

퍼지 ART 알고리즘을 이용한 인쇄 악보의 자동 인식과 연주

김미정* · 김광백* · 박충식**

*신라대학교 컴퓨터정보공학부
**영동대학교 컴퓨터공학과

Automated Recognition of Printed Music
using Fuzzy ART Algorithm and Performance

Mi-jeong Kim* · Kwang-baek Kim* · Choong-shik Park**

*Division of Computer and Information Engineering, Silla University
**Dept. of Computer Engineering, Youngdong University

요 약

음악 연구에 따른 컴퓨터의 역할이 점차 중요한 비중을 차지함에 따라 보다 효과적인 악보 인식 방법이 요구된다. 기존의 악보 인식 방법에서는 특정 수정 프로그램에서 만든 악보만 그 프로그램에서 채수정과 재생이 가능하다는 단점이 있다. 본 논문에서는 이러한 단점을 보완하기 위하여 이미 작성 되어있는 악보들을 자동으로 인식하고 재생을 할 수 있는 방법을 제안한다. 제안된 악보 인식 방법은 수평 히스토그램을 이용하여 악보 이미지의 오선을 제거한 후, Grassfire 알고리즘을 적용하여 잡음을 제거하고 악보 구성 기호들을 추출한다. 추출된 악보 구성 기호들은 악보 구성 기호의 특징을 이용하여 음표와 쉼표, 그 외의 기호들로 분리한다. 분리된 음표 기호들은 박자마다 다른 음표 형태의 특징을 이용하여 다시 세밀하게 분리하고 쉼표와 그 외의 기호들은 퍼지 ART 알고리즘을 적용하여 인식한다. 인식된 악보 구성 기호들을 이용하여 각각 정보를 저장하고 향후에 악보 구성 기호에 해당하는 음의 재생을 용이하게 한다. 제안된 악보 인식 방법의 성능을 평가하기 위해 50장의 악보 영상을 대상으로 실험한 결과, 본 논문에서 제시한 악보 영상의 인식 방법이 실험을 통해서 효율적인 것을 확인하였다.

키워드

악보, 모폴로지, Grassfire 알고리즘, 퍼지 ART

1. 서 론

기술이 발달함에 따라 더욱 신속하고 정확한 정보가 요구되고 있다. 이것은 IT 분야뿐만 아니라 음악과 같은 예능 분야에서도 요구된다. 그에 따라 악보의 인식과 인식 후의 기능 또한 디지털화가 요구된다. 악보가 전용 소프트웨어로 제작되고 있지만, 기존의 악보 제작 소프트웨어들은 다른 소프트웨어들로 작성된 악보들은 인식과 편집을 할 수 없다는 단점이 있다[1].

기존 연구 방법에서는 인쇄 악보의 인식 성

능을 개선하였으나, 같은 박자와 위치의 기호라도 다양하게 표현되는 악보 구성 기호를 배제할 채 인식하는 문제점이 있다[2].

따라서 본 논문에서는 이러한 문제점을 개선하기 위하여 인쇄된 악보 영상에 대한 새로운 추출과 인식 방법을 제안한다. 악보의 정확한 인식을 위하여 획득한 악보 영상에 평균 이진화를 적용하고, 이진화 된 영상을 모폴로지의 팽창을 적용하여 기호들을 연결시킨다. 모폴로지가 적용된 영상에 Grassfire 알고리즘을 이용하여 작은악절영역을 추출하고 포함되지 않는

영역을 잡음으로 간주하고 제거한다. 수평 히스토그램을 이용하여 오선을 제거하고, Grassfire 알고리즘을 이용하여 악보 구성 기호들을 추출한다. 추출된 악보 구성 기호들 중 잇단 음표들은 형태학적 특징을 이용하여 분리한다. 최종적으로 추출된 악보 구성 기호들은 정규화와 세선화를 적용하여 형태학적 특징을 이용하여 분리하고, 분리된 기호들 중 음표 기호를 제외한 쉼표 기호와 그 외의 기호들은 퍼지 ART 알고리즘을 이용하여 인식한다.

