



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0087363
(43) 공개일자 2019년07월24일

(51) 국제특허분류(Int. Cl.)
G10L 15/14 (2006.01) B64C 39/02 (2006.01)
G05D 1/02 (2006.01) G10L 15/20 (2006.01)
G10L 19/038 (2013.01)
(52) CPC특허분류
G10L 15/144 (2013.01)
B64C 39/024 (2013.01)
(21) 출원번호 10-2019-0085039(분할)
(22) 출원일자 2019년07월15일
심사청구일자 2019년07월15일
(62) 원출원 특허 10-2017-0105021
원출원일자 2017년08월18일
심사청구일자 2017년08월18일

(71) 출원인
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
장경희
서울특별시 마포구 백범로 205, 101동 1110호(신공덕동, 마포폰트하우스)
석림
인천광역시 남구 소성로 40, 1기숙사(학익동, 인하대학교기숙사)
하유경
인천광역시 남구 소성로 40, 1기숙사(학익동, 인하대학교기숙사)
(74) 대리인
양성보

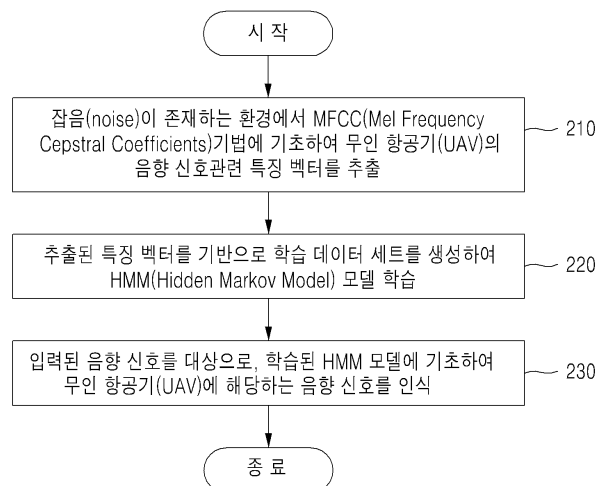
전체 청구항 수 : 총 3 항

(54) 발명의 명칭 실질 잡음 환경에서 MFCC 기법을 이용한 HMM 기반 무인 항공기 음향 인식 방법 및 시스템

(57) 요약

실질 잡음이 많은 환경에서 MFCC 기법을 이용한 HMM 기반 무인 항공기 음향 인식 방법 및 시스템이 개시된다. 무인 항공기(Unmanned Aerial Vehicle)의 음향을 인식하는 방법에 있어서, 잡음(noise)이 존재하는 환경에서 MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기(UAV)의 음향 신호관련 특징 벡터를 추출하는 단계, 추출된 특징 벡터를 기반으로 생성된 학습 데이터 세트에 기초하여 HMM(hidden Markov Model) 모델을 학습시키는 단계, 및 입력된 음향 신호를 대상으로, 학습된 HMM 모델에 기초하여 무인 항공기(UAV)에 해당하는 음향 신호를 인식하는 단계를 포함할 수 있다.

대표도 - 도2



(52) CPC특허분류

G05D 1/0255 (2013.01)

G10L 15/20 (2013.01)

G10L 19/038 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2015R1D1A1A01057190

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 기본연구

연구과제명 [Ezbaro] 불법 드론 감지 및 퇴치를 위한 감시 드론 애드혹 네트워크 설계 연구

기 여 율 1/1

주관기관 인하대학교 산학협력단

연구기간 2016.11.01 ~ 2017.10.31

명세서

청구범위

청구항 1

무인 항공기(Unmanned Aerial Vehicle)의 음향을 인식하는 방법에 있어서,

잡음(noise)이 존재하는 환경에서 MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기의 음향 신호 관련 특징 벡터를 추출하는 단계;

상기 특징 벡터를 기반으로 생성된 학습 데이터 세트에 기초하여 HMM(hidden Markov Model) 모델을 학습시키는 단계; 및

입력된 음향 신호를 대상으로, 상기 학습된 HMM 모델에 기초하여 무인 항공기에 해당하는 음향 신호를 인식하는 단계를 포함하고,

상기 특징 벡터를 추출하는 단계는,

무인 항공기에 해당하는 음향 신호를 기반으로 켈스트럴 계수(Cepstral Coefficient)를 계산하는 단계;

상기 켈스트럴 계수를 기반으로 제 1 델타 계수와 제 2 델타 계수를 계산하는 단계;

상기 켈스트럴 계수와 제 1 델타 계수로 제 1 MFCC를 구성하는 단계;

상기 켈스트럴 계수, 제 1 델타 계수 및 제 2 델타 계수로 제 2 MFCC를 구성하는 단계; 및

상기 제 1 MFCC 및 제 2 MFCC를 통하여 상기 특징 벡터를 추출하는 단계를 포함하고,

상기 켈스트럴 계수를 계산하는 단계는,

상기 음향 신호를 대상으로 고주파 필터링(highpass filtering)을 수행하는 프리 엠퍼시스(pre-emphasis) 처리 단계;

상기 프리 엠퍼시스 처리된 신호를 대상으로 프레임링(framing), 윈도우링(windowing) 및 고속 푸리에 변환(FFT)을 수행하여 상기 음향 신호에 해당하는 특징 벡터를 나타내는 스펙트로그램(spectrogram)을 생성하는 단계;

상기 고속 푸리에 변환된 신호를 대상으로 멜 스케일 필터 뱅크(Mel-scale filter bank)를 통과시키는 단계; 및

상기 멜 스케일 필터 뱅크를 통과한 신호를 대상으로 이산 코사인 변환(DCT)을 수행하여 상기 켈스트럴 계수를 계산하는 단계를 포함하고,

상기 스펙트로그램을 생성하는 단계는,

상기 프레임링을 통해 프레임 쉬프트를 수행하여, 상기 프리 엠퍼시스 처리된 신호에서 인접 프레임들 간의 오버랩을 유도하는 단계;

상기 오버랩된 프레임들을 대상으로, 상기 프리 엠퍼시스 처리된 신호에서 각 프레임에 대해 상기 윈도우링을 수행하는 단계; 및

상기 윈도우링된 신호를 대상으로, 상기 FFT 처리를 수행하여, 상기 스펙트로그램을 생성하는 단계를 포함하고,

상기 HMM 모델을 학습시키는 단계는,

HMM 모델을 초기화하는 단계;

순방향 및 역방향 알고리즘을 이용하여, 상기 특징 벡터를 기반으로 생성된 학습 데이터 세트를 기초로 순방향 우도(forward likelihood) 및 역방향 우도(backward likelihood)를 계산하는 단계; 및

상기 순방향 우도 및 역방향 우도를 기초로 상기 초기화된 HMM 모델의 파라미터들을 재추정하여, 상기 초기화된 HMM 모델을 업데이트하는 단계를 포함하는 무인 항공기의 음향 인식 방법.

청구항 2

제1항에 있어서,

상기 특징 벡터를 추출하는 단계는,

미리 지정된 기준 시간동안 cepstral domain의 포락선을 기반으로 종류 별 무인 항공기의 음향 신호관련 특징 벡터를 추출하는 무인 항공기의 음향 인식 방법.

청구항 3

무인 항공기(Unmanned Aerial Vehicle)의 음향을 인식하는 시스템에 있어서,

잡음(noise)이 존재하는 환경에서 MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기의 음향 신호 관련 특징 벡터를 추출하는 특징 벡터 추출부;

상기 특징 벡터를 기반으로 생성된 학습 데이터 세트에 기초하여 HMM(hidden Markov Model) 모델을 학습시키는 모델 학습부; 및

입력된 음향 신호를 대상으로, 상기 학습된 HMM 모델에 기초하여 무인 항공기에 해당하는 음향 신호를 인식하는 음향 인식부를 포함하고,

상기 특징 벡터 추출부는,

무인 항공기에 해당하는 음향 신호를 기반으로 cepstral 계수(Cepstral Coefficient)를 계산하고,

상기 cepstral 계수를 기반으로 제 1 델타 계수와 제 2 델타 계수를 계산하고,

상기 cepstral 계수와 제 1 델타 계수로 제 1 MFCC를 구성하고,

상기 cepstral 계수, 제 1 델타 계수 및 제 2 델타 계수로 제 2 MFCC를 구성하고,

상기 제 1 MFCC 및 제 2 MFCC를 통하여, 상기 특징 벡터를 추출하고,

상기 특징 벡터 추출부는,

상기 음향 신호를 대상으로 고주파 필터링(highpass filtering)을 수행하는 프리 엠퍼시스(pre-emphasis) 처리하고,

상기 프리 엠퍼시스 처리된 신호를 대상으로 프레이밍(framing), 윈도우링(windowing) 및 고속 푸리에 변환(FFT)을 수행하여 상기 음향 신호에 해당하는 특징 벡터를 나타내는 스펙트로그램(spectrogram)을 생성하고,

상기 고속 푸리에 변환된 신호를 대상으로 멜 스케일 필터 뱅크(Mel-scale filter bank)를 통과시키고,

상기 멜 스케일 필터 뱅크를 통과한 신호를 대상으로 이산 코사인 변환(DCT)을 수행하여 상기 cepstral 계수를 계산하고,

상기 특징 벡터 추출부는,

상기 프레이밍을 통해 프레임 쉬프트를 수행하여, 상기 프리 엠퍼시스 처리된 신호에서 인접 프레임들 간의 오버랩을 유도하고,

상기 오버랩된 프레임들을 대상으로, 상기 프리 엠퍼시스 처리된 신호에서 각 프레임에 대해 상기 윈도우링을 수행하고,

상기 윈도우링된 신호를 대상으로, 상기 FFT 처리를 수행하여, 상기 스펙트로그램을 생성하고,

상기 모델 학습부는,

HMM 모델을 초기화하고,

순방향 및 역방향 알고리즘을 이용하여, 상기 특징 벡터를 기반으로 생성된 학습 데이터 세트를 기초로 순방향 우도(forward likelihood) 및 역방향 우도(backward likelihood)를 계산하고,

상기 순방향 우도 및 역방향 우도를 기초로 상기 초기화된 HMM 모델의 파라미터들을 재추정하여, 상기 초기화된 HMM 모델을 업데이트하는 무인 항공기의 음향 인식 시스템.

발명의 설명

기술 분야

[0001] 본 발명의 실시예들은 잡음이 존재하는 실질 환경에서 여러 음향 신호들 중에서 드론(drone) 등의 무인 항공기(Unmanned Aerial Vehicle)에 해당하는 음향 신호를 인식하여 무인 항공기를 식별하는 기술에 관한 것이다.

