MPQF 비동기 검색 인터페이스를 이용한 허밍 검색 시스템의 설계 및 구현

*이종설 **장세진 ***이석필

전자부품연구원

*leejs@keti.re.kr

Design and Implementation of Query-by-humming system using MPQF asynchronous search interface

*Lee, Jong-Seol **Jang, Se-Jin ***Lee, Seok-Pil

Korea Electronics Technology Institute

요약

반도체 기술 및 방송통신기술의 발달에 따라 소비자는 복잡하고 다양한 서비스 환경을 접하고 있으며, 개인당 소비되는 콘텐츠의 양은 기하급수적으로 증가 되고 있다. 스마트폰과 같은 멀티미디어기기의 발전으로 기존의 단순 시청에서 벗어나 멀티미디어 검색과 추천과 같은 다양한 서비스들이 제공되고 있으며, 웹과 인터넷을 기반으로 하는 정보기술은 모바일 전자기술의 발달에 따라 데스크톱이라는 공간적인 한계를 벗어나 유기적인 유비쿼터스 환경으로 변화하고 있다. 이와 같이 소비자들이 접하는 멀티미디어의 양이 증가됨에 따라, 멀티미디어를 검색하기 위한 다양 서비스들이 사업자들을 통해 제공되고 있다. 그러나 사업자들이 구축한 서버들은 공유되어지지 못하며, 공유되더라도 공유하기 위한 방법이 제각각이다. 이에 MPEG에서는 멀티미디어 콘텐츠를 검색하기 위한 질의 구조를 표준화 하여, 서비스 제공자 및 소비자들에게 유용한 검색 인터페이스를 제공하였다. 이에 본 논문에서는 MPEG Query Format에 기반 하여 임베디드 환경에서 광대한 양의 음악 콘텐츠에서 사용자가 원하는 허밍 질의를 편리하게 검색하기 위한 방법을 설계 및 구현한다.

1. 서론

MPEG, TV-Anytime 등 다양한 표준에서는 멀티미디어 콘텐츠의 메타데이터 기술을 위한 메타데이터 구조를 정의 하였다[1][2]. 이와 같은 메타데이터 정보들은 서비스 제공자들을 통해 다양한 형태로 생산 및 재생산되었으며, 콘텐츠의 내용에 대한 기술뿐만이 아니라 전송 정보, 부가 정보 등 다양한 정보에 대한 기술자를 확장하였다. 메타데이터기술자를 정의한 가장 큰 목적은 멀티미디어 검색서비스 등을 효율적으로 제공하기 위한 멀티미디어 콘텐츠 표현에 대한 표준을 규정하기 위함 이었다. 즉, 멀티미디어 콘텐츠를 표현하기 위한 기술 인터페이스 (description interface)를 표준화하여 멀티미디어 자료를 효율적으로 저장하고 검색할 수 있도록 하는 것이다. 그러나 서비스 제공자들은 자체적인 메타데이터 데이터베이스를 구축하거나, 또는 표준을 준수하는 데이터베이스를 구축하더라도 이를 배포할 방법이 없는 상황이었다. 이에 MPEG 에서는 ad-hoc 그룹으로 Query Format Working Group이 결성되어 질의 포맷의 표준화 작업을 완료하였다[3]. 이 표준을 "Multimedia content description interface" Part 12: Query format(MPQF)로 확정하였다. MPQF는 클라이언트들이 MPEG-7 기반의 데이터베이스를 활용하기 위한 방안이다. 이를 위해 MPEG-7 데이터베이스와의 표준화된 인터페이스를 제공하여 사용자가 자신의 검색조건과 결과 조건을 정의할 수 있으며 질의 관리툴을 이용하여 자신의 질의를 요청할 수 있다. 이를 통하여 서비스 사업자는 멀티미디어 콘텐츠를 검색하고자 하는 사용자들의 요청을 받아들이고, 검색결과의 리턴 하는데 표준화된 인터페이스를 사용할 수 있다.

