



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0077700
(43) 공개일자 2024년06월03일

(51) 국제특허분류(Int. Cl.)
G06V 20/40 (2022.01) G06N 3/045 (2023.01)
G06V 10/774 (2022.01) G06V 10/82 (2022.01)
H04N 21/44 (2024.01) H04N 21/81 (2011.01)
H04N 21/8549 (2011.01)

(52) CPC특허분류
G06V 20/42 (2022.01)
G06N 3/045 (2023.01)

(21) 출원번호 10-2022-0159886
(22) 출원일자 2022년11월25일
심사청구일자 2022년11월25일

(71) 출원인
씨제이올리브네트웍스 주식회사
서울특별시 용산구 한강대로 366, 10층(동자동, 트윈시티)

(72) 발명자
권예하
서울특별시 용산구 한강대로 366, 10층(동자동, 트윈시티)

김지인
서울특별시 용산구 한강대로 366, 10층(동자동, 트윈시티)
(뒷면에 계속)

(74) 대리인
특허법인위솔

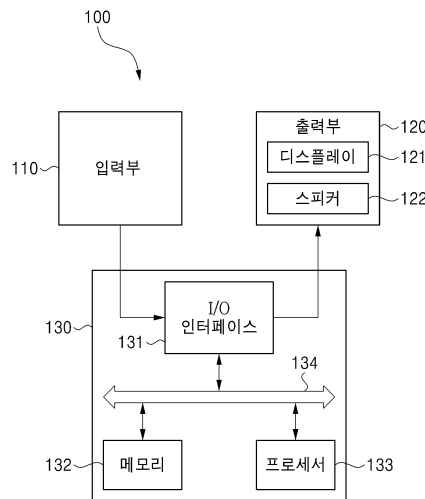
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 영상 편집 지원 방법 및 영상 편집 지원 장치

(57) 요약

영상 편집 지원 방법 및 영상 편집 지원 장치가 제공된다. 본 발명에 따른 영상 편집 지원 방법은, 라운드별 경기 시간이 규정된 종목의 중계 영상을 전처리하여, 상기 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 식별하는 단계, 상기 경기 진행 구간으로부터 복수의 비디오 클립을 추출하는 단계, 및 이벤트 검출 모델로 상기 복수의 비디오 클립을 분석하여, 상기 경기 진행 구간 내 적어도 하나의 유효 구간 - 상기 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응함 - 을 나타내는 편집 가이드 정보를 생성하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06V 10/774 (2023.08)

G06V 10/82 (2022.01)

G06V 20/47 (2022.01)

G06V 20/49 (2022.01)

H04N 21/44008 (2013.01)

H04N 21/8126 (2013.01)

H04N 21/8549 (2013.01)

(72) 발명자

유상기

서울특별시 중구 소월로2길 12 (남대문로5가)

서이안

서울특별시 용산구 한강대로 366, 10층(동자동, 트윈시티)

배종인

서울특별시 용산구 한강대로 366, 10층(동자동, 트윈시티)

이치훈

서울특별시 중구 소월로2길 12 (남대문로5가)

손중수

서울특별시 중구 소월로2길 12 (남대문로5가)

명세서

청구범위

청구항 1

영상 편집 지원 방법에 있어서,

라운드별 경기 시간이 규정된 종목의 중계 영상을 전처리하여, 상기 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 식별하는 단계;

상기 경기 진행 구간으로부터 복수의 비디오 클립을 추출하는 단계; 및

이벤트 검출 모델로 상기 복수의 비디오 클립을 분석하여, 상기 경기 진행 구간 내 적어도 하나의 유효 구간 - 상기 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응함 - 을 나타내는 편집 가이드 정보를 생성하는 단계;

를 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 2

제1항에 있어서,

상기 경기 진행 구간을 획득하는 단계는,

상기 중계 영상으로부터 적어도 하나의 참조 프레임을 샘플링하는 단계;

상기 참조 프레임을 기초로, 상기 중계 영상 내 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나의 추정치를 나타내는 참조 시각 정보를 생성하는 단계; 및

상기 참조 시각 정보를 기초로, 상기 중계 영상으로부터 상기 경기 비진행 구간을 제거하는 단계;

를 포함하는 것을 특징으로 하는, 영상 편집 지원 방법.

청구항 3

제2항에 있어서,

상기 참조 시각 정보를 생성하는 단계는,

상기 참조 프레임에서 중계 상황판을 추출하는 단계;

상기 중계 상황판으로부터, 적어도 한 라운드의 시작 시각으로부터의 경과 시간을 결정하는 단계; 및

상기 경과 시간을 기초로, 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나를 추정하는 단계;

를 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 4

제1항에 있어서,

상기 이벤트 검출 모델은,

동일 종목의 복수의 다른 중계 영상으로부터 추출되어 상기 복수의 이벤트 타입 중 어느 하나로 라벨링된 복수

의 하이라이트 영상을 포함하는 학습 데이터 세트에 의해 학습된 것을 특징으로 하는,
영상 편집 지원 방법.

청구항 5

제1항에 있어서,

상기 복수의 비디오 클립의 인접한 두 비디오 클립 중, 선행 비디오 클립의 종료 시각은 후행 비디오 클립의 시작 시각의 후인 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 6

제1항에 있어서,

상기 편집 가이드 정보를 생성하는 단계는,

상기 이벤트 검출 모델의 제1 딥러닝 모델을 이용하여, 상기 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터로 변환하는 단계; 및

상기 이벤트 검출 모델의 제2 딥러닝 모델을 이용한 다음의 동작들:

상기 복수의 특징 벡터 각각을 복수의 클러스터 - 각 클러스터는 상기 복수의 이벤트 타입 중 적어도 하나를 적어도 부분적으로 표현함 - 중 어느 하나에 맵핑하는 단계;

상기 복수의 특징 벡터를 시간 순으로 그룹핑하여, 복수의 벡터 그룹을 생성하는 단계; 및

상기 복수의 벡터 그룹과 상기 복수의 이벤트 타입 중 적어도 하나 간의 대응 관계로부터 상기 경기 진행 구간 내에서 상기 유효 구간을 식별하는 단계;

를 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 7

제6항에 있어서,

상기 중계 영상으로부터의 하이라이트 영상의 추출에 이용되는 복수의 필터링 항목 중 적어도 하나에 대한 설정 정보를 수신하는 단계;

를 더 포함하되,

상기 제2 딥러닝 모델은, 상기 설정 정보에 따라 동작하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 8

제7항에 있어서,

상기 복수의 필터링 항목은,

이벤트 타입, 이벤트 유사도 및 이벤트 중요도를 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 9

제1항에 있어서,

상기 편집 가이드 정보가 제시되는 영상 편집 인터페이스를 출력하는 단계;

를 더 포함하고,

상기 영상 편집 인터페이스는,

상기 중계 영상 내 상기 유효 구간의 위치 또는 범위를 지시하는 인디케이터를 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 10

제1항에 있어서,

사용자로부터 희망 시간이 지정된 자동 편집 요청이 수신된 것에 응답하여, 상기 적어도 하나의 유효 구간을 가공하여, 상기 희망 시간과 동일 시간 길이를 갖는 추천 하이라이트 영상을 생성하는 단계;

를 더 포함하는 것을 특징으로 하는,

영상 편집 지원 방법.

청구항 11

영상 편집 지원 장치에 있어서,

영상 편집 지원 방법을 실행시키기 위한 명령어들이 기록된 컴퓨터 프로그램 및 라운드별 경기 시간이 규정된 종목의 중계 영상이 저장되는 메모리; 및

상기 메모리에 동작 가능하게 결합되는 프로세서;

를 포함하되,

상기 컴퓨터 프로그램이 상기 프로세서에 의해 실행되는 경우,

상기 프로세서는,

상기 중계 영상을 전처리하여, 상기 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 획득하고,

상기 경기 진행 구간으로부터 복수의 비디오 클립을 추출하고,

이벤트 검출 모델로 상기 복수의 비디오 클립을 분석하여, 상기 경기 진행 구간 내 적어도 하나의 유효 구간 - 상기 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응함 - 을 나타내는 편집 가이드 정보를 생성하도록 구성되는 것을 특징으로 하는,

영상 편집 지원 장치.

청구항 12

제11항에 있어서,

상기 프로세서는,

상기 경기 진행 구간을 식별하기 위해,

상기 중계 영상으로부터 적어도 하나의 참조 프레임을 샘플링하고,

상기 참조 프레임을 기초로, 상기 중계 영상 내 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나의 추정치를 나타내는 참조 시각 정보를 생성하고,

상기 참조 시각 정보를 기초로, 상기 중계 영상으로부터 상기 경기 비진행 구간을 제거하도록 구성되는 것을 특징으로 하는,

영상 편집 지원 장치.

청구항 13

제12항에 있어서,

상기 프로세서는,

상기 참조 시각 정보를 생성하기 위해,

상기 참조 프레임에서 중계 상황판을 추출하고,

상기 중계 상황판으로부터, 적어도 한 라운드의 시작 시각으로부터의 경과 시간을 결정하고,

상기 경과 시간을 기초로, 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나를 추정하도록 구성되는 것을 특징으로 하는,

영상 편집 지원 장치.

청구항 14

제11항에 있어서,

상기 프로세서는,

상기 편집 가이드 정보를 생성하기 위해,

상기 이벤트 검출 모델의 제1 딥러닝 모델을 이용하여, 상기 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터로 변환하고,

상기 이벤트 검출 모델의 제2 딥러닝 모델을 이용하여, 상기 복수의 특징 벡터 각각을 복수의 클러스터 - 각 클러스터는 상기 복수의 이벤트 타입 중 적어도 하나를 적어도 부분적으로 표현함 - 중 어느 하나에 맵핑하고, 상기 복수의 특징 벡터를 시간 순으로 그룹핑하여, 복수의 벡터 그룹을 생성하고, 상기 복수의 벡터 그룹과 상기 복수의 이벤트 타입 중 적어도 하나 간의 대응 관계로부터 상기 경기 진행 구간 내에서 상기 유효 구간을 식별하도록 구성되는 것을 특징으로 하는,

영상 편집 지원 장치.

청구항 15

제14항에 있어서,

상기 프로세서는,

상기 중계 영상으로부터의 하이라이트 영상의 추출에 이용되는 복수의 필터링 항목 중 적어도 하나에 대한 설정 정보를 수신 시,

상기 설정 정보에 따라 상기 제2 딥러닝 모델을 동작시키도록 구성되는 것을 특징으로 하는,

영상 편집 지원 장치.

