

## 디지털 음악정보 검색 시스템의 설계

(Design of a Retrieval System for Digital Music Information)

지정규(숭실대학교)

오해석(숭실대학교)

### 요 약

본 논문은 디지털 도서관에서 임의의 음악정보를 효율적으로 검색하기 위한 시스템을 제안한다. 종래의 전형적인 음악정보 검색 항목인 제목이나 작곡자 또는 주제를 입력하는 것이 아니라, 사용자가 음악 데이터베이스로부터 검색하고자 하는 음악의 일부를 마이크를 통해서 노래한다. 그러면 입력된 선율에 대한 음 신호를 처리하여 음표정보를 인식하고, 이를 바탕으로 음정곡선을 생성하여 이를 탐색 패턴으로 사용한다. 동요를 대상으로 하여 음악 데이터베이스를 생성할 때는 각각의 곡에서 추출한 음정곡선을 이용하여 색인과 메타 데이터를 생성하는데, 색인은 유사 탐색을 용이하게 하기 위해서 Trie 구조를 사용한다. 탐색 패턴을 이용해서 메타 데이터를 탐색할 때는 동적 프로그래밍 방법을 이용하여 유사 탐색을 함으로써 노래의 어느 마디를 부르더라도 쉽게 후보곡을 검색할 수 있도록 한다. 통상 음악의 첫째 동기를 부르는 경우가 많기 때문에 첫째 악절로 구성된 색인을 먼저 탐색토록 하고, 색인에서 탐색을 실패한 경우(음악의 첫째 동기 이후를 부른 경우)에 메타 데이터를 이용한 유사 음표열 탐색을 하도록 하여 효율적인 검색이 되도록 하였다.

## 1. 서론

오늘날 디지털 도서관에 대한 흥미와 관심이 매우 고조되어 국내외의 많은 연구자와 연구기관에서 대규모 지원 프로젝트를 중심으로 연구가 진행되고 있다. 아울러서 단일 주제로서의 국제 학술회의와 전산학 분야는 물론 정보관리학 분야에서의 활발한 학술활동과 새로운 전문잡지의 발간, 그리고 인터넷에서의 많은 활동 등이 두드러진 현상이다. 디지털 도서관은 인쇄자료와 음성, 화상, 동영상으로 된 기존의 여러 매체 전자정보와 복합된 전자정보뿐만 아니라 네트워크 기반의 전자정보를 포함하고 있다. 이것이 내포하는 의미는 이들을 전자미디어에 축적, 조직, 색인하고 검색할 수 있는 새로운 정보관리 기술과 원격지에 송수신하여 유통, 공유시킬 수 있는 통신기술의 결합을 강조하고 있다[Fox].

벨기에의 음악학자 페티스는 “음악이란 음의 배합에 의하여 사람의 감정을 감동시키는 예술이다”라고 정의하고 있다[백병동]. 음악 정보의 검색도 디지털 도서관의 구축과 관련되어 대단히 흥미있는 분야중의 하나로 등장하면서 도전하고 있는 문제이기도 하다. 악보는 전통적으로 제목, 작곡자, 주제분류 등에 의해 목록화되어 있는데, 사서들은 몇 소절의 노래나 콧노래를 기준으로 해당 음악을 찾아 줄 것을 요청 받기도 한다[Smith]. 그런데 디지털 도서관이 본격화 되면서 보관 중인 음악 정보의 양이 많아지면 더욱 힘든 일이 될 것이다.

그래서 본 논문은 미래의 디지털 음악 도서관의 중요한 구성요소로 정형화 될 수 있는 것으로서 몇 악구를 노래한 것을 기준으로 원하는 음악을 쉽고 빠르게 탐색할 수 있는 시스템 구현을 위해, 노래한 선율을 입력으로 받아서 음표 정보로 자동 변환하기 위한 신호 처리에 대해 기술한다. 그리고 추출된 음표 정보를 바탕으로 효율적인 음악정보 데이터베이스를 탐색하기 위한 검색 시스템을 제안한다.