배경 기술

[0002] 무인 항공기(Unmanned Aerial Vehicle, UAV)인 드론(drone)의 이용 분야가 증가하면서, 드론은 군사용뿐만 아니라 상업용, 방송용으로 많이 이용되고 있다. 이러한 무인 항공기는 대부분 소형이며, 소형 무인항공기의 경우, 통제되지 않은 영공(airspace)에서 운행되므로, 공공 안전 및 개인 정보 보호에 심각한 위협이 될 수 있다. 예컨대, 소형 무인 항공기는 화학 무기, 생물학 무기, 또는 핵무기 등을 운반할 수 있으며, 마약을 밀수입하거나 국경 너머로 다른 불법 활동을 수행하기 위해 이용될 수 있다.

[0003] 이처럼, 드론 등의 소형 무인 항공기가 사람의 생명을 위협하거나, 사생활을 위협하는 용도, 마약 밀수입 등의 불법적인 용도 및 공공안전을 위협하는 용도로 활용될 가능성이 존재하기 때문에, 드론을 탐지하고 추적하는 기술이 점차로 중요해 지고 있다. 현재 드론을 탐지 및 추적하는 기술은 레이더(radar)를 기반으로 한다.

[0004] 그러나, 레이더 기반으로 운행 중인 드론의 위치를 탐지 및 추적하는 경우, 드론의 크기가 작으면 지상 가까이 에 위치하는 드론의 위치를 탐지 또는 추적하기 어렵고, 드론이 날아갈 수 있는 사각지대를 내포하는 문제가 존재하며, 또한 상당히 고가의 시스템 설치 및 운영비용이 소요되는 문제가 있다.

[0005] 이에 따라, 드론(drone)이 소형이든, 중형이든, 드론의 크기에 관계없이 드론의 실시간 위치를 보다 정확하게 탐지 및 추적할 수 있는 기술이 요구된다. 즉, 레이더 기반의 드론의 위치를 추적하는 기술 이외 드론의 위치 탐지 및 추적 기술이 요구된다.

[0006] 한국공개특허 제10-2013-0120176호는 무인 항공기의 위치 추적 방법에 관한 것으로, GPS 교란 또는 GPS 수신기의 고장에 의해, 무인 항공기의 위치를 파악할 수 없는 경우, 지상 통제 장비가 영상 감지기를 조정하여 사전에 알고 있는 고정 표적(절대 좌표)을 촬영하도록 함으로써, 무인 항공기의 상태 정보(피치, 롤, 헤딩, 고도)와 영상 감지기의 정보(방위각(Azimuth) 및 고각(Elevation))를 이용하여 고정 표적을 기준점으로 역으로 무인 항공기의 위치를 추적하는 기술을 개시하고 있다.

선행기술문헌

비특허문헌

[0007] (비특허문헌 0001) [1] I. Patel and Y. S. Rao, "Speech recognition using Hidden Markov Model with MFCC-Subband technique," in Proc. International Conference on Recent Trends in Information, Telecommunication and Computing, pp.168-172, Mar. 2010.

(비특허문헌 0002) [2] G. B. Singh and H. Song, "Using hidden Markov models in vehicular crash detection," IEEE Transactions on Vehicular Technology, vol. 58, no. 3, pp. 1119-1128, Mar. 2009.

발명의 내용

해결하려는 과제

[0008] 본 발명은 무인 항공기(UAV)의 크기에 제약받지 않고(예컨대, 드론의 크기에 관계없이), 잡음이 존재하는 실질 환경에서 다양한 음향 신호가 입력으로 들어온 경우에, 입력된 음향 신호 중에서 드론에 해당하는 음향 신호를 정확하게 인식하는 기술에 관한 것이다. 즉, 실질 환경에 존재하는 여러 음향 신호들 중에서 드론에 해당하는 음향 신호를 인식하여 드론을 탐지 및 추적 가능하도록 하는 기술에 관한 것이다.

과제의 해결 수단

[0009] 무인 항공기(Unmanned Aerial Vehicle)의 음향을 인식하는 방법에 있어서, 잡음(noise)이 존재하는 환경에서

MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기(UAV)의 음향 신호관련 특징 벡터를 추출하는 단계, 추출된 상기 특징 벡터를 기반으로 생성된 학습 데이터 세트에 기초하여HMM(hidden Markov Model) 모델을 학습시키는 단계, 및 입력된 음향 신호를 대상으로, 상기 학습된 HMM 모델에 기초하여 무인 항공기(UAV)에 해당하는 음향 신호를 인식하는 단계를 포함할 수 있다.

[0010] 일측면에 따르면, 상기 특징 벡터를 추출하는 단계는, 서로 다른 개수의 캡스트럴 계수로 구성된 제1 MFCC 및 제2 MFCC에 기초하여 상기 특징 벡터를 추출할 수 있다.

[0011] 다른 측면에 따르면, 상기 특징 벡터를 추출하는 단계는, 미리 지정된 기준 시간동안 캡스트럴 도메인(cepstral domain)의 포락선을 기반으로 종류 별 무인 항공기의 음향 신호관련 특징 벡터를 추출할 수 있다.

[0012] 또 다른 측면에 따르면, 상기 특징 벡터를 추출하는 단계는, 무인 항공기에 해당하는 음향 신호를 대상으로 고주파 필터링(highpass filtering)을 수행하는 프리 엠퍼시스(pre-emphasis) 처리 단계, 상기 프리 엠퍼시스 처리된 신호를 대상으로 프레임링(framing), 윈도우링(windowing) 및 고속 푸리에 변환(FFT)을 수행하여 상기 무인 항공기(UAV)의 음향 신호에 해당하는 특징 벡터를 나타내는 스펙트로그램(spectrogram)을 생성하는 단계, 상기 고속 푸리에 변환된 신호를 대상으로 멜 스케일 필터 뱅크(Mel-scale filter bank)를 통과시키는 단계, 상기 멜 스케일 필터 뱅크를 통과한 신호를 대상으로 이산 코사인 변환(DCT)을 수행하여 캡스트럴 계수(Cepstral Coefficient)를 계산하는 단계, 및 계산된 상기 캡스트럴 계수에 기초하여 MFCC를 구성하는 단계를 포함할 수 있다.

[0013] 무인 항공기(Unmanned Aerial Vehicle)의 음향을 인식하는 시스템에 있어서, 잡음(noise)이 존재하는 환경에서 MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기(UAV)의 음향 신호관련 특징 벡터를 추출하는 특징 벡터 추출부, 추출된 상기 특징 벡터를 기반으로 생성된 학습 데이터 세트에 기초하여 HMM(hidden Markov Model) 모델을 학습시키는 모델 학습부, 및 입력된 음향 신호를 대상으로, 상기 학습된 HMM 모델에 기초하여 무인 항공기(UAV)에 해당하는 음향 신호를 인식하는 음향 인식부를 포함할 수 있다.

발명의 효과

[0014] 본 발명은, MFCC(Mel Frequency Cepstral Coefficient) 기법에 기초하여 무인 항공기(UAV)의 음향 신호에서 특징 벡터를 추출하고, 추출된 특징 벡터에 기초하여 HMM(Hidden Markov Model)을 학습시키고, 학습된 모델을 기반으로 실질 환경에서 입력된 음향 신호들 중 무인 항공기의 프로펠러(propeller)에서 방출된 소리를 식별함으로써, 무인 항공기(UAV)의 크기에 제약받지 않고(예컨대, 드론의 크기에 관계없이), 잡음이 존재하는 실질 환경에서 무인 항공기(UAV)를 인식할 수 있다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일실시예에 있어서, 무인 항공기의 음향 인식 시스템의 구성을 도시한 블록도이다.
 도 2는 본 발명의 일실시예에 있어서, 무인 항공기의 음향 인식 방법의 동작을 도시한 흐름도이다.
 도 3은 본 발명의 일실시예에 있어서, 다양한 음향 신호 별 스펙트로그램(spectrogram)을 나타낼 수 있다.
 도 4는 본 발명의 일실시예에 있어서, MFCC 기법에 기초하여 특징 벡터를 추출하는 동작을 도시한 흐름도이다.
 도 5는 본 발명의 일실시예에 있어서, 특징 벡터 추출부의 세부 구성을 도시한 블록도이다.
 도 6은 본 발명의 일실시예에 있어서, BWA(Baum-Welch Algorithm) 기반의 HMM 모델을 학습시키는 동작을 도시한 흐름도이다.
 도 7은 본 발명의 일실시예에 있어서, 학습된 HMM 모델에 기초하여 원하는 음향 신호를 인식하는 동작을 도시한 블록도이다.
 도 8은 본 발명의 일실시예에 있어서, 제1 및 2 MFCC기법을 이용하여 추출된 특징 벡터들을 기반으로 하는 평균 인식률을 도시한 그래프이다.
 도 9는 본 발명의 일실시예에 있어서, SNR에 대한 인식률 성능을 도시한 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0016] 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0018] 본 실시예들은 드론(drone) 등의 무인 항공기(Unmanned Aerial Vehicle)를 인식하는 기술에 관한 것으로서, 특히, 무인 항공기의 프로펠러(propeller)가 구동됨에 따라 방출되는 음향 신호(즉, 프로펠러 소리와 관련된 음향 어레이 신호)를 기반으로 MFCC(Mel Frequency Cepstral Coefficient) 기법에 기초하여 특징 벡터를 추출하고, 추출된 특징 벡터를 이용하여 HMM(Hidden Markov Model)을 학습시켜 생성된 학습 데이터 세트를 이용하여 시스템으로 입력되는 음향 신호가 드론 등 무인 항공기의 프로펠러 소리에 해당하는 것인지를 인식하는 기술에 관한 것이다. 즉, 본 실시예들에서 HMM 모델은 드론 등의 무인 항공기의 음향 신호를 인식하기 위한 분류기(classifier)로 이용될 수 있다.
- [0019] 본 실시예들에서, '무인 항공기(Unmanned Aerial Vehicle)'는 조종사가 탑승하지 않고 지정된 임무를 수행할 수 있도록 제작된 비행체로서, 드론(drone) 등이 포함될 수 있다.
- [0020] 본 실시예들에서, '음향 신호'는 오디오 신호(audio sound)를 나타내는 것으로서, 예컨대, 드론(drone)에 해당하는 음향 신호(sound)는 드론의 프로펠러가 회전하면서 방출하는 음향 신호(acoustic sound)를 나타낼 수 있다.
- [0021] 본 실시예들에서, '제2 MFCC'는 36개 켈스트럴 계수(MFCC)로 구성되는 것으로 설명하나, 이는 실시예에 해당되며, '제2 MFCC'는 36개보다 많은 켈스트럴 계수로 구성될 수 있다.
- [0023] 도 1은 본 발명의 일실시예에 있어서, 무인 항공기의 음향 인식 시스템의 구성을 도시한 블록도이고, 도 2는 본 발명의 일실시예에 있어서, 무인 항공기의 음향 인식 방법의 동작을 도시한 흐름도이다.
- [0024] 도 1에 따르면, 음향 인식 시스템(100)은 특징 벡터 추출부(110), 모델 학습부(120), 및 음향 인식부(130)를 포함할 수 있다. 그리고, 도 2의 각 단계들(210 내지 230 단계)은 특징 벡터 추출부(110), 모델 학습부(120), 및 음향 인식부(130)에 의해 수행될 수 있다.
- [0025] 210 단계에서, 특징 벡터 추출부(110)는 잡음(noise) 및 다양한 음향 신호들이 존재하는 실질 환경에서 MFCC(Mel Frequency Cepstral Coefficients) 기법에 기초하여 무인 항공기(UAV)의 음향 신호와 관련된 특징 벡터들을 추출할 수 있다.
- [0026] 이때, 특징 벡터 추출부(110)는 2개의 MFCC 기법을 이용하여 무인 항공기의 프로펠러가 회전하면서 방출하는 음향 신호로부터 무인 항공기에 해당하는 특징 벡터들을 추출할 수 있다. 즉, 특징 벡터 추출부(110)는 미리 지정된 기준 시간 동안 켈스트럴 도메인(cepstral domain)의 포락선을 기반으로 종류 별 무인 항공기의 음향 신호와 관련된 특징 벡터들을 추출할 수 있다. 여기서, 2개의 MFCC 기법은 서로 다른 개수의 켈스트럴 계수로 구성된 제1 MFCC 및 제2 MFCC를 포함할 수 있으며, 제1 MFCC는 24개 켈스트럴 계수(MFCC)로 구성되고, 제2 MFCC는 36개 켈스트럴 계수(MFCC)로 구성될 수 있다. 예컨대, 제1 및 제2 MFCC는 아래의 표 1과 같이 구성될 수 있다.

