

악보 인식과 연주 기능을 제공하는 음악 플레이어

이승환* · 황보인 · 홍기찬 · 최우석 · 고정국

동명대학교 컴퓨터공학과

A Music player providing a musical score recognition and performance function

Seung-Hwan Lee* · In HwangBo · Kee-Chan Hong · Wo-Suk Choi · Jeong-Gook Koh

Dept. of Computer Engineering, Tongmyong University

E-mail : tmddydk91@naver.com

요 약

컴퓨터의 활용 범위가 다양화되면서 음악 분야에도 컴퓨터의 역할이 점점 커지고 있다. 기존의 악보 인식 방법들은 대부분 특정 프로그램에서 작성된 악보만 편집과 재생이 가능하고 다른 프로그램에서 작성된 악보는 수정과 재생이 어려운 문제가 있었다. 본 논문에서는 이러한 단점을 보완하여 이미 작성된 악보를 자동으로 인식하고 재생할 수 있는 스마트폰 애플리케이션을 설계하고 구현한다. 스마트폰 소지자는 언제 어디서나 원하는 악보를 획득할 수 있으며, 폰에 내장된 카메라를 이용하여 악보를 인식시킨 후 연주 기능을 이용하여 음악을 즐길 수 있다.

키워드

악보 인식, 악보 재생, 연주기능, 스마트 폰, 애플리케이션

1. 서 론

컴퓨터의 활용 범위가 다양화되면서 음악 분야에도 컴퓨터의 역할이 점점 커지고 있다. 그러나 기존의 악보 인식 방법들은 대부분 특정 프로그램에서 작성된 악보만 편집과 재생이 가능하고 다른 프로그램에서 작성된 악보는 수정과 재생이 어려운 문제가 있었다.

본 논문에서는 이러한 단점을 보완하여 이미 작성된 악보를 자동으로 인식하고 재생할 수 있는 스마트폰 애플리케이션을 설계하고 구현한다. 스마트폰 소지자는 언제 어디서나 이미지 검색이나 카메라 촬영 등을 통해 원하는 악보 이미지를 획득할 수 있으며, 폰에 내장된 카메라를 이용하여 악보를 인식시킨 후 연주 기능을 이용하여 음악을 즐길 수 있다[1].

II. 스마트폰 애플리케이션의 설계 및 구현

2.1 설계 목표

본 논문의 설계 목표는 스마트폰 사용자가 언제 어디서나 이미지 검색이나 카메라 촬영 등을 통해 원하는 악보 이미지를 획득하고 스마트폰에 내장된 카메라를 이용하여 악보를 인식시킨 후 연주 기능을 이용하여 음악을 즐길 수 있는 스마

트폰 애플리케이션을 구현하는 것이다. 본 논문에서 설계한 스마트폰 애플리케이션의 세부적인 동작 과정은 그림 1과 같다. 애플리케이션을 실행한 후 Google 이미지 검색이나 카메라 촬영을 통해 악보 이미지를 획득하고, 악보 인식 기능을 이용하여 악보를 인식한다. 그리고 재생 기능을 이용하여 악보를 재생시키며, 인식된 악보를 스마트폰에 저장할 수도 있다.

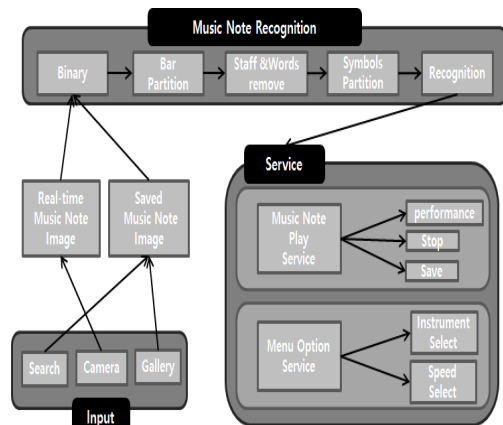


그림 1. 스마트폰 애플리케이션의 동작 과정

2.2 기능 설계

Google 이미지 검색이나 카메라 촬영을 통해 획득한 악보 이미지의 인식 과정은 표 1과 같다.

