

# 내용 기반 음악 정보 검색을 위한 음악 구성 형식을 고려한 대표 선율의 추출 및 색인

구 경 이<sup>†</sup> · 임 상 혁<sup>††</sup> · 이 재 현<sup>†††</sup> · 김 유 성<sup>††††</sup>

## 요 약

최근 내용 기반 음악 정보 검색 시스템에서는 사용자의 응답 시간을 단축시키기 위해 음악의 대표성을 갖는 선율을 추출하여 색인하고, 검색시 이를 이용한다. 음악에서 대표성을 갖는 선율은 해당 음악을 대표하여 사용자가 기억하고 있으며, 사용자가 질의 선율로 사용할 가능성이 높아야 한다. 그러나, 기존의 내용 기반 음악 정보 검색 시스템에서는 음악 구성 형식을 고려하지 않기 때문에 음악 구성 형식에 따라 반복, 대조되는 선율들을 해당 음악을 대표하는 선율로 추출하지 못한다. 본 논문에서는 해당 음악을 대표하는 선율을 추출하기 위하여 한 음악에서 음악 구성 형식을 반영하여 일정한 유사도내에서 반복되는 선율들을 해당 음악의 반복 주제 선율로 추출한다. 또한, 사용자가 일반적으로 인지하는 첫 동기 선율과 절정 선율에 해당되는 선율들을 인식하여 대표 선율에 추가한다. 본 논문에서 제안된 시스템의 성능을 평가하기 위해 프로토타입을 구현하고, 다른 선율 색인과 비교 실험하였다. 실험 결과, 대표 선율 색인은 전체 동기를 색인한 경우에 비해 34%의 적은 저장 공간으로 색인을 구성하여 사용자의 응답 시간을 단축시켰다. 또한, 반복 주제 선율만으로 색인한 경우에 비해 사용자가 질의 가능성이 높은 첫 동기 선율, 절정 선율 등을 대표 선율로 색인하기 때문에 추가적으로 20%의 색인 공간이 필요하였으나, 다양한 사용자의 질의 선율에 대해 검색의 정확성을 증진할 수 있음을 보였다.

## Extraction and Indexing Representative Melodies Considering Musical Composition Forms for Content-based Music Information Retrievals

Kyong-I Ku<sup>†</sup> · Sang-Hyuk Lim<sup>††</sup> · Jae-Heon Lee<sup>†††</sup> · Yoo-Sung Kim<sup>††††</sup>

## ABSTRACT

Recently, in content-based music information retrieval systems, to enhance the response time of retrieving music data from large music database, some researches have adopted the indexing mechanism that extracts and indexes the representative melodies. The representative melody of music data must stand for the music itself and have strong possibility to use as users' input queries. However, since the previous researches have not considered the musical composition forms, they are not able to correctly catch the contrast, repetition and variation of motif in musical forms. In this paper, we use an index automatically constructed from representative melodies such like first melody, climax melodies and similarly repeated theme melodies. At first, we expand the clustering algorithm in order to extract similarly repeated theme melodies based on the musical composition forms. If the first melody and climax melodies are not included into the representative melodies of music by the clustering algorithm, we add them into representative melodies. We implemented a prototype system and did experiments on comparison the representative melody index with other melody indexes. Since, we are able to construct the representative melody index with the lower storage by 34% than whole melody index, the response time can be decreased. Also, since we include first melody and climax melody which have the strong possibility to use as users' input query into representative melodies, we are able to get the more correct results against the various users' input queries than theme melody index with the cost of storage overhead of 20%.

키워드 : 내용 기반 음악 검색(Content-Based Music Retrieval), 대표 선율 색인(Representative Melody Index), 음악 구성 형식(Musical Composition Forms)

### 1. 서 론

최근 증가된 멀티미디어 정보의 요구 중 음악 데이터는

그 쓰임의 다양함이나 활용 범위가 넓은 측면에서 멀티미디어 정보의 독립적인 한 분야로 다루어질 수 있다[1]. 따라서 존재하는 수많은 음악 데이터에 대하여, 음악 데이터 베이스를 구축하고 이를 이용하여 음악 데이터를 효과적으로 저장, 관리함으로써 사용자가 원하는 음악 데이터를 빠르고 정확하게 검색할 수 있도록 하는 음악 데이터 검색의 방법이 지원되어야 한다. 기존의 음악 정보 검색 시스템이

\* 본 연구는 한국과학재단 지정 인천대학교 동북아전자 물류연구센터 지원에 의한 것임.  
† 준 회원 : 인하대학교 대학원 전자계산공학과  
†† 준 회원 : 코난테크놀로지 연구원  
††† 준 회원 : (주)인포클루 연구원  
†††† 통신회원 : 인하대학교 정보통신공학과 교수  
논문접수 : 2003년 11월 17일, 심사완료 : 2004년 1월 19일

나 음악 스트리밍 사이트에서 사용자는 주로 음악 데이터가 가지는 제목이나 가수, 작곡가와 같은 메타데이터를 이용하여 검색한다. 그러나, 이러한 메타데이터 기반 음악 정보 검색 시스템에서 사용자가 자신이 검색하기를 원하는 곡의 제목이나 가수를 모른다면, 원하는 곡을 검색하기 위해 질의를 작성할 수 없다[2]. 이와 달리 실제적으로 사용자는 자신이 찾고자 하는 곡에 대하여 메타데이터보다는 곡이 가지고 있는 첫 동기 선율, 절정 선율, 또는 반복되는 선율을 기억하고 있다[3]. 따라서, 음악 정보 검색에 있어서 질의의 방법으로 사용자가 기억하고 있는 선율의 일부분을 질의 선율로서 입력받고, 입력받은 질의 선율 자체가 가지고 있는 정보를 이용하여 검색하는 내용 기반 음악 정보 검색 시스템이 제안되고 있다[1, 3, 4].

내용 기반 음악 정보 검색 시스템에서 음악 데이터를 검색할 때, 전체 선율 정보를 비교하여 검색 결과를 구성하면, 사용자의 응답 시간이 증가하며 결과적으로 시스템의 실효성이 감소된다. 이와 달리 음악 데이터의 대표성을 갖는 일부분을 선정하여 색인을 구성하고, 색인된 선율을 이용해 검색을 수행하면 사용자 응답 시간을 단축시킬 수 있다[5-9]. 색인된 일부분은 해당 음악을 대표할 수 있고, 사용자가 해당 음악을 검색하기 위한 질의 선율로 사용할 가능성이 높은 부분, 즉 음악의 첫 동기 선율, 절정 선율, 또는 일정 유사도내에서 반복되는 패턴인 반복 주제 선율 이어야 한다. 그러나, 기존의 내용 기반 음악 정보 검색 시스템 [5-8]에서는 첫 동기 선율 또는 정확하게 반복되는 선율만을 색인 선율로 사용하였다. 첫 동기 선율만을 색인한 경우 첫 동기 선율이 아닌 다른 선율로 질의를 제기하면, 색인으로부터 검색할 수 없기 때문에 전체 데이터베이스를 검색해야 원하는 데이터를 정확하게 얻을 수 있다. 또한, 정확하게 반복되는 선율만을 해당 음악의 반복 주제 선율로 선택하는 경우에는 작곡가가 음악의 주제를 위해 일정 범위 내에서 장식적으로 변화시키는 변주(a variation on a theme)를 서로 다른 선율로 인지하게 된다. 이와 달리, 한 음악에서 반복 주제 선율은 일정 유사도내에서 3번 이상 반복된다는 가정 하에 주제 그래프 기반 클러스터링 알고리즘을 이용하여 반복적으로 변화된 선율들을 클러스터링하고, 해당 클러스터로부터 반복 주제 선율을 선택하여 색인하였다[9]. 그러나, 작곡가가 음악 구성을 위해 앞, 뒤악절끼리 반복, 대조시킨 경우, 음악 구성 형식에 따라 일정한 유사도내에서 반복되는 선율의 횟수가 2회 이상일 수 있다. 기존의 그래프 기반 클러스터링 알고리즘을 이용하여 주제 선율 추출시 반복되는 선율의 클러스터에 포함되는 유사한 선율의 최소수가 3이상이어야 하기 때문에 작곡가가 한 선율을 일정한 유사도내에서 2회 반복시킨 경우에는 반복 주

제 선율을 정확히 추출하지 못한다. 또한, 첫 동기 선율이나 절정 선율과 같이 음악 구성 형식에서 사용자가 일반적으로 인지하고 있어 질의 선율로 사용할 가능성이 높은 선율에 대해 질의할 경우 해당 선율을 검색하지 못한다.

이를 해결하기 위해 본 논문에서는 기존의 그래프 기반 클러스터링 알고리즘을 확장하여, 음악 구성 형식에 따라 전체에 그대로 혹은 장단이나 음높이에 약간 변화를 주어 입의의 위치에 두 번 이상 반복되는 주제 선율을 추출할 수 있도록 한다. 또한, 음악 구성 형식에서 해당 음악의 가장 처음으로 주어지는 첫 동기 선율과 음악적 구성 요소들 중 최고 정점으로서 사용자에게 자연스럽게 인식되는 절정 선율이 색인에 포함되도록 한다. 따라서, 본 논문에서는 음악의 전체적인 표현을 단적으로 사용자에게 지각시켜 질의로 사용될 가능성이 높은 대표 선율을 자동적으로 정확하게 추출하여 해당 음악 데이터를 대표하는 색인으로 구성함으로써, 내용 기반 음악 정보 검색에서 사용자 질의 선율에 대해 정확하고 빠른 검색이 가능하도록 하였다. 본 논문에서 제안된 시스템의 성능을 평가하기 위해 다른 선율 색인과 비교 실험하였다. 실험 결과, 대표 선율 색인을 이용한 내용 기반 음악 정보 검색 시스템은 전체 선율로 색인을 구성하는 방법에 비해 34% 적은 양으로 색인 구성이 가능하였다. 따라서, 전체 선율 정보에 대한 비교 없이 대표 선율 색인을 이용한 검색이 수행되기 때문에 사용자의 응답 시간이 단축된다. 또한, 사용자가 질의에 사용할 가능성이 높은 첫 동기 선율과 절정 선율을 포함시켰기 때문에 반복 주제 선율만으로 색인하는 경우에 비해 대표 선율을 색인하는 경우 20%의 저장 공간이 더 요구되었다. 하지만, 사용자가 기억하고 있으며, 사용자가 질의 선율로 사용할 가능성이 높은 선율들을 대표 선율로 구성하였기 때문에 다양한 사용자의 질의에 대해 검색의 정확성을 증진할 수 있음을 보였다.