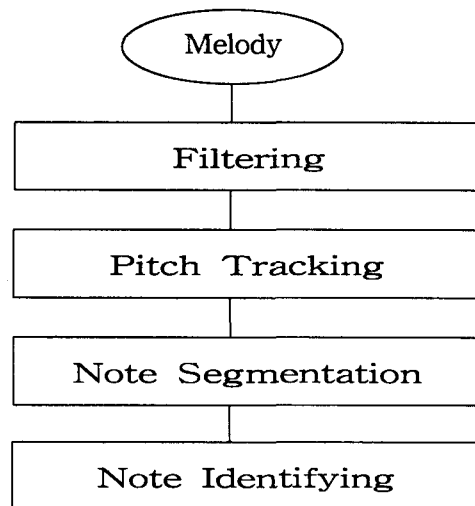
선율 검색 시스템은 적용 범위, 재현율, 정도율, 사용자 노력, 응답시간, 출력 형태 등에 따라서 유용성이 좌우 되는데[Baker], 전통적인 음악 검색 방법은 음악 주제 사전을 찾는 것으로서, “A Dictionary of Opera and Song Themes”와 “The Directory of Tunes and Musical Themes” 등이 그 예다[McNab, 1996b]. 선율 검색에 대한 많은 작업이 이루어졌지만 주로 시스템의 요소에 관계된 것들로서, 반복주기 성분 검출, 악보화, 정확 또는 비정확 문자열 탐색[Landau], 그리고 멀티미디어 장비를 이용한 노래 연주 시스템[Loeb]이 속한다. 그리고 비교적 완전한 선율 검색 시스템의 예는 [Ghias]의 것을 들 수 있는데, 저역 여과기와 자기상관계수법을 이용하여 반복주기 성분을 검출하고, 음 곡선을 추출했다. 그런데 사용자가 노래할 때 음표 사이에 작은 여백을 두도록 하였으며, 리듬은 전혀 계산치 않았다. 그리고 음표의 장단이나 음표의 시작과 끝을 식별하는 절차에 대해서도 언급하지 않았다. [McNab, 1996b, 1997]은 선율의 실시간 인식에 중점을 두고 악보화하는 연구를 했는데, 음악정보를 악보로 하여 MIDI 음표에 의해 음표 비교를 하도록 했다.

본 논문의 구성을 보면, 2장에서 선율이 입력된 음향정보를 분석하여 음표정보를 식별

하는 절차에 대해 기술하고, 인식 방법을 제안한다. 3장에서는 음악정보 검색을 위한 음표열 탐색에 대해서 탐색 기준과 유사 음표열 탐색 방법을 중심으로 기술하고, 4장에서는 음악정보를 효율적으로 검색하기 위해 설계 제안하는 음악 데이터베이스 검색 시스템에 대해 기술한다. 마지막으로 5장에서는 본 연구의 결론과 앞으로의 연구 방향에 대해 기술한다.

## 2. 선율의 인식

음 신호의 인식은 아날로그 신호를 획득하여 디지털 형태로 변환하고, 원하지 않는 주파수들을 제거하기 위하여 여과한 다음 주파수를 식별하게 된다. 식별된 주파수를 분석하여 음표 정보를 인식하게 되는데, 선율 인식을 위한 절차는 <그림 1>과 같다.



<그림 1> 선율 인식의 절차

### 가. 음계와 음표

음계는 어떤 어떤 음을 기점으로 하여 1옥타브 위의 같은 이름의 음에 도달할 때까지 특정된 질서에 의해서 배열된 음렬이고, 옥타브는 음의 높이는 다르나 어떤 음에서 위.아래로 같은 이름의 음인 8번째 음과의 간격으로서[백병동], 두 음높이간의 간격은 2배로 인식된다. 즉, C3은 261.6Hz 주파수를 갖고, 한 옥타브 위인 C4는 523.2Hz, 그리고 한 옥타브 아래인 C2는 130.8Hz이다. 서양 음악에서 바하시대 이후의 음계는 한 옥타브를 12개의 같은 간격의 반음으로 나눈 평균율이다[Martin]. 반음은 서양음악에서 음높이 주기의 가장 작은 단위이지만, 반음의 100분의 1인 센트(cent)로 나누어 사용하기도 하는데, 그러면 1옥타브는 1200센트이다[Backus]. 이처럼 음높이는 옥타브, 반음, 센트 등으로 구분하여 인식할 수 있는데 비해 주파수는 연속적이기 때문에 주어진 주파수에 반복 주기를 할당하는 것은 양자화에 관련된다.

옥타브, 센트 등과 같은 음악적 단위는 상대적인 측정치이기 때문에 센트간의 간격으로 음정을 계산할 수 있으며, 각 음표는 가장 가까운 반음으로 지정되는데, 그것은 음표를 좀 더 세밀하게 표현하기 위해서이다. 따라서 각 음표는 MIDI 음표 0인 8.176Hz 이상의 센트에서 음표의 간격으로 나타낸다. MIDI는 표준 전자음악 악기로서 표준 서양 음계를 표현하는 기능이 있는데, 음계의 각 음표에 상수를 할당한다. C3는 60을 할당하고, 바로 위 음표인 C#3는 61, 그리고 그 아래인 B2는 59를 할당한다. 평균율로 조율된 음계상의 음표들은 A-440에 상대적으로 100의 배수 센트로 표시한다. 즉, C3는 6000센트이고, A3는 4400센트이다[Backus].

## 나. 반복 주기 성분의 검출

음 신호는 1000Hz와 20~40dB에서 여과하고, 22.05KHz로 표본화한 다음 8비트 선형표현으로 양자화 한다. 입력 신호 중에서 기본 주파수의 정수배로 발생하는 배음은 기본음 검출을 혼란스럽게 하므로 가능하면 기본 주파수는 유지한 채 많은 배음을 제거하기 위해 여과한다. 노래하는 음성의 적절한 범위는 보표에 의해 정의되며, 보표의 범위는 F2(87.31Hz)부터 G5(784Hz)까지 이다[McNab, 1996a].