발명의 설명

기술 분야

[0001] 본 발명은 영상 편집 기술에 관한 것으로서, 보다 상세하게는, 딥러닝 기반으로 스포츠 종목의 경기 중계 영상으로부터 편집자로서의 사용자가 희망하는 하이라이트 씬을 자동적으로 선별하여 사용자의 영상 편집 작업을 지원하는 방법 및 장치에 관한 것이다.

배경 기술

[0003] 국내 OTT(Over The Top) 이용률은 2019년 약 41%에서 2021년 약 82%로 가파른 성장을 이루었으며, 이에 따라 수많은 영상 콘텐츠를 언제 어디서나 시청 가능한 환경이 조성되었으나, 상대적으로 스포츠 관련 영상 콘텐츠의 소비가 낮은 측면이 강하다.

[0004] 특히, 팬층이 두터운 몇몇 종목이나 리그 외에는 OTT 플랫폼에서 중계권을 확보하는 노력 대비 중계 영상 콘텐츠의 원활한 소비가 이루어지지 않고, 영화 및 드라마와 같이 보다 대중적 콘텐츠는 그 자체가 높은 수준의 회원 흡입력을 갖는 것과는 대조적으로 축구, 농구 등과 같은 종목의 스포츠 콘텐츠는 OTT 가입 유도용 미끼 상품 정도의 역할에 머무르고 있는 실정이다.

[0005] 영상 콘텐츠의 가장 큰 소비층인 MZ세대는 장시간 영상을 드물게 보는 대신 짧은 영상을 꾸준히 여러 번 시청하는 성향이 두드러지며, 스포츠 경기가 실시간으로 중계 중인 상황이 아니라면, 전체 녹화 영상으로부터 임팩트 있는 주요 장면들로 이루어진 하이라이트 영상에 대한 선호도 또한 MZ세대에 국한되지 않고 점차 증대되고 있다.

[0006] 한편, 중계 영상으로부터 하이라이트 영상의 추출은 사람 즉, 편집자에 의해 수작업으로 이루어 지는 것이 대부분이며, 이러한 편집 방식은 기본적으로 긴 작업 시간이 소요될 뿐만 아니라, 편집자의 주관적 판단, 실수 등이 개입됨으로 인해 일반 콘텐츠 소비자의 기대치보다 낮은 품질의 하이라이트 영상이 제작되는 단점이 있다.

[0007] 진술된 단점을 해소하기 위한 목적으로 영상 편집을 도와주는 소프트웨어 기능들이 몇몇 제시되고 있다. 예컨대, 오디오 데이터와 이미지 데이터가 상호 매칭되도록 음악/음성을 지능적으로 재배열하는 리믹스(Remix) 기능, 고속의 캡션 생성이 가능한 스피치-텍스트 변환(Speech to Text) 기술, GPU 가속을 통한 HDR 콘텐츠 내보내기 기술이 있다.

[0008] 그러나, 종목별로 어느 정도 차이가 있으나 스포츠 경기는 격렬한 움직임을 수반하는 것이 대부분이며, 다수의 참여 선수가 존재하고 여러 대의 중계용 카메라에 따른 앵글 변화가 수시로 발생하므로, 경기 진행 중의 순간적 장면의 특성을 정밀하게 잡아내기가 어렵다.

[0009] 또한, 축구로 예를 들자면 골(goal)과 같이 승패에 결정적 영향을 주는 주요 이벤트는 어느 정도의 시간에 걸친 플레이를 통해 만들어지는데, 동종 이벤트라고 할지라도 하나의 경기 내에서 이벤트별 지속되는 시간 길이가 달라 이벤트별 영상 구간을 일률적으로 추출해내기 어려운 측면이 존재한다.

발명의 내용

해결하려는 과제

[0011] 본 발명은, 상기와 같은 문제점을 해결하기 위해 안출된 것으로서, 스포츠 경기의 중계 영상으로부터 경기 외적인 영상 구간(경기 비진행 구간)을 자동 제거하여 경기가 실제 진행된 영상 구간(경기 진행 구간)을 추출하는 방법 및 장치를 제공하는 것을 목적으로 한다.

[0012] 또한, 본 발명은 중계 영상으로부터 추출된 경기 진행 구간 내에서 특정의 이벤트가 발생된 유효 구간을 식별하여 편집자(사용자)에게 편집 가이드 정보로서 제공하는 방법 및 장치를 제공하는 것을 목적으로 한다.

[0013] 본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있으며, 본 발명의 실시예에 의해 보다 분명하게 알게 될 것이다. 또한, 본 발명의 목적 및 장점들은 특허청구범위에 나타난 수단 및 그 조합에 의해 실현

될 수 있음을 쉽게 알 수 있을 것이다.

과제의 해결 수단

- [0015] 본 발명의 일 측면에 따른 영상 편집 지원 방법은, 라운드별 경기 시간이 규정된 종목의 중계 영상을 전처리하여, 상기 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 식별하는 단계, 상기 경기 진행 구간으로부터 복수의 비디오 클립을 추출하는 단계, 및 이벤트 검출 모델로 상기 복수의 비디오 클립을 분석하여, 상기 경기 진행 구간 내 적어도 하나의 유효 구간 - 상기 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응함 - 을 나타내는 편집 가이드 정보를 생성하는 단계를 포함한다.
- [0016] 상기 경기 진행 구간을 획득하는 단계는, 상기 중계 영상으로부터 적어도 하나의 참조 프레임을 샘플링하는 단계, 상기 참조 프레임을 기초로, 상기 중계 영상 내 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나의 추정치를 나타내는 참조 시각 정보를 생성하는 단계 및 상기 참조 시각 정보를 기초로, 상기 중계 영상으로부터 상기 경기 비진행 구간을 제거하는 단계를 포함할 수 있다.
- [0017] 상기 참조 시각 정보를 생성하는 단계는, 상기 참조 프레임에서 중계 상황판을 추출하는 단계, 상기 중계 상황판으로부터, 적어도 한 라운드의 시작 시각으로부터의 경과 시간을 결정하는 단계 및 상기 경과 시간을 기초로, 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나를 추정하는 단계를 포함할 수 있다.
- [0018] 상기 이벤트 검출 모델은, 동일 종목의 복수의 다른 중계 영상으로부터 추출되어 상기 복수의 이벤트 타입 중 어느 하나로 라벨링된 복수의 하이라이트 영상을 포함하는 학습 데이터 세트에 의해 학습된 것일 수 있다.
- [0019] 상기 복수의 비디오 클립의 인접한 두 비디오 클립 중, 선행 비디오 클립의 종료 시각은 후행 비디오 클립의 시작 시각의 후일 수 있다.
- [0020] 상기 편집 가이드 정보를 생성하는 단계는, 상기 이벤트 검출 모델의 제1 딥러닝 모델을 이용하여, 상기 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터로 변환하는 단계 및 상기 이벤트 검출 모델의 제2 딥러닝 모델을 이용한 다음의 동작들: 상기 복수의 특징 벡터 각각을 복수의 클러스터 - 각 클러스터는 상기 복수의 이벤트 타입 중 적어도 하나를 적어도 부분적으로 표현함 - 중 어느 하나에 맵핑하는 단계, 상기 복수의 특징 벡터를 시간 순으로 그룹핑하여, 복수의 벡터 그룹을 생성하는 단계 및 상기 복수의 벡터 그룹과 상기 복수의 이벤트 타입 중 적어도 하나 간의 대응 관계로부터 상기 경기 진행 구간 내에서 상기 유효 구간을 식별하는 단계를 포함할 수 있다.
- [0021] 상기 영상 편집 지원 방법은, 상기 중계 영상으로부터의 하이라이트 영상의 추출에 이용되는 복수의 필터링 항목 중 적어도 하나에 대한 설정 정보를 수신하는 단계를 더 포함할 수 있다. 이때, 상기 제2 딥러닝 모델은, 상기 설정 정보에 따라 동작할 수 있다.
- [0022] 상기 복수의 필터링 항목은, 이벤트 타입, 이벤트 유사도 및 이벤트 중요도를 포함할 수 있다.
- [0023] 상기 영상 편집 지원 방법은, 상기 편집 가이드 정보가 제시되는 영상 편집 인터페이스를 출력하는 단계를 더 포함할 수 있다. 상기 영상 편집 인터페이스는, 상기 중계 영상 내 상기 유효 구간의 위치 또는 범위를 지시하는 인디케이터를 포함할 수 있다.
- [0024] 상기 영상 편집 지원 방법은, 사용자로부터 희망 시간이 지정된 자동 편집 요청이 수신된 것에 응답하여, 상기 적어도 하나의 유효 구간을 가공하여, 상기 희망 시간과 동일 시간 길이를 갖는 추천 하이라이트 영상을 생성하는 단계를 더 포함할 수 있다.
- [0025] 본 발명의 다른 측면에 따른 영상 편집 지원 장치는, 영상 편집 지원 방법을 실행시키기 위한 명령어들이 기록된 컴퓨터 프로그램 및 라운드별 경기 시간이 규정된 종목의 중계 영상이 저장되는 메모리 및 상기 메모리에 동작 가능하게 결합되는 프로세서를 포함한다. 상기 컴퓨터 프로그램이 상기 프로세서에 의해 실행되는 경우, 상기 프로세서는, 상기 중계 영상을 전처리하여, 상기 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 획득하고, 상기 경기 진행 구간으로부터 복수의 비디오 클립을 추출하고, 이벤트 검출 모델로 상기 복수의 비디오 클립을 분석하여, 상기 경기 진행 구간 내 적어도 하나의 유효 구간 - 상기 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응함 - 을 나타내는 편집 가이드 정보를 생성하도록 구성될 수 있다.
- [0026] 상기 프로세서는, 상기 경기 진행 구간을 식별하기 위해, 상기 중계 영상으로부터 적어도 하나의 참조 프레임을 샘플링하고, 상기 참조 프레임을 기초로, 상기 중계 영상 내 적어도 한 라운드의 시작 시각 및 종료 시각 중 적

어도 하나의 추정치를 나타내는 참조 시각 정보를 생성하고, 상기 참조 시각 정보를 기초로, 상기 중계 영상으로부터 상기 경기 비진행 구간을 제거하도록 구성될 수 있다.

[0027] 상기 프로세서는, 상기 참조 시각 정보를 생성하기 위해, 상기 참조 프레임에서 중계 상황판을 추출하고, 상기 중계 상황판으로부터, 적어도 한 라운드의 시작 시각으로부터의 경과 시간을 결정하고, 상기 경과 시간을 기초로, 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나를 추정하도록 구성될 수 있다.

