



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2020년06월16일

(11) 등록번호 10-2123388

(24) 등록일자 2020년06월10일

(51) 국제특허분류(Int. Cl.)

G06N 99/00 (2019.01)

(52) CPC특허분류

G06N 20/00 (2019.01)

(21) 출원번호 10-2018-0051667

(22) 출원일자 2018년05월04일

심사청구일자 2018년05월04일

(65) 공개번호 10-2019-0127261

(43) 공개일자 2019년11월13일

(56) 선행기술조사문헌

Jin, Qin, and Junwei Liang. "Video description generation using audio and visual cues." Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval. ACM, 2016.\*

Simonyan, Karen, and Andrew Zisserman. Two-stream convolutional networks for action recognition in videos. Advances in neural information processing systems. 2014.\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

변혜란

서울특별시 서대문구 연세로 50, (제4공학관) 12 4동 810호 (신촌동)

김호성

서울특별시 서대문구 연세로 50, (제4공학관) 12 4동 810호 (신촌동)

(뒷면에 계속)

(74) 대리인

특허법인우인

전체 청구항 수 : 총 11 항

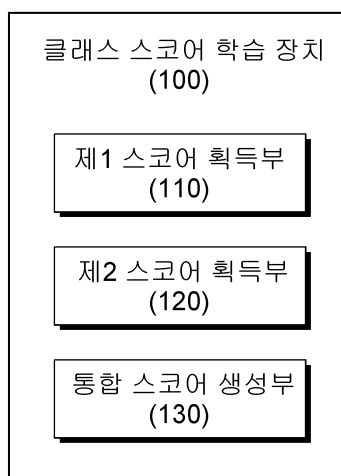
심사관 : 서광훈

(54) 발명의 명칭 행동 인식을 위한 투 스트림 네트워크의 클래스 스코어 학습 방법 및 장치

## (57) 요약

본 실시예들은 복수의 학습 모델의 클래스 스코어 간의 연관성에 상태 및 행동 쌍에 관한 기대 함수를 적용하여 복수의 학습 모델의 클래스 스코어를 통합함으로써, 학습 모델의 행동 인식률을 개선하고 학습 모델의 확장성과 호환성을 향상시킬 수 있는 클래스 스코어 학습 방법 및 장치를 제공한다.

대표도 - 도1



(72) 발명자

**어영정**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

**김태형**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

**홍중광**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

**황선희**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

**기민송**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

**홍용원**

서울특별시 서대문구 연세로 50, (제4공학관) 124  
동 810호 (신촌동)

이 발명을 지원한 국가연구개발사업

과제고유번호 2017-11-1296

부처명 과학기술정보통신부

연구관리전문기관 한국연구재단

연구사업명 차세대정보컴퓨팅기술개발사업

연구과제명 딥러닝 기반 의미론적 상황 이해 원천기술 연구

기 여 율 1/1

주관기관 연세대학교 산학협력단

연구기간 2017.09.01 ~ 2020.12.31

## 명세서

### 청구범위

#### 청구항 1

컴퓨팅 디바이스에 의한 클래스 스코어 학습 방법에 있어서,

미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 단계;

상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 단계; 및

상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 단계를 포함하며,

상기 통합 클래스 스코어를 생성하는 단계는,

상태(State) 및 행동(Action) 쌍에 관한 기대 함수를 정의하고, 상기 상태를 상기 클래스로 설정하고, 상기 행동을 상기 제1 가중치와 상기 제2 가중치로 설정하며,

상기 상태 중 현재 상태에서 상기 행동에 대한 보상을 인식의 정확도로 설정하고, 상기 행동 중 임의의 행동을 선택하는 탐색(Explore) 동작과 상기 현재 상태에 대해 기대 함수가 기 설정된 값보다 큰 값을 갖는 행동을 선택하는 획득(Exploit) 동작을 반복하여 상기 기대 함수를 갱신하는 것을 특징으로 하는 클래스 스코어 학습 방법.

#### 청구항 2

제1항에 있어서,

상기 미디어는 이미지, 비디오, 오디오, 또는 이들의 조합을 포함하는 멀티미디어인 것을 특징으로 하는 클래스 스코어 학습 방법.

#### 청구항 3

제1항에 있어서,

상기 제1 학습 모델은 컨볼루션 네트워크로 구현되어 외형 특징을 기반으로 상기 클래스를 분류하여 상기 제1 클래스 스코어를 생성하고,

상기 제2 학습 모델은 컨볼루션 네트워크로 구현되어 움직임 특징을 기반으로 상기 클래스를 분류하여 상기 제2 클래스 스코어를 생성하는 것을 특징으로 하는 클래스 스코어 학습 방법.

#### 청구항 4

제1항에 있어서,

상기 미디어에 관한 제3 학습 모델의 출력 레이어로부터 상기 클래스에 관한 제3 클래스 스코어를 획득하는 단계를 추가로 포함하며,

상기 통합 클래스 스코어를 생성하는 단계는 상기 제1 클래스 스코어에 관한 제1 가중치, 상기 제2 클래스 스코어에 관한 제2 가중치, 및 상기 제3 클래스 스코어에 관한 제3 가중치를 학습하여 상기 통합 클래스 스코어를 생성하는 것을 특징으로 하는 클래스 스코어 학습 방법.

#### 청구항 5

제4항에 있어서,

상기 제3 학습 모델은 소리 특징을 기반으로 상기 클래스를 분류하여 상기 제3 클래스 스코어를 생성하는 것을 특징으로 하는 클래스 스코어 학습 방법.

## 청구항 6

삭제

## 청구항 7

삭제

## 청구항 8

제1항에 있어서,

상기 통합 클래스 스코어를 기준으로 상기 클래스를 분류하여 미디어 내에서 표현된 행동을 인식하는 단계를 추가로 포함하는 클래스 스코어 학습 방법.