반복 주기 성분의 결정은 신호 처리에서 일반적인 연산이지만 매우 어려운 작업이다. 반복 주기 성분을 검출하는 알고리즘은 대체로 3가지 유형으로 분류되는데, 파형의 표본화를 이용하는 시간 영역 검출과 진폭이나 위상 주파수표를 이용하는 주파수 영역, 그리고 2차 진폭 주파수표를 이용하는 역여과 영역 등이다. 이 중에서 연산 시간의 효율성이 좋은 시간 영역 방법의 하나인 평균차 함수법을 이용하는 것으로 한다.

### (1) 평균차 함수법

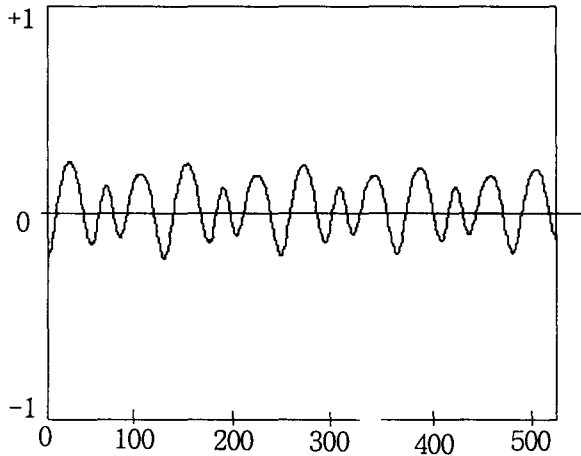
음은 주기인 반복 주기 성분을 가지고 있고, 음의 파형은 분절이나 반복 주기 성분의 반복으로 되어 있다[박경범]. 따라서 음 정보를 추출하기 위해서는 파형의 반복 구조를 발견해 내는 것이 중요하다. 음악에서 반복 주기 성분의 주기는 통상 주기가 시작되는 오름 진폭의 마루에서부터 이므로, 이 마루는 반복 주기 성분 검출기에 의해 사용되는 파형 성분이다. 반복 주기 성분의 검출을 위해 사용하는 평균차 함수는 다음 식으로 정의된다.

$$R(k) = \sum_{n=1}^{m-k} |x(n) - x(n+k)|$$

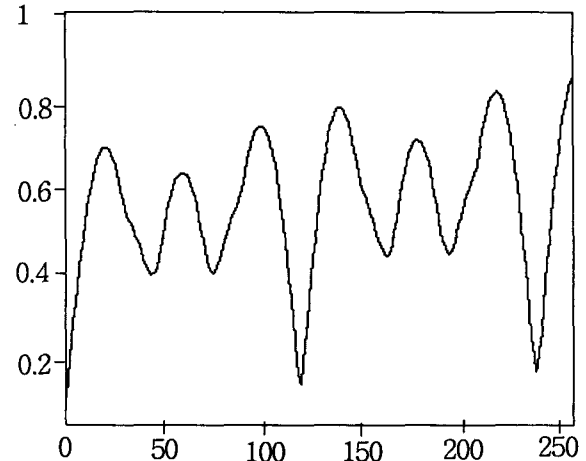
m을 512로 하고, 계수 신호 표본 x(n)의 주변에서 서로 k만큼 떨어져 있는 두 신호값의 차들을 구하여 합한 것으로서, 이 값이 작을수록 k만큼 떨어져 있는 두 신호는 서로 유사하다는 의미이다. 따라서 반복 주기를 P라 할 때 k = 0, P, 2P, 3P, ... 등의 위치에서 R(k)는 극소값을 가진다. R(k)의 값이 20 이하이면 안정된 음 정보로 볼 수 없으므로 제외한다. <그림 2>는 음이름 “솔”음의 파형에서 512Hz를 표본화한 것을 나타낸다.

22.05KHz 표본화 주파수중 512Hz 단위로 구간화하여 연산하는 것은 좀 더 많은 주파

수를 단위 연산에 포함하므로써 반복 주기 성분 검출의 정확성과 안정성을 위해서이다. 이 파형에서 반복 주기 검출을 위해 평균차 함수를 이용하여 계산한 결과 반복 주기값으로 117이 나왔는데, <그림 3>은 <그림 2>의 파형 중  $\frac{1}{2}$ 만 표본으로 하여  $R(0)$ 부터  $R(256)$ 까지 표시되어 있다.



