



(19) 대한민국특허청(KR)  
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0029193  
(43) 공개일자 2023년03월03일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2023.01) G06N 7/00 (2023.01)  
(52) CPC특허분류  
G06N 3/08 (2023.01)  
G06N 3/047 (2023.01)  
(21) 출원번호 10-2021-0111402  
(22) 출원일자 2021년08월24일  
심사청구일자 2021년08월24일

(71) 출원인  
연세대학교 산학협력단  
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)  
(72) 발명자  
이상훈  
서울특별시 영등포구 신봉로 77, 103동 2201호(신길동, 래미안에스티움)  
강지우  
서울특별시 서대문구 연희로8길 31, 101호(연희동)  
(뒷면에 계속)  
(74) 대리인  
민영준

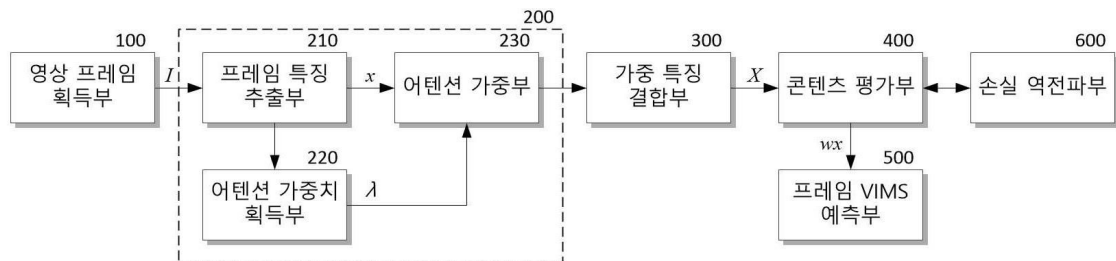
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법

(57) 요약

본 발명은 VR 콘텐츠를 구성하는 VR 영상에 포함된 다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하고, 다수의 특징맵 각각에 각 프레임의 중요도를 나타내는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득하는 어텐션 프레임 특징 추출부, 다수의 가중 특징맵 (뒷면에 계속)

대표도





이 결합된 결합 특징맵을 인가받아, 미리 학습된 신경망 연산 방식에 따라 결합 특징맵이 기지정된 개수의 클래스 각각에 대응하는 수준을 분석하여 다수의 클래스 각각에 대한 다수의 가중 특징맵의 서로 다른 기여도가 반영된 결합 클래스 특징맵을 획득하고, 학습에 의해 미리 설정된 클래스 가중치를 가중하여 VR 영상에 대한 VIMS 점수를 나타내는 클래스 확률을 판별하는 콘텐츠 평가부 및 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 추출되는 클래스 특징맵 각각에 클래스 가중치를 가중하고 모든 클래스에 대해 누산하여, 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 프레임 VIMS 예측부를 포함하여, VR 콘텐츠 영상에 의한 VIMS 점수를 프레임 단위로 정량적으로 예측할 수 있으므로, VR 콘텐츠 개발 비용 및 시간을 크게 저감시킬 수 있도록 하는 가상현실 콘텐츠 멀티 및 어지러움 예측 장치 및 방법을 제공한다.

(52) CPC특허분류

**G06N 7/01** (2023.01)

**G06T 19/003** (2013.01)

**김진우**

서울특별시 서대문구 신촌로1길 42-7(창천동)

(72) 발명자

**이경오**

서울특별시 서대문구 신촌로7안길 75(창천동)

**김우재**

서울특별시 마포구 신촌로 50(동교동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711134754
과제번호	2021-0-00352-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(한국연구재단부설)
연구사업명	정보통신방송연구개발사업
연구과제명	설명 가능한 감성 경험 예측 모델 기반 콘텐츠 평가 기술 개발 및 상용화
기 여 율	1/1
과제수행기관명	주식회사 그루크리에이티브랩
연구기간	2021.04.01 ~ 2021.12.31



## 명세서

### 청구범위

#### 청구항 1

VR 콘텐츠를 구성하는 VR 영상에 포함된 다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하고, 상기 다수의 특징맵 각각에 각 프레임의 중요도를 나타내는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득하는 어텐션 프레임 특징 추출부;

상기 다수의 가중 특징맵이 결합된 결합 특징맵을 인가받아, 미리 학습된 신경망 연산 방식에 따라 상기 결합 특징맵이 기지정된 개수의 클래스 각각에 대응하는 수준을 분석하여 다수의 클래스 각각에 대한 상기 다수의 가중 특징맵의 서로 다른 기여도가 반영된 결합 클래스 특징맵을 획득하고, 학습에 의해 미리 설정된 클래스 가중치를 가중하여 VR 영상에 대한 VIMS(Visually Induced Motion Sickness) 점수를 나타내는 클래스 확률을 판별하는 콘텐츠 평가부; 및

상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 추출되는 클래스 특징맵 각각에 상기 클래스 가중치를 가중하고 모든 클래스에 대해 누산하여, 상기 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 프레임 VIMS 예측부를 포함하는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치.

#### 청구항 2

제1항에 있어서, 상기 어텐션 프레임 특징 추출부는

다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하는 프레임 특징 추출부;

상기 다수의 특징맵에 대응하는 다수의 프레임 각각의 중요도를 나타내는 다수의 어텐션 가중치를 미리 학습된 신경망 연산 방식에 따라 획득하는 어텐션 가중치 획득부; 및

상기 다수의 특징맵 각각에 상기 다수의 어텐션 가중치 중 대응하는 어텐션 가중치를 가중하여 상기 다수의 가중 특징맵을 획득하는 어텐션 가중부를 포함하는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치.

#### 청구항 3

제2항에 있어서, 상기 콘텐츠 평가부는

상기 결합 클래스 특징맵을 인가받아, 상기 결합 클래스 특징맵이 기지정된 개수의 클래스 각각으로 분류될 확률을 나타내는 클래스 확률을 출력하도록 미리 학습된 완전 연결 레이어(Fully Connected Layer: FC)로 구현되는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치.

#### 청구항 4

제2항에 있어서, 상기 프레임 VIMS 예측부는

상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 각각에서 상기 어텐션 가중치를 제거하여 획득되는 다수의 클래스 특징맵 각각에 대응하는 클래스 가중치를 가중하여 각 클래스에 대해 프레임 단위로 획득되는 VIMS 점수를 모두 합산하여 상기 프레임 VIMS를 획득하는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치.

