(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2015년07월23일
(11) 등록번호 10-1539112
(24) 등록일자 2015년07월17일

(51) 국제특허분류(Int. Cl.)
G10L 25/63 (2013.01) *G10L 25/27* (2013.01)
(21) 출원번호 10-2014-0012174
(22) 출원일자 2014년02월03일
심사청구일자 2014년02월03일
(56) 선행기술조사문헌
Qi-rong Mao et al., 'A novel hierarchical speech emotion recognition method based on improved DDAGSVM',ComSIS Col.7, No.1, Special Issue, pp.211~221, February 2010.*
Wansen Wang et al., 'Research of improved SVM model based on FA in e-learning emotion classicication', Proceedings of IEEE CCIS 2012, pp.919~923, 2012*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이승룡
경기도 성남시 분당구 구미로144번길 25,102동 102호 (구미동, 삼환빌라)
르바브이
경기도 용인시 기흥구 덕영대로 1732 (서천동)
(74) 대리인
특허법인 신지

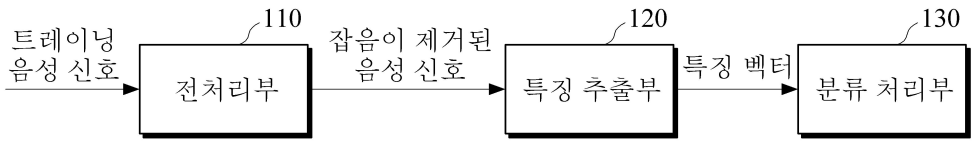
전체 청구항 수 : 총 6 항 심사관 : 정성윤

(54) 발명의 명칭 **음성 인식을 위한 감성 분류 장치 및 분류 방법**

(57) 요 약

본 발명에 따른 감성 분류 장치는 수신된 트레이닝 음성 신호에서 잡음을 제거하는 전처리부, 잡음이 제거된 트레이닝 음성 신호에서 특징 벡터를 추출하는 특징 추출부 및 추출된 특징 벡터를 특징 비트 및 클래스 비트로 구분하고, 클래스 비트를 구성하는 하나 이상의 비트 중에서 유전자 알고리즘(Ganeric Algorithm)을 통해 선택되지 않은 비트를 왼쪽 그룹으로 구분하고, 선택된 비트를 오른쪽 그룹으로 구분하여 솔루션을 생성하고, 서포트 벡터 머신(Support Vector Machine)을 사용하여 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성하는 분류 처리부를 포함한다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업
과제고유번호 1415128832
부처명 지식경제부
연구관리전문기관 정보통신산업진흥원
연구사업명 대학 IT 연구센터 육성지원
연구과제명 동서신의학 u-라이프케어연구센터
기 여 율 0.7/1
주관기관 경희대학교 산학협력단
연구기간 2013.01.01 ~ 2013.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 2011-0030823
부처명 교육과학기술부
연구관리전문기관 한국연구재단
연구사업명 선도연구센터지원사업
연구과제명 디지로그휴먼미디어 감성인지 기술연구개발
기 여 율 0.3/1
주관기관 경희대학교 산학협력단
연구기간 2012.09.01 ~ 2013.08.31

명 세 서

청구범위

청구항 1

수신된 트레이닝 음성 신호에서 잡음을 제거하는 전처리부;

상기 잡음이 제거된 트레이닝 음성 신호에서 특징 벡터를 추출하는 특징 추출부; 및

상기 추출된 특징 벡터를 특징 비트 및 클래스 비트로 구분하고, 상기 클래스 비트를 구성하는 하나 이상의 비트 중에서 유전자 알고리즘(Genetic Algorithm)을 통해 선택되지 않은 비트를 왼쪽 그룹으로 구분하고, 선택된 비트를 오른쪽 그룹으로 구분하여 솔루션을 생성하고, 서포트 벡터 머신(Support Vector Machine)을 사용하여 상기 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성하는 분류 처리부;

를 포함하며,

상기 분류 처리부는,

$$accuracy=\frac{TP+TN}{2}$$

를 통해 상기 생성된 솔류션을 평가하여 정확도를 산출하며, 상기 TP는 트루 포지티브(True Positive)를 나타내고, 상기 TN은 트루 네거티브(True Negative)를 나타내며,

$$fitness=\frac{1}{5}\sum accuracy$$

를 통해 적합성 함수 산출하여 솔루션을 평가하는 것을 특징으로 하는 음성 인식을 위한 감성 분류 장치.

청구항 2

제 1항에 있어서,

상기 분류 처리부는,

오분류 확률(Misclassification Rate)를 사용하여 상기 생성된 솔루션의 평가하여 정확도를 산출하고, 상기 산출된 정확도에 k-폴드 교차 타당화(k-fold Cross Validation)를 적용하여 적합성 함수를 산출하여 상기 솔루션을 평가하여 상기 2진 분류 트리를 생성하는 것을 특징으로 하는 음성 인식을 위한 감성 분류 장치.

