



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0043848
(43) 공개일자 2023년03월31일

(51) 국제특허분류(Int. Cl.)
G06V 40/40 (2022.01) G06V 10/141 (2022.01)
G06V 10/80 (2022.01) G06V 10/82 (2022.01)
G06V 20/64 (2022.01) G06V 40/18 (2022.01)
(52) CPC특허분류
G06V 40/45 (2022.01)
G06V 10/141 (2023.01)
(21) 출원번호 10-2023-7002828
(22) 출원일자(국제) 2021년07월01일
심사청구일자 없음
(85) 번역문제출일자 2023년01월25일
(86) 국제출원번호 PCT/EP2021/068192
(87) 국제공개번호 WO 2022/003107
국제공개일자 2022년01월06일
(30) 우선권주장
20183434.8 2020년07월01일
유럽특허청(EPO)(EP)

(71) 출원인
스마트 아이 아베
스웨덴 413 27 괴테보르그 뢰르스타 룡가탄 28 베
(72) 발명자
릴제스트란드 사이몬
스웨덴 413 27 괴테보르그 뢰르스타 룡가탄 28 베
스마트 아이 아베 씨/오
존슨 케니스
스웨덴 413 27 괴테보르그 뢰르스타 룡가탄 28 베
스마트 아이 아베 씨/오
(뒷면에 계속)
(74) 대리인
박장원

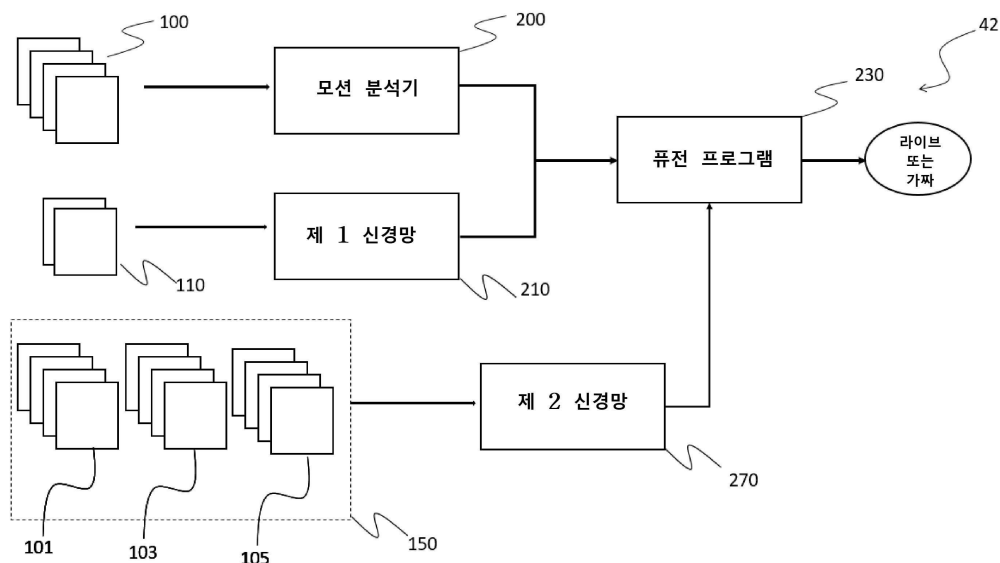
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 스푸핑 방지 시스템

(57) 요약

스푸핑 시도를 감지하도록 구성된 드라이버 모니터링 시스템(DMS)으로서, NIR 이미지 센서 모듈; 광원; 광원과 이미지 센서 모듈을 제어하여 서로 다른 조명 조건에서 시퀀스의 각 연속 이미지를 캡처하도록 구성된 제어기; 프로세서 작동 소프트웨어를 갖는 프로세서로서, 상기 소프트웨어는: 일련의 이미지에서 모션을 분석하고 모션 기반 결과를 출력하도록 작동가능한 모션 분석기; 상기 이미지 시퀀스로부터 연속된 2개 이상의 이미지의 세트를 적층하여 생성된 2개 이상의 채널 이미지로부터 하나 이상의 정보를 추출하고 공간적 결과를 출력하도록 동작하는 제 1 신경망; 및 모션 기반 결과와 공간적 결과를 병합하고 최종 결과를 출력하도록 작동가능한 퓨전 프로그램을 포함하고, 최종 결과는 대상이 라이브인지 가짜인지를 결정한다.

대표도 - 도2



(52) CPC특허분류

G06V 10/811 (2023.01)

G06V 10/82 (2022.01)

G06V 20/64 (2023.01)

G06V 40/18 (2022.01)

G06V 2201/12 (2022.01)

(72) 발명자

피너 존

스웨덴 413 27 괴테보르크 뢰르스타 롱가탄 28 베

스마트 아이 아베 씨/오

클린트베르그 스테판

스웨덴 413 27 괴테보르크 뢰르스타 롱가탄 28 베

스마트 아이 아베 씨/오

명세서

청구범위

청구항 1

스푸핑 시도를 검출하도록 구성된 운전자 모니터링 시스템(Driver Monitoring System: DMS)으로서, 상기 시스템은:

- 타겟의 부분의 이미지 프레임들의 시퀀스를 캡처하기 위한 NIR 이미지 센서 모듈;
- IR 펄스들을 방출하는 광원, 상기 IR 펄스들은 펄스 폭을 가지며 그리고 분리 시간 기간에 의해 시간적으로 분리되며;
- 상이한 조명 조건에서 각각의 연속 이미지 시퀀스를 캡처하도록 상기 광원 및 상기 이미지 센서 모듈을 제어하는 제어기, 상기 상이한 조명 조건은 상기 광원의 상기 IR 펄스들에 대응하고;
- 프로세서, 및 상기 프로세서에서 구현되는 프로세서 동작가능 소프트웨어를 포함하고, 상기 소프트웨어는:
 - o 상기 일련의 이미지에서 모션을 분석하고 모션 기반 결과를 출력하도록 동작가능한 모션 분석기;
 - o 상기 이미지 시퀀스로부터 적어도 2개의 연속적인 이미지들의 세트를 적층함으로써 생성되는 적어도 2개의 채널 이미지로부터 적어도 하나의 유형의 정보를 추출하고 그리고 공간적 결과(spatial result)를 출력하도록 동작가능한 제 1 신경망; 및
 - o 상기 모션 기반 결과와 상기 공간적 결과를 병합하고 최종 결과를 출력하도록 동작가능한 퓨전 프로그램을 포함하고, 상기 최종 결과는 표적이 라이브(live)인지 가짜(fake)인지를 나타내며,

상기 공간적 결과가 2D 형상과 3D 형상을 구별할 수 있도록 상기 적어도 하나의 유형의 정보를 추출하기 위해 상기 제 1 신경망은 3D 형상 분석을 수행하도록 트레이닝되는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 2

제1항에 있어서,

상기 퓨전 프로그램은 분류기이고, 상기 분류기는 모션 기반 결과 및/또는 공간적 결과에 기초하여 상기 최종 결과를 출력하는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 3

제1항에 있어서,

상기 모션 분석기는 안구 추적 프로그램으로서, 타겟의 이미지 프레임들의 눈 영역에서 눈 특징들(eye features)의 세트 및 반짝임 위치(glint position)를 식별하고, 상기 반짝임 위치 및 눈 특징들에 기초하여 시선 방향을 결정하도록 구성된 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 4

제1항 또는 제3항에 있어서,

상기 안구 추적 프로그램은 또한, 상기 이미지들의 시퀀스에서 타겟의 망막에 대한 반사를 검출 및 비교하고 그리고 결과를 퓨전 프로그램에 공급하도록 구성된 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 5

선행하는 청구항들 중 어느 한 항에 있어서,

상기 광원은 제 1 거리만큼 공간적으로 분리되고 IR 광을 교대로 방출하도록 구성된 적어도 2개의 IR 광원을 포함하는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 6

선행하는 청구항들 중 어느 한 항에 있어서,

상기한 조명 조건들 중 하나는 IR 펄스들 사이의 비조명 조건(non-illuminated condition)에 해당하는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 7

선행하는 청구항들 중 어느 한 항에 있어서,

상기 제 1 신경망은 상기 적어도 하나의 유형의 정보를 추출하기 위해 텍스처 정보 분석을 수행하도록 트레이닝되며, 상기 공간적 결과는 타겟이 라이브인지 또는 텍스처적으로 가짜인지에 대한 분류인 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 8

선행하는 청구항들 중 어느 한 항에 있어서,

상기 제 1 신경망이 퓨전 프로그램으로서 부가적으로 기능하여, 상기 공간적 결과가 최종 결과가 되도록 상기 모션 기반 결과가 제 1 신경망에 입력으로서 공급되거나 또는 상기 모션 분석기가 퓨전 프로그램으로서 부가적으로 기능하여, 상기 모션 기반 결과가 최종 결과가 되도록 상기 공간적 결과가 모션 분석기에 대한 입력으로서 공급되는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 9

선행하는 청구항들 중 어느 한 항에 있어서,

상기 프로세서 구현 소프트웨어는 제 2 신경망을 더 포함하고, 상기 제 2 신경망은 복수의 이미지 시퀀스로부터 제 2 유형의 정보를 추출하기 위해 장면 역학 분석을 수행하고 그리고 반복 결과(iteration result)를 출력하도록 트레이닝되며, 상기 반복 결과는 제 2 모션 기반 결과이고, 상기 제 2 신경망에 의해 출력되는 반복 결과는 상기 퓨전 프로그램에 공급되는 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 10

제9항에 있어서,

상기 제 2 신경망과 상기 제 1 신경망은 동일한 신경망인 것을 특징으로 하는 운전자 모니터링 시스템.

