

논문 2006-43SP-3-10

지능형 오디오 그래픽 이퀄라이저 시스템 구현

(Implementation of an Intelligent Audio Graphic Equalizer System)

이 강 규*, 조 윤 호*, 박 규 식**

(Kang-Kyu Lee, Youn-Ho Cho, and Kyu-Sik Park)

요 약

오디오 이퀄라이저의 주목적은 사용자가 원하는 음색을 낼 수 있도록 음악의 주파수 특성을 변환 하는 것으로 대형 오디오 시스템으로부터 휴대용 MP3플레이어까지 음악을 재생하는 대부분의 기기에 사용되고 있다. 현재까지는 각 음악 장르에 적합한 음색을 재생하기 위해 사용자가 각 이퀄라이저 주파수 대역 별로 일일이 수동으로 조절해 주어야 한다는 불편함이 있다. 본 논문에서는 내용기반 음악 장르 분류 기술을 이용하여 음악 장르를 분류하고, 분류된 장르에 알맞은 이퀄라이저를 자동으로 적용하여 주는 지능형 오디오 그래픽 이퀄라이저 시스템을 제안하였다. 이퀄라이저의 경우 초기 음악 재생 중 이퀄라이저 적용에 의한 갑작스런 음의 변화를 방지하기 위하여 Coarse Level과 Fine Level의 2단계 장르분류 절차를 거치면서 각 단계별 이퀄라이저를 적용 하였다. 제안된 시스템의 각 단계별 분류 성공률은 약 80%에 이르며, 2초 안에 모든 동작이 이루어지는 것을 확인 하였다. 제안된 시스템은 PC상에서 비주얼 C++ 을 이용하여 3-밴드 지능형 이퀄라이저의 SW GUI를 구현 하였다.

Abstract

A main objective of audio equalizer is for user to tailor acoustic frequency response to increase sound comfort and example applications of audio equalizer includes large-scale audio system to portable audio such as mobile MP3 player. Up to now, all the audio equalizer requires manual setting to equalize frequency bands to create suitable sound quality for each genre of music. In this paper, we propose an intelligent audio graphic equalizer system that automatically classifies the music genre using music content analysis and then the music sound is boosted with the given frequency gains according to the classified musical genre when playback. In order to reproduce comfort sound, the musical genre is determined based on two-step hierarchical algorithm - coarse-level and fine-level classification. It can prevent annoying sound reproduction due to the sudden change of the equalizer gains at the beginning of the music playback. Each stage of the music classification experiments shows at least 80% of success with complete genre classification and equalizer operation within 2 sec. Simple S/W graphical user interface of 3-band automatic equalizer is implemented using visual C++ on personal computer.

Keywords: 이퀄라이저, 내용기반 장르분류, Equalizer, classifier, content-based

I. 서 론

그래픽 이퀄라이저의 주 기능은 음악의 주파수 밴드를 다양하게 조절하여 사용자가 선호하는 음색의 음악을 만들어 주는 것이다. 디지털 오디오 시스템의 경우,

이퀄라이저의 구현이 간편하여 대부분의 디지털 오디오 기기에 탑재되고 있으며, 대표적인 예로서 우리주변에서 흔히 볼 수 있는 MP3플레이어, CD오디오, 카오디오 등이 있다.

음악을 듣는 사람마다 선호하는 음색이 다르며, 음악 장르에 따라서도 각 장르에 맞는 음색이 있다. 그러나 이퀄라이저를 이용하여 원하는 음색을 만들어 주기 위해서는 적게는 3개에서 많게는 20개가 넘는 버튼을 조작해 주어야 하며, 이러한 조작은 일반 사용자의 입장에서 매우 복잡하고 불편한 작업이다. 이러한 이유로 인하여 여러 주파수 밴드의 음악적 특성이 조합된 다양

* 학생회원, ** 정회원, 단국대학교 정보·컴퓨터학부
(Division of Information and Computer Science,
Dankook University)

※ 이 연구는 2004학년도 단국대학교 대학연구비 지원으로 연구되었음

접수일자: 2005년6월28일, 수정완료일: 2006년1월11일

하고 개성 있는 음색을 낼 수 있음에도 불구하고 대부분의 사용자는 하나의 이퀄라이저 설정으로 모든 장르의 음악을 듣는 경향이 있다. 이퀄라이저의 장점을 보다 편하게 활용하기 위한 방안으로써 WinAmp의 이퀄라이저 적용 방식을 예로 들 수 있다. 윈도우 환경 하에서 MP3파일을 재생하여 주는 WinAmp의 경우 각 음악 파일 마다 이퀄라이저의 수동 설정을 저장 하였다가, 이후에 같은 음악을 들을 때 적용하여주는 방식을 채용하고 있다. 그러나 음악 파일명에 따라 설정을 기억해주는 방식이므로, 음악 파일의 이름이 바뀌게 되면 다시 이퀄라이저를 설정해 주어야 하는 단점이 있다. 이러한 방법은 사용자의 번거로움을 덜어 줄 수는 있으나 근본적인 대책은 되지 못한다.