#### 청구항 5

제2항에 있어서, 상기 프레임 특징 추출부는

미리 학습된 Resnet(Residential network) 또는 CNN(Convolutional Neural Networks) 중 하나로 구현되고,

상기 어텐션 가중치 획득부는

미리 학습된 완전 연결 레이어(Fully Connected Layer: FC)로 구현되는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치.



## 청구항 6

VR 콘텐츠를 구성하는 VR 영상에 포함된 다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 상기 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 가상현실 콘텐츠 멀미 및 어지러움 예측 방법에 있어서,

미리 학습된 인공 신경망을 이용하여 상기 다수의 프레임 각각의 특징을 추출하여 다수의 특징맵을 획득하고, 상기 다수의 특징맵 각각에 각 프레임의 중요도를 나타내는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득하는 단계;

미리 학습된 인공 신경망을 이용하여 상기 다수의 가중 특징맵이 결합된 결합 특징맵을 인가받아, 학습된 방식에 따라 상기 결합 특징맵이 기지정된 개수의 클래스 각각에 대응하는 수준을 분석하여 다수의 클래스 각각에 대한 상기 다수의 가중 특징맵의 서로 다른 기여도가 반영된 결합 클래스 특징맵을 획득하고, 학습에 의해 미리 설정된 클래스 가중치를 가중하여 VR 영상에 대한 VIMS(Visually Induced Motion Sickness) 점수를 나타내는 클래스 확률을 판별하는 단계; 및

상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 추출되는 클래스 특징맵 각각에 상기 클래스 가중치를 가중하고 모든 클래스에 대해 누산하여, 상기 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 단계를 포함하는 가상현실 콘텐츠 멀미 및 어지러움 예측 방법.

## 청구항 7

제6항에 있어서, 상기 다수의 가중 특징맵을 획득하는 단계는

다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하는 단계;

미리 학습된 신경망 연산 방식에 따라 상기 다수의 특징맵에 대응하는 다수의 프레임 각각의 중요도를 나타내는 다수의 어텐션 가중치를 획득하는 단계; 및

상기 다수의 가중 특징맵을 획득하기 위해 상기 다수의 특징맵 각각에 상기 다수의 어텐션 가중치 중 대응하는 어텐션 가중치를 가중하는 단계를 포함하는 가상현실 콘텐츠 멀미 및 어지러움 예측 방법.

## 청구항 8

제7항에 있어서, 상기 클래스 확률을 판별하는 단계는

미리 학습된 완전 연결 레이어(Fully Connected Layer: FC)를 이용하여 인가된 상기 결합 클래스 특징맵이 기지정된 개수의 클래스 각각으로 분류될 확률을 나타내는 클래스 확률을 출력하는 가상현실 콘텐츠 멀미 및 어지러움 예측 방법.

## 청구항 9

제7항에 있어서, 상기 프레임 VIMS를 획득하는 단계는

상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 각각에서 상기 어텐션 가중치를 제거하여 획득되는 다수의 클래스 특징맵 각각에 대응하는 클래스 가중치를 가중하여 각 클래스에 대해 프레임 단위로 획득되는 VIMS 점수를 모두 합산하여 상기 프레임 VIMS를 획득하는 가상현실 콘텐츠 멀미 및 어지러움 예측 방법.

## 청구항 10

제7항에 있어서, 상기 다수의 특징맵을 획득하는 단계는

미리 학습된 Resnet(Residential network) 또는 CNN(Convolutional Neural Networks) 중 하나를 이용하여 수행되고,

상기 다수의 어텐션 가중치를 획득하는 단계는

미리 학습된 완전 연결 레이어(Fully Connected Layer: FC)를 이용하여 수행되는 가상현실 콘텐츠 멀미 및 어지러움 예측 방법.



## 발명의 설명

### 기술 분야

- [0001] 본 발명은 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법에 관한 것으로, 약지도 학습 및 클래스 활성화 맵 기법을 응용하여 프레임별 멀미 수준을 예측할 수 있는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법에 관한 것이다.

### 배경 기술

- [0002] 현재 디스플레이 기술의 성장으로 인하여, 증강현실(Augmented Reality: AR), 가상현실(Virtual Reality: 이하 VR) 및 혼합현실(Mixed Reality: 이하 MR) 콘텐츠가 기하급수적으로 생산되고 있다. 이 중에서도 VR 콘텐츠는 게임산업과 접목되어 매우 활발하게 이용되고 있으며, 사용자로부터 더욱 고품질의 콘텐츠가 요구되고 있다. 그러나 VR 콘텐츠는 좌, 우의 디스플레이가 양쪽 눈으로 확장되어 투영되므로, 넓은 시야(field of view: 이하 FoV), 시간 지연(time latency), 낮은 프레임율(low frame rate), 렌더링 속도(rendering speed) 등의 요인으로 인해, 사용자가 원치 않는 멀미 및 어지러움을 유발할 수 있다는 문제가 있다. 이와 같이 VR 콘텐츠 이용함에 따라 유발되는 멀미와 어지러움을 시각 유도 멀미(Visually Induced Motion Sickness: 이하 VIMS)라 하며, VIMS는 VR에 대한 사용자의 접근을 저해하는 요소가 되어, VR 콘텐츠 개발 산업 발전의 큰 장애로 작용하고 있다.
- [0003] 따라서 사용자가 쾌적하게 VR 콘텐츠를 이용할 수 있도록 하기 위해서는 먼저 VR 콘텐츠 이용 시 발생하는 VIMS를 정량적으로 측정할 수 있어야 하며, 관련된 연구가 기존에도 다양하게 수행되어 왔다.
- [0004] VIMS를 정량적으로 측정하는 대표적인 방법은 대량의 VR 콘텐츠 각각에 대해 다수의 주관적 평가자들을 통해서 주관적 점수를 획득하고, 획득된 주관적 점수를 기반으로 인공 신경망을 학습시킴으로써, 이후 인공 신경망이 새로운 VR 콘텐츠의 VIMS에 대한 정량적 점수를 예측하도록 하는 것이다. 일 예로 주관적 평가자들 각각은 각각의 VR 콘텐츠를 경험 후, VIMS를 미리 지정된 범위(여기서는 일 예로 1 ~ 5점) 내의 점수로 평가하도록 한다. 이렇게 평가된 점수는 신뢰할 수 있는 실측 자료로 간주되며, 측정된 실측 자료를 통해 인공 신경망을 학습시킨다.
- [0005] 다만 이와 같은 학습 방식으로 학습된 인공 신경망은 각 VR 콘텐츠에 대한 VIMS를 예측할 수는 있으나, VR 콘텐츠를 구성하는 영상 내의 각 구간에서 유발되는 VIMS 수준을 구분하여 예측할 수 없다는 한계가 있다. 이로 인해, 기존의 방식에서는 VR 콘텐츠의 VIMS를 예측하더라도, 이를 제대로 활용하기 어렵다. 즉 특정 VR 콘텐츠가 유발하는 멀미나 어지러움 수준이 매우 높다고 판단될지라도, 이는 VR 콘텐츠 전체가 유발하는 멀미나 어지러움 수준을 나타내므로, VIMS를 낮추기 위해 수정되어야 하는 구간을 구분할 수 없다. 이로 인해 최악의 경우, 해당 VR 콘텐츠 전체를 폐기하고 새로이 제작해야 할 수도 있어, VR 콘텐츠 개발 시에 비용 및 시간이 증가되는 요인으로 작용하고 있다.