## 청구항 9

미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 제1 스코어 획득부;

상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 제2 스코어 획득부; 및

상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 통합 스코어 생성부를 포함하며,

상기 통합 스코어 생성부는,

상태(State) 및 행동(Action) 쌍에 관한 기대 함수를 정의하고, 상기 상태를 상기 클래스로 설정하고, 상기 행동을 상기 제1 가중치와 상기 제2 가중치로 설정하며,

상기 상태 중 현재 상태에서 상기 행동에 대한 보상을 인식의 정확도로 설정하고, 상기 행동 중 임의의 행동을 선택하는 탐색(Explore) 동작과 상기 현재 상태에 대해 기대 함수가 기 설정된 값보다 큰 값을 갖는 행동을 선택하는 획득(Exploit) 동작을 반복하여 상기 기대 함수를 갱신하는 것을 특징으로 하는 클래스 스코어 학습 장치.

## 청구항 10

제9항에 있어서,

상기 제1 학습 모델은 컨볼루션 네트워크로 구현되어 외형 특징을 기반으로 상기 클래스를 분류하여 상기 제1 클래스 스코어를 생성하고,

상기 제2 학습 모델은 컨볼루션 네트워크로 구현되어 움직임 특징을 기반으로 상기 클래스를 분류하여 상기 제2 클래스 스코어를 생성하는 것을 특징으로 하는 클래스 스코어 학습 장치.

## 청구항 11

제9항에 있어서,

상기 미디어에 관한 제3 학습 모델의 출력 레이어로부터 상기 클래스에 관한 제3 클래스 스코어를 획득하는 제3 스코어 획득부를 추가로 포함하며,

상기 통합 스코어 생성부는 상기 제1 클래스 스코어에 관한 제1 가중치, 상기 제2 클래스 스코어에 관한 제2 가중치, 및 상기 제3 클래스 스코어에 관한 제3 가중치를 학습하여 상기 통합 클래스 스코어를 생성하는 것을 특징으로 하는 클래스 스코어 학습 장치.

## 청구항 12

삭제

## 청구항 13

삭제

#### 청구항 14

제9항에 있어서,

상기 통합 클래스 스코어를 기준으로 상기 클래스를 분류하여 미디어 내에서 표현된 행동을 인식하는 행동 인식 부를 추가로 포함하는 클래스 스코어 학습 장치.

#### 청구항 15

프로세서에 의해 실행 가능한 컴퓨터 프로그램 명령어들을 포함하고 컴퓨터 판독 가능한 기록 매체에 기록된, 클래스 스코어 학습을 수행하는 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램 명령어들이 컴퓨팅 디바이스의 적어도 하나의 프로세서에 의해 실행되는 경우에,

미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 단계;

상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 단계; 및

상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 단계를 포함한 동작들을 수행하며,

상기 통합 클래스 스코어를 생성하는 단계는,

상태(State) 및 행동(Action) 쌍에 관한 기대 함수를 정의하고, 상기 상태를 상기 클래스로 설정하고, 상기 행동을 상기 제1 가중치와 상기 제2 가중치로 설정하며,

상기 상태 중 현재 상태에서 상기 행동에 대한 보상을 인식의 정확도로 설정하고, 상기 행동 중 임의의 행동을 선택하는 탐색(Explore) 동작과 상기 현재 상태에 대해 기대 함수가 기 설정된 값보다 큰 값을 갖는 행동을 선택하는 획득(Exploit) 동작을 반복하여 상기 기대 함수를 갱신하는 것을 특징으로 하는 컴퓨터 프로그램.

#### 발명의 설명

#### 기술 분야

[0001] 본 발명이 속하는 기술 분야는 디바이스에서 복수의 스트림에 관한 네트워크의 클래스 스코어를 통합하여 학습하는 방법 및 장치에 관한 것이다.

#### 배경 기술

[0002] 이 부분에 기술된 내용은 단순히 본 실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0003] 기계 학습은 학습 방식에 따라 지도 학습(Supervised Learning), 준지도 학습(Semi-supervised Learning), 비지도 학습(Unsupervised Learning, 자율 학습), 강화 학습(Reinforcement Learning)으로 분류된다.

[0004] 지도 학습은 미리 구축된 학습용 데이터(Training Data)를 활용하여 모델을 학습하며, 준지도 학습은 학습용 데이터와 정리되지 않은 데이터를 모두 훈련에 사용하는 방법이다. 비지도 학습은 별도의 학습용 데이터를 구축하는 것이 아니라 데이터 자체를 분석하거나 군집(Clustering)하면서 학습한다. 강화 학습은 학습 수행 결과에 대해 적절한 보상을 주면서 피드백을 통해 학습한다.

[0005] 외형(Appearance)과 움직임(Motion)은 비디오로부터 추출할 수 있는 사람의 행동을 특징짓는 요소이며, 공간 스트림(Spatial Stream)과 시간 스트림(Temporal Stream)으로부터 각각 추출이 가능하다. 공간 네트워크(Spatial Network)는 비디오의 각 프레임으로부터 행동과 물체들이 외양 정보를 추출하고 시간 네트워크(Temporal Network)는 연속하는 프레임 사이의 광학 흐름 영역(Optical Flow Field)로부터 복잡한 행동의 움직임 특징을 추출한다.

#### 발명의 내용

#### 해결하려는 과제

[0006] 본 발명의 실시예들은 복수의 학습 경로를 통하여 학습한 네트워크들의 클래스 스코어를 통합함으로써, 이미 학습을 마친 복수의 학습 모델의 클래스 스코어 간의 연관성을 개선하는 데 주된 목적이 있다.

[0007] 본 발명의 명시되지 않은 또 다른 목적들은 하기의 상세한 설명 및 그 효과로부터 용이하게 추론할 수 있는 범위 내에서 추가적으로 고려될 수 있다.

### 과제의 해결 수단

[0008] 본 실시예의 일 측면에 의하면, 컴퓨팅 디바이스에 의한 클래스 스코어 학습 방법에 있어서, 미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 단계, 상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 단계, 및 상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 단계를 포함하는 클래스 스코어 학습 방법을 제공한다.