다양한 오디오기기에 이퀄라이저가 적용되고 있고 그 구현방식 또한 다양하지만, 근본적인 기능과 동작과정은 각 밴드 별로 나누어진 음악 신호의 이득(gain) 값을 증폭하거나, 감쇠 하여 주파수 밴드별로 나타나는 음색의 특징을 조절하여 준다^[1-3]. 예를 들어 3밴드 이퀄라이저의 경우 저주파 대역을 Bass라고 부르며, 이 밴드를 강조하여줄 경우 음악이 힘 있게 들리며 웅장한 느낌을 준다. 중간 대역 밴드를 Mid라고 부르며, 사람의 귀에 가장 민감한 주파수 대역을 담당하고 있으므로 증폭하여줄 경우 음이 명확하게 들리는 특징이 있다. 고주파 대역을 Treble이라고 부르며 증폭하여줄 경우 음이 맑은 느낌을 주게 된다.

본 연구에서는 내용기반 음악 장르 분류 기술을 이용하여 지능형 그래픽 이퀄라이저 시스템을 제안하였다. 제안된 시스템은 우선 입력 음악으로부터 특징 벡터를 추출하고 이를 바탕으로 특징벡터에 대한 패턴인식 훈련과정을 거쳐 각 음악의 장르 범주를 정해 주며, 새로운 음악 데이터가 입력 될 경우 미리 구축되어 있는 DB와 비교하여 가장 유사한 음악 장르를 결정한다. 이러한 음악 장르 정보를 바탕으로 그래픽 이퀄라이저는 각 장르별 미리 할당되어 있는 주파수 밴드 이득(gain) 값을 적용하여 음악을 재생하게 된다. 이러한 모든 음악 장르 분류 및 이퀄라이저 적용 작업은 자동으로 동작되어 기존 사용자에게 의한 수동 작업의 불편함을 덜어 줄 수 있게 된다.

오디오 정보에 기반 한 내용기반 정보 분류 검색 시스템은 지난 몇 년간 국내·외에서 다양한 연구가 이루어져 왔다. Tzanetakis와 Perry^[4]는 음악 신호로부터 음색 특징, 리듬 정보 그리고 피치 정보에 해당하는 특징 벡터를 추출한 후 10개 음악 장르에 대하여 약 70%의

분류 성공률을 보였고, 리듬과 피치 정보만을 사용하였을 경우에는 23%~28%의 아주 저조한 결과를 나타내었다. 논문 [5]에서 Li는 Daubechies 웨이블릿 계수와 [4]에서 사용된 특징들을 선별·조합하여 여러 가지 비교 실험을 수행하였고, 음악 장르 분류에는 리듬이나 피치 정보 보다는 음색 특징들이 훨씬 효과적이라는 결론을 내렸다. Burred는^[6] 음악을 계층적으로 분류하여 장르별 최적의 특징 벡터를 제시하였다. 13개 음악 장르에 대하여 약 57.8%의 분류 성공률을 얻었다. 또한 Guo와 Li는 [7]에서 이진트리 인식과 support vector machines (SVM)의 사용을 제안하였고, distance-from boundary (DFB)라는 새로운 측정 방법을 사용하여 음악 장르 간의 패턴 유사도를 측정하였다. 이외에도 음악 및 일반 오디오에 대한 특징 추출, 장르 분류 및 검색에 관한 다양한 연구들이 논문 [8-12]에 설명되어있다.

본 연구에서는 실제적인 시스템 구축을 고려하여, 최대한 짧은 시간 안에 정확히 음악 장르를 분류하여 음악 특색에 맞는 이퀄라이저를 적용할 수 있는 방법에 주목하였다. 초기 음악 재생 도중 이퀄라이저 적용에 의한 갑작스런 음색 변화를 완화하기 위하여 Coarse Level, Fine Level의 2단계 장르분류 절차를 거치면서 각 단계별 이퀄라이저를 적용 하였다. 각 단계별 분류 성공률은 약 80%에 이른다. 이렇게 만들어진 시스템은 윈도우OS 상에서 Visual C++을 이용하여 미디어 플레이어 형태의 그래픽 유저 인터페이스로 구현 하였다.

본 논문의 구성은 다음과 같다. 먼저 II장에서는 지능형 그래픽 이퀄라이저 시스템의 전반적인 구성과 핵심 기술에 대하여 설명하였다. III장에서는 다양한 컴퓨터 모의실험을 통해 비교·분석을 수행하고 이를 토대로 최대한 효율적인 이퀄라이저 시스템 구성 방식을 도출 하였다. IV장에서는 윈도우 미디어 형태로 구현된 최종 시스템의 구현 방법에 대하여 설명하였다. 마지막으로 V장에서는 결론으로 끝을 맺는다.