청구항 3

삭제

청구항 4

제 1항에 있어서,

상기 클래스 비트를 구성하는 하나 이상의 비트 각각은 서로 다른 감성 상태를 나타내는 것을 특징으로 하는 음성 인식을 위한 감성 분류 장치.

청구항 5

수신된 트레이닝 음성 신호의 잡음을 제거하는 단계;

잡음이 제거된 상기 트레이닝 음성 신호에서 특징 벡터를 추출하는 단계;

상기 특징 벡터에 포함된 클래스 비트를 구성하는 하나 이상의 비트 중에서 유전자 알고리즘(Genetic Algorithm)을 통해 선택되지 않은 비트를 왼쪽 그룹으로 구분하고, 선택된 비트를 오른쪽 그룹으로 구분하여 솔루션을 생성하는 단계; 및

서포트 벡터 머신(Support Vector Machine)을 사용하여 상기 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성하는 단계;

를 포함하며,

상기 2진 분류 트리(Binary Classification Tree)를 생성하는 단계는,

$$accuracy = \frac{TP+TN}{2}$$

를 통해 상기 생성된 솔루션을 평가하여 정확도를 산출하며, 상기 TP는 트루 포지티브(True Positive)를 나타내고, 상기 TN은 트루 네거티브(True Negative)를 나타내며,

$$fitness = \frac{1}{5}\sum accuracy$$

를 통해 적합성 함수를 산출하여 솔루션을 평가는 것을 특징으로 하는 음성 인식을 위한 감성 분류 방법.

청구항 6

제 5항에 있어서,

서포트 벡터 머신(Support Vector Machine)을 사용하여 상기 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성하는 단계는,

오분류 확률(Misclassification Rate)를 사용하여 상기 생성된 솔루션의 평가하여 정확도를 산출하고, 상기 산출된 정확도에 k-폴드 교차 타당화(k-fold Cross Validation)를 적용하여 적합성 함수를 산출하여 상기 솔루션을 평가하여 상기 2진 분류 트리를 생성하는 것을 특징으로 하는 음성 인식을 위한 감성 분류 방법.

청구항 7

삭제

청구항 8

제 5항에 있어서,

상기 클래스 비트를 구성하는 하나 이상의 비트 각각은 서로 다른 감성 상태를 나타내는 것을 특징으로 하는 음성 인식을 위한 감성 분류 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 음성 인식에 관한 것으로, 보다 상세하게는 음성 인식을 위한 감성 음성 분류 과정에 관한 것이다.

배 경 기 술

[0002] 감성 음성(Emotional Speech) 분류는 자동적으로 화자의 감성 또는 정신 상태를 분류할 수 있는 다양한 음성 처리 방법의 하나이다. 감성 음성은 드러난 의미를 강화하기 위한 문장 내에서 내포된 숨겨진 의미를 탐색하기 위한 중요한 역할을 가진다. 일반적으로 감성 분류의 작업을 위한 종래의 기계 학습(Machine Learning) 방법은 특징 추출과 특징 분류의 두 가지 핵심적인 단계로 구성될 수 있다. 다양한 특징 추출 알고리즘은 음성으로부터 피치(Pitch), 낭독 속도(Speech Rate), 강도(Intensity), 멜 주파스 캡스트럼 계수(Mel Frequency Cepstral Coefficient, MFCC) 및 선형예측 캡스트럼 계수(Linear Phase Cepstral Coefficient, LPCC) 등과 같은 중요한 특징을 추출하기 위해 제안된다.

[0003] 특징 추출 방법에 더하여, 다양한 분류 방법은 음성 신호(Speech Signal)로부터 감성 상태를 분류하기 위해 적용된다. 특징의 서로 다른 유형에 기반한 분류 방법은 최상의 정확도를 성취하거나, 잘못된 오분류율(Misclassification Rate)을 낮추기 위해 선택된다. 특징 벡터의 유형을 위해, 예를 들어, SVM(Support Vector

Machine), KNN(K Nearest Neighbor) 및 GMM(Gaussian Mixture Model)는 변별 방식(Discriminative Way)에서 초평면(Hyper-Plane)을 찾기 위해 적용되거나, 또는 발생 방식(Generative Way)에서 분산 모델을 적합하게 하기 위해 적용된다. 시퀀스 특징 벡터를 위해 MFCC/HMM/CRF는 각각의 감성을 위한 모델을 생성하기 위해 적용된다.

[0004] 최상의 정확도를 얻기 위해 선택된 특징 추출 및 분류 방법은 감성 및 다른 많은 요인(언어, 성별, 나이 등)을 선택한다. 종래의 특징 추출 및 분류 방법은 가장 최상의 정확도를 획득하기 위한 최적화된 특징 세트 및 분류 방법을 찾기 위해 노력한다. 하지만, 이와 같은 과정에서 상관없는 정보를 배제하는 것과 음성 신호로부터 감성 정보를 획득하는 과정은 매우 복잡하며, 비선형을 가지기 때문에 최적화 문제를 해결하는 것이 쉽지 않다.