[0009] 본 실시예의 다른 측면에 의하면, 미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 제1 스코어 획득부, 상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 제2 스코어 획득부, 및 상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 통합 스코어 생성부를 포함하는 클래스 스코어 학습 장치를 제공한다.

[0010] 본 실시예의 또 다른 측면에 의하면, 프로세서에 의해 실행 가능한 컴퓨터 프로그램 명령어들을 포함하는 비일시적(Non-Transitory) 컴퓨터 판독 가능한 매체에 기록되어 클래스 스코어 학습을 위한 컴퓨터 프로그램으로서, 상기 컴퓨터 프로그램 명령어들이 컴퓨팅 디바이스의 적어도 하나의 프로세서에 의해 실행되는 경우에, 미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득하는 단계, 상기 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득하는 단계, 및 상기 제1 클래스 스코어에 관한 제1 가중치와 상기 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성하는 단계를 포함한 동작들을 수행하는 컴퓨터 프로그램을 제공한다.

### 발명의 효과

[0011] 이상에서 설명한 바와 같이 본 발명의 실시예들에 의하면, 복수의 학습 모델의 클래스 스코어 간의 연관성에 상태 및 행동 쌍에 관한 기대 함수를 적용하여 복수의 학습 모델의 클래스 스코어를 통합함으로써, 학습 모델의 행동 인식률을 개선하고 학습 모델의 확장성과 호환성을 향상시킬 수 있는 효과가 있다.

[0012] 여기에서 명시적으로 언급되지 않은 효과라 하더라도, 본 발명의 기술적 특징에 의해 기대되는 이하의 명세서에서 기재된 효과 및 그 잠정적인 효과는 본 발명의 명세서에 기재된 것과 같이 취급된다.

### 도면의 간단한 설명

[0013] 도 1 및 도 2는 본 발명의 실시예들에 따른 클래스 스코어 학습 장치를 예시한 블록도이다.

도 3 및 도 4는 본 발명의 실시예들에 따른 클래스 스코어 학습 장치의 학습 모델 및 클래스 스코어를 예시한 도면이다.

도 5는 본 발명의 일 실시예에 따른 클래스 스코어 학습 장치의 Q-value와 인식 정확도의 관계를 나타낸 그래프이다.

도 6은 본 발명의 다른 실시예에 따른 클래스 스코어 학습 방법을 예시한 흐름도이다.

도 7은 본 발명의 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다.

### 발명을 실시하기 위한 구체적인 내용

[0014] 이하, 본 발명을 설명함에 있어서 관련된 공지기능에 대하여 이 분야의 기술자에게 자명한 사항으로서 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하고, 본 발명의 일부 실시예들을 예시적인 도면을 통해 상세하게 설명한다.

[0015] 도 1 및 도 2는 클래스 스코어 학습 장치를 예시한 블록도이다.

[0016] 도 1에 도시한 바와 같이, 클래스 스코어 학습 장치(100)는 제1 스코어 획득부(110), 제2 스코어 획득부(120), 및 통합 스코어 생성부(130)를 포함한다. 클래스 스코어 학습 장치(100)는 도 1에서 예시적으로 도시한 다양한 구성요소들 중에서 일부 구성요소를 생략하거나 다른 구성요소를 추가로 포함할 수 있다. 예컨대, 클래스 스코

어 학습 장치(200)는 제3 스코어 획득부(230), 행동 인식부(250), 또는 이들의 조합을 추가로 포함할 수 있다.