II. 악보 구성 기호 추출

본 논문에서 제안한 악보 구성 기호 추출 방법은 획득한 영상에서 불필요한 컬러들을 제거하고, 평균 이진화 기법을 적용하여 악보 영상을 이진화 한 후에 모폴로지의 팽창을 적용하고 Grassfire 알고리즘을 이용하여 작은악절영역을 추출한다. 작은악절영역에 포함되지 않는 영역은 잡음으로 간주하여 제거한다. 수평 히스토그램을 이용하여 오선을 제거한 후 Grassfire 알고리즘을 이용하여 악보 구성 기호들을 추출하고 잇단 음표들을 분리한다. 분리되어 추출된 악보 구성 기호들은 인식을 위하여 정규화와 세선화를 적용한다.

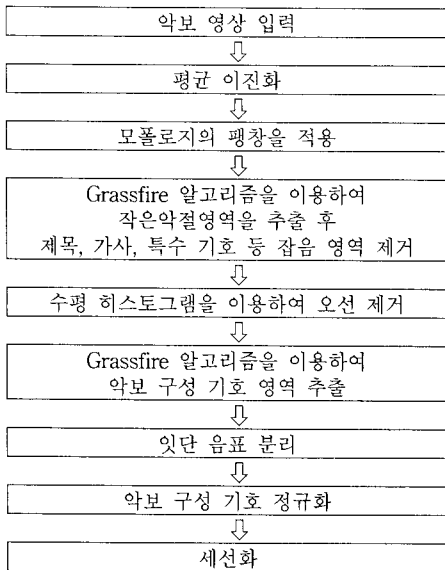


그림 1. 악보 구성 기호 추출 과정

제안된 악보 구성 기호를 추출하는 과정은 그림 1에 나타내었다.

악보 영상에서 불필요한 색상들을 제거 한 뒤, 영상 안에 모든 픽셀에 대한 평균값을 임계치로 하는 평균 이진화 기법을 적용하여 영상을 이진화 한다. 보다 세밀한 기호들 까지 추출하기 위해 이진화 된 영상에 모폴로지의 팽창을 적용하여 기호들을 확장시킨다. 모폴로지의 팽창 기법은 물체의 최외각 픽셀을 확장하고 배경 축소하는 기법으로 홀과 같은 빈 공간을 매우거나, 서로 짧은 거리만큼 끊어진 영역을 연결하는 경우에 사용된다[3]. 평균 이진화 된 영상과 모폴로지가 적용 된 악보 영상은 그림 2와 같다.

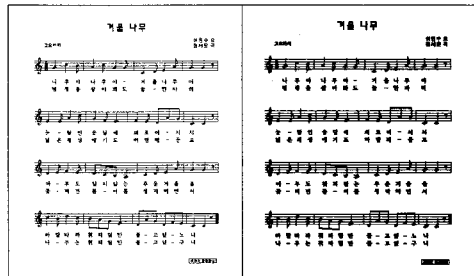


그림 2. 평균 이진화 된 악보 영상(좌)
모폴로지의 팽창이 적용 된 악보 영상(우)

모폴로지가 적용된 영상에 Grassfire 알고리즘을 이용하여 작은악절영역을 추출하고 포함되지 않는 제목, 가사, 특수 기호 등의 잡음을 제거한다. Grassfire 알고리즘은 자기호출을 이용하여 모든 인접 요소가 라벨링 될 때까지 현재 관심 화소의 주변 인접 화소를 차례로 검사하면서 라벨링하는 방법이다. 그러나 오선 영역은 악보 구성 기호 영역과 함께 추출되기 때문에, 악보 구성 기호만 추출하기 위하여 오선 영역을 제거한 후 악보 구성 기호를 추출한다. 오선은 작은악절영역에 수평 히스토그램을 적용하여 오선의 위치를 파악하고, 오선과 악보 구성 요소와의 교차점에 있는 상하 픽셀(Pixel)을 측정하여 흑화소가 없을 경우에는 오선으로 간주하여 제거한다[2,4]. 그림 3은 추출한 작은 악절영역과 오선 및 기타 잡음 영역을 제거한 영상이다.