본 논문의 구성은 다음과 같다. 2장에서는 선율의 일부를 색인한 기존 연구에 대해서 기술한다. 3장에서는 내용 기반 음악 정보 검색을 위해 음악 구성 형식이 고려된 대표 선율의 추출 및 색인 알고리즘을 설명한다. 4장에서는 실험을 통해 제안된 알고리즘의 적합성을 검증한다. 마지막으로 5장에서 결론 및 향후 연구 방향을 기술한다.

## 2. 관련 연구

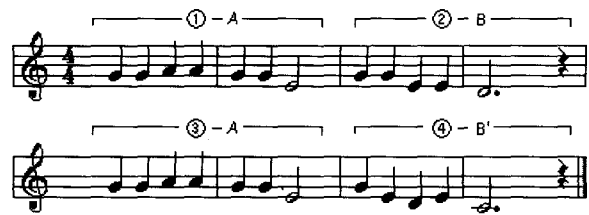
기존에 제안된 내용 기반 음악 정보 검색 시스템에서는 전체 선율 정보를 이용한 색인 대신에 해당 데이터의 일부분을 선정하여 색인을 구성하고, 검색시 구성된 색인을 이용하여 사용자 응답 시간을 단축시키고자 하였다. 선율의

일부를 색인한 연구에서 [5, 10]은 사용자가 음악의 첫 동기 선율을 많이 기억하고 이를 이용하여 질의 선율을 작성할 것이라는 점을 고려하여 질의 선율을 데이터베이스내의 모든 음악의 전체 문자열을 대상으로 비교하기 전에 첫 동기 선율 색인에서 검색하도록 하였다. 그러나, 첫 동기 선율을 제외한 다른 동기 선율로 질의시 전체 데이터베이스를 검색해야만 원하는 곡의 정보를 검색할 수 있고, 따라서 검색 시간이 증가된다. [6-8]에서는 음악에서 정확하게 반복되는 패턴을 색인으로 구성한다. [6]은 음의 게이름을 이용한 문자열을 사용하기 때문에 조옮김 등의 경우, 다른 선율 정보로 인식하는 문제점을 갖는다. 또한, 음악에서 선율이 같은 구조를 지니지만 음만 변화되는 경우와 또는 음의 길이 정보만 바뀌는 경우 등 일정한 유사도내에서 반복시킨 반복 주제 선율을 찾지 못한다. [7]은 상대적인 음높이와 음길이를 가지는 벡터로서 선율 정보를 표시하고, n-그램 방식을 이용해 반복 패턴을 추출하였지만, 유사한 변형 구조의 반복 패턴은 추출하지 못한다. [8]은 반복되는 가장 긴 패턴을 찾고, 이러한 패턴 하나를 색인으로 구성한다. 그러나, 단위 구분 없이 생성된 패턴에 비해 [9]처럼 단위 구분으로 생성된 패턴이 효율적이다[11]. [9]는 [12]에서 제안한 그래프 기반의 클러스터링 알고리즘을 수정하여 유사한 변형 구조의 반복 패턴을 반복 주제 선율로 추출하는데 사용하였다. 데이터베이스 설계에서 단편화(fragmentation)를 위해 [12]는 애트리뷰트 유사도 행렬을 이용해서 비슷한 애트리뷰트들을 클러스터링한다. [12]에서 제안된 클러스터링 알고리즘은 유사도가 낮은 애트리뷰트들도 하나의 클러스터에 포함되지만, 음악의 반복 주제 선율을 추출시 유사도가 낮은 선율들을 클러스터링할 필요가 없으므로 이를 제거한다.

[9]는 곡 전체에 그대로 혹은 장단이나 음높이에 약간 변화를 주어 임의의 위치에 3번 이상 반복되는 선율[13]을 주제 선율로 추출하여 색인하였다. 해당 곡의 반복 주제 선율을 추출하기 위해 동기별로 특성이 추출된 데이터에 대해 동적으로 클러스터를 구성한다. 그러나, [9]는 음악 구성 형식을 고려하지 않고, 처음으로 클러스터링 되어지는 동기의 수를 3이상으로 규정한다. 음악 구성 형식은 음의 연결에 의하여 작곡자의 사상과 감성을 질서 있게 통일시키는 동시에 그것을 변화 있게 전개시키는 음악적인 계획[14]이다. 음악 계층 구조에서 악곡을 구성하는 가장 작은 단위는 동기(motif)이며, 동기 두 개가 모여서 작은악절(phrase)이 되고, 작은악절 두 개가 모여서 큰악절(period)이 된다. 음악 구성 형식에서 큰악절 두 개로 구성된 음악(두 도막 형식, 8개의 동기)의 경우, 앞큰악절의 일부분을 뒤큰악절에서 대조를 주어 변화시키고, 뒤큰악절에서 앞큰악절의 일부분을 모방해 반복시켜 통일성을 갖춘다. 또한, 큰악절 한 개로

구성된 음악(한 도막 형식, 4개의 동기)의 경우, 뒤작은악절은 앞작은악절에 대조 및 반복을 주어 변화시키기도 한다 [15, 16]. 따라서, 한 음악은 음악 구성 형식에 따라 일정한 유사도 내에서 반복되는 선율의 수가 2개 이상일 수 있다.

예를 들어, (그림 1)과 같은 한 도막 형식의 동요 “학교종”은 A-B-A-B'의 음악 구성 형식이다. 대표 선율 추출을 위해서는 유사한 선율간의 클러스터 구성시, A-A와 B-B'로 클러스터링 되어야 하지만, 사이클 형성을 위해 [9]에서는 선율의 수를 3 이상으로 규정하여 A-A-B-B'로 클러스터된다.



(그림 1) 한 도막 형식의 동요 “학교종”

이러한 예는 두 도막 형식의 동요 “고향의 봄-ABCDEB'CD”에서도 동일하다. 또한, 음악 구성 형식에서 사용자가 일반적으로 인지하는 첫 동기 선율과 절정 선율을 색인하지 않기 때문에 만일 첫 동기 선율이 반복 주제 선율로 추출되지 않는 경우에 사용자가 첫 동기 선율로 질의를 구성할 경우 해당 음악 데이터를 색인에서 검색할 수 없다.

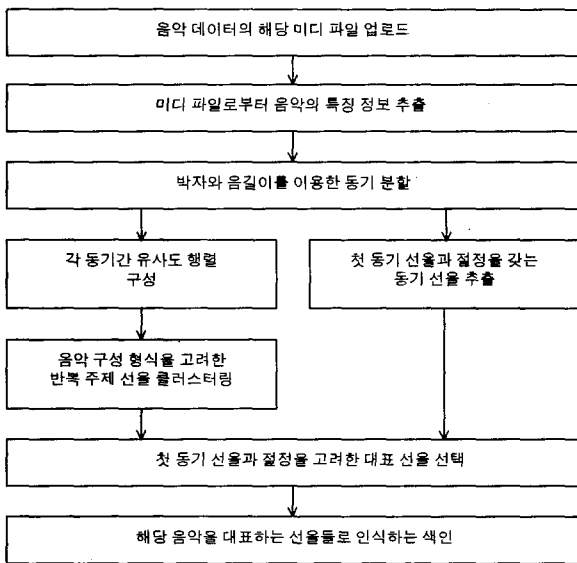
이를 해결하기 위해 본 논문에서는 기존의 그래프 기반 클러스터링 알고리즘[9]을 확장하고, 음악 구성 형식이 고려된 반복 주제 선율 및 첫 동기 선율과 절정 선율을 포함하는 대표 선율을 추출하여 해당 음악 데이터를 대표하는 선율로 색인한다.

### 3. 음악 구성 형식 기반의 대표 선율 추출 및 색인

#### 3.1 음악의 대표 선율 추출 메커니즘

본 논문에서 제안된 음악 구성 형식을 고려한 대표 선율 추출 메커니즘은 어떤 음악의 전체적인 표현을 단적으로 지각시키거나, 그 곡을 듣는 사용자에게 해당 음악을 자연스럽게 기억하도록 하는 유사 반복 주제 선율, 첫 동기 선율, 절정 선율을 해당 음악으로부터 추출한다. 음악에서 동기는 그 자체가 하나의 유형으로서, 곡에서 의미를 갖는 최소의 단위이다. 또한, 동기는 음악의 계층 구조에서 음악을 구성하는 식별 가능한 음악적 단위로서 선율이나 리듬을 움직이는 가장 기본적인 요소가 된다. 즉, 음악 구성에서 동기 단위의 반복, 또는 변주나 확대, 축소에 의해서 발전되어 하나의 곡이 형성된다[16]. 따라서, 본 논문에서 내용

기반 검색을 위한 대표 선율 추출 과정은 동기 단위로 이루어지며, 해당 과정은 (그림 2)와 같다. 대표 선율 추출 과정은 기존에 제안된 그래프 기반 클러스터링 알고리즘[9]을 확장하고, 음악 구성 형식이 고려된 반복 주제 선율 및 첫 동기 선율과 절정 선율을 포함하는 대표 선율을 자동적으로 추출하여, 해당 음악 데이터를 대표하는 선율로 색인한다.



(그림 2) 음악 데이터로부터 대표 선율 추출 과정

먼저, 음악 데이터의 해당 미디 파일로부터 박자, 음높이, 음길이를 특징 정보로 추출한다. 추출된 정보 중에서, 박자와 음길이를 이용하면, 한 음악을 동기별로 분할할 수 있다. 각 동기간 유사도를 계산하고[17], 한 음악 데이터내의 각 동기들에 대해 계산된 유사도 값을 이용하여 유사도 행렬을 구성한다. 또한, 사용자에게 인식되기 쉬운 첫 동기 선율과 절정 선율을 추출한다. 3.2절에 소개하는 음악 구성 형식을 고려한 반복 주제 선율 클러스터링 알고리즘을 이용하여 서로 유사한 동기들을 클러스터링시킨다. 각각의 클러스터는 음악 데이터에서 일정 유사도내에서 변화를 주며 반복적으로 사용되는 선율들의 모임으로서 해당 음악의 반복 주제 선율이라고 할 수 있다. 클러스터링의 후처리로서 의미 없는 클러스터를 제거하게 된다. 클러스터링된 선율들 내에서 대표 선율 추출시 첫 동기 선율과 절정 선율이 포함될 경우 우선 순위를 부여한다. 포함하지 않을 경우, 추출된 대표 선율이 클러스터에 포함된 모든 동기 선율을 가장 잘 대표할 수 있도록 하기 위해 클러스터 내의 모든 동기선율을 다차원 M-tree[18]에 표현할 때 중심에 가까운 선율이 해당 클러스터를 대표하는 선율로 선택되도록 한다. 선택된 대표 선율에 첫 동기 선율과 절정 선율이 포함되지 않은 경우, 이를 포함하여 해당 음악의 대표 선율로 색인되

도록 한다. 즉, 본 논문에서는 음악 구성 형식이 고려된 반복 주제 선율 및 첫 동기 선율과 절정 선율을 포함하는 대표 선율을 추출하여 해당 음악 데이터를 대표하는 선율로 색인한다. 따라서, 사용자가 질의에 사용할 가능성이 높은 대표 선율을 음악 화일로부터 추출하여 색인으로 구성하고, 검색시 전체 데이터를 검색하는 대신 구성된 색인을 이용함으로써 사용자 응답 시간을 단축시킬 수 있다.