- [0017] 클래스 스코어 학습 장치(100, 200)는 복수의 학습 경로를 통하여 학습한 네트워크들의 클래스 스코어에 가중치를 적용하여 클래스 스코어를 통합한다. 클래스 스코어 학습 장치(100, 200)는 클래스에 따른 가중치를 학습하여 통합된 클래스 스코어를 갱신한다.
- [0018] 제1 스코어 획득부(110, 210)는 미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득한다. 제1 학습 모델은 컨볼루션 네트워크로 구현되어 외형 특징을 기반으로 클래스를 분류하여 제1 클래스 스코어를 생성할 수 있다.
- [0019] 제2 스코어 획득부(120, 220)는 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득한다. 제2 학습 모델은 컨볼루션 네트워크로 구현되어 움직임 특징을 기반으로 클래스를 분류하여 제2 클래스 스코어를 생성할 수 있다.
- [0020] 제3 스코어 획득부(230)는 미디어에 관한 제3 학습 모델의 출력 레이어로부터 클래스에 관한 제3 클래스 스코어를 획득한다. 제3 학습 모델은 소리 특징을 기반으로 클래스를 분류하여 제3 클래스 스코어를 생성할 수 있다.
- [0021] 통합 스코어 생성부(130)는 제1 클래스 스코어에 관한 제1 가중치 및 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성한다. 통합 스코어 생성부(240)는 제1 클래스 스코어에 관한 제1 가중치, 제2 클래스 스코어에 관한 제2 가중치, 및 제3 클래스 스코어에 관한 제3 가중치를 학습하여 통합 클래스 스코어를 생성할 수 있다.
- [0022] 통합 스코어 생성부(130)는 상태(State) 및 행동(Action) 쌍에 관한 기대 함수를 정의하고, 상태를 클래스로 설정하고, 행동을 제1 가중치와 제2 가중치로 설정한다. 통합 스코어 생성부(240)는 행동을 제1 가중치, 제2 가중치, 및 제3 가중치로 설정할 수 있다.
- [0023] 통합 스코어 생성부(130, 240)는 상태 중 현재 상태에서 행동에 대한 보상을 인식의 정확도로 설정하고, 행동 중 임의의 행동을 선택하는 탐색(Explore) 동작과 현재 상태에 대해 기대 함수가 기 설정된 값보다 큰 값을 갖는 행동을 선택하는 획득(Exploit) 동작을 반복하여 기대 함수를 갱신한다.
- [0024] 행동 인식부(250)는 통합 클래스 스코어를 기준으로 클래스를 분류하여 미디어 내에서 표현된 행동을 인식할 수 있다.
- [0025] 이하에서는 도 3 내지 도 5를 참조하여, 클래스 스코어 학습 장치가 학습 모델의 클래스 스코어를 통합하는 동작을 설명하기로 한다.
- [0026] 미디어는 이미지, 비디오, 오디오, 또는 이들의 조합을 포함하는 멀티미디어일 수 있다.
- [0027] 제1 학습 모델 내지 제3 학습 모델은 하나 이상의 레이어에서 컨볼루션 연산자를 통해 특징을 추출하여 특징 맵을 생성하고, 하나 이상의 레이어의 노드들은 네트워크로 연결되며, 추출한 특징을 다른 레이어에 전달하고, 서브샘플링을 통해 상기 추출한 특징을 통합하여 공간적 차원을 축소시키는 과정을 수행한다.
- [0028] 클래스 스코어 학습 장치는 제1 학습 모델 내지 제3 학습 모델의 파라미터를 학습한다. 레이어는 파라미터를 포함할 수 있고, 레이어의 파라미터는 학습가능한 필터 집합을 포함한다. 파라미터는 노드 간의 가중치 및/또는 바이어스를 포함한다.
- [0029] 제1 학습 모델은 프레임의 시각적 정보를 기반으로 외형 특징을 추출한다.
- [0030] 제2 학습 모델은 광학 흐름(Optical Flow) 방식을 이용하여 연속하는 프레임 사이의 모션 벡터를 추출할 수 있다. 모션 벡터는 방향과 크기를 갖는다. 제2 학습 모델은 프레임 내부의 영역마다 블록들을 설정한다. 블록 내부에 있는 픽셀들에서의 모션 벡터를 추출한다. 제2 학습 모델은 밀집 광학 흐름(Dense Optical Flow) 방식을 이용하여, 픽셀들에서의 모션 벡터를 추출할 수 있다. 제2 학습 모델은 Farneback Optical Flow를 이용할 수 있다. Dense Optical Flow는 특징 점이 아닌 픽셀마다 모션을 측정한다.
- [0031] 제3 학습 모델은 시간에 따른 소리 정보를 이미지 형태로 변환하고 소리 특징을 추출한다.
- [0032] 아키텍처는 5개의 컨볼루션(Convolutional) 레이어와 3개의 풀리 커넥티드(Fully Connected) 레이어를 갖는 총 8개 레이어로 구성될 수 있다. 첫 번째 컨볼루션 레이어는 11x11x96 (필터 가로 크기x필터 세로 크기x필터 채널 수) 크기의 필터와 스트라이드(Stride)를 4로 설정할 수 있다. 두 번째 컨볼루션 레이어는 5x5x256 크기의 필터를 갖고 맥스 풀링(Max Pooling)을 수행하며, 세 번째 컨볼루션 레이어는 3x3x384 크기의 필터를 갖고 맥스 풀

링을 수행하며, 네 번째 컨볼루션 레이어는 3x3x384 크기의 필터를 갖고, 다섯 번째 컨볼루션 레이어는 3x3x256 크기의 필터를 갖고 맥스 풀링을 수행한다.

[0033] 첫 번째 풀리 커넥티드 레이어는 4,096차원, 두 번째 풀리 커넥티드 레이어는 4,096차원, 세 번째 풀리 커넥티드 레이어는 7차원 또는 15차원으로 구성될 수 있다. 마지막 세 번째 풀리 커넥티드 레이어 이후에는 소프트맥스(Softmax)를 수행하여 예측된 결과와 진실 데이터(Ground Truth) 사이의 오류를 활용하여 SGD(Stochastic Gradient Descent with Momentum) 기반의 역전파(Back Propagation) 방식으로 전체 네트워크의 파라미터들을 학습할 수 있다.

[0034] 공간 네트워크(Spatial Network)의 입력은 224x224x3 (이미지 가로 크기x이미지 세로 크기x채널 수)의 RGB 이미지이며, 시간 네트워크(Temporal Network)의 입력은 224x224x3의 광류(optical flow)로 구성된 이미지이다.

[0035] 공간 네트워크의 미니 배치 사이즈는 64, SGD의 모멘텀(Momentum)은 0.9, SGD의 가중치 감퇴(Weight Decay)는 0.0005, 학습률(Learning Rate)의 감소 스텝 사이즈(Step Size)는 20000, 초기 학습률은 0.001, 매 스텝 사이즈마다 1/10씩 학습률을 감소시키며 50,000회(Iteration)까지 반복하여 학습을 진행한다.

[0036] 시간 네트워크의 미니 배치 사이즈는 64, SGD의 모멘텀(Momentum)은 0.9, SGD의 가중치 감퇴(Weight Decay)는 0.0005, 초기 학습률은 0.001, 50000회 반복할 경우 학습률을 0.0001로 감소시키며, 100,000회 반복할 경우 학습률을 0.00001로 감소시키며, 125,000회까지 반복하여 학습을 진행한다. 시간 네트워크에서는 대용량 데이터셋에서 프리 트레이닝(Pre-Train)된 네트워크가 없기 때문에, 학습 반복 횟수를 2.5배 가량 늘려서 학습할 수 있다.

[0037] 클래스 스코어 학습 장치는 각 행동 클래스마다 외형 의존도와 움직임 의존도가 다르다고 설정한다. 각각의 의존도에 해당하는 각 클래스마다 다른 공간 스트림(Spatial Stream)과 시간 스트림(Temporal Stream)의 가중치를 학습하여 인식 성능을 향상시킨다.

[0038] 데이터셋 중 N개의 샘플( $x_i$ 는 d 차원의 특징,  $y_i \in \{1, \dots, C\}$ 는  $x_i$ 의 Ground Truth 레이블, C는 클래스 개수)을 포함하는 트레이닝셋을  $X = (x_i, y_i) | i = \{1, \dots, N\}$ 라고 할 수 있다.  $w_k \in [0, 1]$ 는 k번째 클래스의 공간 스트림의 가중치로 정의하고 모든 클래스의 공간 스트림의 가중치  $W = [w_1, \dots, w_k, \dots, w_C]$ 라고 표현하면, 스트림 단위 퓨전(Stream-Wise Fusion)의 경우  $w_1 \neq w_2 \neq \dots \neq w_C$ 이 된다.

