음렬 탐색을 위한 주제소절 자동분류에 관한 연구

A Study on the Musical Theme Clustering for Searching Note Sequences

심지영(Ji-Young Shim)*, 김태수(Tae-Soo Kim)**

초 록

본 연구는 음악의 내용에 해당하는 음렬 패턴을 대상으로 분류자질을 선정하고 이를 기준으로 음렬간 유사도를 측정한 후 음렬간 군집을 형성하였다. 이는 내용기반음악검색 시스템에서 유사한 음렬을 검색 결과로 제시함으로써 이용자 탐색을 용이하게 하기 위함이다. 실험문헌집단으로는 「*A Dictionary of Musical Themes*」에 수록된 주제소절의 **kern 형식 파일을 사용하였으며, 음렬 처리 도구로는 Humdrum Toolkit version 1.0을 사용하였다. 음렬의 분절 여부와 시작 위치에 따른 네 가지 형태의 유사도 행렬을 대상으로 계층적 클러스터링 기법을 사용하여 유사한 음렬간 군집을 형성하였다. 이들 결과에 대한 평가는 외적 기준이 되는 수작업 분류표가 있는 경우 WACS 척도를 사용하였고, 음렬 내 임의의 위치에서부터 시작한 음렬을 대상으로 한 경우, 클러스터링 결과로부터 얻어낸 군집 내 공통 자질 패턴 분포를 통해 내적 기준을 마련하여 평가하였다. 평가 결과에 의하면 음렬의 시작 위치와 무관하게 분절한 자질을 사용하여 클러스터링한 결과가 그렇지 않은 것에 비해 뚜렷한 차이를 보이며 높게 나타났다.

ABSTRACT

In this paper, classification feature is selected with focus of musical content, note sequences pattern, and measures similarity between note sequences followed by constructing clusters by similar note sequences, which is easier for users to search by showing the similar note sequences with the search result in the CBMR system. Experimental document was 「*A Dictionary of Musical Themes*」, the index of theme bar focused on classical music and obtained kern-type file. Humdrum Toolkit version 1.0 was used as note sequences treat tool. The hierarchical clustering method is by stages focused on four-type similarity matrices by whether the note sequences segmentation or not and where the starting point is. For the measurement of the result, WACS standard is used in the case of being manual classification and in the case of the note sequences starting from any point in the note sequences, there is used common feature pattern distribution in the cluster obtained from the clustering result. According to the result, clustering with segmented feature unconnected with the starting point is higher with distinct difference compared with clustering with non-segmented feature.

키워드 : 음렬, 자동분류, 내용기반음악검색, 자질, 분절, Humdrum note sequences, clustering, content-based music retrieval, feature, segmentation, Humdrum

* 연세대학교 문헌정보학과 대학원(avon@chollian.net)

** 연세대학교 문헌정보학과 교수(btree@yonsei.ac.kr)

■ 논문 접수일 : 2002. 7. 30

■ 게재 확정일 : 2002. 9. 5

1 서 론

음악에서 악곡을 구성하는 음렬(note sequences) 패턴은 해당 곡을 접해본 사람들이 이후 그 곡을 기억해내고 찾을 수 있도록 하는 단서를 제공해준다. 즉 음렬은 음악 내용으로의 접근점이 되는 것이다(Heinonen 2000).

현재 디지털 음악도서관의 일환으로 활발하게 논의되고 있는 내용기반음악검색(Content-Based Music Retrieval)의 원리가 바로 악곡의 음렬 패턴 일부를 질의로 사용하여 검색하는 기법이다. 특히 Meldex 시스템의 경우, 노래 형태의 질의를 수용하여 이용자 편의를 도모하고 있다(Rodger et al. 1996, 11).

서양음악에서 사용하는 반음계는 1옥타브를 12개의 반음으로 나눈 형태를 띠고 있다. 이 12개의 음을 미국규격협회(ANSI) 표기법에 따라 '가온 다(middle C)' 기준으로 표기하면, 'C4C#4D4D#4E4F4F#4G4G#4A4A#4B4'와 같은 연속적인 음렬을 이룬다. 각각의 음들은 조성 및 장·단조, 각 음에 부여된 옥타브와 리듬에 따라 각기 다른 정보를 가진 음렬 패턴을 구성하며 곡 전체를 통해 적절히 배치되는데, 이것이 곧 음악의 내용이다.

기존의 검색 시스템을 통해서도 음악의 내용에 접근할 수 있었으나, 음렬 패턴에 부여된 각기 다른 음악적 정보를 복합적으로 고려하기보다는 단순히 음고(pitch) 정보를 지닌 12개 음을 대상으로 한 연구가 대부분이었다. 한편 음악의 내용에 접근하는데 있어서도 주로 곡 전체를 단위로 하고 있어, 결과적으로 짧은 음렬을 지닌 민속음악 등이 대상이 되었던 것도 사실이다.

본 연구에서는 이용자가 기억하고 단서로 삼는 음렬이 주제소절이라는 가정하에 이들 음렬과 결부된 조성, 옥타브 등과 같은 음악정보를 고려하여 적절한 분류자질을 선정하고, 이를 기준으로 유사한 음렬끼리 모아줌으로써 궁극적으로는 이용자가 원하는 소절에 가까운 검색 결과를 제시해주는 데 그 목적이 있다.

자동분류의 대상으로 삼은 주제소절은 클래식 기악곡 주제 색인의 일종인 *A Dictionary of Musical Themes*에 근거하여 선정하였으며, 이들 음렬에 대해 음정, 조성, 옥타브, 리듬 정보 등을 포함한 **kern 형식의 파일을 얻기 위해 이를 수록하고 있는 Themefinder 검색 시스템을 사용하였다. 음렬 데이터 처리를 위해 자동화된 음악 분석도구인 Humdrum Toolkit version 1.0을 사용하였으며, 쉘 스크립트로 짜여진 Humdrum 사용을 위한 보조수단으로 MKS Toolkit for Developer version 8.0을 사용하였다.

한편 본 실험에서 자동분류의 대상인 음렬은 의미 검증이 어렵다는 특성을 지닌다. 이러한 한계를 극복하기 위한 방편으로, 악곡 전체의 의미를 함축한 주제소절로 한정하였으나 음렬의 분절 단위 구분에 대한 보다 명확한 검증 기준이 필요하다고 본다.

2 음악 데이터

2.1 음악 텍스트의 선정

음악을 정보처리의 대상으로 보는데 있어서 Alexander(1996)는 음악이 지니는 텍스트와 사운드라는 이원적 속성에 주목했다. 정보검색에서 음악 텍스트를 내용으로 한다는 것은 악보를 대상으로 하는 것이다. 이는 CMN(Common-practice Music Notation) 기보법에 따라 오선(staff)상에 배열된 음렬을 대상으로 기계가 읽어들일 수 있는 일련의 규칙을 부여하여 처리한다. 한편 사운드를 내용으로 할 경우 녹음자료를 대상으로 연속적인 소리 신호의 주파수가 형성해내는 파형에서 음고를 추출해내는 방식을 취한다(Nicholas and Ahmed 1997, 102-103).

현재 음악 데이터를 대상으로 하는 연구는 이 두 가지 속성을 대상으로 하는 것으로 양분되어 있다. 물리적 음고를 주파수로 지각하는 방법은 음렬 패턴의 인식 및 처리 방법 면에서 텍스트에 비해 용이하다. 음렬을 일정한 체계에 맞춰 코딩할 필요없이 소리 신호를 즉각 받아들여 시간영역에서의 신호를 주파수-진폭 표현으로 변환하여 해당 음을 분리해 낼 수 있다. 또한 텍스트 처리의 경우에서처럼 적절한 음렬의 길이를 찾기 위해 고심할 여지없이 일정한 시간 단위로의 분석이 가능하다.

그러나 물리적 주파수로 접근하는 방법의 경우 각 음에 고정된 주파수로 인해 해당 음을 다르게 인식하는 청자의 지각이 반영될 수 없다. 즉 절대음감이 없는 청자의 경우, 정확한 음 인식이 어렵다는 한계를 반영하기 어렵다. 또한 화음과 같이 여러 음이 동시에 결합되는 경우 개개의 음을 정확하게 분리하기 어렵다는 문제를 지니고 있다. 무엇보다 음 고유의 주파수 신호를 분석 대상으로 하기 때문에 소리신호로 전달되는 음고와 음의 세기 이외의 음악적 자질이 개입될 여지가 없어 음악적 특성을 고려한 분류를 하기에 미흡하다. Chin-Chin, Jia-Lien, Arbee (1999)는 음량(loudness), 음고(pitch), 대역폭(bandwidth), 고조파(harmonicity)와 같은 소리 자질(acoustic feature)이 음악적으로 해석되기에는 지나치게 낮은 의미 속성을 지니고 있다고 지적한 바 있다.

이에 비해 음악 텍스트는 사운드를 통한 접근의 경우보다 코딩이 복잡하다는 단점이 있으나, 음악이 내포하는 다양한 형태의 자질을 표현하는데 융통성이 있다. 악보의 중심 내용이라 할 수 있는 선율(melody)을 표현하는데 대강의 음의 흐름을 윤곽으로 나타내는 형상(contour) 표현이 있으며, 옥타브를 감안한 음정(interval)의 사용도 가능하다. 또한 마디와 쉼표, 조표, 셈여림표, 꾸밈음 등과 같이 선율의 내용과 밀접한 관련을 맺고 있는 자질들이 코딩 단계에서 고려될 수 있다. 따라서 본 연구에서는 다양한 음악 데이터의 특성을 분류자질로 사용하는데

용이하다고 판단되는 음악 텍스트를 분석 대상으로 삼았다.

2.2 음렬(note sequences)의 구성

음악 텍스트인 악보를 이루는 핵심 내용은 오선(staff) 상에 배열된 음(note)의 연속체로 구성된다. 이는 다양한 음고와 길이를 가진 음들이 진행되면서 연속적으로 인식되는 것으로 본래 음악용어로 선율에 해당된다.

