논문 2011-3-21

나이브 베이지안 분류기를 이용한 판소리 분류 프로그램 구현

An Implementation of Pan-So-Ri Classification Program Using Naive Bayesian Classifier

김원종*, 이강복*, 김명관**

Won-Jong Kim, Kang-Bok Lee, Myung-Gwan Kim

요 약 판소리는 이야기를 노래로 부르는 우리나라의 전통음악 형식 중 하나로 두 가지 유파(동편제, 서편제)로 나누어진다. 판소리에 대한 지식이 없는 사람은 판소리를 듣고서 이 두 가지 유파를 구별해내기 어렵다. 본 논문에서는 PCD(Pitch Class Distribution)와 나이브 베이지안 분류기를 이용한 판소리 분류 프로그램 구현 과정을 기술한다. 분류기에 사용되는 속성값으로는 각 음계의 출현빈도를 이용하였다. 실험은 확률값을 반올림한 위치를 다르게 하여 두 번 실행하였으며, 그 중 보다 뛰어난 결과로 동편제를 80%, 서편제를 97%, 총 88%의 정확도로 올바르게 분류해 내는 것을 알 수 있었다. 구현한 프로그램에는 이 결과를 적용하였다.

Abstract Pan-So-Ri singing a story as song is one of Korea traditional musics. it divide into two sect(east-sect, west-sect), and it is hard to classify two sect without knowledge about Pan-So-Ri. In this paper, we have propose a Pan-So-Ri classification program using PCD(Pitch Class Distribution) and Naive Bayesian Classifier. Attribute value of classifier is each appearance frequency of pitch. Experiment is conducted two time with different rounding off location of probability value. Better one show correct classification with east-sect 80%, west-sect 97%, and total accuracy of 88%. this result is used our program.

Key Words : 판소리, 나이브 베이지안 분류기(Naive Bayesian Classifier), 데이터마이닝(Data Mining), PCD(Pitch Class Distribution), 지도학습(Supervised Learning)

Ⅰ. 서 론

판소리는 18세기에 우리나라에서 발생하여 현재 까지 전승되어 온 음악 예능의 한 형태이다. 부채를 든 한 명의 창자(唱者)가 고수의 북장단에 맞추어 창(노래. 소리)・아니리(말)・너름새(몸짓)를 섞어가며 이야기를 엮어가는 극적 음악이다.[1][2] 한국 향토의 선율을 토대로 여러 가지 장단에 따라 변화시킨 것으로, 크게 두 가지 유파(동편제, 서편제)로 나누어진다.

판소리에 대한 지식이 없는 일반인이 판소리를 듣게 되면 판소리의 두 가지 유파를 구별해내기 어렵다.

본 논문에서는 판소리 분류 프로그램 구현을 위하여 Chordia의 인디언 음악 분류 논문[3][4]을 참고하였다.

음악의 음계 분산을 도시하는 PCD(Pitch Class Distribution)[3][4][5]와 조건부 확률에 기초하는 분류 알고리즘인 나이브 베이지안 분류기를 이용하여 판소리 분류 프로그램 구현 과정을 기술한다.

*준회원 을지대학교 의료IT마케팅학과
**정회원, 을지대학교 의료IT마케팅학과 (교신저자)
접수일자 2011.2.15, 수정일자 2011.5.7
게재확정일자 2011.6.10

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구를 알아보고, 3장에서는 구현한 프로그램의 설계 과정, 4장에서는 구현한 프로그램의 실험 및 결과를 기술하고, 5장에서는 결론을 맺는다.

II. 관 련 연 구

1. 판소리

판소리는 전승 지역, 전승 계보, 음악적 특성의 세 가지 기준으로 동편제, 서편제, 중고제의 3가지 종류로 나뉜다.[1]

그런데 중고제의 경우, 일제강점기 이후로 전승이 끊어져 존재하지 않는다.[1][6] 뿐만 아니라, 음악적 특성에 관해서도 명확한 정의를 내리지 못하기 때문에 그 실체나 개념이 실제로 존재하고 있다고 보기는 어렵다.[1]

이러한 이유로, 본 논문에서는 분류하는 유파로 동편제, 서편제만을 다루기로 하였다. 그리고 실험에서 사용되는 판소리들의 유파(동편제, 서편제)를 구분하는 데에는 판소리 전승계보[1]를 이용하였다.