[0039] 클래스 스코어 학습 장치는 Q-learning을 이용한 강화 학습(Reinforcement Learning) 방식으로 학습하여 각 클래스의 가중치를 학습할 수 있다. 처음 상태(State) s, 행동(Action) a에 대해서  $Q(s, a) \leftarrow 0$ 되도록 초기화한다. 그 뒤 현재 상태 s를 관측(Observation)하고, 다음 과정을 지속적으로 반복한다. 탐색(Explore) 및 획득(Exploit) 정책(Policy)을 이용하여 행동을 선택하고, 그 때 즉시 발생한 보상(Reward) r을 받은 뒤에 새로운 상태 s'를 관측하고,  $Q(s, a)$ 을 수학적 식 1과 같이 업데이트한다. 마지막으로 현재 상태 s를 새로운 상태 s'로 업데이트한다.

## 수학적 식 1

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a')] \quad [0040]$$

[0041] 수학적 식 1에서  $\gamma$ 는 할인 요소(Discounted Factor)이며, 추후에 발생하는 미래 보상에 대한 값어치를 현재 보상의 값어치에 비해 낮은 가중치를 두기 위한 파라미터이다. 차감 요소는 0.9로 설정할 수 있다. 차감된 미래 보상(Discounted Future Reward) 방식을 적용한 이유는 일반적인 미래 보상(Future Reward)만으로도 결과에 도달할 수 있지만 수렴(Convergence)에 필요한 시간을 단축시킬 수 있는 효과가 있다.

[0042] 일반적인 미래 보상(Future Reward)은 수학적 식 2와 같이 표현된다.

## 수학식 2

$$R = r_1 + r_2 + r_3 + \dots + r_n$$

수학식 2를 시간 t에 대해 확장하면 수학식 3과 같이 표현된다.

## 수학식 3

$$R_t = r_t + r_{t+1} + r_{t+2} + \dots + r_n$$

수학식 3에서 n은 수렴할 때까지 반복되는 무한한 수라고 가정한다.

수학식 3에 대하여 차감된 미래 보상 방식을 적용하면 수학식 4와 같이 표현된다.

## 수학식 4

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots + \gamma^{n-t} r_n$$

수학식 4에 벨만방정식(Bellman Equation)을 적용하면 수학식 5 및 수학식 6과 같이 표현된다.

## 수학식 5

$$R_t = r_t + \gamma (r_{t+1} + \gamma (r_{t+2} + \dots))$$

## 수학식 6

$$R_t = r_t + \gamma R_{t+1}$$

차감된 미래 보상 방식을 적용하여 에이전트(Agent)가 보상을 최대화하는 정책을 학습함으로써 최종값에 수렴하는  $Q(s,a)$ 를 찾고, 이를 바탕으로 학습된 각 상태에 대한 행동을 찾는다.

수학식 1에서  $\alpha$ 는 학습률(Learning Rate)이며, 확률론적인(Stochastic) 환경 혹은 비결정적(Non-Deterministic) 환경에서 발생하는 무작위성(Randomness) 문제를 해결하기 위해 도입하였고, 학습률을 0.1로 설정할 수 있다. 지금까지 학습해 온 과거의 Q에 0.9의 가중치를 두고 현재 획득한 보상에 대해 0.1의 가중치를 더함으로써 증분식 학습(Incremental Learning)을 한다. 이는 학습에 신중을 기한다는 의미를 갖는다.

클래스 스코어 학습 장치는 Q-learning 알고리즘을 통하여 학습된 모델과 기 개발한 투 스트림 컨볼루션 네트워크(Two-Stream Convolutional Neural Network)를 혼합한다. 클래스 기반으로 가중치를 학습하기 위해 Q-learning 알고리즘을 적용하며, 검증 정확도(validation Accuracy)를 Q 값으로 설정하고 각 클래스 기반 가중치를 설정하는 것을 행동(Action)으로 정의하여 학습한다. 학습은 450,000 반복하여 진행하였고, 탐색(Explore) 동작의 임실론 값은 0.1로 설정할 수 있다. 각각의 클래스 스코어에 대응하는 가중치의 합은 일정한 값으로 설정될 수 있으며, 예컨대, 제1 가중치와 제2 가중치의 합이 1로 설정되거나 제1 가중치, 제2 가중치, 및 제3 가중치의 합이 1로 설정될 수 있다.

클래스 스코어 학습 장치의 성능 평가 방법은 인식 정확도(Accuracy)를 사용할 수 있다. 초기 가중치는 모두 동일하게 0.7로 시작할 수 있고, 상태는 각 클래스의 개수인 15개로 설정하고, 행동은  $\{-0.001, 0.001\}$ 로 정의하

여 탐색 과정인 경우 임의로 행동을 선택하여 수행하고, 획득 과정인 경우 현재 상태에 대해 Q-value가 큰 행동을 수행하도록 한다. 현재 상태의 행동에 대한 보상은 인식 정확도로 설정할 수 있다.