3.2 음악 구성 형식을 고려한 반복 주제 선율 클러스터링

음악의 반복 주제 선율 추출 방식은 데이터베이스 설계의 단편화처럼 유사도 행렬로부터 몇 개의 동기 그룹으로 클러스터하는 것과 비슷하기 때문에 수정된 그래프 기반 클러스터링 알고리즘을 적용한다[9]. [12]에서 제안된 클러스터링 알고리즘은 유사도가 낮은 에트리뷰트들도 하나의 클러스터에 포함되지만, 음악의 반복 주제 선율을 추출하기 위한 클러스터링 알고리즘에서는 유사도가 낮은 선율들을 클러스터링할 필요가 없으므로 후처리 과정으로서 이를 제거한다. [9, 12]는 각 동기간 유사도 행렬로부터 유사한 동기를 클러스터링하기 위해, 먼저 해당 행렬을 완전히 연결된 유사도 그래프로 간주한다. 그래프 기반 클러스터링 알고리즘에 사용된 표기와 용어[9, 12]는 다음과 같다.

- 대문자 또는 숫자는 노드를 의미한다. 즉 음악의 동기들이다.
- 소문자나 (1-2)같이 두 개의 괄호로 닫힌 번호들은 두 개의 노드 사이의 간선(edge)을 의미한다.
- $S(e)$ 는 두 개의 동기 사이의 간선  $e$ 의 유사도 값을 의미한다.
- 원시 사이클(Primitive cycle)은 유사도 그래프의 사이클을 의미한다.
- 클러스터 사이클(Cluster cycle)은 한 개의 사이클 노드를 포함한 원시 사이클을 의미한다. 논문에서 다른 설명이 없으면, 사이클은 클러스터 사이클을 의미한다.
- 사이클 완성 간선(Cycle completing edge)은 한 사이클을 완성시키는 간선을 의미한다.
- 사이클 노드(Cycle node)는 사이클 완성 간선 중에서 먼저 선택되어진 노드를 의미한다.
- 전임 간선(Former edge)은 마지막 절단과 사이클 노드사이에 선택되어진 간선을 의미한다.
- 사이클 간선(Cycle edge)은 사이클을 형성하는 간선을 의미한다.
- 사이클의 확장(Extension of a cycle)은 사이클 노드에서 선회해 사이클을 확장한다.
- 후보 클러스터(Candidate cluster)는 클러스터링의 전처리 결과로서 한계치에 의해 제거되기 전의 클러스

터들을 의미한다.

- 사이클의 크기(Size of cycle) : 사이클에 포함되는 노드의 수를 의미한다.

유사도 그래프로부터 사이클의 형성은 간선  $a, b$ 가 선택된 후, 다음으로  $c$ 가 선택되어 원시 사이클( $a, b, c$ )이 형성될 때, 전임 간선이 없거나,  $S(\text{전임 간선}) \leq S(\text{모든 사이클 간선들})$ 이면 한 사이클이 만들어질 가능성이 존재한다. 사이클의 확장은 사이클 노드의 결정후 간선  $d$ 가 고려되면, 사이클의 확장 가능성이 체크된다. 확장의 가능성은  $S(d) \geq \text{MinS}(a), S(b), S(c)$ , 또는  $S(\text{간선 } d \text{에 연결된 다른 노드와 사이클 노드 사이의 사이클 완성 간선}) \geq \text{MinS}(a), S(b), S(c)$ 의 조건을 만족할 경우 확장된다. 클러스터 사이클로부터 분할은 새로 고려된 간선이 확장 가능성을 만족시키지 못하는 경우에 분할하거나, 전임 간선이 확장 가능성을 갖지 못하는 경우 전임 간선과 함께 분할한다. 분할 후에 저장된 한계치 값보다 절단 값이 클 경우 이 절단 값으로 한계치 값을 갱신한다. 후처리로서, 의미 없는 클러스터링 간선들은 한계치 값을 이용하여 후보 클러스터로부터 제거어진다. 한계치 값을 분할 시 분할을 일으킨 간선의 유사도 값을 갖고, 위와 같은 분할 과정을 거치면서 계속적으로 갱신되어진다. 여기서 의미 없는 클러스터링 간선은 최종 한계치 값보다 더 작은 유사도 값을 갖는 간선들을 의미한다. 이러한 과정은 사용자에게 음악의 주요한 선율로 인식되는 반복 주제 선율을 클러스터링 하도록 한다.

그러나, [9]는 음악 구성 형식을 고려하지 않고, 초기에 사이클 형성을 위해 사이클의 크기를 3이상으로 규정한다. (그림 3)(a)는 두 도막 형식의 예로서 동요 “고향의 봄”의 악보이며, (그림 3)(b)는 해당 음악의 동기간 유사도 행렬이다. (그림 3)(b)에 내용 기반 음악 검색 시스템에 색인하기 위해서 [9]을 적용시켜 유사한 선율들로 클러스터링하는 경우, 동기 클러스터 “6-2-1-3-7-4-8”로 클러스터링되며, 한계치 값은 78을 갖는다. (그림 4)에서는 동요 “고향의 봄”의 클러스터링 과정을 보인다. 공간 제약으로 그래프에서 모든 유사도 값을 표시하지는 않고, 필요한 유사도 값은 간선 위에 레이블로 표시한다.

1로 표시된 노드로부터, 클러스터 생성을 시작한다. 노드 1에 인접한 모든 간선들에서 간선 (1-2)는 노드 1과 2사이의 최대 유사도 값을 갖는다. 간선 (1-2), (2-6), (6-1)이 기존의 그래프 기반 클러스터링 알고리즘[9]에 의해 고려된다. 고려된 간선 (6-1)은 원시 사이클(1-2, 2-6, 6-1)을 형성한다. 전임 간선이 없고, 사이클의 조건이 참이기 때문에 원시 사이클은 클러스터 사이클로 간주되고, 간선 1은 사이클 클러스터(1, 2, 6)의 사이클 노드가 된다. 다음에 고려된 간선 (6-3)과 사이클 완성 간선인 (1-3)으로부터 사이클의 확장 가능성을 체크한다. 간선 (1-3)은 (1, 2, 6)으로 구성된



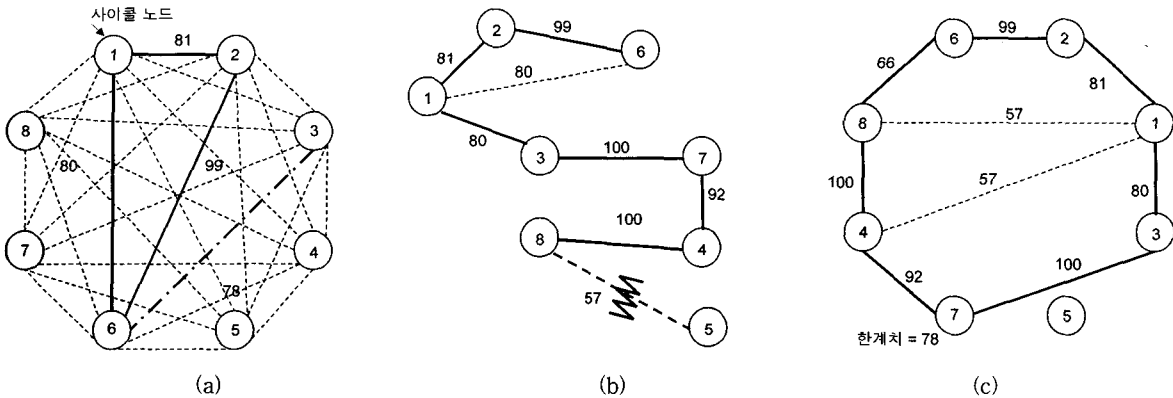
(a) 동요 “고향의 봄”의 악보

동기	1	2	3	4	5	6	7	8
1	100	81	80	57	79	80	80	57
2	81	100	78	66	69	99	78	66
3	80	78	100	92	64	78	100	92
4	57	66	92	100	57	66	92	100
5	79	69	64	57	100	70	64	57
6	80	99	78	66	70	100	78	66
7	80	78	100	92	64	78	100	92
8	57	66	92	100	57	66	92	100

(b) 동요 “고향의 봄”의 유사도 행렬  
(그림 3) 2도막 형식의 동요 “고향의 봄”

후보 클러스터의 유사도 최소값(80)보다 크거나 같기 때문에 확장될 수 있다. 사이클 노드는 6으로 변경된다. (1-3), (3-7), (7-4), (4-8)로 노드들이 유사도 값에 따라 선택된다. 간선 (8-5)의 유사도 값 57은 현재, 클러스터를 이루는 모든 간선들의 최소값(80)보다 작기 때문에 절단이 발생한다. 한계치 값은 변경되지 않는다. 그러나, (그림 4)(c)에서 보듯이 구성된 클러스터내의 간선 (1-4), (1-8)는 한계치 값보다 낮은 유사도 값을 갖음에도 불구하고, 같은 클러스터에 포함됨을 알 수 있다. 동요 “고향의 봄”은 전형적인 두도막 형식으로서 뒤큰악절 구성시 앞큰악절의 앞작은악절 일부를 대조 및 반복시켜 변화시키고, 앞큰악절의 뒤작은악절을 반복하여 통일성을 갖추었다. 즉,  $A-B-C-D-E-B'-C-D$  요소이다. 이러한 곡의 구조에서 음악 구성 형식을 고려할 경우 동요 “고향의 봄”은 클러스터 (2, 6), (3, 7), (4, 8)로 재구성되어야 한다.