정보검색의 대상으로 음악 내용을 다루고 있는 연구에서는 선율(melody)이라는 표현과 더불어 음표(note)의 다른 표현에 해당하는 톤(tone)의 연속체인 튠(tune)이라는 표현을 사용하기도 한다. 튠은 결국 선율에 해당하는 말이지만, 이용자의 허밍 질의라는 측면을 부각시킬 때 이와 같은 구어체적인 표현을 사용하고 있다(서울대학교 서양음악연구소 2001). 한편 음악의 내용을 색인한다는 측면에서 텍스트의 문자열과 같은 맥락에서 음렬(note sequences)이라는 표현을 사용한다(Chul Yong et al. 1999, 125; Alexandra and Justin 1998, 235; Yuen-Hsien 1999, 176).

결국 선율(melody)과 튠(tune), 음렬(note sequences)은 같은 의미를 띠고 있지만, 그것을 어느 관점에서 보는가에 따라 달리 표현된 용어들이다. 본 연구에서는 음악용어로 선율에 해당되는 내용을 컴퓨터 처리 대상으로 보고 음(note)들이 정해진 규칙에 따라 배열되어 있다'는 의미를 내포하는 음렬로 정의하고자 한다.

음렬은 음표(note)를 기본 요소로 한다. 음표는 음악에서 음높이와 길이를 지닌 하나의 소리를 의미한다. 그런데 이 음표는 그 자체로서 고유한 의미를 지니기보다는 음자리표(clef)와 조표(key signature)와의 관련성에 따라 각기 달리 해석되는 변화되는 정보이다. 따라서 음표로 구성된 음렬 해석에 있어 음자리표와 조표 정보를 반영하는 것이 우선되어야만 한다.

특히 조표에 의해 표현되는 조성(tonality)은 12개의 음에 상대적 중요성을 부여하는 기능을 통해 조성 위계(tonal hierarchy)를 성립시키는 특징을 지닌다(김승길 1986, 117; 전지호 1988, 100-102; Krumhansl 1990, 33-35). 고정도법(fixed-doh)과 이동도법(movable do)에 따라 달라지는 음렬 해석 또한 조성과 관련되어 있다. 고정도법은 조성과는 무관하게 C음을 항상 '도'에 고정시켜 읽는 방법으로 절대음 높이의 음이름 그대로 읽는다. 이에 반해 이동도법은 모든 장조의 으뜸음(tonic)을 '도'로 하여 음표를 읽어들이는 방법으로 각 조에 따라 '도'의 위치를 이동시킨 상대음 개념이 적용된다. 이 때, 단조의 경우에는 으뜸음을 '라'로 한다.

한편 전체 곡의 박자를 알려주는 박자표(time signature)가 기본박과 기본박이 한 마디에 들어가는 개수를 지시하지만, 음렬을 구성하는 개별 음표에 이미 4분음표, 8분음표 등과 같이 고유의 길이가 부

여된 상태이므로 음렬의 해석에서 크게 염두에 두지 않았다.

이외에도 쉼표, 꾸밈음, 붙임줄, 이음줄, 셈여림표 등과 같은 부수적인 속성이 음렬의 해석에 개입될 수 있다. 꾸밈음의 경우, 부가적인 성격의 선율이라는 점에서 논외로 하고, 붙임줄이나 쉼표의 경우에는 음렬의 해석에 직접적인 영향을 미칠 수 있는 요소로 간주된다. 특히 같은 높이의 두 음이 연결되어 있는 붙임줄의 경우, 첫 번째 음에서만 소리를 내고 두 음의 길이만큼 소리를 유지한다는 점을 음렬의 처리에서 반영해야만 한다.

2.3 음렬의 분절

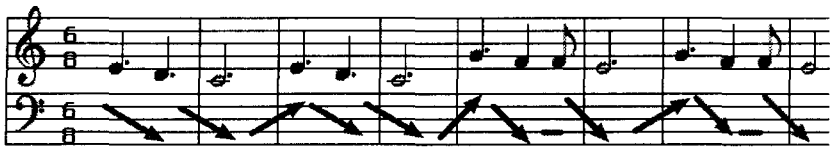
언어에서 특정 단어가 문서를 표현하는 자질이 될 수 있는 이유는 그것이 의미 차원의 검증이 가능하기 때문이다. 한 문장을 구성하고 있는 단어들은 의미에 의해 개별화되기 때문에 동음이의어를 제외하고는 다른 문장에서도 동일한 의미를 지닌 독립적 요소로 작용할 수 있다.

이러한 맥락에서 Stephen(1998)은 n-gram 단위로 분할된 음렬이 하나의 단어와 같이 기능한다는 'musical words' 개념을 제시한 바 있다. 그러나 음렬에서 실제 단어와 같은 개별적 의미를 가진 자질을 분리하기란 쉽지 않다.

백낙호와 서우석(1986)은 인도나 한국의 민속음악의 경우, 단일 음 자체가 의미를 내포하는 자질을 가지며 기호단위로서 성립한다고 했으나 서양음악의 경우, 결합된 음들의 연속체인 음렬 내에서만 의미를 찾을 수 있다는 것이 일반적인 시각이다(Diana 1980, 197).

문제는 구문 구조가 없는 음렬에서 의미를 가진 단위로 분절할 때 그 분절을 정당화해줄 수 있는 명료한 검증 기준이 없다는 것이다. 이에 대해 음악기호학자들은 분포주의적 음악분석을 통하여, 음악에서 구문 구조를 체계적으로 밝혀내려는 시도를 하고 있다.

분포주의적 분석에서는 음렬의 단위를 구분짓는 문맥 집합을 동시출현(co-occurrence)에 근거하여 밝혀내고 있다. 이를 통해 음렬을 구성하는 요소들 사이의 양립성과 비양립성의 관계를 명료히 함으로써 분포적 유(distributional class)를 형성해 낸다. 이러한 유형학적(taxonomic)인 접근을 통해 문장의 분절이 가능하며, 그 결과로부터 성분을 분석할 수 있는 것이다(백낙호, 서우석 1986, 44).

Ruwet(1966)는 음렬에서 유사한 단위들의 범위를 정하기 위한 분절의 단위로 '반복원리'를 적용하였다. 촘스키류의 생성문법 이론에서도 음악의 기본 단위를 정의하는데 반복을 주요한 기준으로 삼고 있다. 음악에서 반복이란 가장 기본적이고 보편적인 진행방식이라고 할 수 있다. '모티브의 반복'이나 '악구의 반복', 또는 음악 형식의 개념을 심어 주는 '큰 부분의 반복'과 같이 외형적으로 쉽게 드러나는 반복의 종류 이외에도 '음정 패턴'이나

*DDUDDUDRDUDRD

〈그림 1〉 선율 형상(contour)을 통한 악보 표현

'리듬 패턴'이 음악에 구조적 통일성을 주기 위하여 숨어 있을 수 있다(이석원 1994, 206-208).

실제로 민속음악 데이터베이스인 Essen Folksong collection의 경우 악곡 전체를 대상으로 색인하고 있지만, 점차 반복되는 음렬 패턴이나 주제를 대상으로 한 연구들이 등장하고 있다. 대용량 데이터베이스를 대상으로 악곡에서 반복되는 음렬 패턴을 사용하여 색인할 것을 제안한 연구의 경우, 음악의 계층적 규칙에서 하나의 악절(sentence)을 이루는 음렬 내 하위 음렬의 출현빈도를 측정하여, 반복되는 패턴을 찾아내는 방안이 제안된 바 있다(Jia-Lien, Chin-Chin, and Arbee 1998, 281-282).

Takashi Yanase(1999)는 악곡을 대표하는 음렬을 색인할 때 주제 악구(theme phrase)를 사용하였으며, Themefinder라는 시스템의 명칭 그대로 주제소절을 대상으로 색인한 내용기반음악검색 시스템이 웹 상에서 서비스되고 있기도 하다.

3 자동분류

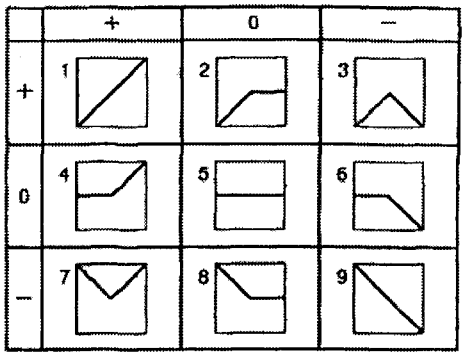
3.1 자질 선정

음렬을 대상으로 한 자질 선정 과정에서는 표현 수준과 구분 단위를 결부시켜 고려할 수 있다(Pickens 2001). 표현 수준은 다시 두 가지로 나누어 생각할 수 있다. 먼저 음렬 패턴이 내포하는 정보량의 수준에 따라 형상(contour), 음고(pitch), 음정(interval), 음도(scale degree) 등으로 표현될 수 있다. 이들 각각은 차례로 형상이 가장 적은 정보량을, 음도가 가장 많은 정보량을 지니고 있다. 다음으로 음렬의 상대적인 높낮이를 해석하는 방법에 따라 이동도법과 고정도법으로 구분된다.

특히 선율 형상을 사용한 음렬 표현은 Parsons(1975)가 처음 음렬 색인에 사용한 이래 그 단순성과 비용 효과적인 측면에서 많은 연구자들에 의해 선호되어 왔다. 이는 첫 음표에 * 표시를 하고, 하강하는 음에 D(down), 상승하는 음에 U(up), 반복되는 음에 R(repeat)을 각각 표시하여 〈그림 1〉과 같이 악보의 선율 정보를 '음고 프로파일(pitch profile)' 형태로 표현

하는 방법이다(Rodger et al. 1996, 14).