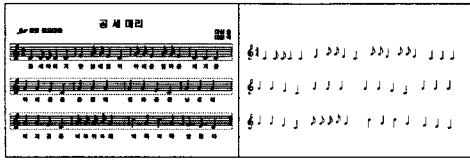


그림 3. 작은악절영역을 추출한 악보 영상(좌)
오선과 잡음 영역이 제거된 악보 영상(우)

오선과 기타 잡음 영역들이 제거된 영상에 Grassfire 알고리즘을 적용하여 악보 구성 기호들을 추출한다. 추출된 기호 중 잇단 음표들은 형태학적 특징을 이용하여 분리하여 최종적으로 악보 구성 기호들을 추출한다. 그림 4는 Grassfire 알고리즘을 적용하여 악보 구성 기호들을 추출한 영상과 추출된 악보 구성 기호 중 잇단 음표가 분리된 영상이고, 그림 5는 잇단 음표 분리한 잇단 음표영역이다.

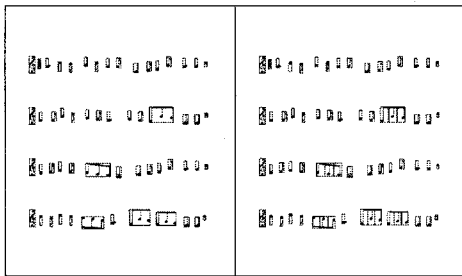


그림 4. Grassfire 알고리즘을 적용하여 악보 구성 기호들을 추출한 영상과 잇단 음표가 분리된 영상

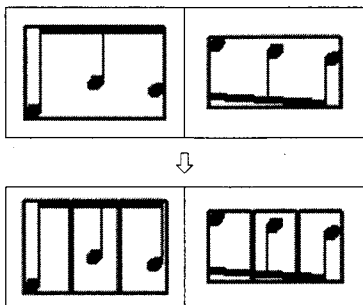


그림 5. 잇단 음표 분리

분할된 악보 구성 요소를 구분하거나 인식하기 위해 정규화를 적용 하는데, 정규화를 하게 될 경우 각각의 다른 영역 비율로 인해 악보 구성 요소의 두께가 임의로 변하는 문제점을 보완하기 위해 정규화 된 영상에 골격만을 추출하는 영상처리 기법인 세션화 방법[5,6,7]을 적용한다. 정규화를 거쳐 세션화 된 기호들은

쉽표기호와 그 외 기호, 음표 기호로 구분하여 쉽표기호와 그 외 기호들은 퍼지 ART 알고리즘을 이용하여 인식하고 음표 기호들은 형태학적 특징을 이용하여 인식한다. 음표 기호들은 수직 히스토그램을 이용하여 기둥 영역을 찾은 후에 머리 영역과 꼬리 영역을 분할한다. 분할된 음표 기호들 중 머리 영역은 오선에서의 위치를 찾아 음의 높낮이를 결정하고, 꼬리가 연결되어 있는 음표 기호들은 꼬리 영역의 형태학적 특징을 이용하여 박자를 검출한다.

III. 퍼지 ART 알고리즘을 이용한 악보 구성 요소 인식

본 논문에서는 추출된 악보 구성 요소들을 인식하기 위하여 퍼지 ART 알고리즘[8]을 적용한다. 퍼지 ART는 자율학습 방법으로 이진 신호뿐만 아니라 아날로그 신호도 처리할 수 있으며 패턴들의 특징을 안정적으로 분류할 수 있다는 장점이 있다.

기존의 ART1에서 유사성 측정은 입력 패턴과 저장 패턴인 하향 가중치와의 곱의 놈(norm)을 적용하여 식 (1)과 같이 계산된다.