[0028] 상기 프로세서는, 상기 편집 가이드 정보를 생성하기 위해, 상기 이벤트 검출 모델의 제1 딥러닝 모델을 이용하여, 상기 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터로 변환하고, 상기 이벤트 검출 모델의 제2 딥러닝 모델을 이용하여, 상기 복수의 특징 벡터 각각을 복수의 클러스터 - 각 클러스터는 상기 복수의 이벤트 타입 중 적어도 하나를 적어도 부분적으로 표현함 - 중 어느 하나에 맵핑하고, 상기 복수의 특징 벡터를 시간 순으로 그룹핑하여, 복수의 벡터 그룹을 생성하고, 상기 복수의 벡터 그룹과 상기 복수의 이벤트 타입 중 적어도 하나 간의 대응 관계로부터 상기 경기 진행 구간 내에서 상기 유효 구간을 식별하도록 구성될 수 있다.

[0029] 상기 프로세서는, 상기 중계 영상으로부터의 하이라이트 영상의 추출에 이용되는 복수의 필터링 항목 중 적어도 하나에 대한 설정 정보를 수신 시, 상기 설정 정보에 따라 상기 제2 딥러닝 모델을 동작시키도록 구성될 수 있다.

발명의 효과

[0031] 본 발명의 실시예들 중 적어도 하나에 의하면, 스포츠 경기의 중계 영상으로부터 경기 외적인 영상 구간(경기 비진행 구간)을 자동 제거하여 경기가 실제 진행된 영상 구간(경기 진행 구간)을 추출할 수 있다. 이에 따라, 스포츠 중계 영상의 전체에 대해 하이라이트 구간을 탐색하는 방식에 비하여 하드웨어 및 소프트웨어적인 컴퓨팅 자원의 낭비를 크게 절감함과 아울러 편집자의 수고로움을 크게 덜어줄 수 있다.

[0032] 또한, 본 발명의 실시예들 중 적어도 하나에 의하면, 중계 영상으로부터 추출된 경기 진행 구간 내에서 특정의 이벤트가 발생된 유효 구간을 식별하여 편집자(사용자)에게 편집 가이드 정보로서 제공할 수 있다. 이에 따라, 편집자는 자신이 희망하는 타입의 이벤트가 발생된 영상 구간의 위치 또는 범위를 직관적으로 파악하여 편집 영상물에 대한 완성본을 제작하기 까지에 소요되는 시간을 단축하고, 경기 주요 장면과의 관련성이 크게 떨어지는 구간에서 하이라이트 영상이 오추출될 가능성을 줄일 수 있다.

[0033] 본 발명의 효과들은 이상에서 언급한 효과들로 제한되지 않으며, 언급되지 않은 또 다른 효과들은 청구범위의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

[0035] 본 명세서에 첨부되는 다음의 도면들은 본 발명의 바람직한 실시예를 예시하는 것이며, 후술되는 발명의 상세한 설명과 함께 본 발명의 기술사상을 더욱 이해시키는 역할을 하는 것이므로, 본 발명은 그러한 도면에 기재된 사항에만 한정되어 해석되어서는 아니 된다.

도 1은 본 발명에 따른 영상 편집 지원 장치의 구성을 예시적으로 나타낸 도면이다.

도 2는 도 1에 도시된 영상 편집 지원 장치에 의해 실행되는, 본 발명의 일 실시예에 따른 영상 편집 지원 방법을 예시적으로 설명하는 데에 참조되는 순서도이다.

도 3 내지 도 5는 도 2에 도시된 단계 S220의 하위 루틴들의 예시적인 실행 과정을 설명하는 데에 참조되는 도면이다.

도 6은 경기 진행 구간으로부터 추출된 복수의 비디오 클립이 복수의 특징 벡터로 변환되는 예시적인 과정을 설명하는 데에 참조되는 도면이다.

도 7은 종목별 사전 정의되는 복수의 이벤트 타입과 이벤트 타입별 검출 기준을 테이블 형식으로 예시하는 도면이다.

도 8은 도 7에 따른 이벤트 타입별 검출 기준에 의해 라벨링된 학습 데이터의 일부분을 예시적으로 보여주는 도면이다.

도 9는 단계 S240의 하위 루틴들의 예시적인 실행 과정을 설명하는 데에 참조되는 순서도이다.

도 10은 복수의 클러스터와 복수의 특징 벡터 간의 대응 관계를 예시적으로 보여주는 도면이다.

도 11은 도 10에 따른 대응 관계 데이터로부터 경기 진행 영상의 하위 영상 구간별로 진행되는 이벤트 검출의 과정을 설명하는 데에 참조되는 도면이다.

도 12는 도 1에 도시된 영상 편집 지원 장치에 의해 실행되는, 본 발명의 다른 실시예에 따른 영상 편집 지원 방법을 예시적으로 설명하는 데에 참조되는 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0036] 본 발명의 목적과 기술적 구성 및 그에 따른 작용 효과에 관한 자세한 사항은 본 발명의 명세서에 첨부된 도면에 의거한 이하의 상세한 설명에 의해 보다 명확하게 이해될 것이다. 첨부된 도면을 참조하여 본 발명에 따른 실시예를 상세하게 설명한다.
- [0037] 본 명세서에서 개시되는 실시예들은 본 발명의 범위를 한정하는 것으로 해석되거나 이용되지 않아야 할 것이다. 이 분야의 통상의 기술자에게 본 명세서의 실시예를 포함한 설명은 다양한 응용을 갖는다는 것이 당연하다. 따라서, 본 발명의 상세한 설명에 기재된 임의의 실시예들은 본 발명을 보다 잘 설명하기 위한 예시적인 것이며 본 발명의 범위가 실시예들로 한정되는 것을 의도하지 않는다.
- [0038] 도면에 표시되고 아래에 설명되는 기능 블록들은 가능한 구현의 예들일 뿐이다. 다른 구현들에서는 상세한 설명의 사상 및 범위를 벗어나지 않는 범위에서 다른 기능 블록들이 사용될 수 있다. 또한, 본 발명의 하나 이상의 기능 블록이 개별 블록들로 표시되지만, 본 발명의 기능 블록들 중 하나 이상은 동일 기능을 실행하는 다양한 하드웨어 및 소프트웨어 구성들의 조합일 수 있다.
- [0039] 제1, 제2 등과 같이 서수를 포함하는 용어들은, 다양한 구성요소들 중 어느 하나를 나머지와 구별하는 목적으로 사용되는 것이고, 그러한 용어들에 의해 구성요소들을 한정하기 위해 사용되는 것은 아니다.
- [0040] 또한, 어떤 구성요소들을 포함한다는 표현은 "개방형"의 표현으로서 해당 구성요소들이 존재하는 것을 단순히 지칭할 뿐이며, 추가적인 구성요소들을 배제하는 것으로 이해되어서는 안 된다.
- [0041] 나아가 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급될 때에는, 그 다른 구성요소에 직접적으로 연결 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 한다.
- [0042] 도 1은 본 발명에 따른 영상 편집 지원 장치(100)의 구성을 예시적으로 나타낸 도면이다.
- [0043] 도 1을 참조하면, 영상 편집 지원 장치(100)는, 입력부(110), 출력부(120) 및 제어부(130)를 포함한다. 영상 편집 지원 장치(100)는, 데스크 탑, 노트북, 스마트폰, 태블릿 PC 등의 형태로 구현될 수 있다.
- [0044] 입력부(110)는, 영상 편집 지원 장치(100)를 통해 임의의 스포츠 중계 영상에 대한 편집을 통해 하이라이트 영상 콘텐츠를 제작하기를 희망하는 사용자(편집자)로부터의 일련의 입력(편집 관련 기능의 실행을 요청하는 액션)을 받아들여, 각각의 입력에 연관된 기능의 실행을 요청하는 신호를 제어부(130)에 전달한다. 입력부(110)는, 예컨대 키보드, 마우스, 터치 패널 등과 같은 공지의 입력 수단 중 어느 하나 또는 둘 이상의 조합일 수 있다.
- [0045] 출력부(120)는, 디스플레이(121) 및 스피커(122)를 포함한다. 디스플레이(121)는, 제어부(130)로부터의 제어 명령에 따라 임의의 스포츠 중계 영상의 편집 툴을 제공하는 영상 편집 인터페이스를 표시한다. 스피커(122)는, 제어부(130)로부터의 제어 명령에 따라 디스플레이(121)에 표시되는 그래픽 정보에 시간 동기화된 오디오 데이터에 대응되는 청각 피드백을 발생시킬 수 있다.
- [0046] 제어부(130)는, 입출력(I/O) 인터페이스(131), 메모리(132) 및 프로세서(133)와, 이들을 통신 가능하도록 접속하는 데이터 버스(134)를 포함한다.
- [0047] 입출력 인터페이스(131)는, 입력부(110)로부터의 사용자 요청을 데이터 버스(134)를 통해 프로세서(133)에 전달하고, 프로세서(133)가 사용자 요청을 처리하여 생성된 출력 신호를 출력부(120)에 전달한다.
- [0048] 메모리(132)는, 임의의 스포츠 중계 영상과, 그로부터 편집자가 희망하는 영상 구간을 포함하는 하이라이트 영상을 제작할 수 있도록 지원하는 데에 요구되는 학습 모델, 컴퓨터 프로그램 및/또는 데이터를 기록한다. 메모리(132)는, 임의의 스포츠 중계 영상과, 그로부터 편집자가 희망하는 영상 구간을 포함하는 하이라이트 영상을 제작할 수 있도록 지원하는 데에 요구되는 학습 모델, 컴퓨터 프로그램 및/또는 데이터를 기록한다. 메모리(132)는, 임의의 스포츠 중계 영상과, 그로부터 편집자가 희망하는 영상 구간을 포함하는 하이라이트 영상을 제작할 수 있도록 지원하는 데에 요구되는 학습 모델, 컴퓨터 프로그램 및/또는 데이터를 기록한다.

리(132)는 하드웨어적으로 플래시 메모리 타입(Flash memory type), 하드디스크 타입(hard disk type), SSD 타입(Solid State Disk type), HDD 타입(Silicon Disk Drive type), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 램(Random Access Memory; RAM), SRAM(Static Random Access Memory), 롬(Read-Only Memory; ROM), EEPROM(Electrically Erasable Programmable Read-Only Memory), PROM(Programmable Read-Only Memory) 중 적어도 하나 또는 둘 이상의 타입의 저장매체를 포함할 수 있다. 메모리(132)는 본 발명에 따른 영상 편집 지원 방법을 실행하는 명령어가 기록된 컴퓨터 프로그램이 저장된 저장매체를 포함할 수 있다.