형상은 음렬을 대상으로 이웃하는 음이 높고, 낮고, 동일하다는 3분법적인 구분을 적용한다. 따라서 색인 과정이 단순하면서도 유사한 형태 정보를 통해 음렬에 접근함으로써 이용자 질의를 융통성있게 수용할 수 있다는 장점을 지닌다. 반면 정보 검색시 지나치게 재현율을 높인다는 한계를 드러낸다(Youngmoo et al. 2000).



〈그림 2〉 3-gram 형상 음렬 패턴

〈그림 2〉와 같이 형상만을 표현 수준으로 하여 3-gram 음렬 패턴을 자질로 선정한 연구(Tsuji et al. 1997)도 있었으나, 형상 정보의 불충분한 식별력을 감안하여, 음고와 음정을 함께 사용함으로써 각각이 지닌 단점을 상호 보완하는 형태의 연구가 보다 일반적이다(Rodger et al. 1996, 16-17; Alexandra and Justin 1998, 235; Cambouropoulos 1998, 3-4).

형상을 사용하는 경우와 대조적으로 음도와 리듬 정보를 함께 고려해 많은 정보량을 포함한 자질을 사용한 연구도 있었으나, 이는 분석 대상을 특정 곡의 제한된 소절로 한정하였다는 점에서 약점이라고 할 수 있다(Marcel and David 1990). 양면성을 지닌 표현 수준은 정보검색에서 정확률과 재현율에 직접적인 영향을 미친다는 점에서 면밀하게 고려해야 할 문제이다.

〈그림 3〉 음정 비를 적용한 상대음정에서 동일하게 간주되는 마디의 예

Kjell, Pauli, Sami(1999)는 음정의 비가 같은 경우를 동일한 음정으로 간주하는 상대음정(RIS: relative interval slope)을 분류자질로 사용하였다. 이 경우에는 〈그림 3〉과 같이 1:2:4의 음정 비를 가진 첫째 마디와 2:4:8의 음정 비를 가진 둘째 마디가 동일한 상대 음정으로 처리된다.

Ta-Chun, Arbee, Chin-Chin(1996)은 〈그림 4〉와 같이 자주 사용되는 화음으로 구성된 '화음 집합'(chord set)을 색인으로 사용하였다.

C1: do C2: do, mi C : do, mi, sol C7 : do, mi, sol, si
D1: re D2: re, fa Dm: re, fa, la Dm7: re, fa, la, do
E1: mi E2: mi, sol Em: mi, sol, si Em7: mi, sol, si, re
F1: fa F2: fa, la F : fa, la, do F7 : fa, la, do, mi
G1: sol G2: sol, si G : sol, si, re G7 : sol, si, re, fa
A1: la A2: la, do Am: la, do, mi Am7: la, do, mi, sol

〈그림 4〉 화음 집합

리듬 정보가 음렬에 반영되는 경우 〈그림 5〉와 같이 선율 면에서는 잡음을 가지고 있더라도 유사한 음렬로 간주될 수 있

다(Kjell, Pauli, and Sami 1999, 320). 그러나 David 등(1999)은 음렬 색인에서 리듬 정보를 고려할 경우 재현율이 지나치게 낮아진다는 이유로 리듬을 배제하기도 하였다.

〈그림 5〉 유사 음렬

한편 구분 단위를 고려할 경우 마디나 악구를 단위로 하여 음렬을 적절한 단위로 나누는 방법을 사용하기도 하였으나, 통상적으로 구분 단위가 명확한 n-gram 방식이 선호되어 왔다. 이 때 가장 핵심적인 과제는 가장 적절한 n의 수치를 찾아내는 것이다. 색인의 크기를 줄일 수 있다는 측면에서 가능한 짧은 단위의 길이를 사용하는 것이 바람직하지만, 이러한 조처는 검색시 정확률을 높이는데 장해 요인이 될 수 있다.

Yuen-Hsien(1999)은 음고를 표현 수준으로 하여 2-gram과 3-gram으로 각각 색인하여 검색 효율성을 판정하는 실험을 통하여 3-gram에서 보다 좋은 성능을 보였다. 유사한 형태를 띤 Alexandra와 Justin(1999)의 연구에서는 4-gram 단위가 본래 음렬을 그대로 색인한 것 다음으로 검색에서 높은 정확률을 보여주었다. 한편 4-gram, 5-gram, 6-gram을 각각 사용한 Stephen(1999)의 연구에서는 5-gram 단위에서 가장 좋은 성능을 나타내었다. 그러나 이러한 결과만으로 특정 n의 수치로 구분 단위를 단정짓기에는 무리가 있다. 대상 문헌집단의 크기와 같은 여타의 환경변수에 대한 제어가 이루어지지 않았을 뿐만 아니라 성능을 평가하는 척도 또한 각 연구마다 달랐기 때문이다.

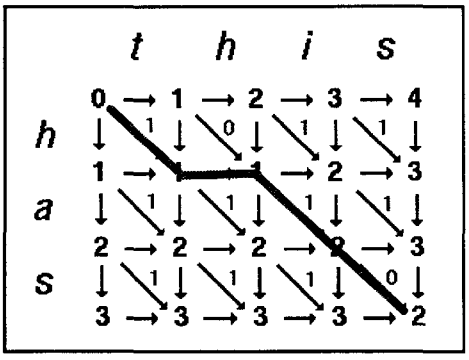
이상의 선행 연구들의 내용을 통하여 음렬을 대상으로 한 자질 선정시 고려해야 할 점은 첫째, 음렬을 이루는 패턴을 어느 요소로 구성할 것인가와 둘째, 음렬의 길이를 어떤 방식으로 구분할 것인가로 요약될 수 있다. 자동분류 실험에서는 이와 같은 두 가지 측면에 초점을 맞추는 동시에, 기존의 표현 수준에서 간과되었던 조성을 반영하기 위해 자질 선정 과정에서 이동도법을 반영하였다.

한편 선행연구에서 언급된 자질들은 엄밀히 말하자면 자동분류의 대상이라기보다는 색인 대상으로서의 자질인 경향이 강하다. 그러나 분류와 관련한 연구(Steven and David 2000)에서도 이에 관한 별다른 구분 없이 자질을 선정하고 있었으므로 색인과 분류 자질을 동일한 맥락에서 살펴보았다.

3.2 유사성 척도

음렬의 유사성을 측정하는 단위로 일반적으로 거리계수를 사용한다. 그 가운데

서도 두 음렬간 최소 변형의 수를 산정하는 edit distance의 원리가 반영되고 있다(Keith and David 1992, 5-8; David et al. 1999, 165; Francu and Nevill-Manning 2000, 890-891; Karin, Dominik, and Christina 2001, 14). 음렬 처리 도구인 Humdrum Toolkit에서 사용하는 simil 프로그램에서도 edit distance의 원리에 따른 Damerau-Levenshtein 거리계수를 사용하고 있다(David 1999).



〈그림 6〉 edit distance의 원리

Damerau-Levenshtein 거리계수는 edit distance의 원리에 따라 비교 대상 문자열이 일치하는 경우는 0값을, 다른 문자가 삽입, 삭제, 치환되었을 경우에 각각 1값의 패널티(penalty)를 주는 방법을 사용하여, 최소 패널티 값을 갖는 거리를 찾아낸다. 예컨대 〈그림 6〉에서 'this'와 'has'의 최소 패널티를 산정할 때 두 문자열의 최단거리가 나올 수 있는 대각선을 중심으로 대상 'this'를 원본 'has'와 대조하는 과정에서, 먼저 문자 't'가 삽입되어 패널티 1을 얻고, 다음으로 'a'대신 'i'로 치환되어 패널티 1을 얻어 총 패널티 2가 되는 것을 알 수 있다(Keith and David 1992, 7-8). 이와 같이 해당 열에서 삽입이나 삭제, 치환을 감안하는 거리 계산은 문자열 비교 탐색에서의 근사 스트링 매칭(approximate string matching)의 원리와 일맥상통하고 있다.

즉 edit distance의 원리에 의하면 소스 음렬이 템플릿 음렬을 기준으로 했을 때의 변형된 정도에 의해 두 음렬간의 유사도 값이 측정되며, 삭제와 삽입, 치환과 같은 변형 정도의 기준에 의해 각각의 값이 부여됨으로써 유사도 산정에 반영된다. 실제 Humdrum의 simil 프로그램에서는 두 개의 음렬을 대조할 때 다음 8개 경우의 연산에 대한 패널티를 디폴트로 모두 1로 부여하고 있다(David 1999).

D1 음렬 1에서 반복되지 않는 음표를 삭제할 경우
D2 음렬 2에서 반복되지 않는 음표를 삭제할 경우
R1 음렬 1에서 반복되는 음표를 삭제할 경우
R2 음렬 2에서 반복되는 음표를 삭제할 경우
S0 음렬 1과 음렬 2 모두에서 반복되지 않는 음표를 치환할 경우
S1 음렬 1에서만 반복되는 음렬을 치환할 경우
S2 음렬 2에서만 반복되는 음렬을 치환할 경우
S3 음렬 1과 음렬 2 모두에서 반복되는 음표를 치환할 경우

한편 Humdrum에서는 음렬의 유사도를

산정하는데 있어, simil 프로그램 이외에 피어슨 상관계수를 사용하는 correl 프로그램을 제공하고 있다. correl 프로그램을 사용하면 완전히 역으로 일치하는 경우의 값을 -1로 나타냄으로써 유사도 산정에 음렬의 역진행 패턴 내용을 반영할 수 있다. 예컨대 '도레미파'라는 내용을 지닌 소스 음렬과 '파미레도'의 템플릿 음렬간 유사도는 correl 프로그램에 의해 -1값을 갖게 된다. 그러나 본 실험에서는 유사성의 기준에서 음렬의 역진행 패턴을 배제하였기 때문에 correl 프로그램은 사용하지 않았다.

