



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0066472
(43) 공개일자 2018년06월19일

(51) 국제특허분류(Int. Cl.)
G10H 1/00 (2006.01) G06N 99/00 (2010.01)
(52) CPC특허분류
G10H 1/0025 (2013.01)
G06N 99/005 (2013.01)
(21) 출원번호 10-2016-0167317
(22) 출원일자 2016년12월09일
심사청구일자 없음

(71) 출원인
반병현
경상북도 안동시 안기1길 39, 102동403호(안기동, 안기동대원아파트)
(72) 발명자
반병현
경상북도 안동시 안기1길 39, 102동403호(안기동, 안기동대원아파트)
(74) 대리인
특허법인충현

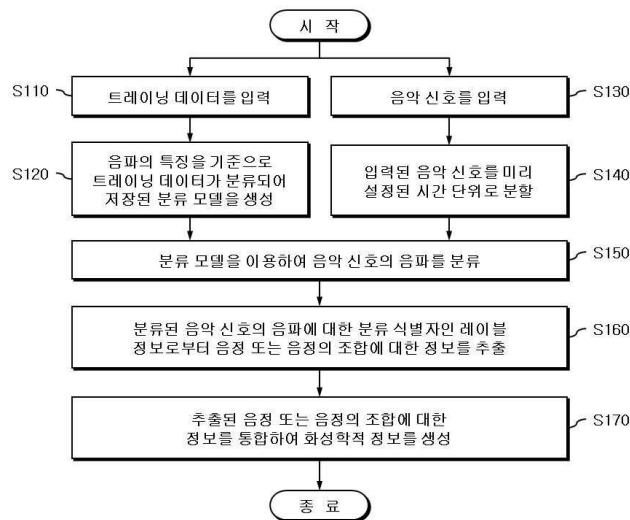
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 기계 학습을 이용한 음악 신호의 처리 방법

(57) 요약

본 발명은 기계 학습을 이용한 음악 신호의 처리 방법에 관한 것으로, 음악 신호를 입력받고, 입력된 음악 신호를 미리 설정된 시간 단위로 분할하고, 음파의 특징을 기준으로 복수 개의 트레이닝 데이터(training data)가 분류되어 저장된 분류 모델을 이용하여 시간 단위로 분할된 음악 신호의 음파를 분류하고, 분류된 음악 신호의 음파에 대한 분류 식별자인 레이블(label) 정보로부터 음정 또는 음정의 조합에 대한 정보를 추출하며, 추출된 음정 또는 음정의 조합에 대한 정보를 통합하여 화성학적 정보를 생성한다.

대표도 - 도1



(52) CPC특허분류

G10H 2210/046 (2013.01)

G10H 2210/071 (2013.01)

G10H 2210/111 (2013.01)

G10H 2210/341 (2013.01)

G10H 2250/311 (2013.01)

명세서

청구범위

청구항 1

음악 신호를 입력받는 단계;

입력된 상기 음악 신호를 미리 설정된 시간 단위로 분할하는 단계;

음파의 특징을 기준으로 복수 개의 트레이닝 데이터(training data)가 분류되어 저장된 분류 모델을 이용하여 상기 시간 단위로 분할된 음악 신호의 음파를 분류하는 단계;

분류된 상기 음악 신호의 음파에 대한 분류 식별자인 레이블(label) 정보로부터 음정 또는 상기 음정의 조합에 대한 정보를 추출하는 단계; 및

추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 통합하여 화성학적 정보를 생성하는 단계를 포함하는 음악 신호의 처리 방법.

청구항 2

제 1 항에 있어서,

적어도 하나 이상의 음정의 조합으로 표현되는 음파로 구성된 복수 개의 트레이닝 데이터를 입력받아 기계 학습(machine learning)을 이용하여 상기 트레이닝 데이터 각각을 미리 분류하여 저장함으로써 분류 모델을 생성하는 단계를 더 포함하는 음악 신호의 처리 방법.

청구항 3

제 2 항에 있어서,

상기 트레이닝 데이터는,

음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파로부터 구축되는 것을 특징으로 하는 음악 신호의 처리 방법.

청구항 4

제 2 항에 있어서,

상기 트레이닝 데이터는,

음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파를 파동 정보로 변환하고, 변환된 상기 파동 정보를 고속 푸리에 변환(fast Fourier transform, FFT)함으로써 다차원의 텐서(tensor)로 생성하며, 생성된 텐서에 분류 식별자인 레이블을 부착하는 것을 특징으로 하는 음악 신호의 처리 방법.

청구항 5

제 1 항에 있어서,

상기 음악 신호의 음파를 분류하는 단계는,

상기 시간 단위로 분할된 음악 신호의 음파를 서포트 벡터(support vector)를 이용하여 학습시킴으로써 상기 분류 모델 내의 이미 분류된 그룹과의 유사 정도에 따라 음파 그룹별로 분류하는 것을 특징으로 하는 음악 신호의 처리 방법.

청구항 6

제 1 항에 있어서,

상기 음악 신호의 음파를 분류하는 단계는,

상기 시간 단위로 분할된 음악 신호의 음파를 신경망(neural network)을 이용하여 계층적으로 학습시킴으로써 상기 분류 모델을 통해 음파 특성 간의 연결 강도에 따라 음파 그룹의 분류 식별자인 레이블(label)을 출력하는 것을 특징으로 하는 음악 신호의 처리 방법.

청구항 7

제 1 항에 있어서,

상기 화성학적 정보를 생성하는 단계는,

추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 나열하여 시계열적으로 연결된 화성학적 정보를 추출하되, 상기 화성학적 정보의 길이는 상기 미리 설정된 시간 단위보다 짧거나 적어도 같은 것을 특징으로 하는 음악 신호의 처리 방법.

