



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년01월05일

(11) 등록번호 10-2198470

(24) 등록일자 2020년12월29일

(51) 국제특허분류(Int. Cl.)  
G06T 7/20 (2017.01) G06T 7/50 (2017.01)  
G06T 7/70 (2017.01)

(52) CPC특허분류  
G06T 7/20 (2013.01)  
G06T 7/50 (2017.01)

(21) 출원번호 10-2019-0061769

(22) 출원일자 2019년05월27일

심사청구일자 2019년05월27일

(65) 공개번호 10-2020-0143534

(43) 공개일자 2020년12월24일

(56) 선행기술조사문헌  
Jun Liu외 3인, "SPatio-Temporal LSTM with Trust Gates for 3D Human Action Recognition", ECCV 2016, (2016.10.14.) 1부.\*

Mude Lin외 4인, "Recurrent 3D Pose Sequence Machines", CVPR 2017. (2017.07.26.) 1부.\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상훈

서울특별시 서대문구 연세로 50, 제3공학관 703호 (신촌동)

이경오

서울특별시 서대문구 연세로 50, 제3공학관 703호 (신촌동)

(74) 대리인

민영준

전체 청구항 수 : 총 12 항

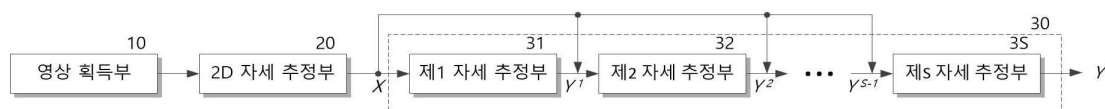
심사관 : 박금옥

(54) 발명의 명칭 관절 상호 의존성 기반 3D 자세 추정 장치 및 방법

## (57) 요약

본 발명은 미리 학습된 패턴 추정 방식에 따라 2D 영상에서 인간의 신체 중 기지정된 다수의 주요 관절의 위치를 나타내는 다수의 2D 특징을 추정하여 2D 특징맵을 획득하는 2D 자세 추정부 및 패턴 추정 방식이 미리 학습된 적어도 하나의 자세 추정부를 포함하여, 2D 특징맵의 다수의 2D 특징 각각에 대한 깊이 정보를 추정하여, 다수의 주요 관절의 3D 위치를 나타내는 다수의 3D 특징을 획득하여 3D 자세를 추정하는 3D 자세 추정부를 포함하고, 적어도 하나의 자세 추정부 각각은 다단 구조로 연결된 다수의 전파 추정부를 포함하고, 다수의 전파 추정부는 다수의 주요 관절의 위치를 기반으로 기지정된 순서에 따라 다수의 3D 특징을 순차적으로 획득하는 3D 자세 추정 장치 및 방법을 제공할 수 있다.

## 대표도



(52) CPC특허분류

**G06T 7/70** (2017.01)

G06T 2207/20044 (2013.01)

G06T 2207/20081 (2013.01)

G06T 2207/30196 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 20160004060042004

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원(한국연구재단부설)

연구사업명 정보통신방송연구개발사업

연구과제명 [주관/숙명여자대학교] (기반SW-창조씨앗2단계)SIAT형 CCTV 클라우드 플랫폼 기술  
개발 (3/5)

기 여 율 1/1

과제수행기관명 숙명여자대학교산학협력단

연구기간 2019.01.01 ~ 2019.12.31

공지예외적용 : 있음

---

## 명세서

### 청구범위

#### 청구항 1

미리 학습된 패턴 추정 방식에 따라 2D 영상에서 인간의 신체 중 기지정된 다수의 주요 관절의 위치를 나타내는 다수의 2D 특징을 추정하여 2D 특징맵을 획득하는 2D 자세 추정부; 및

신경망 연산 방식이 미리 학습되고 다단 구조로 연결된 인공 신경망으로 구현되는 다수의 자세 추정부를 포함하고, 상기 다수의 자세 추정부 중 제1 자세 추정부는 상기 2D 특징맵을 인가받아 학습된 신경망 연산에 의해 다수의 3D 특징을 획득하여 3D 특징맵을 출력하고, 나머지 자세 추정부는 상기 2D 특징맵과 이전 단의 자세 추정부에서 획득된 상기 3D 특징맵을 인가받아, 인가된 3D 특징맵의 3D 특징에 대한 신경망 연산에 의해 반복적으로 재추정하여 3D 특징맵을 업데이트하여 출력하며, 최종 단의 자세 추정부에서 획득된 3D 특징맵에 따라 상기 3D 자세를 추정하는 3D 자세 추정부를 포함하고,

상기 적어도 하나의 자세 추정부 각각은

다단 구조로 연결된 다수의 전과 추정부를 포함하고, 상기 다수의 전과 추정부는 상기 다수의 주요 관절의 위치를 기반으로 기지정된 순서에 따라 상기 다수의 3D 특징을 순차적으로 획득하고,

상기 다수의 전과 추정부는

상기 다수의 주요 관절 중 인간 신체의 중심으로부터 말단 방향으로 각각 기지정된 개수의 주요 관절에 대한 3D 특징을 순차적으로 획득하는 3D 자세 추정 장치.

#### 청구항 2

삭제

#### 청구항 3

삭제

#### 청구항 4

제1 항에 있어서, 상기 다수의 전과 추정부 각각은

상기 2D 특징맵, 상기 3D 특징맵 또는 이전 단의 전과 추정부에서 획득된 자세 깊이 큐를 인가받아 입력 특징으로 전달하는 특징 입력부;

미리 학습된 패턴 추정 방식에 따라 상기 입력 특징의 다수의 2D 또는 3D 특징 중 대응하는 적어도 하나의 특징에 대한 깊이 정보를 추정하여 3D 특징을 획득하는 깊이 추정부; 및

상기 입력 특징에 상기 깊이 추정부에서 획득된 3D 특징을 기지정된 방식으로 병합하여 자세 깊이 큐를 출력하는 깊이 병합부;를 포함하는 3D 자세 추정 장치.

#### 청구항 5

제4 항에 있어서, 상기 깊이 추정부는

패턴 추정 방식이 미리 학습된 LSTM(Long Short Term Memory)으로 구현되는 3D 자세 추정 장치.

#### 청구항 6

제4 항에 있어서, 상기 깊이 병합부는

상기 깊이 추정부에서 추정된 3D 특징으로 상기 입력 특징의 다수의 특징에서 대응하는 특징을 대체하여 상기 자세 깊이 큐를 획득하는 3D 자세 추정 장치.

#### 청구항 7

제4 항에 있어서, 상기 깊이 병합부는

상기 깊이 추정부에서 추정된 3D 특징을 상기 입력 특징과 결합하여 상기 자세 깊이 큐를 획득하는 3D 자세 추정 장치.

#### 청구항 8

제4 항에 있어서, 상기 깊이 병합부는

상기 입력 특징 중 2D 특징을 모두 제거하고, 상기 깊이 추정부에서 추정된 3D 특징을 추가하여 상기 자세 깊이 큐를 획득하는 3D 자세 추정 장치.

#### 청구항 9

제4 항에 있어서, 상기 3D 자세 추정부는

상기 다수의 전과 추정부 중 최종단의 전과 추정부에서 출력되는 상기 자세 깊이 큐를 미리 학습된 방식으로 맵핑하여 상기 3D 특징맵을 출력하는 3D 특징맵 출력부; 를 더 포함하는 3D 자세 추정 장치.

#### 청구항 10

제4 항에 있어서, 상기 3D 자세 추정부는

상기 3D 자세 추정 장치의 학습 시에 수학적식

$$\mathcal{L}_{3D}(Y_{pred}, Y_{GT}) = \sum_s \alpha_s \left[ \sum_t \frac{1}{|J|} \sum_j (Y_{pred}^{t,s} - Y_{GT}^{t,s})^2 + \lambda \sum_k (w_k^s)^2 \right]$$

(여기서  $Y_{pred}$ 는 추정된 3D 자세,  $Y_{GT}$ 는 미리 획득된 3D 자세의 진리값,  $S$ 는 자세 추정부의 개수,  $T$ 는 입력 영상의 프레임 길이,  $\alpha$ 는  $S$ 개의 자세 추정부 각각에 대한 가중치,  $\lambda$ 는 정규화 파라미터,  $w_k$ 는  $K$ 개의 전과 추정부 ( $41 \sim 4K$ ) 중  $k$ 번째 전과 추정부의 가중치 및  $J$ 는 2D 특징맵의 2D 특징 개수)

에 따라 계산되는 손실( $L_{3D}$ )이 역전파되어 학습되는 3D 자세 추정 장치.