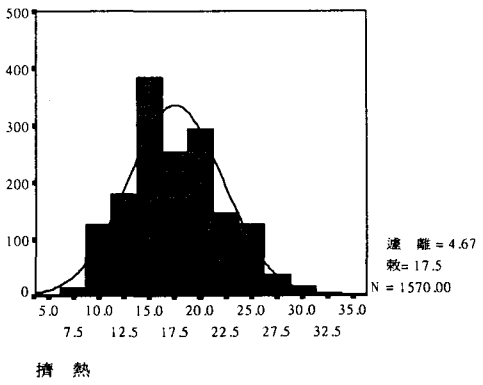
4 주제소절 자동분류

4.1 실험집단

4.1.1 문헌집단 선정

클래식 음악을 대상으로 약 10,000개의 주제소절을 음고에 따라 색인한 A *Dictionary of Musical Themes*에서 고전파 대표 작곡가인 베토벤(560개), 모차르트(587개), 하이든(423개)의 1570개 주제소절 음렬을 Themefinder 검색을 통해 **kern으로 코딩된 파일을 수집하였다. 이 때 각 음렬의 파일명은 사전에 표기되어 있는 코드(ex> B1044.KRN)를 그대로 사용하여 필요할 경우 해당 악보와 대조할 수 있도록 하였다.

1570개 음렬의 길이를 나타내는 각 음표수는 최소 6개, 최대 35개이며, 평균



〈그림 7〉 1570개 음렬을 구성하는 음표수 분포

17.5개로 〈그림 7〉과 같은 분포를 보이고 있다. 단, 음렬의 길이에서 붙임줄로 처리되어 있는 음은 1개로 계산하였으며, 쉼표는 제외하였다.

4.1.2 음렬 패턴 처리

본 실험에서 대상으로 하는 음렬은 단선율이다. 대상 문헌이 주제소절이라는 특성으로 인해 대부분의 경우 단선율 음렬에 해당하였으나, 1570개 파일을 대상으로 Humdrum의 simil 명령어를 적용해 본 결과, 10개의 파일에서 이상을 발견하였다. Humdrum 구문에서는 단선율일 경우 1개의 spine으로 표현되며, 두 개 이상의 성부를 가질 경우 spine이 추가되는 특성을 지니고 있다. simil은 유사도를 구하는 프로그램으로 대상 음렬이 정확하게 1개의 spine일 경우에만 적용되는 특성을 지니기 때문에, 이를 적용함으로써 단선율이 아닌 음렬을 식별할 수 있었다. 이들 10개의 파일에 해당하는 음렬을 사전

상의 악보와 대조한 결과 주선율 이외에도 2성부 혹은 3성부 선율로 이루어진 경우가 포함되어 있었다. 이에 주선율을 제외한 내용을 삭제하고 단선율 형태의 음렬로 수정하였다.

다음으로 **kern 형식의 음렬 데이터에서 리듬 정보를 제외하였으며(kern -x 파일명), 조성을 반영하기 위해서는 deg 명령어(deg 파일명)를 사용하여 음고 정보를 이동도법에 따라 상대 음도로 변환하였다.

마지막으로 deg 표현에 따라 다장조 혹은 가단조로 정규화된 음도 형태의 음렬을 대상으로 붙임줄을 포함한 경우들을 모두 1개의 음으로 수정하였으며, 마디(humsed /^=/d 파일명)와 쉼표(humsed /r/d 파일명)도 각각 humsed 명령어를 사용하여 제거하였다. 마디와 쉼표의 제거는 음렬 패턴을 찾을 때 연속적인 음렬 인식에 도움이 된다. 이와 같은 맥락에서 탭(tab)으로 분리되어있는 음렬의 행벡터를 열벡터로 전환하였다(paste -s 파일명).

4.2 자질 선정

4.2.1 음렬의 분절

음렬 패턴 처리 결과로 얻어진 이동도법의 음도 표현 음렬을 대상으로 출현빈도와 음 진행 형태를 기준으로 음렬을 분절하였다.

〈표 1〉 2-gram 음렬 패턴의 음 진행 형태별 그룹 형성

그룹	음 진행 형태	1열	2열	출현빈도	그룹	음 진행 형태	1열	2열	출현빈도
1	1음 아래 진행	v2	v1	637	9	4음 아래 진행	^2	v5	84
		v4	v3	606			^5	v1	64
		v3	v2	570			^6	v2	34
		v1	v7	512			v2	v5	24
		^5	v4	399			^4	v7	22
		v5	v4	394			^1	v4	20
		^4	v3	388			^3	v6	12
		^6	v5	381			v5	v1	11
		^3	v2	331			v4	v7	5
		v6	v5	300			^7	v3	3
		v7	v6	271			v6	v2	2
		^2	v1	262			v1	v4	1
		^1	v7	223			v3	v6	1
		^7	v6	41			v7	v3	0
2	1음 위 진행	v7	^1	474	10	7음(1옥타브) 위 진행	v5	^5	50
		^2	^3	442			v1	^1	48
		^3	^4	402			^4	^4	43
		^1	^2	382			^1	^1	34
		^4	^5	340			^5	^5	32
		v2	^3	269			^2	^2	16
		v3	^4	251			v6	^6	11
		^7	^1	242			v3	^3	9
		^5	^6	240			v4	^4	6
		v1	^2	217			^3	^3	6
		^6	^7	181			v2	^2	5
		v5	^6	153			v7	^7	1
		v4	^5	75			^7	^7	0
		v6	^7	58			^6	^6	0
3	동일음 반복	5	5	515					
		1	1	464					
		2	2	403					
		3	3	378					
		4	4	267					
		6	6	114					
		7	7	105					

〈표 1〉 2-gram 음렬 패턴의 음 진행 형태별 그룹 형성-계속

그룹	음 진행 형태	1열	2열	출현 빈도	그룹	음 진행 형태	1열	2열	출현 빈도
4	2음 아래 진행	^3	v1	259	11	4음 위 진행	v1	^5	98
		^5	v3	235			v5	^2	46
		^4	v2	229			v7	^4	35
		v3	v1	201			v2	^6	30
		v4	v2	144			^1	^5	20
		^6	v4	122			v4	^1	7
		v2	v7	110			^5	^2	5
		v7	v5	108			v3	^7	4
		^2	v7	97			v6	^3	2
		v5	v3	96			^7	^4	1
		^1	v6	91			^6	^3	1
		v1	v6	37			^2	^6	1
		v6	v4	34			^4	^1	0
		^7	v5	19			^3	^7	0
5	2음 위 진행	^1	^3	256	12	5음 아래 진행	^3	v5	46
		^3	^5	255			^1	v3	27
		v3	^5	155			^6	v1	26
		v7	^2	143			^5	v7	23
		v1	^3	142			^2	v4	18
		v2	^4	124			^4	v6	8
		^4	^6	87			v2	v4	7
		^2	^4	87			v3	v5	6
		v5	^7	45			v1	v3	6
		^7	^2	31			v6	v1	3
		^5	^7	27			v5	v7	2
		v4	^6	27			v7	v2	1
		v6	^1	22			^7	v2	0
		^6	^1	21			v4	v6	0
6	3음 위 진행	v5	^1	184	13	7음(1옥타브) 아래 진행	^1	v1	19
		^5	^1	136			^5	v5	19
		v2	^5	93			^3	v3	8
		v3	^6	71			v7	v7	7
		v1	^4	62			^4	v4	6
		v6	^2	29			^2	v2	5
		^1	^4	24			^6	v6	4
		^2	^5	23			v6	v6	3
		^3	^6	13			v3	v3	3
		v7	^3	12			v4	v4	2
		v4	^7	5			v5	v5	2
		^6	^2	4			v1	v1	1
		^4	^7	4			^7	v7	1
		^7	^3	3			v2	v2	0
7	3음 아래 진행	^1	v5	180	14	6음 위 진행	v5	^4	42
		v1	v5	106			v7	^6	11
		^5	v2	43			^5	^4	5
		^3	v7	30			v6	^5	3
		^6	v3	27			v3	^2	3
		^4	v1	16			v2	^1	3
		v7	v4	12			v1	^7	2
		^2	v6	9			^7	^6	1
		v5	v2	8			^1	^7	1
		v2	v6	5			^6	^5	0
		v6	v3	4			^2	^1	0
		v3	v7	4			v4	^3	0
		v4	v1	5			^4	^3	0
		^7	v4	3			^3	^2	0
8	5음 위 진행	v5	^3	92	15	6음 아래 진행	^6	v7	12
		v3	^1	54			^4	v5	11
		v1	^6	45			^1	v2	11
		v7	^5	39			^3	v4	5
		v6	^4	21			^2	v3	4
		v4	^2	16			^5	v6	4
		^1	^6	11			v4	v5	2
		v2	^7	7			v2	v3	2
		^4	^2	7			v7	v1	1
		^5	^3	6			v5	v6	1
		^3	^1	4			v1	v2	1
		^6	^4	2			^7	v1	0
		^2	^7	2			v6	v7	0
		^7	^5	0			v3	v4	0

출현빈도를 계산할 때, 같은 음렬에서 해당 음이 2회 이상 출현한 경우에 대해서는 가중치를 주지 않고 모두 1값으로 계산하였다. 또한 올림표(#)와 내림표(♭)가 붙은 음의 경우 별도로 분리하지 않고, 조표 유무와 상관없이 원음과 같은 부류로 포함하여 계산하였다. 이는 'v4+(아래로 진행하는 파#)'음과 'v1+(아래로 진행하는 도#)'음을 제외하고는 해당 조표가 붙는 음이 차지하는 비율이 전체 문헌집단에서 0~3% 정도로 미미했기 때문이다.

한편 음렬의 분절에서 n-gram 길이를 특정 수치로 고정시켜 의미 단위를 파악할 경우, 출현빈도 상으로는 같은 비중을 차지하면서도 각기 다른 길이를 지닌 음렬들을 반영할 수 없다. 이 문제를 보완하기 위해 본 실험에서는 동일음 반복이나 순차 진행 패턴과 같이 고빈도에 집중하는 음렬의 경우 n의 수치를 늘려, 보다 긴 음렬 단위로 분절함으로써 n-gram 길이에 융통성을 부여하였다.