발명의 내용

해결하려는 과제

[0005] 본 발명이 해결하고자 하는 과제는 음성 인식에 있어서, 높은 정확도와 효율 결과를 얻을 수 있는 최적화된 감성 분류 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0006] 본 발명에 따른 감성 분류 장치는 수신된 트레이닝 음성 신호에서 잡음을 제거하는 전처리부, 잡음이 제거된 트레이닝 음성 신호에서 특징 벡터를 추출하는 특징 추출부 및 추출된 특징 벡터를 특징 비트 및 클래스 비트로 구분하고, 클래스 비트를 구성하는 하나 이상의 비트 중에서 유전자 알고리즘(Ganeric Algorithm)을 통해 선택되지 않은 비트를 왼쪽 그룹으로 구분하고, 선택된 비트를 오른쪽 그룹으로 구분하여 솔루션을 생성하고, 서포트 벡터 머신(Support Vector Machine)을 사용하여 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성하는 분류 처리부를 포함한다.

[0007] 분류 처리부는 오분류 확률(Misclassification Rate)를 사용하여 생성된 솔루션의 평가하여 정확도를 산출하고, 산출된 정확도에 k-폴드 교차 타당화(k-fold Cross Validation)를 적용하여 적합성 함수를 산출하여 솔루션을 평가하여 2진 분류 트리를 생성한다. 분류 처리부는 $accuracy = \frac{TP+TN}{2}$ 를 통해 정확도를 산출하며, TP는 트루 포지티브(True Positive)를 나타내고, 상기 TN은 트루 네거티브(True Negative)를 나타내며, $fitness = \frac{1}{5} Qaccuracy$ 를 통해 상기 적합성 함수를 산출할 수 있다. 그리고, 클래스 비트를 구성하는 하나 이상의 비트 각각은 서로 다른 감성 상태를 나타낼 수 있다.

[0008] 본 발명에 따른 음성 인식을 위한 감성 분류 방법은 먼저, 수신된 트레이닝 음성 신호의 잡음을 제거하고, 잡음이 제거된 트레이닝 음성 신호에서 특징 벡터를 추출한다. 그리고, 특징 벡터에 포함된 클래스 비트를 구성하는 하나 이상의 비트 중에서 유전자 알고리즘(Ganeric Algorithm)을 통해 선택되지 않은 비트를 왼쪽 그룹으로 구분하고, 선택된 비트를 오른쪽 그룹으로 구분하여 솔루션을 생성한다. 솔루션이 생성되면, 서포트 벡터 머신(Support Vector Machine)을 사용하여 솔루션을 평가하여 2진 분류 트리(Bianary Classification Tree)를 생성한다.

발명의 효과

[0009] 본 발명에 따른 감성 분류 장치 및 방법은 계층적 분류 스키마를 통해 자동적으로 감성 클래스가 분류되어 보다 높은 정확도 및 효율을 가질 수 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명에 따른 음성 인식을 위한 감성 분류 장치의 일 실시예를 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 2진 분류 트리 형식의 계층적 분류 스키마를 나타내는 도면이다.

도 3은 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 2진 분류 트리 형식의 계층적 분류 스

키마를 분류하는 과정을 설명하기 위한 도면이다.

도 4는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 트레이닝 단계를 나타내는 흐름도이다.

도 5는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 방법을 나타내는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하, 본 발명의 실시예를 첨부된 도면들을 참조하여 상세하게 설명한다. 본 명세서에서 사용되는 용어 및 단어들은 실시예에서의 기능을 고려하여 선택된 용어들로서, 그 용어의 의미는 발명의 의도 또는 관례 등에 따라 달라질 수 있다. 따라서 후술하는 실시예에서 사용된 용어는, 본 명세서에 구체적으로 정의된 경우에는 그 정의에 따르며, 구체적인 정의가 없는 경우는 당업자들이 일반적으로 인식하는 의미로 해석되어야 할 것이다.

[0012] 도 1은 본 발명에 따른 음성 인식을 위한 감성 분류 장치의 일 실시예를 나타내는 도면이다.

[0013] 도 1을 참조하면, 본 발명에 따른 감성 분류 장치는 전처리부(110), 특징 추출부(120), 분류 처리부(130)를 포함한다.

[0014] 전처리부(110)는 사용자로부터 수신된 음성 신호(Speech Signal)의 잡음을 제거한다. 특징 추출부(120)는 전처리부(110)에서 잡음이 제거된 음성 신호로부터 특징 벡터(Feature Vector)를 추출한다. 추출된 특징 벡터는 각각의 입력 샘플을 따르는 각각의 열, 각각의 특징을 따르는 행 그리고, 마지막 행은 샘플의 클래스 라벨을 가지는 특징 매트릭스 형태를 가진다.

[0015] 분류 처리부(130)는 추출된 특징 벡터로부터 최적의 분류 스키마(Classification Scheme)를 찾기 위해 2진 트리와 같은 구조를 이용한다. 다수의 노드로 구성되는 2진 트리는 분류의 그룹 및 그에 따른 분류의 솔루션을 나타내고, 각각의 리프 노드(Leaf Node)는 클래스 라벨(Class Label)을 나타낸다.