청구항 11

운전자 모니터링 시스템(DMS)을 사용하여 스푸핑 시도를 검출하는 방법으로서, 상기 방법은:

- a) IR 펄스들을 방출하는 단계, 상기 펄스들은 펄스 폭을 갖고 분리 시간 기간에 의해 시간적으로 분리되며;
- b) 타겟의 적어도 일부에 대한 이미지 프레임들의 시퀀스를 획득하는 단계, 상기 시퀀스의 각각의 연속적인 이미지는 상이한 조명 조건들에서 획득되고, 상기 상이한 조명 조건들은 IR 펄스들에 대응하며;
- c) 상기 이미지 프레임들의 시퀀스에서 모션을 분석하고 모션 기반 결과를 출력하는 단계;
- d) 상기 이미지 프레임들의 시퀀스로부터 적어도 2개의 이미지의 적어도 하나의 세트를 선택하는 단계, 상기 적어도 2개의 이미지는 연속적으로 촬영되며;
- e) 상기 적어도 2개의 이미지의 세트에서 관심 영역을 결정하는 단계;
- f) 상기 적어도 2개의 이미지의 세트를 적층함으로써 적어도 2개의 채널 이미지를 생성하는 단계;
- g) 상기 적어도 2개의 채널 이미지를 3차원 형상 분석을 수행하도록 트레이닝된 제 1 신경망에 공급하는 단계;
- h) 적어도 하나의 유형의 정보를 추출하는 단계;
- i) 상기 추출된 정보에 기초하여 2D 형상과 3D 형상을 구분하는 공간적 결과를 출력하는 단계; 및
- j) 타겟이 라이브인지 가짜인지를 결정하는 최종 결과를 획득하도록 상기 모션 기반 결과와 상기 공간적 결과를

병합하는 단계

를 포함하는 것을 특징으로 하는 방법.

청구항 12

제11항에 있어서,

상기 단계 c의 모션 기반 결과는 상기 제 1 신경망에도 제공되는 것을 특징으로 하는 방법.

청구항 13

제11항 또는 제12항에 있어서,

복수의 이미지 시퀀스로부터 제 2 유형의 정보를 추출하기 위해 장면 역학 분석을 수행하고 그리고 반복 결과를 출력하는 단계, 상기 반복 결과는 제 2 모션 기반 결과이며 그리고 및 최종 결과를 획득하기 위해 단계 j에서 상기 반복 결과를 병합하는 단계를 더 포함하는 것을 특징으로 하는 방법.

청구항 14

제11항 내지 제13항 중 어느 한 항에 있어서,

상기 모션 기반 결과, 상기 공간적 결과, 상기 반복 결과 및/또는 상기 최종 결과 중 적어도 하나의 신뢰 레벨을 계산하는 추가 단계를 포함하는 것을 특징으로 하는 방법.

발명의 설명

기술 분야

[0001] 본 발명은 스푸핑 시도를 탐지하도록 구성된 운전자 모니터링 시스템(DMS)에 관한 것이다.

배경 기술

[0002] 기술이 다양한 생체 인식을 기반으로 신원을 정확하게 판별하는 기술이 점차 향상되면서, 이들 다양한 생체 인식은 인증 요소로서 점점 더 대중화되고 있다. 생체 인식을 이용한 식별의 정확성은 인증 및 부인 방지(non-repudiation)를 보장할 수 있음을 의미하며, 생체 인식이 본질적으로 사람과 연결되어 있다는 사실은 암호를 기억하거나 키와 같은 물리적 토근을 휴대하는 것과 부담을 부여함이 없이 가용성을 보장한다. 하지만, 새로운 인증 요소가 대두됨에 따라 보안을 보장하기 위해 고려되어야 하는 새로운 잠재적 공격 벡터가 나타나고 있다. 하나의 공격 벡터는 스푸핑(spoofing)일 수 있으며, 이는 인증된 사용자의 생체 인식 데이터를 모방하는 위조된 생체 인식 데이터를 제공한다. 예를 들어, 시각적 얼굴 특징이 인증 요소로 사용되는 경우, 승인된 사람의 사진을 제시하거나 또는 승인된 사람과 유사한 마스크를 착용함으로써 얼굴 특징을 스푸핑할 수 있다. 따라서, 생체 인증 요소가 안전하기 위해서는, 이러한 공격에 대처할 수 있는 강력한 방법이 개발되어야 한다.

[0003] 얼굴 인식을 포함하여 생체 인식 인증 시스템에 대한 스푸핑 공격을 탐지하기 위한 다양한 접근법들이 이전에 제시되었으며, 그 중 일부는 NIR 스펙트럼의 이미지를 기반으로 하는 스푸핑 탐지에 기초한다. 2019년 9월 23일자 회로, 시스템, 통신 및 컴퓨터에 관한 국제 학술대회에서 A. Ghoneim, S. Mohamed, 및 A. Youssif 가 발표한 "Visible/infrared face spoofing detection using texture descriptors" 라는 논문에서, NIR 이미지와 가시광 이미지 둘다를 사용하는 안티-스푸핑을 위한 다중-스펙트럼 솔루션이 공개되었다. 또한, 상기 논문은 시각적 데이터와 NIR 데이터에 개별적으로 또는 함께 CNN(컨볼루션 신경망)을 활용하는 것을 포함하는 다양한 접근법들을 공개한다.

발명의 내용

해결하려는 과제

[0004] 하지만, 스푸핑 방지(anti-spoofing)(또는 "안티 스푸핑"이라 함) 시스템은 일상적인 상황, 예를 들어, 안면 인식과 같은 사용자에 대한 시각적 기반 인증이 매우 유용한 차량 내에서 쉽게 이용될 수 있고 원활하게 기능할 수 있어야 한다. 또한, 이러한 스푸핑 방지 시스템을 기존 하드웨어에 최소한의 비용으로 통합할 수 있는 것이 바람직하다.