II. 지능형 오디오 그래픽 이퀄라이저 시스템

1. 시스템 동작 구조

다음의 그림 1 은 제안 시스템의 구조를 나타낸다. 시스템은 각 장르별 음악 파일로부터 WAV형식의 오디오 신호를 입력 받아 음악 장르를 판단하여 이퀄라이저를 적용한 후 음악을 재생한다.

지능형 그래픽 이퀄라이저의 전체 동작과정은 훈련 및 테스트의 2개 과정으로 이루어진다. 훈련 단계에서

는 장르별 특징 테이블을 생성하기 위한 패턴 학습 (Training)을 한다. 패턴 학습은 Classic, Jazz, Hiphop, Rock에 해당하는 오디오 데이터의 각 특성을 모델링하기 위해 많은 수의 음악 파일로부터 특징 벡터를 추출, 분석하여 통계적 모델링을 하는 작업이다. 시스템에 음악 파일이 입력(1~4)되면 Segmentation & Windowing (Hamming)과정(5)을 거쳐 오디오 특징 벡터를 추출하여 최적의 특징 벡터를 선정(6-1)하고 음악 파일의 전체 음악적 특성을 반영한 특징 벡터로 만들기 위하여 MFC(Multi Feature Clustering)알고리즘을 적용한 후 (6-2) 모델 정합과 색인화 과정을 포함하는 학습 단계 (7-1)를 거쳐 특징 벡터 테이블을 생성하여 이를 DB에 저장(7-2)하는 과정이다. MFC과정은 II장 2절에서 설명하였다.

테스트 과정은 사용자 입장에서의 시스템 동작 과정과 일치하는데, 사용자가 음악 파일을 재생(1)시키면 시스템은 이 파일을 받아 WAV 파일을 추출하고 (1-4) Segmentation & Windowing(Hamming)작업(5)을 수행하여 선정된 최적의 특징 벡터를 추출(6)하고 오디오 DB에 이미 구축되어있는 특성 테이블을 참고하여 유사한 특성을 지니는 장르를 결정(8-1~3)하여 각 장르에 맞는 이퀄라이저를 적용(9)한 파일을 재생하게 된다.

본 시스템의 시간에 따른 동작순서는 다음과 같다. 음악이 시작되면 처음 2초간의 오디오 데이터를 이용하여 Coarse level분류를 행하여 대략적인 음악 범주를 결정하며 이에 맞는 이퀄라이저를 적용한다. 이후, 음악의 시작 지점부터 6초 분량의 데이터를 이용하여 Fine level의 분류기를 이용하여 정확한 음악 장르를 결정한 후 좀 더 세밀한 이퀄라이저를 적용하여 준다. 이러한

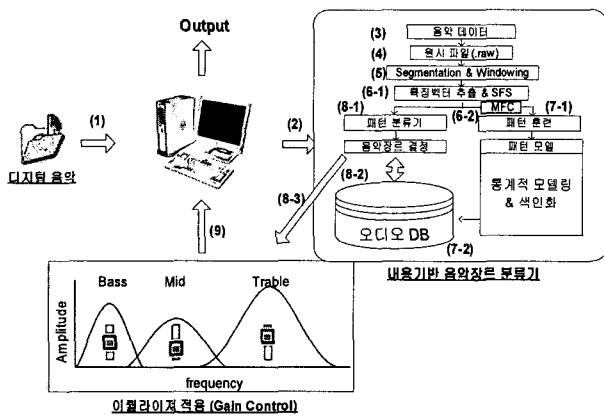


그림 1. 지능형 그래픽 이퀄라이저 시스템 구조도
Fig. 1. Structure of Intelligent Audio Graphic Equalizer System.

2단계 음악 장르 분류는 갑작스러운 음색 변화로 인한 청취자의 거부감을 완화하기 위한 것이다.

2. 최적화된 특징 벡터 추출 및 내용기반 장르 분류

본 연구팀은 이전 논문 [14]에서 내용기반 장르 분류 및 검색 시스템을 구축할 때, 보다 강한 시스템 구축을 위한 MFC (Multi-Feature Clustering) 기술에 관련된 연구를 수행한 바 있다. 본 논문에서는 지능형 오디오 그래픽 이퀄라이저 시스템으로 응용 분야는 다르지만 내용기반 장르 분류 시 기존의 MFC 기술을 이용하여 보다 짧은 질의를 이용하면서도 정확도를 높일 수 있는 시스템을 구축하였다.

가. 특징벡터 추출

특징벡터는 음악 신호의 시간 및 주파수파형으로 부터 추출한 특징 값을 조합하여 구성하게 되며 크게 음색(Timbral)특징 값 과 계수(Coefficient)특징 값 2가지로 분류 할 수 있다. 음악의 음색 특징^[4]으로는 spectral centroid, spectral rolloff, spectral flux, zero crossing rates의 4가지를 사용하였고, 계수 특징으로는 13차 MFCC(Mel frequency cepstral coefficient) 및 Delta-MFCC계수와 12차 LPC(Linear predictive coding)및 Delta-LPC계수를 사용하였다. 음악 파일로부터 특징 값을 추출하여 특징벡터를 구성하기 위한 과정은 그림 2 와 같이 2단계로 이루어진다. 먼저 23ms 단위의 분석 창(Analysis window)으로부터 각각 54개의 특징 값을 얻어 내고, 총 15초 동안의 특징 값을 추출한 후 평균 및 분산을 구하여 총 108개의 특징벡터 로 조합을 하여 DB구축 및 음악 장르분류에 사용하였다.

