



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0042849
(43) 공개일자 2012년05월03일

(51) 국제특허분류(Int. Cl.)
G06T 7/00 (2006.01)
(21) 출원번호 10-2012-7001591
(22) 출원일자(국제) 2010년07월19일
심사청구일자 없음
(85) 번역문제출일자 2012년01월19일
(86) 국제출원번호 PCT/US2010/002028
(87) 국제공개번호 WO 2011/011052
국제공개일자 2011년01월27일
(30) 우선권주장
61/271,381 2009년07월20일 미국(US)

(71) 출원인
톰슨 라이센싱
프랑스 92130 이씨레귤리노 찬 다르크 뤼 1-5
(72) 발명자
바가바티, 시타람
미국 08536 뉴저지주 플레인스보로 헌터스 글렌
드라이브 5910
장, 동 킹
미국 08536 뉴저지주 플레인스보로 레이븐 크레스
트 드라이브 310
(74) 대리인
백만기, 양영준, 전경석

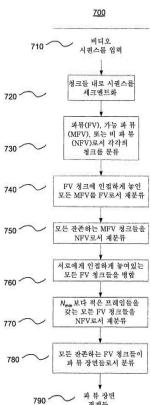
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **스포츠 비디오에서의 패 뷰 장면들에 대한 비디오 프로세싱을 검출하고 적용시키기 위한 방법**

(57) 요 약

특정 이미지 프로세싱 알고리즘들이 적용되어야 하는 때를 결정하기 위해 스포츠 비디오에서 패 뷰 장면들을 검출하도록 비디오를 분석하기 위한 방법이 개시된다. 본 방법은 비디오 신호로부터의 이미지들의 뷰 필드들을 분석하고 분류하는 단계, 연속적인 이미지들의 세트들의 뷰 필드들을 생성하고 분류하는 단계, 및 특정 유형의 뷰 필드를 나타내는 연속적인 이미지들의 세트들에 이미지 프로세싱 알고리즘들을 선택적으로 적용하는 단계를 포함한다.

대 표 도 - 도7



특허청구의 범위

청구항 1

비디오 프로세싱 시스템에서 비디오를 프로세싱하는 방법으로서,

비디오 신호로부터의 복수의 이미지를 분석하는 단계;

상기 이미지의 피쳐(feature)의 적어도 하나의 사이즈에 응답하여 상기 비디오 신호로부터의 상기 복수의 이미지 각각의 뷰 필드를 분류하는 단계;

인접하는 이미지들의 메트릭들(metrics)의 비교들에 기초하여 상기 비디오 신호로부터의 상기 복수의 이미지로부터 연속적인 이미지들의 세트들을 생성하는 단계;

상기 연속적인 이미지들의 세트들 각각의 뷰 필드를 분류하는 단계; 및

적어도 특정 유형의 뷰 필드를 나타내는 상기 연속적인 이미지들의 세트들에 이미지 프로세싱 알고리즘을 선택적으로 적용하는 단계

를 포함하는 비디오 프로세싱 방법.

청구항 2

제1항에 있어서,

상기 비디오 신호로부터의 상기 복수의 이미지 각각은 프레임인 비디오 프로세싱 방법.

청구항 3

제1항에 있어서,

상기 사이즈는 길이, 너비, 또는 면적인 비디오 프로세싱 방법.

청구항 4

제1항에 있어서,

상기 메트릭들은 컬러 히스토그램 데이터를 포함하는 비디오 프로세싱 방법.

청구항 5

제1항에 있어서,

상기 이미지의 뷰 필드를 분류하는 단계는,

상기 이미지가 뷰 샷(view shot)의 파 필드를 나타내는지, 뷰 샷의 파 필드를 나타낼 수 있는지, 또는 뷰 샷의 파 필드를 나타내지 않는지를 표시하는 데이터와 상기 이미지를 연관시키는 단계를 포함하는 비디오 프로세싱 방법.

청구항 6

제1항에 있어서,

상기 이미지의 뷰 필드를 분류하는 단계는,

경기장을 나타낼 수 있는 상기 이미지의 면적들을 나타내는 마스크를 생성하는 단계를 포함하는 비디오 프로세싱 방법.

청구항 7

제6항에 있어서,

상기 이미지의 뷰 필드를 분류하는 단계는,

상기 마스크의 계산된 면적을 적어도 하나의 임계치와 비교하는 단계를 포함하는 비디오 프로세싱 방법.

청구항 8

제6항에 있어서,

상기 이미지의 뷰 필드를 분류하는 단계는,

상기 마스크의 경계들 내의 플레이어 같은 오브젝트들의 면적들을 식별하고 계산하는 단계를 포함하는 비디오 프로세싱 방법.

청구항 9

제8항에 있어서,

상기 이미지의 뷰 필드를 분류하는 단계는,

상기 오브젝트들 중 적어도 하나의 중앙값 계산된 면적 또는 최대값 계산된 면적을 임계치와 비교하는 단계를 포함하는 비디오 프로세싱 방법.

청구항 10

제1항에 있어서,

상기 연속적인 이미지들의 세트의 뷰 필드를 분류하는 단계는,

상기 세트 내의 상기 이미지들의 상기 분류들의 통계적 분석을 포함하는 비디오 프로세싱 방법.

청구항 11

제10항에 있어서,

상기 통계적 분석은,

상기 세트 내의 상기 비디오 신호로부터의 상기 이미지들의 절반 이상에 관련하는 분류를 식별하는 단계를 포함하는 비디오 프로세싱 방법.

청구항 12

비디오 프로세싱 시스템에서 비디오를 프로세싱하기 위한 장치로서,

비디오 신호로부터의 복수의 이미지를 분석하기 위한 수단;

상기 이미지의 피처의 적어도 하나의 사이즈에 응답하여 상기 비디오 신호로부터의 상기 복수의 이미지 각각의 뷰 필드를 분류하기 위한 수단;

인접한 이미지들의 메트릭들의 비교들에 기초하여 상기 비디오 신호로부터의 상기 복수의 이미지로부터 연속적인 이미지들의 세트들을 생성하기 위한 수단;

상기 연속적인 이미지들의 세트들 각각의 뷰 필드를 분류하기 위한 수단; 및

적어도 특정 유형의 뷰 필드를 나타내는 상기 연속적인 이미지들의 세트들에 이미지 프로세싱 알고리즘을 선택적으로 적용하기 위한 수단

을 포함하는 비디오 프로세싱 장치.

청구항 13

제12항에 있어서,

상기 비디오 신호로부터의 상기 복수의 이미지 각각은 프레임인 비디오 프로세싱 장치.

청구항 14

제12항에 있어서,

상기 이미지의 뷰 필드를 분류하는 것은,

상기 이미지가 뷰 샷(view shot)의 과 필드를 나타내는지, 뷰 샷의 과 필드를 나타낼 수 있는지, 또는 뷰 샷의 과 필드를 나타내지 않는지를 지시하는 데이터와 상기 이미지를 연관시키는 것을 포함하는 비디오 프로세싱 장치.