## 선행기술문헌

### 특허문헌

- [0006] (특허문헌 0001) 한국 등록 특허 제10-2141740호 (2020.07.30 공개)

## 발명의 내용

### 해결하려는 과제

- [0007] 본 발명의 목적은 VR 콘텐츠 영상에 의한 VIMS를 프레임 단위로 정량적으로 예측할 수 있는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법을 제공하는데 있다.
- [0008] 본 발명의 다른 목적은 약지도 학습 및 클래스 활성화 맵 기법을 응용하여 VR 콘텐츠 영상의 개별 프레임에 대한 학습 데이터 없이도 정확하게 VIMS를 프레임 단위로 정량적으로 예측하도록 학습되는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법을 제공하는데 있다.



## 과제의 해결 수단

- [0009] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치는 VR 콘텐츠를 구성하는 VR 영상에 포함된 다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하고, 상기 다수의 특징맵 각각에 각 프레임의 중요도를 나타내는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득하는 어텐션 프레임 특징 추출부; 상기 다수의 가중 특징맵이 결합된 결합 특징맵을 인가받아, 미리 학습된 신경망 연산 방식에 따라 상기 결합 특징맵이 기지정된 개수의 클래스 각각에 대응하는 수준을 분석하여 다수의 클래스 각각에 대한 상기 다수의 가중 특징맵의 서로 다른 기여도가 반영된 결합 클래스 특징맵을 획득하고, 학습에 의해 미리 설정된 클래스 가중치를 가중하여 VR 영상에 대한 VIMS(Visually Induced Motion Sickness) 점수를 나타내는 클래스 확률을 판별하는 콘텐츠 평가부; 및 상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 추출되는 클래스 특징맵 각각에 상기 클래스 가중치를 가중하고 모든 클래스에 대해 누산하여, 상기 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 프레임 VIMS 예측부를 포함한다.
- [0010] 상기 어텐션 프레임 특징 추출부는 다수의 프레임을 인가받아 미리 학습된 신경망 연산 방식에 따라 각 프레임의 특징을 추출하여 다수의 특징맵을 획득하는 프레임 특징 추출부; 상기 다수의 특징맵에 대응하는 다수의 프레임 각각의 중요도를 나타내는 다수의 어텐션 가중치를 미리 학습된 신경망 연산 방식에 따라 획득하는 어텐션 가중치 획득부; 및 상기 다수의 특징맵 각각에 상기 다수의 어텐션 가중치 중 대응하는 어텐션 가중치를 가중하여 상기 다수의 가중 특징맵을 획득하는 어텐션 가중부를 포함할 수 있다.
- [0011] 상기 콘텐츠 평가부는 상기 결합 클래스 특징맵을 인가받아, 상기 결합 클래스 특징맵이 기지정된 개수의 클래스 각각으로 분류될 확률을 나타내는 클래스 확률을 출력하도록 미리 학습된 완전 연결 레이어(Fully Connected Layer: FC)로 구현될 수 있다.
- [0012] 상기 프레임 VIMS 예측부는 상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 각각에서 상기 어텐션 가중치를 제거하여 획득되는 다수의 클래스 특징맵 각각에 대응하는 클래스 가중치를 가중하여 각 클래스에 대해 프레임 단위로 획득되는 VIMS 점수를 모두 합산하여 상기 프레임 VIMS를 획득할 수 있다.
- [0013] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 방법은 미리 학습된 인공 신경망을 이용하여 상기 다수의 프레임 각각의 특징을 추출하여 다수의 특징맵을 획득하고, 상기 다수의 특징맵 각각에 각 프레임의 중요도를 나타내는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득하는 단계; 미리 학습된 인공 신경망을 이용하여 상기 다수의 가중 특징맵이 결합된 결합 특징맵을 인가받아, 학습된 방식에 따라 상기 결합 특징맵이 기지정된 개수의 클래스 각각에 대응하는 수준을 분석하여 다수의 클래스 각각에 대한 상기 다수의 가중 특징맵의 서로 다른 기여도가 반영된 결합 클래스 특징맵을 획득하고, 학습에 의해 미리 설정된 클래스 가중치를 가중하여 VR 영상에 대한 VIMS(Visually Induced Motion Sickness) 점수를 나타내는 클래스 확률을 판별하는 단계; 및 상기 결합 클래스 특징맵의 다수의 가중 클래스 특징맵에서 추출되는 클래스 특징맵 각각에 상기 클래스 가중치를 가중하고 모든 클래스에 대해 누산하여, 상기 다수의 프레임 각각에 대한 VIMS를 나타내는 프레임 VIMS를 획득하는 단계를 포함한다.

## 발명의 효과

- [0014] 따라서, 본 발명의 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치 및 방법은 VR 콘텐츠 영상에 의한 VIMS 점수를 프레임 단위로 정량적으로 예측할 수 있으므로, VIMS를 크게 유발하는 프레임 구간을 용이하게 검출하고 수정할 수 있도록 하여 VR 콘텐츠 개발 비용 및 시간을 크게 저감시킬 수 있도록 한다. 또한 약지도 학습 및 클래스 활성화 맵 기법을 적용하여 VR 콘텐츠 영상에 포함된 다수의 프레임 각각에 대한 주관적 VIMS 점수없이 VR 콘텐츠 전체에 대한 주관적 VIMS 점수를 학습 데이터로 이용하여 학습될 수 있어, 프레임 단위로 VIMS를 검출할 수 있도록 학습됨에도 기존과 동일한 수준의 학습 데이터만을 요구하므로 학습 데이터의 획득이 용이하다.