#### 청구항 11

미리 학습된 패턴 추정 방식에 따라 2D 영상에서 인간의 신체 중 기지정된 다수의 주요 관절의 위치를 나타내는 다수의 2D 특징을 추정하여 2D 특징맵을 획득하는 단계; 및

미리 학습된 패턴 추정 방식에 따라 상기 2D 특징맵의 다수의 2D 특징 각각에 대한 깊이 정보를 추정하여, 상기 다수의 주요 관절의 3D 위치를 나타내는 다수의 3D 특징을 획득하여 3D 자세를 추정하는 단계를 포함하고,

상기 3D 자세를 추정하는 단계는

상기 다수의 주요 관절의 위치를 기반으로 기지정된 순서에 따라 상기 다수의 3D 특징을 순차적으로 획득하며,

상기 3D 자세를 추정하는 단계는

신경망 연산 방식이 미리 학습되고 다단 구조로 연결된 인공 신경망 중 제1 단의 신경망을 이용하여 상기 2D 특징맵을 인가받아 다수의 3D 특징을 추출하여 3D 특징맵을 획득하는 단계;

제1 단의 신경망을 제외한 나머지 신경망을 이용하여 기지정된 횟수만큼 상기 2D 특징맵과 이전 획득된 상기 3D 특징맵을 인가받아, 인가된 3D 특징맵의 3D 특징에 대한 신경망 연산에 의해 반복적으로 재추정하여 3D 특징맵을 업데이트하는 단계; 및

최종 획득된 3D 특징맵에 따라 상기 3D 자세를 추정하는 단계를 포함하고,

상기 3D 자세를 추정하는 단계는

상기 다수의 주요 관절 중 인간 신체의 중심으로부터 말단 방향으로 각각 기지정된 개수의 주요 관절에 대한 3D 특징을 순차적으로 획득하는 3D 자세 추정 방법.

## 청구항 12

삭제

## 청구항 13

삭제

## 청구항 14

제11 항에 있어서, 상기 3D 특징맵을 획득하는 단계 각각은

상기 2D 특징맵, 상기 3D 특징맵 또는 이전 획득된 자세 깊이 큐를 입력 특징으로 획득하는 단계;

미리 학습된 패턴 추정 방식에 따라 상기 입력 특징의 다수의 2D 또는 3D 특징 중 대응하는 적어도 하나의 특징에 대한 깊이 정보를 추정하여 3D 특징을 추정하는 단계; 및

상기 입력 특징에 획득된 3D 특징을 기지정된 방식으로 병합하여 자세 깊이 큐를 출력하는 단계; 를 포함하는 3D 자세 추정 방법.

## 청구항 15

제14 항에 있어서, 상기 자세 깊이 큐를 출력하는 단계는

추정된 3D 특징으로 상기 입력 특징의 다수의 특징에서 대응하는 특징을 대체하여 상기 자세 깊이 큐를 획득하는 3D 자세 추정 방법.

## 청구항 16

제14 항에 있어서, 상기 3D 특징맵을 획득하는 단계 각각은

최종 획득된 자세 깊이 큐를 미리 학습된 방식으로 맵핑하여 상기 3D 특징맵을 출력하는 단계; 를 더 포함하는 3D 자세 추정 방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 3D 자세 추정 장치 및 방법에 관한 것으로, 2D 영상으로부터 인간 관절의 상호 의존성에 기반하여 3D 자세를 추정할 수 있는 3D 자세 추정 장치 및 방법에 관한 것이다.

### 배경 기술

[0002] 인간 자세 추정은 컴퓨터 비전 분야에서 광범위하게 연구되어 왔다. 기존의 인간 자세 추정은 대부분 2D RGB 영상에서 인간 관절의 위치를 판별하여 2D 자세를 추정하는 방식으로 수행되었다. 그러나 최근에는 모션 캡처, 가상 훈련, 증강 현실, 재활 및 3D 그래픽 아바타를 비롯한 다양한 응용 분야에 적용 가능하도록 깊이 정보를 더 요구하는 3D 자세 추정에 대한 연구가 활발하게 수행되고 있다.

[0003] 한편 딥 러닝 기술을 이용하여 2D 영상에서 3D 자세를 추정하기 위한 시도 또한 계속되고 있다. 초기에 딥 러닝 기법을 이용한 3D 자세 추정은 깊이 정보가 미리 레이블된 학습 데이터를 기반으로 미리 학습되어 2D 영상에서 직접 3D 자세를 추정하도록 하였다. 그러나 인간의 신체 구조가 고려되지 않은 3D 자세 추정은 정확도 및 신뢰도 측면에서 성능의 한계가 있다는 문제가 있다.

## 선행기술문헌

### 특허문헌

[0004] (특허문헌 0001) 한국 공개 특허 제10-2018-0094532호 (2018.08.24 공개)

## 발명의 내용

### 해결하려는 과제

[0005] 본 발명의 목적은 2D 영상으로부터 인간 관절의 상호 의존성에 기반하여 정확하게 3D 자세를 추정할 수 있는 3D 자세 추정 장치 및 방법을 제공하는데 있다.

### 과제의 해결 수단

[0006] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 3D 자세 추정 장치는 미리 학습된 패턴 추정 방식에 따라 2D 영상에서 인간의 신체 중 기지정된 다수의 주요 관절의 위치를 나타내는 다수의 2D 특징을 추정하여 2D 특징맵을 획득하는 2D 자세 추정부; 및 패턴 추정 방식이 미리 학습된 적어도 하나의 자세 추정부를 포함하여, 상기 2D 특징맵의 다수의 2D 특징 각각에 대한 깊이 정보를 추정하여, 상기 다수의 주요 관절의 3D 위치를 나타내는 다수의 3D 특징을 획득하여 3D 자세를 추정하는 3D 자세 추정부; 를 포함하고, 상기 적어도 하나의 자세 추정부 각각은 다단 구조로 연결된 다수의 전과 추정부를 포함하고, 상기 다수의 전과 추정부는 상기 다수의 주요 관절의 위치를 기반으로 기지정된 순서에 따라 상기 다수의 3D 특징을 순차적으로 획득한다.

[0007] 상기 다수의 전과 추정부는 상기 다수의 주요 관절 중 인간 신체의 중심으로부터 말단 방향으로 각각 기지정된 개수의 주요 관절에 대한 3D 특징을 순차적으로 획득할 수 있다.

[0008] 상기 3D 자세 추정부는 다단 구조로 연결된 다수의 자세 추정부를 포함하고, 상기 다수의 자세 추정부 중 제1 자세 추정부는 상기 2D 특징맵을 인가받아 다수의 3D 특징을 획득하여 3D 특징맵을 출력하고, 나머지 자세 추정부는 상기 2D 특징맵과 이전 단의 자세 추정부에서 획득된 상기 3D 특징맵을 인가받아 다수의 3D 특징을 획득하여 3D 특징맵을 출력하며, 최종 단의 자세 추정부에서 획득된 3D 특징맵에 따라 상기 3D 자세를 추정할 수 있다.

[0009] 상기 다수의 전과 추정부 각각은 상기 2D 특징맵, 상기 3D 특징맵 또는 이전 단의 전과 추정부에서 획득된 자세 깊이 큐를 인가받아 입력 특징으로 전달하는 특징 입력부; 미리 학습된 패턴 추정 방식에 따라 상기 입력 특징의 다수의 2D 또는 3D 특징 중 대응하는 적어도 하나의 특징에 대한 깊이 정보를 추정하여 3D 특징을 획득하는 깊이 추정부; 및 상기 입력 특징에 상기 깊이 추정부에서 획득된 3D 특징을 기지정된 방식으로 병합하여 자세 깊이 큐를 출력하는 깊이 병합부; 를 포함할 수 있다.

[0010] 상기 깊이 추정부는 패턴 추정 방식이 미리 학습된 LSTM(Long Short Term Memory)으로 구현될 수 있다.

[0011] 상기 깊이 병합부는 상기 깊이 추정부에서 추정된 3D 특징으로 상기 입력 특징의 다수의 특징에서 대응하는 특징을 대체하거나, 상기 깊이 추정부에서 추정된 3D 특징을 상기 입력 특징과 결합하거나, 상기 입력 특징 중 2D 특징을 모두 제거하고, 상기 깊이 추정부에서 추정된 3D 특징을 추가하여 상기 자세 깊이 큐를 획득할 수 있다.