본 논문에서는 이러한 문제를 해결하기 위해, [9]의 사이클 가능성을 확장하여 최소 사이클의 크기를 2이상에서 가능하도록 한다. 초기에 유사도 그래프로부터 노드  $A, B$ 가 선택 후, 노드  $C$ 가 고려되어 한 클러스터에 참여될 동기의 수가 3일 경우, 노드  $A, B$  사이의 유사도 값은 상대적으로 높지만, 노드  $C$ 와 다른 동기들 사이의 유사도 값이 상대적으로 낮은 경우 정삼각형의 성질을 이용하여 노드  $A, B$ 를 포함하는 사이클의 형성으로부터 노드  $C$ 를 차단한다.



(그림 4) 동요 "고향의 봄"의 클러스터 생성

완전히 연결된 유사도 그래프에서 세 노드로 구성된 사이클은 삼각형을 이룬다. 예를 들어, 동기 A, B, C가 존재한다고 하자. B는 A의 반복이지만, C는 A에 대한 심한 변형 또는 완전히 새로운 선율이라고 가정하고, 간선 a(=<B, C>), b(=<C, A>), c(=<A, B>)는 유사도 값으로서 60, 60, 100의 값을 갖는다고 가정하자. 세 동기 A, B, C가 이루는 사이클은 정삼각형과 유사하지 않다. 이와 달리, B는 A의 반복이고, C는 A에 대한 약간의 변형이라고 가정하고, 간선 a, b, c는 유사도 값으로서 97, 97, 100의 값을 갖는다고 가정하자. 세 동기 A, B, C가 이루는 사이클은 세 변의 길이가 같은 정삼각형과 유사하다. 즉, 세 개의 동기가 같은 클러스터에 속하기 위해서는 정삼각형을 이루는 세 변의 값처럼 간선 a, b, c의 값이 서로 유사해야 한다[19]. 따라서, 한 클러스터에 속하는 각 동기간 ε-허용 유사 가능성은 [정의 1]과 같이 정의될 수 있다.

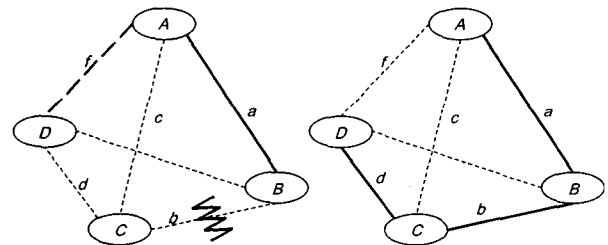
**[정의 1]** ε-허용 유사 가능성: 세 개 이상의 동기가 같은 클러스터내에 존재하기 위해서는 서로 ε-허용 유사 가능해야 하며, 세 동기내의 각 동기간 유사도 값 α = S(a), β = S(b), γ = S(c)는 식 (1)을 만족시킨다.

$$|1 - \frac{\beta + \gamma}{2\alpha}| \leq \epsilon \text{ AND } |1 - \frac{\gamma + \alpha}{2\beta}| \leq \epsilon \text{ AND } |1 - \frac{\alpha + \beta}{2\gamma}| \leq \epsilon, 0 \leq \epsilon \leq 1 \quad (1)$$

**[증명]** 삼각형 ΔABC의 각과 변 사이에는 코사인 정리가 성립된다. α<sup>2</sup> = β<sup>2</sup> + γ<sup>2</sup> - 2βγcosA, β<sup>2</sup> = γ<sup>2</sup> + α<sup>2</sup> - 2γαcosB, γ<sup>2</sup> = α<sup>2</sup> + β<sup>2</sup> - 2αβcosC이다. 세 동기로 이루어진 삼각형 ΔABC가 정삼각형에 ε-허용 유사할 경우, cosA ≅ cosB ≅ cosC ≅ 1/2(1-ε)을 갖는다. 따라서, 코사인 정리는 α<sup>2</sup> ≅ β<sup>2</sup> + γ<sup>2</sup> - (1-ε)βγ, β<sup>2</sup> ≅ γ<sup>2</sup> + α<sup>2</sup> - (1-ε)γα, γ<sup>2</sup> ≅ α<sup>2</sup> + β<sup>2</sup> - (1-ε)αβ로 재구성되며, α ≅ (1-ε) \* (β+γ)/2, β ≅ (1-ε) \* (γ+α)/2, γ ≅ (1-ε) \* (α+β)/2의 값을 얻는다. 이 값들을 식 (1)에 대입

할 경우, 식 (1)이 만족됨을 알 수 있다. 예를 들어, |1 - (α+β)/(2γ)| ≤ ε에 γ의 값 대신 (1-ε) \* (α+β)/2을 입력할 경우, 결과식 ε<sup>2</sup> ≥ 0은 항상 만족한다. 따라서, 세 동기내의 각 동기간 유사도 값 α, β, γ는 서로 ε-허용 유사할 경우 식 (1)을 만족시킨다.

다음은 [정의 1]을 바탕으로 동기간 유사도와 클러스터와의 관계를 보인다. (그림 5)처럼 간선 a가 선택되어 사이클의 크기가 2인 상태에서 간선 b가 선택될 경우, 유사도 행렬은 완전 연결된 그래프로 간주되기 때문에 인접된 노드 A, B, C는 삼각형을 이루며, 다음 두 가지 경우를 갖게 된다.



(a) ε-허용 유사 가능성을 만족하지 않는 경우 (b) ε-허용 유사 가능성을 만족하는 경우

(그림 5) ε-허용 유사 가능성에 의한 동기 분할

① ε-허용 유사 가능성을 만족하지 않는 경우 : 전임 간선이 없고, 노드 B가 사이클 노드인 후보 분할 (A, B)를 생성하고, 간선 b를 절단한다. 기존 한계치 값보다 S(b)가 크면, 한계치 값은 S(b)의 값으로 갱신된다. 후보 분할 (A, B)로부터 간선 f에 대해 확장 가능성을 체크한다. 간선 f 또한, ε-허용 유사 가능성을 만족시키지 않을 경우, 후보 클러스터(A, B)를 생성한다. 이 후, 절단된 간선에 인접된 노드 C에서 시작한다. 후보 분할 (A, B)로부터 간선 f에 대해 노드 A, B, D가 ε-허용 유사 가능성을 만족할 경우 확장하며, 이 후의 과정은 그래프 기반 클러스터링 알고리즘 [9]과 동일하다.

②  $\epsilon$ -허용 유사 가능성을 만족하는 경우 : 노드  $C$ 에 연결된 간선들 중 가장 큰 값을 갖는 간선  $d$ 를 고려하며, 이 후의 과정은 그래프 기반 클러스터링 알고리즘[9]과 동일하다.

따라서, 음악 구성 형식을 고려하기 위해 기존의 그래프 기반 클러스터링 알고리즘[9]과 달리, 최소 사이클의 크기를 2이상에서 가능하도록 하였다. 음악 구성 형식을 고려하여  $\epsilon$ -허용 유사 가능성을 만족하는 동기들을 클러스터링하기 위해 확장된 반복 주제 선율 클러스터링 알고리즘은 (알고리즘 1)과 같다.

입력:  $n \times n$  크기의 유사도 행렬 또는  $n$ 개의 노드가 완전히 연결된 유사도 행렬 그래프  
 출력: 음악 구성 형식에 따라  $\epsilon$ -허용 유사 가능성을 만족하는 후보 클러스터의 집합

step 1. 첫 동기 선율 노드에서 한계치 = 0으로 시작한다.  
 step 2. 다음의 조건을 만족시키는 간선을 고려하며, 모든 노드가 반복에 사용되면, 끝내고 step 7로 간다:  
 (1) 이전의 반복에서 절단이 발생하면, 절단된 간선에 인접된 노드에서 시작한다.  
 (2) 그렇지 않으면, 선택 가능한 간선 중에서 가장 큰 값을 고려한다.  
 step 3. 현재 사이클의 크기가 2이면서, 후보 분할이 존재하지 않으면:  
 (1) ' $\epsilon$ -허용 유사 가능성'을 만족하지 않으면, 유사도가 큰 간선에 접하는 노드들을 후보 분할로 간주하고, 사이클 노드 및 다음 고려 대상 노드를 조정한다. 이전 한계치 값보다 절단 값이 클 경우 절단 값을 한계치 값으로 갱신한다. step 2로 간다.  
 (2) ' $\epsilon$ -허용 유사 가능성'을 만족하면, step 2로 간다.  
 step 4. 현재 사이클의 크기가 2이면서, 후보 분할이 존재하면:  
 (1) ' $\epsilon$ -허용 유사 가능성'을 만족하지 않으면, 다음 고려 대상 노드를 조정 후, step 6으로 간다.  
 (2) ' $\epsilon$ -허용 유사 가능성'을 만족하면, step 2로 간다.  
 step 5. 다음에 고려된 노드가 원시 사이클을 생성하면:  
 (1) 사이클 노드가 존재하지 않으면, 사이클 가능성을 체크한다. 가능성이 존재하면, 그 사이클을 사이클 노드를 갖는 클러스터 사이클로 표시한다. 그리고, 이 사이클을 후보분할로 간주하며, step 2로 간다.  
 (2) 이미 사이클 노드가 존재하면, 이 간선을 제거하고 step 2로 간다.  
 step 6. 다음에 고려된 노드가 사이클을 형성하지 않고, 후보 분할이 존재하면:  
 (1) 전임 간선이 존재하지 않으면, 이 새로운 간선에 의해 사이클의 확장 가능성을 체크한다. 가능성이 없으면, 이 간선을 자르고, 이 사이클을 후보 클러스터로 생성한다. 이전 한계치 값보다 절단 값이 클 경우 절단 값을 한계치 값으로 갱신한다. step 2로 간다.  
 (2) 전임 간선이 존재하면, 사이클 노드를 바꾸고 전임 간선에 의한 확장 가능성을 체크한다. 가능성이 존재하지 않으면, 전임 간선을 자르고 이 사이클을 후보 클러스터로 생성한다. 이전 한계치 값보다 절단 값이 클 경우 절단 값을 한계치 값으로 갱신한다. step 2로 간다.  
 step 7. 후보 클러스터의 집합으로부터 마지막 한계치 값보다 작은 유사도 값을 갖는 의미 없는 간선들을 제거한다.