결과적으로 1570개 주제소절을 대상으로 음의 진행 형태에 따른 2-gram 음렬 패턴 그룹을 〈표 1〉과 같이 구성하였다. 이는 아래로 진행하는 'v'와 결합한 도, 레, 미, 파, 솔, 라, 시를 각각 나타내는 v1, v2, v3, v4, v5, v6, v7과 위로 진행하는 '^'과 결합한 ^1, ^2, ^3, ^4, ^5, ^6, ^7로 구성된 14×14개의 조합과 더불어, 높낮이 정보를 포함하지 않는 음 진행에 해당하는 '11', '22', '33', '44', '55', '66', '77'의 7가지 경우의 수를 함께 고려하여 총 203(14×14+7)개의 음렬 패턴으로 이루어졌다.

4.2.2 후보 자질 선정

Humdrum 구문의 deg 표현에서는 이웃하는 음이 없어 상대적으로 관계가 규명되지 않는 첫 음에 대해서도 그 이전 음으로부터의 높고 낮음의 방향 정보를 부여한다. 따라서 음렬을 구성하는 2-gram 음렬 패턴이 사실상 3-gram의 성격을 띠게 된다. 예컨대 〈표 1〉에서 8그룹의 'v5^3' 형태의 음렬은 그것이 속한 전체 음렬로부터 2-gram으로 분리되었을 경우 결과적으로 '솔' 다음에 위로 진행하는 '미'가 온다는 점에서 '^5^3'과 동일하다. 그러나 'v5^3'과 '^5^3'은 첫 음인 '솔' 이전에 올 수 있는 음의 범위가 분명하게 위와 아래로 양분된다는 점에서 선행음에 대한 정보를 내포하고 있다.

그러나 첫 음의 경우 n-gram 단위로 분절함에 따라 사실상 높고 낮음의 방향 정보를 상실하게 된다. 따라서 자질 음렬 패턴을 구성하는 단계에서는 첫 음이 포함하고 있는 높낮이 방향 정보를 제외시켰다. 한편 2-gram 음렬 패턴에서 음의 진행 형태별로 총 15개 그룹을 형성하였던 내용을 가능한 반영하면서도 지나치게 낮은 빈도를 제외한 다음, 자질 빈도 70에서 200에 해당하는 음렬을 후보 자질로 선정하였다. 이를 충족시키는 자질로는 2-gram에서 6-gram 사이의 총 81개 음렬

〈표 2〉 75개 자질 리스트

75개 자질 리스트

고유번호	음렬	출현빈도	고유번호	음렬	출현빈도	고유번호	음렬	출현빈도
1	2.v1.v7.v6	139	26	3.^4.^5.^6	97	51	5.v3.v1	151
2	2.v1.1	131	27	1.^2.v1	147	52	4.v2.v1	106
3	4.v3.^4	143	28	1.^2.^3.^4	176	53	4.v2.v7	74
4	4.v3.3	134	29	4.^5.v3	85	54	6.v4	158
5	3.v2.v1.v7.^1	84	30	4.^5.v4	104	55	2.v7.^1	76
6	3.v2.v1.^2	87	31	5.^6.v5.v4	123	56	1.v6	134
7	3.v2.^3	174	32	5.^6.^7	158	57	7.v5	130
8	3.v2.^4	79	33	6.^7.^1	177	58	1.^3.v2	158
9	3.v2.2	113	34	5.5.v3	72	59	1.^3.^5	178
10	1.v7.^1.^2	180	35	5.5.^1	115	60	3.^5.^1	88
11	1.v7.^2	111	36	5.5.v4	152	61	3.^5.v4	183
12	1.v7.v6.v5.v4	77	37	5.5.^6	90	62	7.^2.v1	79
13	5.v4.v3.v2.v1.v7	79	38	5.5.5.5	123	63	2.^4.v3	123
14	5.v4.v2	93	39	1.1.^3	76	64	4.^6	116
15	6.v5.v4.v3.v2	112	40	1.1.v7.^1	70	65	5.^7	97
16	6.v5.5	114	41	1.1.^2	101	66	5.^1.^3	146
17	7.^1.^3	112	42	1.1.1	162	67	2.^5	140
18	7.^1.v5	90	43	2.2.^3	80	68	1.^4	108
19	7.^1.v7	75	44	3.3.v2	164	69	3.^6	97
20	7.^1.^2.^3	157	45	3.3.^4	94	70	5.^3	151
21	2.^3.v1	131	46	3.3.3	102	71	2.v5	117
22	2.^3.v2	122	47	4.4.v3	134	72	5.v1	103
23	2.^3.^4.v3	86	48	6.6	114	73	5.^5	102
24	2.^3.^4.^5	129	49	7.7	105	74	1.^1	109
25	3.^4.v2	154	50	3.v1.v5	75	75	1.^5	162

이 해당되었다.

4.2.3 자질간 독립성 검증

선정된 자질 가운데는 자기 자신보다 짧은 길이의 음렬을 완전히 포함하고 있거나, 부분적으로 중복되는 경우가 있을 수 있다. 이와 같이 중복되는 자질의 경우 분류자질에서 제외시킬 필요가 있다. 이에 따라 자질의 중복성 여부를 파악하기 위하여 상관계수를 산출하였다. 1570개 파일에 대한 각 자질의 출현여부를 나타내는 0과 1의 이원벡터 행렬을 사용하여 피어슨 상관계수가 0.4 이상인 자질쌍 총 17개를 대상으로 자질간 중복 여부를 확인하였다. 그 결과 하나의 자질이 다른 자질에 완전히 포함되는 경우가 총 6개 있었으며, 부분적으로 중복되는 자질의 경우가 8쌍에 해당하였다. 이 가운데 다른 자질에 완전히 포함되는 6개 자질은 이보다 긴 음렬에서 이들 의미를 내포하고 있다는 판단 하에 모두 제외하였으나, 부분적으로 중복되는 8개 자질의 경우, 음렬의 연속적인 성질 내에서 고유한 의미를 지닐 수 있다는 가정에서 수용하였다. 이상의 과정을 통해 최종적으로 선정된 75개 자질은 〈표 2〉와 같다.

4.2.4 전체소절 대상

한편 주제소절의 음렬 길이가 6-gram에서 35-gram으로 비교적 짧고, 곡 전체 내에서 주제적 특성을 지녔다는 점은 이들 전체를 자질로 고려할 수 있음을 의미한다. 이처럼 해당 음렬 전체를 대상으로 할 경우 분절로 인한 의미 소실의 위험을 피할 수 있다. 그러나 음렬 전체를 자질로 사용할 경우 분절에 의해 상대적으로 짧은 자질 패턴을 사용한 경우보다 클러스터링 결과가 좋지 않으리란 예상이 가능하다. 패턴의 길이가 길수록 유사하지 않은 패턴이 나올 확률이 크기 때문에,

이에 영향을 받아 자질 패턴과 음렬간에 산정된 유사도의 의미가 약화될 수 있는 것이다.

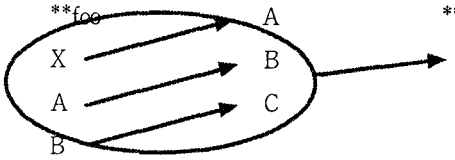
따라서 본 실험에서는 분절하지 않은 전체소절을 자질로 고려하되, 그 결과를 사용하는 목적은 이와의 대비를 통해 음진행 패턴과 중간빈도라는 일련의 기준에 의해 분절한 자질의 사용이 실질적으로 어느 정도의 유의미한 결과를 가져올 수 있는가를 알아내기 위한 것이다.

다음에는 자질의 선정 대상이었던 실험집단 1570개 규모의 약 3분의 1에 해당하는 베토벤의 500개 음렬(파일명 B521.KRN ~ B939.KRN, B1000.KRN~B1058.KRN)로 클러스터링 대상을 축소하여 그 결과를 제시하였다.

4.3 유사도 산출

실험집단을 대상으로 한 유사도 산출은 두 가지 형태로 나뉘어 진행되었다. 먼저 음렬 분석을 통해 유의미한 단위로 가정하고 분절한 75개 자질을 템플릿 음렬로 하여 500개 음렬에 대한 유사도를 구하였다. 다음으로 500개 음렬 자체를 모두 자질로 하여 500×500 조합을 대상으로 유사도를 구하였다. 이 경우 두 개의 음렬을 대조해 나가는 과정에서 상대적으로 길이가 짧은 음렬을 템플릿으로 한 경우에 대해서만 유사도 값이 나온다.

source input template input simil output



**foo	A	**simil
X	B	0.51
A	C	1.00
B		0.51
C		0.37
D		0.51
A		0.72
B		0.72
B		0.51
C		0.51
B		.
A		.
*-		*-

이는 Humdrum의 simil 프로그램이 소스 음렬을 템플릿 음렬과 비교해 나갈 때, 템플릿의 음렬 길이 전체를 하나의 단위로 하여 유사도를 구하기 때문이다. 한편 소스 음렬을 템플릿 음렬과 비교할 때, 첫 음을 시작으로 하여 차례로 이웃하는 음을 기준으로 진행해가며 비교하기 때문에 소스 음렬의 길이가 긴 경우에 산출되는 유사도 값의 개수가 더 많아진다. 이 경우 하나의 음렬에서 구해진 여러 개의 유사도 가운데 특정 유사도를 선택해야 하는 문제에 봉착하게 된다. 이는 음렬 내에서 분석하고자 하는 대상을 지정함으로써 해결될 수 있다.