표 1

24-MFCCs features	36-MFCCs features
12 standard MFCCs	12 standard MFCCs
12 delta MFCCs calculated by Equation (11)	12 delta MFCCs calculated by Equation (11)
	12 delta MFCCs calculated by Equation (12)

- [0027]
- [0028] 위의 표 1은 특징 벡터 추출을 위한 2개의 MFCC 기법을 나타낼 수 있다. 표 1에 따르면, 제1 MFCC 기법은 12개 표준 MFCC 및 12개 델타 MFCC로 구성되고, 제2 MFCC 기법은 12개 표준 MFCC와 24개 델타 MFCC로 구성될 수 있다. 여기서, 24개 MFCC는 후술될 수학적 식 11에 기초하여 계산된 12개 델타 MFCC와 수학적 식 12에 기초하여 계산된 12개 델타 MFCC로 구성될 수 있다.
- [0029] 220 단계에서, 모델 학습부(120)는 추출된 특징 벡터들을 기반으로 생성된 학습 데이터 세트에 기초하여 HMM(Hidden Markov Model) 모델을 학습시킬 수 있다. 여기서, 초기의 학습 데이터 세트는 음향 신호의 종류 별로 미리 지정될 수 있다.

- [0030] 예컨대, 모델 학습부(120)는 HMM 파라미터들을 초기화한 이후, 특징 벡터들을 이용하여 HMM 모델을 학습시킬 수 있다. 즉, 학습을 통해 HMM 파라미터들을 업데이트할 수 있다.
- [0031] 230 단계에서, 음향 인식부(130)는 시스템으로 입력되는 다양한 종류의 음향 신호들을 대상으로, 학습된 HMM 모델에 기초하여 무인 항공기(UAV)에 해당하는 음향 신호를 인식(또는 탐지)할 수 있다. 즉, 학습된 HMM 모델을 분류기(classifier)로 이용하여 입력된 신호들 중에서 무인 항공기(UAV)에 해당하는 음향 신호를 인식할 수 있다.
- [0032] 예컨대, 비행기 소리, 새 소리, 자동차 소리, 비 소리, 드론 소리 등과 같이 실질 환경에 존재하는 다양한 종류의 음향 신호들이 음향 인식부(130)의 음향 입력 센서를 통해 입력될 수 있다. 그러면, 음향 인식부(130)는 상기 무인 항공기와 관련하여 추출된 특징 벡터들을 기반으로 학습된 HMM 모델에 기초하여 다양한 종류의 음향 신호들 중에서 무인 항공기(UAV)에 해당하는 음향 신호를 인식할 수 있다. 이때, 무인 항공기의 종류 별로 모터에 의하여 구동되는 프로펠러의 소리가 상이할 수 있으며, 종류 별로 구분되어 특징 벡터들이 추출 및 학습이 수행될 수 있다. 이에 따라, 상기 음향 인식부(130)는 입력된 음향 신호가 단순히 무인 항공기에 해당함을 인식할 수 있을 뿐만 아니라, 인식된 무인 항공기(UAV)관련 정보(예컨대, 드론의 종류, 제조사, 크기, 드론의 소속 정보 등)를 파악할 수 있다.
- [0033] 그리고, 인식된 무인 항공기관련 정보는 시스템(100)의 디스플레이(미도시)를 통해 표시될 수도 있다. 예컨대, 입력된 음향 신호가 드론(drone)에 해당하는지 여부, 해당한다면 드론의 종류, 크기, 소속 정보(예컨대, 방송사 소속, 군 소속, 기업 소속 등) 등이 디스플레이(미도시)에 표시될 수 있다.
- [0035] 이하에서는 수학적 1 내지 수학적 12를 참고하여, 2개의 MFCC 기법을 이용하여 특징 벡터들을 추출하는 동작에 대해 보다 상세히 설명하기로 한다. 이때, 3가지 학습 데이터 세트가 음향 유형의 증가와 함께 설정될 수 있으며, 각 학습 데이터 세트에 포함된 학습 데이터들(예컨대, 종류별 음향 신호)은 각각의 단일 클러스터(cluster)에 기초하여 HMM 모델을 학습시키기 위해 미리 여러 개의 클러스터로 분류될 수 있다.
- [0036] 드론과 같은 무인 항공기의 프로펠러 소리를 인식하기 위한 프레임워크(즉, 드론 음향 인식 프레임워크)는 음향 신호로부터 특징 벡터를 추출하고, 특징 벡터들을 종류 별로 분류하는 두 단계로 구분될 수 있다. 음향 신호는 이동성, 속도 변화 및 전파 조건으로 인한 시간 영역 또는 주파수 구성 요소, 기타 요소로 인한 주파수 영역에서 자체 기능을 수행할 수 있다. 이때, 특징 벡터들을 포함하는 특징 벡터 집합을 형성하기 위해 시간-주파수 영역에서 음향 특성을 보여주기 위한 STFT(Short Time Fourier Transform)를 통해 스펙트로그램을 생성한 후 MFCC가 구성될 수 있다. STFT는 일종의 푸리에 변환으로서, STFT를 수행하는 것은 시간의 신호를 동일한 길이의 더 짧은 세그먼트로 나눈 다음, 더 짧은 세그먼트마다 별도로 푸리에 변환을 수행하는 것을 나타낼 수 있다. 이때, 연속 시간(continuous-time)의 경우, 음향 신호의 STFT는 아래의 수학적 1과 같이 표현될 수 있다.
- [0037] [수학적 1]
- $$STFT\{x(t)\}(\tau, \omega) \equiv X(\tau, \omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)\exp(-j\omega t)dt$$
- [0038]
- [0039] 수학적 1에서, $w(t)$ 는 윈도우(window) 함수, $x(t)$ 는 푸리에 변환될 신호(예컨대, 드론 등의 음향 신호)를 나타낼 수 있다. 그리고, $X(\tau, \omega)$ 는 시간과 주파수에 따른 신호의 위상과 크기를 나타내는 복소 함수인 $x(t)w(t-\tau)$ 의 푸리에 변환을 나타낼 수 있다.
- [0040] 이때, 이산 시간(discrete-time)의 경우, 푸리에 변환될 신호(즉, 음향 신호에 해당하는 데이터)가 프레임(frame)으로 분할될 수 있으며, 서로 오버랩(overlap)될 수 있다. 신호를 여러 프레임으로 분할 시, 분석하고자 하는 신호의 해상도와 총 지속 시간에 기초하여 프레임의 길이가 분석될 수 있다. 그러면, 이산 시간에서, 각 프레임에 대한 푸리에 변환은 수학적 2와 같이 표현될 수 있다.

[0041] [수학식 2]

$$STFT\{x[n]\}(m, \omega) \equiv X(m, \omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]\exp(-j\omega n)$$

[0042]

[0043] 수학식 2에서, $x[n]$ 은 변환될 신호, $w[n]$ 은 윈도우 함수를 나타낼 수 있다. 이때, 푸리에 변환 시 STFT 대신 FFT가 사용되는 경우, 상기 수학식 2의 변수 m 과 ω 는 이산적(discrete)일 수 있다.

[0044] 이때, 스펙트로그램을 산출하기 위한 STFT의 크기의 제공은 아래의 수학식 3과 같이 표현될 수 있다.

[0045] [수학식 3]

$$spectrogram\{x[n]\}(m, \omega) \equiv |X(m, \omega)|^2$$

[0046]

[0048] 도 3은 본 발명의 일실시예에 있어서, 다양한 음향 신호 별 스펙트로그램(spectrogram)을 나타낼 수 있다.

[0049] 도 3에서는 드론의 녹음된 음향 신호(즉, 드론의 프로펠러 회전 시 방출된 음향 신호를 녹음한 신호, 310) 및 인터넷 상의 데이터베이스로부터 수집된 비행기(320), 자동차(330), 및 새(340) 등의 다양한 종류의 음향 신호 각각에 해당하는 스펙트로그램을 나타낼 수 있다.