[0012] 상기 3D 자세 추정부는 상기 다수의 전과 추정부 중 최종단의 전과 추정부에서 출력되는 상기 자세 깊이 큐를 미리 학습된 방식으로 맵핑하여 상기 3D 특징맵을 출력하는 3D 특징맵 출력부; 를 더 포함할 수 있다.

[0013] 상기 3D 자세 추정부는 상기 3D 자세 추정 장치의 학습 시에 수학적

$$\mathcal{L}_{3D}(Y_{pred}, Y_{GT}) = \sum_s \alpha_s \left[ \sum_t \frac{1}{|J|} \sum_j (Y_{pred}^{t,s} - Y_{GT}^{t,s})^2 + \lambda \sum_k (w_k^s)^2 \right]$$

[0014] (여기서  $Y_{pred}$ 는 추정된 3D 자세,  $Y_{GT}$ 는 미리 획득된 3D 자세의 진리값,  $S$ 는 자세 추정부의 개수,  $T$ 는 입력 영상의 프레임 길이,  $\alpha$ 는  $S$ 개의 자세 추정부 각각에 대한 가중치,  $\lambda$ 는 정규화 파라미터,  $w_k$ 는  $K$ 개의 전과 추정부 (41 ~ 4K) 중  $k$ 번째 전과 추정부의 가중치 및  $J$ 는 2D 특징맵의 2D 특징 개수)에 따라 계산되는 손실( $\mathcal{L}_{3D}$ )이 역전파되어 학습될 수 있다.

[0016] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 3D 자세 추정 방법은 미리 학습된 패턴 추정 방식에

따라 2D 영상에서 인간의 신체 중 기지정된 다수의 주요 관절의 위치를 나타내는 다수의 2D 특징을 추정하여 2D 특징맵을 획득하는 단계; 및 미리 학습된 패턴 추정 방식에 따라 상기 2D 특징맵의 다수의 2D 특징 각각에 대한 깊이 정보를 추정하여, 상기 다수의 주요 관절의 3D 위치를 나타내는 다수의 3D 특징을 획득하여 3D 자세를 추정하는 단계; 를 포함하고, 상기 3D 자세를 추정하는 단계는 상기 다수의 주요 관절의 위치를 기반으로 기지정된 순서에 따라 상기 다수의 3D 특징을 순차적으로 획득한다.

### 발명의 효과

[0017] 따라서, 본 발명의 실시예에 따른 3D 자세 추정 장치 및 방법은 2D 영상에서 미리 학습된 방식으로 인간 관절 위치에 대한 2D 특징을 추출하고, 인간 관절의 상호 의존성에 따라 2D 특징에 대해 기지정된 순서로 깊이 정보를 획득하여 정확한 3D 자세를 추정할 수 있다. 또한 다단 구조로 3D 자세를 반복적으로 추정하여 더욱 정확한 3D 자세를 추정할 수 있도록 한다.

### 도면의 간단한 설명

[0018] 도 1은 본 발명의 일실시예에 따른 3D 자세 추정 장치의 개략적 구조를 나타낸다.  
 도 2는 도 1의 3D 자세 추정 장치의 3D 자세 추정의 개념을 나타낸다.  
 도 3은 도 1의 자세 추정부의 세부 구성의 일예를 나타낸다.  
 도 4는 도 3의 자세 추정부가 다수의 3D 특징을 추출하는 과정을 나타낸다.  
 도 5는 도 3의 전파 추정부의 세부 구성의 일예를 나타낸다.  
 도 6은 도 5의 깊이 병합부가 깊이 정보를 병합하는 방식을 설명하기 위한 도면이다.  
 도 7은 본 발명의 일실시예에 따른 3D 자세 추정 방법을 나타낸다.  
 도 8은 본 실시예에 따른 3D 자세 추정 장치의 성능을 시뮬레이션한 결과를 나타낸다.  
 도 9는 본 실시예에 따른 3D 자세 추정 장치를 실제 적용한 예를 나타낸다.

### 발명을 실시하기 위한 구체적인 내용

[0019] 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.

[0020] 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.

[0021] 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.

[0022] 도 1은 본 발명의 일실시예에 따른 3D 자세 추정 장치의 개략적 구조를 나타내고, 도 2는 도 1의 3D 자세 추정 장치의 3D 자세 추정의 개념을 나타낸다.

[0023] 도 1을 참조하면, 본 실시예에 따른 3D 자세 추정 장치는 영상 획득부(10), 2D 자세 추정부(20) 및 3D 자세 추정부(30)를 포함한다.

[0024] 영상 획득부(10)는 3D 자세를 추정하기 위한 다수 프레임을 갖는 입력 영상을 획득한다. 여기서 영상 획득부(10)는 도 2의 (a)에 도시된 바와 같이, 2D 영상(Raw Image)을 입력 영상으로 획득한다.

[0025] 이때 영상 획득부(10)는 단일 프레임의 2D 영상을 입력 영상으로 획득할 수 있으나, 일반적으로 단일 프레임의 2D 영상에서 정확한 3D 자세를 추정하는 데는 한계가 있다. 특히 입력 영상에 포함된 사람의 자세에 따라서는 폐색 영역이 존재할 수 있으며, 이 경우 정확한 자세를 추정할 수 없다. 이에 반해 연속하는 다수 프레임을 갖는 2D 영상을 입력 영상으로 획득하는 경우, 폐색으로 인한 문제를 극복할 수 있을 뿐만 아니라, 시간에 따른



자세 변화를 고려하여 3D 자세를 더욱 정확하게 추정할 수 있다.