청구항 8

제 1 항에 있어서,

생성된 상기 화성학적 정보 내에 일정 시간 간격 동안 동일한 정보가 반복되는지 여부를 검사하여 반복 구간의 화성학적 정보와 지속 시간의 조합을 추출하는 단계를 더 포함하는 음악 신호의 처리 방법.

청구항 9

제 8 항에 있어서,

입력된 음악 신호로부터 추출된 복수 개의 반복 구간의 화성학적 정보와 지속 시간을 재조합하여 새로운 음악 신호를 생성하는 단계를 더 포함하는 음악 신호의 처리 방법.

청구항 10

제 9 항에 있어서,

상기 새로운 음악 신호를 생성하는 단계는,

재조합된 화성학적 정보와 지속 시간에 기반하되 악기의 특징 정보를 원본 음악 신호와 다른 악기로 치환하거나, 둘 이상의 악기의 음색을 병합하거나, 음악적 구성을 변경하여 음파를 생성하는 것을 특징으로 하는 음악 신호의 처리 방법.

발명의 설명

기술 분야

[0001] 본 발명은 음악 신호를 디지털 신호 처리하는 기술에 관한 것으로, 특히 트레이닝 음원 정보를 기계 학습을 이용하여 학습시킨 후 소스(source) 음악 신호로부터 화성학적 정보를 도출하여 변환되거나, 확장되거나, 치환된 음악을 생성하거나 재구성하는 방법에 관한 것이다.

배경 기술

[0002] 아날로그 방식의 음원 녹음/재생 기술의 시대를 지나 다양한 음원들이 디지털화된 방식으로 표현되고 있다. 또한, 음원의 생성 시점부터 디지털 기반의 악보나 음원 신호를 생성할 수 있는 기술도 널리 보급되고 있다. 이를 위해 아날로그 기반의 음원 신호 또는 음파를 디지털 환경에서 표현하고 다루기 위한 표기 방법이나 악기 고유의 음색이나 특징을 디지털 정보로 표현하고자 하는 노력이 함께 이루어지고 있다. 예를 들어, 이하에서 예시되는 선행기술문헌에는 XML 또는 MIDI 파일을 자바 프로그램에서 사용하는 데이터의 구조로 변환하여, 그 데이터 구조를 이용하여 악보를 분석하는 구성을 소개하고 있다.

[0003] 또는, 아날로그 음원 신호를 직접 마이크로폰을 통해 취득하고, 이를 분석하여 유사 음원을 검색하거나 그에 대한 정보를 안내하는 서비스도 현재 제공되고 있다. 이는 아날로그 음원 신호를 디지털 처리하고, 그에 관한 음악적 구조 내지 특징을 추출하여 미리 구축된 데이터베이스와 대비하는 과정을 통해 수행되게 된다.

[0004] 그러나, 상기된 종래의 기술들은 모두 고정된 형태의 음악으로부터 디지털 데이터를 가공하고, 가공된 데이터의 음악적 특징을 활용하여 동일/유사한 데이터를 검색하여 사용자에게 제공하는데 그치고 있다는 한계를 갖는다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 한국특허공개공보 제10-2005-0005247호, 2005년 01월 13일 공개

발명의 내용

해결하려는 과제

[0006] 본 발명이 해결하고자 하는 기술적 과제는, 종래의 디지털 음악 처리 기술이 아날로그 음악의 디지털화 내지는 아날로그 음원과 동일한 디지털 음원을 검색하거나 재생하기 위한 수단으로 사용되었던 한계를 극복하고, 자동화된 음악 분석에 있어서도 하나의 음악을 기계적인 분해 내지 분석을 통해 얻어진 결과를 전체로서 활용하고 있음으로 인해 음악 처리의 결과의 유연한 활용이 어렵다는 문제를 해소하고자 한다.

과제의 해결 수단

[0007] 상기 기술적 과제를 해결하기 위하여, 본 발명의 일 실시예에 따른 음악 신호의 처리 방법은, 음악 신호를 입력받는 단계; 입력된 상기 음악 신호를 미리 설정된 시간 단위로 분할하는 단계; 음파의 특징을 기준으로 복수 개의 트레이닝 데이터(training data)가 분류되어 저장된 분류 모델을 이용하여 상기 시간 단위로 분할된 음악 신호의 음파를 분류하는 단계; 분류된 상기 음악 신호의 음파에 대한 분류 식별자인 레이블(label) 정보로부터 음정 또는 상기 음정의 조합에 대한 정보를 추출하는 단계; 및 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 통합하여 화성학적 정보를 생성하는 단계를 포함한다.

[0008] 일 실시예에 따른 음악 신호의 처리 방법은, 적어도 하나 이상의 음정의 조합으로 표현되는 음파로 구성된 복수 개의 트레이닝 데이터를 입력받아 기계 학습(machine learning)을 이용하여 상기 트레이닝 데이터 각각을 미리 분류하여 저장함으로써 분류 모델을 생성하는 단계를 더 포함할 수 있다. 상기 트레이닝 데이터는, 음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파로부터 구축될 수 있다. 또한, 상기 트레이닝 데이터는, 음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파를 파동 정보로 변환하고, 변환된 상기 파동 정보를 고속 푸리에 변환(fast Fourier transform, FFT)함으로써 다차원의 텐서(tensor)로 생성하며, 생성된 텐서에 분류 식별자인 레이블을 부착할 수 있다.

[0009] 일 실시예에 따른 음악 신호의 처리 방법에서, 상기 음악 신호의 음파를 분류하는 단계는, 상기 시간 단위로 분할된 음악 신호의 음파를 서포트 벡터(support vector)를 이용하여 학습시킴으로써 상기 분류 모델 내의 이미 분류된 그룹과의 유사 정도에 따라 음파 그룹별로 분류할 수 있다.