2. MPEG Query Format

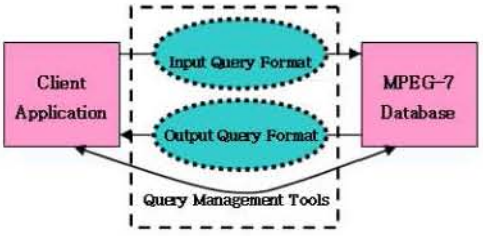

그림 1. MPQF Format

MPEG Query Format의 구성은 3가지로 Input Query Format, Output Query Format, Query Management Tools이다. Input Query Format은 두 가지의 기능을 제공한다. 첫째로 서버와 클라이언트간의 연결과 질의 데이터구조를 위한 인터페이스 구조를 제공한다. 둘째로 질의 결과를 위한 구조를 정의한다. Output Query Format은 서버로부터의 클라이언트를 향한 응답 인터페이스를 정의한다. Query Management Tools는 서버와 클라이언트간의 질의 트랜잭션을 관리하기 위한 기능을 제공한다. TV-Anytime Forum에서의 양방향 메타

데이터 검색 인터페이스는 SOAP과 HTTP 프로토콜에 기반하지만[4], MPEG Query Format은 제한을 두지 않는다.

가. Framework

MPQF는 XML 스키마로 규정되어 있으며 최상위 엘리먼트는 MpegQueryType이다. 그림에서 나타내듯이 질의는 input query 포맷과 output query 포맷으로 나뉘며 질의 관리를 위한 management 엘리먼트는 갖는다.

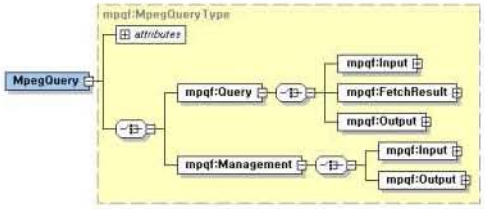

그림 2. Mpeg7Query Element

나. Input Query Format

Input Query 엘리먼트는 InputQueryType를 통해 규정되어 지며, 하나의 질의를 표현하는 콘디션들의 조합과 해당 질의의 결과를 표현하는 방법 등을 갖는다. 즉 질의자는 QueryCondition을 통해 질의할 질문을 표현하고 OutputDescription을 통해 해당 결과를 표현하는 방법을 기술한다.

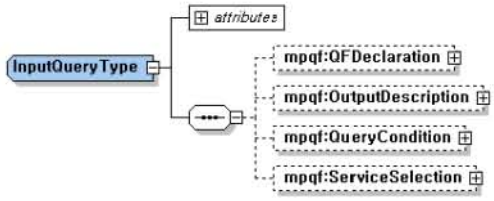

그림 3. InputType type

다. Output Query Format

OutPut Query 엘리먼트는 OutputQueryType를 통해 표현되어지며, 이것은 질의 결과가 표현되어진다. 검색 응답 서버는 검색 결과를 Output Query로 생성해서 질의자에 전달하게 된다.

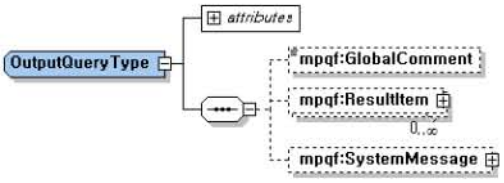

그림 4. OutputType type

라. FetchResult Format

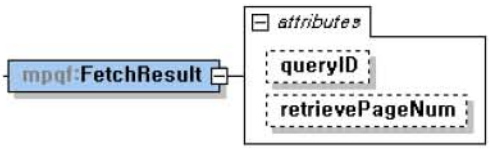

그림 4. FetchResult type

FetchResult 엘리먼트는 사용자로 하여금 이전 질의에 대한 접근성을 부여한다. queryID 에트리뷰트는 접근할 이전 질의의 구분자이다.

3. 비동기 검색 방법

멀티미디어 검색을 위해 기존의 검색 방법에서는 사용자가 서버로의 질의를 요청한 뒤 해당 결과를 대기한다.

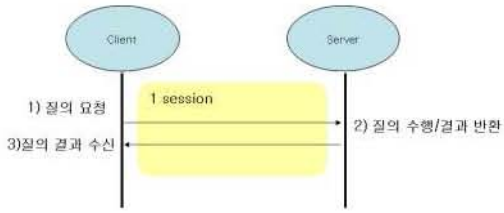

그림 5. 기존의 검색 방법

기존의 검색 방법은 기본적으로 질의를 처리하는 시간이 짧은 경우에는 문제점이 발생하지 않는다. 그러나 이미지, 동영상과 같은 콘텐츠를 기반으로 하는 질의의 경우에는 해당 콘텐츠의 전송, 처리 등에 많은 시간 및 자원이 요구되어 지며, 사용자가 대기하는 시간은 상대적으로 증가한다. 또한, 검색 처리 중에 해당 네트워크 세션이 끊어 질 경우 해당 질의를 최초 질의 요청에서부터 재실행 하여야 한다.