가. 동편제

지리산을 중심으로 경상도 억양에 맞게 형성된 소리제.[7] 섬진강 동쪽 지역인 남원・순창・곡성・구례 등지에 전승된 소리로서, 가왕으로 일컬어지는 운봉 출신의 송흥록의 소리 양식을 표준으로 삼는다. 우조(씩씩한 가락)의 표현에 중점을 두고, 감정을 가능한 절제하며, 장단은 '대마디 대장단'을 사용하여 기교를 부리지 않는다. 발성은 통성을 사용하여 엄하게 하며, 구절 끝마침을 되게 끊어낸다.[1]

나. 서편제

섬진강 서쪽 지역인 광주・나주・담양・화순・보성 등지에 전승된 소리로, 순창 출신이며 보성에서 말년을 보낸 박유전의 소리 양식을 표준으로 삼는다. 계면조(슬픈 가락)의 표현에 중점을 두며, 발성의 기교를 중시하여 다양한 기교를 부린다. 소리가 늘어지는 특징을 지니며, 장단의 운용 면에서는 엇부침이라 하여, 매우 기교적인 리듬을 구사한다. 또한 발림(육체적 표현. 동작)이 매우 세련되어 있다.[1]

2. 나이브 베이지안 분류기

가상 공간 C와 주어진 트레이닝 데이터 D로부터 가장 좋은 가설을 구하고자 할 때 다음과 같은 베이즈 정리를 사용할 수 있다.[8] 식 (1)은 베이즈의 정리를 나타낸다.

가. 베이즈의 정리[9][10]

$$P(C|D) = \frac{P(D|C)P(C)}{P(D)} \quad (1)$$

P(C | D) : 사건 D가 발생한 상태에서 사건 C가 발생할 확률

P(D | C) : 사건 C가 발생한 상태에서 사건 D가 발생할 확률

P(C) : 어떠한 사전 정보도 없는 상태에서 사건 C가 발생할 확률

P(D) : 어떠한 사전 정보도 없는 상태에서 사건 D가 발생할 확률

나. 나이브 베이지안 분류기

나이브 베이지안 분류기는 베이지안 학습방법 중에서 널리 쓰이는 통계적 알고리즘이다. 나이브 베이지안 분류기는 속성 값들의 결합으로 이루어진 각 인스턴스 x와 특정 유한집합 V에서 어떤 값을 갖는 목적함수 $f(x)$가 존재하는 학습 업무에 사용된다.

나이브 베이지안 분류기는 속성 값들이 주어진 목적값에 조건부 독립적(Conditionally Independence)이라는 가정을 기반으로 한다.[8][11] 실제 적용시 이 가정은 맞지 않는다. 특정 위치에서 '분류기'라는 단어가 나올 확률보다 앞의 단어가 '베이지안'일 경우에 그 다음 단어가 '분류기'가 될 확률이 더 크기 때문이다. 하지만, 실제 응용시 위치 독립적이라는 가정으로도 베이지안 분류기는 정상 동작 한다.[8] a1, a2,..., an의 결합 확률이 각 속성들의 곱으로 계산될 수 있다는 가정을 식 (1)에 대입하면 나이브 베이지안에 사용할 공식을 구하게 된다.[9][10][12]

$$v_{NB} = \arg max P(v_j) \prod_i P(a_i|v_j) \quad (2)$$

여기서 v_{NB}(Naive Bayesian)는 나이브 베이지안 분류기가 출력하는 목적값을 나타낸다.

3. PCD(Pitch Class Distribution)[3][4][5]

Chordia가 인디언 음악 분류를 위해 사용한 PCD는 음

의 높낮이 분류 방법으로 실시간으로 연주되는 음악 내의 진동수를 이용하여 인디언 전통음악의 음계를 12가지 음으로 분류 하였다.

실시간으로 연주되는 음악을 토대로 분류해 낸 12개의 음의 분포를 이용, 기존 학습된 인디언 음악과의 음계 출현 횟수를 비교, 분석하여 연주되는 음악이 무엇인지 구별한다. 이와 같은 원리는 본 논문에서도 동일하게 적용된다. 그림 1은 인디언 음악 분류를 위해 사용된 PCD[3]를 나타내는 그림이다.