[0010] 일 실시예에 따른 음악 신호의 처리 방법에서, 상기 음악 신호의 음파를 분류하는 단계는, 상기 시간 단위로 분할된 음악 신호의 음파를 신경망(neural network)을 이용하여 계층적으로 학습시킴으로써 상기 분류 모델을 통해 음파 특성 간의 연결 강도에 따라 음파 그룹의 분류 식별자인 레이블(label)을 출력할 수 있다.

[0011] 일 실시예에 따른 음악 신호의 처리 방법에서, 상기 화성학적 정보를 생성하는 단계는, 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 나열하여 시계열적으로 연결된 화성학적 정보를 추출하되, 상기 화성학적 정보의 길이는 상기 미리 설정된 시간 단위보다 짧거나 적어도 같은 것이 바람직하다.

[0012] 일 실시예에 따른 음악 신호의 처리 방법은, 생성된 상기 화성학적 정보 내에 일정 시간 간격 동안 동일한 정보가 반복되는지 여부를 검사하여 반복 구간의 화성학적 정보와 지속 시간의 조합을 추출하는 단계를 더 포함할 수 있다. 또한, 일 실시예에 따른 음악 신호의 처리 방법은, 입력된 음악 신호로부터 추출된 복수 개의 반복 구간의 화성학적 정보와 지속 시간을 재조합하여 새로운 음악 신호를 생성하는 단계를 더 포함할 수 있다. 나아가, 상기 새로운 음악 신호를 생성하는 단계는, 재조합된 화성학적 정보와 지속 시간에 기반하되 악기의 특징 정보를 원본 음악 신호와 다른 악기로 치환하거나, 둘 이상의 악기의 음색을 병합하거나, 음악적 구성을 변경하여 음파를 생성할 수 있다.

[0013] 나아가, 이하에서는 상기 기재된 음악 신호의 처리 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.

발명의 효과

[0014] 본 발명의 실시예들은, 음악 신호로부터 음파를 분해하여 그로부터 기계 학습을 활용한 분류 모델을 통해 해당 음파 신호를 분류하고, 이로부터 화성학적 정보를 추출함으로써 원본 음악을 넘어서 파편화된 음파에 내재된 음악적 특징 일부만으로도 악곡의 구조, 음색이나 조를 변경하거나, 연주 악기를 달리하는 새로운 음악을 생성하는 것이 가능하다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시예에 따른 기계 학습을 이용한 음악 신호의 처리 방법을 도시한 흐름도이다.

도 2는 본 발명의 일 실시예에 따른 음악 신호의 처리 방법에서 기계 학습을 위해 가공된 트레이닝 데이터를 표현하는 방식의 일례를 나타낸 도면이다.

도 3 및 도 4는 본 발명의 일 실시예에 따른 음악 신호의 처리 방법에서 활용 가능한 기계 학습 방법을 예시한 도면이다.

도 5는 본 발명의 일 실시예에 따른 도 1의 음악 신호의 처리 방법에서 생성된 화성학적 정보를 활용한 음악 신호 처리 방법을 도시한 흐름도이다.

도 6은 본 발명의 일 실시예에 따른 도 1의 음악 신호의 처리 과정을 데이터의 흐름을 중심으로 재구성한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0016] 본 발명의 실시예들을 설명하기에 앞서, 디지털 음악 처리 분야에 활용하는 요소 기술의 특성과 문제점을 검토한 후, 이러한 문제점을 해결하기 위해 본 발명의 실시예들이 채택하고 있는 기술적 수단을 개괄적으로 소개하도록 한다.

[0017] 앞서 논의한 바와 같이, 종래의 디지털 음악 처리 분야의 큰 축은 전통적인 방식으로 녹음된 아날로그 음악을 디지털화거나, 디지털화 과정에서 도출된 특징 정보를 활용하여 아날로그 음원과 동일한 디지털 음원을 검색하거나 재생하기 위한 수단에 관한 것이었다. 이와는 달리 디지털 음악 처리 분야의 또 다른 영역은 음악의 구조를 계량화하거나 분석하는 기술에 관한 것으로, 자동화된 연산 과정을 통해 하나의 음악을 구성하는 구성요소들을 기계적으로 분해하고 이를 분석한 후, 이들 결과를 취합하여 정량적인 데이터를 얻는데 활용하였다. 그 결과 음악 분석의 결과물이 전체로서 활용되고 있을 뿐이며, 분석 과정에서 도출된 다양한 특징은 하나의 결과 내에 통합될 뿐, 각각의 과정으로부터 새로운 결과를 얻을 수 있는 유연함이 부족하였다.

[0018] 따라서, 본 발명의 실시예들은 이러한 종래의 디지털 음악 처리 방식에서 벗어나, 하나의 음악 또는 단편적인 음파를 파편화하고, 파편화된 음악 신호로부터 미리 인공 지능의 여러 방법론을 활용해 구축된 분류 모델을 활용하여 그 범주를 분류하며, 이로부터 화성학적 정보를 추출하는 음악 처리 방법을 제시하고자 한다. 이러한 방법론을 통해 최초의 원본 소스(original source)에 내재된 음악적 특징 일부만으로도 모티프(motif)에 기반한 새로운 음악을 도출할 수 있다.

[0019] 이하에서는, 도면을 참조하여 상기된 기술적 과제를 해결하기 위한 본 발명의 실시예들을 구체적으로 설명한다. 다만, 하기의 설명 및 첨부된 도면에서 본 발명의 요지를 흐릴 수 있는 공지 기능 또는 구성에 대한 상세한 설명은 생략한다. 또한, 도면 전체에 걸쳐 동일한 구성 요소들은 가능한 한 동일한 명칭 및 도면 부호로 나타내고 있음에 유의하여야 한다.