- [0056] 도 5는 클래스 스코어 학습 장치가 450,000회 반복 수행한 전체 실험에서 Q-value가 수렴한 10,000회 반복 수행한 모델에 대한 Q-value와 정확도 결과 그래프를 보여준다. 학습이 진행됨에 따라 Q-value와 정확도가 비슷한 추이를 갖고 증가하는 것을 확인할 수 있다. 비교적 짧은 횟수로 반복 학습하였음에도 3.6%의 뛰어난 성능 향상을 보이며 수렴함을 파악할 수 있다.
- [0057] 클래스 스코어 학습 장치에 포함된 구성요소들이 도 1 및 도 2에서는 분리되어 도시되어 있으나, 복수의 구성요소들은 상호 결합되어 적어도 하나의 모듈로 구현될 수 있다. 구성요소들은 장치 내부의 소프트웨어적인 모듈 또는 하드웨어적인 모듈을 연결하는 통신 경로에 연결되어 상호 간에 유기적으로 동작한다. 이러한 구성요소들은 하나 이상의 통신 버스 또는 신호선을 이용하여 통신한다.
- [0058] 클래스 스코어 학습 장치는 하드웨어, 펌웨어, 소프트웨어 또는 이들의 조합에 의해 로직회로 내에서 구현될 수 있고, 범용 또는 특정 목적 컴퓨터를 이용하여 구현될 수도 있다. 장치는 고정배선형(Hardwired) 기기, 필드 프로그램 가능한 게이트 어레이(Field Programmable Gate Array, FPGA), 주문형 반도체(Application Specific Integrated Circuit, ASIC) 등을 이용하여 구현될 수 있다. 또한, 장치는 하나 이상의 프로세서 및 컨트롤러를 포함한 시스템온칩(System on Chip, SoC)으로 구현될 수 있다.
- [0059] 클래스 스코어 학습 장치는 하드웨어적 요소가 마련된 컴퓨팅 디바이스에 소프트웨어, 하드웨어, 또는 이들의 조합하는 형태로 탑재될 수 있다. 컴퓨팅 디바이스는 각종 기기 또는 유무선 통신망과 통신을 수행하기 위한 통신 모듈 등의 통신장치, 프로그램을 실행하기 위한 데이터를 저장하는 메모리, 프로그램을 실행하여 연산 및 명령하기 위한 마이크로프로세서 등을 전부 또는 일부 포함한 다양한 장치를 의미할 수 있다.
- [0060] 도 6은 본 발명의 다른 실시예에 따른 클래스 스코어 학습 방법을 예시한 흐름도이다. 클래스 스코어 학습 방법은 컴퓨팅 디바이스에 의하여 수행될 수 있으며, 클래스 스코어 학습 장치가 수행하는 동작에 관한 상세한 설명과 중복되는 설명은 생략하기로 한다.
- [0061] 단계 S610에서, 컴퓨팅 디바이스는 미디어에 관한 제1 학습 모델의 출력 레이어로부터 제1 클래스 스코어를 획득한다. 여기서 미디어는 이미지, 비디오, 오디오, 또는 이들의 조합을 포함하는 멀티미디어일 수 있다. 제1 학습 모델은 컨볼루션 네트워크로 구현되어 외형 특징을 기반으로 클래스를 분류하여 제1 클래스 스코어를 생성할 수 있다.
- [0062] 단계 S620에서, 컴퓨팅 디바이스는 미디어에 관한 제2 학습 모델의 출력 레이어로부터 제2 클래스 스코어를 획득한다. 제2 학습 모델은 컨볼루션 네트워크로 구현되어 움직임 특징을 기반으로 클래스를 분류하여 제2 클래스 스코어를 생성할 수 있다.
- [0063] 단계 S630에서, 컴퓨팅 디바이스는 제1 클래스 스코어에 관한 제1 가중치와 제2 클래스 스코어에 관한 제2 가중치를 학습하여 통합 클래스 스코어를 생성한다.
- [0064] 클래스 스코어 학습 방법은 미디어에 관한 제3 학습 모델의 출력 레이어로부터 클래스에 관한 제3 클래스 스코어를 획득하는 단계를 추가로 포함할 수 있다. 제3 학습 모델은 소리 특징을 기반으로 상기 클래스를 분류하여 제3 클래스 스코어를 생성할 수 있다.
- [0065] 통합 클래스 스코어를 생성하는 단계(S630)는 제1 클래스 스코어에 관한 제1 가중치, 제2 클래스 스코어에 관한 제2 가중치, 및 제3 클래스 스코어에 관한 제3 가중치를 학습하여 통합 클래스 스코어를 생성할 수 있다.
- [0066] 통합 클래스 스코어를 생성하는 단계(S630)는 상태 및 행동 쌍에 관한 기대 함수를 정의하고, 상태를 클래스로 설정하고, 행동을 제1 가중치와 제2 가중치로 설정한다. 행동을 제1 가중치, 제2 가중치, 및 제3 가중치로 설정할 수 있다.
- [0067] 통합 클래스 스코어를 생성하는 단계(S630)는 상태 중 현재 상태에서 행동에 대한 보상을 인식의 정확도로 설정하고, 행동 중 임의의 행동을 선택하는 탐색 동작과 현재 상태에 대해 기대 함수가 기 설정된 값보다 큰 값을 갖는 행동을 선택하는 획득 동작을 반복하여 기대 함수를 갱신할 수 있다.
- [0068] 클래스 스코어 학습 방법은 통합 클래스 스코어를 기준으로 클래스를 분류하여 미디어 내에서 표현된 행동을 인식하는 단계를 추가로 포함할 수 있다.
- [0069] 도 7은 본 발명의 실시예들에 따라 수행된 모의실험 결과를 도시한 것이다. 투 스트림 컨볼루션 네트워크에 대

한 대부분의 클래스에서 높은 행동 인식율을 갖는 것을 확인할 수 있다.

[0070] 도 6에서는 각각의 과정을 순차적으로 실행하는 것으로 기재하고 있으나 이는 예시적으로 설명한 것에 불과하고, 이 분야의 기술자라면 본 발명의 실시예의 본질적인 특성에서 벗어나지 않는 범위에서 도 6에 기재된 순서를 변경하여 실행하거나 또는 하나 이상의 과정을 병렬적으로 실행하거나 다른 과정을 추가하는 것으로 다양하게 수정 및 변형하여 적용 가능할 것이다.