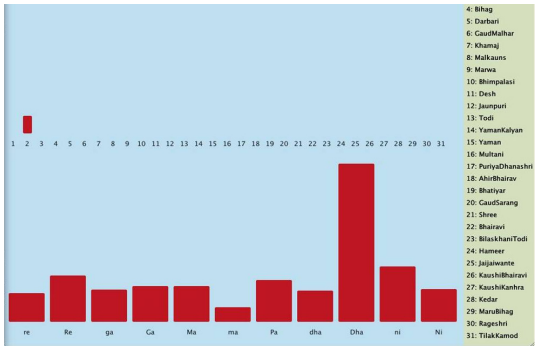

그림 1. 인디언 음악 분류에서 사용된 PCD
Fig. 1. PCD to indian music classification

Ⅲ. 나이브 베이지안 분류기를 이용한 판소리 분류 프로그램 설계

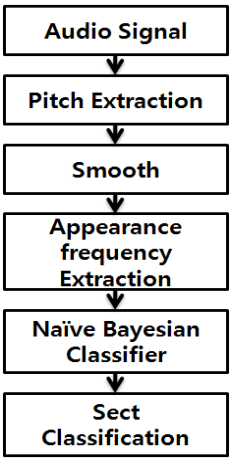

그림 2. 판소리 분류 프로그램의 블록 다이어그램
Fig. 2. Block diagram of Pan-So-Ri classification program

그림 2가 의미하는 것은 프로그램의 블록 다이어그램이다. 구현에 사용된 학습, 실험 데이터인 판소리 음악 파일은 모두 진동수 22050Hz, 1채널 모노 형태의 wav파일을 대상으로 하였다.

판소리 음악 파일을 입력받고 각 음계의 초당 진동수를 이용하여 서양 7음계 세 개의 옥타브(C4~B6), 총 21개 음계를 추출한다.

음계는 입력된 파일의 진동수를 1/10초 단위로 구한 후, 10을 곱하는 방법으로 초당 진동수를 구하여 추출하였다. 진동수를 구하는 방법은 판소리 진폭의 영교차(zero-crossing) 횟수를 구하는 것이다.[13][14] 영교차란 부호가 바뀌는 것을 뜻하는데, 소리에서도 진폭에 0이 존재해 진폭이 0일 때를 기점으로 부호가 바뀌게 된다. 따라서 진폭이 0일 때를 교차하는 횟수를 구하면 된다.

각 음계는 고유의 진동수 영역을 가지고 있는데, 위에서 구한 수치를 각 음계별 진동 영역과 비교하여 각 음을 추출해 낸 것이다. 표 1은 본 논문에서 사용한 세 개 옥타브의 음계별 진동 영역을 나타낸다. 이 진동영역은 참고문헌 [15][15]을 기초로 하였다.

표 1. C4 ~ B6의 음계별 진동영역
Table. 1. Hertz area of C4 ~ B6

음계	진동 영역(Hz)
C4	254 ~ 277
D4	278 ~ 311
E4	312 ~ 338
F4	339 ~ 370
G4	371 ~ 415
A4	416 ~ 466
B4	467 ~ 508
C5	509 ~ 554
D5	555 ~ 622
E5	623 ~ 678
F5	679 ~ 740
G5	741 ~ 831
A5	832 ~ 933
B5	934 ~ 1016
C6	1017 ~ 1110
D6	1111 ~ 1246
E6	1247 ~ 1357
F6	1358 ~ 1481
G6	1482 ~ 1663
A6	1664 ~ 1867
B6	1868 ~ 2033

한 개 옥타브가 아닌 세 개의 옥타브를 추출하는 이유는, 판소리는 음계의 개념이 존재하지 않는 점과, 고유의

음악적 특성이 저음과 고음을 크게 넘나들어 한 개의 옥타브만 추출 할 경우 소리의 높낮이의 최저음과 최고음을 추출해 낼 수 없게 되기 때문이다. 또한, 3옥타브의 범위를 C4~B6로 정한 이유는 실험에 사용된 판소리 파일들의 음계를 도시해 본 결과, 거의 대다수의 음이 이 범위에 속하기 때문이다.