과제의 해결 수단

- [0005] 따라서, 본 발명의 목적은 차량 내에 이미 존재하는 하드웨어와 호환가능한 용이한 스푸핑 방지 시스템을 제공하여, 보다 신뢰할 수 있는 시각적 식별을 가능하게 하는 것이다.
- [0006] 본 발명의 제 1 양태에 따르면, 이러한 목적 및 다른 목적은 스푸핑 시도를 검출하도록 구성된 운전자 모니터링 시스템(DMS: Driver Monitoring System)에 의해 달성되며, 상기 시스템은 다음을 포함한다: 타겟의 부분의 이미지 프레임들의 시퀀스를 캡처하기 위한 이미지 센서 모듈; IR 펄스들을 방출하는 광원, 상기 IR 펄스들은 펄스 폭을 가지며 그리고 분리 시간 기간에 의해 시간적으로 분리되며; 상이한 조명 조건에서 각각의 연속 이미지 시퀀스를 캡처하도록 상기 광원 및 상기 이미지 센서 모듈을 제어하는 제어기, 상기 상이한 조명 조건은 상기 광원의 상기 IR 펄스들에 대응하고; 프로세서, 및 상기 프로세서에서 구현되는 프로세서 동작가능 소프트웨어를 포함한다. 상기 소프트웨어는: 상기 일련의 이미지에서 모션을 분석하고 모션 기반 결과를 출력하도록 동작가능한 모션 분석기; 상기 이미지 시퀀스로부터 적어도 2개의 연속적인 이미지들의 세트를 적층함으로써 생성되는 적어도 2개의 채널 이미지로부터 적어도 하나의 유형의 정보를 추출하고 그리고 공간적 결과(spatial result)를 출력하도록 동작가능한 제 1 신경망; 및 상기 모션 기반 결과와 상기 공간적 결과를 병합하고 최종 결과를 출력하도록 동작가능한 퓨전 프로그램을 포함하고, 상기 최종 결과는 표적이 라이브(live)인지 가짜(fake)인지를 나타낸다.
- [0007] 상기 공간적 결과가 2D 형상과 3D 형상을 구별할 수 있도록 상기 적어도 하나의 유형의 정보를 추출하기 위해 상기 제 1 신경망은 3D 형상 분석을 수행하도록 트레이닝된다. 달리 말하면, 제 1 신경망은 적어도 2개의 연속된 이미지들의 세트의 상이한 조명 조건들에 의해 제공되는 3D 정보를 활용할 수 있고, 텍스처 및 형상 정보의 조합을 추출할 수 있다.
- [0008] 이미지 센서 모듈과 광원은 DMS 시스템의 일부로 통합될 수 있다. 이미지들은 타겟의 적어도 일부로부터 획득될 수 있는바, 예를 들어, 운전자의 눈 영역 또는 운전자의 얼굴 전체에서 획득될 수 있다.
- [0009] 일 실시예에 따르면, 퓨전 프로그램은 모션 기반 결과 및/또는 공간적 결과에 기초하여 최종 결과를 출력하는 분류기이다. 최종 결과는 DMS에 의해 획득된 이미지가 실제 살아있는 사람과 관련이 있는지(라이브) 아닌지(가짜)에 대한 표시로서 이해되어야 한다.
- [0010] 퓨전 프로그램은 유형 분류기일 수 있으며, 예를 들어, 규칙 기반 분류기 또는 대안적으로 미리 등록된 데이터에 대해 트레이닝될 수 있는 신경망 분류기일 수 있지만 이에 제한되지 않는다. 분류기는 모션 분석기 및 제 1 신경망 중 어느 하나 또는 둘 모두의 출력에 기초하여 결과를 결정할 수 있다. 예를 들어, 기본적인 규칙 기반 분류기는 단순히 모션 기반 결과와 공간적 결과 둘 다 또는 둘 중 하나를 최종 결과로서 전달할 수 있다. 대안적으로, 모션 기반 결과와 공간적 결과 모두가 타겟이 "가짜"라고 판단하면, 퓨전 프로그램은 "가짜"라는 최종 결과를 출력할 수 있으며, 그렇지 않고 모션 기반 결과 및/또는 공간적 결과가 타겟이 라이브라고 결정하면 퓨전 프로그램은 "라이브"라는 최종 결과를 출력할 수 있다.
- [0011] 따라서, 퓨전 프로그램은 두 가지 유형의 스푸핑 방지 시스템의 이점을 보장한다. 즉, 모션 분석기를 통해 그리고 제 1 신경망이 하나의 소형 저비용 시스템 내에서 통합되고 활용된다.
- [0012] 모션 분석기는 안구 추적기와 같은 추적기일 수 있으며, 출력된 모션 기반 결과는 시퀀스의 이미지 프레임 이전에 분석되었을 수 있는 이전 이미지 프레임의 결과에 의존할 수 있다.
- [0013] 상이한 조명 조건에서 이미지들의 시퀀스를 획득하는 이유는, 일반적으로 포토메트릭 스테레오로 지칭될 수 있는 상이한 조명 조건들에서 캡처된 다수의 2D 이미지를 기반으로 3D 표면이 적어도 부분적으로 재구성될 수 있기 때문이다. 이는 표면에서 반사되는 빛의 양이 표면, 광원 및 카메라 간의 각도 관계에 따라 달라질 수 있기 때문이다.
- [0014] 본 발명의 제 2 양태에 따르면, 운전자 모니터링 시스템(DMS)을 사용하여 스푸핑 시도를 검출하는 방법이 제공되며, 상기 방법은:
- [0015] a) IR 펄스들을 방출하는 단계, 상기 펄스들은 펄스 폭을 갖고 분리 시간 기간에 의해 시간적으로 분리되며;
- [0016] b) 타겟의 적어도 일부에 대한 이미지 프레임들의 시퀀스를 획득하는 단계, 상기 시퀀스의 각각의 연속적인 이미지는 상이한 조명 조건들에서 획득되고, 상기 상이한 조명 조건들은 IR 펄스들에 대응하며;
- [0017] c) 상기 이미지 프레임들의 시퀀스에서 모션을 분석하고 모션 기반 결과를 출력하는 단계;

- [0018] d) 상기 이미지 프레임들의 시퀀스로부터 적어도 2개의 이미지의 적어도 하나의 세트를 선택하는 단계, 상기 적어도 2개의 이미지는 연속적으로 촬영되며;
- [0019] e) 상기 적어도 2개의 이미지의 세트에서 관심 영역을 결정하는 단계;
- [0020] f) 상기 적어도 2개의 이미지의 세트를 적층함으로써 적어도 2개의 채널 이미지를 생성하는 단계;
- [0021] g) 상기 적어도 2개의 채널 이미지를 3차원 형상 분석을 수행하도록 트레이닝된 제 1 신경망에 공급하는 단계;
- [0022] h) 적어도 하나의 유형의 정보를 추출하는 단계;
- [0023] i) 상기 추출된 정보에 기초하여 2D 형상과 3D 형상을 구분하는 공간적 결과를 출력하는 단계; 및
- [0024] j) 타겟이 라이브인지 가짜인지를 결정하는 최종 결과를 획득하도록 상기 모션 기반 결과와 상기 공간적 결과를 병합하는 단계를 포함한다.
- [0025] 운전자 모니터링 시스템(DMS)이 교통 안전을 위해 예를 들어, 안구 추적을 통해 운전자의 주의를 모니터링하기 위한 표준 구성 요소가 되면서, 눈과 입과 같은 얼굴 특징의 움직임과 관련된 풍부한 정보를 제공한다. 본 발명자들은 차량에 있는 DMS의 기존 하드웨어를 사용함으로써 DMS로부터 이미 제공된 정보를 신경망 사용과 같은 스푸핑 사기에 대처하기 위한 기존 접근 방식과 결합하여, 오류율에 미치는 영향이 적고 통합 비용이 낮으며 실행 속도가 빠른 방식으로 차량에서 스푸핑 방지 시스템을 얻을 수 있음을 깨달았다.
- [0026] 단계 e는 모션 기반 결과에 기초하여 수행될 수 있으며, 이로부터 관심 영역의 위치를 포함하는 특정 위치들이 추출될 수 있다. 단계 e는 프레임 준비라고 지칭될 수 있으며, 여기서 연속적인 이미지 프레임들은 미리 결정된 크기의 관심 영역을 포함하기 위해 예를 들어, 자르기, 회전, 변환 등이 될 수 있다.
- [0027] 세트에 대해 단계 d에서 선택된 이미지들의 개수에 따라, 단계 f에서 각각의 첨부된 이미지가 추가 채널을 부가될 것이다.
- [0028] 이미지 프레임은 NIR 이미지 센서 모듈에 의해 획득될 수 있는 반면, IR 펄스는 적어도 하나의 광원에 의해 방출될 수 있다. 단계 c 내지 k 및 아래에 언급된 대안적인 및/또는 추가적인 단계들은 프로세서 구현 소프트웨어를 포함하는 프로세서에 의해 모두 수행될 수 있다.
- [0029] 이미지 센서 모듈은 카메라 어셈블리 또는 줄여서 카메라라고 지칭될 수 있다. 따라서, 본 출원의 문맥에서 이미지 센서 모듈 및 카메라라는 용어는 상호교환적으로 사용되며 달리 언급되지 않는 한 동일한 구성요소를 지칭함을 의미한다는 점에 유의해야 한다.
- [0030] 일 실시예에서 광원은 IR 광을 연속적으로 방출하도록 구성된다. 즉, 분리 시간 지속 시간은 0이 되도록 구성된다. 이러한 실시예에 따르면 조명 조건이 하나만 존재하기 때문에, 제 1 신경망은 텍스처 정보만을 추출할 수 있고 3D 정보는 추출할 수 없다.
- [0031] 대안적으로, IR 광원은 카메라로부터 일정 거리만큼 떨어진 적어도 2개의 IR 광원을 포함할 수 있으며, 카메라는 타겟으로부터 일정한 거리에 위치한다. 일 실시예에서, 2개의 광원은 제 1 거리만큼 공간적으로 분리되고 IR 광을 교대로 방출하도록 구성된다. 광원들은 가령, 카메라의 양쪽과 같이, 서로 가장 먼 거리에 위치할 수 있다. 대안적으로, 하나의 광원이 타겟에 더 가깝게 카메라 전면에 배치될 수 있는 반면에, 다른 하나의 광원은 카메라의 한쪽에 위치한다. 공간적 분리는 이미지가 획득되는 구별되는 조명 조건을 얻을 수 있도록 허용할 수 있다.
- [0032] 2 이상의 광원의 경우, 각 광원으로부터의 IR 펄스들의 펄스 폭이 동일할 수 있다. 대안적으로, 각각의 광원으로부터의 IR 펄스들의 펄스 폭이 동일하지 않을 수 있다.
- [0033] 이미지들의 시퀀스는 IR 펄스들의 펄스 폭에만 대응하는 상이한 조명 조건들에서 촬영될 수 있다. 후자는 하나 이상의 광원이 있는 실시예에 가장 잘 맞는 구성일 수 있다. 부가적으로 또는 대안적으로, 상이한 조명 조건들 중 하나는 IR 펄스들 사이의 비조명 조건에 대응한다. 즉, 특정 조명 조건이 광원에 의한 비조명이 되도록, 펄스들 사이의 시간적 분리 동안 일부 이미지가 획득될 수 있다. 이 경우 주변 조명이 상이한 조명 조건 중 하나를 제공할 수 있다.
- [0034] 이미지들의 시퀀스로부터의 연속적인 이미지들은 바람직하게는 그들 사이에서 짧은 시간적 거리를 가지며, 이는 그들 사이의 원치 않는 중요한 타겟 이동을 줄이기 위한 것이며, 따라서 네트워크는 라이브 타겟과 가짜 타겟을 구별하는데 집중할 수 있다. 그렇긴 하지만 네트워크는, 예를 들어, 타겟 및/또는 카메라 및/또는 눈꺼풀과 같

은 특징이 연속적인 프레임들 사이에서 움직이는 경우 이미지들의 기하학적 정렬을 통해 프레임들 사이의 임의의 기존 모션을 처리하도록 트레이닝될 수 있다.