청구항 15

제12항에 있어서,

상기 이미지의 뷰 필드를 분류하는 것은,

경기장을 나타낼 수 있는 상기 이미지의 면적들을 나타내는 마스크를 생성하는 것을 포함하는 비디오 프로세싱 장치.

청구항 16

제15항에 있어서,

상기 이미지의 뷰 필드를 분류하는 것은,

상기 마스크의 계산된 면적을 적어도 하나의 임계치와 비교하는 것을 포함하는 비디오 프로세싱 장치.

청구항 17

제15항에 있어서,

상기 이미지의 뷰 필드를 분류하는 것은,

상기 마스크의 경계를 내의 플레이어 같은 오브젝트들의 면적들을 식별하고 계산하는 것을 포함하는 비디오 프로세싱 장치.

청구항 18

제12항에 있어서,

상기 연속적인 이미지들의 세트의 뷰 필드를 분류하는 것은,

상기 세트 내의 상기 이미지들의 상기 분류들의 통계적 분석을 포함하는 비디오 프로세싱 장치.

청구항 19

비디오 프로세싱 시스템에서 비디오를 프로세싱하기 위한 장치로서,

비디오 신호로부터의 복수의 이미지를 분석하고;

상기 이미지의 피처의 적어도 하나의 사이즈에 응답하여 상기 비디오 신호로부터의 상기 복수의 이미지 각각의 뷰 필드를 분류하고;

인접한 이미지들의 메트릭들의 비교들에 기초하여 상기 비디오 신호로부터의 상기 복수의 이미지로부터 연속적인 이미지들의 세트들을 생성하고;

상기 연속적인 이미지들의 세트를 각각의 뷰 필드를 분류하고;

적어도 특정 유형의 뷰 필드를 나타내는 상기 연속적인 이미지들의 세트들에 이미지 프로세싱 알고리즘을 선택적으로 적용하기 위한 프로세서

를 포함하는 비디오 프로세싱 장치.

청구항 20

제19항에 있어서,

상기 연속적인 이미지들의 세트의 뷰 필드를 분류하는 것은,

상기 세트 내의 상기 이미지들의 분류들의 통계적 분석을 포함하는 비디오 프로세싱 장치.

명세서

기술분야

[0001] [관련 출원에 대한 교차 참조]

[0002] 본 출원은 2009년 7월 20일 미국 특허 상표국(United States Patent and Trademark Office)에 출원되고, 일련 번호 제61/271,381호로 지정된 임시 출원에 대한 우선권을, 및 이로부터 발생하는 모든 이점들을 주장한다.

배경기술

[0003] 본 발명은 특정 이미지 프로세싱 알고리즘들이 적용되어야 하는 때를 결정하기 위해 스포츠 비디오에서 파 뷔를 검출하도록 비디오를 분석하기 위한 방법 및 이와 연관된 장치에 일반적으로 관련한다. 본 방법은 비디오 신호로부터의 이미지들의 뷔 필드들(fields of view)을 분석하고 분류하는 단계, 연속적인 이미지들의 세트들의 뷔 필드들을 생성하고 분류하는 단계, 및 이미지 프로세싱 알고리즘들을 특정 유형의 뷔 필드를 나타내는 연속적인 이미지들의 세트들에 선택적으로 적용하는 단계를 포함한다.

[0004] 이 섹션은 독자에게 이하에 기술된 본 발명의 여러 특징에 관련될 수 있는 기술의 여러 특징을 소개하도록 의도된다. 이 논의는 독자에게 본 발명의 여러 특징의 더 양호한 이해를 촉진하도록 배경 정보를 제공하는 데에 있어서 도움이 될 것으로 이해된다. 따라서, 이런 진술들이 종래 기술들의 인정으로서가 아니라 이러한 관점에서 해석되어야 한다는 것을 알아야 한다.

[0005] 모바일 장치들이, 능력이 더 많아지고 모바일 디지털 텔레비전 표준이 진보함에 따라, 그러한 장치들 상에서 비디오 프로그램을 보는 것이 더욱 실체적으로 되었다. 그러나, 이러한 장치들의 작은 스크린들은 특히 스포츠 이벤트들의 뷔에 대해서 몇몇의 제약들을 제시한다. 스포츠 이벤트에서의 공과 같은 작은 오브젝트들은 보기 어려울 수 있다. 높은 비디오 압축 비율의 이용은 파 뷔 장면에서 공과 같은 작은 오브젝트들의 외관을 현저히 퇴화(degrading)시킴으로써 상황을 악화시킬 수 있다.

[0006] 이러한 오브젝트들의 외관을 강화(enhance)시키는 것은 가능하지만, 그러한 알고리즘들은 계산적으로 비용이 많이 들 수 있고 또는 필요하지 않은 경우에 적용되면 전체 이미지 품질을 퇴화시킬 수 있다. 오브젝트 강화 알고리즘들로부터 이익을 얻을 수 있는 장면들의 특정 유형들을 검출할 수 있어서 알고리즘들이 선택적으로 적용될 수 있도록 하는 것이 바람직할 것이다. 여기서 기술된 본 발명은 이런 및/또는 다른 문제를 다룬다.

[0007] [발명의 요약]

[0008] 앞서 기술된 문제들을 해결하기 위해서, 본 발명은 특정 이미지 프로세싱 알고리즘들이 적용되어야 하는 경우를 결정하도록 스포츠 비디오에서 파 뷔 장면들을 검출하도록 비디오를 분석하는 것에 관계한다. 본 방법은 비디오 신호로부터 이미지들의 뷔 필드들을 분석하고 분류하는 단계, 연속 이미지들의 세트들의 뷔 필드들을 생성하고 분류하는 단계, 및 이미지 프로세싱 알고리즘들을 뷔 필드의 특정 유형을 나타내는 연속 이미지들의 세트들에 선택적으로 적용하는 단계를 포함한다. 본 발명의 이런 및 다른 특징들은 첨부된 도면들을 참조하여 상세히 기술될 것이다.

도면의 간단한 설명

[0009] 첨부된 도면들과 함께 취해진 본 발명들의 실시예들의 이하의 설명을 참조함으로써, 본 발명의 앞서 언급된 그리고 다른 특징들 및 이점들 및 이것들을 얻는 방식들이 더 명백해질 것이고, 본 발명이 더 잘 이해될 것이다.

도 1은 오브젝트 로컬라이제이션 및 강화(object localization and enhancement)를 구비한 종래 기술의 비디오 프로세싱 시스템의 다이어그램.

도 2는 파 뷔 장면 검출을 활용하는 오브젝트 로컬라이제이션 및 강화를 구비한 비디오 프로세싱 시스템의 다이어그램.

도 3은 프레임에 의해 나타내어진 뷔 필드를 분류하기 위한 방법의 흐름도.