T는 저장 패턴을 나타내는 하향 가중치이고 X는 입력 패턴이다.

$$\frac{\|X \cdot T\|}{\|X\|} \quad (1)$$

퍼지 ART에서 유사성 측정은 퍼지 논리 교연산자 중에서 Min 연산자를 적용하여 식 (2)과 같이 계산한다. 여기서 \wedge 는 퍼지 논리 Min 연산자이다.

$$\frac{\|X \wedge W\|}{\|X\|} \quad (2)$$

출력값(O_j)은 식 (3)와 같이 계산하고 식 (4)를 이용하여 가장 큰 출력값을 가진 노드를 승자 노드로 선택한다. O_j 는 출력값이고 O_j 는 j 번째 승자노드의 출력값이다. 그리고 α 는 0과 1사이의 선택파라미터(choice parameter)이다

$$O_j = \frac{\|X \wedge W\|}{\alpha + \|W\|} \quad (3)$$

$$O_j = \vee(O_j) \quad (4)$$

퍼지 ART 알고리즘에서 가중치 W 를 조정하는 식은 다음과 같다.

$$W(t+1) = \beta(X \wedge W(t)) + (1-\beta)W(t-1) \quad (5)$$

퍼지 ART 알고리즘에서는 가중치를 조정할 때 적용되는 학습 파라미터 β 는 0과 1사이의 경험적인 값으로 설정한다.

악보 영상에서 추출된 쉽표와 그 외 기호들을 인식하기 위해 적용된 퍼지 ART 알고리즘은 그림 6과 같다.

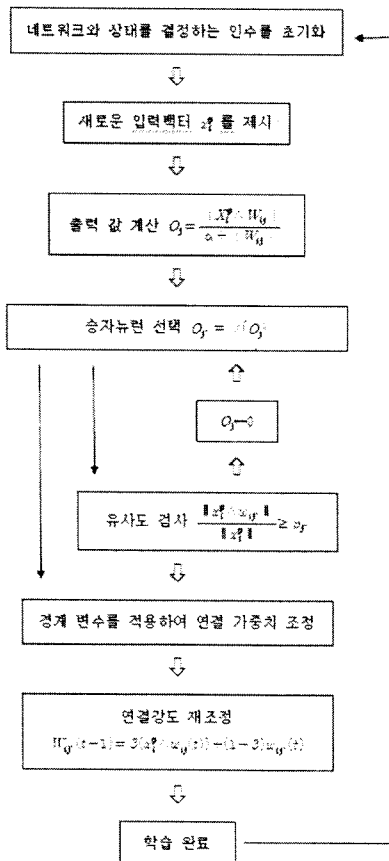


그림 6. 퍼지 ART 알고리즘

본 논문에서는 퍼지 ART 알고리즘을 적용하여 쉽표 기호와 그 외의 기호들을 인식하고, 음표 기호의 형태학적 특징을 이용하여 인식한 후, 음표 기호의 머리영역을 이용하여 음의 높

낮이에 알맞은 소리가 나오도록 연주한다.

IV. 실험 및 결과 분석

실험 환경은 Intel Pentium-IV 2.8GHz CPU 와 512MB RAM이 장착된 IBM 호환 PC상에서 VC++ 6.0으로 구현하였다. 본 논문에서는 인쇄 악보의 인식을 위하여 50장의 악보 영상을 스캔하여 획득된 악보 영상을 대상으로 실험하였다. 실험 방법으로는 음표 기호의 경우에는 음의 높낮이를 이용하여 그에 알맞은 소리가 나오게 하여 인식하였고, 쉽표 기호와 그 외의 기호들은 본 논문에서 제안한 퍼지 ART 알고리즘을 적용하여 50개의 악보 영상에서 추출한 악보 구성 기호 30개를 학습 패턴으로 적용하였다. 음표 기호는 2989개 중 2740개가 알맞게 인식되어 정확히 연주되었고, 쉽표 기호와 그 외의 기호 코드들은 463개 중 407개가 인식되었다. 표 1은 음표 기호의 인식과 쉽표 기호와 그 외의 기호 코드들의 인식을 나타내었다.