[0050] 도 3에 도시된, 비행기, 자동차, 새소리 등에 해당하는 음향 신호(320, 330, 340)는 드론(drone)의 프로펠러 소리에 해당하는 음향 신호를 녹음 및 인식하는데 주된 장애물이 될 수 있다. 도 3에 따르면, 서로 다른 종류의 음향 신호의 특징이 시간 및 주파수 영역에서 모두 나타남을 알 수 있다. 예컨대, 드론(drone)의 경우, 400Hz 내지 8KHz에서 강한 고조파 성분이 나타나는 특징을 가짐을 확인할 수 있으며, 나머지 음향 신호들(320, 330, 340)은 고조파 성분이 다양한 주파수 영역에서 광범위하게 존재함을 확인할 수 있다. 이처럼, 드론의 스펙트로그램에 기초하여, 400Hz 내지 8KHz 영역에서 고조파 성분을 갖는 특징 등을 기반으로 입력된 음향 신호의 스펙트로그램(즉, 패턴)이 상기 드론에 해당하는지 여부가 인식될 수 있다.

[0052] 도 4는 본 발명의 일실시예에 있어서, MFCC 기법에 기초하여 특징 벡터를 추출하는 동작을 도시한 흐름도이고, 도 5는 본 발명의 일실시예에 있어서, 특징 벡터 추출부의 세부 구성을 도시한 블록도이다.

[0053] 도 5에 따르면, 특징 벡터 추출부(500)는 프리 엠퍼시스(pre-emphasis) 처리부(510), 프레이밍(framing, 520), 윈도우(windowing, 530), FFT 처리부(FFT, 540), 멜 스케일 필터 뱅크(Mel-scale filter bank, 550), 대수 처리부(Logarithm operation, 560), DCT(Discrete Cosine Transform, 570), MFCC 구성부(MFCC, 580)를 포함할 수 있다. 도 4에서 각 단계들(410 내지 440 단계는 도 5의 구성 요소인 프리 엠퍼시스(pre-emphasis) 처리부(510), 프레이밍부(framing, 520), 윈도우부(windowing, 530), FFT 처리부(FFT, 540), 멜 스케일 필터 뱅크(Mel-scale filter bank, 550), 대수 처리부(Logarithm operation, 560), DCT(Discrete Cosine Transform, 570), MFCC 구성부(MFCC, 580)에 의해 수행될 수 있다.

[0054] 먼저, 특징 벡터 추출부(500)는 음향 신호 별 특징 벡터(즉, 음향 특징)를 효율적으로 추출하여 음향 신호에 대한 특징 벡터 집합을 생성하기 위해, 대수(logarithm) 영역에서 스펙트럼의 포락선을 추출하는 MFCC 기법을 이용하여 캡스트럴 영역에서 특징 벡터들을 추출할 수 있다. 위의 비특허문헌 [1] I. Patel and Y. S. Rao, "Speech recognition using Hidden Markov Model with MFCC-Subband technique," in Proc. International Conference on Recent Trends in Information, Telecommunication and Computing, pp.168-172, Mar. 2010.에서는 대수 영역에서 스펙트럼의 포락선을 추출하여 캡스트럴 영역에서의 특징 벡터를 추출하는 구성을 제시하고 있다.

[0055] 이때, 특징 벡터 추출을 위해 다양한 종류의 음향 신호가 이용될 수 있다. 예컨대, 동일한 속성의 음향 신호 및 서로 다른 속성의 음향 신호들이 이용될 수 있다. 동일한 속성의 음향 신호는 서로 다른 종류의 드론들 각각의 프로펠러 소리에 해당하는 음향 신호를 나타낼 수 있고, 서로 다른 속성의 음향 신호는 새 소리, 비행기 소리, 자동차 소리 등의 음향 신호를 나타낼 수 있다.

[0056] 먼저, 특징 벡터 추출부(500)는 음향 신호(acoustic signal)을 디지털 신호로 변환할 수 있다. 이때, 모든 이

산 시간(discrete time) 영역에서 신호의 각 레벨을 나타내는 디지털 신호가 생성될 수 있다.

[0057] 410 단계에서, 프리 엠퍼시스(pre-emphasis) 처리부(510)는 디지털 신호를 대상으로 고주파 필터링(highpass filtering)을 수행하는 프리 엠퍼시스(pre-emphasis) 처리를 수행할 수 있다. 예컨대, 프리 엠퍼시스 처리부(510)는 상기 디지털 신호를 대상으로, 기록 매체에 의해 일정 레벨로 압축된 고주파에서의 에너지량을 증가시킴으로써 전체 신호대 잡음비(SNR)를 향상시키는 프리 엠퍼시스 처리를 수행할 수 있다. 즉, 프리 엠퍼시스 처리부(510)는 1차 하이패스 필터(highpass filter)로 구성될 수 있으며, 낮은 주파수의 크기와 관련하여 상대적으로 더 높은 주파수의 크기를 증가시키기 위한 필터링을 수행할 수 있다. 상기 1차 하이패스 필터는 아래의 수학적식 4와 같이 표현될 수 있다.

[0058] [수학적식 4]

[0059]
$$y[n] = x[n] - \alpha x[n-1], \quad 0.9 \leq \alpha \leq 1.0$$

[0060] 수학적식 4에서, $x[n]$ 은 프리 엠퍼시스 처리부(510)로 입력되는 음향 신호(즉, 디지털 신호)를 나타내고, $y[n]$ 은 출력 신호를 나타낼 수 있다. 그리고, α 는 필터 계수를 나타낼 수 있다.

[0061] 420 단계에서, 특징 벡터 추출부(500)는 프리 엠퍼시스 처리된 신호를 대상으로 프레임링(framing, 520), 윈도우링(windowing, 530), FFT 처리부(FFT, 540)를 통해 스펙트로그램(spectrogram)을 생성할 수 있다.

[0062] 예컨대, 프레임링(framing, 520) 과정에서, 프레임 쉬프트(frame shift)가 수행될 수 있으며, 프레임 쉬프트는 프레임 길이의 절반에 해당할 수 있다. 이때, 프레임 쉬프트가 수행되면서, 인접한 프레임들 간에 오버랩(overlap)이 발생할 수 있다. 즉, 프레임링(520)을 통해 인접 프레임들 간의 오버랩을 유도할 수 있다. 이처럼, 프레임 쉬프트를 통해 오버랩된 프레임들(즉, 디지털화된 음향 신호는 복수개의 프레임들로 구성됨)을 대상으로, 각 프레임에 대해 윈도우링(windowing, 530)이 수행될 수 있다. 이때, 프레임링(520) 및 윈도우링(530) 과정을 통해 불연속을 피하면서 윈도우 경계에서 신호값이 0으로 감소할 수 있다. 그러면, FFT 처리부(540)는 윈도우링된 신호를 대상으로 FFT 처리를 수행하여 스펙트로그램을 생성할 수 있다.

[0063] 즉, FFT 처리부(540)는 윈도우링된 신호를 고속 푸리에 변환(Fast Fourier Transform, FFT)할 수 있으며, 고속 푸리에 변환(FFT)된 출력 신호는 아래의 수학적식 5와 같이 표현될 수 있다.

[0064] [수학적식 5]

[0065]
$$R_i^{fft}(f) = FFT(r_i(n) \times w_i(n)), \quad 1 \leq i \leq I, 1 \leq n \leq N$$

[0066] 수학적식 5에서, $r_i(n)$ 는 i번째 프레임(frame)에 해당하는 음향 데이터를 나타내고, $w_i(n)$ 는 i번째 프레임(frame)에 해당하는 윈도우 함수(window function)를 나타낼 수 있다. 그리고, I는 전체 프레임들의 개수를 나타내고, N은 프레임의 길이, $R_i^{fft}(f)$ 는 주파수 영역에서 윈도우링된 신호를 나타낼 수 있다. 예컨대, 해밍 윈도우(hamming window)가 이용될 수 있으며, N-포인트 윈도우 함수(N-point window function)는 아래의 수학적식 6과 같이 표현될 수 있다.

[0067] [수학적식 6]

[0068]
$$w(n) = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right), \quad \alpha = 0.54, \beta = 0.46, \quad 0 \leq n \leq N-1$$

[0069] 430 단계에서, FFT 변환된 신호는 멜 스케일 필터 뱅크(Mel-scale filter bank, 550)를 통과하여 출력될 수 있다. 멜 스케일 필터 뱅크(550)는 FFT 처리부(540)로부터 FFT 변환된 신호를 입력받아 멜 스케일(Mel scale)에 주파수를 맵핑(mapping)할 수 있다. 예컨대, 아래의 수학적식 7에 기초하여 멜 스케일에 주파수가 맵핑(mapping)될 수 있다.

[0070] [수학적식 7]

[0071]
$$mel(f) = 1127 \ln\left(1 + \frac{f}{700}\right)$$

[0072] 그리고, 멜 스케일 필터 뱅크의 m번째(m^{th}) 필터를 구현하기 위한 방정식은 아래의 수학식 8과 같이 표현될 수 있다.

[0073] [수학식 8]

[0074]

[0075] 수학식 8에서, $f(\cdot)$ 는 멜 스케일 주파수를 나타내고, M은 주파수 대역에서 필터의 총 개수를 나타낼 수 있다.

[0076] 각 프레임의 음향 신호(즉, 음향 시퀀스(sequence))가 멜 스케일 필터 뱅크(550)를 통과하고 난 후, 각 필터에 가중 스펙트럼 값이 더해지면, 필터당 하나의 단일 값이 얻어지고, 모든 계수는 멜 스펙트럼을 형성할 수 있다. 이때, 아래의 수학식 [9]에 기초하여 멜 스펙트럼의 대수가 계산될 수 있다.

[0077] [수학식 9]

[0078]

[0079] 440 단계에서, DCT(Discrete Cosine Transform, 570)는 멜 스케일 필터 뱅크(550)를 통과한 신호를 대상으로 이산 코사인 변환(DCT)을 수행하여 캡스트럴 계수를 계산할 수 있다. 예컨대, 아래의 수학식 10에 기초하여 이산 코사인 변환을 수행함으로써, n번째(n^{th}) 캡스트럴 계수가 계산될 수 있다.

[0080] [수학식 10]

[0081]

[0082] 수학식 10에서, N은 캡스트럴 계수의 총 수를 나타낼 수 있다.

[0083] 440 단계에서, MFCC 구성부(MFCC, 580)는 계산된 캡스트럴 계수에 기초하여 MFCC를 구성할 수 있다. 예컨대, 위의 수학식 10을 기반으로 계산된 캡스트럴 계수를 대상으로, 처음 12개의 캡스트럴 계수는 MFCC를 구성하기 위해 유지될 수 있다.