[0035] 추가적으로 또는 대안적으로, 안구 추적 프로그램은 상기 이미지 시퀀스에서 타겟의 망막에 대한 반사를 검출 및 비교하고 결과를 퓨전 프로그램에 공급하도록 구성된다. 이러한 실시예에서, 타겟의 동공의 외관이 다를 수 있는 조명 조건이 생성 및 비교될 수 있다. 가까운 거리에서 광 펄스가 망막을 비추면, "적목 현상(red eye effect)"이라 지칭되는 효과가 발생할 수 있다. 이러한 실시예에서, 모션 분석기는 그 자체로 안면 모션을 반드시 또는 명시적으로 검출하지 않을 수 있지만, 상이한 조명 조건들에서 망막으로부터의 반사의 존재 또는 부존재를 검출할 수 있는바, 예를 들어, 모션 분석기는 동공이 어두운지 밝은지 감지할 수 있다. 이러한 실시예는 타겟의 2D 프린트와 같은 스푸핑을 검출하거나, 또는 빠르고 효율적인 방식으로 반복되는 비디오 루프를 탐지하는데 매우 유용할 수 있다. 추가적으로 또는 대안적으로, 동공이 밝거나 어둡다는 정보는 제 1 신경망에 입력으로 공급될 수 있다. 모션 분석기로부터의 정보가 입력으로서 제 1 신경망에 공급될 수 있는 대안적 실시예에 대한 더 많은 정보는 아래에서 보다 상세히 설명될 것이다.

[0036] 신경망은 이미지 분석에서 분류 문제를 위해 성공적으로 사용된다. 일 실시예에서, 제 1 신경망은 적어도 하나의 유형의 정보를 추출하기 위해 텍스처 정보 분석을 수행하도록 트레이닝되며, 공간적 결과는 진짜(live), 또는 텍스처적으로 가짜, 또는 예를 들어, 2D 사진의 일부, 가짜 물체 또는 재생 비디오 루프의 재생과 같은 스푸핑인 타겟에 대한 분류이다. 가짜 물체는 마스크, 인형 머리와 같은 가짜 인간 머리 또는 마네킹일 수 있다.

[0037] 텍스처 정보를 추출하고 분석하기 위해 신경망은 컬러, 라인(line), 실루엣 등과 같은 텍스처 정보를 인식하고 그리고 진짜 타겟과 가짜 타겟의 텍스처 및/또는 형상을 구별하도록 트레이닝될 수 있다.

[0038] 제 1 신경망은 중간 레벨에서 "실제(live)" 또는 "가짜"를 구별하도록 트레이닝될 수 있다. 이를 위해 적어도 2개의 연속적인 이미지들의 세트가 함께 결합되어 제 1 신경망에 공급될 수 있다. 제 1 신경망은 공급된 적층 이미지들로부터 텍스처 정보를 추출할 수 있다.

[0039] 일 실시예에서, 모션 기반 결과 및 공간적 결과는 상기 모션 기반 결과 및 공간적 결과를 직접 병합함으로써 최종 결과를 출력하기 위해 퓨전 프로그램에 대한 직접 입력으로서 공급된다.

[0040] 대안적으로, 일 실시예에서, 상기 공간적 결과가 상기 최종 결과가 되도록 제 1 신경망은 추가로 퓨전 프로그램으로서 기능하며 및/또는 모션 기반 결과, 즉 모션 분석기로부터의 검출 및 분석된 모션 정보가 제 1 신경망에 대한 입력으로서 공급된다. 이러한 실시예에서, 제 1 신경망은 타겟이 라이브인지 가짜인지에 대한 정보를 추출하기 위해 모션 데이터 분석으로부터의 정보에 대해 작용하도록 추가로 트레이닝된다. 이러한 실시예에서, 공간적 결과 및 최종 결과는 동일할 수 있으며 그리고 스푸핑 검출 시스템으로부터의 최종 출력된 결과와 동일할 수 있다. 이러한 실시예에서, 적어도 2개의 이미지 세트에서 관심 영역을 결정하기 위해 모션 기반 결과의 적어도 서브세트가 사용될 수 있는바, 예를 들어, 적어도 2개의 이미지 세트의 이미지 프레임을 잘라내고 그리고 타겟의 얼굴 영역을 추출하기 위해 사용될 수 있다.

[0041] 다른 대안적인 실시예에서, 모션 분석기는 제 1 신경망으로부터 출력된 상기 공간 결과가 모션 분석기에 입력으로서 공급되도록 퓨전 프로그램으로서 부가적으로 기능한다. 이러한 대안적인 실시예에서, 모션 분석기는 타겟이 라이브인지 가짜인지에 대한 정보를 추출하기 위해 제 1 신경망으로부터의 정보에 대해 작용하도록 추가적으로 트레이닝된다. 이러한 실시예에서, 모션 기반 결과와 최종 결과는 동일할 수 있으며, 스푸핑 검출 시스템에서의 최종 출력된 결과와 동일할 수 있다.

[0042] DMS 시스템에서, 물론 상기 NIR 이미지 센서 모듈에 의해 타겟의 적어도 일부에 대해 하나 이상의 이미지 프레임 시퀀스가 연속적으로 캡처될 수 있다. 이는 DMS 시스템의 비디오 획득 기능과 잘 일치할 수 있다. 이 경우, 모션 분석기는 상기 복수의 이미지 시퀀스의 각 시퀀스에서 모션을 분석하고 하나 이상의 모션 기반 결과를 출력하도록 동작할 수 있으며, 각각의 모션 기반 결과는 각각의 분석된 이미지 시퀀스에 대응한다. 추가적으로 또는 대안적으로, 제 1 신경망은 상기 2 이상의 이미지 시퀀스들 각각으로부터의 적어도 2개의 연속 이미지의 세트로부터 적어도 하나의 유형의 정보를 추출하고 그리고 하나 이상의 공간적 결과를 출력하도록 동작할 수 있으며, 각각의 공간적 결과는 각각의 이미지 시퀀스에 대응한다.

[0043] 이를 통해 시스템은 더 많은 수의 스푸핑 분석으로 인해 보다 안정적인 스푸핑 검출의 장점을 얻을 수 있다. 후자는 운전과 같은 사용 기간 동안 주어진 빈도로 수행될 수 있다.

[0044] 추가적으로 또는 대안적으로, 일 실시예에서 상기 프로세서 구현 소프트웨어는 제 2 신경망을 더 포함하고, 상기 제 2 신경망은 복수의 이미지 시퀀스로부터 제 2 유형의 정보를 추출하기 위해 장면 역학 분석을 수행하고

그리고 반복 결과(iteration result)를 출력하도록 트레이닝되며, 상기 반복 결과는 제 2 모션 기반 결과이고, 상기 제 2 신경망에 의해 출력되는 반복 결과는 상기 퓨전 프로그램에 공급된다.

[0045] 이미지 프레임들의 복수의 시퀀스들을 획득하는 경우, 연장된 시간 기간 동안 타겟을 분석함으로써 스푸핑 검출에 유용한 다른 정보를 추출할 수 있다. 이를 위해, 각 시퀀스의 이미지 프레임들은 하나의 스택으로 함께 첨부되며, 다중 스택의 각 스택은 하나의 이미지 시퀀스에 대응할 수 있다. 이것은 또 다른 분석 계층을 스푸핑 검출에 추가할 수 있으며, 여기서 불일치 또는 반복이 감지될 수 있으며 반복 결과가 출력되고 퓨전 프로그램에 공급될 수 있다.

[0046] 이것은 또한 스푸핑 시도를 검출하는데 도움이 될 수 있는바, 스푸핑 검출의 추가 계층이 계산의 오류율을 감소시키고 스푸핑 검출을 보다 안정적으로 만들 수 있다.