<그림 2> “솔”음 파형의 예

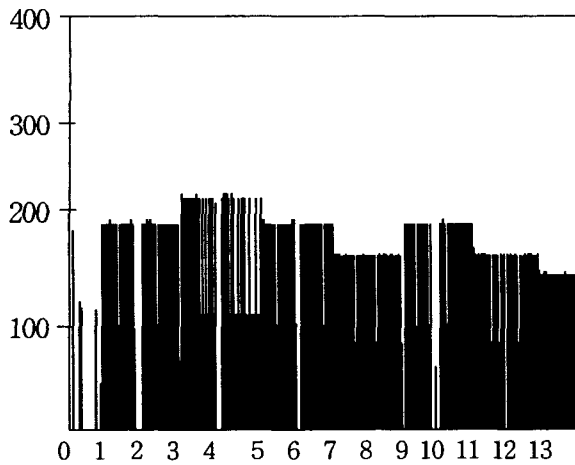


<그림 3> 평균차 함수에 의한 주기 검출

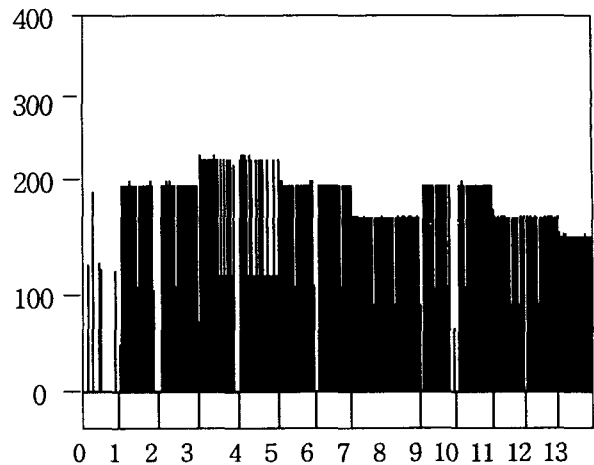
## (2) 후처리

완전한 결과를 돌려주는 반복 주기 성분 검출기는 없다. <그림 4>는 반복 주기 성분 검출기에 의해 출력된 음원 정보를 나타내는데, 모든 표본 위치에서 반복 주기 성분의 계산으로 이루어진다.

일반적으로 음의 시작과 끝, 또는 음표 조옮김의 경우에는 신호가 불안정하게 나타나서 오류가 발생하기 쉽다. 그리고 반복 주기 성분 추출을 위한 계산 결과가 실제 주파수의 2배 혹은 반으로 나타날 때는 가장 일반적인 옥타브 오류가 발생하는데, 이같은 원인은 반복 주파수의 시작이 첫 번째 배음으로 되었거나, 반복 주기 성분 길이 계산이 실패했을 경우이다. 본 연구에서는 오류로 인한 오인식을 최소화하기 위하여 후처리를 행함으로써 파형의 안정성 지역을 확장하고, 옥타브 오류를 바르게 하며, 음표 인식 전에 오류를 보완하여 음원 정보의 구별을 명확히 하고 있다. <그림 4>를 보면 음표간의 경계가 불명확한 부분이 있다. 그래서 앞에서 검출된 반복 주기 성분을 이용하여 반복 주기별로 표본 수열의 합을 구하여 다음 반복 구간의 합과 비교한다. 하나의 음이 끝나갈 때는 음파가 0에 가까워 지다가 새로운 음이 시작되면 다시 급격히 높아지는 파형의 성질을 이용하여 다음의 표본 수열의 합이 바로 전의 표본 수열의 합보다 확연히 커지면 다음 음표의 시작으로 간주해서 구별 한다. 이와 같이 표본 수열의 계산에 의해 음표간의 경계를 구분한 것이 <그림 5>이다.



<그림 4> 음원정보의 출력



<그림 5> 음표간의 경계 인식

#### 다. 음표의 인식

반복 주기 성분에 의한 방법을 이용하여 음표를 분할하는데, 이는 검출된 반복 주기값을 이용하여 반복 프레임을 집단화하고 이들을 비교 연산함으로써 후처리된 음원 정보로부터 직접 음표를 분할하는 것이다.

분할된 음표 구간 내에는 유사한 주파수 값이 다수 놓이게 된다. 이 중에서 하나의 주파수를 결정하는 방법에는 평균을 구하는 법과, 빈도수에 의한 법, 그리고 통상적으로 신호의 안정 단계로 볼 수 있는 3~4 번째의 값을 취하는 방법 등이 있는데, 본 연구에서는 음 높이가 낮아질 수 있는 경우를 대비하여 빈도수에 의한 방법을 택하였다.

