# Backpropagation을 이용한 악보인식

박현준* · 차의영*

## Recognition of Music using Backpropagation Network

Hyun-Jun Park* · Eui-Young Cha*

요 약

본 논문에서는 신경회로망 알고리즘 중 하나인 backpropagation network을 이용한 악보인식 기법과 그에 필요한 악보 영상에 대한 전처리 기법을 제안한다. 전처리과정으로 이진화, 기울기 보정, 오선제거 등의 과정을 수행하여 인식에 필요한 음악 기호와 음표를 분리한다. 분리된 음악 기호와 음표들은 backpropagation 알고리즘을 사용하여 구성된 음표 인식 신경망과 비음표 인식 신경망을 통해 각각 음표와 비음표 인식과정을 거친다. 다양한 복잡도를 가진 악보를 대상으로 한 실험 및 분석 결과를 통해 제안한 악보 인식 기법의 정확도를 기술하였다.

ABSTRACT

This paper presents techniques to recognize music using back propagation network, one of the neural network algorithms, and to preprocess technique for music image. Music symbols and music notes are segmented by preprocessing such as binarization, slope correction, staff line removing, etc. Segmented music symbols and music notes are recognized by music note recognizing network and non-music note recognizing network. We proved correctness of proposed music recognition algorithm through experiments and analysis with various kind of musics.

키워드

악보(Music), 악보인식(Music Recognition), 오선제거(Staff line removing), 신경망(Neural network)

## Ⅰ. 서 론

현재 국내에는 많은 시청각적인 매체들이 발달, 초고속 통신망의 대중화가 이루어졌다. 정보의 바다라 하는 인터넷의 접근성이 좋아지고, 어렵지 않게 많은 음악을 접할 수 있게 되었다. 그에 따라 자연스럽게 음악 문화가 발달하게 되었고, 여러 음악적인 문화가 발달함에 따라서 음악에 대한 대중의 관심이 높아졌다. 음악에 대한 관심은 대중의 음악적 수준을 향상시켰다.

이러한 움직임은 음악에 대한 전문적인 지식을 가진 이들만이 아니라 일반인들도 악보 편집을 하는 경우가 늘어나게 만들었다. 전문가들이 아닌 일반인들의 경우, 이미 존재하는 악보를 수정하는 작업을 많이 하게 되는 경향이 있고, 그에 따라 이미 존재하는 악보 인식의 필요성이 증가하였다.

새로운 악보의 생성, 또는 기존 악보의 편집 등의 작업에 대해서 예전과 달라진 점이 있다. 컴퓨터의 발달과 음악 작업을 좀 더 쉽고, 빠르고, 정확하게 할 수 있도록

* 부산대학교 컴퓨터공학과
** 부산대학교 컴퓨터공학과 교수

접수일자 : 2007. 6. 1

도와주는 많은 악보 편집 툴이 개발됨으로, 과거와 달리 현재 작성되는 악보의 대부분은 손으로 그려진 악보가 아닌, 악보 편집 프로그램을 사용하여 만들어진, 비교적 정형화된 악보가 만들어진다는 점이다. 이는 컴퓨터를 이용한 악보인식에 있어서 고무적인 일이다.

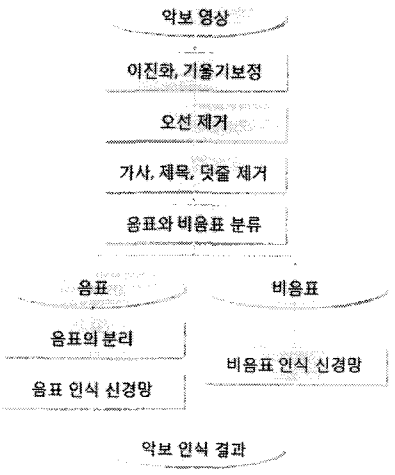


그림 1. 전체적인 악보 인식 단계
Fig 1. Music Recognition steps

90년대 초반부터 악보인식에 대한 많은 연구가 진행되었지만, 만족할 만큼의 악보 인식률을 보이지는 못하였고, 현재의 기술을 실제 악보에 적용 시 인식률이 많이 떨어진다. 실제로 상용화된 악보 인식 프로그램이 존재했었지만, 성공하지 못한 것은 인식률이 사용자들의 요구를 만족시키지는 못하였기 때문으로 판단된다. 이는 과거의 손으로 그려진, 정형화 되지 않은 악보를 대상으로 연구를 진행하였기 때문일 것이다.