음계 추출 중 고수의 북소리와 창자의 소리가 섞여 나오는 소리 등 비정상적으로 높은 음계가 추출 될 수 있는데, 이러한 경우 그 음계를 삭제(Smoothing)한다. 음계 삭제 과정으로는 전후 음계의 연속성을 고려하여 5음계 이상 차이가 나게 될 경우 비정상적인 음계로 가정하여 삭제한다.

이렇게 추출된 각 음계들의 출현 빈도를 구하고, 이 출현 빈도를 이용하여 나이브 베이지안 분류기에 적용시킨다. 식 (3)은 나이브 베이지안 분류기를 적용시킨 판소리 분류식이다.

$$v_{NB} = \arg max P(\mathrm{sect}) \prod_{i=1}^{21} P(appearance_i | \mathrm{sect}) \tag{3}$$

식 (3)에서, $P(\mathrm{sect})$는 특정 유파가 선택될 확률을 나타내는데, 동편제, 서편제의 두 가지 유파가 존재하기 때문에 1/2로 고정된다. $P(appearance_i|\mathrm{sect})$는 특정 유파에서 한 음계의 출현빈도가 선택될 확률을 나타낸다.

이 계산 과정에서 실험 데이터의 한 음계 출현빈도가 어떠한 학습 데이터와도 일치하지 않을 경우(실험 데이터와 학습 데이터 간에 일치하는 값이 없는 경우), 계산 과정에서 0으로 처리되어 $P(appearance_i|\mathrm{sect})$의 값이 0이 나오는 경우가 있다. 그런데 나이브 베이지안 분류기는 확률의 곱으로 이루어져 있기 때문에 계산 결과 값이 0이 나오게 되면 제대로 된 결과를 유추할 수 없다. 때문에 0이 나오는 경우는 그 값을 0.01로 치환하여 계산하도록 하였다.

이렇게 설계하여 구해진 sect = 동편제 일 때의 v_{NB} 값과 sect = 서편제 일 때의 v_{NB}값을 비교하였을 때, 나이브 베이지안 분류기에 의해 값이 보다 큰 쪽이 해당 판소리의 유파로 분류된다. 예를 들어, v_{NB}(sect=동편제) > v_{NB}(sect=서편제) 일 때, 분류기는 해당 판소리를 동편제로 분류한다.

실제 프로그램 화면에 표시되는 내용은 그림 3과 같다. 판소리 음악 파일의 기본적인 주파수 파형, 음 높이 파형, PCD를 이용한 음계 출현 빈도의 그래프, 해당 판소리의 재생 시간과 유파 등이다.

나이브 베이지안을 이용한 판소리 분류 프로그램은 Microsoft사의 Visual C# 2008 Express Edition을 사용하여 구현하였다.

그림 3. 판소리 분류 프로그램 실행 화면
Fig. 3. Implementation screen of Pan-So-Ri classification program

그림 3은 판소리 분류 프로그램의 실행화면이다. ①은 판소리 파일 불러오기, PCD로 분류된 내용 저장하기, 파일 재생 등의 기능을 수행하는 버튼이 위치해 있고 해당 판소리가 동편제인지 서편제인지 알려주는 부분이다. ②는 불러온 판소리 파일의 진폭을 보여준다. ③은 판소리 파일의 음 높낮이를 보여주며, ④는 PCD로 얻어진 결과가 화면에 출력되는 부분이다.

Ⅳ. 실험 및 결과

판소리 유파 분류 실험은 110개의 학습 데이터(동편제

55, 서편제 55)와 60개의 실험 데이터(동편제 30, 서편제 30)가 각 유파별 5:5 비율로 사용되었다.

학습 데이터와 실험 데이터의 판소리 음악 파일의 포맷은 wav 파일, 초당 진동수는 22050Hz, 채널은 1채널 모노로 통일 하였다.

판소리의 아니리(명창이 말로 이야기를 하는 판소리의 한 부분) 부분은 적절한 음높이가 존재하지 않아 실험 대상에서 제외하고 실제 노래하는 부분인 창 부분으로만 학습, 실험 데이터를 구성하였다. 실험에 사용된 음원은 표 2와 같다.