음표에 대한 주파수가 결정되면 해당 주파수의 음높이를 할당해야 한다. 음높이의 할당은 한 옥타브가 12개의 반음으로 구성되므로, 으뜸음을  $Cf$ 라 하고 다음 으뜸음을  $C'f$ 라 하면  $Cf$ 와  $C'f$  사이에 12개의 반음이 놓이게 된다. 따라서  $C'f = 2Cf$  이므로 다음 식에 의해 단위음을 구한다.

$$\log_2 C'f = \log_2 Cf + \frac{k}{12}$$

$$C'f = 2^{\log_2 Cf + \frac{k}{12}}$$

$$k = 12 \times (\log_2 C'f - \log_2 Cf)$$

$k$ 의 값은 0에서부터 12까지로 옥타브내의 단위음을 나타내는데,  $C0(0)$ ,  $C\#(1)$ ,  $D(2)$ , ...,  $A\#(10)$ ,  $B(11)$ ,  $C1(12)$  등이다. 위 식에 의해 으뜸음의 주파수만 결정되면 나머지 음의 음높이를 구할 수 있다.

### 3. 음표열 탐색

음표열의 탐색은 정확성과 빠른 탐색 시간이 요구되는데, 먼저 정확한 탐색을 위해서는 탐색코자하는 패턴이 정확해야 한다. 그런데 본 논문에서 다루고자 하는 패턴은 정확성이 낮다고 보아야 한다. 왜냐하면 대부분 전문 음악교육을 전혀 받지 않은 사용자들이 악보도 없이 패턴을 생성해야 하기 때문이다. 따라서 노래하는 형태가 다양하고 원래의 음악에 일치한다는 것이 어렵기 때문에 정확성 탐색을 하게 되면 검색률은 형편없을 것이다. 이와같이 음표열 탐색을 위한 탐색 패턴의 환경이 다양하기 때문에 본 시스템에서는 유연성이 효율성과 깊은 관련이 있다. 따라서 탐색의 유연성을 높이기 위해서 음표열의 정확성 탐색보다 유사성 탐색 방법을 택한다.

#### 가. 탐색 기준

입력된 선율의 조(key)와 관계없이 인식이 가능토록 하여 사용자들이 구애받지 않도록 해야 한다. 이를 위해서 음정을 기준으로 탐색하도록 한다. 그런데 사용자에게 따라 음정에 대한 정확도가 다양하기 때문에 음정의 크기와 관계없이 방향을 이용하는 것이 중요하다 [Dowling]. [Parsons]은 선율의 방향을 기준으로 선율 윤곽선이라는 선율 색인을 만들었는데, 첫번째 음표를 \*로, 내림 음정을 D로, 올림 음정을 U로, 그리고 반복을 R로 표시했다. 예를 들어 동요 “비행기”는 “\*DDUURRDRRUURDDDUURRDRUDD”로 표현한다. 이러한 방법은 사용자들이 정확한 음정으로 노래하지 않아도 된다는 장점이 있지만, 표현 기호가 단순하여 레코드의 키로 할 때 중복의 문제가 발생한다. 따라서 본 논문에서는 음계에서 음정의 크기에 관계없이 위치만으로 음의 윤곽선을 만드는 것이 아니라 두 음간의 음정을 계산하여 그 결과값에 의한 음정곡선을 만든다.

#### 나. 유사 음표열의 탐색

문자열 탐색 문제는 매우 일반적인 문제로서, 커다란 텍스트 파일  $T=t_1t_2\cdots t_n$  내에서 문자열  $P=p_1p_2\cdots p_m$ 을 찾는 것이다[Wu]. 이 문제에 대한 가장 유명한 알고리즘은 Knuth Morris Pratt 알고리즘[Knuth]과 Boyer-Moore 알고리즘[Boyer]이다. 이들 알고리즘을 중심으로 여러 변형 알고리즘들이 발표되고, 정확성 대조로 시작했으나 유사성 대조 기능을 포함하기도 한다. 사용자들이 크게 제약에 얽매이지 않고 쉽게 탐색 패턴을 만들고 이를 토대로 후보곡을 찾는 것이기 때문에 정확성 탐색 보다는 유사성 탐색이 바람직할 것으로 본다. 그런데 유사 음표열의 탐색은 Trie 탐색 알고리즘을 이용하거나 동적 프로그래밍을 적용하면 효율적으로 탐색할 수 있다. Trie를 이용한 유사 음표열 탐색은 다음과 같이 Trie를 이용한 철자 정정 알고리즘[Muth]을 사용하는데, 일반적으로 두 음표열의 탐색을 위한 연산은 치환, 삽입, 삭제 등이 있다.

```

WHILE i ≤ m DO
  CASE pat(i) : trie-lev(j)
    WHEN pat(i) = trie-lev(j) i=i+1, j=j+1;
    WHEN error(substitution) i=i+1, j=j+1;
    WHEN error(deletion) j=j+1;
    WHEN error(insertion) i=i+1;