## 도면의 간단한 설명

- [0015] 도 1은 본 발명의 일 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 개략적 구조를 나타낸다.
- 도 2는 도 1의 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 각 구성별 동작을 설명하기 위한 도면이다.
- 도 3은 클래스 활성화 맵 기법을 설명하기 위한 도면이다.



도 4는 본 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치가 VR 콘텐츠 영상의 각 프레임에 대해 예측한 VIMS 점수를 그래프로 표현한 결과를 나타낸다.

도 5는 본 발명의 일 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 방법을 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시예에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0017] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 설명함으로써, 본 발명을 상세히 설명한다. 그러나, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 설명하는 실시예에 한정되는 것이 아니다. 그리고, 본 발명을 명확하게 설명하기 위하여 설명과 관계없는 부분은 생략되며, 도면의 동일한 참조부호는 동일한 부재임을 나타낸다.
- [0018] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라, 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈", "블록" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0019] 도 1은 본 발명의 일 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 개략적 구조를 나타내고, 도 2는 도 1의 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 각 구성별 동작을 설명하기 위한 도면이며, 도 3은 클래스 활성화 맵 기법을 설명하기 위한 도면이다.
- [0020] 도 1 및 도 2를 참조하면, 본 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치는 영상 프레임 획득부(100), 어텐션 프레임 특징 추출부(200), 가중 특징 결합부(300), 콘텐츠 평가부(400) 및 프레임 VIMS 예측부(500)를 포함할 수 있다.
- [0021] 영상 프레임 획득부(100)는 VIMS가 측정되어야 하는 VR 콘텐츠를 구성하는 VR 영상을 획득한다.
- [0022] 영상 프레임 획득부(100)는 VR 콘텐츠를 획득하거나, VR 콘텐츠에 포함된 VR 영상만을 별도로 획득할 수 있다. 상기한 바와 같이, VR 콘텐츠는 VR 영상과 음향 등을 포함하여 구성되므로, 영상 프레임 획득부(100)는 VR 콘텐츠가 획득되면 획득된 VR 콘텐츠에 포함된 VR 영상을 별도로 추출할 수 있다. 그리고 영상 프레임 획득부(100)는 획득된 VR 영상에 포함된 다수의 프레임( $I_1 \sim I_T$ )을 각각 구분하여 어텐션 프레임 특징 추출부(200)로 전달할 수 있다. 이때 영상 프레임 획득부(100)는 VR 영상에서 구분된 다수의 프레임( $I_1 \sim I_T$ )을 영상 내 시간의 흐름에 따라 순차적으로 어텐션 프레임 특징 추출부(200)로 전달할 수 있다.
- [0023] 어텐션 프레임 특징 추출부(200)는 영상 프레임 획득부(100)에서 구분되어 전달되는 다수의 프레임( $I_1 \sim I_T$ ) 각각에 대해 특징을 추출하여 다수의 특징맵( $x_1 \sim x_T$ )을 획득하고, 획득된 각 프레임( $I_1 \sim I_T$ )에 대한 특징맵( $x_1 \sim x_T$ )에 각 프레임의 중요도를 나타내는 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 가중하여 가중 특징맵을 획득한다.
- [0024] 어텐션 프레임 특징 추출부(200)는 프레임 특징 추출부(210), 어텐션 가중치 획득부(220) 및 어텐션 가중부(230)를 포함할 수 있다.
- [0025] 프레임 특징 추출부(210)는 영상 프레임 획득부(100)에서 인가되는 다수의 프레임( $I_1 \sim I_T$ ) 각각에 대해 미리 학습된 신경망 연산 방식에 따라 특징을 추출하여 다수의 특징맵( $x_1 \sim x_T$ )을 획득한다. 프레임 특징 추출부(210)는 미리 학습된 인공 신경망으로 구현될 수 있으며, 도 2에서는 일 예로 Resnet(Residential network)로 구현되는 경우를 가정하여 도시하였으나 CNN(Convolutional Neural Networks)과 같은 다른 신경망으로 구현될 수도 있다.
- [0026] 어텐션 가중치 획득부(220)는 프레임 특징 추출부(210)에서 획득된 다수의 특징맵( $x_1 \sim x_T$ )을 인가받고, 인가된 다수의 특징맵( $x_1 \sim x_T$ ) 각각의 중요도를 나타내는 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 미리 학습된 신경망 연산 방식에 따라 추출한다. 어텐션 가중치 획득부(220)는 도 2에 도시된 바와 같이, 완전 연결 레이어(Fully Connected Layer: FC)로 구현되어 인가된 특징맵( $x_1 \sim x_T$ )에 따른 중요도를 분류하여 대응하는 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 획득할 수 있다. 여기서 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )는 스칼라 값으로 획득된다.