[0049] 프로세서(133)는, 입출력 인터페이스(131) 및 메모리(132)에 동작 가능하게 결합되어, 영상 편집 지원 장치(100)의 전체적인 동작을 제어한다. 프로세서(133)는 하드웨어적으로, ASICs(application specific integrated circuits), DSPs(digital signal processors), DSPDs(digital signal processing devices), PLDs(programmable logic devices), FPGAs(field programmable gate arrays), 마이크로 프로세서(microprocessors), 기타 기능 수행을 위한 전기적 유닛 중 적어도 하나를 포함할 수 있다.

[0050] 도 2는 도 1에 도시된 영상 편집 지원 장치에 의해 실행되는, 본 발명의 일 실시예에 따른 영상 편집 지원 방법을 예시적으로 설명하는 데에 참조되는 순서도이다.

[0051] 도 2를 참조하면, 단계 S210에서, 프로세서(133)는, 입력부(110)를 통해 수신되는 편집자로부터의 요청에 응답하여, 메모리(132)에 저장된 중계 영상을 편집 대상으로 설정한다. 본 발명에 따른 편집 지원의 대상이 되는 종목의 스포츠(경기)는, 라운드별 경기 시간이 규정된 종목이다. 여기서, 라운드란, 스포츠의 전체 진행 구간을 순서 상으로 구분해 놓은 것으로서, 종목에 따라 그 지칭이 다를 수 있다.

[0052] 일 예로, 축구의 전반전과 후반전, 농구의 1 내지 4 쿼터가 본 발명에 따른 '라운드'에 해당한다. 천연하자면, 축구는 각각의 라운드가 45분으로 규정되어 있고, 농구는 각각의 쿼터가 10분 또는 12분으로 규정되어 있다. 또한, 2 이상의 라운드를 포함하는 종목의 경우, 인접한 두 라운드 간의 휴식 시간 또한 규정될 수 있는데, 이러한 라운드별 경기 시간과 라운드들 사이의 휴식 시간은 단계 S210에서 선택된 중계 영상의 종목에 관련된 경기 규칙 정보로서 메모리(132)에 미리 기록되어 있을 수 있다.

[0053] 단계 S220에서, 프로세서(133)는, 중계 영상을 전처리하여, 중계 영상으로부터 경기 비진행 구간이 제거된 경기 진행 구간을 식별한다.

[0054] 구체적으로, 임의의 종목의 경기에 대한 중계 영상은, 해당 경기에 관련된 녹화본일 수 있는데, 이러한 중계 영상은 해당 경기가 실제로 진행되는 중에 촬영되는 영상뿐만 아니라 경기 시작 전부터 경기 종료 후의 부분도 포함한다.

[0055] 축구 경기를 예로 들자면, 전반전의 시작 전에는 해설진 소개와 양팀 선수 소개가 이루어지고, 전반전과 후반전 사이의 휴식 시간에는 전반전의 경기 내용에 대한 해설이나 광고가 진행되며, 후반전 종료 후에는 경기 결과나 향후 일정(토너먼트 대진표) 등에 대한 요약 코멘트 등이 진행된다. 즉, 단계 S220에 의해, 경기가 실제로 진행되는 시간 범위 외의 영상 부분인 경기 비진행 구간이 중계 영상으로부터 제거되는 것이다. 중계 영상에서 경기 비진행 구간을 제거한다는 것은, 경기 비진행 구간을 삭제하는 것을 의미할 수 있으나, 경기 비진행 구간 외의 나머지 구간인 경기 진행 구간을 식별한다는 의미까지도 포괄할 수 있다. 단계 S220의 상세 과정에 대해서는 도 3 내지 도 5를 참조하여 이하에서 보다 구체적으로 설명하기로 한다.

[0056] 단계 S230에서, 프로세서(133)는, 경기 진행 구간으로부터 복수의 비디오 클립(도 6 참조)을 추출한다. 각 비디오 클립은 경기 진행 구간 내 하나의 영상 조각으로서, 단계 S230에서 추출되는 복수의 비디오 클립은 모두 서로 동일한 시간 길이를 갖을 수 있고, 대안적으로 복수의 비디오 클립 중 적어도 하나는 나머지 비디오 클립과는 시간 길이가 상이할 수 있다.

[0057] 복수의 비디오 클립에 있어서, 서로 인접한 비디오 클립은 하나의 클립 쌍을 이루며, 임의의 클립 쌍의 선행 비디오 클립의 종료 시각은 후행 비디오 클립의 시작 시각과 동일할 수 있다.

[0058] 또는, 임의의 클립 쌍의 선행 비디오 클립의 종료 시각은 후행 비디오 클립의 시작 시각에 앞설 수 있다. 즉, 임의의 클립 쌍의 두 비디오 클립은 일정 시간 범위에서 서로 중첩되는 부분을 가질 수 있다. 이 경우, 임의의 클립 쌍의 두 비디오 클립이 경기 진행 구간의 서로 별개의 부분을 표현하는 것이 아니라, 서로 공통된 부분을 중복 표현하게 되므로, 특정 비디오 클립에는 그에 선행 또는 후행하는 다른 비디오 클립과의 관련성이 담기게 된다.

[0059] 프로세서(133)는, 경기 진행 구간의 전체적인 시간 길이에 따라, 그로부터 추출될 비디오 클립의 시간 길이를

가변 조정할 수 있다. 예컨대, 비디오 클립의 시간 길이는 경기 진행 구간의 시간 길이의 1/10000일 수 있다. 물론, 경기 진행 구간의 시간 길이에 무관하게 비디오 클립의 시간 길이는 고정된 값으로 설정되어 있을 수 있다.

- [0060] 단계 S240에서, 프로세서(133)는, 이벤트 검출 모델로 복수의 비디오 클립을 분석하여, 경기 진행 구간 내 적어도 하나의 유효 구간을 나타내는 편집 가이드 정보를 생성한다. 편집 가이드 정보는 메모리(132)에 저장될 수 있다. 여기서, 유효 구간은 복수의 이벤트 타입 중 적어도 하나에 대응하는 경기 진행 구간의 일 부분일 수 있다. 복수의 이벤트 타입은 중계 영상의 경기 종목에 따라 사전 정의되어 있을 수 있다.
- [0061] 프로세서(133)는, 기능적으로 학습부 및 추론부를 포함한다. 학습부는 학습 데이터 세트를 이용하여 이벤트 검출 모델을 사전 학습시키는 구성이고, 추론부는 학습부에 의해 학습 완료된 이벤트 검출 모델을 이용하여 경기 진행 구간 내 특정 이벤트가 발생된 구간을 검출하는 구성이다. 이벤트 검출 모델에 의한 경기 진행 구간으로부터의 이벤트 검출 과정에 대해서는 도 6 내지 도 11을 참조하여 상세히 후술하기로 한다.
- [0062] 단계 S250에서, 프로세서(133)는, 편집 가이드 정보가 제시되는 영상 편집 인터페이스를 출력한다. 즉, 디스플레이(121)는 프로세서(133)의 명령에 응답하여 영상 편집 인터페이스를 표시한다. 영상 편집 인터페이스는, 중계 영상에 대한 일종의 편집 툴로서, 중계 영상 내 유효 구간별 위치, 범위, 대응되는 이벤트 타입 등을 지시하는 적어도 하나의 그래픽 인디케이터를 포함할 수 있다.
- [0063] 한편, 프로세서(133)는, 단계 S240 이전에, 입력부(110)를 통해, 중계 영상으로부터의 하이라이트 영상의 추출에 이용되는 복수의 필터링 항목 중 적어도 하나에 대한 설정 정보를 수신할 수 있다.
- [0064] 복수의 필터링 항목은, 이벤트 타입, 이벤트 유사도 및 이벤트 중요도 중 적어도 하나를 포함하고, 단계 S1210에서 수신되는 설정 정보는 필터링 항목별 설정값을 나타낼 수 있다. 가령, 중계 영상의 경기 종목이 축구인 경우, 설정 정보의 이벤트 타입은 축구와 관련된 여러 이벤트 타입 중 편집자가 희망하는 몇몇 이벤트 타입(예, 골, 옐로우 카드)만을 지정하는 것일 수 있다. 또한, 이벤트 유사도는 설정 정보로 지정된 이벤트 타입에 대응하는지 여부(즉, 특정의 이벤트에 대한 정답 여부)를 판정할 때에 이용되는 신뢰도(confidence level)로서의 후술된 임계 수준(레벨)을 지정하는 것일 수 있다. 이벤트 중요도는, 설정 정보로 지정된 이벤트별 중요도를 지정하는 것일 수 있다.
- [0065] 지금부터 도 3 내지 도 5을 참조하여 단계 S220의 하위 루틴들의 예시적인 실행 과정을 설명한다.
- [0066] 도 3을 참조하면, 단계 S310에서, 프로세서(133)는, 중계 영상으로부터 참조 프레임을 샘플링한다. 여기서, 중계 영상을 구성하는 정적 이미지 프레임들 중의 어느 하나 또는 둘 이상이 참조 프레임으로서 샘플링될 수 있다. 프로세서(133)는, 중계 영상의 전체 시간 구간 중에서 해당 종목의 경기 규칙에 기초한 시간 범위 내의 프레임을 샘플링할 수 있다.
- [0067] 가령, 단계 S210에서 축구에 대한 중계 영상이 편집 대상으로 설정되어 있다고 해보자. 도 4를 참조하면, 축구의 중계 영상은, 경기 준비 구간(선수 소개 등), 전반전 진행 구간, 휴식 구간, 후반전 진행 구간, 및 경기 마무리 구간(경기 내용 및 결과 안내 등)으로 크게 구분될 수 있고, 각각의 구간의 시간 길이 및/또는 구간들 간의 시간 길이의 비율 등은 통계치로서 미리 주어질 수 있다.
- [0068] 프로세서(133)는, 동일한 시간 간격을 두고 여러 프레임을 참조 프레임으로 샘플링할 수 있다. 예컨대, 5개의 구간에서 하나씩 총 5개의 참조 프레임(#1~#5)이 중계 영상으로부터 샘플링될 수 있다. 또는, 프로세서(133)는, 5개 구간 중에서 경기 진행 구간에 해당하는 전반전 진행 구간과 후반전 진행 구간의 시간 범위를 추정한 다음, 전반전 진행 구간과 후반전 진행 구간 중 적어도 하나로부터 참조 프레임(#2, #4)을 추출할 수 있다.
- [0069] 단계 S320에서, 프로세서(133)는, 단계 S310에서 샘플링된 참조 프레임 내에 중계 상황판이 존재하는지 여부를 판정한다. 구체적으로, 중계 영상에는 화면의 소정 영역에 중계 상황판이 위치하는 바, 프로세서(133)는 중계 영상의 소정 영역을 크로핑하고, 크로핑된 영역에 중계 상황판이 존재하는지 여부를 판정할 수 있다. 단계 S320의 값이 "예"인 경우, 단계 S330으로 진행된다. 단계 S320의 값이 "아니오"인 경우, 단계 S310으로 돌아갈 수 있다.
- [0070] 단계 S330에서, 프로세서(133)는, 참조 프레임 내 중계 상황판으로부터, 적어도 한 라운드의 시작 시각으로부터의 경과 시간을 결정한다.
- [0071] 도 5는 중계 상황판(501)이 표시된 참조 프레임(예, 도 4의 참조 프레임 #2)을 예시한다. 도 5를 참조하면, 참조 프레임(#2)의 좌측 상단에는 중계 상황판(501)이 위치하고 있는 바, 프로세서(133)는 참조 프레임(#2)으로부

