



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0108409
(43) 공개일자 2023년07월18일

- | | |
|---|---|
| <p>(51) 국제특허분류(Int. Cl.)
 <i>G06T 7/73</i> (2017.01) <i>G06N 20/00</i> (2019.01)
 <i>G06N 3/04</i> (2023.01) <i>G06V 40/10</i> (2022.01)</p> <p>(52) CPC특허분류
 <i>G06T 7/73</i> (2017.01)
 <i>G06N 20/00</i> (2021.08)</p> <p>(21) 출원번호 10-2022-0003775
 (22) 출원일자 2022년01월11일
 심사청구일자 2022년01월11일</p> | <p>(71) 출원인
 울산과학기술원
 울산광역시 울주군 언양읍 유니스트길 50</p> <p>(72) 발명자
 임성훈
 울산광역시 울주군 언양읍 유니스트길 50
 최재경
 울산광역시 울주군 언양읍 유니스트길 50
 (뒷면에 계속)</p> <p>(74) 대리인
 이재록</p> |
|---|---|

전체 청구항 수 : 총 16 항

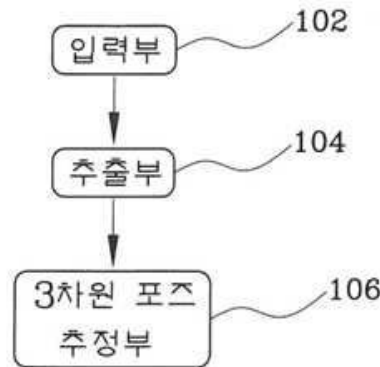
(54) 발명의 명칭 **가시성 지표를 활용한 3차원 포즈 추정 시스템 및 방법**

(57) 요약

가시성 지표를 활용한 3차원 포즈 추정 시스템 및 방법이 제공된다. 본 발명의 일 실시예에 따른 3차원 포즈 추정 시스템은, 사용자 이미지 또는 영상을 입력 받는 입력부; 입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절영역에 대한 가시성 지표(visibility score)를 추출하고, 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행한 후 상기 2차원 키 포인트 및 상기 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성하는 추출부; 및 상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출하는 3차원 포즈 추정부를 포함한다.

대표도 - 도1

100



(52) CPC특허분류

G06N 3/045 (2023.01)

G06V 40/10 (2022.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

(72) 발명자

김선준

울산광역시 울주군 언양읍 유니스트길 50

구민주

울산광역시 울주군 언양읍 유니스트길 50

이 발명을 지원한 국가연구개발사업

과제고유번호	1711134405
과제번호	2021-0-01139-001
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	ICT융합산업원천기술개발
연구과제명	3D 데이터와 AI의 기술융합을 기반한 3D Pose Estimation Motion Data 개발
기 여 율	1/1
과제수행기관명	울산과학기술원
연구기간	2021.04.01 ~ 2022.03.31

명세서

청구범위

청구항 1

사용자 이미지 또는 영상을 입력 받는 입력부;

입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절영역에 대한 가시성 지표(visibility score)를 추출하고, 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행한 후 상기 2차원 키 포인트 및 상기 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성하는 추출부; 및

상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출하는 3차원 포즈 추정부를 포함하는, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 2

청구항 1에 있어서,

상기 2차원 키 포인트는, 상기 관절영역의 x 좌표 및 y 좌표를 포함하며,

상기 가시성 지표는, 상기 관절영역의 가시성 정도에 따라 0 ~ 1 사이의 값으로 표현되는 값인, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 3

청구항 1에 있어서,

상기 추출부는, 상기 2차원 키 포인트 및 상기 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 ReLU(Rectified Linear Unit) 활성 함수를 순차적으로 적용함으로써 상기 전처리 작업을 수행하는, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 4

청구항 1에 있어서,

상기 학습 모델은, 확장된 합성곱 신경망(Dilated Convolutional Neural Network)인, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 5

청구항 4에 있어서,

상기 3차원 포즈 추정부는, 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신하는, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 6

청구항 5에 있어서,

상기 확장된 합성곱 신경망의 목적 함수(objective function)는, 아래의 수학적 식 1과 같이 표현되는, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

[수학적 식 1]

$$L(\theta) = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \|f(x_i, \theta)^{(t)} - \hat{y}_i^{(t)}\|_2$$

(여기서, L은 상기 확장된 합성곱 신경망의 목적 함수, T는 전체 데이터의 개수, N은 하나의 객체에 대한 관절 좌표의 개수, θ 는 상기 확장된 합성곱 신경망의 파라미터, x_i 는 상기 입력 데이터, $f(x_i, \theta)$ 는 상기 파라미터와 상기 입력 데이터를 기반으로 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트, $\hat{y}_i$ 는 실제 3차원 관절 좌표 값을 각각 나타냄)

청구항 7

청구항 6에 있어서,

상기 3차원 포즈 추정부는, 아래의 수학적 식 2와 같이 상기 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시키는, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

[수학적 식 2]

$$\operatorname{argmin}_{\theta} L(\theta)$$

청구항 8

청구항 7에 있어서,

상기 파라미터의 최적화 방식은, Adam(Adaptive moment estimation) 기법인, 가시성 지표를 활용한 3차원 포즈 추정 시스템.