[0020] 도 1은 본 발명의 일 실시예에 따른 기계 학습을 이용한 음악 신호의 처리 방법을 도시한 흐름도로서, 입력은 크게 트레이닝 데이터와 처리 대상인 음악 신호로 구분할 수 있다. 물론 트레이닝 데이터 역시 음악 신호이나, 여기서는 음악 신호의 처리 이전에 미리 기계 학습 알고리즘을 통해 분류 모델을 생성하기 위해 활용되는 기초 데이터를 의미하는 것으로 명시하였다. 또한, 도 1에서 제시되는 음악 신호 처리 방법은 적어도 하나의 프로세서(processor)를 구비하는 음악 신호의 처리 장치가 일련의 처리 과정을 정의하는 알고리즘을 구동함으로써 구현될 수 있다. 이를 위해 본 발명의 실시예들을 통해 주어지는 입력 신호는 비록 최초 취득시에 아날로그 데이터 또는 디지털 데이터일 수 있으나, 적어도 디지털 처리기를 통해 입력되는 시점에는 디지털 정보로 가공되었

음을 전제로 한다.

- [0021] S110 단계에서, 음악 신호의 처리 장치는, 적어도 하나 이상의 음정의 조합으로 표현되는 음파로 구성된 복수 개의 트레이닝 데이터를 입력받고, S120 단계를 통해 기계 학습(machine learning)을 이용하여 음파의 특징을 기준으로 상기 트레이닝 데이터 각각을 미리 분류하여 저장함으로써 분류 모델을 생성한다. 이를 위해 다양한 인공지능 이론에 기반한 지도 학습(supervised learning) 모델이 활용될 수 있다.
- [0022] 도 2는 본 발명의 일 실시예에 따른 음악 신호의 처리 방법에서 기계 학습을 위해 가공된 트레이닝 데이터를 표현하는 방식의 일례를 나타낸 도면이다.
- [0023] 본 발명의 실시예들에서는 트레이닝 데이터 또는 입력 데이터의 처리를 위해 음악 신호를 다차원 배열로 표현할 수 있는데, 예를 들어 구글(Google)의 텐서플로우(TensorFlow) 등에서 활용할 수 있는 데이터 객체인 텐서(Tensor)를 채택하여 활용 가능하다.
- [0024] 텐서는 딥 러닝 등에서 그래프 내의 오퍼레이션 간에 데이터를 전달하는 객체로서, 본 발명의 실시예들에서 트레이닝 데이터는, 음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파로부터 구축되는 것이 바람직하다. 구현의 관점에서, 상기 트레이닝 데이터는, 음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파를 파동 정보로 변환하고, 변환된 상기 파동 정보를 고속 푸리에 변환(fast Fourier transform, FFT)함으로써 다차원의 텐서(tensor)로 생성하며, 생성된 텐서에 분류 식별자인 레이블(label)을 부착할 수 있다. 도 2를 참조하면, 레이블이 부착된 음악 신호 데이터(여기서는 트레이닝 데이터를 의미한다)를 다차원 배열로 표현하였으며, 각각의 트레이닝 데이터를 차원을 기준으로 펼치면 도식된 바와 같은 텐서로서 표현될 수 있음을 예시하였다.
- [0025] 나아가, 본 발명의 실시예들에서 분류 모델은 통상적인 피드-포워드 네트워크(Feedforward Networks) 구조를 채택하는 인공 신경망뿐만 아니라, RNN (Recurrent Neural Networks) 방식의 신경망, 특히 Long Short-Term Memory Units (LSTM)으로 구현될 수 있다. 이 경우, 상술된 텐서로 표현되는 트레이닝 데이터의 활용이 더욱 용이하다.
- [0026] 일반적으로 음악에서 사용되는 음계는 총 88개로, 피아노의 건반 수와 일치한다. 따라서 본 발명의 실시예들에서는 실시의 범용성을 위해 선택 가능한 음정의 개수 P를 88로 제한하여 예시하였고, 동시에 선택 가능한 음정의 개수 k를 3으로 제한하여 예시하였으며, 단음 또는 화음의 조합의 개수인 q를 10으로 제한하여 예시하였다. 그러나, 이는 설명의 편의와 프로토타입 구현의 편의를 위해 설정된 값으로서, 본 발명의 기술적 사상이나 권리 범위를 제한하는 것이 아님을 밝혀둔다.
- [0027] 이제, P개의 음정 중에서 동시에 선택 가능한 최대 k개의 음정으로 화음을 구축하는 경우의 수 n은 수학적 1과 같다.

수학적 1

$$n = \sum_{i=1}^k p C_i$$

- [0028]
- [0029] 상기된 전제에 따라, P= 88, k = 3일때 n은 113,652가 된다. 여기에 각각의 단음/화음 조합을 q개의 서로 다른 악기로 연주하여 파형을 디지털 데이터로 취한다면 그 데이터는 1,113,652개의 열(column)을 가진 행렬 M으로 표현될 수 있다.
- [0030] 상술된 파라미터는 본 발명의 실시예들을 통해 구현되는 음악 신호 처리 소프트웨어에 활용될 전자기기의 성능에 따라 제한될 수 있으며, 구현의 편의를 위해 한 개의 옥타브에 해당하는 13개의 음만 구비할 경우 n은 858이라는 작은 값으로 고정될 수 있다.
- [0031] 총 n 종류의 레이블이 부착된 10n개의 데이터를 n개의 클러스터로 분류하는 과정은 기계 학습 영역 안에서도 지도 학습이라는 영역에 속한다. 지도 학습 과정을 통하여 트레이닝 데이터를 학습한 지도 학습기(Supervised Learner)는 추가로 입력된 벡터(본 발명의 실시예들에서는 텐서로 표현된 음악 신호를 의미한다)가 어떤 종류의 값인지 표시하는 분류(Classification)작업을 수행할 수 있다.
- [0032] 10n개의 빅데이터를 학습한 인공지능은 p개의 가능한 음정 영역 내에서 최대 k개까지 선택된 음정 또는 음정의

조합을 입력받아 그 입력값이 트레이닝 데이터 중 어떤 값과 가장 일치하는지를 분류할 수 있다.