- [0027] 어텐션 가중부(230)는 다수의 특징맵( $x_1 \sim x_T$ ) 각각에 어텐션 가중치 획득부(220)에서 획득된 다수의 어텐션 가중치( $\lambda_1 \sim \lambda_T$ ) 중 대응하는 어텐션 가중치를 가중하여 다수의 가중 특징맵을 획득한다. 어텐션 가중부(230)는 각 특징맵( $x_1 \sim x_T$ )의 다수의 픽셀 각각에 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 곱하여 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )을 획득할 수 있다.
- [0028] 본 실시예에서 어텐션 프레임 특징 추출부(200)가 다수의 프레임( $I_1 \sim I_T$ )에 대한 다수의 특징맵( $x_1 \sim x_T$ )을 획득할 뿐만 아니라, 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 추가로 획득하여 대응하는 특징맵( $x_1 \sim x_T$ )에 가중하는 것은 다수의 특징맵( $x_1 \sim x_T$ )을 기반으로 각 프레임( $I_1 \sim I_T$ )의 중요도를 분석하여 강조할 수 있도록 하기 위함이다.
- [0029] 가중 특징 결합부(300)는 어텐션 프레임 특징 추출부(200)에서 획득된 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )을 결합(concatenation)하여 결합 특징맵(X)을 획득한다.
- [0030] 콘텐츠 평가부(400)는 결합 특징맵(X)을 인가받아 미리 학습된 신경망 연산 방식에 따라 결합 특징맵(X)을 분석하여 VIMS 점수( $S_{pred}$ )를 예측한다. 콘텐츠 평가부(400)는 학습된 방식에 따라 결합 특징맵(X)을 분석하여, 결합 특징맵(X)이 미리 지정된 개수(여기서는 일 예로 m개)의 클래스 각각에 대응하는 수준, 즉 결합 특징맵(X)에 따라 VR 영상이 각 클래스로 분류될 클래스 분류 확률을 추정할 수 있다. 여기서 결합 특징맵(X)의 클래스는 VR 영상에 대한 VIMS 점수 분포일 수 있다.
- [0031] 예를 들면 콘텐츠 평가부(400)는 1점에서 5점까지 이산 분류된 VIMS 점수 각각을 분류 가능한 클래스로 설정하고, 학습된 방식에 따라 결합 특징맵(X)이 5개의 클래스( $k = 1 \sim 5$ ) 각각으로 분류될 확률을 추정하여, 가장 높은 확률로 분류된 클래스를 VR 영상의 VIMS 점수( $S_{pred}$ )로 예측할 수 있다.
- [0032] 여기서 콘텐츠 평가부(400)는 일 예로 도 2에 도시된 바와 같이, 완전 연결 레이어(FC)와 시그모이드 함수(Sigmoid) 등으로 구현되는 활성화 함수(Activation Function)로 구현될 수 있으나, 활성화 함수는 시그모이드 함수 이외에 알려진 다양한 함수로 구현될 수도 있으며, 경우에 따라서는 생략될 수도 있다.
- [0033] 학습된 콘텐츠 평가부(400)는 학습된 방식에 따라 결합 특징맵(X)을 구성하는 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ ) 각각이 각 클래스( $k = 1 \sim m$ )에 대응하는 수준을 분석하여, 다수의 클래스 각각에 대한 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )의 서로 다른 기여도가 반영된 결합 클래스 특징맵( $X(k)$ )을 획득하고, 획득된 결합 클래스 특징맵( $X(k)$ )에 학습에 의해 미리 설정된 클래스 가중치( $w(k)$ )를 가중하여 결합 특징맵(X)의 클래스 확률을 판별함으로써 VIMS 점수( $S_{pred}$ )를 예측하는 동작을 수행하는 것으로 볼 수 있다.
- [0034] 따라서 콘텐츠 평가부(400)에서 예측된 VIMS 점수( $S_{pred}$ )는 수학적 식 1과 같이 표현될 수 있다.

### 수학적 식 1

$$S_{pred} = \sum_{k=1}^m w(k)X(k)$$

- [0035]
- [0036] 한편, 프레임 VIMS 예측부(500)는 콘텐츠 평가부(400)가 결합 특징맵(X)으로부터 VIMS 점수( $S_{pred}$ )를 예측하는 과정에서 다수의 클래스 각각으로 분류하기 위해 결합 특징맵(X)의 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ ) 각각에 대해 주목한 수준과 각 클래스(k)에 대한 클래스 가중치( $w(k)$ )를 기반으로 VR 영상에 포함되는 다수의 프레임( $I_1 \sim I_T$ ) 각각의 VIMS를 예측할 수 있다.
- [0037] 상기한 바와 같이, 결합 특징맵(X)은 가중 특징 결합부(300)에서 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )이 결합되어 획득되었으므로, 결합 클래스 특징맵( $X(k)$ )은 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ ) 각각의 클래스별 기여도가 반영된 다수의 가중 클래스 특징맵( $\lambda_1 x_1(k) \sim \lambda_T x_T(k)$ )이 결합된 것으로 볼 수 있으므로, 수학적 식 2와 같이 표현될 수 있다.



## 수학식 2

$$X(k) = \sum_{k=1}^m \sum_{t=1}^T \lambda_t x_t(k)$$

수학식 2를 수학식 1에 대입하면, 수학식 3과 같이 다시 표현될 수 있다.

## 수학식 3

$$S_{pred} = \sum_{k=1}^m w(k) \sum_{t=1}^T \lambda_t x_t(k) = \sum_{t=1}^T \lambda_t \sum_{k=1}^m w(k) x_t(k)$$

수학식 3에 따르면, 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )에서 시간(t)에 대한 각 특징맵( $x_1 \sim x_T$ )을 개별 클래스(k)에 대한 기여도가 반영된 클래스 특징맵( $x_1(k) \sim x_T(k)$ )에 따라 구분하여 클래스 가중치(w(k))를 가중한 후, 대응하는 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 가중하여도 동일한 결과로 나타나게 됨을 알 수 있다.

상기한 바와 같이, 콘텐츠 평가부(400)는 다수의 클래스 각각에 대한 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ ) 각각의 기여도가 반영된 결합 클래스 특징맵(X(k))을 획득하고, 이에 각 클래스에 대한 가중치로서 클래스 가중치(w(k))를 가중하여 각 클래스에 대응할 확률을 판별한다. 따라서 프레임 VIMS 예측부(500)는 결합 클래스 특징맵(X(k))과 설정된 클래스 가중치(w(k))가 클래스 활성화 맵(Class Activation Mapping: 이하 CAM)과 유사하게 이용할 수 있다.

클래스 활성화 맵(Class Activation Mapping: CAM) 기법은 딥 러닝 모델의 추론과 결과 생성 과정을 분석하는 학문으로 알려진 설명가능 인공지능(Explainable AI: 이하 XAI)의 대표적인 기법이다. CAM 기법은 주로 컴퓨터 비전(Computer Vision) 분야의 인공 신경망을 설명하기 위해 이용되는 방법으로, 인공 신경망에 의해 구분되는 각 클래스를 분류하기 위해 인공 신경망이 주목한 영역을 시각적으로 표시하는 기법이다.

도 3에 도시된 바와 같이, CAM 기법은 적어도 하나의 레이어를 포함하는 인공 신경망에서 마지막 레이어가 분류된 클래스값을 획득하기 위해 주목한 영역을 추적할 수 있도록, 마지막으로 추출된 특징맵에 대해 각 클래스에 대응하는 클래스 가중치( $w_1 \sim w_n$ )를 가중하고 누적 가산하여 CAM을 획득한다. 그리고 획득된 CAM은 입력된 이미지에서 해당 클래스를 식별하기 위해 주목한 영역이 강조되는 형태로 나타나게 된다.