청구항 9

입력부에서, 사용자 이미지 또는 영상을 입력 받는 단계;

추출부에서, 입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절 영역에 대한 가시성 지표(visibility score)를 추출하는 단계;

상기 추출부에서, 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행하는 단계;

상기 추출부에서, 상기 2차원 키 포인트 및 상기 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성하는 단계; 및

상기 3차원 포즈 추정부에서, 상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출하는 단계를 포함하는, 가시성 지표를 활용한 3차원 포즈 추정 방법.

청구항 10

청구항 9에 있어서,

상기 2차원 키 포인트는, 상기 관절영역의 x 좌표 및 y 좌표를 포함하며,

상기 가시성 지표는, 상기 관절영역의 가시성 정도에 따라 0 ~ 1 사이의 값으로 표현되는 값인, 가시성 지표를 활용한 3차원 포즈 추정 방법.

청구항 11

청구항 9에 있어서,

상기 전처리 작업을 수행하는 단계는, 상기 2차원 키 포인트 및 상기 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 ReLU(Rectified Linear Unit) 활성화 함수를 순차적으로 적용함으로써 상기 전처리 작업을 수행하는, 가시성 지표를 활용한 3차원 포즈 추정 방법.

청구항 12

청구항 9에 있어서,

상기 학습 모델은, 확장된 합성곱 신경망(Dilated Convolutional Neural Network)인, 가시성 지표를 활용한 3차원 포즈 추정 방법.

청구항 13

청구항 12에 있어서,

상기 3차원 키 포인트를 도출하는 단계 이후,

상기 3차원 포즈 추정부에서, 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신하는 단계를 더 포함하는, 가시성 지표를 활용한 3차원 포즈 추정 방법.

청구항 14

청구항 13에 있어서,

상기 확장된 합성곱 신경망의 목적 함수(objective function)는, 아래의 수학식 1과 같이 표현되는, 가시성 지표를 활용한 3차원 포즈 추정 방법.

[수학식 1]

$$L(\theta) = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \|f(x_i, \theta)^{(t)} - \hat{y}_i^{(t)}\|_2$$

(여기서, L은 상기 확장된 합성곱 신경망의 목적 함수, T는 전체 데이터의 개수, N은 하나의 객체에 대한 관절 좌표의 개수, θ 는 상기 확장된 합성곱 신경망의 파라미터, x_i 는 상기 입력 데이터, $f(x_i, \theta)$ 는 상기 파라미터와 상기 입력 데이터를 기반으로 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트, $\hat{y}_i$ 는 실제 3차원 관절 좌표 값을 각각 나타냄)

청구항 15

청구항 14에 있어서,

상기 확장된 합성곱 신경망의 파라미터를 갱신하는 단계는, 아래의 수학식 2와 같이 상기 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시키는, 가시성 지표를 활용한 3차원 포즈 추정 방법.

[수학식 2]

$$\operatorname{argmin}_{\theta} L(\theta)$$

청구항 16

청구항 15에 있어서

상기 파라미터의 최적화 방식은, Adam(Adaptive moment estimation) 기법인, 가시성 지표를 활용한 3차원 포즈 추정 방법.

발명의 설명

기술 분야

[0001] 본 발명은 가시성 지표를 이용하여 3차원 포즈를 추정하는 기술과 관련된다.

배경 기술

[0003] 포즈 추정 기술은 이미지, 영상 등을 학습 모델에 입력하여 사람의 관절에 대한 좌표를 추정하는 기술로서, 추정하고자 하는 관절의 범위, 관절의 개수, 추정하고자 하는 사람의 인원 수에 따라 상기 학습 모델의 구성이 달라질 수 있다. 이러한 포즈 추정 기술은 추정 방법과 추정 인원 수에 따라 Top-Down 방식, Bottom-Up 방식, Single Person 방식, Multi Person 방식 등으로 나뉘어질 수 있다. 구체적으로, Top-Down 방식은 사람을 먼저 감지한 후 경계 박스 내부에서 사람의 관절을 추정하는 방식이며, Bottom-Up 방식은 관절 좌표를 먼저 감지하여 이들을 상호 연결하여 모든 객체에 대한 자세를 추정하는 방식이다. 또한, Single Person 방식은 추정하고자 하는 사람의 인원 수가 한 명인 경우 적용될 수 있는 방식이며, Multi Person 방식은 추정하고자 하는 사람의 인원 수가 여러 명인 경우 적용될 수 있는 방식이다.

[0004] 그러나, 종래의 포즈 추정 기술은 추정하고자 하는 객체가 다른 객체에 의해 가려졌을 때 좌표 추정의 정확도가 매우 낮아질 수 있다는 문제점(Occlusion Problem)이 있다. 이러한 문제점 해결을 위해 시간적 정보를 활용하는 LSTM(Long Short Term Memory) 모델에 L2 규제를 더하여 가려짐 문제를 완화하려는 시도가 있었으나, 이 경우 상기 모델의 파라미터가 지나치게 많아져 모델이 무거워지는 단점이 있다. 상기 L2 규제는 딥러닝 모델의 문제점 중 하나인 과적합을 해결하기 위한 방법으로서, 목적 함수에 정규항을 추가하여 학습에 규제를 가하는 방식이다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 한국등록특허공보 제10-2194282호(2020.12.16)

발명의 내용

해결하려는 과제

[0007] 본 발명의 실시예들은 포즈 추정 과정에서 다른 객체에 의한 가려짐으로 인한 성능 저하 문제 및 학습 모델의 과도한 계산으로 인한 비용 문제를 해결하기 위한 것이다.