- [0033] 서포트 벡터 머신(support vector machine, SVM), 인공 신경망(artificial neural network), 나이브 베이즈 분류(Naive Bayes Classification) 방법 등 레이블이 부착된 대량의 데이터를 학습하여 분류 인공지능을 구축할 수 있는 알고리즘이 다양하게 공개되어 있으므로 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 P와 k의 크기와 인공지능 학습에 사용할 컴퓨터 장치의 성능에 따라 다양한 알고리즘을 취사 선택하여 인공 지능을 구축할 수 있을 것이다.
- [0034] 도 1로 돌아와서, 음악 신호의 처리 장치는, S130 단계를 통해 새로운 음악 신호, 즉 처리 대상인 음악 신호를 입력받고, S140 단계를 통해 상기 음악 신호를 미리 설정된 시간 단위로 분할한다. 이때의 시간 단위는 적어도 음악의 일부 소절로부터 화성학적 특징이 도출될 수 있을 정도의 충분한 길이를 갖는 것이 바람직하다.
- [0035] 보다 구체적으로, 사용자로부터 외부 음파를 입력받는 수단은 마이크로폰과 같은 물리적 장치일 수도 있으며, 사용자의 단말기에 저장된 소리 파일을 전자적으로 불러들이는 모듈일 수도 있다. 물리적 장치의 경우, 입력받은 외부 음파의 파형을 검출하여 아날로그 오디오 신호를 획득한 후, 획득된 아날로그 오디오 신호를 디지털 오디오 신호로 변환하는 과정이 필요하다. 도 1에서는 간단히 S130 단계를 통해 디지털화된 음악 신호가 입력되었다고 가정하자. 반면, 사용자의 단말기에 저장된 소리 파일을 전자적으로 불러들여 사용한다면 그 데이터가 바로 디지털 오디오 신호 형태로 변환된 음파 데이터에 해당한다. 이제, 입력된 디지털 신호를 단위 시간 간격으로 쪼개어 분할된 하나의 조각(fragment)을 분류 모델의 입력 신호로서 준비한다.
- [0036] S150 단계에서, 상기 음악 신호의 처리 장치는, 음파의 특징을 기준으로 복수 개의 트레이닝 데이터(training data)가 분류되어 저장된 분류 모델을 이용하여 상기 시간 단위로 분할된 음악 신호의 음파를 분류한다. 보다 구체적으로 이미 S120 단계를 통해 학습이 진행된 인공지능은 분류장치(Classifier)로써의 역할을 수행할 수 있다. 사용자가 임의의 입력값을 주입하면 인공지능은 이를 트레이닝 데이터와 비교하여 가장 유사한 클러스터(cluster)에 배치하고, 그 클러스터에 부착된 레이블을 출력할 수 있다. 가장 간단한 예로서, {도/미/솔}이 입력될 경우, 음악 신호의 처리 장치 내의 인공 지능이 트레이닝 데이터 내부의 클러스터 중 이와 가장 일치하는 클러스터 {도/미/솔}에 입력값 {도/미/솔}을 배치하고, 클러스터 {도/미/솔}에 부착되어 있던 레이블 "도/미/솔"을 출력하게 된다.
- [0037] 즉, 상기 분류 모델은, 레이블이 부착된 트레이닝 데이터를 학습하여 레이블을 따라 트레이닝 데이터를 분류하고, 레이블이 부착되지 않은 입력 데이터를 분석하여 어떤 레이블을 부착할 것인지, 어떤 클러스터에 배정할 것인지 또는 어떤 레이블을 따라 분류할 것인지를 계산한다.
- [0038] 앞서 S120 단계를 통해 설명한 바와 같이, 트레이닝 데이터는 "음파+레이블"의 사상으로 구축될 수 있다. 이때, 상기 레이블은 최대 k개의 음계의 조합으로 구성될 수 있으며, 상기 트레이닝 데이터는 88개의 음계 중 1개 이상 k개 이하의 음정을 임의로 선택하는 모든 경우의 수 n을 취하고, 그 음정 또는 화음을 악기로 연주하여 취한 음파로 구축될 수 있다. 또한, 상기 악기는 실물 악기는 물론 전자악기(MIDI 등)도 포함할 수 있으며, 상기 트레이닝 데이터는 동일한 음정 또는 화음을 다양한 악기로 연주하여 녹음된 데이터를 포함할 수 있다.
- [0039] 따라서, 본 발명의 실시예들에 따른 음악 신호의 처리 방법에서 구현하는 인공 지능은 상기된 음악 신호의 음파를 레이블에 따라 분류하기 위해 기계 학습 또는 클러스터링 알고리즘을 활용한다. 일례로서, 상기 인공 지능은 내부적으로 n개의 클러스터를 구축할 수 있으며, 입력 데이터를 최대 n개까지의 클러스터에 배정할 수 있다. 상기 입력 데이터는 디지털 신호로 가공된 음파 정보일 수 있는데, 상기 인공 지능 분류 모델을 통해 계산 결과 도출된 가장 적절한 클러스터 정보를 출력하게 된다.
- [0040] S160 단계에서, 상기 음악 신호의 처리 장치는, 분류된 상기 음악 신호의 음파에 대한 분류 식별자인 레이블(label) 정보로부터 음정 또는 상기 음정의 조합에 대한 정보를 추출한다. 레이블 정보에 따르면, 클러스터(cluster)를 지칭할 수 있는데, 이를 통해 포인터 주소, 객체명, 클러스터명 등의 정보를 얻을 수 있다. 따라서, S160 단계는 분류된 상기 음악 신호의 음파에 대해 클러스터 정보를 사상(mapping)시킴으로써, 음정 또는 화음 정보를 출력할 수 있다.
- [0041] S170 단계에서, 상기 음악 신호의 처리 장치는, 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 통합하여 화성학적 정보를 생성한다. 이러한 화성학적 정보에는 일정 시간 동안 지속되는 음정과 음정들의 조합에 대한 정보가 포함될 수 있으므로, 이로부터 동일한 음악적 구조를 넘어서는 새로운 악곡 특징을 도출하는 것이 가능하다.
- [0042] 또한, 화성학적 정보를 생성하는 S170 단계는, 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 나열하