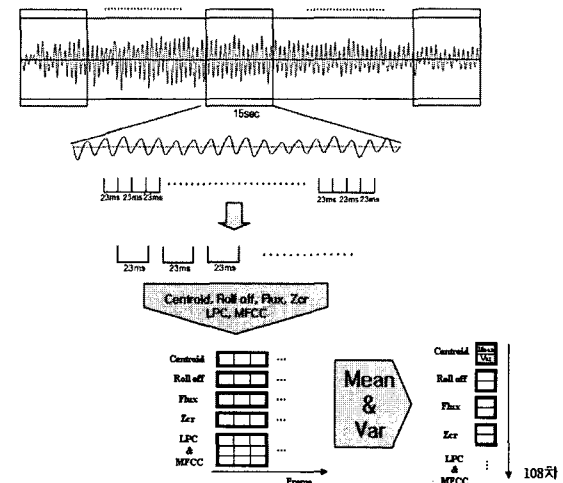


그림 2. 특징벡터 추출 과정
Fig. 2. Feature vector extraction.

나. SFS 알고리즘을 이용한 특징 벡터의 최적화

내용기반 음악 장르 분류는 입력받은 음악 데이터로부터 특징 값을 추출하여 각 음악 장르별로 DB를 구축하여 장르 분류에 사용한다. 본 연구에서 사용한 특징 벡터는 총 108차 이므로, 각 음악 파일 당 108개의 특징 값이 생성되며 이는 음악 장르분류 시스템의 계산 복잡도를 높여지게 한다. 이를 해결하기 위하여 본 연구는 SFS(Sequential Forward Selection) 알고리즘^[13]을 이용하여 특징벡터간의 상관관계를 제거하고 특징 벡터의 차수를 줄여 최적화된 특징벡터의 특징 값만을 사용하게 된다.

다. MFC(Multi Feature Clustering) 알고리즘

특징 벡터는 음악의 전체적 특성을 반영하기 위하여 총 15초 분량을 추출하여 평균과 분산 값을 구하고 SFS 알고리즘을 통하여 최적화 된다. 그러나 질의(Query)로 입력된 음악 데이터가 음악 파일의 어느 부분에서 추출되어 들어오는지에 따라 장르 분류 성공률의 편차가 발생하게 되며 특히 음악의 첫 시작 부분에서는 비교적 음악의 특징이 잘 담겨 있지 않기 때문에 음악 장르 분류의 성공률을 낮아지게 하는 경향이 있다.

본 논문에서는 이러한 성공률의 편차를 극복하고 안정적인 성능을 낼 수 있는 시스템 구축을 위하여 논문 [14]에서 제안한 MFC(Multi Feature Clustering) 기법을 도입하여 음악질의 위치와 질의 길이에 상관없이 안정적인 성공률을 보여 줄 수 있는 시스템을 구축하였다. MFC는 음악 파일의 시작부분부터 끝부분까지 순차적으로 특징벡터를 추출한 후, K-Means 클러스터링 알고리즘을 사용하여 최종적으로 1개의 음악 파일 당 4개의 특징벡터열을 만들어 음악 메타데이터 DB를 구축하며 음악 장르 분류에 사용한다. 그림 3 은 이러한 MFC

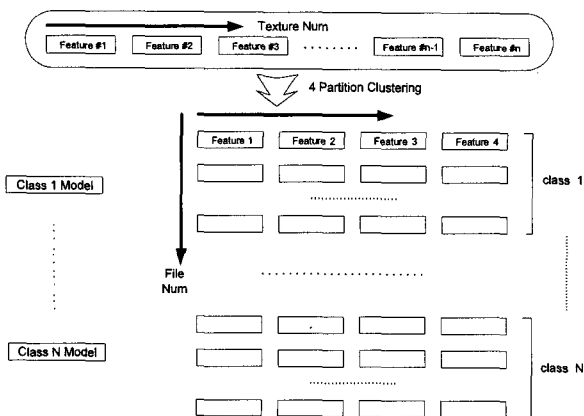


그림 3. MFC 알고리즘을 이용한 DB 구축
Fig. 3. Construction DB using MFC.

알고리즘을 이용하여 장르별 모델링을 하는 과정을 보여준다.