```

동적 프로그래밍 알고리즘은 더 큰 부프로그램을 해결하기 위해 계산했던 것들이 필요하다면 그것들을 다시 계산하지 않고 한 단위씩 저장하여 놓고 거슬러 올라감으로써 해결한다. 그러한 동적 프로그래밍 알고리즘은 재귀 알고리즘이 쓰일 수 있는 문제에 적절하다 [Cormen]. 패턴의 길이를  $m$ , 문자열의 길이를  $n$ 이라 하고,  $0 \leq i \leq m$ 과  $0 \leq j \leq n$ 일 때  $D_{ij}$ 는  $pat_1 \dots pat_i$ 와  $str_j$ 에서 끝나는  $str$ 의 모든 요소 문자열 사이의 최소 편집 간격을 갖는다. 탐색을 위한 초기 조건은 다음과 같다.

$$D_{0j} = 0, \quad 0 \leq j \leq n \quad \dots\dots\dots ①$$

$$D_{i0} = D_{i-1,0} + 1, \quad 1 \leq i \leq m \quad \dots\dots\dots ②$$

①은 문자열의 어디에서도 탐색을 시작할 수 있는 것으로, 패턴의 첫번째 기호 앞에 어떤 문자를 삽입하기 위한 것이다. ②는 문자열의 첫 번째 기호 앞에 있는 모든 패턴을 삭제하기 위한 것이다. 나머지 값들은 ③의 재귀 관계를 이용하여 계산한다. ①과 ②는 패턴이 문자열의 어디에서나 발생할 수 있기 때문에 요소 문자열의 탐색을 허용한다.

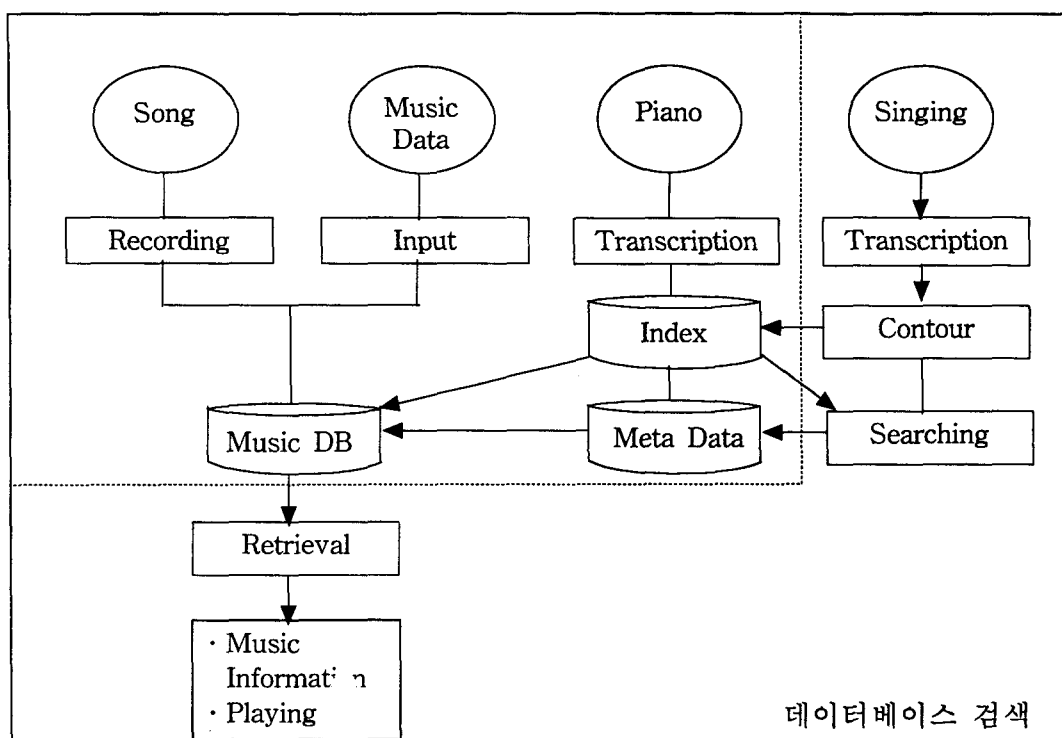
$$D_{ij} = \min \begin{cases} D_{i-1,j-1} + \delta_{ij} & \text{(replacement)} \\ D_{i,j-1} + 1 & \text{(insertion)} \\ D_{i-1,j} + 1 & \text{(deletion)} \end{cases} \quad \dots\dots\dots ③$$

③에서 만일  $pat_i = str_j$ 이면  $\delta$ 는 0이고, 그렇지 않으면 1이다.

#### 4. 음악 데이터베이스 검색 시스템

설계 제안하는 음악 데이터베이스 시스템은 크게 데이터베이스 생성부와 데이터베이스 검색부로 나누어 지며, 검색의 유연성과 효율성을 높이기 위해서 음표열에 의한 유사성 탐색을 한다. 그리고 탐색 시간을 보다 빠르게 하기 위하여 색인과 메타 데이터 탐색으로 구분한다.





<그림 6> 음악 데이터베이스 검색 시스템의 구성

### 가. 시스템의 구성

음악 데이터베이스 검색 시스템의 구성을 보면 <그림 6>과 같다.