터 중계 상황판(501)을 추출한다. 그 다음, 프로세서(133)는, OCR(Optical Character Recognition, 광학식 문자 인식) 등의 텍스트 검출 알고리즘을 중계 상황판(501)에 적용하여, 중계 상황판(501)으로부터 경기의 진행 상황 정보를 취득한다.

[0072] 중계 상황판(501)으로부터 집적적으로 검출 가능한 경기 진행 상황 정보로는 라운드 넘버(2 이상의 라운드 중 현재 라운드를 나타냄), 특정 라운드의 진행 시간(특정 라운드의 시작 시각으로부터 경과된 시간), 양 팀의 점수 등이 있다. 예컨대, 도 5의 중계 상황판(501)에는 양 팀명('대한민국', '코스타리카'), 라운드 넘버('전반'), 라운드(또는 경기)의 시작 시각으로부터의 경과 시간('9:19'), 양 팀의 득점('0-0')이 표시되어 있다.

[0073] 단계 S340에서, 프로세서(133)는, 경과 시간을 기초로, 적어도 한 라운드의 시작 시각 및 종료 시각 중 적어도 하나를 추정하여, 추정된 시각을 포함하는 참조 시각 정보를 생성한다. 프로세서(133)는, 중계 상황판(501)에서 검출된 '전반'의 '경과 시간 9:19'를 기초로, 중계 영상 내 전반전 진행 구간의 시작 위치를 추정(식별)할 수 있다. 즉, 프로세서(133)는 참조 프레임(#2)의 시간 코드로부터 9분 19초만큼 역산함으로써, 중계 영상 내 전반전 진행 구간의 시작 위치에 해당하는 시간 코드를 갖는 프레임을 특정할 수 있다.

[0074] 또한, 중계 상황판(501)으로부터 집적적으로 검출 가능한 경기 진행 상황 정보로는 특정 라운드의 종료까지의 잔여 시간, 다음 라운드의 시작 시각까지의 잔여 시간 등이 있다. 예컨대, 프로세서(133)는, 참조 프레임(#2)의 중계 상황판(501)에서 직접 취득한 정보를 기초로, '전반'의 종료 시각은 물론 추가적으로는 '후반'의 시작 시각, '후반'의 종료 시각 등을 추정할 수 있다. 축구 등과 같이 라운드별 정규 경기 시간이 규정되어 있는 한편 심판의 재량에 따라 어느 정도의 추가 시간이 주어질 수 있는 종목의 경우, 종목별 추가 시간의 평균 등과 같은 통계치(α)가 메모리(132)에 미리 저장되어 있을 수 있고, 프로세서(133)는 종목별 추가 시간 정보와 경기 규칙 정보를 이용하여, 샘플링된 참조 프레임이 속하는 라운드의 종료 시각의 추정치($35:41+\alpha$)는 물론 그에 후행하는 라운드에 관련된 시간 정보(예, 후반 시작까지의 잔여 시간 $50:41+\alpha$, 후반 종료까지의 잔여 시간 $95:41+\alpha$ 등)를 추론할 수 있다.

[0075] 단계 S350에서, 프로세서(133)는, 참조 시각 정보를 기초로, 중계 영상으로부터 경기 비진행 구간을 제거한다. 즉, 프로세서(133)는, 참조 시각 정보를 기초로 중계 영상 내 적어도 한 라운드의 경기 진행 구간과 경기 비진행 구간 간의 경계를 식별하고, 식별된 경계를 기준으로 경기 진행 구간이 아닌 부분을 경기 비진행 구간으로 설정할 수 있다. 도 4를 재참조하면, 전술된 바와 같이, '경기 준비', '휴식' 및 '마무리'는 경기 미진행 구간으로 각각 식별된다. 이로써, 전반전과 후반전에 해당하는 구간이 경기 진행 구간으로서 식별된다. 프로세서(133)는, 경기 진행 구간의 라운드별 시작 위치(RS1, RS2) 및 종료 위치(RE1, RE2)를 나타내는 시간 코드를 메모리(132)에 기록할 수 있다.

[0076] 이어서, 도 6 내지 도 11을 참조하여 이벤트 검출 모델이 이용되는 단계 S240의 하위 루틴들의 예시적인 실행 과정을 설명한다.

[0077] 먼저, 이벤트 검출 모델은 제1 딥러닝 모델 및 제2 딥러닝 모델을 포함하며, 각각의 딥러닝 모델은 딥러닝 기반으로, 단계 S210에서 선택된 중계 영상과 동일 종목의 경기에 대한 복수의 다른 중계 영상으로부터 추출되어 복수의 이벤트 타입 중 어느 하나로 라벨링된 복수의 하이라이트 영상을 포함하는 학습 데이터 세트에 의해 학습 완료된 것일 수 있다.

[0078] 도 6은 경기 진행 구간(도 4의 RS1~RE1의 구간 및 RS2~RE2의 구간 중 어느 하나 또는 둘의 연결)으로부터 추출된 복수의 비디오 클립(VC_1~VC_m)이 이벤트 검출 모델의 제1 딥러닝 모델에 의해 복수의 특징 벡터(FV_1~FV_m)로 변환되는 과정을 설명하는 데에 참조되는 도면이다. 복수의 비디오 클립(VC_1~VC_m)은 경기 진행 구간을 시작 시점으로부터 k초(예, 2초)마다 순차 구획한 것일 수 있다. 영상의 프레임 레이트가 30fps인 경우, 각 비디오 클립은 30k개의 프레임을 포함하게 된다. 마지막 비디오 클립의 시간 길이는 k초 미만일 수 있다.

[0079] 제1 딥러닝 모델은 프로세서(133)에 의해 실행 중에, 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터(FV_1~FV_m)로 변환한다. 즉, 각각의 비디오 클립은 제1 딥러닝 모델을 통과하면서 다차원의 벡터로 변형된다. 이때, 각 특징 벡터의 차수는 d (2 이상의 소정값)일 수 있으며, d 의 값은 학습 과정을 통해 정해질 수 있다. 본 발명에 따르면, 프레임(frame) 단위로 영상 특징을 추출하는 대신, 경기 중계 영상에 담긴 동적 모션에 대한 시간적 정보가 내포(반영)된 클립 단위의 벡터를 획득한다는 점에서 의의가 있다. 특징 벡터는 피쳐 맵(feature map)의 일종일 수 있으며, 비디오 클립별로 그에 포함된 2차원 이미지 프레임들 각각에 패딩(padding) 등의 로직을 적용한 다음 그 결과값을 프레임 시간 코드 순 등의 일정 규칙에 따라 정렬함으로써 생성될 수 있다.