[0047] 예를 들어, 다중 스택이 공급될 수 있는 제 2 신경망으로서 리커런트 신경망(RNN: Recurrent Neural Network) 구조가 사용될 수 있다. 일 실시예에서, 제 2 신경망은 제 2 유형의 정보를 추출하기 위해 장면 역학을 분석하도록 트레이닝된다. 장면 역학은 예를 들어, 타겟의 주변 환경을 지칭할 수 있으며, 이는 차량의 실내 환경 및/또는 예를 들어, 차량 창을 통해 볼 수 있는 외부 환경을 지칭할 수 있다. 이것은 카메라에 제시되는 타겟의 2D 프린트와 같은 스푸핑 시도를 검출하는데 유용할 수 있으며, 여기서 장면은 절대 변하지 않는다. 추가적으로, 또는 대안적으로, 그리고 더 중요하게는 이것은 비디오 루프의 재생을 검출하는데 완전히 유용할 수 있으며, 여기서 장면 변화는 시간이 지남에 따라 반복되며, 제 1 신경망 분석(모션 기반 및 공간적 결과)과 함께 스푸핑 검출의 오직 하나의 계층, 즉 모션 분석은 검출을 하지 못할 수 있다.

[0048] 제 2 신경망은 모션 분석기 및/또는 제 1 신경망과 독립적으로 작동할 수 있다.

[0049] 실시예는 상기 모션 기반 결과, 상기 공간적 결과, 상기 반복 결과 및/또는 상기 최종 결과 중 적어도 하나의 신뢰 레벨을 계산하는 추가적인 단계를 포함한다. 이는 시스템, 측정치들 및/또는 분석에서 발생할 수 있는 오류를 해결하는데 도움이 될 수 있다.

도면의 간단한 설명

[0050] 본 발명은 본 발명의 현재 바람직한 실시예를 도시하는 첨부된 도면을 참조하여 보다 상세하게 설명될 것이다.

도 1a 및 도 1b는 스푸핑 시도를 감지하도록 구성된 DMS 시스템을 개략적으로 도시한다.

도 2는 스푸핑 방지 시스템의 실시예의 블록도를 도시한다.

도 3은 스푸핑 방지 시스템의 실시예의 블록도를 도시한다.

도 4는 스푸핑 시도를 탐지하는 방법의 흐름도를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0051] 다음의 상세한 설명에서, 본 발명의 몇몇 실시예가 설명될 것이다. 하지만, 다음을 유의해야 하는바, 다른 내용이 구체적으로 언급되지 않는 한, 상이한 실시예들의 특징들은 실시예들 사이에서 교환가능하며, 상이한 방식으로 조합될 수 있다. 이하의 설명에서, 본 발명의 보다 철저한 이해를 제공하기 위해 다수의 특정 세부사항들이 제시될 것이지만, 이러한 특정 세부사항이 없이도 본 발명이 실시될 수 있다는 것은 당업자에게 명백할 것이다.

[0052] 도 1a는 스푸핑 시도를 검출하도록 구성된, 즉 DMS에 의해 획득된 이미지가 예를 들어, 사진 또는 안면 마스크와 같은 일부 아티팩트가 아니라 실제 살아있는 사람과 관련되는지를 결정하도록 구성된 DMS 시스템(1)을 개략적으로 도시한다.

[0053] 예시된 일례에서, 시스템은 카메라(10) 및 2개의 적외선(IR) 광원(20a, 20b)을 포함한다. 카메라(10)는 활성 영역을 포함하는 이미지 센서를 일반적으로 포함할 수 있다. 활성 영역은 실질적으로 스펙트럼의 IR 부분에서 광선을 감지하도록 동작가능한 복수의 픽셀들을 포함한다. 광학 어셈블리(11)는 이미지 센서의 활성 영역 위에 배치된다. 광학 어셈블리(11)는 제어기(30)에 의해 제어되고, 이미지 시퀀스에 대응하는 노출 시퀀스에 이미지 센서를 연속적으로 노출하도록 구성된다. 광학 어셈블리(11)는 셔터, 복수의 픽셀들로부터 출력 신호를 획득하는 판독 회로, 복수의 픽셀들로부터의 출력 신호에 기초하여 일련의 이미지를 생성하도록 구성된 하나 이상의 이미지 프로세서를 포함할 수 있다. 카메라(10)는 IR 또는 근적외선(near-IR: NIR) 필터(13)를 더 포함한다. 필터(13)는 예를 들어, IR 펄스들 사이의 시간 기간의 시간적 분리 동안 이미지를 캡처하는 경우 인출가능하도록(withdrawable) 구성될 수 있다. 광원(20a, 20b)은 LED와 같은 고체 광원일 수 있다. 예시된 일례에서, 광원

(20a, 20b)은 약 850 또는 940nm(NIR)를 중심으로 하는 50nm 대역에 집중된 광 스펙트럼을 갖는 광을 방출하도록 구성된 LED이다. 필터(13)는 예를 들어, 간섭 필터와 같은 대역 통과 필터일 수 있다. 필터(13)는 광원(20a, 20b)의 발광 스펙트럼에 실질적으로 대응하는 통과 대역을 갖도록 구성될 수 있다. 따라서, 상술한 일례에서, 필터(13)는 약 825~875nm 또는 915~965nm의 통과 대역을 가져야 한다. 협대역 조명과 좁은 대역 통과 필터링의 조합으로 인해, 이미지 획득 시스템은 가령, 햇빛과 같은 주변광에 대해 덜 민감해진다.

[0054] 도 1a 및 도 1b에서, 서로 맞/또는 타겟(50, 9)으로부터 후자의 구성요소들의 일부의 상대적인 거리는 개략적으로 예시된다. 시스템(1)은 차량(미도시) 내에 설치되는 것으로 간주되며, 그리고 거리 Dc에서 카메라(10)가 운전자의 얼굴(50) 및/또는 운전자의 눈 영역(간단하게는 운전자로 지칭됨)(9)인 타겟을 향하도록 배치된다. 반면에, 2개의 IR 광원(20a, 20b)은 서로 거리 L 만큼 이격되어 카메라의 양측에 배치되는바, 이들 광원들은 카메라(10)로부터 측방(lateral) 거리 1a 및 1b를 가지며 그리고 운전자(50, 9)로부터는 거리 D1 및 D2를 갖는다. 도 1에 실시예 1에서, 2개의 광원(20a, 20b)은 카메라 및/또는 운전자(50, 9)에 대해 동일한 측방 평면 상에 위치되고 있다는 점에 유의한다. 카메라(10)와 광원(20a, 20b) 사이의 수직 거리는 Lc이다. 하지만, DMS 시스템(1)의 일부 실시예는 도 1에 도시된 실시예와 달리, 2개의 광원(20a, 20b)이 카메라(10) 및/또는 타겟(50, 9)에 대하여 비대칭적으로 배치된다는 점을 유의해야 한다. 카메라(10)는 타겟(50, 9)으로부터 일련의 이미지 프레임을 캡처하도록 동작할 수 있는 반면, IR 광원(20a, 20b)은 IR 광의 펄스들을 교대로 방출하도록 구성된다. 카메라(10) 및 IR 소스(20a, 20b)는 제어기(30)에 전자적으로 연결된다. 제어기(30)는 광원들(20a, 20b) 중 어느 하나로부터 방출되는 IR 펄스들의 펄스 폭에 대응되도록 및/또는 연속적인 각각의 펄스 사이의 시간적 분리로부터 이미지 프레임들의 획득을 제어하도록 구성된다.

[0055] 실제 일례로서, 광원들(20a, 20b)은 하나의 이미지 프레임을 획득하기 위해 충분한 조명을 제공하도록 선택된 듀티 사이클을 갖는 펄스에 의해 트리거된다. 전술한 바와 같이, 광원들(20a, 20b) 중 하나 또는 둘 모두가 이미지 프레임에 대해 활성화될 수 있다. 바람직하게는, 카메라(10)가 하나의 이미지를 획득하는데 필요한 시간에 펄스 폭이 대응함을 보장하도록, 광원(20a, 20b)은 듀티 사이클을 갖는 펄스 신호에 의해 제어된다. 이후, 카메라(10)는 이러한 펄스 신호와 동기화된다.

[0056] 제어기(30)는 프로세서(40)에 전자적으로 연결되며, 이는 프로세서 구현 소프트웨어를 포함할 수 있다. 시스템은 또한, 예를 들어 프로세싱 회로(15)에 의해 실행될 소프트웨어를 저장하는 가령, RAM 메모리와 같은 메모리를 갖는다.

[0057] 도 2 및 3은 프로세서가 소프트웨어 기능을 구현하는 방법의 예시적인 실시예(42, 44)를 보여준다.

[0058] 도 2의 프로세서 구현 소프트웨어 실시예(42)에서, 연속적인 이미지들(100)의 시퀀스는 시퀀스의 각 이미지가 상이한 조명 조건에 대응하도록, 타겟(50, 9)으로부터 카메라(10)에 의해 캡처되도록 구성된다. 시퀀스(100)는 DMS 시스템(1)의 카메라(10)에 의해 캡처된 비디오 클립의 이미지 프레임의 선택된 시퀀스일 수 있다.

[0059] 앞서 언급한 바와 같이, 서로 다른 조명 조건들에서 이미지 시퀀스를 획득하는 이유는, 서로 다른 조명 조건들에서 캡처된 다수의 2D 이미지들에 기초하여 3D 표면이 적어도 부분적으로 재구성될 수 있기 때문이다. 재구성 그 자체는 본 발명의 목적을 위해 이용되지 않을 수 있지만, 신경망(210)은 재구성 프로세스를 통해 3D 정보를 추출할 수 있으며 이는 라이브 타겟(50, 9)을 가짜로부터 구별하는데 유용할 수 있다.