과제의 해결 수단

[0009] 본 발명의 예시적인 실시예에 따르면, 사용자 이미지 또는 영상을 입력 받는 입력부; 입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절영역에 대한 가시성 지표(visibility score)를 추출하고, 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행한 후 상기 2차원 키 포인트 및 상기 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성하는 추출부; 및 상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출하는 3차원 포즈 추정부를 포함하는, 가시성 지표를 활용한 3차원 포즈 추정 시스템이 제공된다.

[0010] 상기 2차원 키 포인트는, 상기 관절영역의 x 좌표 및 y 좌표를 포함하며, 상기 가시성 지표는, 상기 관절영역의 가시성 정도에 따라 0 ~ 1 사이의 값으로 표현되는 값일 수 있다.

[0011] 상기 추출부는, 상기 2차원 키 포인트 및 상기 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 ReLU(Rectified Linear Unit) 활성화 함수를 순차적으로 적용함으로써 상기 전처리 작업을 수행할 수 있다.

[0012] 상기 학습 모델은, 확장된 합성곱 신경망(Dilated Convolutional Neural Network)일 수 있다.

[0013] 상기 3차원 포즈 추정부는, 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신할 수 있다.

[0014] 상기 확장된 합성곱 신경망의 목적 함수(objective function)는, 아래의 수학식 1과 같이 표현될 수 있다.

[0015] [수학식 1]

$$L(\theta) = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \|f(x_i, \theta)^{(t)} - \hat{y}_i^{(t)}\|_2$$

[0016]

[0017] (여기서, L은 상기 확장된 합성곱 신경망의 목적 함수, T는 전체 데이터의 개수, N은 하나의 객체에 대한 관절 좌표의 개수, θ 는 상기 확장된 합성곱 신경망의 파라미터, x_i 는 상기 입력 데이터, $f(x_i, \theta)$ 는 상기 파라미터와 상기 입력 데이터를 기반으로 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트, $\hat{y}_i$ 는 실제 3차원 관절 좌표 값을 각각 나타냄)

[0018] 상기 3차원 포즈 추정부는, 아래의 수학식 2와 같이 상기 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시킬 수 있다.

[0019] [수학식 2]

$$\operatorname{argmin}_{\theta} L(\theta)$$

[0020]

[0021] 상기 파라미터의 최적화 방식은, Adam(Adaptive moment estimation) 기법일 수 있다.

[0022] 본 발명의 다른 예시적인 실시예에 따르면, 입력부에서, 사용자 이미지 또는 영상을 입력 받는 단계; 추출부에서, 입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절영역에 대한 가시성 지표(visibility score)를 추출하는 단계; 상기 추출부에서, 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행하는 단계; 상기 추출부에서, 상기 2차원 키 포인트 및 상기 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성하는 단계; 및 상기 3차원 포즈 추정부에서, 상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출하는 단계를 포함하는, 가시성 지표를 활용한 3차원 포즈 추정 방법이 제공된다.

[0023] 상기 2차원 키 포인트는, 상기 관절영역의 x 좌표 및 y 좌표를 포함하며, 상기 가시성 지표는, 상기 관절영역의 가시성 정도에 따라 0 ~ 1 사이의 값으로 표현되는 값일 수 있다.

[0024] 상기 전처리 작업을 수행하는 단계는, 상기 2차원 키 포인트 및 상기 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 ReLU(Rectified Linear Unit) 활성화 함수를 순차적으로 적용함으로써 상기 전처리 작업을 수행할 수 있다.

[0025] 상기 학습 모델은, 확장된 합성곱 신경망(Dilated Convolutional Neural Network)일 수 있다.

[0026] 상기 가시성 지표를 활용한 3차원 포즈 추정 방법은, 상기 3차원 키 포인트를 도출하는 단계 이후, 상기 3차원 포즈 추정부에서, 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신하는 단계를 더 포함할 수 있다.

[0027] 상기 확장된 합성곱 신경망의 목적 함수(objective function)는, 아래의 수학식 1과 같이 표현될 수 있다.

[0028] [수학식 1]

$$L(\theta) = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \|f(x_i, \theta)^{(t)} - \hat{y}_i^{(t)}\|_2$$

[0029]

[0030] (여기서, L은 상기 확장된 합성곱 신경망의 목적 함수, T는 전체 데이터의 개수, N은 하나의 객체에 대한 관절 좌표의 개수, θ 는 상기 확장된 합성곱 신경망의 파라미터, x_i 는 상기 입력 데이터, $f(x_i, \theta)$ 는 상기 파라미터와 상기 입력 데이터를 기반으로 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트, $\hat{y}_i$ 는 실제 3차원 관절 좌표 값을 각각 나타냄)

[0031] 상기 확장된 합성곱 신경망의 파라미터를 갱신하는 단계는, 아래의 수학식 2와 같이 상기 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시킬 수 있다.

[0032] [수학식 2]

$$\operatorname{argmin}_{\theta} L(\theta)$$

[0033]

[0034] 상기 파라미터의 최적화 방식은, Adam(Adaptive moment estimation) 기법일 수 있다.

발명의 효과

[0036] 본 발명의 일 실시예에 따르면, 2차원 키 포인트뿐 아니라 가시성 지표를 학습 모델에 추가적으로 입력하여 3차원 키 포인트를 도출하도록 함으로써 포즈 추정의 정확도를 향상시키고 다른 객체에 의한 가려짐 문제를 최소화시킬 수 있다.