음악 데이터베이스 검색 시스템은 크게 데이터베이스 생성부와 데이터베이스 검색부로 이루어 지는데, 데이터베이스 생성부는 음악 데이터베이스와 메타 데이터, 그리고 색인으로 구성된다. 반면에 데이터베이스 검색부는 탐색과 검색, 그리고 검색된 음악 정보의 출력이나 연주 등으로 구성된다.

### 나. 데이터베이스의 생성

음악 데이터베이스를 생성하는데, 탐색을 효율적으로 하기 위해서 메타 데이터와 색인을 추가로 생성한다.

#### (1) 음악 데이터베이스

먼저 음악 데이터베이스에 수록 할 대상 동요를 CD 연주기로 연주하면 음악 편집기인 GoldWave로 녹음하여 .WAV 파일을 생성한다. 그리고 해당 동요에 대한 제목, 작곡자, 첫째 악절의 가사 등을 입력하여 .WAV 파일과 함께 <그림 7>의 스키마 형식으로 음악 데이터베이스를 생성한다.

순번	제목	작곡자	첫째 악절의 가사	전 곡
----	----	-----	-----------	-----

<그림 7> 음악 데이터베이스 스키마

## (2) 메타 데이터와 색인

메타 데이터와 색인은 음악 데이터베이스를 효율적으로 검색하기 위해서 생성한다. 음반에 저장된 동요곡은 다중음으로 되어있기 때문에 피아노로 단음 연주한 다음, 이를 음표 인식기를 통해 음표정보로 인식하여 만들어진 음정 곡선을 가지고 메타 데이터를 구성한다. 그리고 메타 데이터 중 첫 악절에 해당하는 부분을 가지고 색인을 구성하는데, 색인의 구조는 유사 키 탐색을 용이하게 하기 위해서 TRIE 구조로 한다. 음정 곡선은 인식된 음표들을 차례로 한 다음, 두 음표 간의 차를 음계에 의해 구한다. 그 차를 반음으로 구분하여 양의 반음 수만큼은 A에서부터 M까지 차례로 할당하고, 음의 반음 수만큼은 O에서부터 Z까지 할당하며, 음정의 차가 0이면 N을 할당하여 차례로 놓는다.

### 다. 데이터베이스의 검색과 출력

사용자가 검색을 원하는 노래 중 알고 있거나 기억나는 부분을 마이크를 통해 노래한다. 그러면 음표 인식기의 선율 인식 절차를 거쳐 음표 정보가 인식되고, 이를 이용하여 음정 곡선을 만들어 탐색 패턴을 구성한다. 구성된 탐색 패턴을 가지고 먼저 색인의 유사 탐색을 시도한다. 탐색이 성공하면 키 정보를 이용하여 바로 음악 데이터베이스를 검색하면 되고, 색인에서 탐색이 실패 했으면 메타 데이터를 기준 음표열로 하여 탐색 패턴에 의한 유사 음표열 탐색을 시도한다. 여기서 유사 음표열이 탐색되면 해당 곡의 키 정보를 이용하여 음악 데이터베이스를 검색한다.

음악 데이터베이스로부터 검색된 음악정보 중 제목, 작곡자, 첫째 악절의 가사, 정확률 등을 정확률의 내림차순으로 화면에 출력한다. 사용자는 화면에 출력된 결과를 토대로 검색하고자 하는 음악을 선택하여 볼 수 있고, 연주가 되게 하여 들을 수 있다.

## 5. 결론 및 향후 연구 방향

본 논문은 음향 입력으로부터 음 신호를 처리한 다음 선율을 인식하고, 인식된 음정 곡선을 탐색 패턴으로 하여 적절한 음악 정보를 효율적으로 검색하기 위해 설계한 시스템을 제안했다. 마이크를 통해서 노래한 음악의 일부분에 해당하는 음 신호를 분석하여 대응되는 음표를 인식하기 위해서 원래의 아나로그 신호는 표본율을 22.05KHz로 하여 양자화하고, 저역 통과 대역과 통과 대역 여과기를 이용하여 1000Hz 이상의 배음과 20~40dB 범위 밖의 음을 잡음으로 간주하여 골라 내었다. 그리고 반복 주기 성분을 검출하기 위

하여 시간 영역 중심의 평균차 함수를 이용하였다. 반복 주기표로부터 출력된 음원 정보로부터 음표를 분할하기 위하여 반복 주기 성분을 적용하는데, 표본화 구간별로 표본 수 열의 합을 구하여 구간 값을 비교한다. 분할된 음표를 식별하기 위하여 동일한 음표 구간 내에서 발생한 다수의 주파수 값 중에서 빈도수가 가장 많은 값을 해당 음표의 주파수로 결정한다.