[0084] 수학식 10에서, 계산된 캡스트럴 계수에 기초하여 구성된 MFCC(즉, 특징 벡터)는 단일 프레임의 전력 스펙트럴(power spectral) 포락선 만을 나타내지만, 동적 특징들을 나타내기 위해 델타 계수가 더해질 수 있다. 예컨대, 아래의 수학식 11에 기초하여 MFCC 구성 시 이용되는 델타 계수가 계산될 수 있다.

[0085] [수학식 11]

[0086]

[0087] 즉, 처음 12개의 캡스트럴 계수들은 수학식 10에 기초하여 계산되고, 이후 12개의 델타 계수들은 수학식 11에

기초하여 계산될 수 있으며, 처음 12개의 캡스트럴 계수들과 이후 12개의 델타 계수들로 구성된 MFCC는 제1 MFCC에 해당할 수 있다. 그리고, 제2MFCC는 다른 방법으로 12개의 델타 계수들이 계산될 수 있으며, 상기 제1MFCC에 해당하는 계수들과 다른 방법으로 계산된 12개의 델타 계수들에 기초하여 제2 MFCC가 구성될 수 있다. 예컨대, 아래의 수학적 식 12에 기초하여 제1MFCC와는 다른 12개의 델타 계수들이 계산될 수 있다.

[수학적 식 12]

$$d_n = \begin{cases} c_{n+1} - c_n, & n \leq L \\ \frac{\sum_{\delta=1}^L (c_{n-\delta} + c_{n+\delta}) - 2LC_n}{2L}, & \text{others} \\ c_n - c_{n-1}, & n > N-L \end{cases}$$

수학적 식 12에서 델타(δ)는 계수의 차를 계산하기 위해 이용될 수 있다.

이처럼, 특징 벡터 추출을 위해 2개의 MFCC 기법(예컨대, 24개 MFCC 및 36개 MFCC)이 이용될 수 있으며, 위의 표 1과 같이, 24개 MFCC는 12개 표준 MFCC 및 수학적 식 11 기반의 12개 델타 MFCC를 포함하고, 36개 MFCC는 12개 표준 MFCC, 수학적 식 11 기반의 12개 델타 MFCC 및 수학적 식 12 기반의 12개 델타 MFCC를 포함할 수 있다. 이러한 최적화는 실시간 인식 프로세스 이전에 수행되므로, 특징 추출 및 학습의 최적화와 관련된 복잡도는 크게 문제가 되지 않을 수 있다.

이하에서는, 2개의 MFCC 기법에 기초하여 추출된 특징 벡터들, 예컨대, 24개 MFCC(즉, 특징 벡터), 및 36개 MFCC(즉, 특징 벡터)를 기반으로 HMM 모델을 학습시켜 드론 등의 무인 항공기(UAV)를 인식하는 동작에 대해 설명하고자 한다. 본 실시예들에서 다차원 HMM(Hidden Markov Model)이 이용될 수 있다.

HMM 모델은 파라메트릭 랜덤 프로세스(Parametric random process)로 특성화되는 변수의 정렬된 시퀀스(sequence)에 대한 통계 모델을 나타낼 수 있다. HMM에서 상태(states)는 은닉(hidden)되는 반면, 입력은 관측/관찰 가능할 수 있다. 관측 벡터의 시퀀스(sequence)는 아래의 수학적 식 13과 같이 표현될 수 있으며, HMM 모델은 아래의 수학적 식 14와 같이 표현될 수 있다.

[수학적 식 13]

$$O = \{o_t\}, \quad 1 \leq t \leq T$$

[수학적 식 14]

$$HMM = (N, M, A, B, \Pi)$$

수학적 식 14에서, N은 은닉 상태(hidden state)의 개수를 나타내고, M은 상태(state) 당 관측 수를 나타내고, A는 상태 천이(state transition) 행렬을 나타낼 수 있다. 그리고, a_{ij} 는 상태 i에서 상태 j로 천이될 확률을 나타내고, B는 상태당 방출(emission) 확률 분포를 나타낼 수 있다. $b_i(k)$ 는 HMM 모델이 상태 i에 있을 때 관측 k를 방출할 확률을 나타낼 수 있다. 상태 i에서의 원소가 π_i 인 경우, Π 는 초기 상태 분포의 확률로서, 압축 표기법에 따라 아래의 수학적 식 15와 같이 표현될 수 있다.

[수학적 식 15]

$$\lambda = (A, B, \Pi)$$

그리고, HMM 모델을 학습시키기 위해 학습 데이터 세트가 미리 지정될 수 있다. 예컨대, 음향 신호의 종류 별로 초기 학습 데이터 세트가 미리 지정될 수 있다. 여러 개의 음향 신호 각각에 해당하는 클러스터(cluster) 측면에서 초기 학습 데이터 세트가 정의될 수 있다.

[0103] 그리고, 각각의 음향 신호에 해당하는 클러스터는 HMM 모델로 표현될 수 있다. S개 클러스터로 구성된 HMM은 아래의 수학적식 16과 같이 표현될 수 있다.

[0104] [수학적식 16]

[0105]
$$\lambda_s = (A_s, B_s, \Pi_s), \quad 1 \leq s \leq S$$

[0106] 수학적식 16에서, S는 s번째(s^{th}) 클러스터에 해당하는 s번째(s^{th}) HMM 모델을 나타낼 수 있다. 그리고, 특징 벡터들을 기반으로 생성된 학습 데이터 세트를 이용하여 HMM 모델(즉, 분류기(classifier))가 학습될 수 있다. 그러면, 학습된 HMM 모델을 사용하여 테스트 데이터 세트(예컨대, 입력되는 음향 신호) 내의 음향 신호가 드론의 프로펠러 소리에 해당하는지, 자동차, 새 소리 등에 해당하는지 여부가 인식될 수 있다.

[0107] HMM 모델의 파라미터는 학습 데이터 세트에 포함된 학습 데이터에 의해 결정되며, 상기 학습 데이터는 HMM 모델로 입력되는 입력 데이터로서, 특징 벡터 추출부(110)에서 추출된 특징 벡터를 나타낼 수 있다. 학습된 HMM 모델은 가장 가능성이 높은, 즉, 원하는 신호와 가장 일치하는 가능성이 높은 음향 신호를 찾기 위해 이용될 수 있다. 예컨대, 입력된 테스트 데이터 세트에 포함된 음향 신호를 대상으로, 드론의 음향 신호일 가능성이 높은 신호를 찾기 위해 이용될 수 있다.

[0108] 학습 데이터 세트는 아래의 수학적식 17과 같이 표현될 수 있다.

[0109] [수학적식 17]

[0110]
$$Dataset_{train} = \{D_{train}^s\}, \quad 1 \leq s \leq S$$

[0111] 관측 데이터 0의 세트를 제공할 가능성이 높은 HMM 파라미터를 찾아내는 것이 바람직할 수 있으며, 상기 가능성이 높은 HMM 파라미터를 찾아내는 것이 HMM 모델 최적화를 위한 학습 시 중요 문제가 될 수 있다. 여기서, 관측 데이터 0는 각 클러스터에 해당하는 학습 데이터 세트에서 추출된 특징 벡터를 나타낼 수 있다.

[0112] 도 6은 본 발명의 일실시예에 있어서, BWA(Baum-Welch Algorithm) 기반의 HMM 모델을 학습시키는 동작을 도시한 흐름도이다.

[0113] 도 6은 BWA에 기초하여 해결되는 HMM 모델의 학습 과정을 나타낼 수 있다.

[0114] 도 6에서, 각 단계들(610 내지 630 단계)은 도 1의 모델 학습부(120)에 의해 수행될 수 있다.

[0115] 610 단계에서, 초기 학습 데이터 세트가 지정된 이후, 특징 벡터를 기반으로 학습 데이터 세트가 생성되고, HMM 모델이 초기화될 수 있다.

[0116] 620 단계에서, 모델 학습부(120)는 순방향 및 역방향(Forward/Backward) 알고리즘에 기초하여 시간 t에서의 상태 i에 대한 순방향 우도(forward likelihood) $\alpha_t(i)$, 및 역방향 우도(backward likelihood) $\beta_t(i)$ 를 계산할 수 있다. 위의 비특허문헌 [2] G. B. Singh and H. Song, "Using hidden Markov models in vehicular crash detection," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1119-1128, Mar. 2009.에서는 HMM 모델의 상태(state) 별로 우도(likelihood)를 계산하는 자세한 구성을 제시하고 있다.

[0117] 그러면, 관측된 배열 0와 HMM 모델 파라미터 i가 주어지면, 시간 t에서 상태 i에 해당할 확률은 베이즈(Bayes)의 정리에 기초하여 아래의 수학적식 18과 같이 표현될 수 있다.

[0118] [수학적식 18]

[0119]
$$\begin{aligned} \gamma_t(i) &= P(x_t = i | O, \lambda) \\ &= \frac{\alpha_t(i) \cdot \beta_t(i)}{P(O | \lambda)} \end{aligned}$$

[0120] 이때, 상태(state) i에서 상태(state) j로의 천이 우도(transition likelihood)는 아래의 수학적식 19와 같이 표현될 수 있다.

[0121] [수학식 19]

$$\begin{aligned}\zeta_t(i, j) &= P(x_{t-1} = i, x_t = j | O, \lambda) \\ &= \frac{P(x_{t-1} = i, x_t = j, O | \lambda)}{P(O | \lambda)} \\ &= \frac{\alpha_{t-1}(i) a_{ij} b_j(o_t) \beta_t(j)}{P(O | \lambda)}\end{aligned}$$

[0122]

[0123] 630 단계에서, 순방향 및 역방향 우도에 기초하여 HMM 파라미터가 업데이트될 수 있다.

[0124] 예컨대, 아래의 수학식 20 내지 22에 기초하여, HMM 모델의 파라미터들(즉, HMM 파라미터들)이 재추정될 수 있으며, 재추정된 값으로 HMM 파라미터들이 업데이트될 수 있다. 여기서, 재추정(re-estimated)은, 새로운 모델 $\hat{\lambda}$ 에 대한 학습 시퀀스의 우도(likelihood)를 증가시킬 수 있으며, 상기 새로운 모델 $\hat{\lambda}$ 은 아래의 수학식 23과 같이 표현될 수 있다.