도 4는 프레임에 의해 나타내어진 뷔 필드를 분류하기 위한 방법의 부분들을 적용한 결과들의 예시 도면.

도 5는 비디오 시퀀스를 청크들(chunks)로 세그멘트화(segmenting)하는 방법의 흐름도.

도 6은 비디오 시퀀스의 청크의 뷰 필드를 분류하는 방법의 흐름도.

도 7은 청크 레벨 분류들을 활용하는 과 뷰 장면 검출의 방법의 흐름도.

여기서 나열한 예시들은 본 발명의 바람직한 실시예들을 설명한다. 그러한 예시들은 어떤 방식으로도 본 발명의 범위를 제약하는 것으로 해석되어서는 안된다.

[바람직한 실시예들의 기술]

여기서 기술된 바와 같이, 본 발명은 특정 이미지 프로세싱 알고리즘들이 적용되어야 할 때를 결정하기 위해 스포츠 비디오에서 과 뷰 장면들을 검출하도록 비디오를 분석하기 위한 방법 및 이와 연관된 장치를 제공한다. 본 방법은 비디오 신호로부터의 이미지들의 뷰 필드들을 분석하고 분류하는 단계, 연속적 이미지들의 세트들의 뷰 필드들을 생성하고 분류하는 단계, 및 이미지 프로세싱 알고리즘들을 특정 유형의 뷰 필드를 나타내는 연속적 이미지들의 세트들에 선택적으로 적용하는 단계를 포함한다.

본 발명이 바람직한 설계를 갖는 것으로서 기술되었지만, 본 발명은 본 개시의 사상 및 범위 내에서 추가로 수정될 수 있다. 본 출원은 그에 따라 이것의 일반적인 원리들을 이용하여 본 발명의 임의의 변형들, 이용들 또는 적용들을 커버하도록 의도된다. 추가로, 본 출원은, 본 발명이 관계하는 그리고 첨부된 청구항들의 경계들 내에 들어가는 기술 분야에서 알려진 또는 관습적인 실시의 범위 내에 있는 본 개시로부터의 그러한 변경들(departures)을 커버하도록 의도된다.

바람직한 실시예에서, 본 발명은 텔레비전 프로덕션 또는 송신 환경 내에서의 신호 프로세싱 하드웨어에서 구현될 수 있다. 본 방법은 스포츠 비디오에서, 예시적 응용으로 축구 비디오에서 과 뷰 장면들을 검출하기 위해 이용된다. 과 뷰 장면들은 경기장의 광각 카메라 뷰들에 대응하는 것들인데, 여기서, 예를 들어 플레이어들 및 공과 같은 관심 있는 오브젝트들은 비디오 압축에 의해 쉽게 퇴화되거나 선명하게 보이지 않기에 충분하다.

도 1은, 특정 장면의 뷰 필드에 무관하게, 비디오의 모든 부분들에 적용된 오브젝트 로컬라이제이션 및 강화를 구비한 전형적인 이미지 프로세싱 시스템(100)을 도해한다. 입력 비디오(110)는 처음으로 오브젝트 로컬라이제이션 알고리즘들(120)을 통해 공들 또는 퍽들(pucks)과 같은 검출된 오브젝트들에 대한 정보(130)를 산출하기 위해 프로세싱된다. 관심 있는 오브젝트의 특성들에 대응하는 필터들의 이용과 같은, 그러한 오브젝트들을 검출하기 위한 여러 기법이 기술분야에서 알려져 있다.

식별된 오브젝트들에 대한 정보(130)는 오브젝트 강화 스테이지(150)로, 그리고 오브젝트 인식 인코더(object aware encoder)(170)로 넘겨진다. 이런 정보는, 예를 들어 공의 위치, 사이즈, 궤적(trajecotry), 또는 마스크를 포함할 수 있다. 비디오 프레임들(140)이 식별된 오브젝트들에 대한 정보(130)를 이용하여 오브젝트 강화 스테이지(150)에서 프로세싱된다. 예를 들어, 하이라이트 컬러가 공 또는 퍽의 위치 위에 놓여, 뷰어가 그것의 위치를 더 쉽게 식별하는 것을 허용할 수 있다.

검출된 오브젝트들에 강화가 적용된 결과 비디오 프레임들(160)이 그 후 오브젝트 인식, 인코더(170)에 의해 인코딩되고, 그 결과 출력 비트스트림(180)이 생긴다. 오브젝트 인식 인코더(170)에 의한 오브젝트 정보(130)의 이용은, 플레이어들 또는 공들과 같은, 과 뷰 장면들에서 식별된 오브젝트들의 가시성 및 외관을 보존하도록 인코딩이 조정되는 것을 허용할 수 있다. 예를 들어, 공이 작은 오브젝트로서 나타나는 장면들 또는 볼이 나타나는 프레임들의 특정 영역들에 대해 보다 낮은 압축 비율이 이용될 수 있다.

시스템(100)과 같은 시스템에서, 오브젝트 로컬라이제이션 및 강화는 프로세싱되고 있는 프레임들에서 나타내어지는 뷰의 유형에 관계없이 수행된다. 따라서, 불필요한 프로세싱이 장면들의 몇몇 유형들에 대해 수행되어, 시간 낭비, 프로세싱 리소스들의 낭비, 또는 이미지 품질 저하를 잠재적으로 야기한다.

도 2는 본 발명에 따른 비디오 프로세싱 시스템에서의 데이터 흐름을 도해한다. 입력 비디오(210)는 처음으로 이하에서 상세하게 기술될 과 뷰 장면 검출 알고리즘(220)에 의해 프로세싱된다. 과 뷰 장면 검출(220)의 추가로, 과 뷰로서 분류된 장면들(230)은 과 뷰가 아니라고 분류된 장면들(240)과는 개별적으로 프로세싱될 수 있다. 이러한 점은 시간 절약, 프로세싱 리소스들의 절약, 또는 이미지 품질 저하를 잠재적으로 야기한다.

과 뷰 장면들(230)은 오브젝트 로컬라이제이션 및 강화 프로세싱(250)에게 보내진다. 이 프로세싱은 검출된 공을 하이라이팅하는 것, 공 궤적의 일러스트레이션, 또는 다른 강화들을 포함할 수 있다. 비 과 뷰 장면들(non-far-view scenes)(240)은 오브젝트 로컬라이제이션 및 강화 스테이지를 우회하고 오브젝트 인식 인코더(280)에

게 바로 보내진다. 오브젝트 로컬라이제이션 및 강화 스테이지(250)에 의해 산출된 오브젝트 정보(260) 및 강화된 파 뷔 장면들(270)이 또한 오브젝트 인식 인코더(280)에게 보내지고, 이것은 인코딩된 출력 비트스트림(290)을 산출한다.