특정 음렬을 질의로 사용할 경우 해당 음렬의 처음부터 시작하거나, 그렇지 않은 경우로 양분된다는 점에 착안해 음렬 내 위치에 따른 두 가지 분석 대상을 상

정하였다. 먼저 음렬의 시작인 첫 음을 기준으로 한 패턴을 대상으로 할 경우, 해당 음렬에서 구한 첫 번째 유사도를 대상으로 유사도 행렬을 구성하였다. 다음으로 특정 패턴의 시작 위치를 음렬 전체를 대상으로 할 경우 가장 두드러진 패턴을 기준으로 반영하기 위해, 여러 개의 유사도 가운데 최대값을 대상으로 하였다.

이상의 내용을 바탕으로 분절 유무에 따른 자질의 형태와 음렬 내 특정 패턴의 위치에 따른 네 가지 형태의 유사도 행렬을 〈표 3〉과 같이 구성하였다.

4.4 클러스터링

〈표 3〉 클러스터링을 위한 유사도 행렬 형태

	75개 분류자질 사용	전체음렬 사용
음렬의 시작 대상	①	②
음렬의 전체 대상	③	④

〈표 3〉의 네 가지 형태의 유사도 행렬을 대상으로 계층적 클러스터링 기법의 하나인 Ward 방법을 사용하였다. 75개 분류 자질을 대상으로 한 경우는 피어슨 상관계수를 측도 간격으로 사용하여 각 자질과 대상 음렬간의 상관관계를 구하였으며, 전체 음렬을 자질로 한 경우는 simil 프로그램에 의해 구해진 유사도를 그대로 사용하였다. 한편 소속 군집으로는 클러스터링 평가 기준표에 근거해 50 군집으로 지정하였다.

4.5 분석 결과

4.5.1 평가 기준

음렬의 시작 부분을 대상으로 클러스터링한 결과를 평가하기 위한 기준은 실험 집단의 실제 악보를 수록하고 있는 *A Dictionary of Musical Themes*의 Notation Index를 이용하였다. 이는 사전에 수록된 주제소절 가운데 첫 번째 음을 기준으로 4-gram으로부터 13-gram 길이로 발췌하여 이동도법으로 코딩한 권말 색인이다.

Notation Index는 'C', 'D', 'E', 'F', 'G', 'A', 'B'로 표현된 각 음과 내림표와 올림표의 조합으로 형성된 음렬의 알파벳순 배열로써, 이는 음렬의 속성상 유사한 음렬끼리 모여 있다는 특성을 지닌다. 단, 여기에는 실험집단 이외의 주제소절에 대한 내용들을 담고 있으므로, 이를 제외한 내용들을 추출하여 알파벳순 정렬을 하고, 수작업으로 최소한 첫 소절에 해당하는 2-gram 이상 동일한 음렬을 기준으로 총 50 여개의 유사한 군집을 형성하였다.

한편 음렬의 전체 부분을 대상으로 특정 패턴을 클러스터링한 결과를 평가하기 위한 기준은 음렬 분석을 통해 얻어진 75개 자질 패턴이 음렬 내에서 분포하는 정도를 기준으로 삼았다. 이를 위해 500개 음렬이 각각 포함하고 있는 자질 패턴 분포를 〈표 4〉와 같이 나타내었다. 이 때 75개 자질의 고유번호는 〈표 2〉에서 부여한 것을 그대로 사용하기로 한다.

〈표 4〉 음렬 내 자질 패턴 분포

파일명	자질 패턴 분포										
B1000.KRN	10	27									
B1001.KRN	7	22	25								
B1002.KRN	20	28	29	49	50	51					
B1003.KRN	52	68									
B1004.KRN	18	67	71								
B1005.KRN	8	44	47	48	69						
B1006.KRN	5	10	13	20	28	64					
B1007.KRN	9	21	25	43	44						
B1008.KRN	14	17	21	57	59	61	65				
B1009.KRN	19	33	49	51							
B1010.KRN	14	50	53	61	62	73					
B1011.KRN	2	14	20	32	33	36	39	44	59		
					<중략>						
B930.KRN	34	39	51	58							
B931.KRN	40	55	74	75							
B932.KRN	11	14	34	52	58	59	61	63			
B933.KRN	17	36	38	53							
B934.KRN	2	4	10	16	20	24	28	36	40	44	73
B935.KRN	29										
B936.KRN	26	54	64	68							
B937.KRN	25	28	52	67							
B938.KRN	21	26	27	32	33						
B939.KRN	3	25	61	67	72						

4.5.2 평가 척도

음렬의 시작 부분을 대상으로 한 클러스터링 결과 평가에는 Notation Index라는 외적 기준이 있으므로, 평가 척도는 클러스터링 결과와 Notation Index 기준과의 대비를 통한 형태를 선택하였다.

클러스터링 결과를 평가하기 위한 기준이 되는 수작업 분류표와 대비할 때, 비교 대상이 되는 결과의 군집과 분류표의 군집을 동일선상에 놓고 평가하기에는 무리가 있다. 이는 두 개의 비교 대상이 완전히 일치하지 않는 한, 각각의 군집 수와 각 군집에 속하는 요소들의 구성이 다르기 때문이다.

이러한 문제점을 해결하기 위한 방편으로 클러스터링 결과와 수작업 분류표를 동일한 조건에서 비교하기 위하여 정규화된 평가척도가 제시되어 왔다. 이 가운데 클러스터링 결과를 기준으로 각 요소가 수작업 분류표의 군집에 속한 정도의 비율을 반영하여 가중치를 준 WACS (Weighted Average Cluster Similarity)를 사용하여 실험 결과를 평가하였다(정영미, 이재윤 2001, 208-209).

$$WACS(C) = \frac{1}{D}\sum_{j=1}^{n}\sum_{i=1}^{m}\frac{2|M_i \cap C_j|^2}{|M_i| + |C_j|}$$

반면 음렬 전체를 대상으로 한 경우,

특정 위치에 출현하는 패턴을 기준으로 군집을 형성해 놓은 절대적 기준이 없기 때문에, 각 음렬 내에서 75개 자질 패턴이 분포한 형태를 간접적인 기준으로 삼은 바 있다. 이 때, 클러스터링 결과로 군집으로 묶여진 음렬을 기준으로 동일한 자질 패턴들이 어느 정도 함께 출현하고 있느냐에 착안했다는 사실에 주목할 필요

〈표 5〉 75개 분류 자질을 사용한 결과로 나온 군집별 자질 분포

군집	자질	자질 음렬 패턴	출현빈도(F)	군집 크기(Cn)	F/Cn
2	44	3.3.v2	8	10	0.80
	7	3.v2.^3	5	10	0.50
	9	3.v2.2	5	10	0.50
	43	2.2.^3	5	10	0.50
3	20	7.^1.^2.^3	5	8	0.63
	28	1.^2.^3.^4	4	8	0.50
	49	7.7	4	8	0.50
4	52	4.v2.v1	4	8	0.50
	68	1.^4	4	8	0.50
	75	1.^5	4	8	0.50
6	44	3.3.v2	8	11	0.73
	46	3.3.3	6	11	0.55
7	5	3.v2.v1.v7.^1	8	11	0.73
	13	5.v4.v3.v2.v1.v7	7	11	0.64
	10	1.v7.^1.^2	6	11	0.55
	15	6.v5.v4.v3.v2	6	11	0.55
9	49	7.7	7	8	0.88
10	48	6.6	6	11	0.55
11	47	4.4.v3	6	8	0.75
	48	6.6	4	8	0.50
12	43	2.2.^3	5	7	0.71
	9	3.v2.2	4	7	0.57
	61	3.^5.v4	4	7	0.57
13	61	3.^5.v4	8	11	0.73
	13	5.v4.v3.v2.v1.v7	6	11	0.55
14	36	5.5.v4	8	11	0.73
	44	3.3.v2	8	11	0.73
	47	4.4.v3	7	11	0.64
17	6	3.v2.v1.^2	6	10	0.60
18	28	1.^2.^3.^4	7	9	0.78
19	10	1.v7.^1.^2	6	11	0.55
20	61	3.^5.v4	7	9	0.78
21	39	1.1.^3	6	11	0.55
	42	1.1.1	6	11	0.55
22	47	4.4.v3	5	6	0.83
24	2	2.v1.1	6	9	0.67
	62	7.^2.v1	6	9	0.67
25	22	2.^3.v2	4	8	0.50
28	33	6.^7.^1	16	22	0.73
	32	5.^6.^7	15	22	0.68
	28	1.^2.^3.^4	14	22	0.64
	20	7.^1.^2.^3	12	22	0.55
29	51	5.v3.v1	12	23	0.52
30	33	6.^7.^1	4	7	0.57
	32	5.^6.^7	4	7	0.57
	17	7.^1.^3	4	7	0.57

- 후략 -

가 있다. 단순히 특정 자질 패턴의 출현 여부를 조건으로 하는 것은 해당 패턴과 일치하는 것을 찾는 것으로 복잡한 클러스터링 과정 없이도 쉽게 그 결과를 얻을 수 있기 때문이다. 한편, 이와 같이 자질 패턴을 단서로 삼은 것은 연속적인 음렬 내에서 해당 자질 패턴을 지배적으로 가지고 있는 음렬이 유사할 확률이 높다는 전제로부터 비롯된다.

이에 따라 음렬 전체를 대상으로 한 클러스터링 평가 척도는 클러스터링 결과에 해당하는 각 군집 내의 음렬들이 해당 군집 내에서 50%이상 공통 자질 패턴으로 출현하는 음렬을 1개 이상 지닌 군집을 적합한 군집으로 삼고, 이들이 전체 군집에서 차지하는 비율을 반영하는 한편, 적합한 군집의 내용을 반영하기 위해 각 군집의 크기를 고려한 출현빈도 비율을 함께 적용하여 평가하였다.