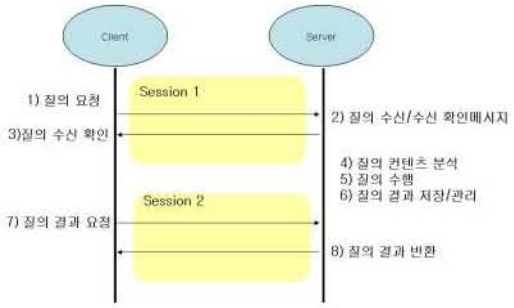

그림 6. 비동기 검색 방법

비동기 검색은 1)질의의 요청 및 질의 저장, 2) 질의 결과 요청으로 구성되어진다. 사용자는 서버로의 질의를 요청하며 서버에서는 해당 질의를 수신완료 한 뒤 수신 확인 메시지를 사용자에게 전송한다. 사용자는 수신확인 메시지를 수신한 뒤 서버와의 연결을 끊는다. 시간이 지난 뒤 사용자는 자신의 예약 질의 결과를 요청하게 되며 서버에서는 해당 질의 결과, 또는 검색 상태를 사용자에게 전송한다.

1)질의의 요청을 위해 InputQueryType의 immediateResponse와 timeOut 에트리뷰트를 활용한다. immediateResponse는 질의의 모드

를 결정하며 기본 값은 "true"이다. timeOut 값동안 질의를 수행하며, 클라이언트는 수행 결과를 대기한다. immediateResponse가 "false"라면 클라이언트는 "mpqfID"를 응답 결과로 전달 받게 된다. 2) 질의 결과 요청을 위해 InputQueryType 에 FetchResult를 조합하여 서버에 요청하게 되며, 이때 mpqfID를 추가하여 서버에 요청을 하게 된다.

```
<mpqf:Query>
    <mpqf:Input immediateResponse="false" timeout="30M">
        <mpqf:QueryCondition>
            ……
        </mpqf:QueryCondition>
    </mpqf:Input>
</mpqf:Query>
```

```
<mpqf:Query>
    <mpqf:FetchResult queryID="??????" retrievePageNum="" />
</mpqf:Query>
```

그림 8. 비동기 검색 질의

멀티미디어 내용검색을 위해서는 여러 가지 질의 방법이 존재한다. 이중 멀티미디어 콘텐츠를 검색 파라미터로 할 경우 해당 콘텐츠의 전송 및 정보 추출 등에 많은 시간을 요구하게 되며 연결 세션이 끊어질 경우 질의를 새로 시작하여야 한다는 단점을 갖는다. MPQF의 FetchResult와 immediateResponse 는 질의의 요청 이후 서버의 검색 수행 중에는 사용자와의 연결을 끊음으로써 세션을 보호 하고 사용자는 지속적인 대기 시간 없이 결과가 예측되는 시점에서 해당 결과를 획득한다.