앞서 기술된 바와 같이, 오브젝트 인식 인코더(280)에 의한 오브젝트 정보(260)의 이용은, 파 뷔 장면들에서 예를 들어 축구공들과 같은 식별된 오브젝트들의 가시성 및 외관을 보존하도록 조정되도록 인코딩하는 것을 허용한다. 예를 들어, 공이 나타나는 장면들에 대해, 또는 공이 나타나는 프레임들의 특정 영역들에 대해 보다 낮은 압축 비율이 이용될 수 있다.

파 뷔 장면들의 검출 스텝(220)은 이하의 스테이지들을 포함한다:

1. 도 3에 대해 이하에서 기술되는, 파 뷔(FV: far view), 가능 파 뷔(MFV: maybe far-view), 또는 비 파 뷔(NFV: not far-view)로서 시퀀스 내의 각각의 프레임의 분류.
2. 도 5에 대해 이하에서 기술되는, 시퀀스를 다수의 청크들로 세그멘트화, 여기서 각각의 청크는 "유사하게 보이는" 프레임들의 인접하는(contiguous) 세트임.
3. 도 6에 대해 이하에서 기술되는, FV, MFV, 또는 NFV로서 각각의 청크의 분류.
4. 도 7에 대해 이하에서 기술되는, 청크들을 병합하고 재분류함으로써 파 뷔 장면들을 검출.

도 3은 축구 비디오의 예시적 응용에 대해 기술되는 파 뷔 프레임 분류의 방법(300)의 흐름도이다. 축구 비디오로부터의 입력 프레임(310)이 축구 필드 검출 스텝(320)에서 2진 필드 마스크로 프로세싱된다. 간단한 경우에, 축구 필드 픽셀들은, 필드가 일반적으로 잔디 또는 잔디 색을 띤 물질로 커버된다는 지식을 이용하여 식별된다. 그 결과는 1의 값을 갖는 모든 필드 픽셀들 및 0의 값을 갖는, 필드 내의 오브젝트들을 포함하는 모든 비 필드 픽셀들을 분류하는 2진 마스크이다. 그 후 경기장의 경계들을 식별하고, 전경 오브젝트들을 무시하고, 또한 견고한 필드 마스크를 생성하기 위해 여러 이미지 프로세싱 기법들이 이용될 수 있다. 예를 들어, 특정 면적 임계치를 초과하는 필드 픽셀들의 모든 인접하는 구역들을 둘러싸는 간단한 바운딩 박스(bounding box) 내의 모든 픽셀들이 필드 마스크 내에 포함될 수 있다. 필드를 식별하고 필드 마스크로부터 전경 오브젝트들을 제거하기 위해 필터들의 이용을 포함하는 다른 기법들이 이용될 수 있다.

스텝 330에서, "플레이어 같은" 오브젝트들이 전경 오브젝트들의 분석을 통해 식별되는데, 이는 스텝 320에서 식별된 필드 경계 내의 비 필드 픽셀들의 연결된 세트들이다. 구성 픽셀들{(xi, yi)}을 갖는 전경 오브젝트 o에 대해, 이하의 오브젝트 파라미터들이 계산된다:

$$a_o = \sum_i l_i, \quad (\text{면적: } l_i, \text{ 즉, 오브젝트 내의 픽셀들의 수})$$

$$\text{높이: } h_o = \max_i y_i - \min_i y_i + 1,$$

$$\text{폭: } w_o = \max_i x_i - \min_i x_i + 1,$$

밀집도(compactness): $c_o = a_o / (h_o \times w_o)$, 및

종횡비(aspect ratio): $r_o = h_o / w_o$.

면적, 밀집도, 및 종횡비 각각이 축구 비디오에서 플레이어들의 알려진 특성들에 관련하는 임계치를 넘는 경우에 오브젝트들이 "플레이어 같다"고 여겨진다. 달리 말해서, 만약 ($a_o > t_a$), ($c_o > t_c$), 및 ($r_o > t_r$)이면, 오브젝트(o)는 "플레이어 같다"고 여겨진다. 320x240 픽셀 축구 비디오들로 지향되는 바람직한 실시예에서, $t_a=10$, $t_c=0.1$, 및 $t_r=0.5$ 의 임계치가 이용된다.

스텝 340에서, 모든 플레이어 같은 오브젝트들의 최대 면적 A_{\max} 및 중앙값 면적 A_{med} 이 계산된다. 앞서 기술된 바와 같이, 특정 오브젝트의 면적은 오브젝트를 포함하는 픽셀들의 수로서 간단하게 계산될 수 있다. 마찬가지로, 필드 A_{field} 의 면적은 필드 마스크를 포함하는 픽셀들의 수로서 계산될 수 있다.

스텝 350에서, 필드의 면적, 플레이어 오브젝트들의 중앙값 면적, 및 플레이어 오브젝트들의 최대 면적이 파-뷰

장면에서 기대되는 면적들에 관련하는 임계치들과 비교된다. 만약 ($A_{\text{field}} > T_{\text{field}}$), ($A_{\text{med}} < T_{\text{med}}$), 및 ($t_{\text{max}} < A_{\text{max}} < T_{\text{max}}$)이면, 프레임은 스텝 360에서 FV로서 라벨링된다. 즉, 만약 프레임에서의 필드 면적이 넓으면, 중앙값 플레이어 면적은 작고, 최대 플레이어 면적은 기대되는 범위 내에 있고, 장면의 뷰 필드는 넓거나 멀다.

만약 프레임이 FV로서 라벨링되지 않으면, 프레임이 과 뷰일 수 있는지, 아니면 과 뷰가 아닐 가능성이 큰지에 대한 추가적인 결정들이 이루어진다. 만약 기준 ($A_{\text{field}} > t_{\text{field}}$) 및 ($A_{\text{max}} \leq t_{\text{max}}$)가 스텝 370에서 만족되면, 이 프레임은 스텝 380에서 MFV로서 분류된다. 다르게 말하면, 만약 필드 면적이 낮은 임계치보다 높지만 플레이어 같은 오브젝트의 최대 면적이 최소 임계치보다 높지 않으면, 단일 프레임에 기초하여 믿을 수 있는(reliable) 결정이 이루어질 수 없다. 만약 프레임이 FV 또는 MFV로서 라벨링되지 않으면, 이것은 스텝 390에서 NFV로서 분류된다.

320x240 픽셀 축구 비디오들로 지향되는 바람직한 실시예에서, 임계값들 $T_{\text{field}}=0.4 \times H \times W$ (프레임 면적의 40%), $t_{\text{field}}=0.3 \times H \times W$ (프레임 면적의 30%), $T_{\text{med}}=600$, $t_{\text{max}}=100$, 및 $T_{\text{max}}=2400$ 이 사용되었는데, 여기서 H 및 W 는 픽셀들로 프레임의 높이 및 폭이다.

방법(300)의 스텝들은 시퀀스 내의 각각의 프레임에 대해 반복된다. 따라서, 프레임당 하나씩 뷰 필드의 분류들의 시퀀스가 획득된다.