식별된 음표들간의 음정을 계산하여 음정 곡선을 구한 다음 이를 탐색 패턴으로 이용한다. 피아노로 단음 연주한 곡을 자동 인식하여 메타 데이터를 생성하고, 첫째 동기를 이용하여 Trie 구조의 색인을 구성한다. 탐색 패턴을 가지고 색인을 참조하여 곡을 검색한 다음 검색이 실패했을 경우 메타 데이터를 다시 탐색한다. 색인이나 메타 데이터에서 탐색된 키값을 기준으로 음악 데이터베이스로부터 음악정보가 검색되어 그 결과가 출력되고, 선택에 따라 연주가 이루어 진다. 색인이나 메타 데이터 탐색시에는 유사성 탐색을 하여 불완전한 입력 음 신호로부터 효율적인 검색이 이루어 지도록 했다.

이 시스템은 디지털 음악 도서관에서 감상을 필요로 하는 음악을 검색하는 시스템으로 활용하기 위한 것이다. 통상적인 음악 정보 검색 항목인 제목, 작곡자, 주제분야 등에 대한 정보를 알지 못한 상태에서 단지 몇 소절의 노래만 부름으로써 원하는 음악 정보를 검색할 수 있도록 하는 시스템으로서, 수 많은 음악을 저장해 둔 대규모 음악 데이터베이스로부터 사용자는 쉽고 빠르게 음악 감상을 위한 곡의 검색을 할 수 있을 것이다.

현재는 음표의 높이와 길이의 정밀도를 보장하는 작업과 함께 검색 알고리즘의 실험이 병행되고 있는데, 전체적인 시스템의 구현이 마무리 되어 다양한 평가 결과를 분석하는 것이 우선 과제이다. 그리고 선율 인식의 실시간 처리와 복잡음으로 된 선율의 인식에 대한 연구도 계속 되어야 할 것으로 생각된다.

## [References]

Backus, J., *The Acoustical Foundations of Music*, John Murray, 1969.

Baker, S.L., Lancaster, F.W., *The Measurement and Evaluation of Library Services*, Information Resources Press, 1991.

Boyer, R.S., Moore, J.S., "A Fast String Searching Algorithm", *Communication of the ACM*, Vol. 20, No. 10, PP. 762-772, 1977.

Cormen, T.H., Leiserson, C.E. and Rivest, R.L., *Introduction to Algorithms*, MIT Press, 1990.

Dowling, W.J., "Scale and Contour : Two Components of a Theory of Memory for Melodies", *Psychological Review* 85: 341-354, 1974.

Fenske, D.E., "The VARIATIONS Project at Indiana University's Music Library", *D-Lib Magazine*, June 1996.

Fox, E.A., "Digital Libraries", *Communication of The ACM*, Vol.38, No.4, Apr.1995.

Ghias, A., Logan, J., Chamberlin, D. and Smith, B.C., "Query by Humming: Musical Information Retrieval in an Audio Database", *Proc. ACM Multimedia '95*, 1995.

Harter, S.P., "What is a Digital Library? Definitions, Content, and Issues", *Proceedings of the ICDL*, 1996.

Knuth, D.E., Morris, J.H. and Pratt, V.R., "Fast Pattern Matching in Strings", *SIAM Journal of Computing*, Vol. 6, No. 2, PP. 323-350, 1977.

Landau, G.M., Vishkin, U., "Efficient String Matching with K Mismatches", *Theoretical Computer Science*, Vol. 43, PP. 239-249, 1986.

Loeb, S., "Architecting Personalized Delivery of Multimedia Information", *Communications of the ACM*, Vol. 35, No. 12, PP. 39-48, 1992.

Martin, D.W., "Musical Scales since Pythagoras", *Sound*, Vol. 1, No. 3, PP.22-24, 1962.

McNab, R.J., Smith, L.A., Witten, I.H. "Signal Processing for Melody Transcription", *Proc Australasian Computer Science Conference*, PP.301-307, 1996a.

McNab, R.J., *Interactive Applications of Music Transcription*, Thesis of Master, University of Waikato, New Zealand, 1996b.

McNab, R.J., Smith, L.A., Bainbridge, D. and Witten, I.H., "The New Zealand Digital Library MELody inDEX", *D-Lib Magazine*, May 1997.

Muth, Frank, E., and Tharp, A.L., "Correcting Human Error in Alphanumeric Terminal Input", Information Processing and Management, Vol. 11, PP. 329-337, 1977.

Parsons, D., The Directory of Tunes and Musical Themes, Spenser Brown, 1975.

Smith, L.A., McNab, R.J. and Witten, I.H., "Music Information Retrieval Using Audio Input", AAAI Symposium, 1996.

Steiglitz, K., A Digital Signal Processing Primer, Addison-Wesley Publishing, 1996.

Wu, S., Mander, U., "Fast Text Searching Allowing Errors", Communication of the ACM, Vol. 35, No. 10, PP. 83-91, 1992.

박경범, 음성의 분석 및 합성 가 그 응용, 도서출판 그린, 1997.

백병동, 개정 대학음악이론, 현대음악출판사, 1997.