[0016] 분류 처리부(130)는 트레이닝 단계 및 분류 단계로 구분될 수 있다. 트레이닝 단계는 실제 사용자의 음성 신호를 분류하는 분류 단계를 수행하기 이전에 수행되어 분류 트리 모델을 생성함으로써, 분류 단계 수행 시 사용자의 음성 신호를 정확하게 분류할 수 있도록 한다.

[0017] 먼저, 트레이닝 단계에서,분류 처리부(130)는 추출된 특징 벡터를 선택된 특징을 보여주는 특징 비트와 입력 클래스를 2 그룹으로 분류하기 위한 클래스 비트 형태로 처리한다. 그리고, 분류 처리부(130)는 특징 추출부(120)에서 추출된 특징 벡터를 유전자 알고리즘(Genetic Algorithm, 이하 GA라 칭함)을 이용하여 분류하기 전에, GA의 큐(Queue)를 초기화한다. 큐 내의 첫 번째 GA는 가장 먼저 처리된다. GA를 통해 감성 분류의 반복을 수행하기 전에, GA 큐를 초기화한다. 본 발명에서 생성된 GA를 저장하는 큐를 사용한다.

[0018] 다음으로, 분류 처리부(130)는 클래스 비트의 값에 기초하여 추출된 특징 벡터의 데이터 세트를 GA를 이용하여 왼쪽 그룹 및 오른쪽 그룹의 두 그룹으로 분류한다. 그리고, 분류 처리부(130)는 GA를 이용하여 분류된 솔루션을 정확도를 평가한다. 분류 처리부(130)의 분류 과정 및 정확도 평가 과정은 후술하는 도 2 및 도 3에서 설명하도록 한다. 상술한 과정을 통해 분류 처리부(130)는 각 리프 노드에서 SVM 분류기의 파라미터인 클래스 비트 및 특징 비트로 구성된 각각의 GA로부터 최적 솔루션을 사용하여 구성된 2진 분류 트리를 생성한다.

[0019] 분류 단계에서, 분류 처리부(130)는 트레이닝 단계에서 생성된 2진 분류 트리에 따라 특징 벡터를 순차적으로 분류한다. 분류 처리부(130)는 적용된 SVM 분류기로 구성한 입력 신호를 분류하고, 입력 신호를 분류하는 분류 트리로부터 원본 클래스 라벨의 하나까지 각 리프 노드에 적용한 SVM 분류기, 클래스 비트 및 특징 비트로 구성된 입력 신호를 분류한다.

[0020] 도 2는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 2진 분류 트리 형식의 계층적 분류 스키마를 나타내는 도면이다.

[0021] 도 2를 참조하면, 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치는 최고의 정확도 및 효율을 가지는 분류 결과를 얻기 위하여 2진 분류 트리(Bianary Classification Tree)와 같은 최적의 분류 스키마를 찾기 위해 설계된다. GA와 SVM(Support Vector Machine)은 최적의 솔루션을 검색 및 평가하는데 활용되는 두 개의 핵심 방법이다. 2진 분류 트리는 최종 결과를 대표하도록 사용될 수 있다.

[0022] 다수의 노드로 구성되는 2진 트리는 분류의 그룹 및 그에 따른 분류의 솔루션을 나타내고, 각각의 리프 노드(Leaf Node)는 클래스 라벨(Class Label)을 나타낸다. 도 2a의 2진 트리의 일 실시예는 7개의 서로 다른 클래스

로 구성되는 데이터 세트(201)의 구성을 나타낸다. 데이터 세트(210)에 포함된 7개의 서로 다른 클래스(1,2,3,4,5,6,7) 각각은 서로 다른 감성 상태를 나타낸다. 예를 들어, 데이터 세트(210)의 1은 분노(Angry)를 나타내는 클래스이고, 2는 기쁨(Happy)을 나타내는 클래수 일 수 있다.

[0023] 도 2의 2진 트리에서 7개의 서로 다른 클래스로 구성되는 데이터 세트(210)는 2개의 그룹(220,230)으로 결합 및 분류된다. 2개의 그룹(220,230)은 클래스 1,2,4로 구성된 왼쪽 그룹(220) 및 남아있는 클래스 3,5,6,7로 구성된 오른쪽 그룹(230)을 나타낸다.

[0024] 다음 단계에서, 왼쪽 그룹(220) 및 오른쪽 그룹(230)은 각각 다시 2개의 그룹으로 분류된어, 총 4개의 그룹(221,222,231,232)로 분류된다. 이러한 과정은 각각의 그룹이 단지 하나의 클래스를 가지거나 또는 모든 클래스들이 리프 노드로 배열될 때까지 연속적으로 수행된다. 2진 트리를 두 개의 그룹으로 분류하는 과정을 반복적으로 수행하여 개별 리프 노드로 분류하는 방법은 GA를 통해 분류하게 되며, 구체적인 방법은 후술하는 도 2b를 통해 설명하도록 한다.