[0037] 또한, 본 발명의 일 실시예에 따르면, 학습 모델로서 확장된 합성곱 신경망을 이용하여 시간적으로 멀리 떨어져 있는 프레임의 키 포인트에 대한 정보도 활용 가능하게 되며, 이에 따라 종래 시간적 정보를 활용하던 LSTM 모델 또는 일반적인 합성곱 신경망과 비교하여 학습 모델의 경량화 및 성능 개선의 효과를 달성할 수 있다.

[0038] 또한, 본 발명의 일 실시예에 따르면, 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신함으로써, 확장된 합성곱 신경망의 성능을 개선하고 최적화시킬 수 있다.

도면의 간단한 설명

[0040] 도 1은 본 발명의 일 실시예에 따른 3차원 포즈 추정 시스템의 상세 구성을 나타낸 블록도

도 2는 본 발명의 일 실시예에 따른 입력 데이터의 생성 과정을 나타낸 블록도

도 3은 본 발명의 일 실시예에 따른 사용자 이미지 또는 영상의 예시

도 4는 일반적인 합성곱 신경망과 본 발명의 일 실시예에 따른 확장된 합성곱 신경망을 나타낸 예시

도 5는 본 발명의 일 실시예에 따른 확장된 합성곱 신경망의 각 레이어에서 처리되는 데이터의 흐름을 나타낸 예시

도 6은 본 발명의 일 실시예에 따른 3차원 포즈 추정부에서 출력되는 사용자 이미지 또는 영상에 대한 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 나타낸 예시

도 7은 본 발명의 일 실시예에 따른 3차원 포즈 추정 방법을 설명하기 위한 흐름도

도 8은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도

발명을 실시하기 위한 구체적인 내용

[0041] 이하, 도면을 참조하여 본 발명의 구체적인 실시형태를 설명하기로 한다. 이하의 상세한 설명은 본 명세서에서 기술된 방법, 장치 및/또는 시스템에 대한 포괄적인 이해를 돕기 위해 제공된다. 그러나 이는 예시에 불과하며 본 발명은 이에 제한되지 않는다.

[0042] 본 발명의 실시예들을 설명함에 있어서, 본 발명과 관련된 공지기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다. 그리고, 후술되는 용어들은 본 발명에서의 기능을 고려하여 정의된 용어들로서 이는 사용자, 운용자의 의도 또는 관례 등에 따라 달라질 수 있다. 그러므로 그 정의는 본 명세서 전반에 걸친 내용을 토대로 내려져야 할 것이다. 상세한 설명에서 사용되는 용어는 단지 본 발명의 실시예들을 기술하기 위한 것이며, 결코 제한적이어서는 안 된다. 명확하게 달리 사용되지 않는 한, 단수 형태의 표현은 복수 형태의 의미를 포함한다. 본 설명에서, "포함" 또는 "구비"와 같은 표현은 어떤 특성들, 숫자들, 단계들, 동작들, 요소들, 이들의 일부 또는 조합을 가리키기 위한 것이며, 기술된 것 이외에 하나 또는 그 이상의 다른 특성, 숫자, 단계, 동작, 요소, 이들의 일부 또는 조합의 존재 또는 가능성을 배제하도록 해석되어서는 안 된다.

[0044] 도 1은 본 발명의 일 실시예에 따른 3차원 포즈 추정 시스템(100)의 상세 구성을 나타낸 블록도이다. 도 1에 도시된 바와 같이, 본 발명의 일 실시예에 따른 3차원 포즈 추정 시스템(100)은 입력부(102), 추출부(104) 및 3차원 포즈 추정부(106)를 포함한다.

[0045] 입력부(102)는 사용자 이미지 또는 영상을 입력 받는다. 여기서, 사용자 이미지 또는 영상은 사람에 대한 RGB 이미지, 영상 등이 될 수 있다. 입력부(102)는 예를 들어, 오픈 데이터 셋인 Human 3.6M Dataset 으로부터 상기 사용자 이미지 또는 영상을 입력 받을 수 있다. 이러한 사용자 이미지 또는 영상은 예를 들어, 11명의 배우가 취하는 17가지 동작에 대한 영상으로부터 추출되는 데이터일 수 있다. 상기 영상은 예를 들어, 50Hz의 고해상도 mp4 영상일 수 있다.

[0046] 또한, 입력부(102)는 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 입력 받을 수 있다. 후술할 바와 같이, 상기 실제 3차원 관절 좌표 값은 포즈 추정을 위한 학습 모델을 평가하고 상기 학습 모델의 파라미터를 갱신하는 데 활용될 수 있다.

[0047] 추출부(104)는 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트(key point) 및 상기 관절영역에 대한 가시성 지표(visibility score)를 추출한다. 여기서, 상기 2차원 키 포인트는 예를 들어, 상기 관절영역의 x 좌표 및 y 좌표를 포함할 수 있다. 추출부(104)는 예를 들어, Mask R-CNN 모델을 이용하여 상기 관절영역의 2차원 키 포인트, 즉 2차원 좌표를 추출할 수 있다. 또한, 상기 가시성 지표는 상기 관절영역의 가시성 정도에 따라 0 ~ 1 사이의 값으로 표현되는 값일 수 있다. 추출부(104)는 예를 들어, Mask R-CNN, Alphapose 등과 같은 2D 관절좌표 탐지 모델을 활용하여 상기 관절영역의 가시성 지표를 추출할 수 있다.

[0048] 추출부(104)는 상기 2차원 키 포인트와 상기 가시성 지표를 결합하여 입력 데이터를 생성한다.