도 4는 스텝들 320 및 330의 필드 검출 및 플레이어 같은 오브젝트 검출 프로세스들을 도해한다. 프레임(410)이 주어지면, 축구 필드 검출은 첫 번째로 이진 마스크(420)를 생성하여, 값 1을 갖는 모든 필드 컬러 픽셀들(all field-colored pixels) 및 필드에서의 오브젝트들을 포함하는, 값 0을 갖는 모든 비 필드 픽셀들을 분류한다. 플레이어들, 라인들, 및 공과 같은 필드상에서의 오브젝트들은 마스크에서 구멍들과 같이 나타나는데, 왜냐하면 이것들은 기대되는 필드의 컬러가 아니기 때문이다. 축구 필드의 경계들의 결정의 결과가 430에서 도시된다. 플레이어들, 라인들, 및 공으로부터의 마스크 내의 구멍들이 제거되어, 넓은 연속적 필드 마스크를 생성한다. 이미지(440)는 430에서 도시된 필드 경계 내의 비 필드 픽셀들의 연결된 세트들인 검출된 전경 오브젝트들을 도시한다. 이것들은, 예를 들어 플레이어들, 필드 라인들, 및 공의 이진 마스크들을 포함한다. 이미지(450)는 플레이어 같은 오브젝트들의 검출의 결과를 나타낸다. 면적, 밀집도, 또는 종횡비에 대한 임계치들을 만족하지 않는, 필드 라인들과 같은 오브젝트들은 제거되었다.

방법(300)의 프레임 레벨 분류는 몇몇의 에러가 있는 프레임 분류들을 일반적으로 산출할 것이다. 그러한 에러들의 효과들을 완화시키기 위해, 비디오는 "청크들(chunks)"이라고 불리는 연속적인 "유사하게 보이는(similar-looking)" 프레임들의 세트들로 세그먼트화된다. 청크들을 식별하는 프로세스는 도 5에 대하여 이하에서 기술된다. 도 6에 대하여 이하에서 기술되는 바와 같이, 각각의 청크의 뷰 필드는 그것의 구성 프레임들의 분류들의 통계치에 기초하여 전체로서 분류된다.

도 5는, 비디오 시퀀스(510)를 프레임들의 연속적인 시퀀스들의 컬러 히스토그램들의 유사성에 기초하여 다수의 청크들로 세그멘트화하기 위해 이용되는 방법(500)을 약술한다. 컬러 히스토그램은 컬러 값들의 여러 범위 내에서 프레임 내의 픽셀들의 수의 카운트들로 구성된다. 같은 장면으로부터의 프레임들은, 카메라가 일반적으로 일정한 컬러를 갖는 동일한 오브젝트들에 포인팅될 것이기 때문에, 유사한 히스토그램들을 가질 가능성이 있다. 장면의 변화는 뷰 필드 내에 다른 컬러들의 다른 오브젝트들을 위치시킬 것이고, 그에 따라 일반적으로 컬러 히스토그램을 변화시킨다.

스텝 515에서, 청크들 C의 리스트는 빈 상태이다. 구성 중에 있는 청크에서의 첫 번째 프레임인 시작 프레임 번호 j 는 비디오 시퀀스에서 첫 번째 프레임인 값 1로 초기화된다.

스텝 520에서, j 번째 프레임의 컬러 히스토그램이 계산된다. 바람직한 실시예에서, H_j 는 프레임 j 에서의 픽셀들의 그레이스케일 값들의 256 빈(bin) 히스토그램이다. 히스토그램 비교의 계산적 강도를 줄이기 위해 더 적은 수의 더 큰 히스토그램 빈들(bins)이 활용될 수 있다. j 번째 프레임의 컬러 히스토그램은 청크의 조립을 위한 기초로서 이용된다.

스텝 525에서, 루프가 프레임들 $i = j+1, \dots, N$ 에 대해 개시되는데, 여기서 N 은 시퀀스에서의 마지막 프레임의 번호이다. 프레임 j 를 뒤따르는 프레임들은 이것들이 프레임 j 로 시작되는 청크에 포함되어야 할지를 결정하도록 프레임 j 에 대한 유사성에 대해 한번에 하나씩 분석될 것이다.

스텝 530에서, i 번째 프레임의 컬러 히스토그램 H_i 가 스텝 520에서 이용된 동일한 기법을 이용하여 계산된다.

그러면 스텝 535에서, i 번째 프레임과 j 번째 프레임 간의 히스토그램 차이 d_{ij} 가 계산된다. 차이는 $d_{ij} = \|H_i - H_j\|_1$ 로서 계산될 수 있는데, 여기서 $\|\cdot\|_1$ 는 두 벡터 간의 절대값 차이들의 합(SAD: sum of absolute differences) 또는 1-놈(norm)이라고 한다.

단계 540에서, 프레임 i 가 프레임 j 로 시작되는 청크에 포함되도록 프레임 i 의 컬러 히스토그램이 프레임 j 의 그것과 충분히 유사한지 여부에 관하여 스텝 535로부터의 차이 d_{ij} 에 기초하여 결정이 이루어진다. 만약, 그것들의 컬러 히스토그램들 H_i 와 H_j 간의 거리 d_{ij} 가 임계치 T_{chunk} 보다 작다면, 프레임들은 충분히 유사하다고 여겨진다. 바람직한 실시예에서, $T_{chunk}=0.4$ 이다.

만약 $d_{ij} > T_{chunk}$ 라면, 즉, 만약 프레임 j 으로부터 프레임 i 의 컬러 히스토그램 차이가 프레임 i 가 청크에 포함되기에에는 너무 크다면, 스텝 545에서 간격[j, i-1]이 청크들의 리스트에 더해진다. 따라서, 현재 청크는 유사성 임계치를 충족시키는 마지막 프레임인 이전 프레임, 프레임 $i-1$ 에서 종료된다. 새로운 청크에 대한 시작 프레임 번호 j 가 스텝 565에서 현재 값 i 로 설정되고, 이 프로세스는 다음 청크를 만들기 위해 스텝 520으로 되돌아간다.

그러나, 만약 $d_{ij} \leq T_{chunk}$ 라면, 즉, 만약 프레임 i 는 현재 청크에 더해질 청크의 처음 프레임 j 와 충분히 유사하다고 여겨진다. 그러면 스텝 555에서 $i=N$ 인지(즉, 비디오에서 마지막 프레임에 도달하였는지) 여부에 관하여 결정이 이루어진다. 만약 아닐 경우, 프로세스는 스텝 525에서 루프의 시작으로 되돌아가고, 현재 청크에서의 포함에 대해 다음 프레임이 고려된다.

구성 중인 청크의 종료 프레임 번호는 그에 따라 충분히 유사하지 않은 프레임이 발견될 때까지 또는 마지막 프레임에 도달할 때까지 증가된다. 만약 마지막 프레임에 도달하면, 즉, $i=N$ 이면, 최종 프레임들을 구성하는 간격[j N]이 스텝 570에서 청크들(575)의 리스트에 추가되고 프로세스는 종료한다.