- [0026] 따라서 본 실시예에 따른 3D 자세 추정 장치에서 영상 획득부(10)는 다수 프레임을 갖는 2D 영상을 입력 영상으로 획득하는 것으로 가정한다.
- [0027] 2D 자세 추정부(20)는 도 2의 (b)에 도시된 바와 같이, 영상 획득부(10)에서 획득된 입력 영상의 다수의 프레임 각각에서 다수의 2D 특징을 추출하여 2D 특징맵(X)을 3D 자세 추정부(30)로 전달한다. 여기서 다수의 2D 특징은 인간의 신체 중 기지정된 주요 관절의 위치를 나타낸다. 2D 자세 추정부(20)가 2D 영상에서 주요 관절의 위치에 대응하는 다수의 2D 특징을 추출하는 것은 관절의 위치가 인간 자세 추정에 가장 적합한 정보를 제공할 수 있기 때문이다. 일례로 2D 자세 추정부(20)는 각 프레임에서 인체의 머리, 목, 양어깨, 고관절, 팔꿈치, 손목, 무릎 및 발목의 14개의 주요 관절의 위치를 2D 특징으로 추출하여 2D 특징맵(X)을 획득할 수 있다.
- [0028] 2D 자세 추정부(20)는 미리 학습된 인공 신경망으로 구현될 수 있으며, 일례로 컨볼루션 신경망(Convolutional Neural Network: 이하 CNN)으로 구현될 수 있다. CNN과 같은 인공 신경망을 이용하여 2D 영상에서 주요 관절의 위치에 대응하는 다수의 2D 특징을 추출하는 것은 공지된 기술이므로, 여기서는 상세하게 설명하지 않는다.
- [0029] 3D 자세 추정부(30) 또한 미리 학습된 인공 신경망으로 구현되어, 2D 자세 추정부(20)에서 획득된 2D 특징맵(X)을 인가받고, 인가된 2D 특징맵(X)에 대해 학습된 방식으로 신경망 연산을 수행하여 2D 특징맵(X)의 다수의 2D 특징 각각에 대한 깊이 정보를 추정함으로써 도 2의 (c)에 도시된 바와 같이 3D 자세(Y)를 추정한다.
- [0030] 3D 자세 추정부(30)는 다단 구조로 연결되는 다수(여기서는 일례로 S개)의 자세 추정부(31 ~ 3S)를 포함한다. 다단 구조의 다수의 자세 추정부(31 ~ 3S) 중 제1 자세 추정부(31)는 2D 자세 추정부(20)로부터 2D 특징맵(X)을 인가받고, 다수의 2D 특징 각각에 대한 깊이 정보를 추정하고, 2D 특징에 추정된 깊이 정보가 반영된 다수의 3D 특징을 포함하는 제1 3D 특징맵( $Y^1$ )을 출력한다. 그리고 제1 자세 추정부(31)를 제외한 나머지 자세 추정부(32 ~ 3S)는 각각 이전 단계에서 출력되는 3D 특징맵( $Y^1 \sim Y^{S-1}$ )과 2D 특징맵(X)을 함께 인가받아, 인가받은 3D 특징맵( $Y^1 \sim Y^{S-1}$ )에서 다수의 2D 특징에 대해 추정되었던 깊이에 대해 신경망 연산을 수행하여 다시 깊이를 추정함으로써 3D 특징맵( $Y^2 \sim Y^S$ )을 출력한다.
- [0031] 여기서 3D 특징맵( $Y^1 \sim Y^S$ )은 2D 특징맵(X)의 다수의 2D 특징에 깊이 정보가 추가된 3차원 좌표 정보를 포함하는 다수의 3D 특징을 추출하므로, 3D 특징맵( $Y^1 \sim Y^S$ )은 인간의 3차원 자세에 대한 정보를 나타내는 3D 자세로 볼 수 있다. 본 실시예에서 3D 자세 추정부(30)의 다수의 자세 추정부(31 ~ 3S) 중 마지막 단의 자세 추정부(3S)에서 출력되는 3D 특징맵( $Y^S$ )을 3D 자세(Y)로 출력한다.
- [0032] 3D 자세 추정부(30)는 하나의 자세 추정부만을 포함하여 3D 자세(Y)를 획득하도록 구성될 수 있으나, 다단 구조의 다수의 자세 추정부(31 ~ 3S)를 포함하여 구성되는 경우, 제1 자세 추정부(31)를 제외한 나머지 자세 추정부(32 ~ 3S)가 이전단계에서 추정한 3D 특징맵( $Y^1 \sim Y^{S-1}$ )과 2D 특징맵(X)을 기반으로 다수의 3D 특징에 대한 깊이를 더욱 정확하게 추정할 수 있다.
- [0033] 한편, 본 실시예에서 3D 자세 추정 장치가 입력 영상에서 도 2의 (b)와 같이 다수의 2D 특징을 우선 추출하여 2D 특징맵(X)을 획득하고, 이후 2D 특징맵(X)으로부터 도 2의 (c)와 같은 3D 자세(Y)를 추정하는 것은 2D 영상에서 곧바로 3D 자세를 추정하는 것보다 정확하게 3D 자세를 추정할 수 있도록 하기 위함이다.
- [0034] 도 3은 도 1의 자세 추정부의 세부 구성의 일례를 나타내고, 도 4는 도 3의 자세 추정부가 다수의 3D 특징을 추출하는 과정을 나타낸다.
- [0035] 도 3을 참조하면, 다수의 자세 추정부(31 ~ 3S) 각각은 다수(여기서는 일례로 K개)의 전과 추정부(41 ~ 4K)와 하나의 3D 특징맵 출력부(50)를 포함할 수 있다.
- [0036] 직렬로 연결된 다단 구조의 다수의 전과 추정부(41 ~ 4K) 각각은 패턴 추정 방식이 미리 학습된 인공 신경망을 포함하여 구현되어 2D 특징맵(X)의 다수의 2D 특징 또는 이전 단계에서 전달되는 3D 특징맵( $Y^1 \sim Y^{S-1}$ )의 다수의 3D 특징 중 대응하는 특징의 깊이 정보를 추정하고, 추정된 깊이 정보가 기지정된 방식으로 반영된 자세 깊이 큐(pose depth cue)( $\hat{X}^1 \sim \hat{X}^K$ )를 출력한다.
- [0037] 다단 구조의 다수의 전과 추정부(41 ~ 4K) 각각은 일례로 LSTM(Long Short Term Memory)을 포함하여 구현될 수



있다. LSTM은 인공 신경망의 하나로서, 순환 신경망(Recurrent Neural Network: RNN)이 장기간(Long Term) 특징을 반영할 수 있도록 개선한 구조를 갖는다. 즉 LSTM은 시간 상관 특성을 가지므로, 이전 프레임에 대해 추출된 3D 특징이 고려된 현재 프레임에 대한 3D 특징을 추출할 수 있다. 따라서 다수 프레임을 갖는 입력 영상에서 높은 신뢰도를 갖고 3D 특징을 추출할 수 있다.

[0038] 전과 추정부(41 ~ 4K)의 개수는 2D 특징맵(X)의 2D 특징의 개수에 대응하는 개수로 구비될 수 있으며, 각각 지정된 2D 특징 또는 3D 특징에 대한 깊이 정보를 추정한다. 여기서 각각의 전과 추정부(41 ~ 4K) 각각은 하나의 2D 특징 또는 3D 특징에 대한 깊이 정보를 추정할 수도 있으나, 다수의 특징에 대응하여 다수의 2D 특징 또는 3D 특징에 대한 깊이 정보를 동시에 추정할 수도 있다.

[0039] 도 4를 참조하면, 다수의 전과 추정부(41 ~ 4K) 중 제1 전과 추정부(41)는 (a)에 도시된 바와 같이, 6개의 2D 특징 또는 3D 특징에 대한 깊이 정보를 추정하는 반면, 이후, 나머지 전과 추정부(42 ~ 4K) 각각은 (b) 내지 (i)에 도시된 바와 같이, 하나의 2D 특징 또는 3D 특징에 대한 깊이 정보를 순차적으로 추정할 수 있다.

[0040] 본 실시예에서 자세 추정부(31 ~ 3S)가 다단 구조의 다수의 전과 추정부(41 ~ 4K)를 포함하여 2D 특징맵(X)의 다수의 2D 특징 또는 이전 단계에서 전달되는 3D 특징맵( $Y^1 \sim Y^{S-1}$ )의 다수의 3D 특징 전체에 대한 깊이 정보를 동시에 추정하지 않고, 순차적으로 깊이 정보를 추정하는 것은 2D 특징 또는 3D 특징과 더불어 이전 추정된 깊이 정보에 기초하여 각 특징에 대한 정확한 깊이 정보를 인간 관절의 상호 의존성(joint interdependency)에 기반하여 정확하게 추정할 수 있도록 하기 위함이다.

[0041] 특히 도 4에 도시된 바와 같이, 제1 전과 추정부(41)가 인간의 신체 중 몸통의 주요 관절에 대응하는 다수의 2D 특징에 대한 깊이 정보를 우선 추정하고, 이후 나머지 전과 추정부(42 ~ 4K) 각각이 이전 추정된 깊이 정보에 기반하여 신체 말단 방향의 관절에 대응하는 2D 특징에 대한 깊이 정보를 순차적으로 추정함으로써 정확한 3D 자세를 추정할 수 있다.

[0042] 인간의 신체 구조상 몸통의 목, 어깨 및 고관절과 몸통에 근접한 머리는 독립적으로 움직이기 어려울 뿐만 아니라 근접하여 위치하는 반면, 신체 말단 방향의 팔꿈치와 손목, 무릎 및 발목과 같은 관절은 상대적으로 독립적으로 움직일 수 있다. 일례로 오른쪽 손목이 움직이는 경우, 오른쪽 팔꿈치와 오른쪽 어깨가 일부 함께 움직일 수 있으나, 왼쪽 팔꿈치나 손목 및 양 다리의 관절은 전혀 영향을 받지 않는다.

[0043] 따라서 제1 전과 추정부(41)는 도 2의 (a)에 도시된 바와 같이, 3D 자세 추정을 위한 기준위치로서 몸통의 주요 관절의 깊이 정보를 일괄로 우선 추정하고, 나머지 전과 추정부(42 ~ 4K) 각각은 독립적으로 움직일 수 있는 나머지 주요 관절에 대한 2D 특징 중 대응하는 하나의 특징에 대한 깊이 정보를 순차적으로 추정함으로써, 다수의 3D 특징을 정확하게 추정할 수 있다. 특히 몸통 부분으로부터 신체의 말단 방향의 특징에 대한 깊이 정보를 순차적으로 추정함에 따라 이전 획득된 특징의 깊이 정보가 반영되도록 하여 정확한 깊이 정보를 추정할 수 있다.

[0044] 그러나 경우에 따라서는 모든 전과 추정부(41 ~ 4K)가 각각 대응하는 하나의 특징에 대한 깊이 정보를 추정하도록 구성되거나, 대응하는 둘 이상의 깊이 정보를 추정하도록 구성될 수도 있다.