[0049] 이를 위해, 추출부(104)는 먼저 상기 2차원 키 포인트 및 상기 가시성 지표에 대한 전처리 작업을 수행할 수 있다. 구체적으로, 추출부(104)는 상기 2차원 키 포인트(즉, 2차원 관절 좌표) 및 상기 가시성 지표에 각각 합성

곱 신경망, 정규화 함수 및 ReLU(Rectified Linear Unit) 활성화 함수를 순차적으로 적용함으로써 상기 전처리 작업을 수행할 수 있다.

[0050] 다음으로, 추출부(104)는 상기 전처리 과정을 거친 2차원 키 포인트 및 가시성 지표를 연결(concatenation)하여 입력 데이터(X)를 생성할 수 있다. 상기 2차원 키 포인트(2D) 및 상기 가시성 지표(V)는 행렬로 표현될 수 있으며, 아래와 같은 연산을 통해 입력 데이터(X)가 생성될 수 있다.

[0052] $X = [C(2D); C(V)]$

[0053] (여기서, X는 입력 데이터, C는 상기 합성곱 신경망, 정규화 함수 및 ReLU 활성화 함수를 포함하는 전처리 수행함수, 2D는 상기 2차원 키 포인트, V는 상기 가시성 지표, [;]는 연결(concatenation) 연산자를 각각 의미함)

[0055] 이러한 2차원 키 포인트(2D), 가시성 지표(V) 및 입력 데이터(X)를 구체적인 예시를 들어 설명하면 아래와 같다.

[0057] [전처리 수행함수 C를 통과한 17개의 2차원 키 포인트(C(2D))]

[0058] = [[0.34, 0.50],

[0059] [0.44, 0.75]

[0060] ...,

[0061] [0.56, 0.12]] // (17×2 행렬)

[0063] [전처리 수행함수 C를 통과한 17개의 가시성 지표(C(V))]

[0064] = [[0.1],

[0065] [0.2]

[0066] [0.3]

[0067] ...

[0068] [0.7]] // (17×1 행렬)

[0070] [입력 데이터(X)]

[0071] = [[0.34, 0.50, 0.1],

[0072] [0.44, 0.75, 0.2],

[0073] ...

[0074] [0.56, 0.12, 0.7]] // (17×3 행렬)

[0076] 3차원 포즈 추정부(106)는 상기 입력 데이터를 학습 모델에 입력하여 상기 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 도출한다. 상술한 바와 같이, 관절영역에 대한 가시성 지표가 반영된 입력 데이터가 학습 모델에 입력되며, 상기 가시성 지표를 통해 관절영역의 깊이 정보(즉, z 좌표)에 대한 추정이 가능해진다. 예를 들어, 여러 프레임에 대한 2차원 키 포인트가 입력 데이터에 포함되어 학습 모델에 입력되는데, 가시성 지표가 상대적으로 높은 프레임의 2차원 키 포인트가 포즈 추정에 더 많은 영향을 주게 된다.

[0077] 일 예시로서, 관절영역의 2차원 키 포인트(2D) 및 가시성 지표(V)가 아래와 같다고 가정하는 경우 입력 데이터(X)는 아래와 같이 표현될 수 있다.

- [0079] [전처리 수행함수 C를 통과한 2차원 키 포인트($C(2D)$)]
- [0080] = $[[a, b]$
- [0081] $[c, d]]$
- [0083] [전처리 수행함수 C를 통과한 가시성 지표($C(V)$)]
- [0084] = $[[0.1],$
- [0085] $[0.9]]$
- [0087] [입력 데이터(X)]
- [0088] = $[[a, b, 0.1],$
- [0089] $[c, d, 0.9]]$
- [0091] 이때, $[a, b]$ 키 포인트보다는 $[c, d]$ 키 포인트에 대한 가시성 지표가 상대적으로 더 크므로, $[c, d]$ 키 포인트가 $[a, b]$ 키 포인트에 비해 3차원 포즈 추정에 상대적으로 더 큰 영향을 줄 수 있다.
- [0092] 이와 같이, 학습 모델에 입력되는 입력 데이터(X)는 합성곱 신경망, 정규화 함수 및 ReLU 활성화 함수를 포함하는 전처리 수행함수(C)를 거친 2차원 키 포인트(2D) 및 가시성 지표(V)가 연결(concatenation)된 형태의 행렬로 이루어지며, 이렇게 생성된 입력 데이터(X)는 하나의 함축적 가중치를 갖는 데이터가 된다. 본 발명의 일 실시예에 따르면, 2차원 키 포인트뿐 아니라 가시성 지표를 학습 모델에 추가적으로 입력하여 3차원 키 포인트를 도출하도록 함으로써 포즈 추정의 정확도를 향상시키고 다른 객체에 의한 가려짐 문제를 최소화시킬 수 있다.
- [0093] 한편, 본 실시예들에 있어서, 상기 학습 모델은 예를 들어, 확장된 합성곱 신경망(Dilated Convolutional Neural Network)일 수 있다. 3차원 포즈 추정부(106)는 상기 입력 데이터를 상기 확장된 합성곱 신경망에 입력할 수 있다.
- [0094] 상기 확장된 합성곱 신경망은 특정 형태의 필터를 가지고 있는 합성곱 신경망으로서, 필터가 가지고 있는 파라미터 사이에 0을 삽입하여 수용 영역(reception field)을 넓히면서 기존 계산 비용(model computational cost)은 그대로 유지하는 형태의 합성곱 신경망이다. 본 발명의 실시예들에 따르면, 확장된 합성곱 신경망을 활용함으로써 학습 모델의 경량화 및 성능 개선의 효과를 달성할 수 있다.
- [0095] 또한, 3차원 포즈 추정부(106)는 상술한 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)를 이용하여 상기 학습 모델을 평가할 수 있다. 3차원 포즈 추정부(106)는 상기 학습 모델, 즉 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값을 비교하여 상기 학습 모델에서 예측된 3차원 키 포인트를 평가하고, 상기 학습 모델의 파라미터를 갱신하여 최적화시킬 수 있다. 예를 들어, 3차원 포즈 추정부(106)는 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값 간의 평균 좌표 오차(MPJPE : Mean per joint position error)로부터 상기 학습 모델 및 상기 학습 모델에서 예측된 3차원 키 포인트를 평가할 수 있다.
- [0096] 이때, 상기 확장된 합성곱 신경망의 목적 함수(objective function)(또는 손실 함수)는 아래의 수학식 1과 같이 표현될 수 있다. 상기 목적 함수는 상술한 평균 좌표 오차를 사용하였으며, 상기 평균 좌표 오차는 각 관절 좌표 간의 유클리드 거리(Euclidean distance)를 기반으로 계산될 수 있다.