본 논문에서는 인공신경망 알고리즘 중 하나인 backpropagation 학습 알고리즘을 이용한 악보 영상 인식 방법을 제시한다. 인공신경망을 이용한 방법은 기존의 형판(template) 매칭 방법이나 통계적(statical) 방법, 구조적(structural) 방법에 비해서 잡음에 의한 입력 패턴의 변형에도 잘 적응하고, 병렬처리가 가능하므로 처리 속도를 향상시킬 수 있다.

본 논문에서 제시하는 악보 인식 방법은 다음과 같다. 먼저 악보 영상의 이진화와 기울기보정을 하고, 오선에 의해 연결되어 있는 각 음표와 비음표가 하나로 인식 되는 것을 방지하기 위해 오선을 제거하고, 그 외의 인식 대상이 아닌 것들을 제거한다. 오선이 제거된 악보 영상에서 각각의 음표와 비음표를 분류하여 연결된 음표를 분리한 뒤, 음표의 머리 위치 정보를 이용하여 음표의 음 인식을 수행하고 인공신경망을 이용하여 박자를 인식한다.

본 논문의 2장에서는 악보 영상의 전처리과정에 대해 기술한다. 3장에서 backpropagation 학습 알고리즘의 적용에 대해 기술하고, 4장에서는 인식 결과를 비교, 분석하고 마지막 5장에서는 결론과 향후 연구 방향을 제시한다.

## Ⅱ. 악보 인식을 위한 전처리과정

Backpropagation network을 이용한 악보 인식을 위해 먼저 각각의 음악 기호들과 음표를 분리할 필요가 있다. 그러므로 영상의 이진화, 기울기보정, 오선제거, 덧줄 제거, 연결된 음표의 분리, 음표와 비음표의 구분, 음표의 머리 위치 찾기 등의 전처리과정을 거쳐야 한다.

### 2.1. 오선 및 불필요 영역 제거

악보 인식 방법에 따라서 오선을 제거 할 수도 있고 제거 하지 않을 수도 있다. 오선을 제거하면 오선을 제외한 다른 음악기호들이 쉽게 분리되기 때문에 인식을 용이하게 수행할 수 있다. 그러나 오선을 제거할 때 음악기호와 오선이 겹친 부분을 구별하여 오선 선분만을 제거하는 데는 어려움이 따른다.

오선 외에도 악보의 연주에 있어서 필요하지 않은 부분이 존재한다. 가사, 제목, 덧줄 등을 제거함으로 악보 인식에 필요한 연산과정을 줄이고, 좀 더 정확한 악보인식을 수행할 수 있도록 한다.

① 오선 제거 알고리즘

먼저 영상에서 오선의 위치를 찾는다. 오선은 가로의 긴 선으로 이루어진다. 이를 이용하여 horizontal projection기법을 사용한다.

Projection 결과가 영상의 가로 크기의 일정 이상인 곳을 오선으로 가정한다. 만약 너무 큰 값을 잡으면 노이즈가 많거나 기울어진 악보의 경우 오선을 인식하지 못하는 경우가 생기며, 너무 작은 값을 잡으면 오선이 아닌 음표나 기호가 밀집되어 있는 부분을 오선으로 오인식

하는 경우가 생길 수 있다. 본 논문에서는 악보 영상 가로 크기의 50%이상인 곳을 오선으로 가정하였다.

찾은 오선을 제거하기 위해 두 가지 방법을 제안한다. 첫 번째 방법은 오선 위치의 상하 일정영역에 대해 vertical projection을 수행한다. 오선에 다른 음표나 기호가 존재하는 경우, 오선의 두께 이상의 값을 가지게 된다. 이 점을 이용하여, vertical projection값이 일정 값 이하인 경우 오선만 존재하는 것으로 가정하고, 그 부분을 삭제를 한다. 두 번째 방법으로 오선 위치의 일정 영역에 대해 median filter를 적용하는 것이다. 제시된 두 가지 방법은 모두 만족할만한 결과를 보였다. 본 논문에서는 median filter를 이용한 오선 제거를 하였다.