여 시계열적으로 연결된 화성학적 정보를 추출하되, 상기 화성학적 정보의 길이는 앞서 도 1의 S140 단계를 통해 미리 설정된 시간 단위보다 짧거나 적어도 같은 것이 바람직하다. 이를 통해 파편화된 소절에서도 화성학적 특징을 도출하는 것이 가능하며, 이들의 시계열적인 배치를 통해 음악 전체를 관통하는 조성의 변화와 흐름을 파악할 수 있게 된다.

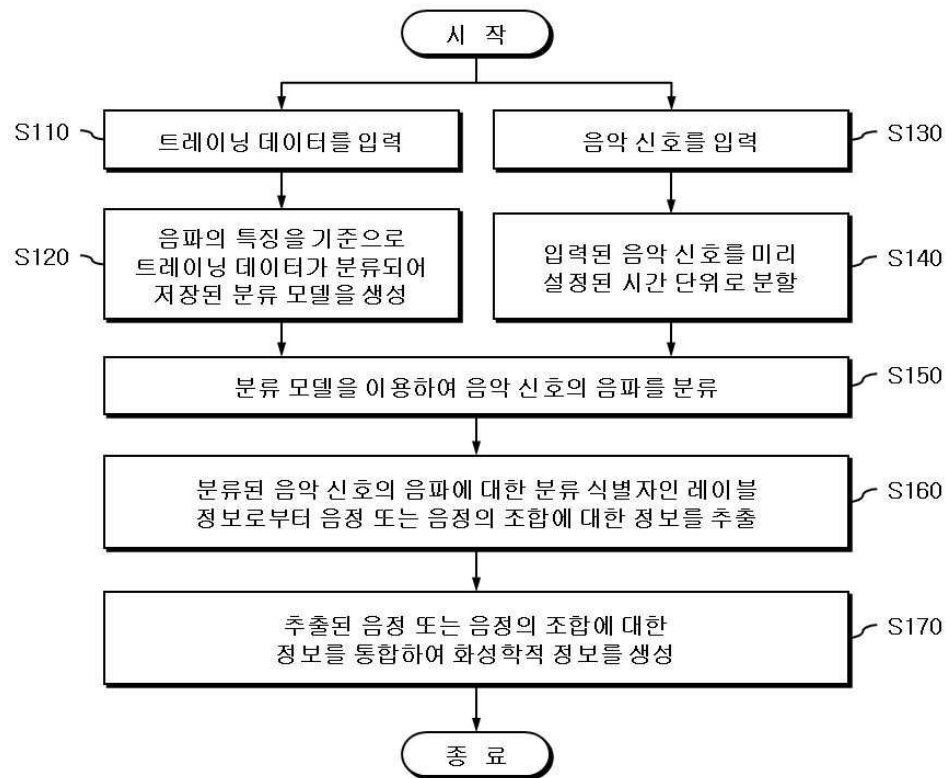
- [0043] 도 3 및 도 4는 본 발명의 일 실시예에 따른 음악 신호의 처리 방법에서 활용 가능한 기계 학습 방법을 예시한 것으로서, 각각 기계 학습의 다양한 방법론 중 서포트 벡터 머신과 신경망을 설명하기 위한 도면이다.
- [0044] 서포트 벡터 머신(support vector machine)은 서포트 벡터를 이용하여 분류 모델을 생성, 갱신할 수 있다. 서포트 벡터는 패턴 인식과 자료 분석을 위한 지도 학습(supervised learning) 모델로서, 복수 개의 그룹 데이터를 분류하거나 회귀분석에 유용한 방법이다. 예를 들어, 서포트 벡터를 활용하여 전체 트레이닝 데이터에 대하여 레이블 간의 음파 특징을 고려한 분류가 가능하다.
- [0045] 도 3에 예시된 바와 같이, 시간 단위로 분할된 음악 신호의 음파를 흑과 백의 도트(dot)로 구분하는 서포트 벡터(support vector)를 이용하여 학습시킴으로써 분류 모델 내의 이미 분류된 그룹과의 유사 정도에 따라 음파 그룹별로 음악 신호의 음파를 분류할 수 있다.
- [0046] 또한, 신경망(neural network)을 이용하여 분류 모델을 생성, 갱신할 수 있다. 신경망은 생물학적 뉴런의 동작 원리와 뉴런간의 연결 관계를 모델링한 것으로 노드(node) 또는 처리 요소(processing element)라고 명명된 다수의 뉴런들이 층(Layer) 구조의 형태로 연결된 정보처리 시스템이다. 이러한 인공 신경망 모델, 예를 들어 일반적인 다층 신경망은 입력층과 은닉층, 출력층으로 구성되는데, 입력층(input layer)은 외부의 자료들을 받아들이는 층으로서 입력층의 뉴런 수는 입력되는 변수의 수와 동일하며, 은닉층(hidden layer)은 입력층과 출력층 사이에 위치하며 입력층으로부터 신호를 받아 특성을 추출하여 출력층으로 전달한다. 출력층(output layer)은 은닉층으로부터 신호를 받아 외부로 출력한다. RNN의 경우 시계열적인 입력을 받아 종래의 기억으로 활용할 수 있으며, 과거에 입력되어 처리된 데이터를 재차 입력으로 고려하게 된다. 뉴런간의 입력신호는 각각의 연결 강도와 승산(또는 LSTM의 경우 피드백의 합산)된 후 취합되며 이 값이 뉴런의 임계치보다 크면 뉴런이 활성화되어 활성화 함수를 통하여 출력값으로 구현된다.
- [0047] 도 4에 예시된 바와 같이, 시간 단위로 분할된 음악 신호의 음파를 신경망(neural network)을 이용하여 계층적으로 학습시킴으로써 분류 모델을 통해 음파 특성 간의 연결 강도에 따라 음파 그룹의 분류 식별자인 레이블(label)을 출력함으로써 음악 신호의 음파를 분류할 수 있다.
- [0048] 도 5는 본 발명의 일 실시예에 따른 도 1의 음악 신호의 처리 방법에서 생성된 화성학적 정보를 활용한 음악 신호 처리 방법을 도시한 흐름도로서, 화성학적 정보를 생성하는 도 1의 S170 단계에 연속하는 과정을 기술하고 있다. 따라서, 도 5에서는 S170 단계를 통해 화성학적 정보보다 도출되었음을 전제로 이후의 과정을 설명하도록 한다.
- [0049] S180 단계에서, 음악 신호의 처리 장치는, 앞서 생성된 상기 화성학적 정보 내에 일정 시간 간격 동안 동일한 정보가 반복되는지 여부를 검사하여 반복 구간의 화성학적 정보와 지속 시간의 조합을 추출한다.
- [0050] S190 단계에서, 상기 음악 신호의 처리 장치는, 최초에 입력된 음악 신호로부터 추출된 복수 개의 반복 구간의 화성학적 정보와 지속 시간을 재조합하여 새로운 음악 신호를 생성한다. 이러한 과정은, 재조합된 화성학적 정보와 지속 시간에 기반하되 악기의 특징 정보를 원본 음악 신호와 다른 악기로 치환하거나, 둘 이상의 악기의 음색을 병합하거나, 음악적 구성을 변경하여 음파를 생성함으로써 수행될 수 있다. 특히 MIDI와 같은 디지털 음악의 경우 일부 파라미터의 변경만으로 악기의 변경이 가능하므로, 음악적 구조를 동일하게 유지하면서도 연주를 위한 악기를 다르게 선택함으로써 청자에게 새로운 느낌을 주는 음악을 생성하는 것이 가능하다.
- [0051] 만약, 악곡의 조(key) 전개나 어울리는 조 바꿈에 대한 데이터베이스를 미리 보유하고 있다면, 이러한 데이터베이스를 참고하여 이전의 과정을 통해 파편화된 음악 신호로부터 획득된 화성학적 정보를 재조합하여 조를 변경하거나 새로운 전개 방식으로 변환하는 것이 가능하다.
- [0052] 도 6은 본 발명의 일 실시예에 따른 도 1의 음악 신호의 처리 과정을 데이터의 흐름을 중심으로 재구성한 도면으로서, 여기서는 설명의 중복을 피하기 위해 각각의 구성만을 약술하도록 한다.
- [0053] 우선, 마이크로폰 등의 하드웨어 장치를 통해 아날로그 음원 신호를 취득한다(11 단계). 이렇게 취득된 신호는 적어도 하나의 프로세서를 구비하는 음악 신호의 처리 장치에 입력될 수 있도록 디지털 신호로 변환된다(13 단계). 그런 다음, 음악의 전체가 아닌 일부 소절에서도 음악적 모티프나 화성학적 특징을 도출할 수 있도록 해당