[0045] 여기서는 도 4에 도시된 바와 같이, 2D 자세 추정부(20)가 추출한 머리, 목, 어깨, 고관절, 팔꿈치, 손목, 무릎 및 발목의 14개의 주요 관절 중 제1 전과 추정부(41)가 몸통과 연관된 주요 관절인 머리, 목, 어깨 및 고관절의 6개의 2D 특징에 대한 깊이 정보를 추출하고, 나머지 전과 추정부(42 ~ 4K)가 오른쪽 팔꿈치, 왼쪽 팔꿈치, 오른쪽 무릎, 왼쪽 무릎, 오른쪽 손목, 왼쪽 손목, 오른쪽 발목 및 왼쪽 발목의 순서로 하나씩의 2D 특징에 대한 깊이 정보를 추출하는 것으로 가정한다. 따라서 다수의 자세 추정부(31 ~ 3S) 각각은 9개의 전과 추정부(41 ~ 49)를 포함할 수 있다.

[0046] 한편, 3D 특징맵 출력부(50)는 다수의 전과 추정부(41 ~ 4K) 중 마지막 단의 전과 추정부(4K)에서 출력되는 자세 깊이 큐( $\bar{X}^N$ )를 인가받아 미리 학습된 방식으로 맵핑하여 3D 특징맵( $Y^1$ )을 출력한다. 여기서 3D 특징맵 출력부(50)는 일례로 완전 연결 레이어(Fully Connected layer)로 구현될 수 있다.

[0047] 도 5는 도 3의 전과 추정부의 세부 구성의 일례를 나타내며, 도 6은 도 5의 깊이 병합부가 깊이 정보를 병합하는 방식을 설명하기 위한 도면이다.

[0048] 도 5를 참조하면, 다수의 전과 추정부(41 ~ 4K) 각각은 특징 입력부(61), 깊이 추정부(62) 및 깊이 병합부(63)를 포함한다. 특징 입력부(61)는 대응하는 특징맵을 인가받아 깊이 추정부(62)와 깊이 병합부(63)로 전달한다. 특징 입력부(61)는 전과 추정부(41 ~ 4K)가 다수의 자세 추정부(31 ~ 3S) 중 제1 자세 추정부(31)의 제1 전과

추정부(41)이면, 2D 자세 추정부(20)에서 획득된 2D 특징맵(X)을 인가받고, 나머지 자세 추정부(32 ~ 3S)의 제1 전과 추정부(41)이면, 이전 자세 추정부(31 ~ 3(S-1))에서 획득된 3D 특징맵( $Y^1 \sim Y^{S-1}$ )을 인가받고, 제1 전과 추정부(41)가 아닌 나머지 전과 추정부(42 ~ 4K)이면, 이전 전과 추정부(41 ~ 4(K-1))에서 출력되는 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^{K-1}$ )를 인가받아 입력 특징으로서 깊이 추정부(62)와 깊이 병합부(63)로 전달한다.

[0049] 깊이 추정부(62)는 미리 학습된 인공 신경망으로 구현되어, 특징 입력부(61)에서 전달된 입력 특징에서 대응하는 특징의 깊이 정보를 추정하여 깊이 정보가 반영된 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 깊이 병합부(63)로 전달한다. 깊이 추정부(62)는 상기한 바와 같이, 시간 상관 특성을 갖는 LSTM으로 구현될 수 있다.

[0050] 깊이 병합부(63)는 특징 입력부(61)에서 전달된 입력 특징에 깊이 추정부(62)에서 전달된 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 기지정된 방식으로 병합하여 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득하여 출력한다.

[0051] 여기서 깊이 병합부(63)는 도 6에 도시된 바와 같이 일예로 3가지 방식으로 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득할 수 있다.

[0052] 깊이 병합부(63)는 첫번째 방식으로 1)과 같이 특징 입력부(61)에서 전달된 입력 특징에서 2D 특징( $x_1 \sim x_{14}$ )을 모두 제거하고, 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )을 추가할 수 있다. 상기한 바와 같이, 제1 전과 추정부(41)가 기지정된 개수의 주요 관절에 대한 깊이 정보를 추출한 경우, 나머지 전과 추정부(42 ~ 4K)는 학습된 패턴 추정 방식에 따라 2D 특징에 무관하게 추출된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )으로부터 나머지 3D 특징을 추출할 수도 있다. 즉 깊이 병합부(63)는 전달된 입력 특징에서 2D 특징( $x_1 \sim x_{14}$ )이 존재하면, 2D 특징을 모두 제거하고, 3D 특징만을 남기며, 깊이 추정부(62)에서 새로이 추정된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )을 추가하여 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득할 수 있다.

[0053] 두번째 방식으로 깊이 병합부(63)는 2)와 같이 특징 입력부(61)에서 전달된 입력 특징에서 다수의 2D 특징( $x_1 \sim x_{14}$ )을 그대로 유지하고, 깊이 추정부(62)에서 새로이 추정된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )을 추가하여 결합(Concatenation)하는 방식으로 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득할 수 있다.

[0054] 세번째 방식에서 깊이 병합부(63)는 3)과 같이 깊이 추정부(62)에서 새로이 추정된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )으로 특징 입력부(61)에서 전달된 입력 특징의 다수의 특징( $x_1 \sim x_{14}$ ) 중 대응하는 특징( $x_1, x_2, x_3, x_6, x_9, x_{12}$ )을 대체(Replacement)하여, 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득할 수 있다.

[0055] 본 실시예에 따른 3D 자세 추정 장치에서 다수의 2D 특징을 추출하여 2D 특징맵(X)을 획득하는 2D 자세 추정부(20)에 대해서는 이미 많은 연구가 수행되어 있다. 따라서 기존에 미리 학습된 2D 자세 추정부(20)를 그대로 적용할 수 있다. 그러나 3D 자세 추정부(30)가 2D 특징맵(X)에서 3D 자세(Y)를 추정하기 위해서는 학습이 수행되어야 한다.

[0056] 3D 자세 추정부(30)는 각각 인공 신경망을 포함하여 구성되는 다수의 전과 추정부(41 ~ 4K)를 포함하는 다수의 자세 추정부(31 ~ 3S)를 가지므로, 3D 자세 추정부(30)를 학습시키기 위한 손실 함수( $L_{3D}$ )는 수학적 1과 같이 표현될 수 있다.

# 수학식 1

$$\mathcal{L}_{3D}(Y_{pred}, Y_{GT}) = \sum_s \alpha_s \left[ \sum_t \frac{1}{|J|} \sum_j (Y_{pred}^{t,s} - Y_{GT}^{t,s})^2 + \lambda \sum_k (w_k^s)^2 \right]$$

[0057]

[0058]

여기서  $Y_{pred}$ 는 본 실시예의 3D 자세 추정 장치에서 추정된 3D 자세로서 지정된 주요 관절에 대한 3D 위치 좌표를 나타내고,  $Y_{GT}$ 는 미리 획득된 3D 자세의 진리값(Ground Truth)을 나타내며,  $S$ 는 자세 추정부(31 ~ 3S)의 개수,  $T$ 는 입력 영상의 프레임 길이,  $\alpha$ 는  $S$ 개의 자세 추정부(31 ~ 3S) 각각에 대한 가중치,  $\lambda$ 는 정규화 파라미터를 나타낸다. 그리고  $w_k$ 는  $K$ 개의 전파 추정부(41 ~ 4K) 중  $k$ 번째 전파 추정부(4k)의 가중치를 나타내고,  $J$ 는 2D 특징맵(X)의 2D 특징 개수를 나타낸다.

[0059]

3D 자세 추정 장치는 학습시에 진리값( $Y_{GT}$ )이 미리 레이블된 2D 영상인 다수의 학습 데이터를 인가받아 3D 자세( $Y_{pred}$ )를 추정하고, 추정된 3D 자세( $Y_{pred}$ )와 진리값( $Y_{GT}$ ) 사이의 손실을 수학식 1에 따라 획득하여 역전파함으로써, 3D 자세 추정부(30)를 학습시킬 수 있다.

[0060]

도 7은 본 발명의 일실시예에 따른 3D 자세 추정 방법을 나타낸다.

[0061]

도 1 내지 도 6을 참조하여, 도 7의 3D 자세 추정 방법을 설명하면, 우선 3D 자세를 추정하기 위한 입력 영상을 획득한다(S11). 여기서 입력 영상은 2D 영상으로 특히 다수의 프레임을 갖는 2D 영상일 수 있다.