[0125] [수학식 20]

$$\hat{\pi}_i = \gamma_1(i)$$

[0126]

[0127] [수학식 21]

$$\hat{a}_{ij} = \frac{\sum_{t=2}^T \zeta_t(i, j)}{\sum_{t=2}^T \gamma_{t-1}(i)}$$

[0128]

[0129] [수학식 21]

$$\hat{b}_j(k) = \frac{\sum_{t=1}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}$$

[0130]

[0131] [수학식 23]

$$P(O_{train} | \hat{\lambda}) \geq P(O_{train} | \lambda)$$

[0132]

[0134] 도 7은 본 발명의 일실시예에 있어서, 학습된 HMM 모델에 기초하여 원하는 음향 신호를 인식하는 동작을 도시한 블록도이다.

[0135] HMM 모델 학습 과정에서, S개 클러스터로 구성된 HMMs, λ_s 를 혼합한 분류기(classifier)가 완성될 수 있다. 그리고, 테스트 과정에서, HMM 모델을 기반으로 주어진 시퀀스(sequence)의 확률을 최대화하는 상태 시퀀스를 찾기 위해 비터비(Viterbi) 알고리즘이 이용될 수 있다. 이때, 인식 과정에서 특징 벡터 O_{test} 의 시퀀스로 표현되는 입력 음향 신호를 검색하는 것이 중요하며, 인식과정은 시퀀스가 주어지면 가장 높은 확률로 HMM을 찾는 과정을 나타내는 것으로서 아래의 수학식 24와 같이 표현될 수 있다.

[0136] [수학식 24]

$$g^* = \arg \max_{all s \in S} P(\lambda_s | O_{test})$$

[0137]

[0138] 예컨대, HMM 모델에 테스트 데이터 세트(710)에 포함된 음향 신호(즉, 시퀀스)에 해당하는 특징 벡터(720)가 입력 데이터로 이용될 수 있으며, 입력된 특징 벡터(720)와 가장 높은 확률로 일치하는 HMM을 찾음으로써, 입력된

음향 신호가 드론관련 음향 신호인지, 비행기관련 음향 신호인지 여부가 인식될 수 있다. 즉, 입력된 다양한 종류의 특징 벡터들(720)을 대상으로, 가장 높은 확률로 일치하는 HMM 모델이 속하는 클러스터로 상기 특징 벡터들 각각이 분류될 수 있으며(730), 분류를 통해 해당 특징 벡터가 어떤 소리에 해당하는지 여부가 인식될 수 있다(740).

[0140] 이하에서는 드론(drone)의 음향 인식에 대한 실험 결과 및 성능 평가 결과에 대해서 설명하고자 한다. 실험을 위해 학습 데이터 세트와 테스트 데이터 세트로 구성된 데이터베이스가 설정되며, 데이터베이스에는 각 데이터 세트에 대한 음향 신호의 클러스터들이 다수 포함될 수 있다. 그리고, 상기 두 개의 MFCC 기법을 사용하여 특징 벡터를 추출하는 동작은 각 클러스터 별로 서로 다른 음향 신호를 기반으로 수행될 수 있다. 이후, HMM 모델이 초기화되고, 학습 데이터 세트의 클러스터로부터의 특징 벡터가 초기화된 HMM 모델에 적용되어 HMM 파라미터를 최적화할 수 있다. 즉, 특징 벡터를 기반으로 HMM 파라미터를 최적화하고, 최적화된 파라미터로 HMM 파라미터를 갱신하는 과정을 통해 HMM 모델을 학습시킬 수 있다. 그리고, 입력되는 소리(즉, 실질 환경에 존재하는 다양한 종류의 음향 신호)를 테스트 데이터 세트로 설정하고, 테스트 데이터 세트를 대상으로 HMM 모델을 분류기(classifier)로 이용하여 테스트 데이터 세트에 포함된 음향 신호가 어떤 음향 신호인지 여부가 인식될 수 있다.

[0141] 이때, 드론에 해당하는 음향 신호의 인식을 위해 플랫폼(platform)이 설계될 수 있다. 학습 데이터 세트 및 테스트 데이터 세트는 5개의 클러스터(즉, 드론, 공항에서의 항공기, 자동차, 새, 및 비 소리에 해당하는 클러스터 1 내지 클러스터 5)를 포함할 수 있다. 아래의 표 2는 클러스터 관점에서의 테스트 데이터 세트에 포함된 음향 신호를 나타낼 수 있다.

표 2

	D^1_{train}	D^2_{train}	D^3_{train}	Remarks
Cluster 1	Drone 1	Drone 1, Drone 2	Drone 1, Drone 2, Drone 3	Drone 1: SYMA X5SW Drone 2: BYROBOT Drone Fighter Drone 3: WLtoys Skywalker Plane 1: Helicopter Plane 2: Boeing737 Plane 3: Boeing747
Cluster 2	Plane 1	Plane 1, Plane 2	Plane 1, Plane 2, Plane 3	
Cluster 3	Car 1, Car 2	Car 1, Car 2, Car 3	Car 1, Car 2, Car 3, Car 4, Car 5	
Cluster 4	Bird 1	Bird 1, Bird 2	Bird 1, Bird 2, Bird 3	
Cluster 5	Rain 1	Rain 1, Rain 2	Rain 1, Rain 2, Rain 3	

[0142] 표 2에서, 3개의 학습 데이터 세트는 $Dataset_{train} = \{D^s_{train}\}, 1 \leq s \leq 3$ 로 표현될 수 있으며, 3개의 학습 데이터 세트 별 차이는 클러스터 1, 클러스터 2, 클러스터 4, 클러스터 5에 학습을 위하여 1가지, 2가지 또는 3가지 종류의 음향 신호들이 포함되는지 여부에 해당할 수 있다. 여기서, 클러스터 3의 경우, 2가지, 3가지, 또는 5가지 종류의 음향 신호가 수집될 수 있으며, 자동차에 해당하는 음향 신호의 짧은 지속시간으로 인해 각 클러스터에 동일한 음향 데이터 길이가 유지될 수 있다. 그리고, 3개의 학습 데이터 세트 내의 서로 다른 클러스터에 해당하는 각 음향 신호에 대한 특징 벡터들이 HMM 모델로 입력되기 위해 추출될 수 있다.

[0144] 이때, 2개의 MFCC 기법을 이용하여 특징 벡터들이 추출될 수 있으며, 아래의 표 3은 특징 벡터 추출을 위해 이용되는 파라미터들을 나타낼 수 있다.

표 3

Parameter	Value
Acoustic data length per instance	1 s
Sample frequency f_s	44.1 kHz
Frame length T_f	10 ms
Frame shift ΔT_f	5 ms ($= 1/2 T_f$)
FFT size per frame	512
Number of filters in filter bank	24
Number of features	24 / 36

[0145]

[0146]

표 3의 파라미터들을 이용하여 특징 벡터들이 추출되면, 추출된 특징 벡터들을 이용하여 분류기(즉, HMM 모델)를 구성하기 위해, 3개의 학습 데이터 세트 각각에 해당하는 3개의 분류기(즉, HMM 모델)가 초기화될 수 있으며, $H = \{H_n\}$, $1 \leq n \leq 3$ 로 표현될 수 있다. 여기서, n은 분류 기호 인덱스(index)를 나타낼 수 있으며, 각각의 H_n 내에서 5개의 클러스터에 대해 HMM 모델이 모델링될 수 있다($\{\lambda_s\}$, $1 \leq s \leq 5$). 그러면, 각 HMM 파라미터들은 상기 특징 벡터들에 기초하여 학습되고, 각각 5개의 HMM으로 구성된 3개의 분류기를 이용하여 입력되는 음향 신호가 인식될 수 있다.

[0147]

일례로, 24개 MFCC에 기초하여 드론에 해당하는 음향 신호를 인식하는 경우, 각 학습 데이터 세트 내의 각 클러스터에 대해 100개의 샘플(sample) 데이터가 수집될 수 있다. 총 입력되는 음향 신호의 총 N_I 개가 테스트 샘플인 경우, 분류기를 통해 성공적으로 인식된 샘플 수를 계산하여 N_D 로 표현하고, 인식률이라 불리는 확률은 $P_D = N_D / N_I \times 100\%$ 로서 계산될 수 있다. 일반적인 음향 신호의 인식률은 아래의 수학적식 25와 같이 표현될 수 있다.

[0148]

[수학적식 25]

[0149]

$$P_D = \Pr(\text{incoming sound has been classified to } \lambda_s \mid \text{sound of cluster } s \text{ existence})$$

[0150]

아래의 표 4는 제1 MFCC 기법을 이용하여 인식된 음향 신호의 실험 결과를 나타낼 수 있다.

표 4

Training dataset	D ¹ _{train}	D ² _{train}	D ³ _{train}
Test dataset	Drone 2: 58% Cluster 1, 32% Cluster 2, 10% Cluster 3 Drone 3: 64% Cluster 1, 36% Cluster 2 Drone 4: 62% Cluster 1, 38% Cluster 2	Drone 3: 86% Cluster 1, 14% Cluster 2 Drone 4: 89% Cluster 1, 11% Cluster 2	Drone 4: 100% Cluster 1
	Plane 4: 60% Cluster 2, 28% Cluster 1, 12% Cluster 3	Plane 4: 78% Cluster 2, 14% Cluster 1, 8% Cluster 3	Plane 4: 86% Cluster 2, 8% Cluster 1, 6% Cluster 3
	Car 6: 78% Cluster 3, 14% Cluster 1, 8% Cluster 2	Car 6: 90% Cluster 3, 10% Cluster 1	Car 6: 96% Cluster 3, 4% Cluster 1
	Bird 4: 92% Cluster 4, 8% Cluster 2	Bird 4: 95% Cluster 4, 5% Cluster 2	Bird 4: 100% Cluster 4
	Rain 4: 90% Cluster 5, 10% Cluster 2	Rain 4: 94% Cluster 5, 6% Cluster 2	Rain 4: 100% Cluster 5

[0151]

[0152]

표 4에서는 드론 4 모델이 DJI Phantom 3 프로(Pro)이고, 비행기 4(Plane 4)가 에어버스(Airbus) A320인 경우, 24개 MFCC를 이용하여 각기 다른 학습 데이터 세트 별 인식률을 나타낼 수 있다.

[0153]

표 4에 따르면, 학습 데이터 세트 내에 포함된 음향 신호의 종류가 많을수록 인식률이 증가함을 확인할 수

있다. 학습 데이터 세트 D^1_{train} 가 클러스터 당 단일 종류의 음향 신호를 포함할 때, 드론(Drone)의 음향 신호에 대한 인식률은 64%에 불과하지만, 클러스터 1에 속하지 않는 다른 입력 음향 신호는 해당 클러스터에서 잘못 분류되지 않을 수 있다.