라. 장르분류기

본 논문에서는 최적화된 시스템 구축을 위하여 k-NN(k-nearest neighbor), Gaussian, GMM(Gaussian Mixture Model), HMM(Hidden Markov Model), SVM(Support Vector Machine) 등 다양한 패턴 분류기를 검토하였으며 이중 장르 분류 성공률과 알고리즘 복잡도를 감안하여 SVM과 k-NN분류기를 각각 1단계 coarse level 과 2단계 fine level 분류기에 사용하였다. k-NN분류기는 Euclidean 거리 함수를 이용하여 질의로 입력된 특징벡터와, DB에 저장되어 있는 특징 벡터열 간의 거리 값을 구하여 데이터의 유사도를 측정하는 방식이다. SVM은 훈련데이터를 이용하여 고차원 선형방정식을 얻어내고, 이를 이용하여 2개의 장르를 구분한다. SVM의 경우 기본적으로 2개의 장르만을 구분할 수 있으나, 여러개의 SVM분류기를 조합하여 다중클래스를 구분할 수 있도록 하였다.

III. 컴퓨터 모의실험

1. 실험 구성

실험을 위해 사용된 음악 DB는 인터넷 전문 음악 사이트의 MP3파일, 음악CD 그리고 라디오로부터 Classic, Hiphop, Jazz 그리고 Rock의 4개 음악 장르에 대해 각 장르별로 60곡을 선정하여 총 240곡을 사용하였다. 이중 매 실험 마다 랜덤하게 장르별로 45곡 총 180곡을 훈련용 DB로 사용하고 나머지 60곡을 질의 데이터로 사용하였다. 실험에 사용된 음악 파일은 22050Hz, 16bits, mono WAV형식의 파일로 변환하여 사용하였다.

질의 데이터는 음악파일들을 1초부터 10초까지 1초 단위로 늘려가면서 특징벡터를 추출하여 질의 시간별 성공률을 측정하는데 사용하였다. 본 연구에서 사용된 분류기는 k-NN과 SVM을 사용하였으며, 특징벡터의 차수와 질의 시간별로 성공률을 측정하여 가장 높은 성공률을 보이는 분류기를 선정하여 사용하였다.

제안된 시스템은 그림 4 와 같이 2단계 - coarse level과 fine level - 분류 단계를 거쳐 질의 음악의 장르를 분류하게 된다. Coarse level분류 단계에서는 음악의 4개 장르를 비교적 음색 특징이 유사한 두 개 장르 (Classic & Jazz, Hiphop & Rock)로 구성하여 질의 음

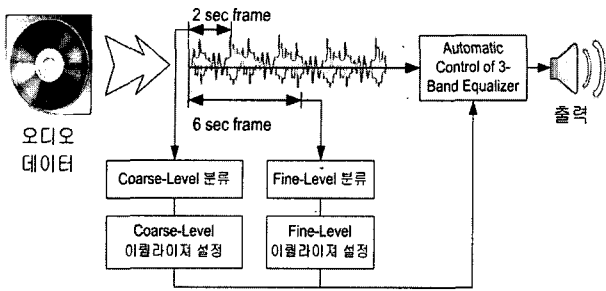


그림 4. 2단계 이퀄라이저 시스템
Fig. 4. Two-step hierarchical Equalizer system.

악을 분류하고 이퀄라이저의 평균 gain 값을 적용함으로써 최종적으로 fine level의 이퀄라이저가 적용될 때까지 자연스러운 음색의 변화를 유도하는 효과를 얻을 수 있다. 이러한 시스템 특성을 얻어 내기 위하여 Coarse level의 분류는 질의 음악 데이터의 길이가 3초 미만인 경우 약 80%의 성공률을 얻을 수 있는 지점을 찾아내는 것을 목표로 하였다. 2단계 Fine level분류기에서는 이를 세분화하여 질의 음악의 해당 4개 장르를 각각 분류 하였으며, 질의 시간이 7초미만 일 때 80% 이상의 성공률을 보이는 지점으로 선정하였다.

2. 실험 결과

표 1 은 II장의 SFS와 MFC기법을 이용하여 구해낸 최적 특징벡터에 대한 1단계 Coarse level에서의 질의 시간별 분류 성공률이다. SVM의 경우 34차 특징벡터를 이용할 경우 최고의 성공률을 보여 주었으며, k-NN의 경우 36차 특징벡터를 사용할 경우 최고의 성공률을 보여 주었다. SVM의 경우 2초간의 질의 데이터만을 사용하여도 충분히 만족할 만한 성공률을 보여 주고 있으며, k-NN의 경우 1-NN을 사용하여 4초 이상의 질의 데이터를 사용해야만 80% 이상의 성공률을 보여 주었다.

표 2 는 fine level에서의 질의 시간별 분류 성공률을 보여 준다. SVM의 경우 2진 분류를 여러번 하여 나온 결과를 토대로 다중 클래스 분류를 하기 때문에, 다중 클래스 분류를 할 경우 성공률이 떨어짐을 볼 수 있다.

표 1. Coarse level에서의 질의 시간별 장르분류 성공률

Table 1. Coarse-level classification accuracy.