[0060] 시퀀스(100)는 DMS 시스템의 모션 분석기와 같은 모션 분석기(200)에 공급되도록 구성된다. 모션 분석기는 모션 이벤트를 감지하고 분석하도록 구성되는바, 예를 들어, 가령 타겟(50, 9)의 눈 영역 및/또는 입 영역과 같은 얼굴의 움직임 검출 및 분석하고 및/또는 타겟(50, 9)의 머리 자세 및 시선을 추정하고 예를 들어, 졸음(drowsiness)을 추정하고, 그리고 타겟(50)의 자연스러운 움직임의 존재 또는 부존재를 출력하도록 구성된다.

[0061] 이것은 예를 들어, 카메라(10)에 고정된 타겟(50, 9)의 2D 프린트, 또는 예를 들어, 고정된 입을 갖는 전체 얼굴 역학에서 허용되지 않는 일부 얼굴 마스크와 같은 스푸핑 공격을 탐지하는데 매우 유용할 수 있다. 모션 분석기(200)는 시퀀스의 이미지들을 서로 비교할 수 있으며 그리고 타겟(50, 9)이 합리적인 및/또는 자연스러운 움직임을 나타내는지 여부를 검출할 수 있다.

[0062] 모션-기반 결과는 이러한 움직임들의 존재 또는 부존재를 결정할 수 있다. 이것은 여러 가지 방식으로 구현될 수 있는바, 예를 들어, 프로세서 구현 소프트웨어는 신뢰 레벨을 계산할 수 있으며 이는 다음을 의미할 수 있다. 예를 들어, 프로세서가 합리적인 움직임을 확실하게 검출한 경우에는 예를 들어 0이라는 점수가 부여되는 반면에, 합리적인 움직임이 전혀 검출되지 않는 경우에는 예를 들어, 1이라는 점수가 부여된다. 전형적으로, 분석된 이미지 시퀀스가 수신하는 점수는 0과 1 사이에 놓일 수 있다. 미리정의된 임계값이 0과 1 사이에 있을 수

있으며, 이러한 임계값 보다 낮은 출력된 모션 기반 결과는 합리적 및/또는 자연스러운 모션의 존재일 수 있으므로 스푸핑이 검출되지 않는 반면에 임계값 이상의 점수에 대해 출력된 모션 기반 결과는 합리적이고/또는 자연스러운 움직임이 부존재를 의미할 수 있으므로, 이는 스푸핑이 감지되었음을 의미한다.

[0063] 또한, 상이한 조명 조건들 하에서 획득된, 시퀀스(100)로부터의 적어도 2개의 연속적인 이미지들의 세트(110)는 프로세서 구현 소프트웨어에 의해 선택되도록 구성된다.

[0064] 세트(110)의 이미지들이 첨부되는바, 즉, 함께 적층되고, 제 1 신경망, 예를 들어, 프로세서 구현 소프트웨어에 의한 텍스처 및 3D 형상 분석을 위한 컨볼루션 신경망(210)에 공급된다. 이미지 스택, 또는 달리 말하면 2채널 이미지는 제 1 신경망(210)이 3D 정보를 추출할 수 있는 재료를 제공하고 및/또는 세트(110)의 각각의 2D 이미지들의 상이한 조명 조건들로부터 타겟(50, 9)의 적어도 부분적인 3D 재구성을 형성할 수 있는 가능성을 제공한다. 제 1 신경망(210)은 데이터를 다음과 같이 분류할 수 있다: 라이브, 2D 프린트, 또는 재생 비디오 루프, 또는 마스크, 또는 분석 후 임의의 다른 유형의 스푸핑 시도. 모션 분석기(200) 및 제 1 신경망(210)으로부터의 출력은 안티스푸핑 분류기 정보 퓨전 프로그램일 수 있는 퓨전(fusion) 프로그램(230)에 공급되도록 구성되며, 그리고 "라이브" 또는 "가짜"인 타겟(50)에 대한 최종 결과를 출력하도록 병합된다.

[0065] 컨볼루션 신경망(CNN: Convolutional Neural Network)은 예를 들어, 텍스처 정보와 같은 정보를 추출하고 분석하기 위해 사용될 수 있다. 텍스처 분석의 일례를 제공하기 위해, 예를 들면 자연스럽게 보이는 특징과 아티팩트를 구별하는 능력을 참조할 수 있다. 예를 들어, 일부 안면 마스크의 눈 개구부에서, 더블 에지들(double edges)이 검출될 수 있는데, 왜냐하면 마스크 개구부의 텍스처와 아래에 있는 실제 눈 사이에 약간의 불일치가 있을 수 있기 때문이다.

[0066] 상이한 조명 조건들에서 획득된 이미지는 타겟(50, 9)으로부터 상이한 정보를 나타낼 수 있다. 이미지 세트가 신경망(210)에 공급되기 전에, 적어도 2개의 연속적인 이미지들의 세트(110)는 서로 적층되도록 구성되며, 따라서 이들은 적어도 2개의 채널 이미지를 생성한다. 이것은 다음을 의미하는바, 세트(110)의 각 이미지 프레임의 정보가 깊이 방식으로 결합되어 타겟(50, 9)의 적어도 부분적인 3D 구성을 제공한다는 것을 의미한다.

[0067] 다음을 유의해야 하는바, 사용하기 전에, 신경망(210)이 트레이닝된다. 이를 위해, 예를 들어, 프레임들이 데이터 컬렉션에서 만들어진 레코딩으로부터 추출되고, 예를 들어, 미리 결정된 사이즈에서 관심 영역을 포함하도록 연속적인 이미지 프레임들이 자르기, 회전, 변환 등이 될 수 있도록 준비된다. 스푸핑 레코딩에서 추출된 프레임은 0으로 라벨링될 수 있는 반면, 라이브 레코딩에서 추출된 프레임은 1로 라벨링될 수 있다. 다음으로, 랜덤으로 선택된 대상들의 모든 레코딩들이 트레이닝 데이터로 사용되고 나머지는 검증 데이터로 사용되도록 프레임들이 분할된다. 일부 마스크들을 갖는 프레임들은 트레이닝에 사용되고 다른 마스크들을 갖는 프레임들은 검증에 사용되도록, 마스크 레코딩으로부터의 프레임들이 분할된다. 트레이닝 및 검증 전에, 트레이닝 데이터 세트의 통계를 기반으로 샘플 방식 및/또는 기능 방식으로 데이터가 정규화된다는 점에 유의해야 한다. 트레이닝 데이터 세트가 상대적으로 작은 경우, 트레이닝 데이터의 증가에 의해 변동이 발생할 수 있다. 활용되는 옵티마이저는 예를 들어 "Adam" 또는 "SGD" 분류기일 수 있다. 이진 교차-엔트로피가 비용 함수로 사용될 수 있다. 정확도 향상이 중단됨에 따라 학습률을 낮추기 위해, 학습률 스케줄러를 사용하여 반복적으로 속도를 업데이트할 수 있다.

[0068] 모션 기반 결과와 유사하게, 공간적 결과(spatial result)를 결정하고 출력하는 것은 여러 방식으로 구현될 수 있는바, 예를 들어 프로세서 구현 소프트웨어는 신뢰 레벨을 계산할 수 있으며 이는 예를 들어, 텍스처 불일치를 감지하거나 및/또는 2D와 3D를 확실하게 구분하는 경우 프로세서는 예컨대, 1이라는 점수를 부여할 수 있으며, 반면에 텍스처 불일치가 전혀 감지되지 않고 및/또는 타겟의 3D 구조가 확실하게 결정되는 경우 예를 들어, 0이라는 점수가 부여됨을 의미한다. 일반적으로, 이미지들의 분석된 시퀀스가 수신하는 점수는 0과 1 사이에 있을 수 있다. 미리정의된 임계값이 0과 1 사이에 있을 수 있으며, 이러한 임계값을 초과하면 출력된 공간적 결과가 텍스처 불일치의 존재 및/또는 타겟의 2D 구조의 존재일 수 있으므로 스푸핑이 검출될 수 있다. 반면에, 임계값 미만의 점수의 경우 출력된 공간적 결과는 최소한의 텍스처 불일치 또는 타겟의 3D 구조일 수 있으므로 스푸핑 시도는 발생하지 않았다. 예를 들어, 무시될 수 있는 시스템의 오류를 고려하는 경우, 신뢰 레벨을 계산하는 것이 도움이 될 수 있다.