이런 프로세스의 마지막에서, 청크들 C의 리스트가 산출된다. 각각의 청크는 쌍(pair)[b e]으로서 나타내어지는데, 여기서 b는 청크의 시작 프레임이고 e는 종료 프레임이다.

도 6은 각각의 청크의 뷰 필드를 FV, MFV, 또는 NFV로서 분류하기 위해 이용되는 방법(600)을 도시한다. 바람직한 실시예에서, 청크 분류는 청크 내의 프레임들의 라벨들 중 다수결의 엄격한 형태에 기초하여, 프레임 레벨 분류에서 발생한 라벨링 에러들을 완화시킨다.

스텝 620에서, 입력 비디오 청크(610)의 각각의 프레임은 FV, MFV, 또는 NFV로서 분류된다. 이런 프레임 레벨 분류는 앞서 기술된 방법(300)을 이용하여 수행될 수 있다.

스텝 630에서, 청크에 대해 FV 프레임들의 비율(percentage)이 계산된다. 만약, 청크 내의 프레임들의 50% 이상이 스텝 640에서 FV인 것으로 결정된다면, 전체 청크는 스텝 650에서 FV로서 분류된다. 즉, 만약 구성 프레임들의 대다수가 FV이면, 청크는 FV로 여겨진다.

만약 FV 프레임들의 비율이 50%보다 높지 않으면, 청크 내의 MFV 프레임들의 비율이 스텝 660에서 계산된다. 프레임들 중 50%가 넘게 MFV인 것으로 스텝 670에서 결정되면, 청크는 스텝 680에서 MFV로서 분류된다. 만약 두 기준 모두 만족되지 않으면, 청크는 스텝 690에서 NFV로서 분류된다. 대안적 실시예에서, 청크들은, 프레임 카운트가 특정 임계치 미만이면 NFV로서 분류될 수 있다.

도 7은 전체 FV 장면 검출 프로세스에서 이용되는 방법(700)의 흐름도이다. 입력 비디오 시퀀스(710)가, 도 5에 대하여 앞서 기술된 바와 같이, 스텝 720에서 청크들로 세그멘트화된다. 스텝 730에서, 도 6에 대하여 기술된 바와 같이, 각각의 청크는 FV, MFV, 또는 NFV로서 분류된다.

스텝 740에서, 만약 MFV 청크가 FV 청크에 인접하게 있으면, 이것은 FV로서 재분류된다. 즉, 뷰 필드에 관한 결정이 스텝 730에서 이루어질 수 없다면, 청크는, 이것이 FV인 경우에 FV로 여겨질 것이다. $b_1=e_2+1$ 또는 $e_1=b_2-1$ 인 경우에, 청크 [b₁ e₁]는 청크 [b₂ e₂]에 인접해 있다고 한다. 바람직한 실시예에서, 단지 원래의 FV 청크들에 인접한 MFV 청크들만이 재분류되고, 다른 재분류된 FV 청크들에의 인접성에 기초한 재분류는 허용되지 않는다.

스텝 750에서, 모든 잔존하는 MFV 청크들은 NFV로서 재분류된다. 즉, 만약 청크의 뷰 필드와 관련한 결정이 스텝 620에서 이루어질 수 없고 또한 청크가 FV로서 식별된 청크에 인접하지 않다면, 청크는 FV로 재분류된다.

가정될 것이다.

스텝 760에서, 프로세스는 서로 인접하게 있는 모든 FV 청크들을 더 큰 청크들로 병합한다. $e_1=b_2-1$ 인, 두 개의 인접한 FV 청크 $[b_1 e_1]$ 및 $[b_2 e_2]$ 를 병합함으로써, 이러한 두 개의 청크를 제거하고 이것들 대신에 $[b_1 e_2]$ 를 더함으로써 C가 수정된다. 새로운 청크는 그것의 구성 부분들의 FV 라벨을 인계받는다. 이런 병합 프로세스는 반복적이고 또한 어떤 인접한 FV 청크들도 남아있지 않을 때까지 수행된다.

만약 FV 청크가 N_{min} 개의 프레임보다 더 적은 프레임들을 갖는다면, 이것은 NFV로서 재분류될 수 있다. 하나의 특정한 실시예에서, N_{min} 은 30개의 프레임으로 선택된다. 따라서 짧은 장면들의 프로세싱이 회피될 수 있다.

마지막으로, 스텝 780에서, C에서 모든 잔존하는 FV 청크들은 파 뷰 장면들로서 분류된다. 각각의 FV 청크의 시작 및 종료 프레임들은 대응하는 파 뷰 장면의 경계들(790)을 나타낸다. 유사하게, 모든 잔존하는 NFV 청크들은 (앞서 기술된 바와 같이 인접한 것들을 병합한 후에) 비 파 뷰 장면들로서 분류된다.

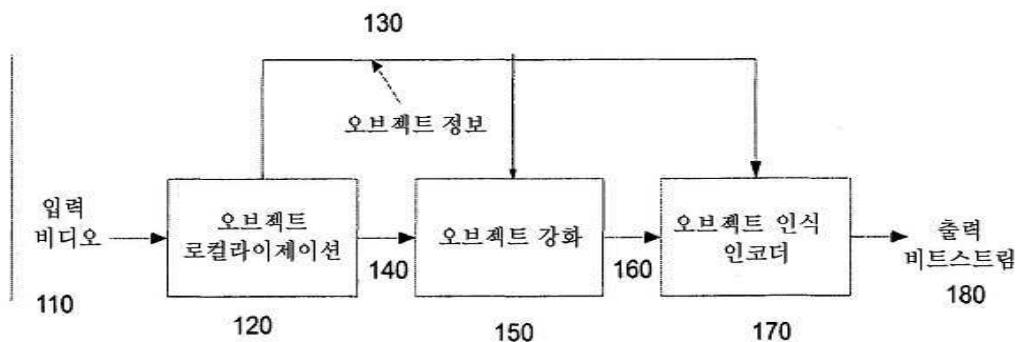
이 프로세스의 마지막에서, 파 뷰 장면들 S_{FV} 의 리스트 및 비 파 뷰 장면들 S_{NFV} 의 리스트가 획득된다. 각각의 장면은 쌍 $[b e]$ 에 의해 나타내어지는데, 여기서 b는 장면의 시작 프레임이고 e는 이것의 종료 프레임이다. 이러한 장면 경계들은 그 후 특정 뷰들에게만 적용 가능한 이미지 프로세싱 기법들의 적용을 제어하기 위해 도 2에서 도시된 오브젝트 하이라이팅 시스템에 의해 이용될 수 있다. 파 뷰 장면들은, 관심 대상인 오브젝트들을 강화시키고 오브젝트 인식 인코더(object-aware encoder)를 위해 오브젝트 메타데이터를 생성하는 오브젝트 로컬라이제이션 및 강화 모듈들에게 보내진다. 비 파 뷰 장면들은 오브젝트 하이라이팅이 없이 인코딩되도록 인코더에게 직접 보내질 수 있다.