[0025] 도 3은 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 2진 분류 트리 형식의 계층적 분류 스키마를 분류하는 과정을 설명하기 위한 도면이다.

[0026] 도 2 및 3을 참조하면, 2진 분류 트리의 일 실시예를 분류하는 과정은 유전자 알고리즘(GA)를 이용한다. GA의 중요한 두 가지 요구 조건은 다음 세대를 위해 유지되기 위해서 어떻게 특정한 문제의 해결을 나타내는가와 어떻게 생성된 솔루션을 평가하는가이다. GA는 두 그룹 사이에서 최고의 구분을 찾는 두 개의 임무를 가진다. 그리고, GA는 도 2a와 나타난 바와 같이 두 개의 관련 파트로 구성되는 솔루션(염색체)를 위한 가장 적절한 특징을 가진다.

[0027] GA의 중요한 첫 번째 요구를 위한 방법으로, 왼쪽 파트(310)는 추출된 특징 벡터의 특징 비트를 나타내며, 왼쪽 파트(310)에서 각각의 클래스를 나타내는 비트 1(311)은 선택된 비트를 나타내고, 비트 0(312)는 선택되지 않은 비트를 나타낸다. 왼쪽 파트(310)의 길이는 도 1의 특징 추출부(120) 이후의 특징의 개수이다.

[0028] 오른쪽 파트(320)는 도 2의 데이터 세트(210) 그룹의 클래스 비트를 나타내며, 오른족 파트(320)에서 각각의 클래스를 나타내는 비트 0(321)은 선택되지 비트로서 왼쪽 그룹(210)에 속하고, 각각의 클래스를 나타내는 비트 1(322)는 오른쪽 그룹(220)에 속한다. 오른쪽 파트(320)의 길이는 클래스들의 개수이다. 도 2의 데이터 세트(210)에 포함된 7개의 클래스(1,2,3,4,5,6,7)은 오른쪽 파트(320)에 나타난 비트와 같이 0,0,1,0,1,1,1에 대응하여, 7개의 클래스(1,2,3,4,5,6,7) 중 1,2,4는 비트 0으로 왼쪽 그룹(210)으로 분류되고, 7개의 클래스(1,2,3,4,5,6,7) 중 3,5,6,7은 비트 1로 오른쪽 그룹(220)으로 분류된다.

[0029] GA의 중요한 두 번째 요구는 어떻게 생성된 솔루션 또는 GA 적합성 함수의 정의를 평가하는가이다. GA의 일반적인 임무는 적합성 함수를 최소화하기 위해 최적의 솔루션을 찾는다. 본 발명에서, 분류된 그룹을 선택된 특징과 함께 분류하기 위해 SVM을 사용한다. 그리고, 생성된 솔루션의 효율을 평가하기 위한 측정으로 오분류율(Misclassification Rate)을 사용한다. 도 2에 기재된, 2개의 클래스로 분류하는 문제에서, 결과는 표 1에 따르는 오차 행렬에 의해 나타내어진다.

표 1

[0030] 오차 행렬

Classes	Class 1	Class 2
Class 1	True Positives (TP)	False Negative (FN)
Class 2	False Positives (FP)	True Negative (TN)

[0031] 정확도(accuracy)는 수학식 1과 같이 트루 포지티브(True Positive) 및 트루 네거티브(True Negative)의 평균값에 의해 연산 된다.

수학식 1

$$accuracy = \frac{TP + TN}{2}$$

[0032]

[0033] 각각의 SVM의 초기화 과정은 무작위로 수행되어 정확도는 일관되지 않기 때문에, 평가의 신뢰도를 증가시키기 위해 k-폴드 교차 타당화(k-fold Cross Validation)를 적용한다. 이를 통해 적합성 함수(Fitness Function)는 수학식 2와 같이 교차 타당화의 정확도 평균으로 정의된다.

수학식 2

$$fitness = \frac{1}{5}\sum accuracy$$

[0034]

[0035] 도 4는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 장치의 트레이닝 단계를 나타내는 흐름도이다.

[0036] 도 1 내지 도 4를 참조하면, 전처리부(110)는 사용자로부터 수신된 트레이닝을 위한 음성 신호의 잡음을 제거한다(401). 그리고, 특징 추출부(120)는 전처리부(110)에서 잡음이 제거된 음성 신호로부터 특징 벡터(Feature Vector)를 추출한다(402). 추출된 특징 벡터는 각각의 입력 샘플을 따르는 각각의 열, 각각의 특징을 따르는 행 그리고, 마지막 행은 샘플의 클래스 라벨을 가지는 특징 매트릭스 형태를 가진다. 그리고, GA 큐 초기화를 초기화한다(403).