(알고리즘 1) 음악 구성 형식을 고려하기 위해 확장된 반복 주제 선율 클러스터링 알고리즘

동요 “고향의 봄”은  $\epsilon=0.1$ 일 때 (알고리즘 1)에 따라 각

각 클러스터 (2, 6), (3, 7), (4, 8)을 생성하며, 후보 클러스터 (1, 5)는 한계치 값보다 작기 때문에 제거된다. 동요 “고향의 봄”은 반복 주제 선율 클러스터링 결과로서 (2, 6), (3, 7), (4, 8) 클러스터를 갖는다.



(a) 동요 “가을”의 악보

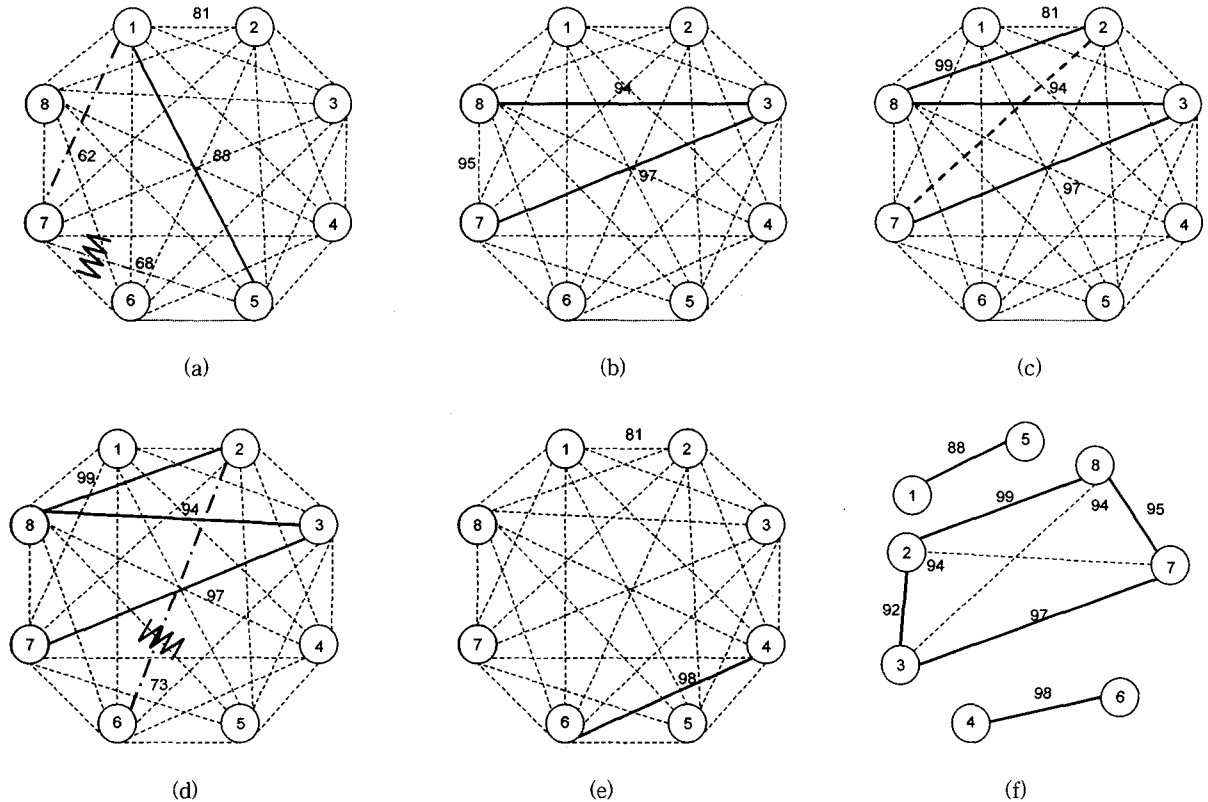
동기	1	2	3	4	5	6	7	8
1	100	55	63	84	88	84	62	61
2	55	100	92	66	50	73	94	99
3	63	92	100	78	60	81	97	94
4	84	66	78	100	64	98	74	67
5	88	50	60	64	100	63	68	56
6	84	73	81	98	63	100	78	74
7	62	94	97	74	68	78	100	95
8	61	99	94	67	56	74	95	100

(b) 동요 “가을”의 유사도 행렬

(그림 6) 2도막 형식의 동요 “가을”

이 논문에서 제안된 클러스터링 알고리즘의 정확성을 예로서 보인다. (그림 6)(a)는 동요 “가을”의 악보이며, (그림 6)(b)의 유사도 행렬에 제안된 반복 주제 선율 클러스터링 알고리즘( $\epsilon=0.1$ )을 적용한 것은 (그림 7)과 같다. (그림 6)(a)의 악보에 따르면, 동요 “가을”은 음악 구성을 위해 앞큰악절 구성시, 앞작은악절의 일부분을 뒤작은악절에서 반복시켰다. 또한, 뒤큰악절의 뒤작은악절 구성시, 앞큰악절의 앞작은악절의 일부를 변형, 반복하여 통일성을 갖게 하였기 때문에 음악 구성 형식에 따라 (1, 5), (2, 3, 7, 8), (4, 6)으로 클러스터링되는 것이 타당하다.

노드 1로부터 인접한 모든 간선들에서 간선 (1-5)는 노드 1과 5사이의 최대 유사도 값을 갖는다. 간선 (1-5), (5-7)이 (알고리즘 1)에 의해 고려된다. 고려된 간선 (5-7)은 step 3의 조건을 만족시킨다. (그림 7)(a)처럼  $S(a) > S(b)$ 이면서,  $\epsilon$ -허용 유사 가능성을 만족하지 않기 때문에 노드 5를 사이클 노드로 하는 후보 분할을 생성 후, step 2로 간다. 한계치 값은 0에서 68로 변경된다. 다음으로 간선 (1-6)이 고려된다. step 4에서 후보 클러스터가 존재하지만,  $\epsilon$ -허용



(그림 7) 동요 “가을”의 반복 주제 선을 클러스터링( $\epsilon = 0.1$ )

유사 가능성을 만족하지 않기 때문에 현재 고려된 간선 (1-6)을 무시하고, 다음에 고려되어야 할 노드를 분할 이전에 고려된 노드인 7로 조정하여 step 6으로 간다. 후보 클러스터 (1,5)를 생성 후, step 2로 간다. (그림 7)(b)처럼 노드 7로부터 (7-3), (3-8)이 (알고리즘 1)에 의해 고려된다. 고려된 간선 (3-8)은 step 3의 조건을 만족시킨다. 그러나, 고려된 세 동기 사이에  $\epsilon$ -허용 유사 가능성을 만족시키기 때문에 분할이 발생되지 않고, step 2로 간다. 다음 고려된 간선은 (8-2)이다. 노드 2에서 사이클의 확장성을 갖는 다른 노드가 없기 때문에 (그림 7)(c)처럼 클러스터 (7, 3, 8, 2) 순으로 원시 사이클을 형성한다. step 2로 간다. 다음 고려된 간선은 (2-6)이다. S(2-6)의 값 73이 사이클의 확장 가능성을 만족하지 않기 때문에 즉, 사이클의 최소 유사도 값 94보다 작기 때문에 step 6에서 (그림 7)(d)처럼 절단이 발생한다. 절단 값이 기존에 저장된 값보다 크기 때문에 한계치 값은 73으로 갱신된다. (그림 7)(e)처럼 남은 간선 (4-6)이 클러스터링을 이룬다. step 7에 따라 한계치 값보다 작은 의미 없는 클러스터를 제거한다. 마지막으로, (그림 7)(f)와 같은 클러스터링 결과를 얻는다. 동요 “가을”은 반복 주제 선을 클러스터링 결과로서 (1, 5), (2, 3, 7, 8), (4, 6)을 얻는다. 동요 “가을”의 음악 구성 형식에 따라 직접 클러스터링한 결과와 같은 결과를 (알고리즘 1)의 실행 결과로 얻을 수 있음을 알 수 있다. (그림 7)(f)의 클러스터 (2, 3, 7,

8)처럼 한 클러스터내에 존재하는 각 동기들은  $\epsilon$ -허용 유사 가능성을 만족시킨다. 즉, 음악 구성 형식을 고려한 반복 주제 선을 클러스터링의 최종 결과 클러스터를 대표하는 선들과 일정한 유사도내에서 변형된 동기 선을들을 포함한다.

본 논문에서 제안된 (알고리즘 1)의 음악 구성 형식을 고려한 반복 주제 선을 클러스터링은 사이클의 가능성을 확장하여 최소 사이클의 크기를 2이상에서 가능하도록 하여, 장단이나 음높이에 약간 변화를 주어 임의의 위치에 두 번 이상 반복된 반복 주제 선을들을 클러스터링할 수 있다. 즉, [9]와 달리 최소 사이클의 크기를 2이상에서 가능하게 함으로써, 반복 주제 선율이 2번 이상 반복되는 음악 구성 형식에 대해서도 정확하게 반복 주제 선을을 클러스터링할 수 있다. 또한, 한계치 값을 이용하여 의미 없는 클러스터를 제거하여 일정 유사도(1- $\epsilon$ )내에서 반복되는 클러스터의 집합을 얻게 된다. 이러한 클러스터링 알고리즘은 반복적인 이중 분할에 대한 부담이 없다. 또한, 특정한 경험상의 목적 함수가 필요 없다[12]. 단지, 음악 구성 형식을 고려하면서 일정 유사도내에서 반복되도록  $\epsilon$ -허용 유사 가능성을 정하는 것이 필요하지만, 이는 클러스터링 과정의 목적 함수가 아니라 사용자 임의의 부가적인 값으로서 작곡가가 음악 구성 형식으로서 사용자에게 자연스럽게 기억되도록 하기 위해 약간의 변형을 주어 반복시키는 선율에



대한 최대 변형 허용치로 간주될 수 있다. 음악 구성 형식을 고려한 반복 주제 선율 클러스터링 알고리즘의 최종 결과로 얻은 클러스터들에 대해 클러스터에 포함된 동기 선율들을 대표하는 선율을 추출해서 색인을 구성해야 한다.