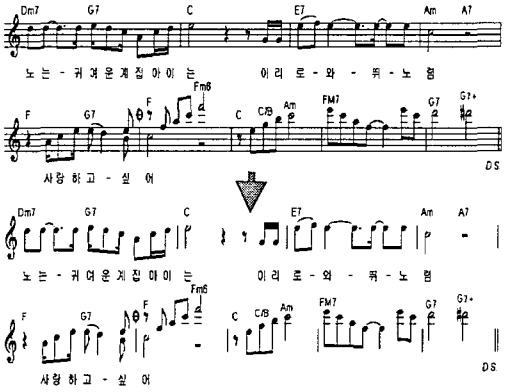

그림 2. 오선 제거 결과
Fig 2. Result of removing the staffs

② 제목 및 가사 제거 알고리즘

제목과 가사는 사람이 악보를 인식함에 있어서 중요한 정보이다. 하지만 본 논문의 인식 대상은 아니므로 제거함으로써 인식 속도와 정확도를 높일 수 있도록 한다.

이는 악보 내에서의 위치적 특성을 이용한다. 제목과 가사는 항상 오선사이 또는 악보의 가장 위, 아래에 존재한다. 그러므로 각 오선의 상하 일정영역을 제외한 부분을 삭제 대상 범위로 지정하고, 삭제한다.

③ 덧줄 제거

덧줄은 사람이 음을 인식하기에는 용이하나, 신경회로망을 이용하여 음표를 인식할 경우, 노이즈로 작용하여 신경회로망의 인식률을 떨어뜨리는 원인이 될 수 있다.

오선의 간격은 비교적 일정하고 덧줄의 위치는 오선의 간격만큼의 차이를 두고 위치한다는 특성을 이용하면, 덧줄의 존재 가능한 위치를 예상할 수 있고, 그 부분에 덧줄이 존재하면 삭제한다.
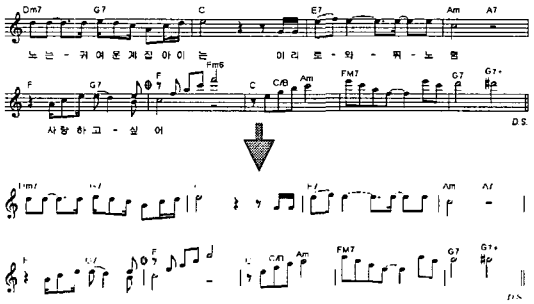
그림 3. 전처리과정 후 악보 영상
Fig 3. Music image after preprocessing

### 2.2. 음표와 비음표의 분류

악보 인식에 있어서 음표와 비음표를 분류하면, 음표의 박자를 인식하기 위한 신경회로망을 간단히 할 수 있다. 인식 대상이 늘어나면 신경회로망의 학습 대상이 늘어나고 연산량이 기하급수적으로 늘어나기 때문이다. 분류 방법으로는 신경회로망을 이용하는 방법, 음표의 특징을 이용하는 방법이 있다. 전자의 경우, 신경회로망을 하나 더 구성해야하고, 학습 대상이 많아지므로 연산시간이 늘어나는 단점이 있다.

① 음표의 특징

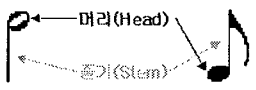


그림 4. 음표의 구성
Fig 4. Composition of the music score

온음표( : 4박자동안 같은 음을 연주)를 제외한 음표는 항상 줄기(Stem)을 가지고, 비교적 일정한 크기를 가진다. 음표 머리의 세로 크기는 항상 오선의 간격과 같으며, 줄기는 오선의 간격의 3배 이상이 된다. 가로크기는 꼬리가 없는 음표가 최소크기를 가지며, 이는 오선 간격의 1.5배정도의 크기를 가진다.

### 2.3. 음표의 분리와 머리 위치 찾기

하나 이상의 음표가 하나로 묶여서 표현된 경우 backpropagation network을 이용한 인식에 어려움이 있

다. 음표를 분리 하여, 신경회로망에 각각의 음표 하나씩을 인식대상으로 만듦으로써 인식률을 높일 수 있고, 학습 시간을 단축시킬 수 있다.

음표의 머리 위치는 오선에서 음의 높이를 표현한다. 그러므로 악보 인식을 위해서 반드시 음표의 머리 위치를 찾아야 한다.