다만, 본 발명의 경우, 도 2에 도시된 바와 같이, 다수의 프레임( $I_1 \sim I_T$ )으로 구성되는 VR 영상이 입력되고, 결합 특징맵(X)이 프레임 단위( $I_1 \sim I_T$ )로 추출된 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )을 결합하여 획득되므로, 개별 프레임 내에서 영역별로 강조되는 형태가 아니라, VR 영상내에서 각 프레임 단위로 강조되는 형태의 CAM이 획득될 수 있으며, 획득된 CAM은 개별 프레임( $I_1 \sim I_T$ )의 VIMS로 볼 수 있다.

프레임 VIMS 예측부(500)는 개별 프레임( $I_1 \sim I_T$ ) 각각에 대한 프레임 VIMS( $a_t$ )를 클래스 가중치(w(k))와 클래스 특징맵( $x_1(k) \sim x_T(k)$ )을 이용하여 수학식 4에 따라 획득할 수 있다.

## 수학식 4

$$a_t = \sum_{k=1}^m w(k) x_t(k)$$

수학식 4에 따라 계산되는 프레임 VIMS( $a_t$ )는 각 프레임( $I_1 \sim I_T$ )이 VR 영상 전체에 대해 예측된 VIMS 점수에 어



는 정도 연관되어 있는지를 나타낸다.

- [0049] 수학식 4에 따르면, 프레임  $VIMS(a_t)$ 는 결합 클래스 특징맵( $X(k)$ )의 다수의 가중 클래스 특징맵( $\lambda_1 x_1(k) \sim \lambda_{T_T} x_{T_T}(k)$ ) 각각에서 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 제거하여 획득되는 다수의 클래스 특징맵( $x_1(k) \sim x_T(k)$ ) 각각에 대응하는 클래스 가중치( $w(k)$ )를 가중하여 각 클래스에 대해 프레임 단위로 획득되는 VIMS 점수를 모두 합산하여 획득되는 것으로 볼 수 있다.
- [0050] 한편 본 실시예의 가상현실 콘텐츠 멀미 및 어지러움 예측 장치는 손실 역전파부(600)를 더 구비할 수 있다. 손실 역전파부(600)는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치를 학습 시에 구비되고, 실제로 이용 시에는 생략될 수 있다. 손실 역전파부(600)는 다수의 학습 VR 영상과 다수의 학습 VR 영상 각각에 대해 다수(예를 들면 100명)의 사람들이 주관적으로 VIMS 점수를 평가한 학습 VIMS를 학습 데이터로 획득할 수 있다. 여기서 다수의 학습 VR 영상 각각에 대한 학습 VIMS는 다수의 사람들의 평가 결과를 통계적으로 취합하여 획득된 점수의 형태로 나타날 수 있으며, 미리 지정된 클래스에 대응하여 이산된 값으로 획득될 수 있다. 그리고 손실 역전파부(600)는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치의 영상 프레임 획득부(100)에 학습 VR 영상을 제공하고, 콘텐츠 평가부(400)에서 예측되어 출력되는 VIMS와 미리 획득된 학습 VIMS 사이의 차이를 계산하여 역전파함으로써 가상현실 콘텐츠 멀미 및 어지러움 예측 장치를 학습시킬 수 있다.
- [0051] 가상현실 콘텐츠 멀미 및 어지러움 예측 장치는 실제 이용되기 이전에 미리 학습이 수행되어야 한다. 일반적으로 다수의 프레임 각각에 대한 VIMS를 추정하기 위해 학습시키기 위해서는, 개별 프레임( $I_t$ ) 각각의  $VIMS(a_t)$ 가 미리 주석된 대량의 VR 영상이 학습 데이터로 요구된다. 그러나 상기한 바와 같이, 본 실시예의 가상현실 콘텐츠 멀미 및 어지러움 예측 장치에서는 프레임 VIMS 예측부(500)가 VR 영상에 대한 VIMS를 예측하는 과정에서 획득되는 클래스 가중치( $w(k)$ )와 클래스 특징맵( $x_1(k) \sim x_T(k)$ )을 이용하여 프레임  $VIMS(a_t)$ 를 획득하므로, VR 영상 전체의 VIMS가 미리 주석된 학습 데이터로 학습되어도 각 프레임( $I_t$ )에 대한 프레임  $VIMS(a_t)$ 를 획득할 수 있다.
- [0052] 따라서 본 실시예의 가상현실 콘텐츠 멀미 및 어지러움 예측 장치에서는 상기한 바와 같이, 약지도 학습 방식에 따라 개별 프레임( $I_t$ )이 아닌 VR 영상에 대한 VIMS를 기반으로 학습이 수행될 수 있다.
- [0053] 도 4는 본 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치가 VR 콘텐츠 영상의 각 프레임에 대해 예측한 VIMS 점수를 그래프로 표현한 결과를 나타낸다.
- [0054] 도 4에서 (a)는 VR 영상의 회전이나 이동에 따른 각 프레임에서의 VIMS 변화를 나타내고, (b)는 배경과 해상도 등에 의한 각 프레임에서의 VIMS 변화를 나타낸다. (a)와 (b)에서는 VIMS와 함께 어텐션 가중치( $\lambda$ )도 함께 도시하였다.
- [0055] 그리고 도 4에서는 가상현실 콘텐츠 멀미 및 어지러움 예측 장치가 예측한 VIMS 점수가 높은 프레임, 즉 멀미나 어지러움을 크게 유발하는 프레임에 대해 밝게 표현하였고, VIMS 점수가 낮은 프레임에 대해서는 어둡게 표시하였다. 도 4에 도시된 바와 같이, 본 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 장치는 회전이나 고속 이동, 또는 해상도 저하 등으로 멀미를 더 크게 유발하게 되는 프레임에서 밝게 표현되는 반면, 변화가 적은 프레임에서는 어둡게 표현되어 각 프레임별로 VIMS가 정확하게 예측됨을 알 수 있다.
- [0056] 도 5는 본 발명의 일 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 방법을 나타낸다.
- [0057] 도 1 내지 도 3을 참조하여, 본 실시예에 따른 가상현실 콘텐츠 멀미 및 어지러움 예측 방법을 설명하면, 우선 VR 콘텐츠에 포함된 VR 영상을 획득하고, 획득된 VR 영상을 구성하는 다수의 프레임( $I_1 \sim I_T$ ) 각각을 구분하여 획득한다(S10). 그리고 구분된 다수의 프레임( $I_1 \sim I_T$ ) 각각에 대해 미리 학습된 신경망 연산 방식에 따라 특징을 추출하여 다수의 특징맵( $x_1 \sim x_T$ )을 획득한다(S20).
- [0058] 다수의 특징맵( $x_1 \sim x_T$ )이 획득되면, 다수의 특징맵( $x_1 \sim x_T$ ) 각각의 중요도를 나타내는 어텐션 가중치( $\lambda_1 \sim \lambda_T$ )를 미리 학습된 신경망 연산 방식에 따라 추출한다(S30). 그리고 추출된 어텐션 가중치( $\lambda_1 \sim \lambda_T$ ) 각각을 다수의 특징맵( $x_1 \sim x_T$ ) 중 대응하는 특징맵에 가중하여 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )을 획득한다(S40).
- [0059] 이후 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )을 결합(concatenation)하여 결합 특징맵( $X$ )을 획득한다(S50). 그리고