본 발명이 특정적인 실시예에 관하여 기술되었지만, 본 발명의 범위 내에 포함될 수정들이 이루어질 수 있다고 이해될 것이다. 예를 들어, 여러 프로세싱 스텝이 개별적으로 구현되거나 또는 조합될 수 있고, 또한 범용 또는 전용의 데이터 프로세싱 하드웨어에서 또는 소프트웨어에서 구현될 수 있다. 방법의 전체적인 복잡도는, 필드 내의 오브젝트들이 판정에서 고려되는 기준들을 완화함으로써 줄일 수 있다. 예를 들어, 플레이어 같은 오브젝트들을 검출하는 대신에, 임계 면적보다 더 넓은 모든 오브젝트들이 고려될 수 있다. 또한, 청크 세그멘트화 동안에 히스토그램들을 계산하기 위해 그레이스케일 픽셀 값들을 이용하는 것 대신에, 풀 컬러 값들(예를 들면, RGB, YUV)을 이용하는 것도 가능하다. 추가로, SAD가 아닌 거리 척도들이 히스토그램들의 계산을 위해 이용될 수 있다. 프레임 분류를 위해 임계치에 기초한 기준을 이용하는 것 대신에, 분류자[예를 들어, 서포트 벡터 머신(support vector machine)]가 FV, MFV, 및 NFV 프레임들의 라벨링된 트레이닝 세트로부터 알려질 수 있다. 추가로, 제안된 방법은 관심 있는 움직이는 오브젝트들을 갖는 다른 스포츠들 또는 이벤트들에 적용될 수 있다. 마지막으로, 본 방법은 특화된 프로세싱을 위해 파 뷰가 아닌 장면들의 다른 유형들을 검출하기 위해 이용될 수 있다.