[0071] 본 실시예들에 따른 동작은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능한 매체에 기록될 수 있다. 컴퓨터 판독 가능한 매체는 실행을 위해 프로세서에 명령어를 제공하는 데 참여한 임의의 매체를 나타낸다. 컴퓨터 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 또는 이들의 조합을 포함할 수 있다. 예를 들면, 자기 매체, 광기록 매체, 메모리 등이 있을 수 있다. 컴퓨터 프로그램은 네트워크로 연결된 컴퓨터 시스템 상에 분산되어 분산 방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수도 있다. 본 실시예를 구현하기 위한 기능적인(Functional) 프로그램, 코드, 및 코드 세그먼트들은 본 실시예가 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있을 것이다.

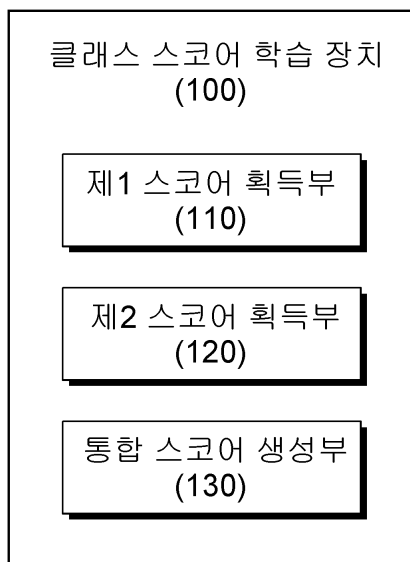
[0072] 본 실시예들은 본 실시예의 기술 사상을 설명하기 위한 것이고, 이러한 실시예에 의하여 본 실시예의 기술 사상의 범위가 한정되는 것은 아니다. 본 실시예의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 실시예의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

### 부호의 설명

[0073] 100, 200: 클래스 스코어 학습 장치  
110, 210: 제1 스코어 획득부  
120, 220: 제2 스코어 획득부  
230: 제3 스코어 획득부  
130, 240: 통합 스코어 생성부