### 3.3 음악 구성 형식을 고려한 대표 선율 선택

음악 구성 형식을 고려하여 일정 유사도내에서 반복되는 선율들에 대해 클러스터링한 후에 각 클러스터내에서 해당 클러스터를 대표할 수 있는 선율을 추출해야 한다. 음악 구성 형식에서 첫 동기 선율은 그 곡을 듣는 사람에게 일반적으로 해당 음악을 자연스럽게 기억하도록 해준다. 이러한 것은 사용자가 음악의 첫 동기 선율을 많이 기억하고 이를 이용하여 질의할 것이라는 점이 고려된 연구들[5, 10]에서도 알 수 있다. 음악 구성 형식에서 절정은 선율을 작곡하는데 가장 중요하게 고려해야 할 하나의 사항[13-15]으로서 일반적으로 절정은 선율을 이루고 있는 음악적 구성 요소들 중 하나 또는 그 이상의 최고 정점이다[15]. 음악의 선율적 방향은 시작에서부터 절정을 향하여 움직이며, 절정에서부터 음악의 끝을 향한다. 절정 또한 음악 구성 형식에서 그 곡을 듣는 사람에게 일반적으로 해당 음악을 자연스럽게 기억하도록 해준다. 따라서, 첫 동기과 절정을 포함하는 절정 선율은 내용 기반 음악 검색에서 해당 음악을 대표하는 대표 선율로서 색인에 포함되어야 한다.

본 논문에서는 음악 구성 형식을 고려한 반복 주제 선율 클러스터링 후, 최종 결과 클러스터로부터 대표되는 선율을 추출시 첫 동기 선율과 절정 선율에 가중치를 둔다. 즉, 첫 동기 선율과 절정 선율이 포함될 경우 우선 순위를 부여하여 해당 클러스터를 대표하는 선율로 추출한다. 만약, 해당 클러스터에 첫 동기 선율과 절정 선율이 포함되지 않은 경우, 추출된 선율이 클러스터에 포함된 모든 동기 선율을 가장 잘 대표할 수 있도록 하기 위해 클러스터 내의 모든 동기 선율을 다차원 공간에 표현할 때 중심에 가까운 선율이 해당 클러스터를 대표하는 선율로 선택되도록 한다[20].

동요 “고향의 봄”은 반복 주제 선율 클러스터링 결과로서 (2,6), (3,7), (4,8)을 갖는다. 대표 선율 클러스터 결과로부터 (3,7) 클러스터의 경우, 절정 동기인 3동기를 포함하기 때문에 3동기를 클러스터의 대표로 선출한다(3동기와 7동기는 음악적으로 동일한 선율이다). (2,6), (4,8)의 경우는 다차원 공간에 표현할 때 중심에 가까운 선율이 해당 클러스터를 대표하는 선율로 선택되도록 각각 2, 8동기를 대표로 선출한다. 마지막으로 추출된 대표 선율에 첫 동기 선율 선율이 존재하지 않기 때문에 첫 동기를 추가한다.

동요 “가을”은 (그림 7)의 클러스터 결과로부터 (1,5) 클러스터의 경우, 첫 동기 선율이 포함되었기 때문에 첫 동기 선율을 해당 클러스터의 대표로 선출한다. (2,3,7,8) 클러스터의 경우, 절정 선율인 2동기를 포함하기 때문에 2동기

를 클러스터의 대표로 선출한다. (4,6) 클러스터의 경우는 다차원 공간에 표현할 때 중심에 가까운 선율이 해당 클러스터를 대표하는 선율로 선택되도록 6동기가 추출되었다.

각 클러스터들로부터 추출된 선율을 해당 음악 데이터를 대표하는 선율로서 내용 기반 음악 정보 검색 시스템에서 색인한다. 추출된 대표 선율로 대표 선율 색인을 구성하기 위해 다차원 공간 색인 기법으로 널리 알려진 M-tree[18]를 이용한다. M-tree는  $M=(D, d)$ 로 표현되며 D는 메트릭 공간, d는 메트릭 공간내의 거리 함수이다. 대표 선율 색인을 위한 메트릭 공간은 음악의 주요 특성인 음의 높이 변화량을 축으로, 음의 길이 변화량을 축으로 사용한 2차원 공간이 되며, 거리 함수로는 삼각부등식을 만족하는 유클리드 거리 함수를 이용한다. M-tree에서의 객체 검색은 질의 객체와 데이터 객체 사이의 거리를 기준으로 검색하며, 이는 질의 객체와 근접한 k개의 객체를 검색하는 k-근접 질의(k-nearest neighbor query)와 질의 객체로부터 일정한 거리 내에 존재하는 모든 데이터 객체를 검색하는 범위 질의(range query)로 구분할 수 있다. 추출된 대표 선율들을 M-tree의 색인 기법을 이용하여 대표 선율 색인을 구성하는 자세한 내용은 [20]을 참조하기 바란다.

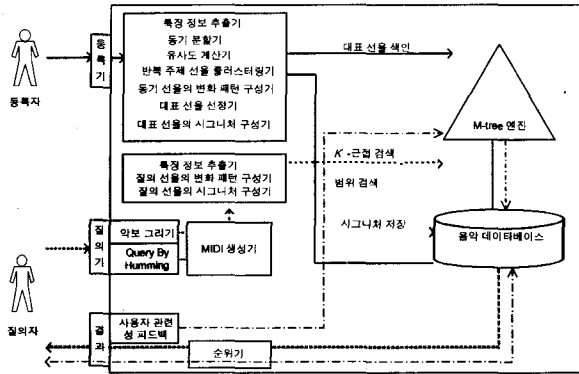
따라서, 사용자가 질의로 입력할 가능성이 높은 선율을 해당 음악 데이터의 대표 선율로 선정하여 색인을 구성하고, 색인된 대표 선율을 이용한 검색을 수행하면 전체 데이터를 검색하는 대신 색인을 이용함으로써 사용자 응답 시간을 단축시킬 수 있다.

## 4. 실험 및 결과

본 논문에서 제안한 음악 구성 형식을 고려한 대표 선율의 추출 및 색인 기법에 대해 검색의 정확성과 효율성을 증명하기 위해 내용 기반 음악 정보 검색 시스템을 구현하고, 실험하였다. 제안된 대표 선율을 이용한 내용 기반 음악 정보 검색 시스템은 Windows 2000을 운영 체제로 하는 개인용 컴퓨터(Pentium-4 1.5G, 512RAM)에서 Microsoft Visual C++ 6.0으로 구현되었다. 사용자 질의는 사용자가 원하는 데이터를 선율 정보로 입력하는 형태이고, 다차원 색인인 M-tree에 대표 선율 색인을 구성하였다. M-tree의 한 노드의 크기를 256바이트라고 가정하였고, 따라서 한 노드에 저장되는 대표 선율 객체의 최대 수(M)는 4이다. 사용자가 질의와 예상 결과 개수(k)를 입력하면 질의는 k-근접 질의를 이용하여 결과를 반환하도록 구성되었다. 대표 선율을 이용한 내용 기반 음악 정보 검색 시스템의 전체 구조를 표시하면 (그림 8)과 같다.

사용자는 내용 기반 음악 정보 검색 시스템에 새로운 음악 데이터를 등록하기 위해 등록 인터페이스를 이용해 해당 미디 화일을 업로드한다. 3장의 대표 선율 추출 및 색인

처럼 해당 midi 파일로부터 대표 선율을 추출해 M-tree에 색인하고, 대표 선율의 시그니처를 음악 데이터베이스에 삽입한다.

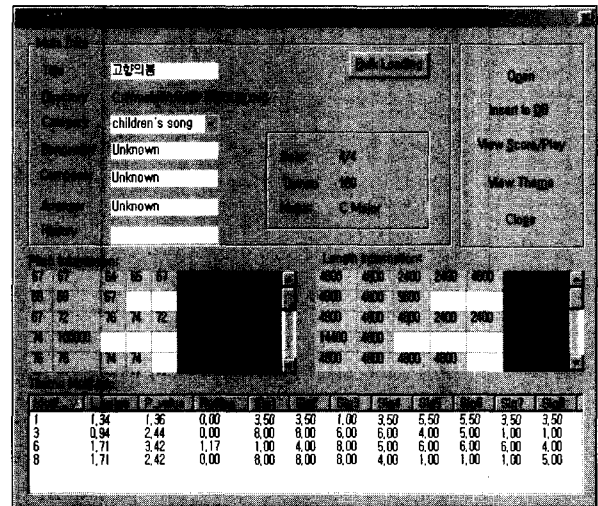


(그림 8) 대표 선율을 이용한 내용 기반 음악 정보 검색 시스템의 구조

음악 데이터를 검색하기 원하는 사용자는 내용 기반 음악 정보 검색 시스템의 질의 인터페이스에 접속한다. 사용자가 원하는 음악을 내용 기반 검색하기 위해 기억하고 있는 선율을 ‘악보 그리기’ 인터페이스를 이용해 질의 선율로 표현한다. 그리고, 검색 결과 예상 개수(k)를 입력한다. 만약, 질의 선율이 한 동기 이상일 경우, 대표 선율 색인 구성 과정과 동일하게 질의 선율로부터 대표 선율을 추출해야 한다. 질의 선율로부터 변화 패턴과 시계열 데이터 형태로 시그니처를 구성한다. 사용자가 요청한 결과 예상 개수를 만족하는 유사한 음악 데이터를 검색하기 위해 M-tree로 구현된 대표 선율 색인에서 k-근접 검색을 실시한다. k개의 대표 선율 색인들에 대해 시그니처를 이용해 유사도를 계산하여 순위를 부여하고 사용자에게 순위에 따라 결과를 제공한다. 결과 제공시 해당 대표 선율들을 포함하는 음악 화일을 사용자에게 제공한다. 사용자로부터 관련성 판단 정보를 이용해 사용자 관련성 피드백(user relevance feedback) 기능을 제공할 수 있다. 관련성 판단 정보를 이용해 해당 유사도 범위내의 음악 데이터를 찾기 위해 범위 검색을 수행하고, 해당 결과 음악 데이터를 사용자에게 제공한다. 사용자 관련성을 검사하여 사용자가 질의 결과 선율에 대해 만족하지 못할 경우, 범위 검색 과정으로 이동하여 사용자 확신도의 변화에 따른 반복적인 검색 과정을 수행하여 사용자 피드백을 지원한다.

대표 선율을 이용한 내용 기반 음악 정보 검색 시스템에서 음악 데이터의 등록 인터페이스는 (그림 9)(a), 내용 기반 질의 인터페이스는 (그림 9)(b)와 같다. (그림 9)(a)의 아랫부분에 등록을 위해 추출된 대표 색인 선율의 동기 번호(Motif), 해당 동기 선율의 음의 길이 변화량(L\_value), 음의 높이 변화량(P\_value), 그리고, 동기 선율의 변화를 요약한 시그니처(Sig1~Sig8)를 표시하고 있다. (그림 9)(b)의

경우, 악보 그리기를 통해 내용 기반 질의시 M-tree에서 질의 선율과 가까운 거리순으로 순위 부여한 결과와 시그니처를 이용한 순위 부여 결과를 함께 제공한다. 추출된 대표 선율들을 이용한 검색의 자세한 내용은 [20]을 참조하기 바란다.