[0069] 도 2의 실시예(42)에서, 스푸핑 검출의 추가 계층이 묘사된다. 다음을 유의해야 하는바, 스푸핑 검출은 이러한 추가 계층과 독립적으로 기능할 수 있으며, 이러한 계층의 존재는 이미 설명된 실시예(42)에 추가적인 정도의 스푸핑 검출을 부가할 수 있다. 추가 계층은 타겟(50)으로부터 이미지 프레임들(101, 103, 105)의 여러 시퀀스들을 캡처하는 것을 포함한다. 각 시퀀스(101, 103, 105)의 각 이미지는 서로 다른 조명 조건에서 촬영되도록

구성된다. 이러한 시퀀스는 비디오 획득과 마찬가지로 서로에 대해 시간적으로 연속적으로 획득되도록 구성될 수 있다. 각 시퀀스(101, 103, 105)의 이미지는 서로 결합되도록 구성되어 이미지의 다중 스택(150)을 생성한다. 이후, 이러한 다중 스택(150)은 프로세서 구현 소프트웨어에 의해 제 2 신경망(270)에 대한 입력으로서 공급된다. 제 2 신경망은 예를 들어, 리커런트 신경망(RNN: Recurrent Neural Network)일 수 있으며, 불일치들이나 반복들이 검출 및 분석되고 그리고 반복 결과들이 출력될 것이다. 그런 다음 출력은 퓨전 프로그램(230, 250)에 제공되며, 퓨전 프로그램(230, 250)은 반복 결과를 다른 출력과 병합하고 타겟(50)이 "라이브" 또는 "가짜"인 최종 결과를 출력하도록 구성된다.

[0070] 제 2 신경망(270)은 사용 전에 트레이닝될 수 있다는 것이 다시 주목된다.

[0071] 다시 말하지만, 최종 결과의 결정 및 출력은 예를 들어, 프로세서 구현 소프트웨어가 최종 계층에서 신뢰 레벨을 계산하는 것과 같이 여러 방식으로 구현될 수 있다.

[0072] 도 3의 대안적인 프로세서 구현 소프트웨어 실시예(44)에서, 도 2와 유사하게, 타겟(50)으로부터 캡처된 일련의 이미지(100)는 모션 분석기(200)로 공급되며, 모션 분석기(200)는 모션 이벤트를 검출 및 분석하고 그리고 타겟(50)의 자연스러운 움직임의 존재 또는 부존재를 출력하도록 구성된다. 하지만, 후자의 실시예(42)와 달리, 도 3의 실시예(44)에서, 모션 분석기(200)로부터의 출력은 시퀀스(100)로부터의 적어도 2개의 선택된 연속 이미지들(110)의 조합 세트와 함께 제 1 신경망(250)에 대한 입력으로 제공되며, 이는 입력된 데이터에 대한 텍스처 및 3D 형상 분석을 수행할 뿐만 아니라 퓨전을 또한 수행하여 타겟(50)이 "라이브" 또는 "가짜"인 최종 결과를 출력한다.

[0073] 도 2의 실시예(42)와 유사하게, 도 3의 실시예(44)에서 스푸핑 검출의 추가 계층이 설명된다. 다음을 다시 유의해야 하는바, 스푸핑 검출은 이러한 추가 계층과 독립적으로 기능할 수 있으며, 이러한 계층의 존재는 이미 설명된 실시예(42)에 추가적인 정도의 스푸핑 검출을 부가할 수 있다. 추가 계층은 실시예(42)의 단계들에 대해 언급된 것과 동일한 단계들을 수행하도록 구성되므로 해당 설명을 참조한다.

[0074] 도 4는 도 1 내지 도 3의 시스템을 사용하여 스푸핑 시도를 검출하는 방법의 순서도를 도시한다. 단계 S1에서 광원(20a, 20b)은 소정의 시간 기간 및 소정의 시간적 간격을 갖는 IR 펄스들을 방출한다. 단계 S2에서, 각각의 IR 펄스들과 동시에 및/또는 각각의 연속적인 IR 펄스 사이의 시간 기간의 간격에서, 카메라(10)는 교번하는 IR 펄스들에 의해 생성된 상이한 조건들에서 타겟(50, 9)의 이미지들(100)의 시퀀스를 획득한다. 따라서, S1 단계와 S2 단계는 서로 연계하여 수행되어야만 한다. 일례로서, 초당 120 프레임의 속도를 갖는 하나의 구성 및 광원(20a 및 20b)으로부터의 교번하는 IR 펄스들이 상이한 조명 조건들을 형성하는 2-플래시 구성에서, IR 펄스들의 펄스 폭, 따라서 이미지는 카메라의 획득 시간은 350 μ s 로 설정된다. 각 프레임에 대한 이용가능한 기간이 8333 μ s인 이러한 구성에서, 각 IR 펄스는 해당 기간의 500 μ s에서 시작되고 350 μ s 후에 종료된다. 따라서, 초당 60 프레임의 유효 프레임 속도가 전술한 2-플래시 구성에 대해 달성될 것이다. 120의 동일한 프레임 속도와 4-플래시 구성을 갖는 구성에서 초당 30프레임의 유효 프레임 속도에 도달할 수 있다.

[0075] 단계 S3에서, 프로세서 구현 소프트웨어의 모션 분석기(200)는 타겟(50, 9)의 자연스럽고/또는 합리적인 모션의 존재를 검출하기 위해 이미지들(100)의 시퀀스를 분석한다. 언급된 모션의 존재 또는 부존재를 검출하고 그리고 시스템 및/또는 측정치의 가능한 에러를 보상하기 위해, 모션 분석기(200)는 단계 S4에서 신뢰 레벨을 계산한다. 다음을 유의해야 하는바, 단계 S4는 도 4의 예시적인 실시예에 도시된 바와 같은 방법을 수행하기 위한 필수적인 단계가 아니며, 무시되거나 대안적으로 다른 수단으로 대체될 수 있다. 단계 S5에서 모션 분석기(200)는 모션 기반 결과를 출력한다.

[0076] 단계 S6에서 상이한 조명 조건들에서 획득된 적어도 2개의 연속적인 이미지들의 세트(110)가 이미지 시퀀스(100)로부터 선택된다. 단계 S7에서 적어도 2개의 이미지 세트(110)에서 관심 영역이 결정된다. 이러한 관심 영역은 예를 들어, 타겟(50, 9)의 눈 영역 또는 전체 얼굴 영역이 될 수 있다. 이러한 단계는 프레임 준비라고 지칭될 수 있으며, 여기서 연속적인 이미지 프레임은 미리결정된 원하는 사이즈에서 관심 영역을 포함하기 위해 예를 들어 자르기(cropping), 회전, 변환 등이 될 수 있다. 다음을 유의해야 하는바, 프레임 준비에서, 예를 들어, 자르기는 일반적으로 모션 기반 결과에 기초하여 수행된다. 트래커는 예를 들어, 자르기를 결정하는데 사용될 수 있는 얼굴 특징 좌표를 제공할 수 있다. 대안적으로, 얼굴 검출기를 활용하는 것과 같은 다른 방법이 후자를 달성하기 위해 사용될 수 있으며, 이는 모션 분석기와 독립적인 안티스푸핑 기능을 렌더링한다.

[0077] 단계 S8에서, 적어도 2개의 연속 이미지 세트(110)의 이미지들은 적어도 2개의 채널 이미지를 생성하기 위해 이전에 설명된 바와 같이 깊이 방식으로 서로 적층된다. 다음을 유의해야 하는바, 세트의 서로 다른 조명 조건들

에서 촬영한 각 연속 이미지는 채널을 추가한다. 예를 들어, 2개 이상의 연속 이미지, 예를 들어 서로 다른 조명 조건에서 촬영된 4개의 이미지 중 3개가 이미지 세트(110)에 대해 선택되고, 서로 적층되며, 그리고 3 채널 또는 4 채널 이미지가 생성될 것이다. 단계 S9에서, 스택은 제 1 신경망에 공급되며, 제 1 신경망은 텍스처 정보를 분석하거나 및/또는 2D 구조와 3D 구조를 구별함으로써 세트에 대한 분류를 수행할 것이다. 제 1 신경망은 단계 S10에서 텍스처 불일치 및/또는 타겟의 2D 구조의 존재에 대한 신뢰 레벨을 계산하고 그리고 시스템 및/또는 측정치의 가능한 에러를 보상한다. 다음을 유의해야 하는바, 단계 S4와 유사하게, 단계 S10은 도 4의 예시적인 실시예에 도시된 바와 같은 방법을 수행하기 위한 필수적인 단계가 아니며, 무시되거나 대안적으로 다른 수단으로 대체될 수 있다. 대안적으로, 신뢰 레벨을 계산하는 것이 사후 처리 단계(post-processing step)일 필요는 없다. 이 경우, 신뢰도는 제 1 신경망 자체에서 계산할 수 있으며, 이는 S10과 S9를 하나의 단계로 합칠 수 있음을 의미한다. 단계 S11에서, 제 1 신경망은 공간적 결과를 출력한다.