[0062]

입력 영상이 획득되면, 획득된 입력 영상의 각 프레임에서 미리 학습된 패턴 추정 방식에 따라 다수의 2D 특징을 추출하여 2D 특징맵(X)을 획득한다(S12). 여기서 2D 특징은 입력 영상에서 인간의 신체 중 기지정된 주요 관절의 위치를 나타내며, 일례로 각 프레임에서 인체의 머리, 목, 양어깨, 고관절, 팔꿈치, 손목, 무릎 및 발목의 14개의 주요 관절의 위치 2D 특징으로 추출할 수 있다. 따라서 다수의 2D 특징을 포함하여 구성되는 2D 특징맵(X)은 추정된 2D 자세를 나타낼 수 있다.

[0063]

한편 미리 학습된 패턴 추정 방식에 따라 2D 특징맵(X)의 다수의 2D 특징 중 기지정된 순서에 따라 적어도 하나의 2D 특징에 대응하는 적어도 하나의 3D 특징( $\hat{Y}^k$ )을 획득한다(S13). 일례로 인체의 몸통 부위에 위치하는 머리, 목, 양어깨, 고관절의 6개의 주요 관절에 대한 2D 특징에 대응하는 깊이 정보를 추정하여 반영된 6개의 3D 특징( $\hat{Y}^k$ )을 우선 획득할 수 있다. 그리고 추정된 적어도 하나의 3D 특징( $\hat{Y}^k$ )을 기지정된 방식으로 2D 특징맵(X)에 병합하여 자세 깊이 큐( $\hat{X}^k$ )를 획득한다(S14).

[0064]

이후 2D 특징맵(X)의 모든 특징에 대한 3D 특징을 획득하였는지 판별한다(S15). 만일 모든 특징에 대한 3D 특징이 획득되지 않은 것으로 판별되면, 이전 획득된 자세 깊이 큐( $\hat{X}^k$ )로부터 다시 기지정된 순서에 따라 적어도 하나의 2D 특징에 대응하는 적어도 하나의 3D 특징( $\hat{Y}^k$ )을 추정하여 획득한다(S13).

[0065]

그러나 모든 특징에 대한 3D 특징이 획득된 것으로 판별되면, 최종적으로 획득된 자세 깊이 큐( $\hat{X}^k$ )를 3D 특징맵( $Y^1 \sim Y^{S-1}$ )으로 획득한다(S16).

[0066]

그리고 획득된 3D 특징맵( $Y^1 \sim Y^{S-1}$ )에 대해 3D 특징을 재추정할지 여부를 판별한다(S17). 만일 3D 재추정하는 것으로 판별되면, 2D 특징맵(X)과 이전 획득된 3D 특징맵( $Y^1 \sim Y^{S-1}$ )으로부터 다수의 3D 특징에 대응하는 다수의 3D 특징( $\hat{Y}^k$ )을 기지정된 순서에 따라 다시 추정하여 3D 특징맵( $Y^2 \sim Y^S$ )을 획득한다(S13 ~ S16).

[0067]

즉 본 실시예에 따른 3D 자세 추정 방법은 3D 자세 추정 성능을 향상 시키기 위해, 2D 특징맵(X)에서 3D 특징을 추정하여 3D 특징맵( $Y^1$ )을 추정할 뿐만 아니라, 이미 획득된 3D 특징맵( $Y^1$ )에서 다시 3D 특징맵( $Y^2 \sim Y^S$ )을 추정하여 획득할 수 있다.

[0068] 한편, 3D 재추정하지 않는 것으로 판별되면, 최종적으로 획득된 3D 특징맵( $Y^S$ )을 3D 자세(Y)로 출력한다(S18).

[0069] 도시하지 않았으나, 본 실시예에 따른 3D 자세 추정 방법에서 2D 특징맵(X) 또는 3D 특징맵( $Y^1 \sim Y^{S-1}$ )로부터 다수의 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 추정하기 위해서는 패턴 추정 방식이 미리 학습되어야 한다. 이에 본 실시예에 따른 3D 자세 추정 방법은 3D 자세를 추정하기 이전에, 진리값( $Y_{GT}$ )이 미리 레이블된 2D 영상인 다수의 학습 데이터를 인가받아 3D 자세( $Y_{pred}$ )를 추정하고, 추정된 3D 자세( $Y_{pred}$ )와 진리값( $Y_{GT}$ ) 사이의 손실을 수식 1에 따라 획득하여 역전파하여 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 추정하기 위한 패턴 추정 방식이 사전에 학습될 수 있다.

[0070] 도 8은 본 실시예에 따른 3D 자세 추정 장치의 성능을 시뮬레이션한 결과를 나타낸다.

[0071] 도 8에서 기준 라인(baseline)은 2D 특징맵(X)에서 다수의 3D 특징을 동시에 추출하여 3D 특징맵을 생성한 결과를 나타내고, 3D 자세 추정부(30)의 제1 스테이지(Stage-1)는 제1 자세 추정부(31)에서 획득된 제1 3D 특징맵( $Y^1$ )을 나타내고, 제2 스테이지(Stage-2)는 제2 자세 추정부(32)에서 획득된 제2 3D 특징맵( $Y^2$ )을 나타내며, 제3 스테이지(Stage-3)는 제3 자세 추정부(33)에서 획득된 제3 3D 특징맵( $Y^3$ )을 나타낸다.

[0072] 도 8에 도시된 바와 같이, 본 실시예에 따른 3D 자세 추정 장치에서는 3D 자세 추정부(30)가 다단 구조의 다수의 자세 추정부(31 ~ 3S)를 포함하여 3D 자세를 추정함으로써 3D 자세를 매우 정확하게 추정할 수 있음을 알 수 있다.

[0073] 표 1은 깊이 병합부(63)가 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득하는 3가지 방법(D1, D2, D3)과 자세 추정부(31 ~ 3S)의 다수의 전과 추정부(41 ~ 4K)가 3D 특징을 추정하는 순서에 따라 추정된 3D 자세의 정확도를 시뮬레이션한 결과를 나타낸다.

표 1

Method (Protocol 1)	Head	Neck	R.shld	R.elbow	R.wrist	L.shld	L.elbow	L.wrist	R.hip	R.knee	R.ankle	L.hip	L.knee	L.ankle	Avg.
Single LSTM (baseline)	79.1	39.1	88.9	103.5	130.3	90.0	105.4	140.6	79.5	98.5	197.8	75.7	90.1	200.4	108.5
p-LSTMs ( $D=1, o$ )	65.1	33.4	83.5	100.4	110.1	87.6	97.4	108.1	74.5	80.0	117.1	74.3	80.5	116.2	87.6
p-LSTMs ( $D=2, o$ )	63.6	28.6	81.5	95.9	103.8	80.4	87.3	102.6	66.1	74.3	99.6	64.1	73.0	97.7	79.8
p-LSTMs ( $D=3, o$ )	<b>52.4</b>	<b>16.2</b>	<b>34.6</b>	<b>56.6</b>	<b>74.3</b>	<b>34.1</b>	<b>58.9</b>	<b>74.9</b>	<b>44.2</b>	<b>63.5</b>	<b>83.1</b>	<b>48.4</b>	<b>58.6</b>	<b>81.4</b>	<b>55.8</b>
p-LSTMs ( $D=3, i$ )	69.6	35.4	85.1	85.5	96.3	86.6	84.9	97.5	76.8	81.7	105.8	76.9	83.1	105.3	83.6

[0074]

[0075] 도 6에서는 깊이 병합부(63)가 자세 깊이 큐( $\hat{X}^1 \sim \hat{X}^K$ )를 획득하기 위해 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 병합하는 3가지 방법을 제시하였다. 표1 을 참조하면, 2D 특징( $x_1 \sim x_{14}$ )을 모두 제거하고, 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )을 추가하여, 이후로는 추가된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )으로부터 나머지 3D 특징을 추정하는 제1 방식(D1)과 다수의 2D 특징( $x_1 \sim x_{14}$ )을 그대로 유지한 채로 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )을 결합하는 제2 방식(D2)보다 새로이 추정된 3D 특징( $y_1, y_2, y_3, y_6, y_9, y_{12}$ )으로 이전 2D 특징( $x_1 \sim x_{14}$ )을 대체하는 제3 방식(D3)으로 3D 특징( $\hat{Y}^1 \sim \hat{Y}^K$ )을 병합하는 경우 가장 우수한 성능을 나타냄을 알 수 있다.

[0076] 또한 도 4에 도시된 바와 같이, 인체의 몸통으로부터 말단, 즉 외곽 방향으로 3D 특징을 추정하는 경우(표 1의 o), 인체의 말단에서 몸통 방향, 즉 내측 방향으로 3D 특징을 추정하는 경우(표 1의 i)에 비해 더 정확하게 3D 자세를 추정할 수 있음을 확인할 수 있다.