(a) 음악 데이터 등록(ε=0.1)



(b) 내용 기반 질의 인터페이스

(그림 9) 대표 선율을 이용한 내용 기반 음악 정보 검색 시스템

구현된 시스템에서 대표 선율 색인을 이용한 내용 기반 검색의 정확성을 알아보기 위해 “고향의 봄”, “학교종”과 같은 동요 265곡(전체 동기의 수는 1,910개)을 가지고 색인을 구성하였다.

먼저, ε 값에 따라 음악 구성 형식을 고려한 반복 주제 선율 클러스터링 알고리즘을 분석하였다. 동요 265곡에 대해 ε 값을 0.1, 0.15, 0.2로 할 경우, 음악 구성 형식을 고려한 반복 주제 선율 클러스터링은 ε 값이 커질수록 음악 구성 형식에 따라 유사하지 않은 동기들 사이에도 클러스터

링되어 색인되는 대표 선율의 수가 작아진다. 동요 265곡에 대해  $\epsilon = 0.1$ 일 때, 색인되는 동기의 수는 660,  $\epsilon = 0.15$ 일 때, 색인되는 동기의 수는 612,  $\epsilon = 0.2$ 일 때, 색인되는 동기의 수는 587이다. 유사 허용치가 0.1와 0.15일 경우에 검색의 정확성에 있어서 큰 차이를 보이지 않았지만, 0.1, 0.2사이에는 음악 구성 형식상 선율이 2번만 반복되거나, 형식상 대조로 구성된 음악에 대해 검색을 제대로 수행하지 못한다. 본 논문에서는 사용자에게 자연스럽게 기억시키기 위해 약간의 변형을 주어 반복시키는 선율에 대한 허용 유사 가능치( $\epsilon$ )를 0.1로 한다.

동요 265곡에 대해 허용 유사 가능치를 0.1로 하고 반복 주제 선율 클러스터링을 수행하였다. 첫 동기 선율과 절정 선율이 일치하는 곡은 59개이며, 첫 동기 선율과 절정 선율이 일정 유사도 내에서 반복된 경우가 77곡이었다. 즉, 265곡 중에 51.3%가 절정 선율과 첫 동기 선율이 일치하거나, 일정 유사도내에서 반복되도록 구성되어 그 곡을 듣는 사람에게 해당 음악을 자연스럽게 기억해 줄을 알 수 있다. 첫 동기 선율(31곡) 또는 절정 선율(44곡)이 일정 유사도 내에서 반복되지 않은 경우는 48.7%이다.

동요 “고향의 봄”의 첫 동기 선율, 절정 선율, 그리고 일정 유사도내에서 반복되는 유사 반복 주제 선율에 대해 질의 인터페이스의 ‘악보 그리기’를 이용하여 질의하였다. k-근접 질의(k=3)로서 내용 기반 검색을 수행한 결과는 <표 1>과 같다.

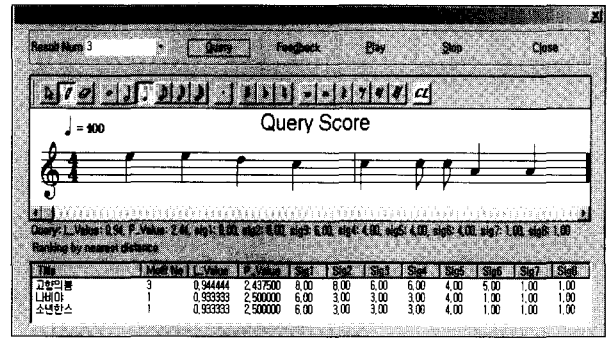
<표 1> 동요 “고향의 봄”에 대한 검색 결과

질의 선율 선율 색인	첫 동기 선율	절정 선율	반복 주제 선율
기존의 주제 선율 색인	1 	나비아, 1동기 	고향의 봄, 8동기 
	2 	소년한스, 1동기 	누가누가잠자나, 1동기 
	3 	우리유치원, 1동기 	바둑이방울, 9동기 
제안된 대표 선율 색인	1 	고향의 봄, 3동기 	고향의 봄, 8동기 
	2 	나비아, 1동기 	누가누가잠자나, 1동기 
	3 	소년한스, 1동기 	누가누가잠자나, 9동기 

<표 1>의 검색 결과 및 검색된 동기 선율의 악보에 따르면, 색인된 대표 선율은 기존의 클러스터링 알고리즘에

의해 추출된 주제 선율 색인에 비해 다양한 사용자의 질의에 대해 정확하게 검색함을 알 수 있다.

사용자의 부정확한 질의 선율에 대한 검색 결과를 알아보기 위해, 동요 “고향의 봄”의 절정 선율인 3동기의 변형 선율로 대표 선율 색인을 이용한 시스템에 내용 기반 검색한 결과는 (그림 10)과 같다.



(그림 10) 동요 “고향의 봄” 절정의 변형 선율에 대한 k-근접 질의 결과 화면

<표 2>의 악보에 따르면, 사용자의 기억하고 있는 질의 선율에 대해 정확하게 내용 기반 검색함을 알 수 있다. 따라서, 대표 선율을 이용한 색인 및 검색에서 내용 기반 검색이 정확히 수행됨을 알 수 있다.

이와 달리 기존의 주제 선율 색인의 경우 “고향의 봄” 절정의 변형 선율에 대한 질의 결과로서 “곰돌이” 8동기, “나비아” 1동기, “소년한스” 1동기를 반환한다. 따라서, 다양한 선율 색인에 대해 첫 동기 선율, 절정 선율, 그리고 일정 유사도내에서 반복되는 주제 선율에 대해 내용 기반 검색을 수행한 결과, <표 3>을 얻을 수 있다.

<표 2> “고향의 봄” 절정의 변형 선율에 대한 k-근접 검색 결과 악보

선 율	악 보
질의 선율	
고향의 봄, 3동기	
나비아, 1동기	
소년한스, 1동기	

<표 3> 질의 선율과 색인된 대표 선율 사이의 k-근접 검색 결과

	첫 동기 선율	절 정 선 율	반복 주제 선율	그 외 선율
기존의 주제 선율 색인	△	△	△	×
제안된 대표 선율 색인	○	○	○	×
전체 동기 선율 색인	○	○	○	○

첫 동기 선율에 대한 검색에서 반복 주제 선율이 첫 동기 선율과 비슷한 경우, 해당 선율에 대해 모든 색인 기법에서 검색될 수 있다. 또한, 절정 선율이 반복 주제 선율과 비슷한 경우에도 검색될 수 있다. 그러나, 서로 유사하지 않을 경우, 기존의 주제 선율 색인에서는 첫 동기 선율과 절정 선율을 질의 선율로 이용하는 경우 정확한 검색 결과를 얻을 수 없다. 기존의 주제 선율 색인에 비해 음악 구성 형식이 고려된 대표 선율 색인의 경우 형식상 대조된 음악에 대해 검색에 더 정확성을 보인다. 예를 들어, 4동기로 구성된 “학교중”이 형식 A-B-A-B'을 갖는 경우, 기존의 주제 선율 색인에서는 색인된 “학교중” 3동기에 대해서만 유사한 선율의 검색이 가능하지만, 음악 구성 형식을 고려한 색인에서는 모든 경우에 대해 정확하게 “학교중”을 검색한다. 따라서, 본 논문에서 제안된 음악 구성 형식을 고려한 대표 선율을 이용한 색인이 다른 색인 기법에 대해 전체 선율 색인과 같이 전체 음악 데이터를 색인하지 않아도 검색에 있어서 정확함을 알 수 있다. 첫 동기 선율, 절정 선율, 그리고 반복 주제 선율을 제외한 선율에 대해서는 전체 동기 선율을 색인한 경우가 검색에 있어서 정확하다. 하지만, 첫 동기 선율, 절정 선율, 그리고 반복 주제 선율을 제외한 선율들은 사용자가 질의로 사용할 가능성이 적다.

동일한 265개의 동요에 제안된 대표 선율 색인과 전체 동기 선율을 색인시, 색인 대상에 따라 선율의 변화 패턴을 이용하여 M-tree에 색인한 결과는 <표 4>와 같다.

<표 4> 색인 대상에 따른 색인 비용 비교

	색인된 동기의 수	전체 노드의 수	트리의 높이
기존의 주제 선율 색인	506	263	5
제안된 대표 선율 색인	660	360	5
전체 동기 선율 색인	1,910	994	6

<표 4>에 따르면, 반복 주제 선율만으로 색인한 경우에 비해 대표 선율 색인시 사용자의 질의 가능성이 높은 첫 동기 선율, 절정 선율등을 대표 선율로 색인하기 때문에 추가적으로 20%의 색인 공간이 필요하였다. 그러나, 음악 구성 형식을 고려한 대표 선율 색인이 기존의 주제 선율 색인 기법에 비해 색인 오버헤드가 있었지만, 대표 선율로 색인한 경우 <표 3>에 기술한 바와 같이 음악 구성 형식의 고려가 필요한 반복 주제 선율뿐만 아니라, 사용자가 기억하는 첫 동기 선율이나, 절정 선율과 같이 다양한 사용자의 질의에 대해 검색의 정확성을 증진할 수 있다. 또한, 전체 동기 선율로 색인을 구성하는 경우보다 34% 적은 양으로 음악 데이터베이스를 위한 대표 선율 색인을 구성한다. 또한, 트리의 높이가 전체 동기 선율을 색인한 경우에 비해 더 낮기 때문에 검색에 있어서 검색을 위한 평균 디스크

입출력의 횟수가 작아지기 때문에 빠른 응답 시간을 얻을 수 있음을 알 수 있다.