- [0080] 복수의 특징 벡터(FV_1-FV_m)가 일단 취득되면, 프로세서(133)는 제2 딥러닝 모델을 이용하여 복수의 특징 벡터(FV_1-FV_m)로부터 경기 중계 영상 내 특정의 이벤트가 발생된 부분을 탐색한다. 이를 위해서는, 중계 영상이 표현하는 종목에 관련된 복수의 이벤트 타입에 대한 식별 정보가 미리 정의되어 있어야 하는 바, 지금부터 자세히 살펴본다.
- [0081] 도 7은 종목별 사전 정의되는 복수의 이벤트 타입과 이벤트 타입별 검출 기준을 테이블 형식으로 예시하는 도면이고, 도 8은 도 7에 따른 이벤트 타입별 검출 기준에 의해 라벨링된 학습 데이터의 일부분을 예시적으로 보여주는 도면이다.
- [0082] 도 7을 참조하면, 파울/선수 교체/킥오프/옐로우 카드/골/유효슈팅/볼아웃 등 각각이 '축구'에 관련된 복수의 이벤트 타입으로서 설정되어 있으며, 각 이벤트 타입의 검출 기준이 마련되어 있다. '축구' 등의 각종 종목에 관련된 복수의 이벤트 타입은 자유롭게 정해질 수 있는데, 예컨대 특정 선수가 영상 프레임의 일정 비율 이상 클로즈 업되는 것 등이 독립된 이벤트 타입으로서 설정되는 것도 무방하다.
- [0083] 동일한 이벤트 타입이라도 하더라도, 판단하는 사람에 따라 그 발생 시점을 다르게 해석할 여지가 있으므로, 도 7에서와 같이 이벤트 타입별 검출 기준이 사전에 특정됨으로써, 학습 데이터 세트에 대한 라벨링 과정부터 정확한 이벤트 검출을 위한 지표적 성격이 강화될 수 있다.
- [0084] 학습 데이터 세트로서 제공된 임의의 다른 중계 영상으로 제2 딥러닝 모델을 트레이닝하는 과정에서 제2 딥러닝 모델로부터 학습 결과 데이터가 취득될 수 있다. 도 8을 참조하면, 축구 경기의 시작 시로부터 16분 23초에 패 널티 이벤트, 46분 21초에 골 이벤트, 6분 28초에 프리킥 이벤트, 그리고 37분 29초에 파울 이벤트가 발생하여 각각에 대응하는 라벨링이 이루어진 것을 확인할 수 있다.
- [0085] 한편, 도 7 및 도 8에서는 설명의 편의를 위해 경기 종목을 '축구'로 특정하였으나, 지극히 당연하게 본 발명이 '축구'에 대한 중계 영상에 관련된 것으로 국한되지 않는다는 점을 당업자라면 쉽게 이해할 수 있을 것이다.
- [0086] 도 9는 단계 S240의 하위 루틴들의 예시적인 실행 과정을 설명하는 데에 참조되는 순서도이고, 도 10은 복수의 클러스터와 복수의 특징 벡터 간의 대응 관계를 예시적으로 보여주는 도면이고, 도 11은 도 10에 따른 대응 관계 데이터로부터 경기 진행 영상의 하위 영상 구간별로 진행되는 이벤트 검출의 과정을 설명하는 데에 참조되는 도면이다.
- [0087] 도 9를 참조하면, 단계 S910에서, 프로세서(133)는, 이벤트 검출 모델의 제1 딥러닝 모델을 이용하여, 복수의 비디오 클립을 그에 일대일 대응하는 복수의 특징 벡터로 변환한다.
- [0088] 단계 S920 내지 S940에서는 이벤트 검출 모델의 제2 딥러닝 모델이 이용된다.
- [0089] 단계 S920에서, 프로세서(133)는, 복수의 특징 벡터 각각을 복수의 클러스터 중 어느 하나에 맵핑한다. 각 클러스터는 상기 복수의 이벤트 타입 중 적어도 하나를 적어도 부분적으로 표현한다. 도 10을 참조하면, 제2 딥러닝 모델의 트레이닝 과정에서, 클러스터의 개수와 클러스터별 범위 그리고 복수의 이벤트 타입과의 관련도가 결정될 수 있다. 프로세서(133)는, 제2 딥러닝 모델을 이용하여, 복수의 특징 벡터(FV_1-FV_m) 중 유사한 특징을 갖는 벡터들을 동일 클러스터로 분류할 수 있다.
- [0090] 단계 S930에서, 프로세서(133)는, 복수의 특징 벡터를 시간 순으로 그룹핑하여, 복수의 벡터 그룹을 생성한다. 상세하게는, 단일 특징 벡터는 그 자체만으로도 복수의 이벤트 타입 중에서 임의의 이벤트를 지시(충분히 설명)할 수도 있으나, 특징 벡터별로 대응하는 비디오 클립은 어떠한 이벤트 타입을 전체적으로 표현하는 데에는 충분치 않은 제한적인 시간 범위를 가지는 것이 일반적이다.
- [0091] 도 11을 참조하면, 경기 진행 구간으로부터 취득된 복수의 특징 벡터가 제2 딥러닝 모델에 입력되면, 제2 딥러닝 모델에 의해 복수의 특징 벡터는 시간 순으로 소정 개수씩 그룹핑되어 복수의 벡터 그룹(VG_1, VG_2, VG_n)이 생성된다.
- [0092] 단계 S940에서, 프로세서(133)는, 복수의 벡터 그룹과 복수의 이벤트 타입 중 적어도 하나 간의 대응 관계로부터, 경기 진행 구간 내에서 유효 구간을 식별한다.
- [0093] 복수의 비디오 클립(VC_1-VC_m)은 복수의 특징 벡터(FV_1-FV_m)로 일대일 변환되고, 복수의 특징 벡터(FV_1-FV_m)는 복수의 클러스터 중 적어도 하나로 맵핑된다. 따라서, 복수의 특징 벡터와 복수의 클러스터 간의 맵핑 관계를 복수의 비디오 클립의 시간 위치를 기준으로 정렬한다면, 시간 순으로 소정 개수의 특징 벡터의 집합별로 동일 개수의 클러스터가 연속되는 특유의 조합을 얻을 수 있고, 벡터 그룹별 클러스터 조합은 복수의 이

벤트 타입 중 어느 한 이벤트 타입을 나머지 이벤트 타입보다 강하게 표현할 것이다. 본 발명에 있어서, 임의의 벡터 그룹이 특정의 이벤트 타입에 대응한다는 것은, 해당 벡터 그룹이 복수의 이벤트 타입 중 그 특정의 이벤트 타입을 임계 수준 이상으로 충분히 표현하고, 그 특정의 이벤트 타입에 대한 표현 수준이 나머지 이벤트 타입에 대한 표현 수준보다 높음을 의미한다. 일 예로, 벡터 그룹이 임의의 이벤트 타입을 얼마나 높으나 수준으로 표현하는지는 제2 딥러닝 모델에 의해 수치화될 수 있고, 복수의 이벤트 타입에 대한 해당 벡터 그룹의 복수의 수치값 중 최대값이 임계치 이상인 경우 해당 벡터 그룹이 특정의 이벤트 타입에 대응하는 것으로 판정될 수 있다. 이때, 특정의 이벤트 타입에 대응하는 벡터 그룹의 시간 범위 즉, 해당 벡터 그룹을 이루는 소정 개수의 특징 벡터에 연관된 소정 개수의 비디오 클립의 시작 시각과 종료 시간 사이의 구간이 바로 유효 구간이다.

[0094] 서로 인접한 두 유효 구간이 동일 이벤트 타입에 대응하는 경우, 프로세서(133)는 이들 두 유효 구간을 단일의 유효 구간으로 취급(관리)할 수 있다.

[0095] 프로세서(133)는, 제2 딥러닝 모델에 복수의 벡터 그룹(VG_1, VG_2, VG_n)을 입력하며, 제2 딥러닝 모델은 공통의 벡터 그룹에 속하는 각각의 특징 벡터들에 시간 순으로 맵핑된 클러스터들의 조합을 기초로 복수의 이벤트 타입 중 벡터 그룹별 시간 범위에서 발생된 이벤트 타입이 있는지와, 그러한 이벤트 타입이 무엇인지를 식별할 수 있다. 복수의 벡터 그룹(VG_1, VG_2, VG_n)에 대한 이벤트 타입 식별 과정이 완료되면, 그 결과로의 편집 가이드 정보가 생성되는 것이다. 도 11에서는, 벡터 그룹(VG_1)에 대응되는 이벤트 타입은 "볼아웃", 벡터 그룹(VG_2)에 대응되는 이벤트 타입은 미존재, 벡터 그룹(VG_n)에 대응되는 이벤트 타입은 "골"인 것으로 예시되어 있다.

[0096] 프로세서(133)는, 도 2를 참조하여 전송된 설정 정보에 따라 제2 딥러닝 모델을 동작시킴으로써, 경기 진행 구간 내 설정 정보에 부합하는 유효 구간만을 선별적으로 식별할 수 있다. 즉, 편집자로부터의 설정 정보에 "볼아웃" 및 "골"만이 이벤트 타입으로 지정된 경우, 벡터 그룹(VG_2)이 실제로는 "옐로우 카드" 이벤트에 대응하더라도, 도 11에서와 같이 그에 대응되는 이벤트 타입이 없는 것("N/A")으로 식별될 수 있다.

[0097] 도 12는 도 1에 도시된 영상 편집 지원 장치에 의해 실행되는, 본 발명의 다른 실시예에 따른 영상 편집 지원 방법을 예시적으로 설명하는 데에 참조되는 순서도이다. 도 12의 방법은 도 2의 방법과 병렬적으로 실행 가능할 수 있다.

[0098] 도 12를 참조하면, 단계 S1210에서, 프로세서(133)는, 입력부(110)를 통해 자동 편집 요청을 수신한다. 자동 편집 요청에는 사용자(편집자)에 의해 지정된 희망 시간이 포함될 수 있다. 희망 시간은 사용자가 최종적으로 제작하길 원하는 하이라이트 영상의 시간 길이를 나타낸다. 한편, 단계 S1210에 따른 자동 편집 요청은 도 2를 참조하여 전송된 필터링 항목에 대한 설정 정보와 함께 수신될 수도 있으며, 이 경우 단계 S1210은 생략될 수 있다.

[0099] 단계 S1220에서, 프로세서(133)는, 단계 S240에서 생성된 편집 가이드 정보에 따른 적어도 하나의 유효 구간을 가공하여, 희망 시간과 동일한 시간 길이를 갖는 하이라이트 영상을 생성한다.

[0100] 구체적으로, 편집 가이드 정보에 따른 유효 구간이 하나인 경우, 유효 구간의 시간 길이와 희망 시간을 비교하여, 유효 구간을 가공한다. 예컨대, 유효 구간의 시간 길이가 희망 시간보다 짧은 경우, 프로세서(133)는 유효 구간의 재생 속도를 저하시키거나, 유효 구간의 시간 길이와 희망 시간 간의 차이와 동일한 시간 길이를 갖도록 유효 구간의 재생 속도를 증가시킨 비디오 클립을 유효 구간의 시작 위치 및 종료 위치 중 적어도 하나에 연결하여, 하이라이트 영상을 생성할 수 있다.

[0101] 다음으로, 편집 가이드 정보에 따른 유효 구간이 둘 이상인 경우, 설정 정보에 지정된 이벤트 유형별 이벤트 중요도에 맞춰, 각 유효 구간의 시간 길이를 결정한 다음, 결정된 시간 길이를 가지도록 가공된 유효 구간을 연결 시킴으로써, 하이라이트 영상을 생성할 수 있다.

[0102] 일 예로, 희망 시간은 120초, 편집 가이드 정보에 따른 제1 내지 제3 유효 구간의 시간 길이는 각각 60초, 제1 내지 제3 유효 구간의 중요도는 각각 1, 2 및 3(클수록 높은 중요도)이라고 해보자. 이 경우, 유효 구간들의 시간 길이는 동일하므로, 중요도에 따라, 제1 내지 제3 유효 구간에 대한 할당 시간은 각각 $120/(1+2+3)=20$ 초, $120*2/(1+2+3)=40$ 초 및 $120*3/(1+2+3)=60$ 초이다. 결과적으로, 제1 유효 구간은 3배속으로, 제2 유효 구간은 1.5배속으로, 그리고 제3 유효 구간은 그대로 서로 연결된 하이라이트 영상이 생성될 수 있다.

[0103] 다른 예로, 희망 시간은 120초, 편집 가이드 정보에 따른 제1 내지 제3 유효 구간의 시간 길이는 각각 150초, 100초 및 200초, 제1 내지 제3 유효 구간의 중요도는 각각 2, 5 및 3(클수록 높은 중요도)이라고 해보자. 이 경우, 유효 구간별 시간 길이 및 중요도에 따라, 제1 내지 제3 유효 구간에 대한 할당 시간은 각각 $120 * (150*2)$

$(150*2+100*5+200*3) = \text{약 } 25.7\text{초}$, $120 * (100*5) / (150*2+100*5+200*3) = \text{약 } 42.9\text{초}$ 및 $120 * (200*3) / (150*2+100*5+200*3) = \text{약 } 51.4\text{초}$ 이다. 결과적으로, 제1 유효 구간은 150/25.7배속으로, 제2 유효 구간은 100/42.9배속으로, 그리고 제3 유효 구간은 200/51.4배속으로 조정된 후 서로 연결된 하이라이트 영상이 생성될 수 있다.