[0077] 도 9는 본 실시예에 따른 3D 자세 추정 장치를 실제 적용한 예를 나타낸다.

[0078] 도 9에 도시된 바와 같이, 본 실시예에 따른 3D 자세 추정 장치가 실제 사람에 대한 2D 영상을 입력받아 3D 자세를 추정한 결과는 만족할수 있는 수준의 성능을 나타냄을 확인할 수 있다.

[0079] 본 발명에 따른 방법은 컴퓨터에서 실행 시키기 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.

[0080] 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.

[0081] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

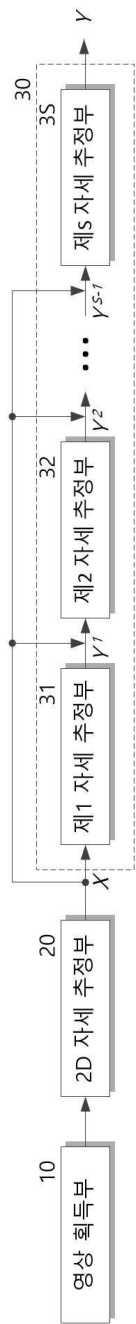
## 부호의 설명

[0082]

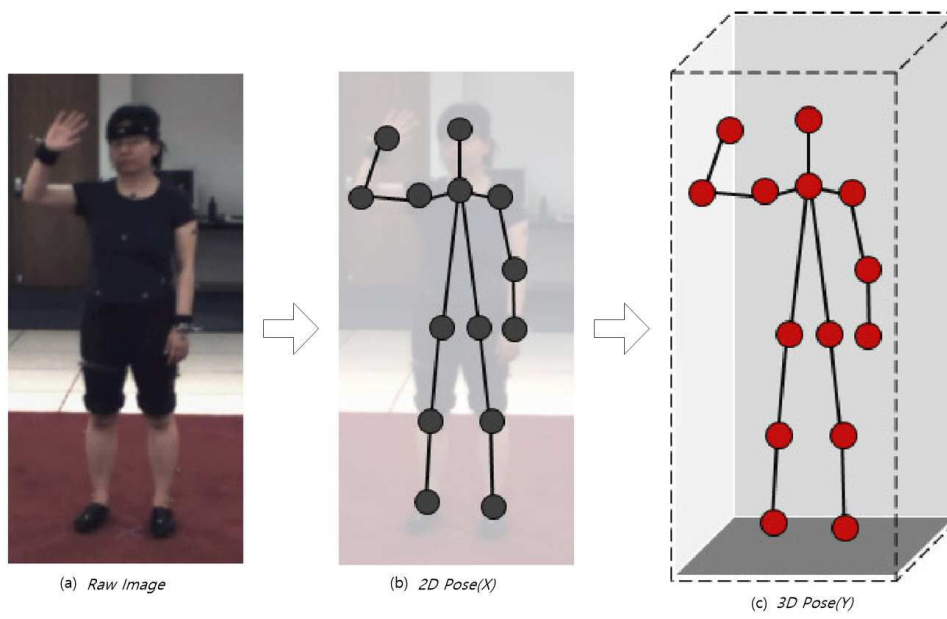
10: 영상 획득부	20: 2D 자세 추정부
30: 2D 자세 추정부	31 ~ 3S: 제1 ~ 제S 자세 추정부
41 ~ 4K: 제1 ~ 제K 전파 추정부	50: 3D 특징맵 출력부
61: 특징 입력부	62: 깊이 추정부
63: 깊이 병합부	