〈표 5〉는 75개 분류 자질을 사용한 결과로 나온 각 군집을 대상으로 자질 패턴 분포를 분석한 결과 F/C_n 가 0.5 이상인 자질을 포함한 군집, 즉 해당 군집 내에서 50% 이상 공통 자질 패턴을 가지는 군집만을 선별한 내용이다. 결과로 제시된 전체 50군집 가운데 총 40개의 군집이 해당 기준에 의해 적합한 군집으로 판명되었다. 실제 클러스터링 결과는 이를 전체 군집으로 나눈 40/50 값과, 40개의 적합한 군집에 포함된 F/C_n의 모든 수치를 평균 낸 0.66 값을 곱하는 방식으로 계산하였다.

〈표 6〉은 전체 음렬을 사용한 결과로 나온 각 군집을 대상으로 자질 패턴 분포를 분석한 결과로서, 이 경우 75개 자질을 사용한 결과에 비해 기준치 내에 포함되는 군집이 17개로 사실상 전체 군집 50개 가운데 적합한 군집으로 판명된 것이 절반 이하임을 알 수 있다. 한편, 0.5 이상인 F/C_n의 모든 수치를 평균 낸 값은 0.67로 적합한 군집 내 공통 자질 패턴의 포함 정도는 〈표 5〉의 경우와 거의 비슷하였다.

4.5.3 평가 결과

먼저 WACS를 사용하여 음렬의 시작 부분을 중심으로 클러스터링한 결과를 평가한 내용은 다음과 같다.

① 75개 분류 자질을 사용하고 음렬의 시작을 대상으로 한 경우: 0.6539
② 전체 음렬을 사용하고 음렬의 시작을 대상으로 한 경우: 0.1505

한편 부분적으로 중복된 자질을 모두 제외하고 75개 자질을 67개로 축소한 것을 음렬의 시작을 대상으로 한 클러스터링에 적용해 본 결과의 값이 0.6431로 나와, 부분적으로 중복되는 자질을 그대로 사용하여 자동분류해도 무방하며 오히려 더 좋은 결과가 나올 수 있음을 알 수 있었다.

다음으로 각 군집에 속한 음렬 내 자질 패턴 분포를 기준으로, 음렬 전체의 임의

〈표 6〉 전체 음렬을 사용한 결과로 나온 군집별 자질 분포

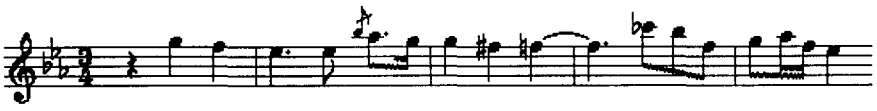
군집	자질	자질 음렬패턴	출현빈도(F)	군집 크기(Cn)	F/Cn
2	7	3.v2.^3	3	6	0.50
6	10	1.v7.^1.^2	7	14	0.50
7	43	2.2.^3	2	3	0.67
13	24	2.^3.^4.^5	3	5	0.60
14	44	3.3.v2	4	8	0.50
17	63	2.^4.v3	8	15	0.53
22	33	6.^7.^1	4	5	0.80
24	10	1.v7.^1.^2	4	8	0.50
25	32	5.^6.^7	5	6	0.83
	33	6.^7.^1	5	6	0.83
	20	7.^1.^2.^3	3	6	0.50
35	4	4.v3.3	3	5	0.60
	44	3.3.v2	3	5	0.60
37	20	7.^1.^2.^3	3	3	1.00
	32	5.^6.^7	3	3	1.00
	33	6.^7.^1	3	3	1.00
	22	2.^3.v2	2	3	0.67
38	34	5.5.v3	3	5	0.60
43	58	1.^3.v2	5	9	0.56
44	5	3.v2.v1.v7.^1	2	2	1.00
	58	1.^3.v2	2	2	1.00
45	31	5.^6.v5.v4	3	4	0.75
	8	3.v2.^4	2	4	0.50
	63	2.^4.v3	2	4	0.50
47	59	1.^3.^5	3	6	0.50
49	5	3.v2.v1.v7.^1	2	4	0.50

의 위치에 나타나는 특정 패턴을 대상으로 한 클러스터링 결과를 평가한 내용은 다음과 같다.

③ 75개 분류 자질을 사용하고 음렬의 전체를 대상으로 한 경우: 0.5280
④ 전체 음렬을 사용하고 음렬의 전체를 대상으로 한 경우: 0.2278

4.5.4 결과 요약 및 논의

본 연구에서 사용한 실험집단은 클래식 음악의 기악곡 가운데 주제소절로 곡 전체에 비해 극히 짧은 음렬 단위를 대상으로 하였다. 이 때 주제소절 자체를 곡 전체를 대표하는 의미 자질로 사용할 수 있다. 그러나 이를 곡 전체적인 맥락에서 사용하지 않고 독립적으로 분리하여 자동분류할 경우, 음렬 길이의 영향으로 인해 분절한 자질을 사용한 경우보다 상대적으로 군집 결과가 좋지 않다는 사실이 입증되었다.

음이 진행하는 속성은 특정 음렬 내에서 첫 음으로부터 이웃하는 음으로 차례로 진행해가며 각기 다른 형태의 음렬을 형성해간다. 이 때문에 하나의 단위로 간주되는 음렬 내에도 여러 개의 음렬이 내포되기 마련이다. 예컨대 〈그림 8〉의 악보를 계이름으로 나타내었을 때, '미레도도솔파미미레#레라#솔레미파레도'로 총 17개의 음으로 구성되어 있음을 알 수 있

<그림 8> B567 음렬에 해당하는 악보

〈표 7〉 음렬 내 시작 위치에 따라 내포된 음렬 형태

음렬의 시작 위치	음 렬
첫 번째 음	'미레도도솔파미미레#레라#솔레미파레도'
두 번째 음	'레도도솔파미미레#레라#솔레미파레도'
세 번째 음	'도도솔파미미레#레라#솔레미파레도'
네 번째 음	'도솔파미미레#레라#솔레미파레도'
다섯 번째 음	'솔파미미레#레라#솔레미파레도'
여섯 번째 음	'미미레#레라#솔레미파레도'
이하 생략	

다. 이는 음렬의 시작 위치를 첫 번째 음에서 하나씩 진행해감에 따라 각각 한 음씩 길이가 짧은 음렬을 〈표 7〉과 같이 형성해 나가는 것이다.

위와 같은 음렬의 속성은 하나의 주제소절에서 내포하는 의미 단위를 분절할 때 각 음렬 내에서 시작 위치를 고려하게 한다. 예컨대 검색 시스템의 이용자가 노래를 기억해 낼 때 '노래의 시작' 부분부터 부를 수도 있지만, 그 곡의 '임의의 위치'에서도 시작할 수 있는 것이다. 따라서 본 실험의 자질 선정 단계에서는 해당 자질 패턴이 출현하는 위치를 음렬의 시작으로 고정시키지 않고 음렬 내 모든 부분을 반영하도록 하였다. 또한 자동분류를 할 때, 주제소절이 시작하는 부분을 중심으로 유사한 진행 패턴을 보이는 음렬을 모아줌과 동시에, 주제소절 내 특정 위치에서부터 발견되는 유사한 음렬 패턴끼리 모이도록 하였다.

한편 자동분류에 앞서 음렬 내에서 의미를 지닌 단위로 구분할 수 있는 가능한 방법을 찾아내는 것이 주요한 과제였다. 본 실험에서는 음렬의 진행 패턴과 출현빈도 정보를 동시에 반영하여 의미 단위로 간주하고 분절한 자질을 선정하였다. 이와 같은 방법으로 선정된 자질을 사용한 군집결과는 주제소절에 대해 어떠한 통제를 가하지 않고 그 자체를 그대로 자질로 사용한 경우와 뚜렷한 차이를 보이고 있음이 드러났다. 이 때 음렬의 진행 패턴이나 출현빈도라는 일차적 요인과는 무관하게 분절로 인한 짧은 음렬 단위라는 변수가 군집결과에 보다 직접적 요인

으로 작용하였을지도 모른다. 그럼에도 불구하고 음 진행 패턴과 출현빈도를 사용한 분절 방법은 비교적 좁은 음역(音域) 내에 12개의 음들이 밀집해 있는 음렬에서 유사성이 발견될 수 있는 단위를 일차적으로 검증해 주었다는데 그 의미가 있다.

5 결 론

음렬을 대상으로 자동분류를 하는 것은 결국 음렬 내에 의미를 내포한 분류 자질이 있다는 가정으로부터 비롯된다. 언어 체계와 달리 구문 구조가 없는 음악에서 의미 단위를 명료하게 정의할 수 있는 것은 아니다. 그러나 음악 기호학과 형태주의(gestalt) 원리에 따르면 음악도 언어와 마찬가지로 의미 단위와 유사성의 형태가 발견될 수 있다는 의견이 지배적이다. 본 연구에서는 이와 같은 내용을 수용하여 음렬에서 '음의 반복'과 '음의 진행 형태'를 반영하여 분절한 자질을 자동분류에 적용하였다.

자동분류 실험 과정에서 음렬을 선정할 때 곡 전체를 대상으로 하기보다는 주제적 함축성이 높은 주제소절을 사용하였다. 이 때 6-gram에서 35-gram 범위 내의 주제소절 단위에서도 유사성의 단위가 발견되는지의 여부를 알아내기 위해, 주제소절 전체를 자질로 사용하여 보았다. 이는 주제소절 내에서 의미 단위를 고려하여 분절한 자질을 사용하여 클러스터링한 결과와의 비교 대상으로 사용할 수 있었다.

한편 음렬을 텍스트 형태로 표현해주는 구문 가운데 유닉스 환경에서 음렬 패턴 처리를 할 수 있도록 **kern 형식으로 표현된 Humdrum 구문과, 이를 처리해주는 프로그램 패키지인 Humdrum Toolkit을 사용하여 음렬 패턴 처리와 유사도를 구하였다. 이 때 음렬이 내포하는 여러 가지 정보 가운데 조성과 옥타브를 포함한 음도 형태의 음렬로 가공하여 가능한 높은 수준의 선율 정보를 포함하고자 하였다. 단, 실험문헌 집단의 크기가 크지 않은 점을 고려하여 리듬 정보는 제외하였다. 리듬 정보까지 반영할 경우 유사한 단위로 묶이는 주제소절의 수가 적어져 군집의 수가 지나치게 많아질 것으로 예상되었기 때문이다.