[0098] [수학식 1]

$$L(\theta) = \frac{1}{T} \frac{1}{N} \sum_{t=1}^T \sum_{i=1}^N \|f(x_i, \theta)^{(t)} - \hat{y}_i^{(t)}\|_2$$

[0099]

[0100] (여기서, L은 상기 확장된 합성곱 신경망의 목적 함수, T는 전체 데이터의 개수, N은 하나의 객체에 대한 관절 좌표의 개수, θ 는 상기 확장된 합성곱 신경망의 파라미터, x_i 는 상기 입력 데이터, $f(x_i, \theta)$ 는 상기 파라미터와 상기 입력 데이터를 기반으로 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트, $\hat{y}_i$ 는 실제 3차원 관절 좌표 값을 각각 나타냄)

[0102] 3차원 포즈 추정부(106)는 아래의 수학식 2와 같이 상기 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시킬 수 있다.

[0104] [수학식 2]

$$\operatorname{argmin}_{\theta} L(\theta)$$

[0105]

[0107] 이때, 상기 파라미터의 최적화 방식은 예를 들어, 경사하강법(gradient descent)일 수 있다. 구체적으로, 상기 경사하강법은 예를 들어, Adam(Adaptive moment estimation) 기법일 수 있다. 상기 Adam 기법은 간단한 구현과 효율적인 연산을 장점으로 하고 있다. 즉, 3차원 포즈 추정부(106)는 상기 목적 함수(또는 손실 함수)를 최소화시키는 방향으로 최적화된 파라미터 θ 를 탐색하게 된다. 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트인 $f(x, \theta)$ 와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값인 $\hat{y}_i$ 간의 차이에 따라 파라미터 θ 의 갱신 방식이 달라질 수 있다.

[0108] 예를 들어, 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트인 $f(x, \theta)$ 와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값인 $\hat{y}_i$ 간의 차이가 클수록 파라미터 θ 의 갱신이 목표값을 향해 크게 일어나게 되며, 상기 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트인 $f(x, \theta)$ 와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값인 $\hat{y}_i$ 간의 차이가 작은 경우 파라미터 θ 의 갱신이 목표값을 향해 일어나지만 앞선 경우에 비해 상대적으로 작게 일어날 수 있다.

[0109] 이와 같이, 3차원 포즈 추정부(106)는 상술한 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)를 이용하여 상기 학습 모델을 검증하고, 상기 학습 모델의 파라미터를 최적화시킴으로써 상기 학습 모델의 성능을 개선할 수 있다.

[0111] 도 2는 본 발명의 일 실시예에 따른 입력 데이터(X)의 생성 과정을 나타낸 블록도이다.

[0112] 상술한 바와 같이, 추출부(104)는 2차원 키 포인트 및 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 활성화 함수인 ReLU(Rectified Linear Unit) 활성화 함수를 순차적으로 적용함으로써 전처리 작업을 수행할 수 있다.

[0113] 다음으로, 추출부(104)는 상기 전처리 과정을 거친 2차원 키 포인트 및 가시성 지표를 연결(concatenation)하여 입력 데이터(X)를 생성할 수 있다. 상기 2차원 키 포인트(2D) 및 상기 가시성 지표(V)는 행렬로 표현될 수 있으며, 아래와 같은 연산을 통해 입력 데이터(X)가 생성될 수 있다.

[0115] 도 3은 본 발명의 일 실시예에 따른 사용자 이미지 또는 영상의 예시이다.