① 음표의 분리

온음표를 제외한 모든 음표는 줄기(Stem)가 존재한다. 온음표는 분리가 필요하지 않은 음표이므로, 분리 대상인 모든 음표는 반드시 하나 이상의 줄기를 가지고 있다. 음표를 분리해야 하는 개수와 줄기의 수는 항상 동일하다. 줄기들의 위치 정보를 이용하여 줄기의 좌우 일정 영역을 잘라내고, 새로운 음표로 만들어내는 방법을 사용하였다.

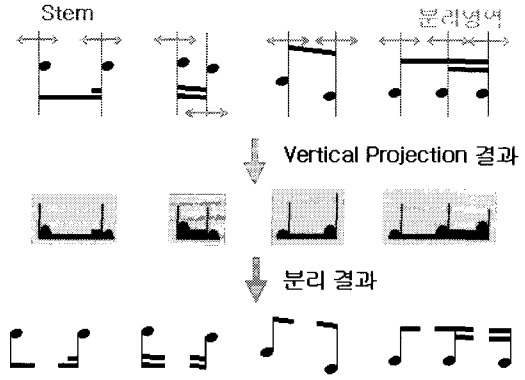


그림 5. 음표 분리 알고리즘
Fig 5. Music score separation algorithm

② 음표 머리 위치 찾기

음표 머리는 항상 줄기의 양쪽 끝에 위치한다. 그러므로 줄기의 양쪽 끝을 projection하여 머리의 존재 여부를 판단하고, 위치를 찾는다.

그림 6. 음표 머리 찾기 알고리즘
Fig 6. Music score head finding algorithm

### 2.4. 음표의 음인식 알고리즘

음의 높이 인식은 음표 머리와 오선의 상대적인 위치를 이용한다. 오선의 각 선들의 위치를 직접 비교하는 방법, 중심선 하나를 선택하여 중심선과의 상대적인 위치를 이용하는 방법을 제시한다. 전자의 경우, 오선 내에서의 음 인식률은 좀 더 나은 결과를 보였다. 하지만 오선 사이의 간격이 일정하지 않은 악보와 덧줄에 위치하는 음표의 경우, 오히려 더 좋지 않은 성능을 보였다. 후자의 방법은 오선 내에서의 인식률은 조금 떨어질 수 있으나, 전자에 비해서 노이즈나 오류, 덧줄에 위치한 음표의 음 인식에 강하다.

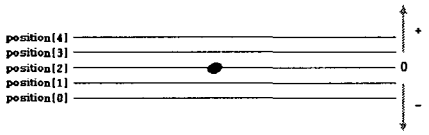


그림 7. 음표의 음 인식 방법
Fig 7. Tone recognition method

## Ⅲ. Backpropagation 학습 알고리즘을 이용한 악보 인식

전처리과정을 수행하면서 각 기호와 음표의 모양은 불완전할 수밖에 없다. 신경회로망은 입력이 주어지면 그 입력을 스스로 조정하여 일치 되는 답을 형성할 수 있도록 학습을 할 수 있다. 이는 왜곡되거나 noise를 가진 입력에 대해서도 바른 인식이 가능하도록 하고(일반화), 불완전한 입력에서 이상적인 것의 추출이 가능(추상화)하다. 이런 특징이 있기에 악보 인식에 신경회로망을 사용가능하게 한다.

본 논문에서는 문자나 물체 인식 등에 좋은 성능을 보이고 있는 backpropagation 학습 알고리즘을 사용하였다. 다층 구조를 가지는 알고리즘으로, 하나의 24개의 뉴런을 가지는 은닉층을 두었다. 입력층은 24 × 24의 영상, 총 576개의 입력을 받고, 출력층은 인식대상의 수로 20개의 출력을 가진다.

본 논문에서는 아래의 입력패턴으로 음표 인식 신경망과 비음표 인식 신경망을 구성하였고, 최대허용오차를 0.05%로 학습을 수행하였다.

그림 8. 인식대상
Fig 8. Recognition target

## Ⅳ. 실험 결과

표 1은 제안된 악보인식 기법의 정확도를 분석하기 위해 가우시안 노이즈를 추가한 악보와 복잡함이 다른 악보에 대한 실험결과이다.

