하였다. 제안된 알고리즘은 기존에 음높이 차만을 가지고 유사도를 계산하는 경우 방법에 비해서 음길이 차를 유사도 측정을 위한 요소로 적용함으로써 음높이 차의 관점에서 동일한 유사도

값을 갖는 경우에도 이를 구분할 수 있는 방법론을 제시하였다. 향후에는 본 논문에서 제안된 유사도 계산 알고리즘을 이용하여 내용 기반 음악 검색 시스템을 개발함으로써 제안된 알고리즘의 활용성을 보이고자 한다.

참고문헌

- [1] 네이버 국어사전, <http://krdic.naver.com>, 2006. 7
- [2] 나가오 마코토 외 5인, “문자와 소리의 정보처리”, 한국 학술정보, 2000. 8
- [3] Richard Moore, Paul A. Wheeler, “Science of Sound, 3rd edition”, Addison-Wesley, San Francisco, 2002
- [4] Ghias, A., Logan, J., Chamberlin, D., and Smith, B.C., “Query By Humming, Musical Information Retrieval in an Audio Database,” Proc. of ACM Multimedia Conf., pp.231-236, 1995.
- [5] 모종식, 김소영, 구경이, 한창호, 김유성, “선율의 음높이와 리듬 정보를 이용한 음악의 유사도 계산 알고리즘”, 한국정보처리학회 논문지, 제7권, 제12호, pp. 3762-3774, 2000. 12.
- [6] Lu, L., You, H., and Zhang, H.-J. “A new Approach to query by humming in music retrieval”, In Proceedings of IEEE International Conference on Multimedia and Expo, 2001.
- [7] M. Grachten, J.-L. Arcos, and R. Lopez deM-antaras. “Melodic Similarity: Looking for a Good Abstraction Level”, Proc. ISMIR, 2004.
- [8] E. Narmour. “The Analysis and cognition of basic melodic structures : the implication - realization model.”, University of Chicago Press, 1990.
- [9] 조효임, 이동남, 주대창, “새로운 음악 통론”, 학문사, 1998.
- [10] Daniel McEnnis, Cory Mckay, Ichiro Fujinaga, Philippe Depalle, “JAudio : A Feature Extraction Library”, In Proceeding of ISMIR 6th Conf., pp.600-603, 2005.

저자 소개



오 동 열

1999년 : 경희대학교 컴퓨터 학사
 2002년 : 숭실대학교 컴퓨터 공학석사
 2004년 : 숭실대학교 컴퓨터 공학박사
 수료
 관심분야 : 유비쿼터스 컴퓨팅, P2P, 멀티미디어



오 해 석

1975년 :서울대학교 응용 수학과 학사
 1981년 :서울대학교 계산 통계학 석사
 1989년 : 서울대학교 계산통계학 박사
 1982년 ~ 2003년 : 숭실대학교 정보 과학대학 교수
 2003년 ~ 현재 : 경원대학교 소프트 웨어대학 교수
 관심분야 : 멀티미디어, 데이터베이스, 정보보호