- [0116] 상술한 바와 같이, 본 발명의 일 실시예에 따른 입력부(102)는 예를 들어, 오픈 데이터 셋인 Human 3.6M Dataset 으로부터 상기 사용자 이미지 또는 영상을 입력 받을 수 있다. 상기 사용자 이미지 또는 영상은 예를 들어, 11명의 배우가 취하는 17가지 동작에 대한 영상으로부터 추출되는 데이터일 수 있다.
- [0118] 도 4는 일반적인 합성곱 신경망과 본 발명의 일 실시예에 따른 확장된 합성곱 신경망을 나타낸 예시이며, 도 4는 본 발명의 일 실시예에 따른 확장된 합성곱 신경망의 각 레이어에서 처리되는 데이터의 흐름을 나타낸 예시이다.
- [0119] 먼저, 도 4의 (a)는 일반적인 합성곱 신경망을 나타내며, 도 4의 (b)는 본 발명의 일 실시예에 따른 확장된 합성곱 신경망을 나타낸다. 상술한 바와 같이, 상기 확장된 합성곱 신경망은 특정 형태의 필터를 가지고 있는 합성곱 신경망으로서, 필터가 가지고 있는 파라미터 사이에 0을 삽입하여 수용 영역(reception field)을 넓혀 시간적으로 떨어져 있는 프레임의 정보를 활용 가능하게 하며, 기존 계산 비용은 그대로 유지하는 형태의 합성곱 신경망이다.
- [0120] 다음으로, 도 5에 도시된 바와 같이 확장된 합성곱 신경망은 예를 들어, 제1 레이어(first layer), 제2 레이어(second layer) 및 제3 레이어(third layer)로 이루어질 수 있다. 이때, 확장된 합성곱 신경망의 레이어가 많아질수록 더 많은 프레임을 입력으로 받을 수 있다. 예를 들어, 3개의 레이어를 쌓은 확장된 합성곱 신경망은 27개의 프레임, 4개의 레이어를 쌓은 확장된 합성곱 신경망은 81개의 프레임, 5개의 레이어를 쌓은 확장된 합성곱 신경망은 243개의 프레임을 각각 입력으로 받을 수 있다.
- [0121] 도 5에서 확장된 합성곱 신경망의 레이어가 3개라고 가정하는 경우, $t_1 \sim t_9$ 는 2차원 키 포인트, $c_1 \sim c_7$ 는 중간에서 계산된 히든 노드들(hidden nodes), p_1 은 확장된 합성곱 신경망에서 출력되는 3차원 키 포인트를 각각 나타낸다.
- [0122] 이때, 도 4의 (a)에 도시된 일반적인 합성곱 신경망의 경우 p_1 을 출력하기 위해 c_1, c_2, c_3 노드를 사용하게 되며, 이 경우 p_1 을 계산하기 위해 활용되는 입력의 범위는 c_1, c_2, c_3 를 계산하기 위해 사용되었던 $(t_1, t_2, t_3), (t_2, t_3, t_4), (t_3, t_4, t_5)$ 가 된다. 즉, 일반적인 합성곱 신경망에서의 입력의 범위는 $t_1 \sim t_5$ 까지 포함된다.
- [0123] 반면, 도 4의 (b)에 도시된 확장된 합성곱 신경망의 경우 p_1 을 출력하기 위해 c_1, c_4, c_7 노드를 사용하게 되며, 이 경우 p_1 을 계산하기 위해 활용되는 입력의 범위는 c_1, c_4, c_7 를 계산하기 위해 사용되었던 $(t_1, t_2, t_3), (t_4, t_5, t_6), (t_7, t_8, t_9)$ 인 $t_1 \sim t_9$ 으로 일반적인 합성곱 신경망에 비해 더 넓다. 즉, 일반적인 합성곱 신경망에 따른 필터에서의 파라미터 벡터를 $[a, b, c]$ 라 가정하는 경우 확장된 합성곱 신경망에 따른 필터에서의 파라미터 벡터는 $[a, 0, 0, b, 0, 0, c]$ 가 될 수 있다. 이와 같이, 확장된 합성곱 신경망은 필터가 가지고 있는 파라미터들의 중간에 0을 삽입함으로써 기존 계산 비용은 그대로 유지하되 포즈 추정 과정에서 시간적으로 멀리 떨어져 있는 프레임(사용자 이미지 또는 영상)의 키 포인트에 대한 정보도 활용할 수 있으며, 이에 따라 입력 범위가 넓어지게 된다. 한편, 일반적인 합성곱 신경망 또한 모델의 레이어를 많이 쌓아 올리면 입력 범위를 넓힐 수 있으나, 이 경우 모델의 파라미터가 기하급수적으로 늘어나 모델의 속도가 느려지게 된다.
- [0125] 도 6은 본 발명의 일 실시예에 따른 3차원 포즈 추정부(106)에서 출력되는 사용자 이미지 또는 영상에 대한 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 나타낸 예시이다.
- [0126] 도 6의 (a)는 3차원 포즈 추정부(106)에서 출력되는 사용자 이미지 또는 영상에 대한 3차원 키 포인트를 나타내며, 도 6의 (b)는 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)을 나타낸다.
- [0127] 상술한 바와 같이, 3차원 포즈 추정부(106)는 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신할 수 있다. 구체적으로, 3차원 포즈 추정부(106)는 상기 확장된 합성곱 신경망의 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시킬 수 있다. 이때, 상기 파라미터의 최적화 방식은 예를 들어, 경사하강법(gradient descent) 중의 하나인 Adam(Adaptive moment estimation) 기법일 수 있다. 본 발명의 일 실시예에 따르면, 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값을 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신함으로써, 확장된 합성곱 신경

망의 성능을 개선하고 최적화시킬 수 있다.