### 도면

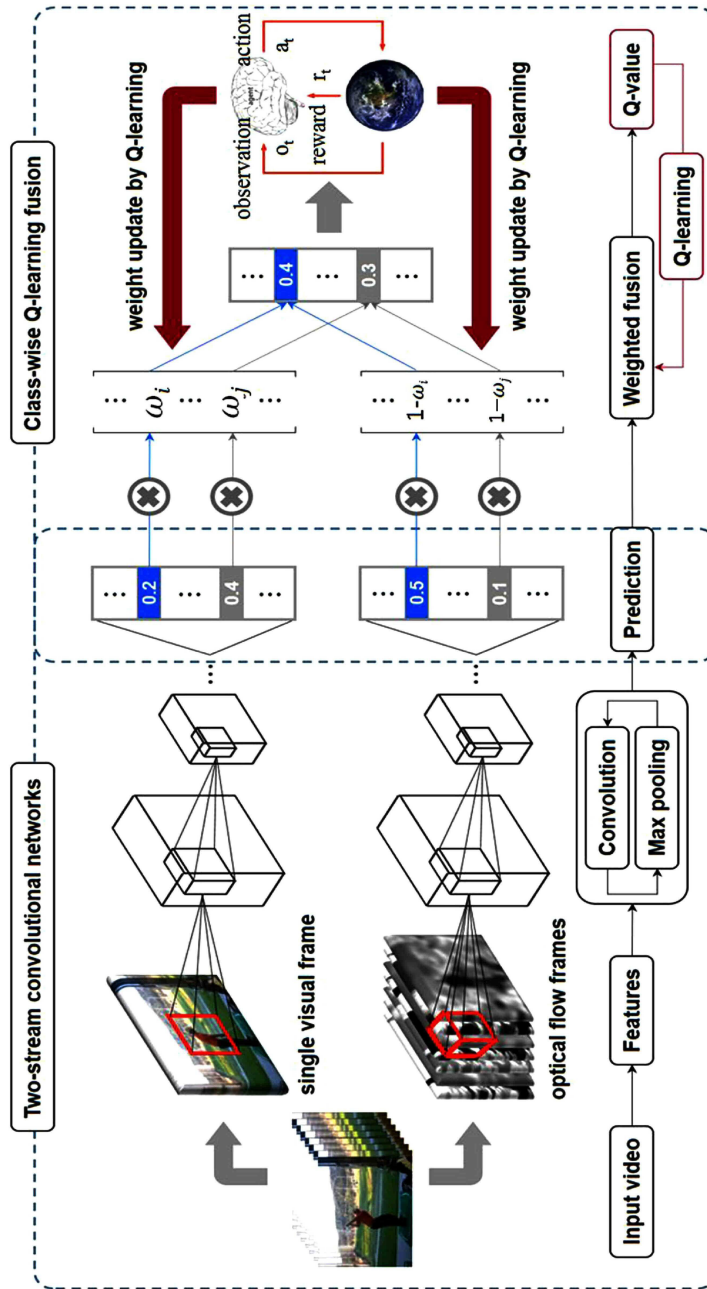
#### 도면1



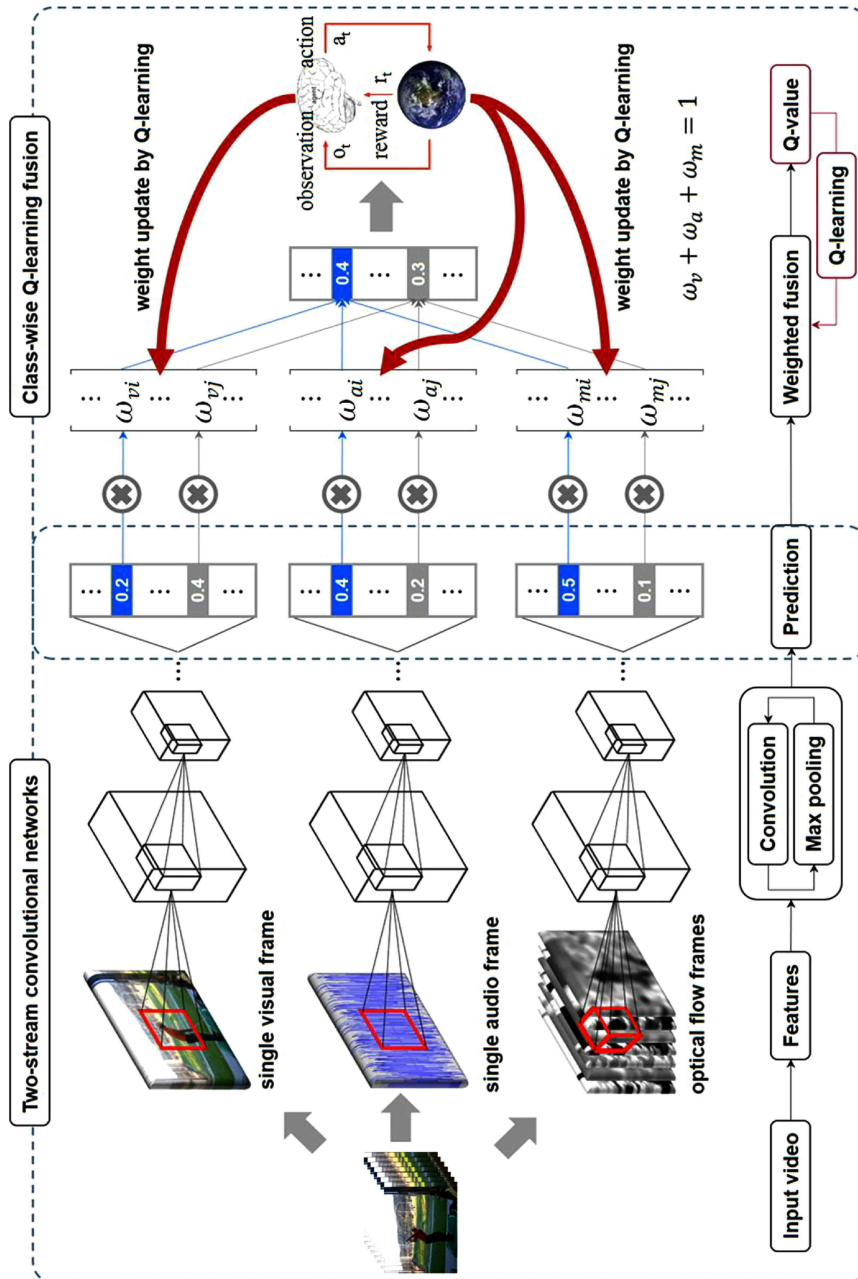
도면2



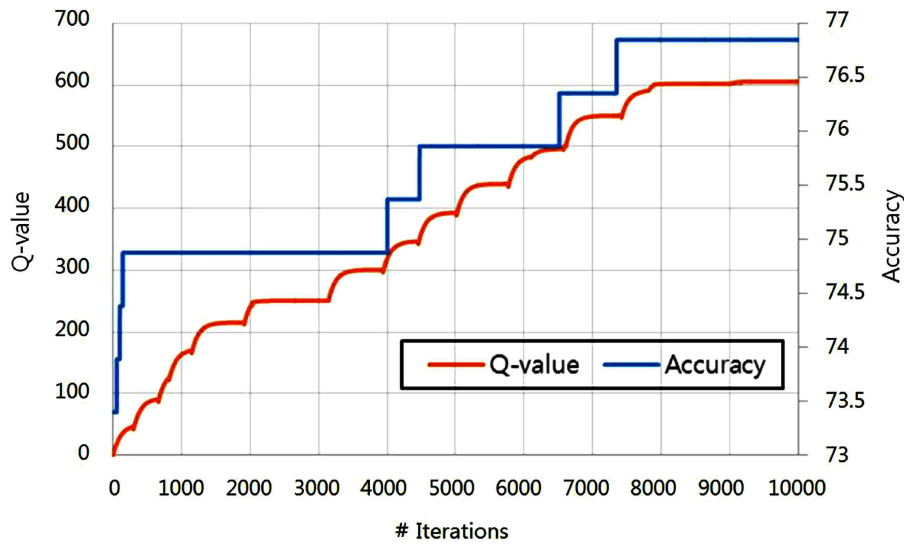
도면3



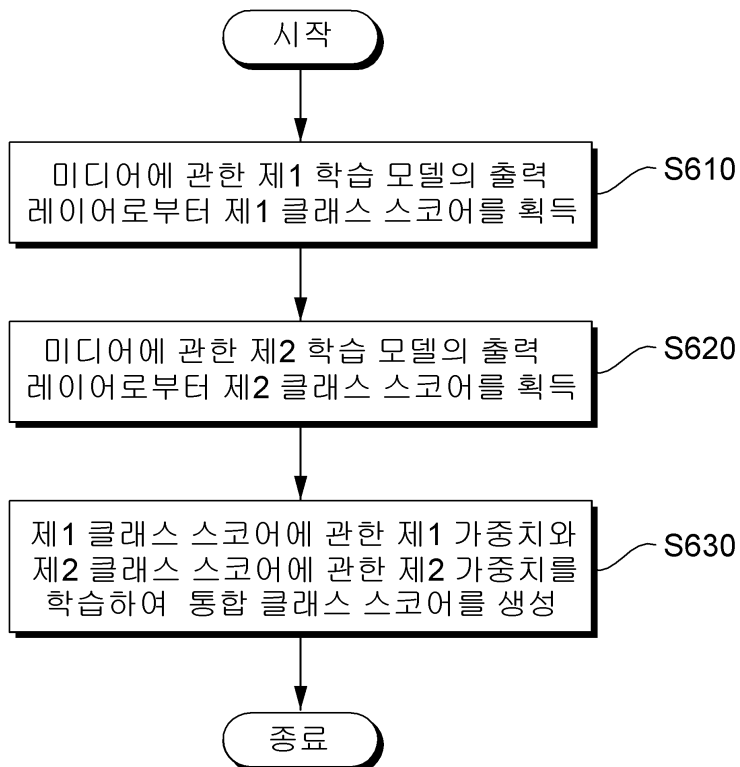
도면4



도면5



도면6



도면7