표 1. 악보 구성 기호 인식 결과

	음표	쉽표와 그 외 코드
인식개수	2740 / 2989	407 / 463

악보 구성 기호의 인식에 실패한 경우는 추출한 음표 기호들이 함께 추출된 경우와 제안한 기준의 범위를 넘어서 소리를 낼 수 없는 경우의 음표 기호이다. 그림 7은 개별 문자 인식에 실패한 악보 구성 기호를 나타내었다.

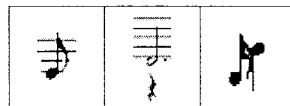


그림 7. 인식에 실패한 악보 구성 기호

VI. 결 론

본 논문에서는 악보 구성 기호를 추출하고 인식하는 방법을 제안하였다. 악보 구성 기호의 추출은 불필요한 컬러 잡음들을 제거한 영상을 그레이 레벨로 변환하고 평균 이진화 방법을 적용하여 이진화 하였다. 이진화 된 영상에

모폴로지의 팽창 기법을 사용하여 객체를 연결시킨 후, Grassfire 알고리즘을 이용하여 작은악절 영역을 추출하였다. 추출된 작은악절 영역 이외의 제목, 가사 등은 잡음으로 간주하여 제거하였고, 가로 히스토그램을 이용하여 오선을 제거하였다. 잡음과 오선이 제거된 영상에 Grassfire 알고리즘을 이용하여 작은악절영역에 포함되는 악보 구성 기호들을 추출하였다. 추출된 기호들 중 잇단 음표는 서로 분리한 후 최종적으로 기호들을 추출하였다. 추출된 기호들은 정규화를 거쳐 세션화 하였고 악보 구성 기호들을 최종적으로 추출하였다. 추출된 악보 구성 기호들은 형태학적 특징과 퍼지 ART 알고리즘을 적용하여 인식 하고 연주하였다.

향후 연구 과제로는 비교적 기호들의 구성이 간단한 악보의 경우에는 충분한 인식 결과를 얻을 수 있었지만 일반적으로 시중에서 사용되는 악보는 다양한 구조를 가지므로 다양한 구조를 가지는 악보를 추출하고 인식하는 방법에 대해 연구할 것이다.

참고문헌

- [1] 박충식, 장경식, 전정우, 김재희 등, “악보인식을 위한 전처리 과정,” 인공지능, 신경망 및 퍼지관련 학술발표회 논문집, pp.169- 175, 1993.
- [2] 김미정, 김재군, 박충식, 김광백, Hierarchical ART2 알고리즘을 이용한 악보 영상 인식, 한국해양정보통신학회추계학술대회 논문집, 12권 1호, pp.369-374, 2008.
- [3] Harely R. M., Arthur R. W., Computer Imaging Recipes in C, PTR Prentice Hall, 1993.
- [4] 손화정, 김수형, 오성열, “카메라 기반 악보 영상 인식을 위한 오선 검출 및 삭제 알고리즘,” 한국콘텐츠학회논문지, 제 7권 11호, pp. 34-42, 2007.
- [5] 조형재, 조경은, “골격선과 런 길이 정보를 이용한 피아노 악보 인식,” 정보과학회논문지, 제2권, 4호, pp.461-473, 1996.
- [6] 성기중, 유근호, 이철희, “이진영상의 새로운 세션화 알고리즘,” 대한전자공학회 학술대회 논문집, pp.289-291, 1985.
- [7] 한동협, 조성배, “오프라인 필기한글을 위한 지식기반 세션화 알고리즘”, 한국정보과학회 정보과학회논문지, pp.1381-1388.
- [8] 김광백, 조재현, “퍼지 신경망을 이용한 자동차 번호판 인식 시스템”, 한국컴퓨터정보학회 논문지, pp. 313 ~ 319, 2007.