표 2. 판소리 분류 실험에 사용된 음원
Table. 2. Used sound for Pan-So-Ri classification test

유파	소리꾼 - 판소리 제목
동편제	김소희 - 춘향가, 김연수 - 수궁가, 김연수 - 흥보가, 박봉술 - 수궁가, 박봉술 - 흥보가
서편제	김소희 - 심청가, 김연수 - 적벽가, 김연수 - 춘향가, 한애순 - 심청가, 정정렬 - 춘향가

학습 데이터를 토대로 동편제, 서편제 각각에 PCD를 적용하였다. 그림 4, 그림 5는 PCD를 적용한 동편제, 서편제의 음계별 출현빈도를 그래프로 표현한 것이다.

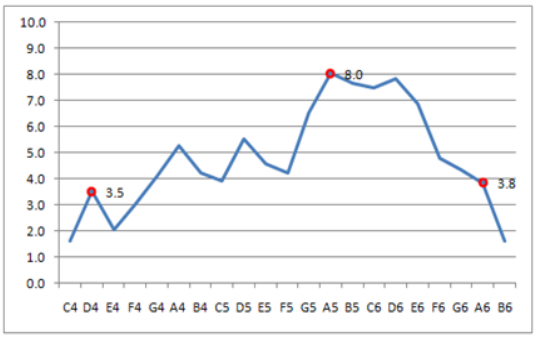

그림 4. PCD를 적용한 동편제의 음계별 출현 빈도
Fig. 4. Appearance frequency of east-sect using PCD

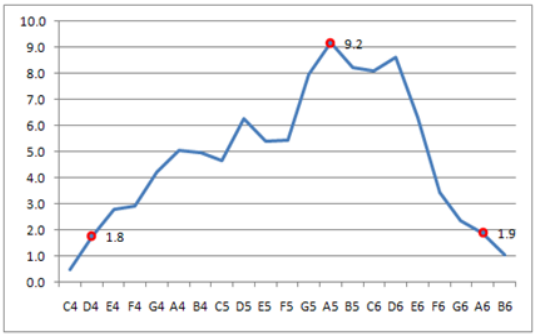

그림 5. PCD를 적용한 서편제의 음계별 출현 빈도
Fig. 5. Appearance frequency of west-sect using PCD

동편제의 경우 서편제와 비교하여 저음에서부터 고음까지의 음의 빈도가 고르게 분포되어있지만, 서편제의 경우 중고음(G5~E6)의 빈도가 타 음에 비해 상당히 높게 나타나는 특징이 있다. 그림에 수치를 표기한 세 점은 동편제와 서편제 간에 차이가 가장 많이 나는 세 음의 출현 확률을 나타낸 것이다.

추출된 각 음계의 출현 빈도에 PCD를 적용하여 도시하고, 나이브 베이지안 분류기를 통해 입력된 판소리 음악 파일의 유파를 구별해 낸다.

위 과정을 토대로 60개 실험 데이터의 유파를 올바르게 구별해 낼 수 있는지 실험하였다. 실험은 추출된 확률값을 소수점 두 번째 자리에서 반올림해서 한 번(실험 1), 소수점 첫 번째 자리에서 반올림해서 한 번(실험 2), 총 두 번에 걸쳐 진행되었다. 표 3과 표 4는 각각 실험 1과 실험 2의 판소리 분류 결과표이다.

표 3. 실험 1의 판소리 분류 결과
Table. 3. Classification result of experiment 1

실험 데이터	분류 결과
동편제 30개	50% 15/30
서편제 30개	83% 29/30
전체 60개	72% 43/60

표 4. 실험 2의 판소리 분류 결과
Table. 4. Classification result of experiment 2

실험 데이터	분류 결과
동편제 30개	80% 24/30
서편제 30개	97% 29/30
전체 60개	88% 53/60

표 3, 표 4에서 보여지듯 실험1 보다 실험2에서 보다 우수한 정확도로 판소리가 분류되었고, 동편제 30개 중 24개, 서편제 30개 중 29개, 전체 60개의 실험 데이터 중 7개를 제외한 나머지 데이터들의 유파를 알맞게 구별 해냈다.