도면

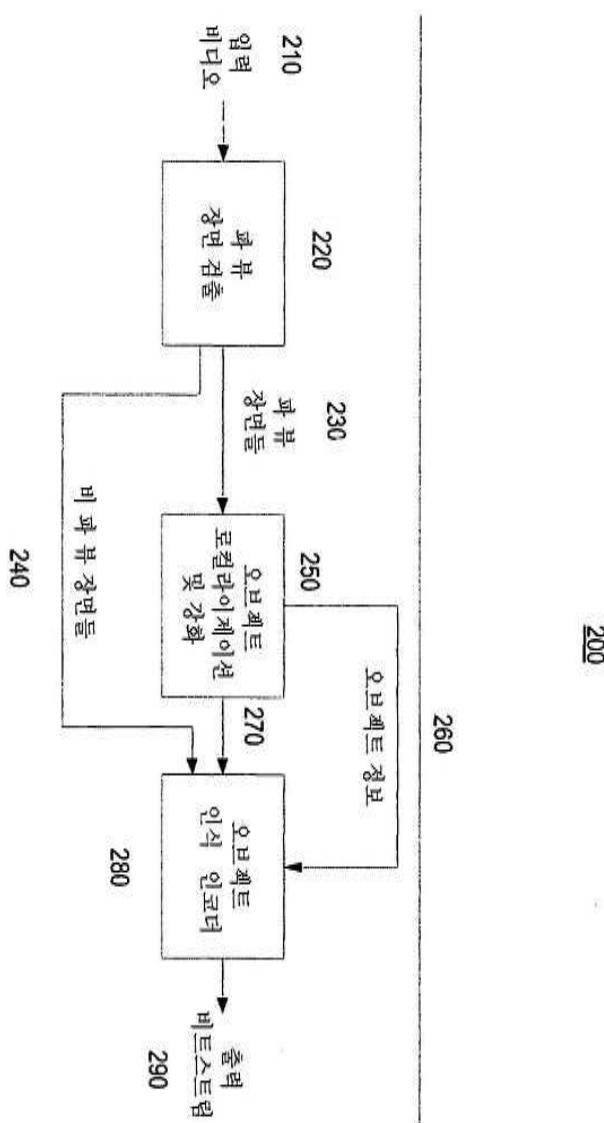
도면1

100

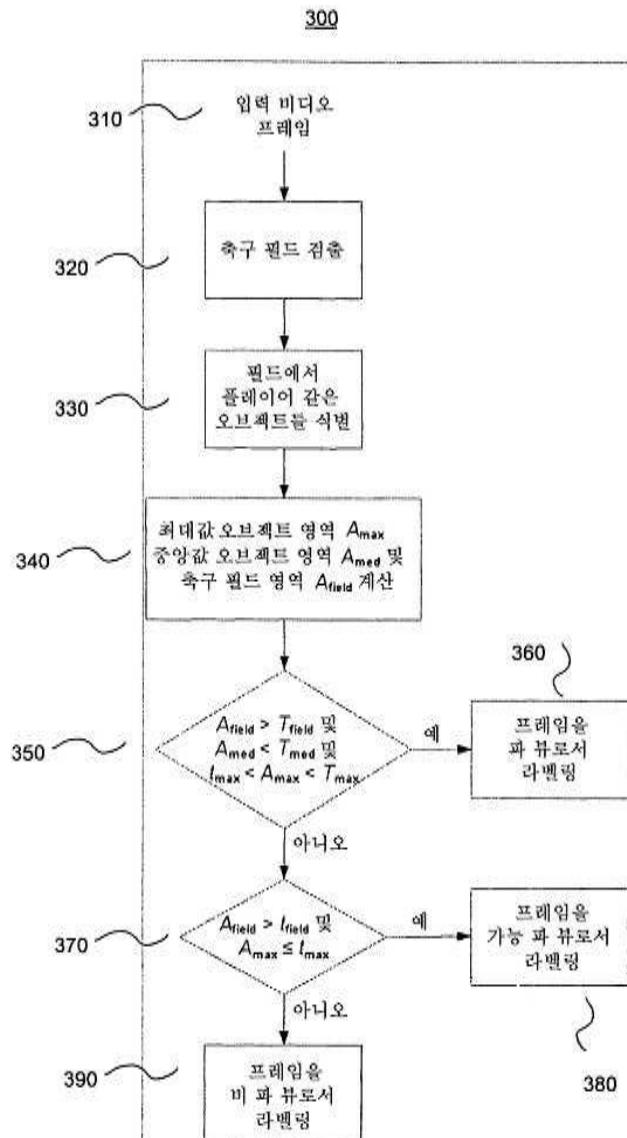


종래 기술

도면2

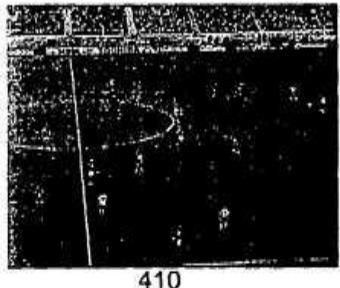


도면3

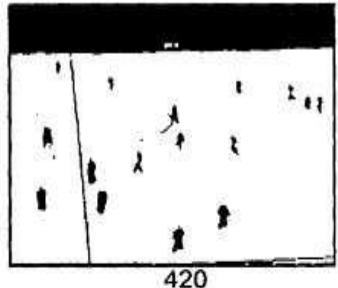


도면4

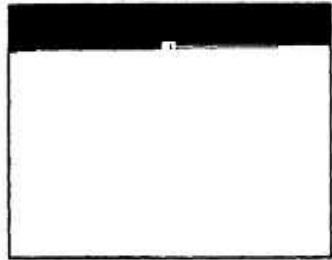
400



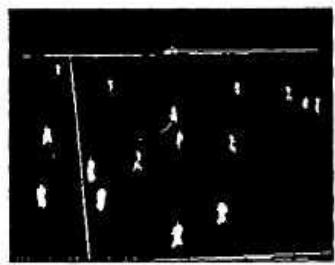
410



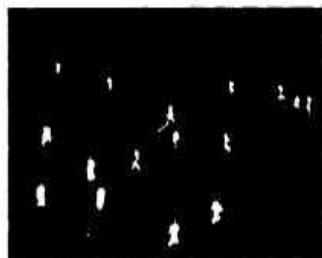
420



430

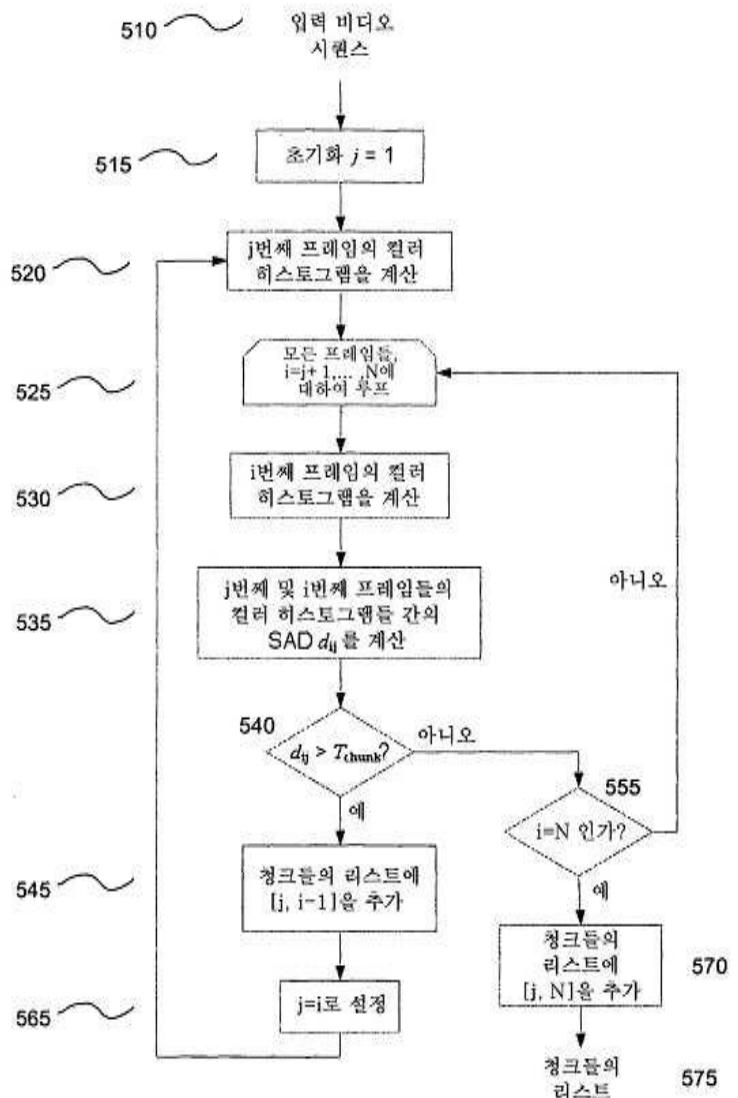


440

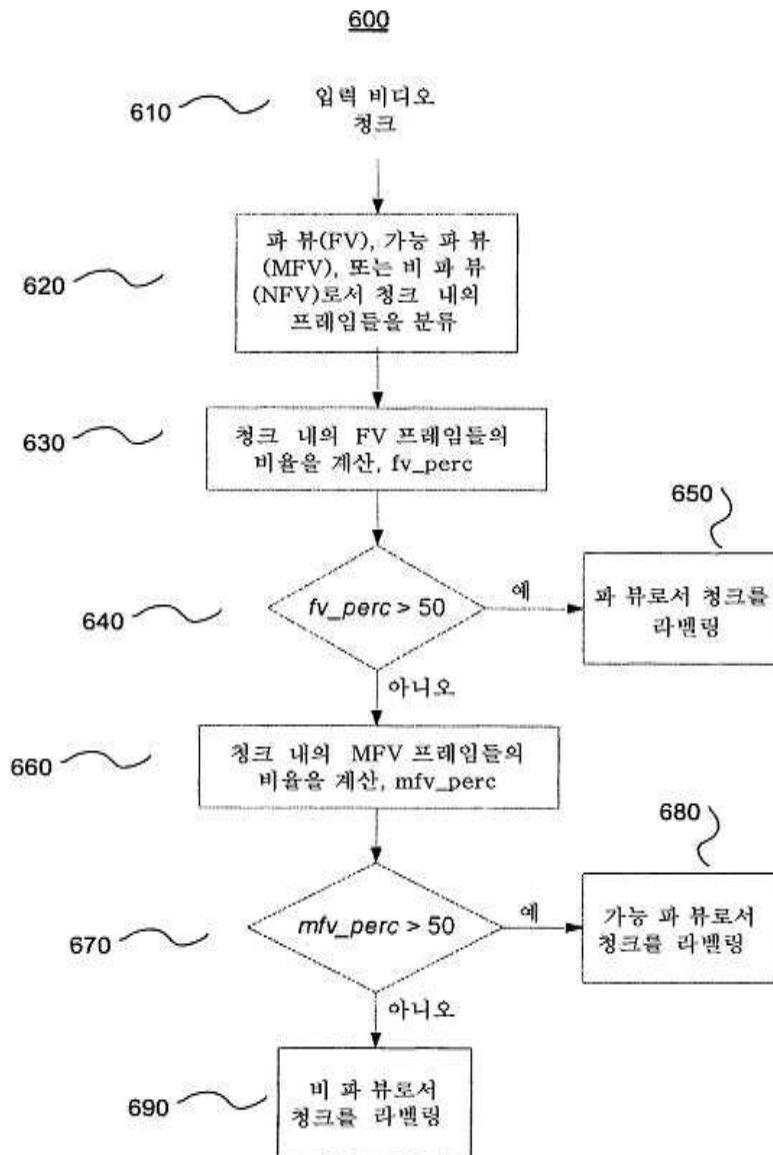


450

도면5

500

도면6



도면7