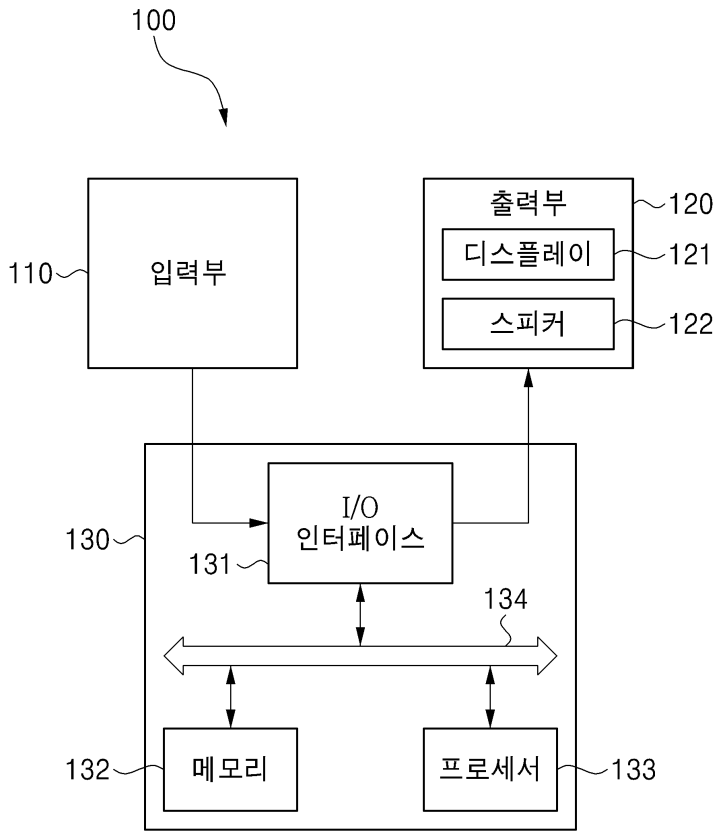
- [0104] 전술된 유효 구간의 가공을 위한 할당 시간의 연산 방식은 하나의 예시일 뿐이다. 즉, 임의의 유효 구간의 할당 시간이 해당 유효 구간의 시간 길이와 중요도에 양의 상관 관계를 갖도록 정해지는 방식이라면 변형되어도 무방하다.
- [0105] 본 발명은 상술한 특정의 실시예 및 응용예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 구별되어 이해되어서는 안 될 것이다.
- [0106] 특히, 본 명세서에 첨부된 도면에 도시된 블록도와 순서도에 포함된 본 발명의 기술적 특징을 실행하는 구성들은 상기 구성들 사이의 논리적인 경계를 의미한다. 그러나 소프트웨어나 하드웨어의 실시 예에 따르면, 도시된 구성들과 그 기능들은 독립형 소프트웨어 모듈, 모놀리식 소프트웨어 구조, 코드, 서비스 및 이들을 조합한 형태로 실행되며, 저장된 프로그램 코드, 명령어 등을 실행할 수 있는 프로세서를 구비한 컴퓨터에서 실행 가능하게 매체에 저장되어 그 기능들이 구현될 수 있으므로 이러한 모든 실시 예 역시 본 발명의 권리범위 내에 속하는 것으로 보아야 할 것이다.
- [0107] 따라서, 첨부된 도면과 그에 대한 기술은 본 발명의 기술적 특징을 설명하기는 하나, 이러한 기술적 특징을 구현하기 위한 소프트웨어의 특정 배열이 분명하게 언급되지 않는 한, 단순히 추론되어서는 안 된다. 즉, 이상에서 기술한 다양한 실시 예들이 존재할 수 있으며, 그러한 실시 예들이 본 발명과 동일한 기술적 특징을 보유하면서 일부 변형될 수 있으므로, 이 역시 본 발명의 권리범위 내에 속하는 것으로 보아야 할 것이다.
- [0108] 또한, 순서도의 경우 특정한 순서로 도면에서 동작들을 묘사하고 있지만, 이는 가장 바람직한 결과를 얻기 위하여 도시된 것으로서, 도시된 특정한 순서나 순차적인 순서대로 그러한 동작들을 반드시 실행되어야 한다거나 모든 도시된 동작들이 반드시 실행되어야 하는 것으로 이해되어서는 안 된다. 특정한 경우, 멀티 태스킹과 병렬 프로세싱이 유리할 수 있다. 아울러, 이상에서 기술한 실시형태의 다양한 시스템 컴포넌트의 분리는 그러한 분리를 모든 실시형태에서 요구하는 것으로 이해되어서는 안되며, 설명한 프로그램 컴포넌트와 시스템들은 일반적으로 단일의 소프트웨어 제품으로 함께 통합되거나 다중 소프트웨어 제품에 패키징될 수 있다는 점을 이해하여야 한다.

부호의 설명

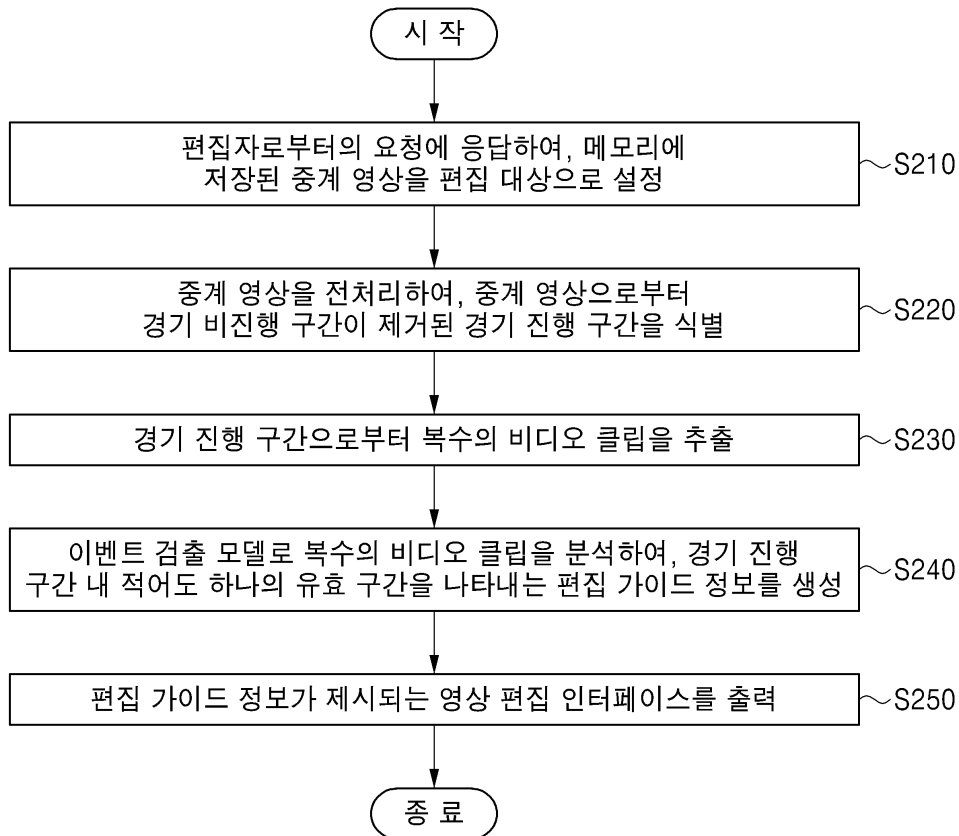
- [0110] 100: 영상 편집 지원 장치
- 110: 입력부
- 120: 출력부
- 122: 스피커
- 130: 제어부
- 132: 메모리
- 134: 데이터 버스
- 121: 디스플레이
- 131: 입출력 인터페이스
- 133: 프로세서