V. 결 론

본 논문에서는 나이브 베이지안 분류기를 이용한 판소리 분류 프로그램을 구현하였으며, 동편제 80%, 서편제 97%로 전체 판소리 유파를 88%의 정확도로 구별해 낼 수 있었다.

현재 구현된 프로그램의 문제점은 판소리 파일의 입력이 22050Hz, 1채널 모노의 wav 파일로 제한된다는 점과 크기가 큰 파일의 경우 로딩 시간이 상당히 오래 걸린다는 점이다.

향후 과제로는 MP3포맷과 스테레오 형식을 추가 지원하여 판소리 파일 입력의 유연성을 높이고, 나이브 베이지안 이외의 마이닝 기법 사용과 판소리의 박자(중모리, 자진모리 등) 개념을 적용하여 분류하는 것이다. 이렇게 한다면 보다 향상된 성능을 기대할 수 있을 것이다.

참 고 문 헌

[1] 최동현, 판소리 이야기, 도서출판 인동, pp.44-49, 1999.

[2] 장윤정, 주제중심의 판소리 지도방안 연구, 한국국악교육연구학회 제 3회 정기총회 및 학술세미나, pp. 129~149, 2009.

[3] P. Chordia and A. Rae, Real-time Raag Recognition for interactive music, New Interfaces for Musical Expression, 2008.

[4] P. Chordia and A. Rae, Raag recognition using pitch-class and pitch-class dyad distributions. In Proceedings of International Conference on Music Information Retrieval, 2007.

[5] P, Chordia and A. Rae, Automatic raag Classification of pitch-tracked performances using pitch-class and pitch-class dyad distributions. In Poceedings of International Conference on Music Information Retrieval, 2007.

[6] 서종문, 김석배, 판소리 '중고제'의 역사적 이해, 국어교육학회, 24권, pp.33-64, 1992.

[7] 이중훈, 경상도 지방을 중심으로 발달한 동편제 판소리와 송문일가인 국창 송만갑, 한국음반학, 4호, 1994.

[8] 조한철, 조근식, 나이브 베이지안 분류자와 메세지 규칙을 이용한 스팸메일 필터링 시스템, 한국 정보과학회 봄 학술논문 발표집, 29권, 1호, pp.223-225, 2002.

[9] Mitchell Tom, Machine Learning, McGraw-Hill, Chapter 6 : Bayesian Learning, 1997.

[10] Jon Kagstrom, Improving Naive Bayesian Spam Filtering, Mittuniversitetet, 2005

[11] 문현구, Hybrid Naive Bayes HMM 기법을 사용한 텍스트로부터의 감정 분류, 석사 학위논문, 2002.

[12] P. Domingos and M. Pazzani. Beyond Independence: Conditions for the Optimality of the simple Bayesian Classfier. In Proc. of the 13th International Conference on Machine Learning in the New Information Age, ECML 2000.

[13] Lawrence R. Rabiner, Michael J. Cheng, Aaron E. Rosenberg, Carol A. Mcgonegal, A Comparative Performance Study of Several Pitch Detection Algorithms, IEEE Transactions on Acoustics, Speech, and Signal Processing, VOL. ASSP-24,NO. 5, 1976.

[14] Kristoffer Jensen, Declan Murphy, Segmenting Melodies into Notes, In Olsen, S., ed.: Proc. 10th Danish Conf. on Pattern Recognition and Image Analysis, 2001.

[15] 김형태, 보이스 오디세이, 북로드, pp135~135, 2007.

저자 소개

김 원 종(준회원)

• 을지대학교 의료IT마케팅학과 재학생

김 명 관(정회원)

• 1981년 3월~1985년 2월 숭실대학교 전자계산학과 학사
1985년 3월~1987년 2월 숭실대학원 전자계산학과 석사
1996년 9월~2004년 2월 숭실대학원 컴퓨터학과 박사
• 1989년 8월~1993년 2월 한국전자통신연구소 인공지능연구실 연구원
1993년 3월~2007년 2월 서울보건대학 컴퓨터정보과 부교수
2007년 3월~현재 : 을지대학교 의료IT마케팅학과 부교수
<주관심분야 : 인공지능, 자연어처리, 질의응답시스템, 시멘틱웹>

이 강 복(준회원)

• 을지대학교 의료IT마케팅학과 재학생