질의 시간(sec)	1	2	3	4	5
SVM	75.833	79.167	79.167	80.833	81.25
k-NN	65.833	72.917	77.917	82.917	83.75

표 2. Fine level에서의 시간별 장르분류 성공률
Table 2. Fine-level classification accuracy .

질의 시간(sec)	1	2	3	4	5	6	7	8
SVM	35.42	43.33	52.92	57.08	60.83	63.33	68.75	72.5
k-NN	45.83	62.92	72.08	76.25	78.75	81.67	85	86.67

k-NN은 66차 최적 특징벡터 에서 6초의 질의 데이터를 이용할 경우 81.67%의 성공률을 보여 주었다.

1단계 Coarse level분류의 경우 SVM은 2초길이의 질의 데이터에 대해 약 79%에 이르는 성공률을 보여 주고 있다. k-NN이 4초의 질의를 입력 해야지만 80%가 넘는 성공률이 나오는 것에 비해 약 2배 정도 짧은 질의를 넣었음에도 불구하고 만족할 만한 결과를 보여 주었다. 이는 단순히 전체 시스템이 시간적으로 빠른 결과를 얻을 수 있도록 해줄 뿐 만 아니라, 특징 벡터를 얻어내기 위한 연산횟수를 절반 이하로 줄여 줌으로써 전체적인 시스템의 부담을 줄여주게 된다. 본 논문에서는 총 240개의 음악 파일로부터 질의로 사용할 60개의 파일을 제외하고 180개의 훈련용 파일을 이용하여 720개 (그림 3에서 1개 음악 파일 당 4개 특징 벡터)의 특징벡터를 추출 하여 DB로 저장하고 있다. 공간 복잡도의 경우 n을 특징벡터의 차수라고 하면 SVM은 선형방정식 연산에 필요한 기울기(n)와 질의로부터 추출한 특징벡터의 저장 공간(n), Bias 값(1) 및 결과를 저장할 공간(1)을 합하여 총 [2n + 2]의 연산 공간만을 확보하면 되지만 k-NN의 경우 본 시스템을 기준으로 720개의 특징벡터(720n) 및 질의로 들어온 벡터(n)와 각 벡터간의 거리 값이 저장될 공간(n) 과 최종 결과 값이 저장될 변수(1) 를 총 합하여 [722n + 1]의 연산 공간을 필요로 한다.

2단계 Fine level 의 분류성공률은 k-NN의 성공률이 SVM 보다 약 20%가량 높은 성공률을 보여준다. 이러한 결과는 SVM이 가지는 여러 가지 장점에도 불구하고 Fine level 의 장르 분류에서는 k-NN을 사용하는 것이 보다 안정된 시스템 구축을 가능하게 함을 볼 수 있다.

위의 계산 복잡도의 경우 PC상에서는 커다란 차이를 느낄 수 없는 정도의 계산량이지만 모바일 시스템에서 구현 시 여러 가지 장점을 가지게 된다.

IV. 시스템 구현

본 연구의 목표는 지능형 그래픽 이퀄라이저 알고리즘을 구현 하고, 이를 비주얼 C++을 이용하여 소프트웨어로 구현하는 것이다. 이를 위하여 이전 장의 결과를 토대로 시스템 구동 방식을 그림 5와 같이 결정하였다. 그림 6은 지능형 그래픽 이퀄라이저의 프로그램 구조도를 나타내고 있다. 지능형 그래픽 이퀄라이저 시스템은 이퀄라이저 루틴과 분류기 루틴으로 구성되어 있다.

이퀄라이저 루틴에서는 음악 파일이 23ms의 frame 단위로 나뉘어져서 이퀄라이저가 적용된 후, overlap add 방식으로 재조립 되어 출력된다.

분류기 루틴에서는 음악 파일이 재생되기 시작한 시점부터 이후로 6초간의 음악 데이터를 버퍼에 저장하고 coarse level과 fine level의 장르 분류에 필요한 연산을 순차적으로 한 후, 분류 결과를 이퀄라이저 루틴에 알려준다.

본 시스템의 기능은 오디오 시스템에서 음악이 재생 되면, 각 장르에 알맞은 음색을 내줄 수 있도록 이퀄라

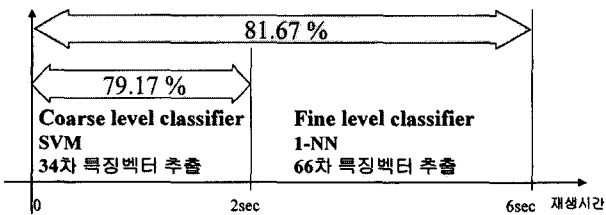


그림 5. 2단계 분류 알고리즘
Fig. 5. Two-step hierarchical classification algorithm.