표 1. 악보 인식 과정

단계	세부 동작
1단계	이미지 획득
2단계	이진화
3단계	소절 분할 및 가사 제거
4단계	오선 제거 (음표 영역 확보)
5단계	음표 분할(개별 음표 확보)
6단계	악보 인식
7단계	악보 연주

우선, 획득한 악보 이미지를 인식하기 위해서는 이진화 과정을 수행한다. 그 다음에는 악보의 소절을 분할하고 오선과 가사를 제거하며, 음표 분할을 통해 개별 음표를 확보한다. 그러나 일반적인 템플릿 매칭으로는 만족스러운 인식 성능을 얻기 힘들기 때문에 본 논문에서는 이미지의 픽셀을 이용하여 보다 정교하고 세밀한 작업을 수행하는 방법을 제안한다. 음표 인식 과정에서 음표는 머리와 꼬리 부분으로 구분된다는 점을 감안하여 머리 부분은 픽셀을 이용한 구분 방법을, 꼬리 부분은 템플릿 매칭을 이용한 방법을 적용한다. 악보 인식이 완료되면 음악 플레이어의 연주 기능을 이용하여 악보를 연주하게 된다.

스마트폰 애플리케이션의 메인 화면은 그림 2와 같이 검색(Search), 갤러리(Gallery), 카메라(Camera), 그리고 게임(퍼즐) 기능을 제공한다.

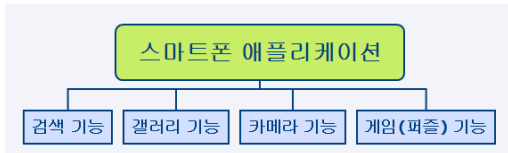


그림 2. 스마트폰 애플리케이션의 주요 기능

2.3 악보 인식 기능 구현

첫 번째 단계는 이미지의 이진화이다. 이를 위해 OpenCV의 cvtColor 함수를 이용하여 원본 이미지를 COLOR_RGBA2GRAY화 시킨 후 threshold 함수를 이용하여 이진화된 영상을 구한다.

두 번째 단계는 소절 분할 및 가사 제거이다. 악보에는 제목, 가사 외에도 다양한 기호들이 존재하지만, 제목, 가사 등은 인식에 사용되지 않기 때문에 제거해야 한다. 그림 3은 소절 분할과 가사가 제거된 영상을 보여준다.



그림 3. 소절분할 및 가사를 제거한 영상

세 번째 단계는 오선 제거이다. 오선을 제거하려면 각 소절의 오선 위치와 오선의 두께 정보가 필요하다. 그림 4는 그림 2에서 오선을 제거한 영상이다.

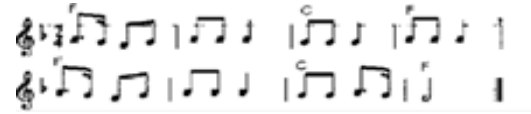


그림 4. 오선을 제거한 영상

네 번째 단계는 음표 분할이다. 오선이 제거된 소절 단위의 악보 이미지를 토대로 음표를 그림 5와 같이 분할하며, 분할된 이미지에 대한 추가적인 가공을 통해 불필요한 요소를 제거한다. 그림 6은 80X80 크기로 비트맵 크기를 조절하여 정규화된 음표 영상이다.



그림 5. 분할된 음표 그림 6. 정규화된 음표

다섯 번째 단계는 악보 인식 및 후처리이다. 악보 인식을 위해서는 음표와 음표가 아닌 것의 구분, 음표 머리 부분의 홀 여부, 음표 꼬리 부분의 구분, 계명 추출이 필요하다. 본 논문에서는 분할된 이미지에서 음표와 음표가 아닌 것들을 구분하기 위해 OpenCV의 matchTemplate을 사용하였다[1]. matchTemplate을 사용하려면 비교 이미지를 미리 등록해야 한다. 등록된 이미지와 비교 대상의 유사도를 측정하여 음표 여부를 구분한다. 구분 작업이 완료되면 음표로 구분된 이미지에 대한 인식 작업이 시작된다. 음표는 머리와 꼬리 부분으로 구성되며, 머리 부분은 위쪽을 향하는 것과 아래쪽을 향하고 것으로 구분된다. 머리 부분의 위치 검사는 2/3 가량으로 잘라 나뉜 두 개 이미지의 픽셀 값을 검사하며, 머리 부분의 검은 픽셀 값이 더 크므로 해당 부분을 머리 위치로 판단한다. 2분/4분/8분/16분 음표의 꼬리 부분 검사는 음표 구분에 사용한 방법과 동일하게 유사도를 측정하며, 가장 높은 유사도에 해당하는 기호에 대응시킨다. 머리 부분의 홀 여부 검사는 꼬리 부분을 지우고 머리의 중심을 찾은 후 중심에서 8방향을 검사하여 홀의 존재 여부를 판단한다. 계이름은 꼬리 부분을 제외한 음표에서 머리 부분의 가장 위와 아래 지점, 그리고 오선의 위치와 두께를 사용하여 계이름을 구한다.