디지털 신호를 미리 설정된 시간 단위로 분할한다(15 단계). 이렇게 분할된 신호가 신호 입력 수단을 통해 입력된다(20 단계).

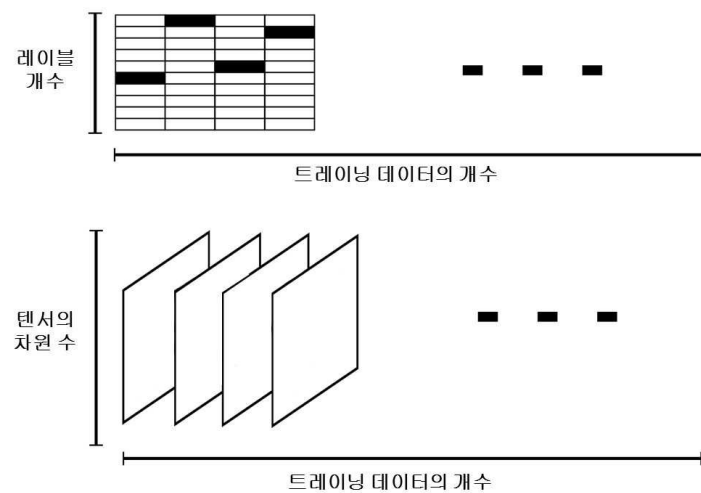
- [0054] 그에 앞서, 도 6의 음악 신호의 처리 장치는 분류 모델의 학습(31 단계)을 위해 미리 복수의 트레이닝 데이터를 입력받는다(17 단계). 구현의 관점에서 이러한 트레이닝 데이터는 음계 중 동시에 선택 가능한 조합에 의해 형성되는 음정 또는 화음을 적어도 하나의 악기를 이용하여 취득한 음파로부터 생성된 다차원의 텐서(tensor)와 그에 부속된 분류 식별자인 레이블의 조합으로 입력되는 것이 바람직하다.
- [0055] 이제, 앞서 입력된 음원 신호로부터 시간 단위로 분할된 음악 신호는 분류 모델을 통해 분류되며(33 단계), 클러스터 정보 사상을 통해 분류된 음악 신호의 음파에 대한 분류 식별자인 레이블(label) 정보로부터 음정 또는 상기 음정의 조합에 대한 정보를 추출한다(40 단계).
- [0056] 그런 다음, 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 통합하여 화성학적 정보를 생성한다(50 단계). 이 과정에서, 추출된 상기 음정 또는 상기 음정의 조합에 대한 정보를 나열하여 시계열적으로 연결된 화성학적 정보를 추출하게 된다.
- [0057] 마지막으로, 생성된 상기 화성학적 정보 내에 일정 시간 간격 동안 동일한 정보가 반복되는지 여부를 검사하여 반복 구간의 화성학적 정보와 지속 시간의 조합을 추출하며, 입력된 음악 신호로부터 추출된 복수 개의 반복 구간의 화성학적 정보와 지속 시간을 재조합하여 새로운 음악 신호를 생성한다(60 단계).
- [0058] 상기된 본 발명의 실시예들에 따르면, 음악 신호로부터 음파를 분해하여 그로부터 기계 학습을 활용한 분류 모델을 통해 해당 음파 신호를 분류하고, 이로부터 화성학적 정보를 추출함으로써 원본 음악 신호에 내재된 음악적 특징 일부만으로도 악곡의 구조, 음색이나 조를 변경하거나, 연주 악기를 달리하는 새로운 음악을 생성하는 것이 가능하다.
- [0059] 한편, 본 발명의 실시예들은 컴퓨터로 읽을 수 있는 기록 매체에 컴퓨터가 읽을 수 있는 코드로 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다.
- [0060] 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등을 포함한다. 또한, 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술 분야의 프로그래머들에 의하여 용이하게 추론될 수 있다.
- [0061] 이상에서 본 발명에 대하여 그 다양한 실시예들을 중심으로 살펴보았다. 본 발명에 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