4. 허밍 검색 시스템

본 논문에서 구현한 허밍 검색 시스템 구성도는 아래 그림 9와 같다. 시스템은 MPQF를 적용하기 위한 전통적인 서버-클라이언트 모델이다. 서버는 음원의 특징정보 추출, 매칭엔진, 추천엔진, 스트리밍서비스, 음원 및 특징 DB의 생성, 관리 및 갱신 등을 수행한다. 각각의 클라이언트에서는 사용자의 허밍을 입력 받고, 입력 받은 허밍신호에서 특징을 추출하고 서버에 질의를 하게 된다. 이렇게 질의된 내용을 기반으로 서버에서는 20개의 후보를 추천하게 되고 클라이언트는 추천된 후보를 사용자에게 출력한다 클라이언트는 기본적으로 10초간의 허밍을 사용자로부터 입력 받게 된다. 이때 입력되는 허밍신호는 8000Hz 샘플링주파수를 가지고 각 샘플 당 2바이트 즉, 16bit 크기의 부호화 있는 신호가 된다. 이렇게 녹음된 신호는 배경 잡음을 제거하고 특징 추출 모듈로 전달된다. 특징 추출 모듈은 프레임단위로 특징을 추출하게 되는데 이때 프레임 단위는 16ms이다. 즉 16ms 마다 하나의 특징을 추출하게 되고 이렇게 추출된 특징은 MPQF 포맷으로 구조화하여 서버로 전송하게 된다. 서버는 전송 받은 MPQF형식의 질의 내용을 기반으로 특징DB에 존재하는 내용과 유사도 판별을 하게 된다. 매칭 엔진은 DWT를 기반으로 하고 있다. 추천엔진은 매칭엔진의 유사도를 기반으로 20개의 최적 후보를 선별하게 된다. 선별된 후보는 클라이언트로 전송되고 사용자가 원하는 곡을 선택하게 되면 서버는 클라이언트에게 스트리밍서비스를 제공한다. 스트리밍서비스는 RTP/RTSP기반의 실시간 스트리밍을 제공한다. 검색 결과를 클라이언트에게 제공하기 위하여 각각의 MP3음원에 ID3를 기반으로 하는 각 음원의 제목, 가수, 장르, 앨범자켓, 그리고 시작부분으로 구성된 5가지 메타데이터 정보를 제공한다. 이는 사용자가 찾는 노래를 쉽게 인지할 수 있도록 제공되는 정보들이다.

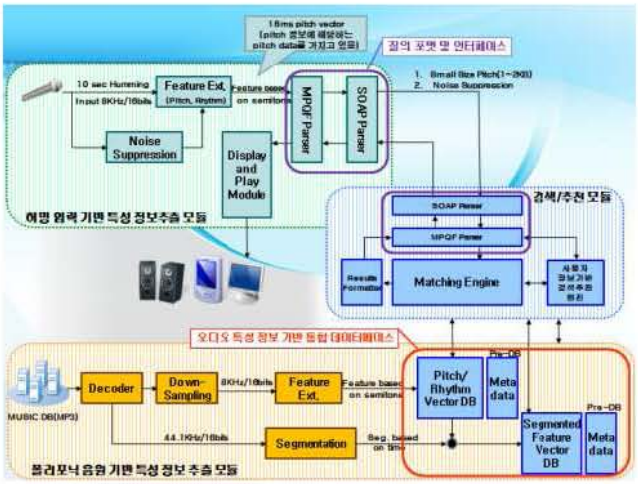

그림 9. 허밍 검색 시스템 구조

그림 10. 로그인 화면

그림 11. 음악 재생 화면

그림 12. 검색 결과 목록

5. 결론

본 논문에서는 MPEG Query Format에 기반 하여 허밍 질의와 같은 멀티미디어 검색 시스템에서 적용이 가능한 비동기 검색 방법에 대해 구현 하였다. 일반적으로 멀티미디어 콘텐츠의 내용에 기반을 둔 검색 엔진의 경우 검색 요소가 상당히 많으며 또한 검색해야 될 콘텐츠가 방대하므로 일반 검색과 같은 즉각적인 검색결과를 요청하기는 상당히 무리가 있다. 특히, 허밍 질의의 경우 모바일 단말을 주로 활용하므로 사용자에게 접속을 유지하고 대기를 유지하는 것은 많은 비용을 초래한다. 본 논문에서는 MPEG Query Format을 통해 비동기 검색을 통한 새로운 검색 인터페이스를 구현하였다.

참고문헌

[1]ISO/IEC 15938-2, Information technology – Multimedia content description interface – Part 2: Description definition language

[2] ETSI TS 102 822-3-1: "Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 3: Metadata; Sub-part 1: Phase 1 - Metadata schemas".

[3] ISO/IEC SC29 WG11 N8780, MPEG-7 Query Format Requirements, January 2007

[4] ETSI, ETSI 102 822-6-1 v1.3.1 (2006-01), Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 6: Delivery of metadata over a bi-directional network Sub-part 1: Service and transport, European Telecommunications Standard Institute, Jan. 2006

[5] W3C, XML Query (XQuery), available at http://www.w3c.org/TR/xquery/, 2006

[6] ISO/IEC JTC 1/SC 29/WG 11 N 4320, MPEG-7 Requirements Document V.15, F. Pereira ed., Sydney,July2001.

[7] ISO/IEC JTC 1/SC 29/WG 11 N 2730, Results of MPEG-7 Technology Proposal Evaluations and Recommendations, MPEG7 Requirements Group, Seoul, March 1999.