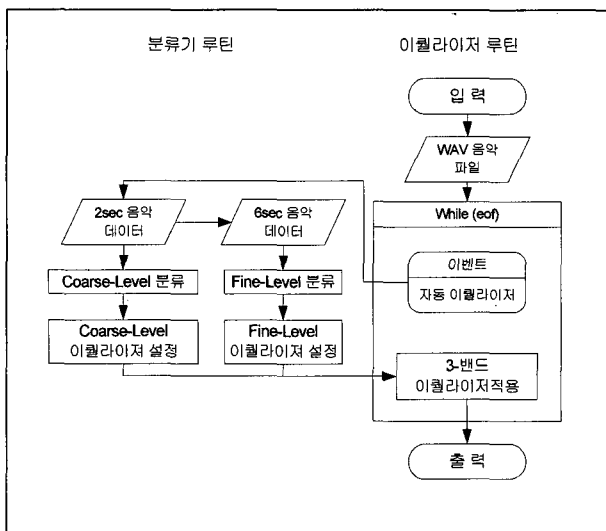


그림 6. 지능형 그래픽 이퀄라이저 프로그램 구조
Fig. 6. Overall intelligent audio graphic equalizer program.

이저를 자동으로 설정해 주는 것이다. 이러한 이유로 본 논문에서 구분한 4개의 음악 장르에 알맞은 이퀄라이저 설정을 정의해 주어야 한다. 본 논문에서 사용한 음악 장르는 Classic, Hiphop, Jazz, Rock의 4개 장르로써, 각 장르별 이퀄라이저 설정 값은 윈앰프 등의 범용 프로그램에서 설정한 값을 참고하여 표 3 과 같이 주었다. 표 안의 값들은 원본 음악 소스의 주파수 밴드별 진폭 크기에 곱해주는 이득(gain) 값들이다.

Coarse level의 경우 Slow Beat과 Fast Beat으로 장르를 구분하여 주었다. 이퀄라이저 설정 값은 Slow Beat장르는 Classic과 Jazz의 평균값을 이용하였고, Fast Beat 장르는 Hiphop과 Rock의 평균값을 이용하였다.

그림 7 은 비주얼 C++의 MFC를 이용하여 구현된 유저 인터페이스이다. 주요 기능으로는 불러오기, 재생, 정지, 반복, 수동 이퀄라이저 설정 등이 있다. 자동 이퀄라이저가 활성화 되면 활성화된 시점으로부터 6초간의 데이터를 이용하여 Coarse level, Fine level의 분류를 실행하여 이퀄라이저를 적용해준다.

표 3. 3 밴드 이퀄라이저의 주파수 gain
Table 3. Frequency gains in 3-band audio graphic equalizer.

음악 장르		BASS (LowPass)	MID (BandPass)	TREBLE (HighPass)
Coarse	Slow Beat	1	0.9	0.9
	Fast Beat	1.4	1	1.2
Fine	Classical	1	1	0.6
	Jazz	1	0.8	1.2
	Hiphop	1.5	1.3	1
	Rock	1.3	0.7	1.4

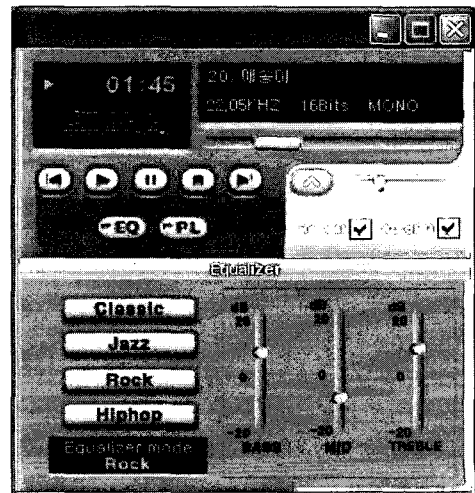


그림 7. 그래픽 유저 인터페이스
Fig. 7. Graphic User Interface.

지능형 그래픽 이퀄라이저 시스템을 구동하는데 사용된 컴퓨터의 사양은 AMD 2500+, 1GigaRAM이다. 매번 실험할 때 마다 약간의 차이는 있으나, 평균적으로 Coarse level의 경우 분류 결과가 나오는데 약 0.7초가 소비 되었으며, fine level의 경우 약 1.1초의 시간이 소비 되었다. 여기서 0.7 초 와 1.1초는 각 2초 와 6초 분량의 음악 신호 데이터를 받아 들여 장르 분류 처리 하여 실제 각 단계에서 이퀄라이저가 적용되기까지의 시간을 말한다.

IV. 결 론

본 연구 에서는 내용기반 음악장르분류 시스템을 이용하여 자동으로 각 음악 장르에 맞는 이퀄라이저를 적용하여 주는 시스템을 제안하였다.