마지막 단계는 악보 연주이며, 본 논문에서는 소리를 출력하기 위해 안드로이드에서 제공하는 SoundPool을 이용하였다. 박자는 핸들러를 사용하여 4분 음표는 0.5초, 2분 음표는 1초, 8분 음표는 0.25초, 16분 음표는 0.125초 동안 출력하여 구분

한다. 계이름은 미리 만들어둔 wav 파일을 SoundPool에 적재한 후 인식 과정에서 산출된 계이름으로 해당 음을 출력한다.

2.4 구현 환경

본 논문에서는 악보 자동 인식과 재생 기능을 제공하는 스마트폰 애플리케이션을 구현하기 위해 윈도우8.1이 설치된 PC에서 개발 언어로 JAVA를 사용하였다. 스마트폰 애플리케이션의 개발도구로는 Eclipse-jee-luna-SR2-win32-x86_64와 OpenCV-2.4.10, Android-SDK-24.3.3, Android-AVD, Android NDK-r10d, JSON을 사용하였다. 구현된 스마트폰 애플리케이션의 동작 여부를 확인하기 위해 갤럭시 S2 HD와 갤럭시 S3, 옵티머스 LTE2 단말기에서 기능 시험을 실시하였다.

2.5 스마트폰 애플리케이션의 사용자 인터페이스 구현

본 논문에서 구현한 스마트폰 애플리케이션의 사용자 인터페이스는 그림 7과 같다. 그림 7(a)는 메인 화면이며, 그림 7(b)는 악보 검색 화면을, 그림 7(c)는 카메라 구동 화면을, 그림 7(d)는 악보 인식 기능이 동작 중임을 보여준다.



그림 7. 스마트폰 애플리케이션의 UI

2.6 성능 평가

본 논문에서는 구현된 스마트폰 애플리케이션의 성능 평가를 위해 표 2와 같이 3개의 악보를 대상으로 인식률을 측정하였다[2].

표 2. 악보 인식률 측정 결과

제목	심볼 수	오인식 심볼 수	인식률 (%)
작은별	42	1	98
새신	21	3	86
학교중	24	0	100
평균인식률	87	4	95.4

표 2에서 보듯이 미리 등록시킨 음표와 음표

가 아닌 기호들로 구성된 3종의 악보에 대한 인식률 실험 결과는 평균 95.4%에 달하였다. 그러나 다른 악보를 사용했을 때는 전처리 과정 외에도 다양한 처리 과정을 거치면서 발생한 잡음으로 인해 등록시킨 음표와 다르게 인식하는 상황이 연출되었다. 본 논문에서는 현재까지 8개의 음표와 5개의 음표가 아닌 기호를 등록시켰으며, 앞으로 다른 기호와 음표를 추가로 등록한다면 보다 높은 인식률을 보일 것으로 예상된다.

III. 결론

컴퓨터의 활용 범위가 확대되면서 음악 분야에서도 컴퓨터의 역할이 점점 커지고 있다. 본 논문에서는 스마트폰 사용자가 언제 어디서나 이미지 검색이나 카메라 촬영을 통해 원하는 악보 이미지를 획득하고, 해당 악보를 인식하여 재생할 수 있는 스마트폰 애플리케이션을 설계하고 구현하였다. 구현된 애플리케이션은 현재 인식 가능한 음표들(2분/4분/8분/16분 음표)과 음표가 아닌 기호(높은음자리표, 쉼표, 박자)들이 제한적이다.

향후 지속적인 기능 시험과 인식 알고리즘 개선을 통해 미비점을 보완하고 인식률을 개선해 나간다면 복잡한 악보들도 인식하고 연주할 수 있을 것이며, 음악 애호가나 음악 전공자들에게 많은 도움이 될 것이다.

참고문헌

- [1] 유재명, 이귀상, 김기홍, “부분적 템플릿 매칭을 활용한 악보 인식,” 한국컨텐츠학회논문지, 제8권 제11호, pp. 85-93, 2008.
- [2] 박건희, 오성열, 손화정, 유재명, 김수형, 이귀상, “휴대폰 카메라로 촬영한 악보영상 인식을 위한 의사트리 알고리즘,” 한국컨텐츠학회 논문지, 제8권 제6호, pp. 16-25, 2008.