미리 학습된 신경망 연산 방식에 따라 결합 특징맵(X)을 구성하는 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ ) 각각이 VIMS 점수를 나타내는 기지정된 개수의 클래스( $k = 1 \sim m$ ) 각각에 대응하는 수준을 분석하여, 다수의 클래스 각각에 대한 다수의 가중 특징맵( $\lambda_1 x_1 \sim \lambda_T x_T$ )의 서로 다른 기여도가 반영된 결합 클래스 특징맵(X(k))을 획득하고, 획득된 결합 클래스 특징맵(X(k))에 학습에 의해 미리 설정된 클래스 가중치(w(k))를 가중하여 결합 특징맵(X)의 클래스 확률을 판별함으로써 VR 영상에 대한 VIMS 점수( $S_{pred}$ )를 예측한다(S60).

[0060] VR 영상에 대한 VIMS 점수( $S_{pred}$ )가 획득되면, 결합 클래스 특징맵(X(k))의 다수의 가중 클래스 특징맵( $\lambda_1 x_1(k) \sim \lambda_T x_T(k)$ )에서 추출되는 클래스 특징맵( $x_1(k) \sim x_T(k)$ ) 각각에 클래스 가중치(w(k))를 가중하여 모든 클래스(k)에 대해 누산함으로써, 각 프레임( $I_1 \sim I_T$ )에 대한 VIMS를 나타내는 프레임 VIMS( $a_i$ )를 수학적 식 4에 따라 획득한다(S70).

[0061] 본 발명에 따른 방법은 컴퓨터에서 실행시키기 위한 매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다. 여기서 컴퓨터 판독가능 매체는 컴퓨터에 의해 액세스 될 수 있는 임의의 가용 매체일 수 있고, 또한 컴퓨터 저장 매체를 모두 포함할 수 있다. 컴퓨터 저장 매체는 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 임의의 방법 또는 기술로 구현된 휘발성 및 비휘발성, 분리형 및 비분리형 매체를 모두 포함하며, ROM(판독 전용 메모리), RAM(랜덤 액세스 메모리), CD(컴팩트 디스크)-ROM, DVD(디지털 비디오 디스크)-ROM, 자기 테이프, 플로피 디스크, 광데이터 저장장치 등을 포함할 수 있다.

[0062] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다.

[0063] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

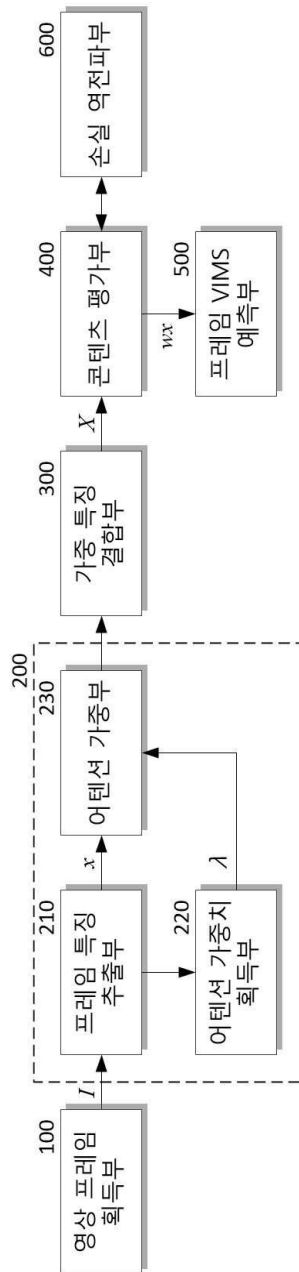
### 부호의 설명

[0064] 100: 영상 프레임 획득부    200: 어텐션 프레임 특징 추출부  
210: 프레임 특징 추출부    220: 어텐션 가중치 획득부  
230: 어텐션 가중부    300: 가중 특징 결합부  
400: 콘텐츠 평가부    500: 프레임 VIMS 예측부  
600: 손실 역전파부