도면

도면1

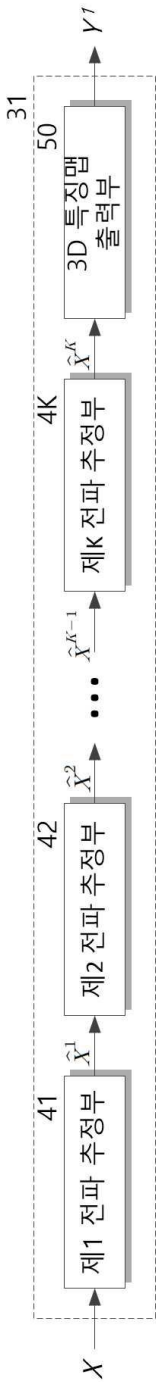


도면2

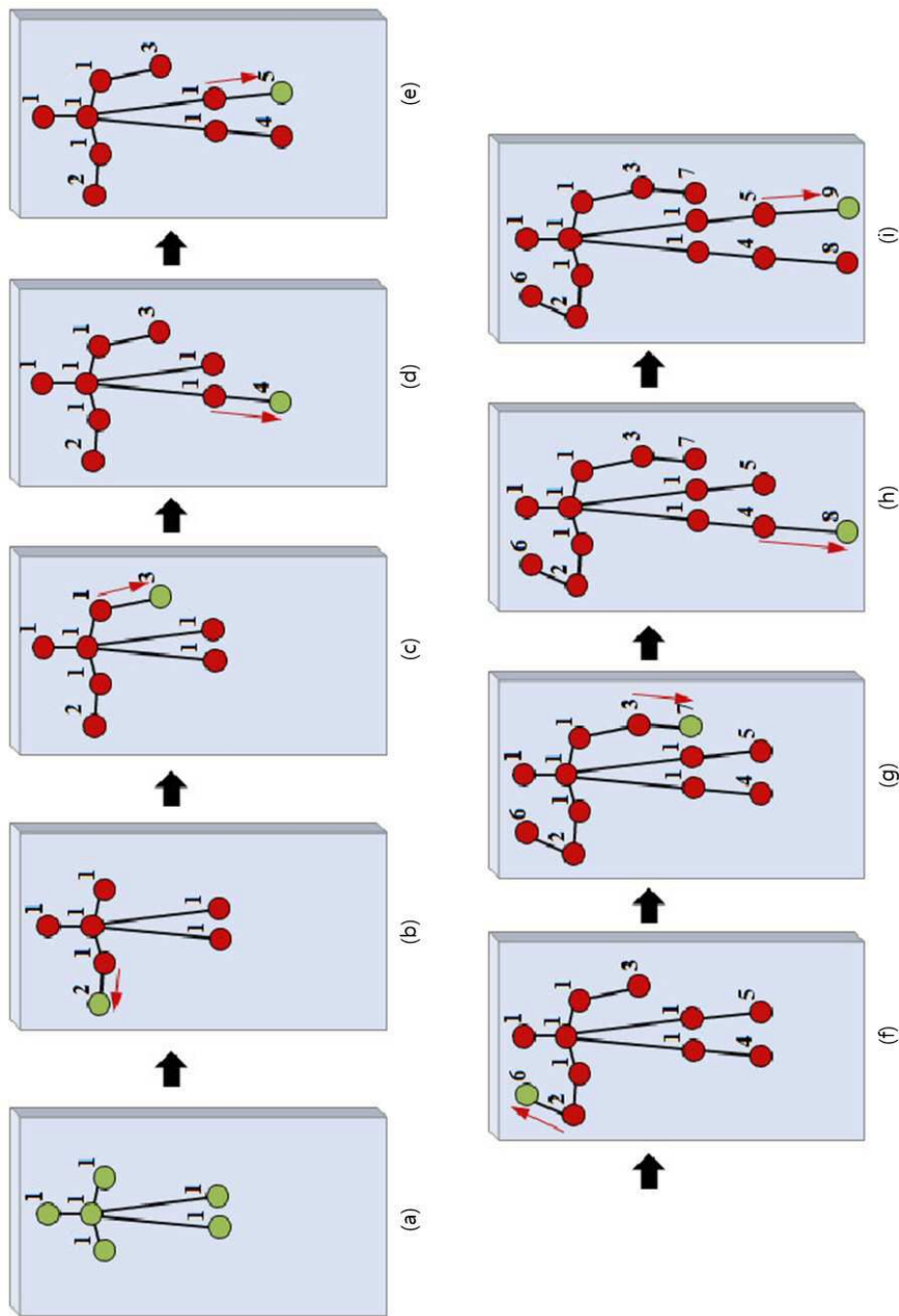




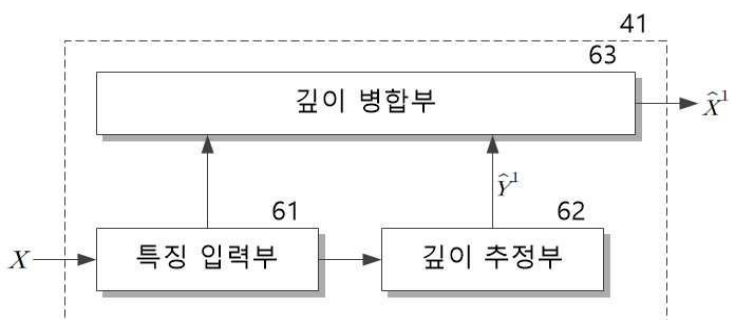
도면3



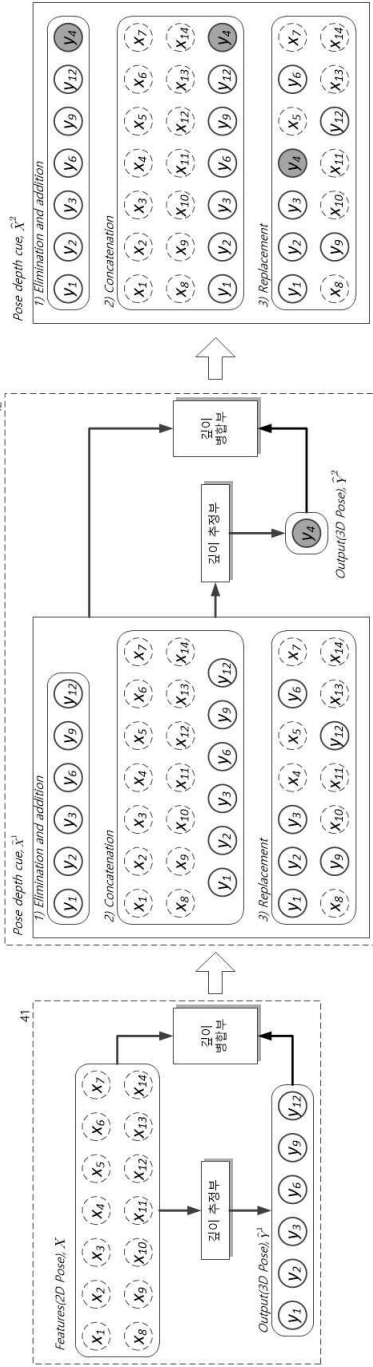
도면4



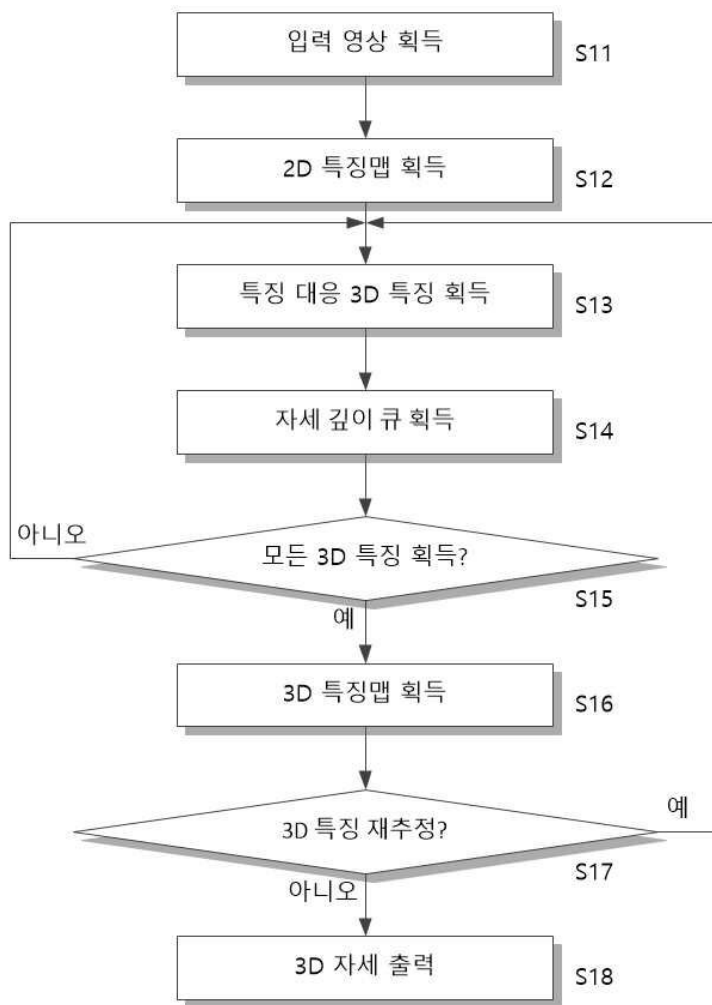
도면5



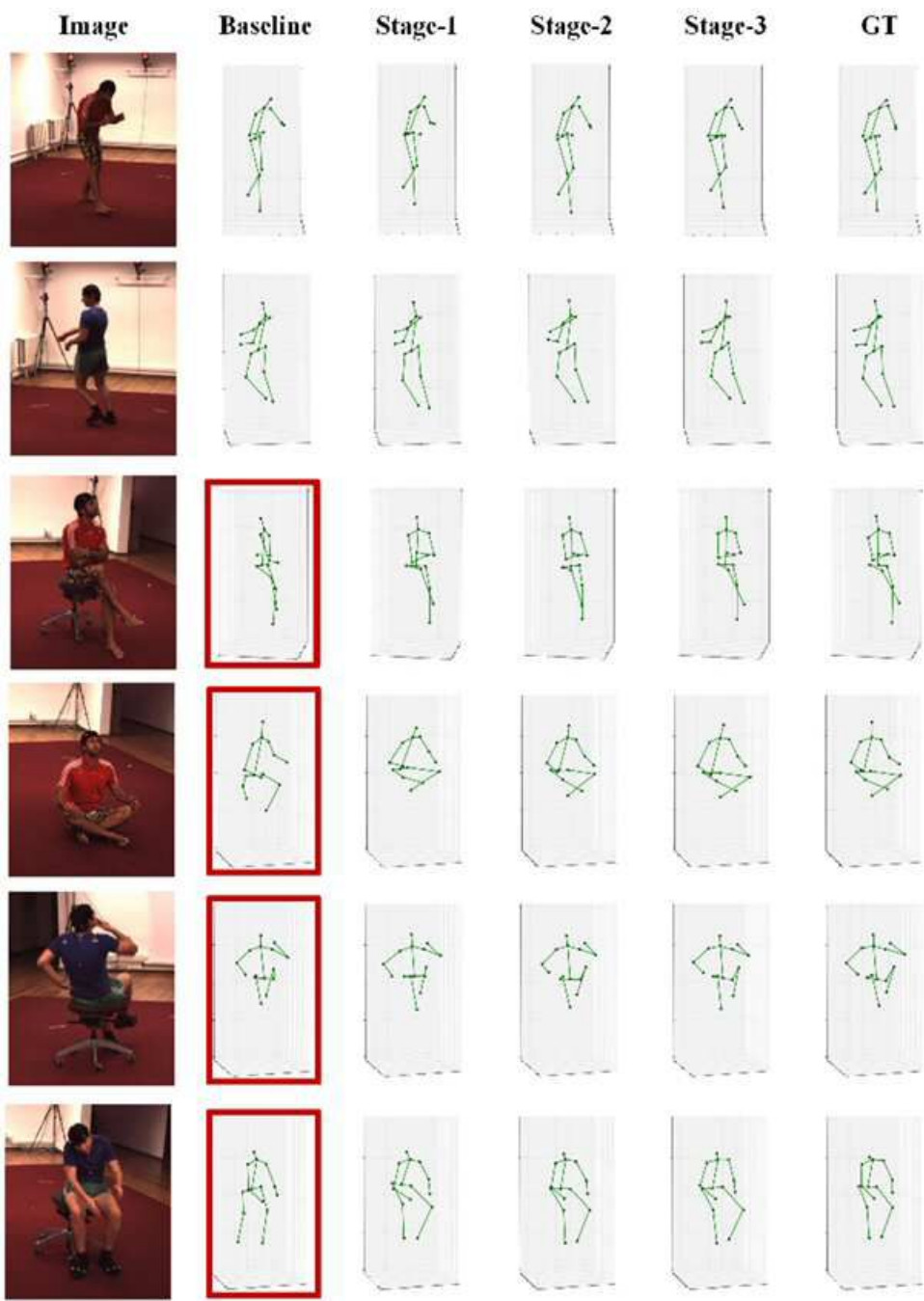
도면6



도면7



도면8



도면9