표 1. 정확도 실험 결과
Table 1. Result of the Correctness test

| 악보상태 / 인식률 | 노이즈 10% | 노이즈 30% | 단순한 악보 | 복잡한 악보 |
|---|---|---|---|---|
| 음 높이 | 95.78% | 90.32% | 93.75% | 78.12% |
| 2분 음표 | 88.82% | 80.00% | 94.07% | 69.75% |
| 4분 음표 | 88.75% | 85.56% | 90.90% | 65.83% |
| 8분 음표 | 91.36% | 86.27% | 96.10% | 81.24% |
| 16분 음표 | 94.28% | 86.67% | 94.33% | 72.96% |
| 비 음표 | 77.18% | 75.25% | 77.25% | 68.75% |

실험결과 노이즈보다 악보의 복잡함에 더 민감하게 반응함을 확인할 수 있었다. 잘 분리된 음표는 신경회로망에 의해서 정확히 인식됨을 확인할 수 있었고, 그러므로 이는 악보가 복잡해짐에 따라 정확한 세그멘테이션을 수행하지 못하여 복잡한 음표들을 정확히 분리하지 못한 결과였다.

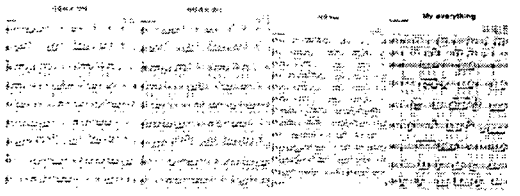

그림 9. 실험에 사용된 악보 영상
Fig 9. Music images used in experiment

## Ⅴ. 결 론

정보 통신의 발달로 인한 음악 문화의 발달은 일반인들이 기존의 악보를 편집하고자하는 욕구를 증가시켰고, 그로인해 자동화된 악보 인식의 필요성이 높아졌다.

본 논문은 악보 영상을 인식하기 위한 전처리방법과 backpropagation network을 적용한 악보 인식방법을 제시하였다.

실험결과 악보 영상의 노이즈의 정도, 전처리과정의 정확도, 악보의 복잡함의 정도에 따라서 차이가 많았다. 노이즈가 적고, 비교적 덜 복잡한 악보 영상의 경우 만족할 만큼의 인식이 가능했다. 역으로 아직은 복잡하고 노이즈가 많은 악보에 대한 정확한 인식은 할 수 없었다.

향후 악보의 복잡함에도 강한 전처리 방법과 악보 편집 프로그램의 특성을 잘 분석하여, 정확한 전처리과정이 수행 된다면 악보 인식이라는 문제를 해결함에 많은 도움이 될 것이다.

## 참고문헌

[1] 박충식, 장경식, 전정우, 악보 인식을 위한 전처리 과정, 대한전자공학회, 대한전자공학회 학술발표회 논문집, pp.168~175, 1993. 1.

[2] 이성기, 신재욱, 신경망을 이용한 악보인식, 한국정보과학회, 정보과학회논문지 제21권 제7호, pp.1358~1366, 1994. 7.

[3] S. Marinai, P. Nesi, Projection Based Segmentation of Musical Sheets, IEEE Computer Society, pp.515~519, 1999

[4] H.Miyao, and Y.Nakano, "Note Symbol Extraction for printed Piano Scores Using Neural Networks", IEICE Trans. & Syst. vol.E79-D, no.5, pp.548~554, May. 1996.

[5] Sicard. E. "An Efficient Method for the Recognition of Printed Music", Pattern Recognition, 1992. Vol.III. pp.573~576, 1992

## 저자소개

**박 현 준(Hyun-Jun Park)**

2007년 부산대학교 정보컴퓨터공학과 졸업.
2007년 부산대학교 컴퓨터공학과 석사 재학 중.
※관심분야: 패턴인식, 영상처리, 컴퓨터비전, 인공신경망.

**차 의 영(Eui-Young Cha)**

1979년 경북대학교 전자공학과 졸업.
1982년 서울대학교 전자계산학과 석사 졸업.
1998년 서울대학교 컴퓨터공학과 박사졸업.
1981년~1985년 한국전자기술연구소 연구원.
1995년~1996년 University of London 방문교수.
1985년~현재 부산대학교 컴퓨터공학과 교수
※관심분야: 패턴인식, 영상처리, 컴퓨터비전, 인공신경망.