도면

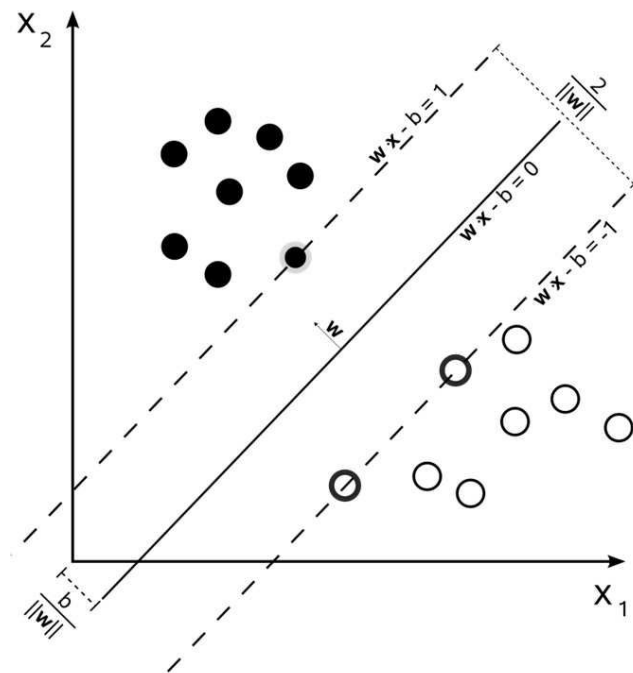
도면1



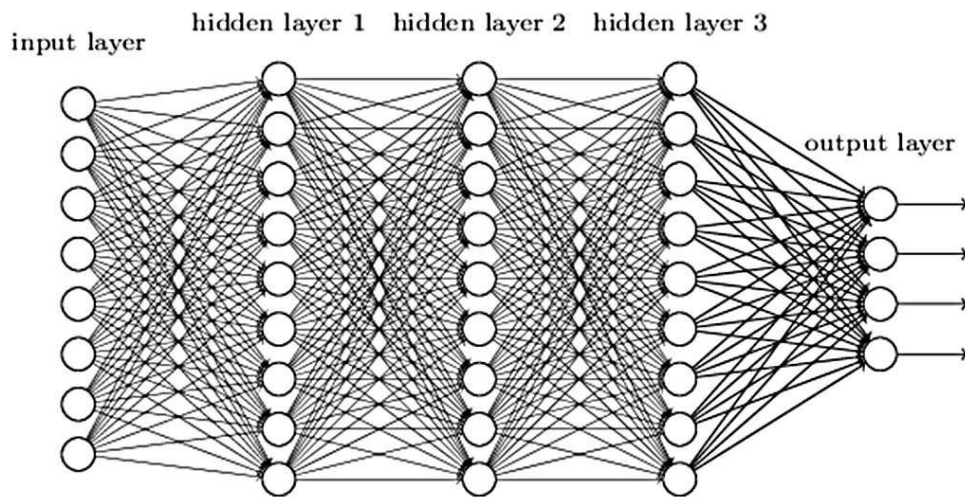
도면2



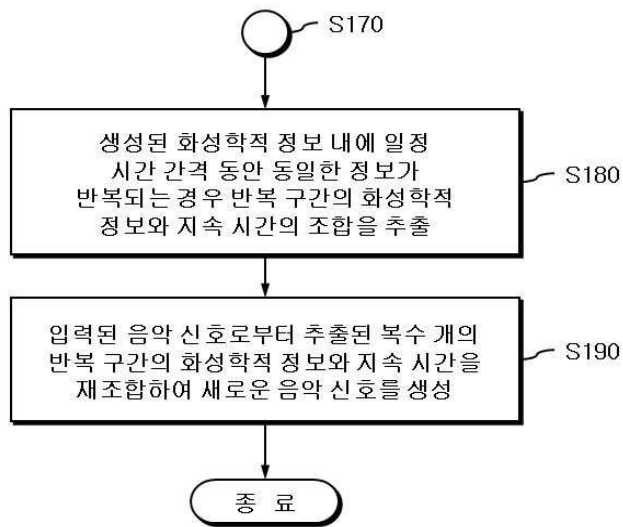
도면3



도면4



도면5



도면6