동요 “고향의 봄”의 첫 동기 선율에 대해서 M-tree내에서 k-근접 검색 및 범위 검색을 하기 위한 거리 계산, I/O Read 비용을 비교한 것은 <표 5>와 같다. 첫 동기 선율에 대해 결과 예상 개수는 3으로 하였으며, 범위 검색은 해당 선율로부터 반경 1.0내에 있는 모든 색인된 선율을 검색한다. 주제 선율 색인을 이용한 경우와 음악 구성 형식을 고려한 대표 선율로 색인한 경우, k-근접 검색 결과 객체의 수는 3개이며, 범위 검색 결과 객체의 수는 25개였다. 전체 동기 선율을 색인한 경우, k-근접 검색 결과 객체의 수는 3개이며, 범위 검색 결과 객체의 수는 60개였다.

<표 5> 색인 기법에 따른 검색 비용 비교

	k-근접 검색		전체 노드의 수	트리의 높이
	거리 계산	IO Read	거리 계산	IO Read
기존의 주제 선율 색인	123	57	146	71
제안된 대표 선율 색인	124	67	184	98
전체 동기 선율 색인	247	146	381	198

<표 5>에서 기존의 주제 선율 색인의 경우에 비해 제안된 대표 선율 색인이 색인 비용에 약간의 오버헤드가 있기 때문에 거리 계산 및 디스크 액세스로 인해 검색 비용도 늘어난다. 그러나, 기존의 주제 선율 색인은 다양한 사용자의 질의에 대해 검색의 정확성을 갖지 못한다. 또한, 전체 동기 선율을 색인으로 구성하는 경우보다 대표 선율 색인을 이용하는 경우 적은 거리 계산 및 디스크 액세스 횟수를 필요로 하기 때문에 효율적인 검색을 지원한다. 따라서, 대용량을 지원해야 하는 내용 기반 음악 정보 검색 시스템에서 음악 구성 형식을 고려하여 반복 주제 선율을 정확하게 추출하고, 사용자가 질의 선율로 사용할 가능성이 높은 선율을 추출하여 색인한 경우, 검색의 정확성 및 효율성을 증진시킬 수 있다.

### 5. 결론 및 향후 연구 방향

멀티미디어 자료의 증가와 활용의 증대로 멀티미디어 자료에 대한 검색의 필요성이 증가되고 있다. 그 중에서도 음악의 종류의 다양성과 방대함에 의해 음악의 검색과 색인의 필요성이 증가되고 있다. 내용 기반 음악 정보 검색 시스템에서 사용자의 응답 시간을 단축시키기 위해 음악의 첫 동기 선율, 단위의 구분 없이 일정하게 반복되는 패턴, 또는 일정한 유사도내에서 반복되는 주제 선율과 같이 음악 데이터의 일부를 색인하고 색인된 선율을 이용한 내용 기반 검색을 수행한다. 그러나, 기존의 연구들에서 내용 기반 검색을 위해 추출한 대표 선율이 음악의 대표성을 정확

하게 보장하지 못하는 단점이 있다. 또한, 음악 구성 형식을 고려하지 않기 때문에 일정한 유사도내에서 반복되는 주제 선율 추출이 충분히 않다. 그리고, 음악 구성 형식에서 사용자가 일반적으로 인지하는 첫 동기 선율과 절정 등을 고려하지 않기 때문에 사용자가 일반적으로 기억하고 있는 선율로 질의할 경우 해당 선율을 검색하지 못한다.

이를 해결하기 위해 본 논문에서는 일정한 유사도내에서 반복되는 동기 선율을 클러스터링하는 그래프 기반 클러스터링 알고리즘에 대해 음악 구성 형식을 고려하도록 확장한다. 또한, 클러스터된 선율 중에서 첫 동기 선율과 절정을 고려한 선율을 해당 음악 데이터를 대표하는 선율로 추출한다. 다차원 색인 중에서 성능이 우수한 M-tree를 이용하여 대표 선율 색인을 구성하고, 해당 사용자 질의에 대해 내용 기반 검색이 가능하도록 하였다. 실험 결과, 음악 구성 형식을 고려한 대표 선율 색인이 기존의 주제 선율 색인 기법에 비해 색인 오버헤드가 있었지만, 사용자가 인지하는 첫 동기 선율, 절정 선율 또는 일정한 유사도내에서 반복되는 주제 선율에 대해 정확한 검색 결과를 제공하며, 전체 음악 데이터의 색인보다 색인 비용 및 사용자 응답 시간이 빠름을 알 수 있다.

향후 연구로는 동기보다 작거나 큰 값이 들어올 경우, 즉 사용자의 입력 질의가 색인된 동기 단위와 서로 다른 경우에도 유사한 음악을 검색할 수 있는 서브 시퀀스 질의가 지원되도록 확장이 필요하다. 실험에 사용된 데이터는 동요가 많이 이용되었다. 동요의 경우 한 음악을 결정짓기 위해 음길이와 음높이를 이용하면 유리하지만, 창작 음악의 경우, 음성(tone)이 음악을 결정짓는 특성을 갖는다[21]. 따라서, 다양한 장르의 음악들이 갖는 특성을 고려하여 색인 및 검색이 효율적으로 이루어지도록 하는 내용 기반 음악 정보 검색 시스템의 확장이 필요하다.

### 참 고 문 헌

[1] P. Y. Rolland, "Music information retrieval : A Brief Overview of Current and Forthcoming Research," Workshop Engineering and Music, 2002.  
 [2] L. A. Smith, R. J. McNab and I. H. Witten, "Music Information Retrieval Using Audio Input," Proc AAAI Spring Symposium on Intelligent Integration and Use of Text, Image, Video, and Audio Corpora, Stanford, CA., 1997.  
 [3] E. Selfridge-Field, "What Motivates a Musical Query?," Proceedings of ISMIR 2000 1st Annual International Symposium on Music Information Retrieval, 2000.  
 [4] P. Y. Rolland, G. Raskinis and J. G. Ganascia, "Musical Content-based Retrieval : An Overview of the Melodiscov Approach and System," Proceedings of the 7th ACM international conference on Multimedia, pp.81-84, 1999.

[5] K. Andreas, "Themefinder : A Web-based Melodic Search Tool," Melodic Similarity : Concepts, Procedures and Applications, Computing in Musicology, 1998.  
 [6] C. Liu, J. Hsu and A. L. P. Chen, "Efficient Theme and Non-trivial Repeating Pattern Discovering in Music Databases," Proceedings of the 15th International Conference on Research Issues in Data Engineering, 1999.  
 [7] A. Pienimaki, "Indexing Music Databases Using Automatic Extraction of Frequent Phrases," Proceedings of ISMIR 2002 3th Annual International Symposium on Music Information Retrieval, 2002.  
 [8] K. Lemstrom, G. A. Wiggins and D. Meredith, "A Three-Layer Approach for Music Retrieval in Large Databases," Proceedings of ISMIR 2001 2nd Annual International Symposium on Music Information Retrieval, 2001.  
 [9] Y. Kang, K. Ku and Y. Kim, "Extracting Theme Melodies by Using a Graphical Clustering Algorithm for Content-based Music Information Retrieval," LNCS 2151, Springer-Verlag, 2001.  
 [10] 지정규, 오해석, "선율을 이용한 음악 정보 검색 시스템의 설계 및 구현," 정보처리논문지, 제5권 제1호, 1998.  
 [11] M. Melucci and N. Orio, "A Comparison of Manual and Automatic Melody Segmentation," Proceedings of ISMIR 2002 3th Annual International Symposium on Music Information Retrieval, 2002.  
 [12] S. B. Navathe and M. Ra, "Vertical Partitioning for Database Design : A Graphical Algorithm," Proceedings of the ACM SIGMOD, 1989.  
 [13] 이병욱, 백기풍, '3번만 읽으면 누구나 작곡할 수 있다', 도서출판 작은 우리, 1989.  
 [14] 이성천, '음악통론과 그 실습', 음악예술사, 1994.  
 [15] 김순란, '현대 음악 작곡법', 세광음악출판사, 1994.  
 [16] 김정양, 이종록, 한광희, '누구나 작곡할 수 있는 작곡법', 엘멘출판사, 2000.  
 [17] 임상혁, 구경이, 김유성, "Cosine Measure를 이용한 음악 동기 유사도 계산", 한국정보처리학회 춘계학술대회, 2003.  
 [18] P. Ciaccia, M. Patella and P. Zezula, "M-tree : An Efficient Access Method for Similarity Search in metric Spaces," 23rd VLDB Conference, 1997.  
 [19] 김태성, '수학독본', 한길사, 1994.  
 [20] C. Shin, K. Ku and Y. Kim, "Automatic Construction of Theme Melody Index from Music Database for Fast Content-based Retrievals," Proceedings of 25th European Conference on Information Retrieval Research, LNCS 2633, Springer-Verlag, 2003.  
 [21] E. Selfridge-Field, "What Motivates a Musical Query?," Proceedings of ISMIR 2000 1st Annual International Symposium on Music Information Retrieval, 2000.



**구 경 이**

e-mail : g2001419@inhavision.inha.ac.kr  
1997년 인하대학교 전자계산공학과(공학사)  
1999년 인하대학교 전자계산공학과 석사  
(공학석사)  
2000년~현재 인하대학교 전자계산공학과  
박사과정

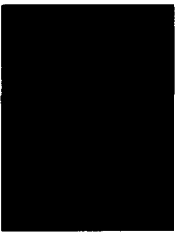
관심분야 : 멀티미디어 데이터베이스, 음악 데이터베이스, 이동  
데이터베이스



**이 재 현**

e-mail : c2021022@inhavision.inha.ac.kr  
2002년 인하대학교 전자전기컴퓨터공학부  
(공학사)  
2004년 인하대학교 정보통신대학원 석사  
과정(공학석사)  
2004년~현재 (주)인포클루 연구원

관심분야 : 멀티미디어 데이터베이스, 정보 검색, 다차원 색인  
기법



**임 상 혁**

e-mail : c2021024@inhavision.inha.ac.kr  
2002년 인하대학교 전자전기컴퓨터공학부  
(공학사)  
2004년 인하대학교 정보통신대학원 석사  
과정(공학석사)  
2004년~현재 코난테크놀로지 연구원

관심분야 : 멀티미디어 데이터베이스, 정보 검색, 다차원 색인  
기법



**김 유 성**

e-mail : yskim@inha.ac.kr  
1986년 인하대학교 전자계산학과(이학사)  
1988년 한국과학기술원 전산학과(공학석사)  
1992년 한국과학기술원 전산학과(공학박사)  
1996년~1997년 미국, 퍼듀대학교 전산학과  
방문연구원

1992년~현재 인하대학교 정보통신공학과 부교수  
관심분야 : 멀티미디어 정보검색, 이동 데이터베이스, 트랜잭션  
관리