[0037] 다음으로, 분류 처리부(130)는 큐가 비어있는지 여부를 판단한다(404). 만약, 큐가 비어있지 않다면, 첫 번째 GA는 처리하기 위해 선택된다(405). 분류 처리부(130)는 선택된 GA의 파라미터를 초기화한다(406). 이때, GA 파라미터의 염색체 개체군(Population of Chromosome)은 무작위로 생성된다. 그리고, 분류 처리부(130)는 생성된 염색체 또는 솔루션을 적용된 SVM 분류기를 통해 평가한다(407). 분류 처리부(130)는 클래스 비트를 사용하여 클래스 라벨을 왼쪽 그룹/오른쪽 그룹 및 특징 비트 사용으로 분열시킨다. 그리고, 분류를 위해 후보 특징을 선택한다. 그리고, GA의 적합성은 도 3에 기재된 SVM의 정확도 활용을 통해 추정된다. 405 단계 내지 407 단계는 도 2 및 도 3의 과정을 나타낸다.

[0038] 그리고, 분류 처리부(130)는 적합성 값이 정지 상태를 만족하는지 여부를 판단한다(408). 만약 적합성 값이 정지 상태를 만족하지 않는다면, 2진 분류 트리가 부적합한것으로 판단하여 염색체의 새로운 개체군은 적용된 크로스 오버(Crossover) 과정(409) 및 변화(Mutation) 과정(410)에 의해 재생성된다. GA의 내부 반복 처리과정은 정지 상태가 만족될 때까지 새로운 솔루션을 평가하기 위해 계속된다.

[0039] 만약, 408 단계에서, 적합성 함수의 값이 현저하게 변하지 않거나 또는 생성의 수가 사전에 정의된 임계값에 도달하면, 2진 분류 트리가 적합한 것으로 판단하여 정지 상태는 만족되고, 최적 솔루션은 새로운 왼쪽 GA 및 오른쪽 GA를 생성하도록 처리된다(411). 만약 왼쪽 또는 오른쪽 그룹이 단지 하나의 클래스를 가진다면, 하나의 클래스는 결국 하나의 리프 노드이기 때문에 이후의 단계(412,414)를 처리하지 않는다. 그렇지 않으면, 분류 처리부(130)는 새로운 GA를 큐에 넣고 새로운 반복을 수행한다(413,415).

[0040] 404 단계에서 큐가 비어있다면, GA로부터의 최적 솔루션을 사용한 결과 처리(416)을 통해 상태 분류를 위해 사용되는 2진 분류 트리의 모델을 생성한다(417).

[0041] 도 5는 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 방법을 나타내는 흐름도이다.

[0042] 도 5를 참조하면, 본 발명의 일 실시예에 따른 음성 인식을 위한 감성 분류 방법은 먼저, 수신된 트레이닝을 위한 음성 신호에서 잡음을 제거한다(501). 그리고, 잡음이 제거된 음성 신호에서 감성 인식을 위한 특징 벡터를 추출한다(502).

[0043] 특징 벡터가 추출되면, 추출된 특징 벡터를 선택된 특징을 보여주는 특징 비트 및 입력 클래스를 2 그룹으로 분류하기 위한 클래스 비트 형태로 처리한다(503). 그리고, 클래스 비트를 GA를 통해 왼쪽 그룹 및 오른쪽 그룹으

로 분류한다(504). 분류 처리부(130)는 클래스 비트의 값에 기초하여 추출된 특징 벡터의 데이터 세트를 GA를 이용하여 왼쪽 그룹 및 오른쪽 그룹의 두 그룹으로 분류한다. 2진 트리는 서로 다른 클래스로 구성되는 데이터 세트의 구성을 나타낸다. 그리고, 데이터 세트에 포함된 서로 다른 클래스 각각은 서로 다른 감성 상태를 나타낸다. 클래스 비트를 분류하는 과정은 유전자 알고리즘(GA)를 이용한다. 클래스 비트에서 각각의 클래스를 나타내는 비트 0은 선택되지 비트로서 왼쪽 그룹에 속하고, 각각의 클래스를 나타내는 비트 1는 오른쪽 그룹에 속한다. 오른쪽 파트의 길이는 클래스들의 개수이다.

[0044] 다음으로 GA를 통해 왼쪽 그룹 및 오른쪽 그룹으로 분류된 솔루션을 평가한다(505). GA의 중요한 두 번째 요구는 어떻게 생성된 솔루션 또는는 GA 적합성 함수의 정의를 평가하는가이다. GA의 일반적인 임무는 적합성 함수를 최소화하기 위해 최적의 솔루션을 찾는다. 본 발명에서, 분류된 그룹을 선택된 특징과 함께 분류하기 위해 SVM을 사용한다. 그리고, 생성된 솔루션의 효율을 평가하기 위한 측정으로 오분류율(Misclassification Rate)을 사용한다. 2개의 클래스로 분류하는 문제에서, 결과는 도 3의 표 1에 따르는 오차 행렬에 의해 나타내어진다. 그리고, 정확도(accuracy)는 도 3의 수학식 1과 같이 트루 포지티브(True Positive) 및 트루 네거티브(True Negative)의 평균값에 의해 연산된다. 각각의 SVM의 초기화 과정은 무작위로 수행되어 정확도는 일관되지 않기 때문에, 평가의 신뢰도를 증가시키기 위해 k-폴드 교차 타당화(k-fold Cross Validation)을 적용한다. 이를 통해 적합성 함수(Fitness Function)는 도 3의 수학식 2와 같이 교차 타당화의 정확도 평균으로 정의된다.