도면

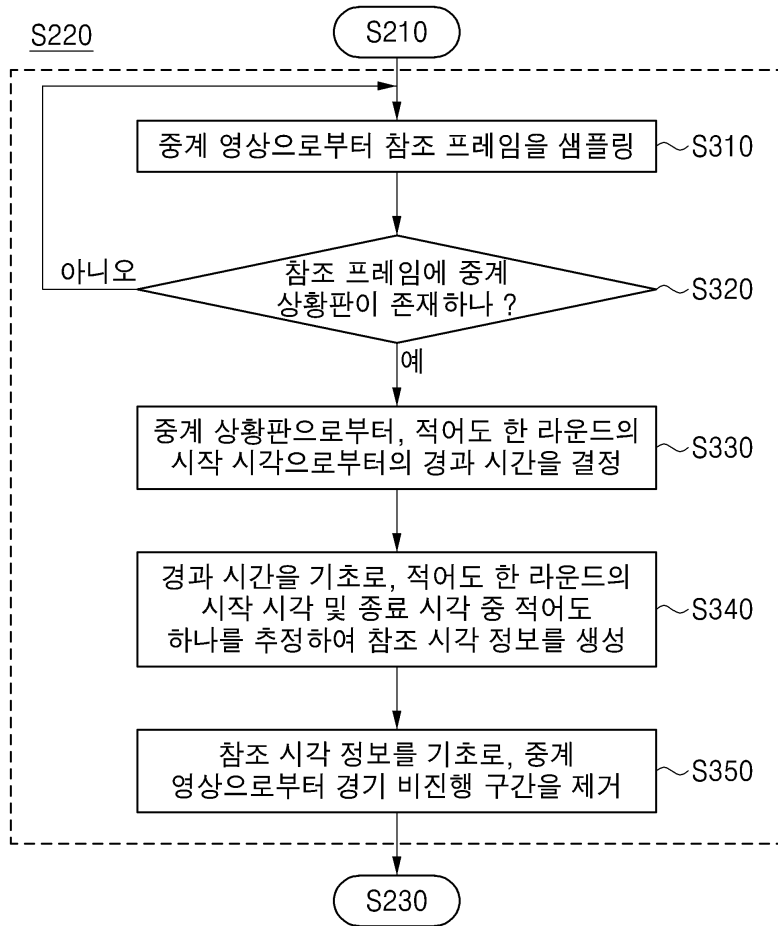
도면1



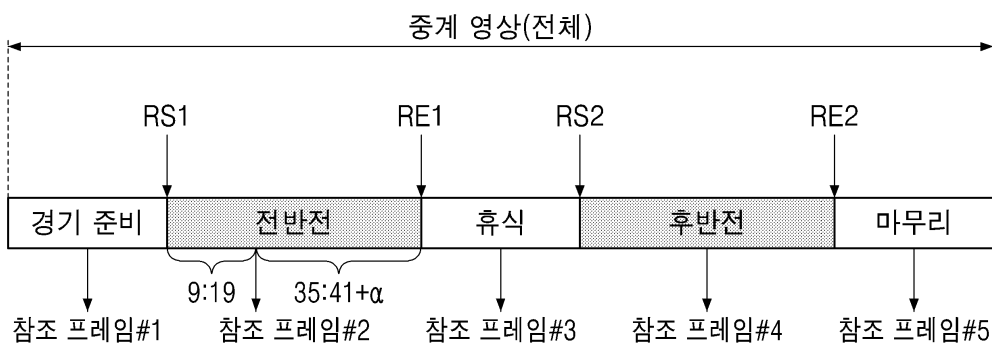
도면2



도면3



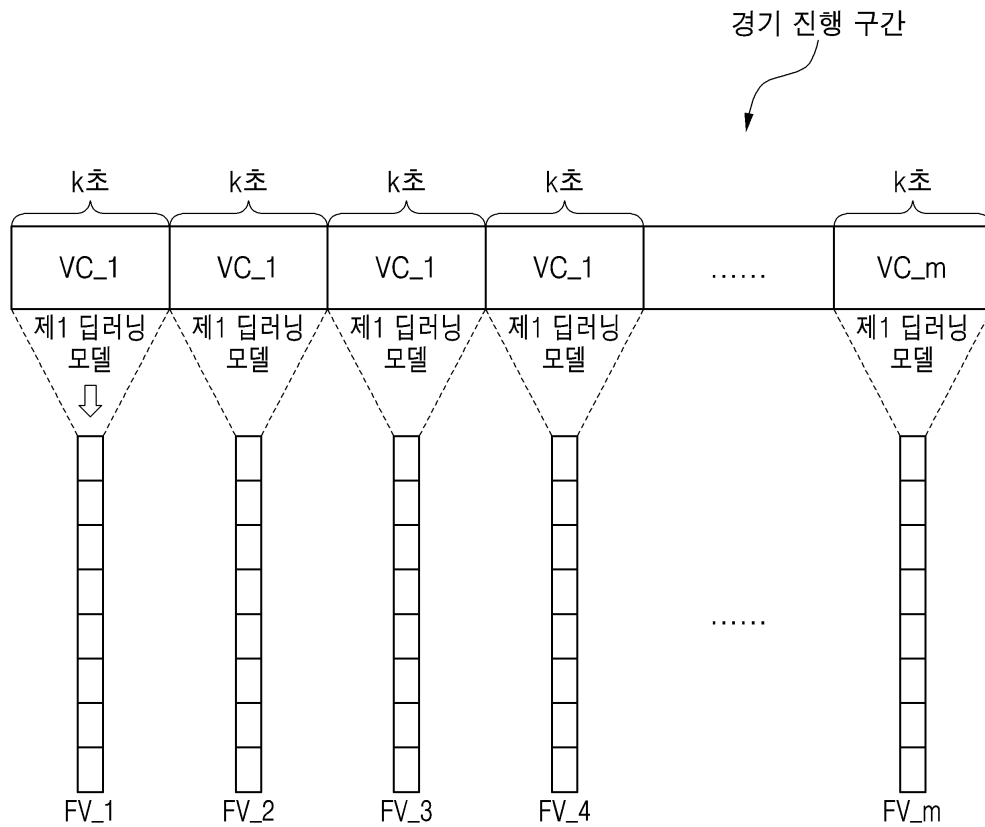
도면4



도면5



도면6



도면7

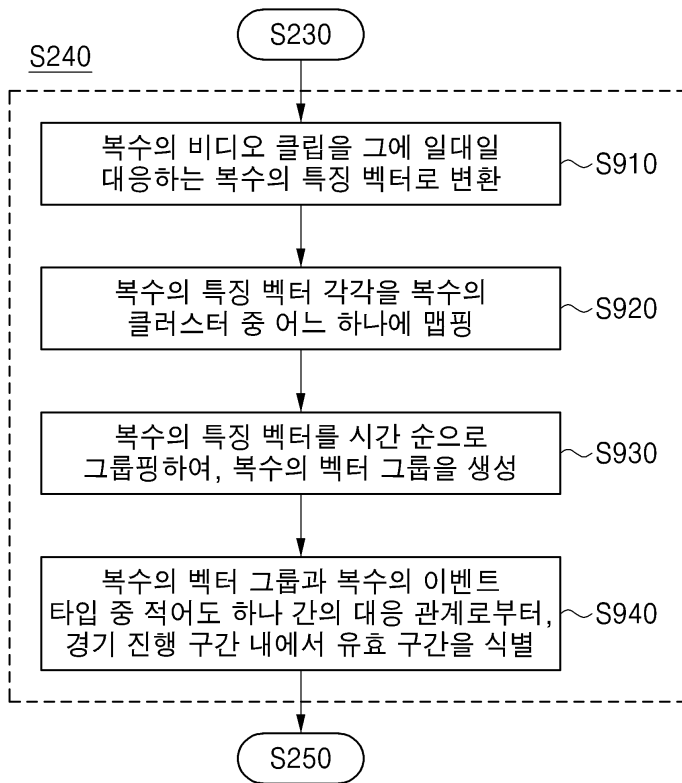
이벤트 타입	검출 기준
파울	심판으로부터 파울이 선언된 시점
선수 교체	교체 대기 선수가 경기장 라인 안으로 들어서는 시점
킥오프	하프타임의 시작점이나 골 이후, 두 명의 선수가 경기 중앙의 원에서 첫 번째 패스를 시작하는 시점
옐로우카드	심판이 선수에게 옐로우카드를 꺼내보이는 시점
골	공이 골라인을 지나치는 시점
유효슈팅	골을 성공시키기 위한 의도를 가지고 골대 방향으로 공을 찬 시점
볼아웃	경기장 외곽 라인을 공이 넘어가는 시점

⋮

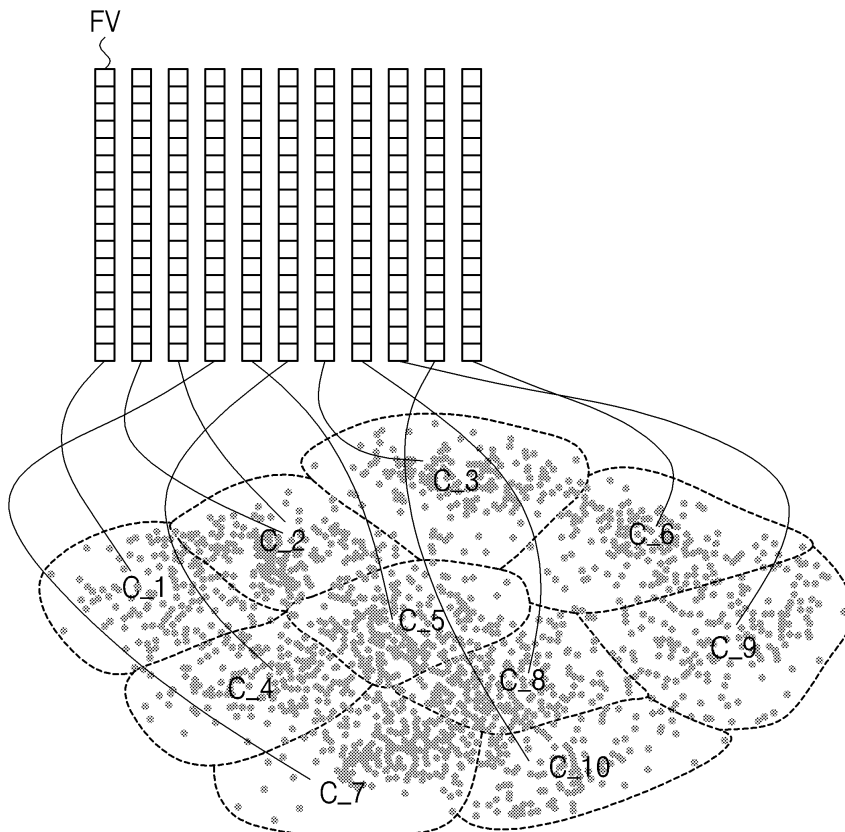
도면8

```
{
  ⋮
  "gameTime" : "16:23",
  "label" : "Penalty"
},
{
  ⋮
  "gameTime" : "46:21",
  "label" : "Goal"
},
{
  ⋮
  "gameTime" : "6:28",
  "label" : "Free-kick"
},
{
  ⋮
  "gameTime" : "37:29",
  "label" : "Foul"
},
⋮
⋮
⋮
```

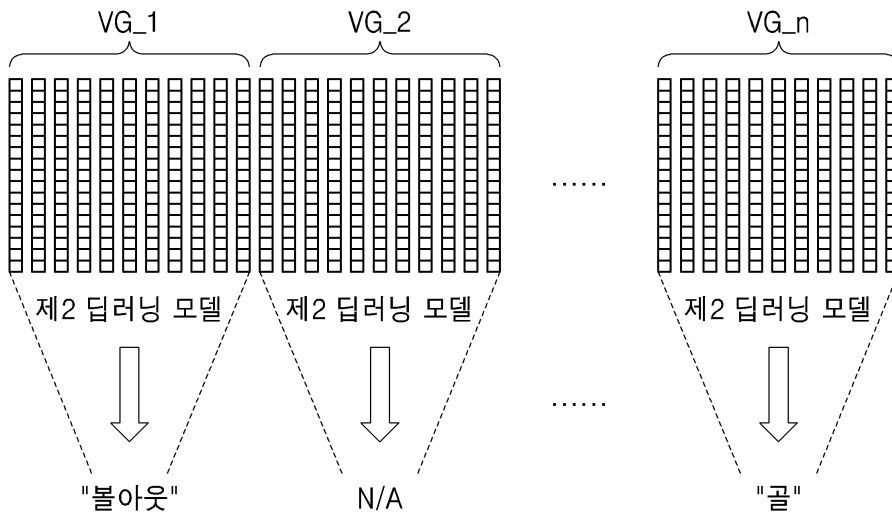
도면9



도면10



도면11



도면12