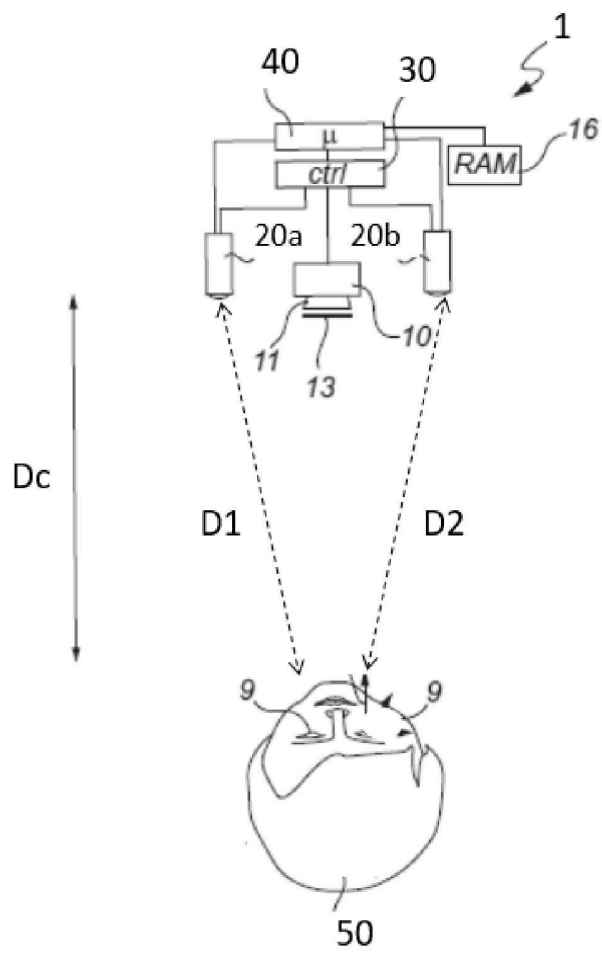
[0078] 순서도에 도시된 하나의 대안적인 경로에서, S5의 모션 기반 결과는 이미지 스택과 함께 S9에서 신경망에 대한 입력으로 공급될 수 있다. 대안적인 경로에서, S5의 모션 기반 결과는 S12에서 공간적 결과와 함께 퓨전 프로그램으로 공급되고, 이는 입력 정보를 병합하고 그리고 가능하면 S13에서 신뢰 레벨을 계산하고, 그리고 타겟(50)을 "라이브" 또는 "가짜"로 분류하는 최종 결과를 S14에서 출력할 것이다.

[0079] 단계 S21 - S26의 독립적인 세트는 도 4에 도시된 순서도의 왼쪽에 표시되며, 스푸핑 검출의 추가 계층을 상기 방법에 추가한다. S21에서, 이미지 프레임들(101, 103, 105)의 다수의 시퀀스가 타겟(50, 9)으로부터 캡처되며, 각 시퀀스(101, 103, 105)의 각각의 이미지는 상이한 조명 조건에서 촬영된다. 이러한 시퀀스는 시간적으로 서로 결과적으로(temporally consequent) 획득된다. S22에서 각 시퀀스(101, 103, 105)의 이미지는 서로 적층되어 S23에서 이미지의 다중 스택(150)이 된다. 이러한 다중 스택(150)은 S24에서 제 2 신경망(270)에 대한 입력으로서 공급된다. 제 2 신경망은 예를 들어, 리커런트 신경망일 수 있으며, 리커런트 신경망은 불일치 또는 반복들을 감지 및 분석할 수 있으며 그리고 S25에서와 같이 신뢰 레벨을 계산하고 반복 결과를 S26에서 출력할 것이다. 상기 출력은 S12의 퓨전 프로그램에 공급되고, S11 및 S5의 다른 출력과 병합되어, 타겟(50, 9)이 "라이브" 또는 "가짜"인 최종 결과를 출력한다.

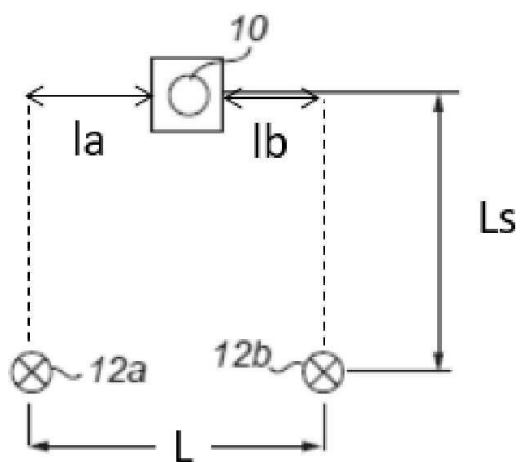
[0080] 당업자는 본 발명이 전술한 바람직한 실시예에 결코 제한되지 않는다는 점을 인식할 것이다. 오히려, 첨부된 청구 범위 내에서 많은 수정예들 및 변형예들이 가능하다. 예를 들어, DMS의 세부 사항은 여기에 개시된 것과 다를 수 있다. 또한, 펄스 폭 및 펄스 분리는 신경망에 대한 입력으로 사용되는 프레임 또는 프레임 세트의 수와 같이 구현예에 따라 조정될 수 있다.

도면

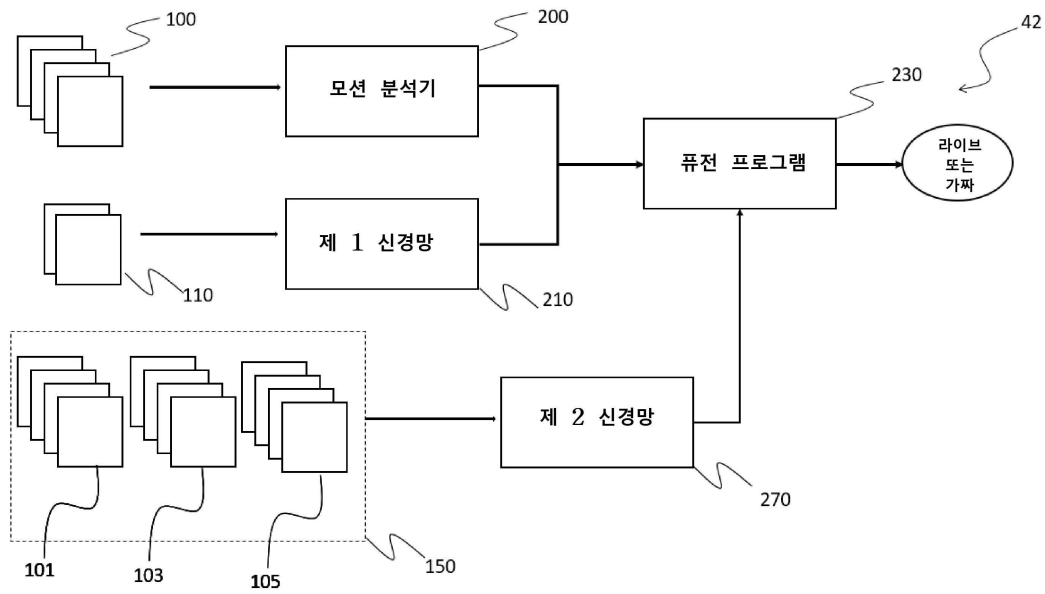
도면1a



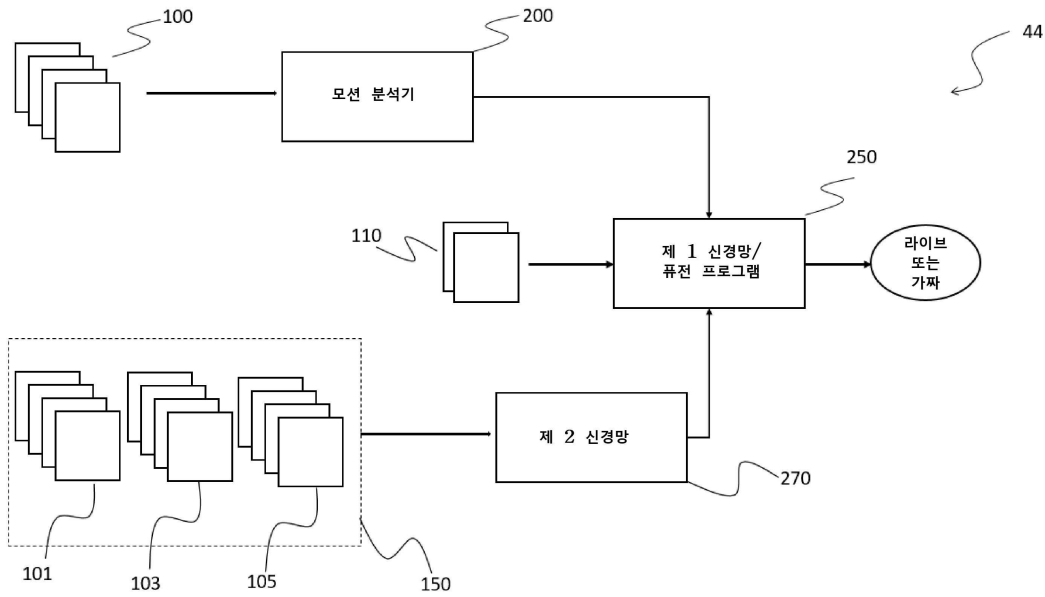
도면1b



도면2



도면3



도면4