[0154] 그리고, 학습 데이터 세트 D^2_{train} 내에 포함된 음향 신호의 종류가 클러스터 당 2 종류로 증가하면, 최대 정확도는 89%로 증가하고, 학습 데이터 세트 D^3_{train} 내 음향 신호의 종류가 클러스터 당 3가지 유형으로 증가하면, 정확도에 대한 확률은 100%로 됨을 확인할 수 있다. 즉, 학습 데이터 세트 D^3_{train} 의 경우, 학습 데이터 세트 D^1_{train} 및 D^2_{train} 를 사용할 때 보다 정확도가 상당히 높음을 알 수 있다. 표 4의 실험 데이터를 참조하면, 학습 데이터 세트에 포함된 학습 데이터의 양이 많이 추가될수록 인식률 성능이 향상됨을 알 수 있다.

[0155] 마찬가지로, 공항의 비행기, 자동차, 새, 또는 비 등의 소리에 해당하는 음향 신호가 입력되면, 상기 음향 신호들 각각을 자신의 클러스터 모델로 분류 가능하고, 해당 음향 신호와 관련된 학습 데이터가 증가할수록 해당 음향 신호를 인식하는 정확도(즉, 인식률)도 마찬가지로 높아질 수 있다.

[0156] 아래의 표 5는 제2 MFCC 기법을 이용하여 인식된 음향 신호의 실험 결과를 나타낼 수 있다.

표 5

Training dataset	D^1_{train}	D^2_{train}	D^3_{train}
Test dataset	Drone 2: 62% Cluster 1, 30% Cluster 2, 8% Cluster 3 Drone 3: 67% Cluster 1, 33% Cluster 2 Drone 4: 64% Cluster 1, 36% Cluster 2	Drone 3: 88% Cluster 1, 12% Cluster 2 Drone 4: 94% Cluster 1, 6% Cluster 2	Drone 4: 100% Cluster 1
	Plane 4: 62% Cluster 2, 24% Cluster 1, 14% Cluster 3	Plane 4: 83% Cluster 1, 11% Cluster 2, 6% Cluster 3	Plane 4: 92% Cluster 2, 6% Cluster 1, 2% Cluster 3
	Car 6: 78% Cluster 3, 16% Cluster 1, 6% Cluster 2	Car 6: 93% Cluster 3, 7% Cluster 1	Car 6: 100% Cluster 3
	Bird 4: 90% Cluster 4, 8% Cluster 2, 2% Cluster 5	Bird 4: 95% Cluster 4, 5% Cluster 2	Bird 4: 100% Cluster 4
	Rain 4: 92% Cluster 5, 8% Cluster 2	Rain 4: 96% Cluster 5, 4% Cluster 2	Rain 4: 100% Cluster 5

[0157]

[0158] 표 5는 36개 MFCC를 이용하여 각기 다른 학습 데이터 세트 별 인식률을 나타낼 수 있다. 표 5에 따르면, 표 4와 비교하여 36개 MFCC를 사용하는 경우의 인식률이 2% 내지 6% 향상되는 것을 확인할 수 있다.

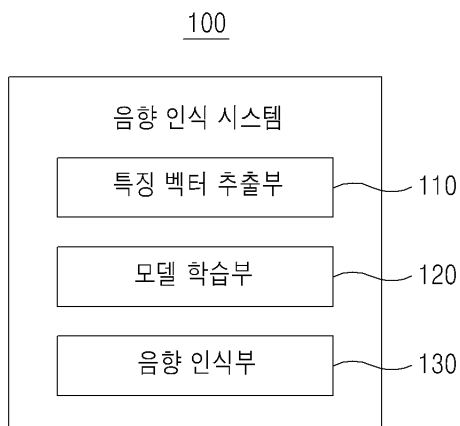
[0160] 도 8은 본 발명의 일실시예에 있어서, 제1 및 2 MFCC기법을 이용하여 추출된 특징 벡터들을 기반으로 하는 평균 인식률을 도시한 그래프이다.

[0161] 도 8을 참고하면 24개 MFCC를 사용하는 제1 MFCC 기법을 이용할 때 보다 36개 MFCC를 사용하는 제2 MFCC 기법을 이용하는 경우에 추출된 특징 벡터를 기반으로 학습된 분류기 H_1 을 학습 데이터 세트(training set 1) D^1_{train} 과 함께 사용할 때 평균 인식률이 3% 향상되는 것을 확인할 수 있다. 그리고, 분류기 H_2 를 학습 데이터 세트(training set 2) D^2_{train} 과 함께 사용할 때 성능이 2.5% 향상되는 것을 확인할 수 있다. 그리고, 2개의 MFCC 기법 모두에서, 학습된 분류기 H_3 를 학습 데이터(training set 3) D^3_{train} 와 함께 이용한 드론의 음향 신호에 대한 인식률 성능은 모두 최적화된 분류기로 인해 100% 달성함을 확인할 수 있다. 동일한 분류기를 사용하는 두 개의 MFCC 기법들 간의 성능 차이는 주로 음향 신호의 특징들 및 더 긴 유효 특징 벡터들의 동적인 정보와 같은 음향 신호의 특징들과 관련된 세부 사항들이 더 많이 이용되므로, 보다 효율적이고 높은 인식 정확도가 달성될 수 있다.

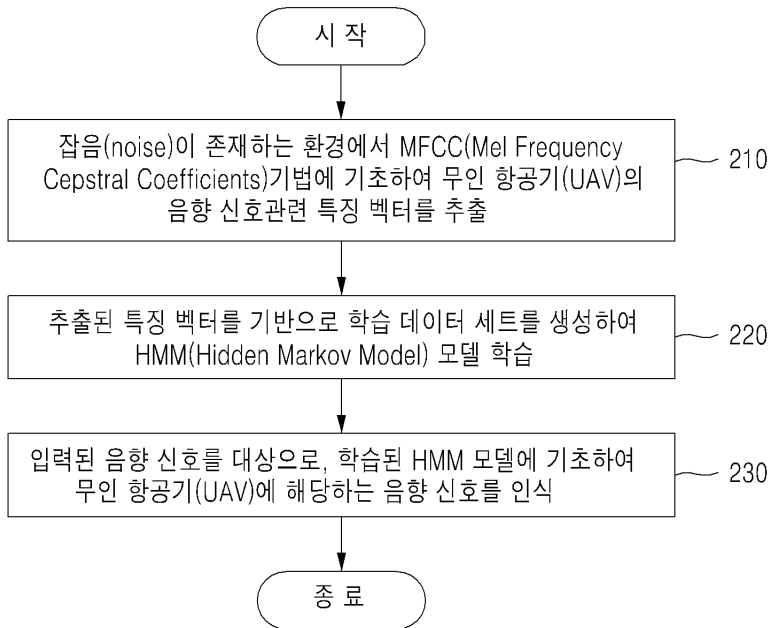
- [0163] 도 9는 본 발명의 일실시예에 있어서, SNR에 대한 인식률 성능을 도시한 그래프이다.
- [0164] 도 9에서는 잡음이 많은 실질 환경에서 제1 및 제2 MFCC 기법을 이용하여 특징 벡터를 추출하여 학습된 HMM 모델을 기반으로 인식된 음향 신호에 대한 인식률의 견고성을 테스트하기 위해 AWGN(Adaptive White Gaussian Noise) 환경이 고려되었다.
- [0165] 도 9에 따르면, 인식률의 성능을 향상시키기 위해서는 SNR을 증가시키는 것이 합리적임을 확인할 수 있다. 그리고, 36개 MFCC 및 학습 데이터 세트 3을 이용하는 경우 최상의 성능을 제공하며, 동일한 학습 데이터 세트(즉, 학습 데이터 세트 3)와 24개 MFCC를 이용하는 경우, 두 번째로 좋은 성능을 제공함을 확인할 수 있다. 36개 MFCC와 24개 MFCC의 성능 차이는 0.01% 내지 0.03%이며, 36개 MFCC의 경우 SNR이 20dB일 때, 24개 MFCC의 경우 SNR이 16dB일 최대 정확도를 달성한다. 그리고, SNR이 5dB와 같이 낮은 경우에도, 인식률이 약 80%이상임을 확인할 수 있다. 결과적으로 도 9에 따르면, 36개 MFCC를 이용하는 제2 MFCC 기법 및 상대적으로 많은 학습 데이터를 가지는 HMM 모델을 이용하여 학습된 분류기가 실질 잡음이 많은 환경에서 드론(drone) 등의 무인 항공기(UAV)의 음향 신호를 인식하는데 적합함을 알 수 있다.
- [0166] 이상에서 설명한 바와 같이, 무인 항공기(UAV)의 음향을 인식하는 시스템 및 방법은 드론(drone)의 프로펠러 소리를 이용하여 드론을 인식하기 위한 특징 벡터를 추출하고, 분류 접근법을 이용함으로써, 잡음이 많은 실질 환경에서 보다 정확하게 드론에 해당하는 음향 신호를 인식할 수 있다. 여러 클러스터에 대한 학습 데이터 세트 및 테스트 데이터 세트가 다양한 종류의 음향 신호를 위한 유형으로 설정되고, 24개 MFCC 및 36개 MFCC를 이용하여 특징 벡터를 추출하고, 추출된 특징 벡터를 이용하여 HMM 모델을 학습시키고, 학습된 HMM 모델에 기반한 분류기를 이용하여 3개의 학습 데이터 세트에 대한 인식률을 살펴보면, 많은 유형의 음향 신호가 학습 데이터 세트에 포함될수록 분류기가 점점 최적화되어 드론에 해당하는 음향 신호의 평균 인식률이 거의 100%에 다다를 수 있다. 36개 MFCC의 경우, 24개 MFCC 보다 동적인 정보를 더 많이 고려한 것이므로(즉, 동적인 정보인 12개의 델타 계수를 더 포함함), 인식률이 평균적으로 3 내지 3.5% 우수할 수 있다.
- [0167] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.
- [0168] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