[0045] 다음으로, 상술한 과정을 통해 각 리프 노드에서 SVM 분류기의 파라미터인 클래스 비트 및 특징 비트로 구성된 각각의 GA로부터 최적 솔루션을 사용하여 구성된 2진 분류 트리를 생성한다(506).

[0046] 이상 바람직한 실시예를 들어 본 발명을 상세하게 설명하였으나, 본 발명은 전술한 실시예에 한정되지 않고, 본 발명의 기술적 사상의 범위 내에서 당분야에서 통상의 지식을 가진자에 의하여 여러 가지 변형이 가능하다.

부호의 설명

[0047] 110: 전처리부

120: 특징 추출부

130: 분류 처리부

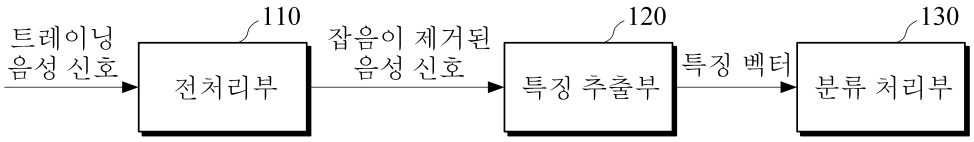
210: 데이터 세트

220: 왼쪽 그룹

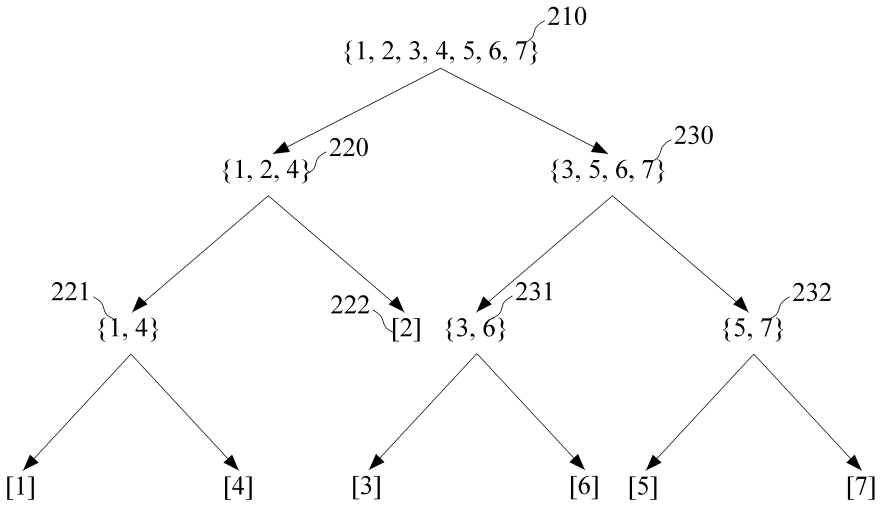
230: 오른쪽 그룹

도면

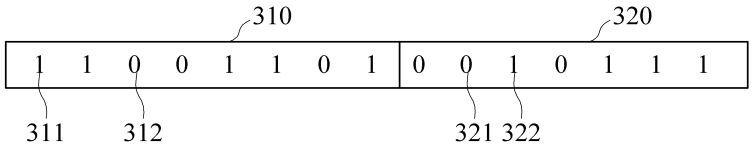
도면1



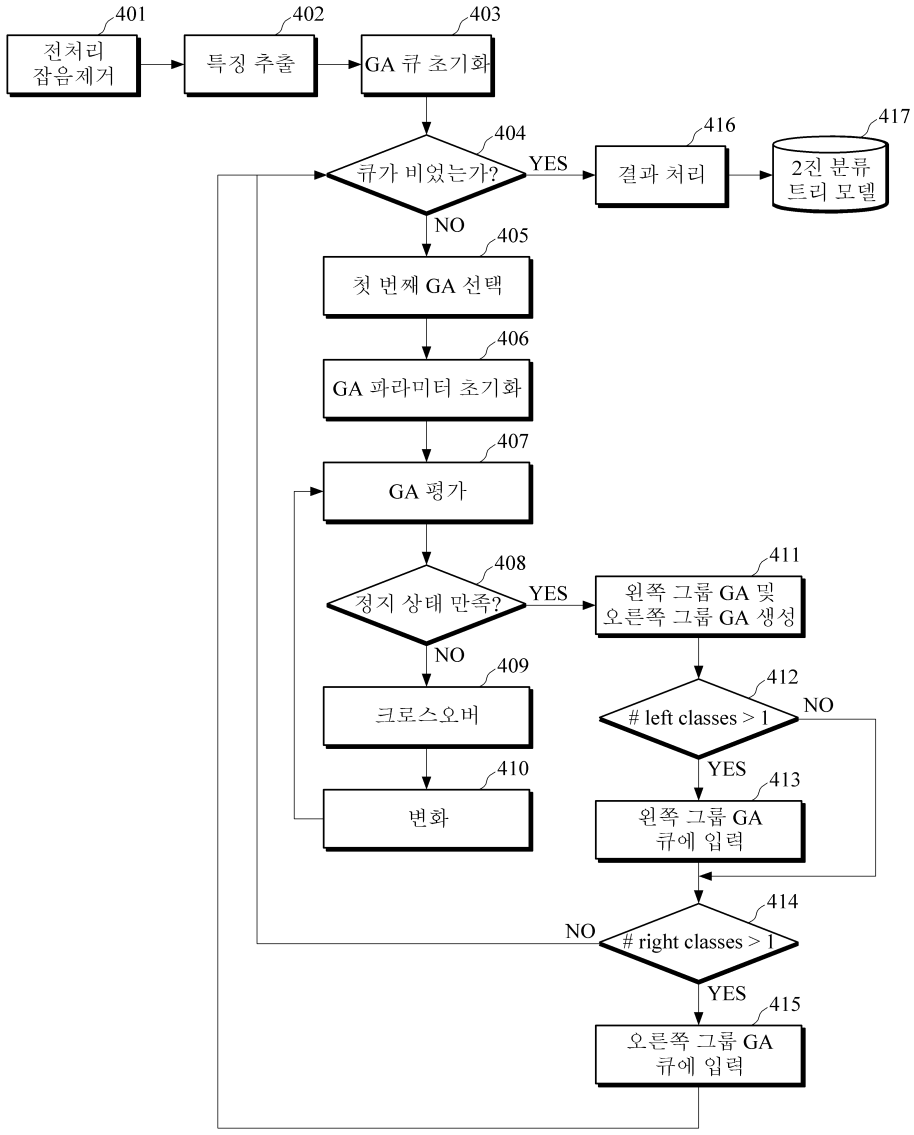
도면2


210
{1, 2, 3, 4, 5, 6, 7}
220
{1, 2, 4}
230
{3, 5, 6, 7}
221
{1, 4}
222