자질 선정 과정에서는 음렬의 표현 수준과 구분 단위를 함께 고려하였다. Humdrum의 deg 프로그램을 사용하여 음렬의 표현 수준을 음도 형태로 표현하였으며, 구분 단위의 경우 음의 반복과 진행 형태를 기준으로 하여 grep 명령어와 regular expression을 사용하여 75개 분류 자질을 선정하였다.

주제소절과 자질 패턴간의 유사도를 산정하는데 있어서는 edit distance의 원리에 따른 Damerau-Levenshtein 거리계수를 사용하였다. 실제로 Humdrum의 simil 프로그램을 사용하여 소스 음렬과 템플릿 음렬간의 유사도를 구하였다. 이러한 방

식을 사용하는 것은 기본적으로 음렬을 구성하는 모든 음에 대한 중요도를 동일하게 취급하는 것을 전제로 한다. 또한 음렬간 유사도를 판별하는 기준이 각 음의 진행 패턴 거리상의 근접성을 통한 것이라고 볼 수 있다.

클러스터링을 위한 유사도 행렬은 네 가지 형태로 구성하였다. 먼저 자질로 사용하는 템플릿 음렬을 구성할 때, 주제소절로부터 분절한 경우와 주제소절 전체를 사용한 경우의 두 가지 형태가 있다. 다음으로 주제소절 내에서 유사한 단위로 묶을 위치에 따라 주제소절의 시작과 전체를 대상으로 한 두 가지 형태의 행렬을 만들었다. 주제소절의 처음 부분부터 시작되는 음렬인 경우에는 첫 번째 유사도 값만을 반영한 행렬을, 주제소절 내에서 가장 뚜렷하게 유사성이 발견되는 임의의 지점으로부터 시작되는 음렬인 경우에는 해당 유사도 값들 중에서 최대값을 반영하여 행렬을 만들었다.

음렬의 분절 여부와 시작 위치에 따른 네 가지 형태의 유사도 행렬을 대상으로 계층적 클러스터링 기법을 사용하여 클러스터링을 하였다. 이들 결과에 대한 평가는 외적 기준이 되는 수작업 분류표가 있는 경우 WACS 척도를 사용하였고, 음렬 내 임의의 위치에서부터 시작한 음렬을 대상으로 한 경우, 클러스터링 결과로부터 얻어낸 군집 내 공통 자질 패턴 분포를 통해 내적 기준을 마련하여 평가하였다.

평가 결과에 의하면 음렬의 시작 위치와 무관하게 분절한 자질을 사용하여 클러스터링한 결과가 그렇지 않은 것에 비해 분명한 차이를 보이며 높게 나타났다. 따라서 곡 내에서 의미를 함축하는 주제소절일지라도 그 자체만을 독립적으로 자질로 사용하여 유사한 음렬끼리 모이게 하는 것은 사실상 무의미한 것으로 드러났으며, 분절한 자질 사용의 필요성을 입증하였다.

본 연구에서는 음렬 내에서 의미를 지닌 단위로 구분하기 위해 패턴의 반복과 음의 진행 형태를 감안한 바 있다. 그럼에도 불구하고 음렬의 의미 검증을 위해서는 분명한 의미 단위를 발견해내고 이를 척도화하기 위한 노력이 요구되고 있다. 이에 음악적 진행에 관한 일련의 규칙들과 형태주의 원리—근접의 원리(principle of proximity), 유사성의 원리(principle of similarity), 원만한 연결의 원리(principle of good continuation)—와 같은 음악분석의 원리를 이해하는 것이 선행되어야 할 것이다. 이를 통해 음렬의 자동분류 과정, 특히 자질 선정 과정에서 야기되는 의미의 불명확성이 해결될 수 있을 것으로 보며, 이는 결과적으로 음렬의 자동분류 결과에 대한 타당성에 직접적으로 영향을 미칠 것이리라 본다.

참고문헌

김승길. 1986. 음악 분석의 통계적 방법:

음계와 조성의 추출을 중심으로. 『음악 연구』, 제5집: 115-147.

백낙호, 서우석. 1986. 음악기호학 연구. 『음악 연구』, 제5집: 33-53.

이석원. 1994. 『음악심리학: 음악적 경험의 과학』. 서울: 심설당.

전지호. 1988. 조성음악에 있어서의 음정 연쇄에 대한 구조적 표현. 『음악 연구』, 제4집: 96-147.

정영미, 이재윤. 2001. 지식 분류의 자동화를 위한 클러스터링 모형 연구. 『정보관리학회지』, 18(2): 203-230.

Alexander, McLane. 1996. "Music as Information." *ARIST*, 31: 225-262.

Alexandra, L. Uitdenbogered and Justin, Zobel. 1998. "Manipulation of Music For Melody Matching." *Proceedings of the ACM multimedia 98*: 235-240.

Cambouropoulos, E. 1998. *Towards a General Computational Theory of Musical Structure*. Ph.D. diss., University of Edinburgh.

Chin-Chin, Liu, Jia-Lien, Hsu and Arbee, L. P. Chen. 1999. "An Approximate String Matching Algorithm for Content-Based Music Data Retrieval." *Proceedings of the IEEE International Conference on Multimedia Computing and Systems Volume I.*

ChulYong, Yang, JongTak, Shin, JinWook, Kim and HangJoon, Kim. 1999. "Korean Folk Song Retrieval using Rhythm Pattern Classification." Fifth International Symposium on Signal Processing and its Applications: 123-126.

David, Bainbridge, Craig, G. Nevill-Manning, Ian, H. Witten, Lloyd, A. Smith and Rodger, J. McNab. 1999. "Towards a Digital Library of Popular Music." *Proceedings of the Fourth ACM International Conference on Digital Libraries*: 161-169.

David, Huron. 1999. "Music Research Using Humdrum A User's Guide." [cited 2001.11.25]. 〈http://www.music-cog.ohio-state.edu/Humdrum/guide.toc.html〉.

Francu, C. and Nevill-Manning, C. G. 2000. "Distance Metrics and Indexing Strategies for a Digital Library of Popular Music." In *Proc. of IEEE International Conference on Multimedia and Expo*: 889-892.

Heinonen, Yrjo. 2000. "Mi-re-do-la-so : A pentatonic pattern and its variations in the music of the Beatles." [cited 2002.2.13]. 〈http://www.jyu.fi/musica/henkilosto/heinonen_abstraktit.html.〉

Jia-Lien, Hsu, Chin-Chin, Liu and Arbee L. P. Chen. 1998. "Efficient repeating pattern finding in music databases." *Proceedings of the 1998 ACM 7th international conference on Information and knowledge management*: 281-288.

Karin, Höthker, Dominik, Hönel and Christina, Anagnostopoulou. 2001. "Investigating the Influence of Representations and Algorithms in Music Classification." *Computers and the Humanities*, 35: 65-79.

Keith, Orpen and David, Huron. 1992. "Measurement of similarity in music: A quantitative approach for non-parametric representations." *Computers in Music Research*, Vol. 4: 1-44.

Kjell, Lemström, Pauli, Laine and Sami, Perttu. 1999. "Using relative interval slope in music information retrieval." In *Proceedings of the International Computer Music Conference*: 317-320.

Krumhansl, C. L. 1990. *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.

Marcel, Mongeau and David, Sankoff. 1990. "Comparison of musical sequences." *Computers and the Humanities*, 24: 161-175.

Pickens, Jeremy. 2001. "A Survey of Feature Selection Techniques for Music Information Retrieval." CIIR Technical Report. [cited 2002.1.8]. <http://citeseer.nj.nec.com/460365.html>.

Rodger, J. M., Lloyd, A. S., Ian, H. W., Clare L. H. and Sally J. C. 1996. "Towards the Digital Music Library: Tune Retrieval from Acoustic Input." *Proceedings of the ACM Digital Libraries*: 11-18.

Stephen, J. Downie. 1999. "Music Retrieval as Text Retrieval: Simple Yet Effective." *ACM SIGIR*, 99: 297-298.

Steven, Blackburn and David, De Roure. 2000. "Musical Part Classification in Content Based Systems." Papers from the 6th Workshop on Open Hypermedia Systems. May 30-June 3, San Antonio, Texas, USA. [cited 2002.1.8]. <http://www.ifs.uni-linz.ac.at/ifs/staff/reich/ohs6/docs/pp/blackburn/blackburn.pdf>.

Ta-Chun, Chou, Arbee, L. P. Chen and Chin-Chin, Liu. 1996. "Music Databases: Indexing Techniques and Implementation." *Proceedings*

of the 1996 International Workshop on Multi-Media Database Management Systems: 46-53.

Takashi, Yanase, A. Takasu, T. Kanazawa and J. Adachi. 1999. "Music Structure Analysis and Its Application to Theme Phrase Extraction." 3rd European Conference on Digital Library: 92-105.

Tsuji, Y., M. Hoshi and T. Ohmori. 1997. "Local Patterns of a Melody and Its Applications to Retrieval by Sensitive Words." Technical Report of IEICE, SP96(124): 17-24.

Youngmoo, E. Kim, Wei, C., Ricardo, G and Barry, Vercoe. 2000. "Analysis of a Contour-based Representation for Melody." Music IR 2000 International Symposium on Music Information Retrieval: Posters. [cited 2001.10.2].
〈http://web.media.mit.edu/~chaiwei/papers/MusicIR_melody.pdf〉

Yuen-Hsien, Tseng. 1999. "Content-based retrieval for music collections." *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*: 176-182.