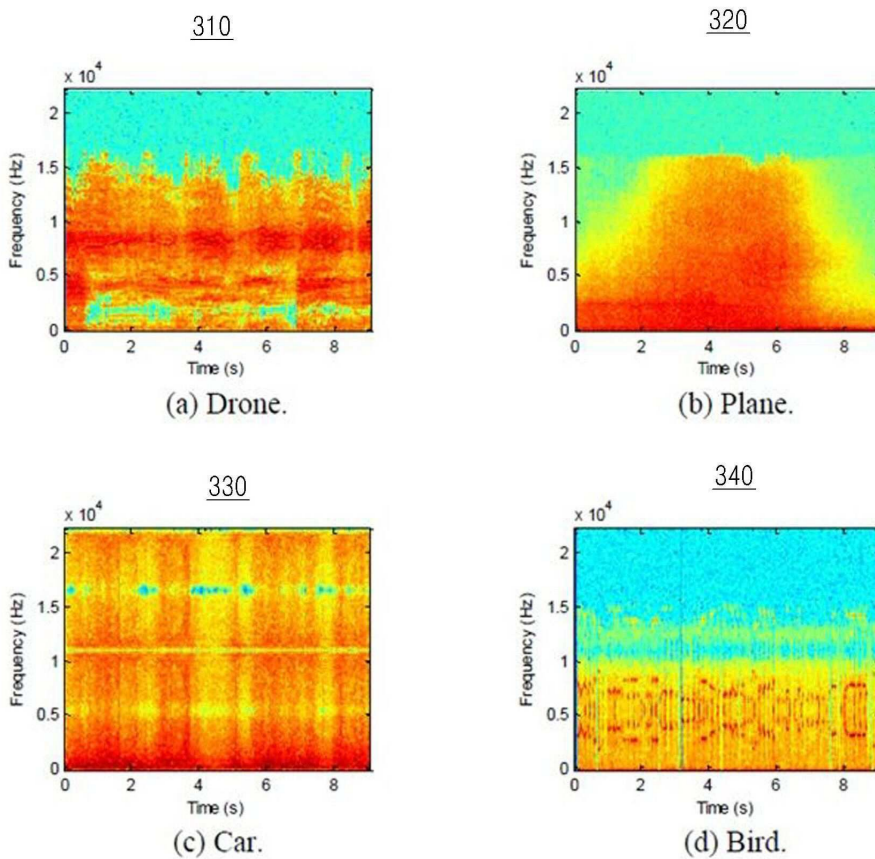
도면1



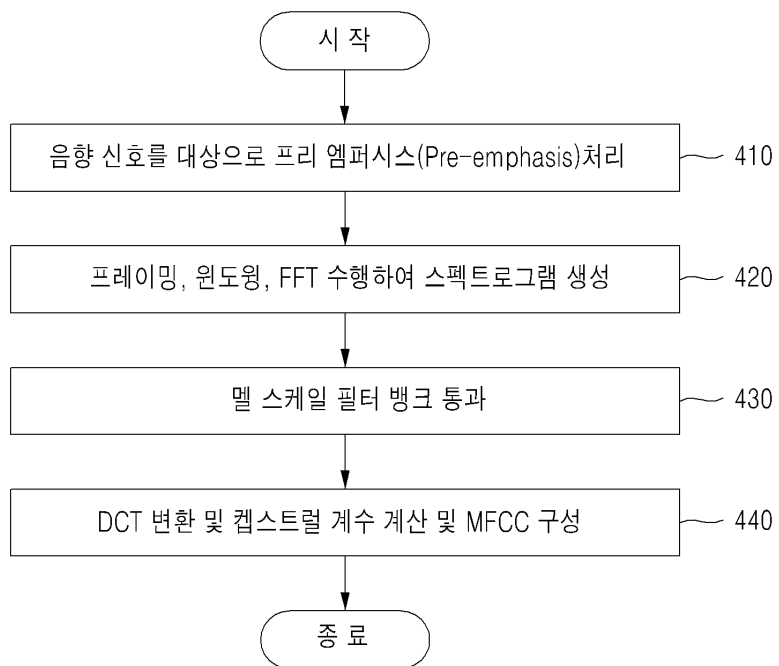
도면2



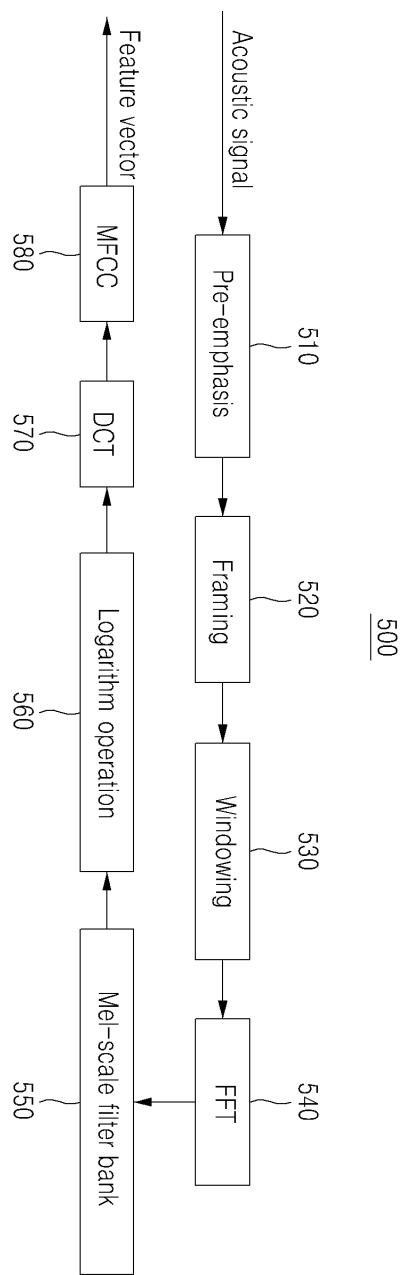
도면3



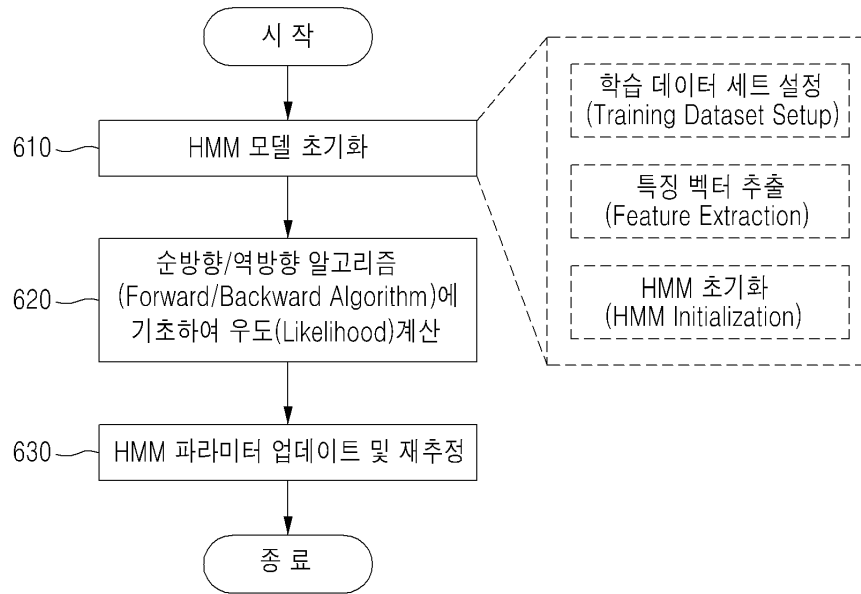
도면4



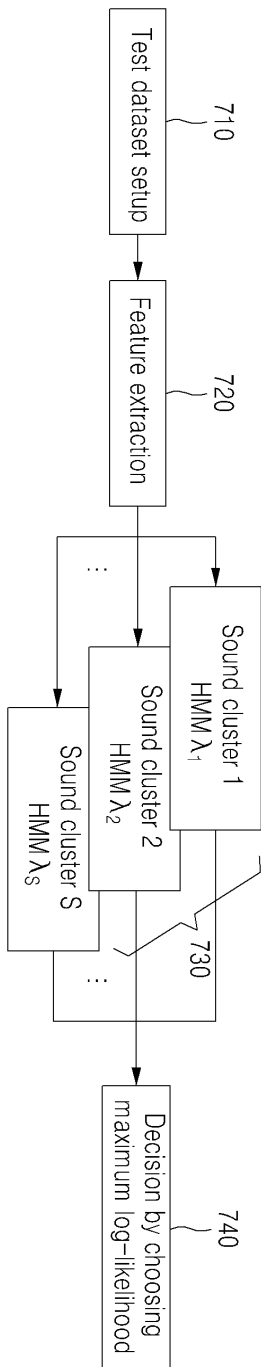
도면5



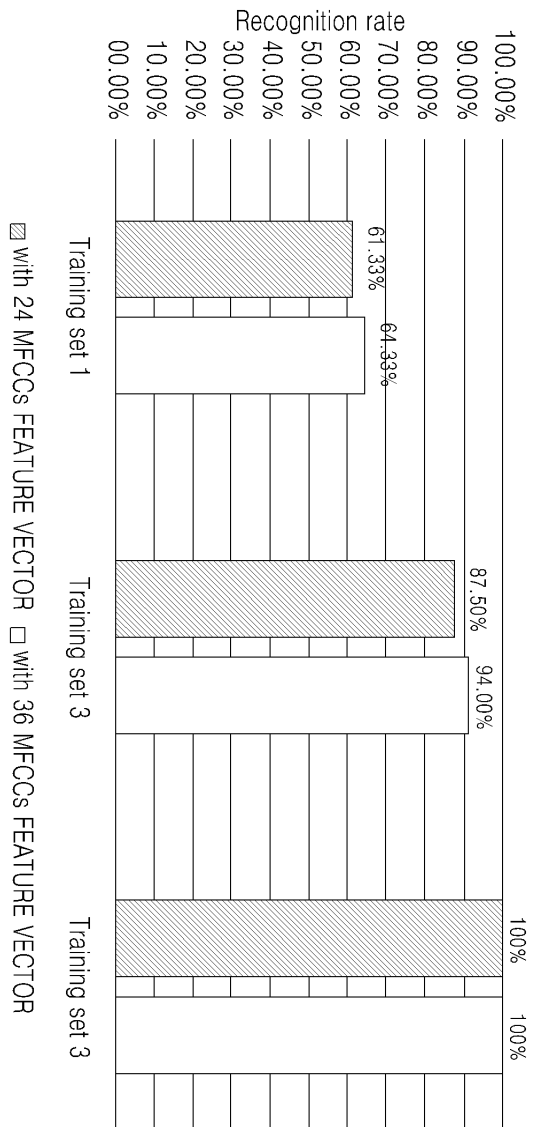
도면6



도면7



도면8



도면9