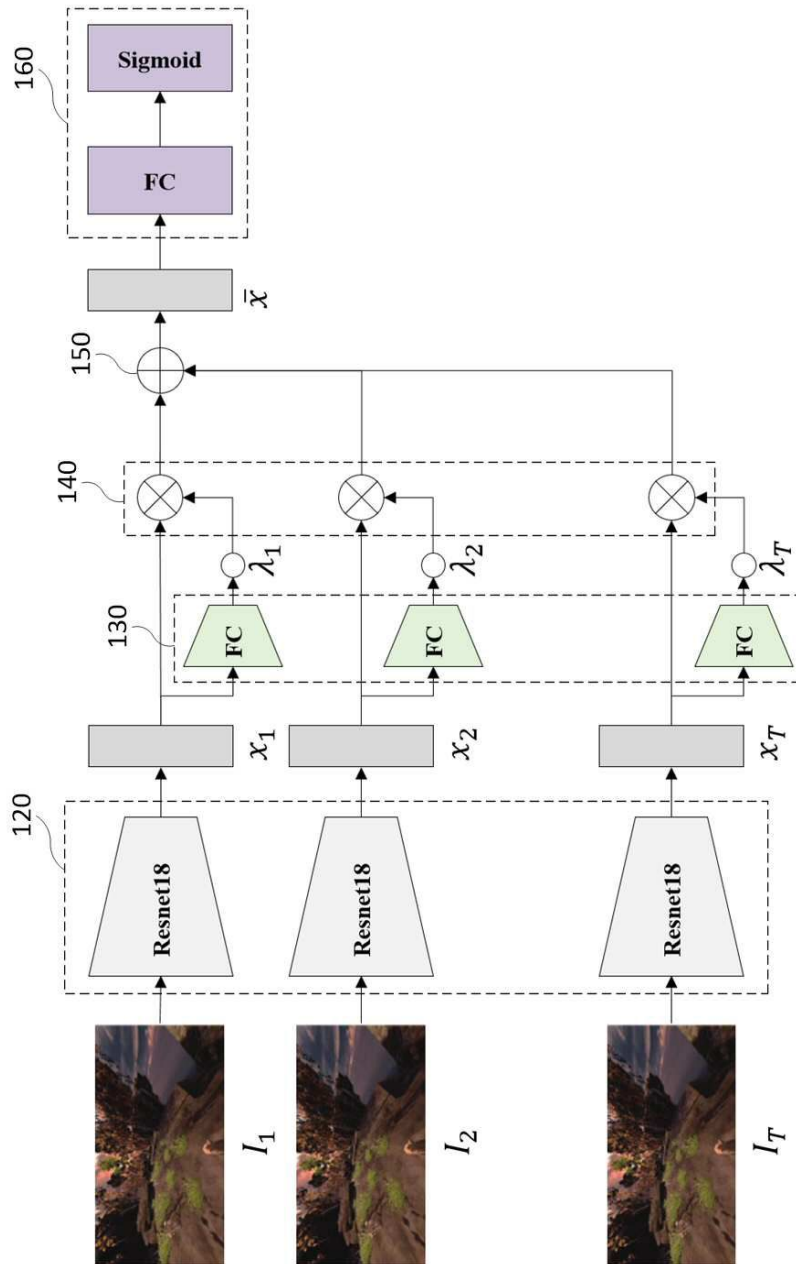
도면

도면1



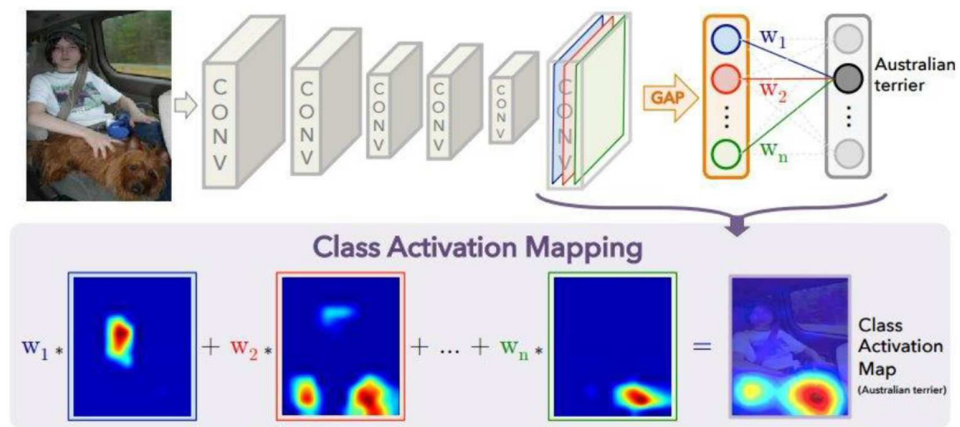


도면2

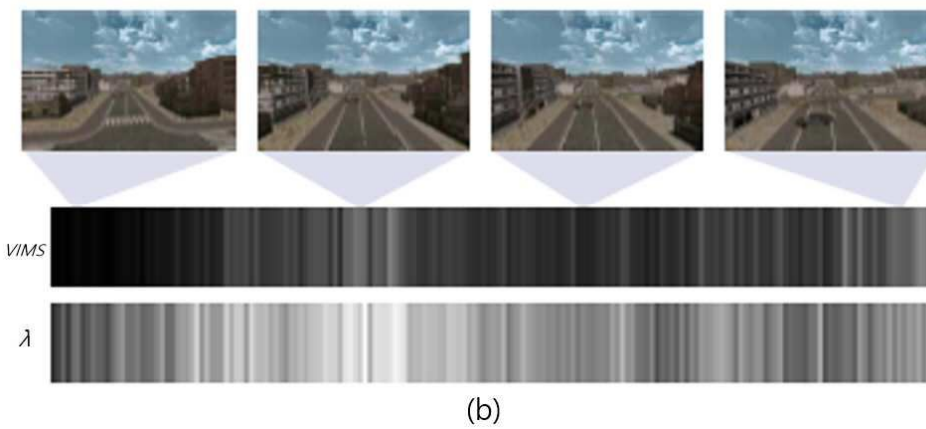
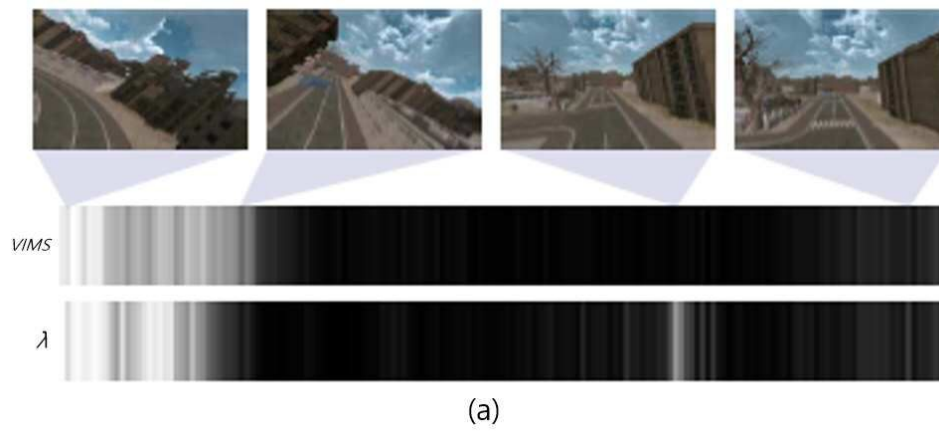




도면3



도면4





도면5