[2]
231
{3, 6}
232
{5, 7}
[1]
[4]
[3]
[6]
[5]
[7]

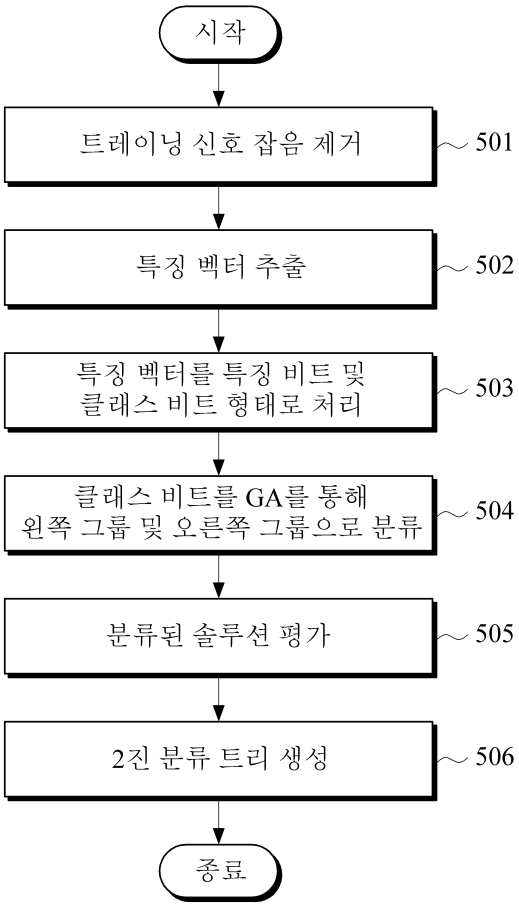
도면3


310
320
1 1 0 0 1 1 0 1
0 0 1 0 1 1 1
311
312
321
322

도면4


401
전처리
잡음제거
402
특징 추출
403
GA 큐 초기화
404
큐가 비었는가?
YES
416
결과 처리
417
2진 분류
트리 모델
NO
405
첫 번째 GA 선택
406
GA 파라미터 초기화
407
GA 평가
408
정지 상태 만족?
YES
411
왼쪽 그룹 GA 및
오른쪽 그룹 GA 생성
NO
409
크로스오버
412
left classes > 1
NO
YES
410
변화
413
왼쪽 그룹 GA
큐에 입력
414
NO
right classes > 1
YES
415
오른쪽 그룹 GA
큐에 입력

도면5


시작
트레이닝 신호 잡음 제거
501
특징 벡터 추출
502
특징 벡터를 특징 비트 및
클래스 비트 형태로 처리
503
클래스 비트를 GA를 통해
왼쪽 그룹 및 오른쪽 그룹으로 분류
504
분류된 솔루션 평가
505
2진 분류 트리 생성
506
종료