SFS기법을 사용하여 시스템의 계산 복잡도를 줄여 주고 분류성공률을 높여주었으며, MFC기법을 적용하여 DB를 구축하여준 결과 음악질의를 위치에 상관없이 안정적인 성공률을 보여주는 시스템을 구축하였다. 시스템의 반응속도를 극대화 하고, 자연스러운 음색의 변화를 유도하기 위하여 Coarse level에서 Fine level로 이어지는 2단계 이퀄라이저 적용 방식을 택하였다. Classic, Jazz, Hiphop, Rock의 4개 장르에 대하여 coarse level과 fine level의 2단계 분류기를 구성하고, 각각 SVM과 k-NN패턴 인식 알고리즘을 사용하였다. 완성된 시스템의 각 level 별 장르 분류 성공률은 80%에 달하며, 2초 안에 장르 분류 및 이퀄라이저 적용 등의 모든 동작이 완료된다.

이러한 시스템은 PC상에서 동작하는 음악 파일 재생기뿐만 아니라 CD플레이어, MP3플레이어, CAR오디오 시스템 등 다양한 플랫폼으로 구현 될 수 있다.

참 고 문 헌

- [1] A. Bellini, G. Cibelli, E. Ugolotti, et al, "Non-linear digital audio processor for dedicated loudspeaker systems", IEEE Trans. On Consumer Electronics, Vol. 44, pp. 1024-1031, August 1998.
- [2] W. Klippel. "Compensation for nonlinear distortion of horn loudspeakers by digital signal processing", Journal of AES, vol. 44, no. 11, PP. 964-972, Nov. 1996.
- [3] J.N. Mourjopolos, "Digital equalization of room acoustics", Journal of AES, vol. 42, no. 11, pp. 884-900, Nov. 1994.
- [4] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals" IEEE Trans. on Speech and Audio Processing, vol. 10, no. 5, pp. 293-302, July 2002.
- [5] T. Li, M. Ogihara and Q. Li, "A comparative study on content-based music genre classification" in Proc. of the 26th annual internal ACM SIGIR, pp. 282-289, ACM Press, July 2003.
- [6] J. J. Burred and A. Lerch, "A hierarchical approach to automatic musical genre classification" in Proc. DAFx03, pp. 308-311, 2003.
- [7] G. Guo and S. Z. Li, "Content-based audio classification and retrieval by support vector machine" IEEE Trans. on neural networks, vol. 14, no. 1, pp. 209-215, Jan. 2003.
- [8] J. Foote et al, "An overview of audio information retrieval" ACM-Springer Multimedia Systems, vol. 7, no. 1, pp. 2-11, Jan. 1999.
- [9] Y. Wang, Z. Liu and J. Huang, "Multimedia content analysis: using both audio and visual clues" IEEE Signal Proc. Mag., Nov. 2000.
- [10] S. Blackburn, "Content based retrieval and navigation of music" Mini-thesis, University of Southampton, 1999.
- [11] S. Z. Li, "Content-based classification and retrieval audio using the nearest feature line method" IEEE Trans. on Speech Audio Processing, vol. 8, pp.619-625, Sept. 2000.
- [12] T. Zhang and C. Kuo, "Hierarchical system for content-based audio classification and retrieval" Proceedings of SPIE98, vol.3527, pp. 398-409, Boston, Nov, 1998.
- [13] D.Ververidis and C.Kotropoulos, "Sequential Forward Feature Selection with Low Computational Cost" FP6 European Union Network of Excellence MUSCLE "Multimedia Understanding through Semantics, Computation and LEarning" (FP6-507752)
- [14] Kyu-Sik Park, Won-Jung Yoon, Kang-Kue Lee, "A Robust Approach to Content-Based Musical Genre Classification and Retrieval Using Multi-Feature Clustering" ASIAN2004, LNCS Thailand, Dec. 2004.

저 자 소 개



이 강 규(학생회원)
 2003년 상명대학교 정보통신학과
 학사 졸업.
 2005년 단국대학교 컴퓨터 과학
 및 통계학과 석사졸업
 2005년~현재 단국대학교 컴퓨터
 과학 및 통계학과 박사과정
 <주관심분야 : 음성 및 음향신호처리, 멀티미디어
 신호처리, DSP 시스템 구현>



조 윤 호(학생회원)
 1994년 단국대학교 농과대학
 농학과 학사 졸업.
 2004년 단국대학교 멀티미디어
 대학원 멀티미디어학과
 석사졸업
 2004년~현재 단국대학교 컴퓨터
 과학 및 통계학과
 박사과정
 <주관심분야 : 음성 및 음향신호처리, 멀티미디어
 신호처리, DSP 시스템 구현>



박 규 식(정회원)
 1986년 Polytechnic University
 전자공학과 학사 졸업.
 1988년 Polytechnic University
 전자공학과 석사 졸업.
 1993년 Polytechnic University
 전자공학과 박사 졸업.
 1994년~1996년 삼성전자 마이크로사업부,
 선임 연구원
 1996년~2001년 상명대학교 컴퓨터·정보통신
 공학부 조교수
 2001년~현재 단국대학교 정보컴퓨터학부 부교수
 <주관심분야 : 음성 및 음향신호처리, 멀티미디어
 신호처리, DSP 시스템 구현>