- [0129] 도 7은 본 발명의 일 실시예에 따른 3차원 포즈 추정 방법을 설명하기 위한 흐름도이다. 도시된 흐름도에서는 상기 방법을 복수 개의 단계로 나누어 기재하였으나, 적어도 일부의 단계들은 순서를 바꾸어 수행되거나, 다른 단계와 결합되어 함께 수행되거나, 생략되거나, 세부 단계들로 나뉘어 수행되거나, 또는 도시되지 않은 하나 이상의 단계가 부가되어 수행될 수 있다.
- [0130] 단계 S102에서, 입력부(102)는 사용자 이미지 또는 영상을 입력 받는다. 입력부(102)는 예를 들어, 오픈 데이터 셋인 Human 3.6M Dataset 으로부터 상기 사용자 이미지 또는 영상을 입력 받을 수 있다.
- [0131] 단계 S104에서, 추출부(104)는 입력 받은 상기 사용자 이미지 또는 영상 내 관절영역의 2차원 키 포인트 및 상기 관절영역에 대한 가시성 지표를 추출한다.
- [0132] 단계 S106에서, 추출부(104)는 상기 2차원 키 포인트와 상기 가시성 지표를 결합하여 입력 데이터를 생성한다. 이때, 추출부(104)는 상기 2차원 키 포인트 및 상기 가시성 지표에 각각 합성곱 신경망, 정규화 함수 및 ReLU 활성화 함수를 순차적으로 적용함으로써 전처리 작업을 수행하고, 상기 전처리 작업이 수행된 2차원 키 포인트 및 가시성 지표를 연결(concatenation)하여 입력 데이터를 생성할 수 있다.
- [0133] 단계 S108에서, 3차원 포즈 추정부(106)는 상기 입력 데이터를 학습 모델에 입력한다. 여기서, 학습 모델은 예를 들어, 확장된 합성곱 신경망일 수 있다.
- [0134] 단계 S110에서, 3차원 포즈 추정부(106)는 상기 관절영역에 대한 3차원 키 포인트를 도출한다.
- [0135] 단계 S112에서, 3차원 포즈 추정부(106)는 상기 확장된 합성곱 신경망을 통해 출력되는 3차원 키 포인트와 상기 사용자 이미지 또는 영상에 대한 실제 3차원 관절 좌표 값(ground truth)를 비교하여 상기 확장된 합성곱 신경망의 파라미터를 갱신한다. 구체적으로, 3차원 포즈 추정부(106)는 상기 확장된 합성곱 신경망의 목적 함수를 최소화시키는 방향으로 상기 파라미터를 최적화시킬 수 있다.
- [0137] 도 8은 예시적인 실시예들에서 사용되기에 적합한 컴퓨팅 장치를 포함하는 컴퓨팅 환경을 예시하여 설명하기 위한 블록도이다. 도시된 실시예에서, 각 컴포넌트들은 이하에 기술된 것 이외에 상이한 기능 및 능력을 가질 수 있고, 이하에 기술되지 것 이외에도 추가적인 컴포넌트를 포함할 수 있다.
- [0138] 도시된 컴퓨팅 환경(10)은 컴퓨팅 장치(12)를 포함한다. 일 실시예에서, 컴퓨팅 장치(12)는 3차원 포즈 생성 시스템(100), 또는 3차원 포즈 생성 시스템(100)에 포함되는 하나 이상의 컴포넌트일 수 있다.
- [0139] 컴퓨팅 장치(12)는 적어도 하나의 프로세서(14), 컴퓨터 판독 가능 저장 매체(16) 및 통신 버스(18)를 포함한다. 프로세서(14)는 컴퓨팅 장치(12)로 하여금 앞서 언급된 예시적인 실시예에 따라 동작하도록 할 수 있다. 예컨대, 프로세서(14)는 컴퓨터 판독 가능 저장 매체(16)에 저장된 하나 이상의 프로그램들을 실행할 수 있다. 상기 하나 이상의 프로그램들은 하나 이상의 컴퓨터 실행 가능 명령어를 포함할 수 있으며, 상기 컴퓨터 실행 가능 명령어는 프로세서(14)에 의해 실행되는 경우 컴퓨팅 장치(12)로 하여금 예시적인 실시예에 따른 동작들을 수행하도록 구성될 수 있다.
- [0140] 컴퓨터 판독 가능 저장 매체(16)는 컴퓨터 실행 가능 명령어 내지 프로그램 코드, 프로그램 데이터 및/또는 다른 적합한 형태의 정보를 저장하도록 구성된다. 컴퓨터 판독 가능 저장 매체(16)에 저장된 프로그램(20)은 프로세서(14)에 의해 실행 가능한 명령어의 집합을 포함한다. 일 실시예에서, 컴퓨터 판독 가능 저장 매체(16)는 메모리(랜덤 액세스 메모리와 같은 휘발성 메모리, 비휘발성 메모리, 또는 이들의 적절한 조합), 하나 이상의 자기 디스크 저장 디바이스들, 광학 디스크 저장 디바이스들, 플래시 메모리 디바이스들, 그 밖에 컴퓨팅 장치(12)에 의해 액세스되고 원하는 정보를 저장할 수 있는 다른 형태의 저장 매체, 또는 이들의 적합한 조합일 수 있다.
- [0141] 통신 버스(18)는 프로세서(14), 컴퓨터 판독 가능 저장 매체(16)를 포함하여 컴퓨팅 장치(12)의 다른 다양한 컴포넌트들을 상호 연결한다.
- [0142] 컴퓨팅 장치(12)는 또한 하나 이상의 입출력 장치(24)를 위한 인터페이스를 제공하는 하나 이상의 입출력 인터페이스(22) 및 하나 이상의 네트워크 통신 인터페이스(26)를 포함할 수 있다. 입출력 인터페이스(22) 및 네트워크

크 통신 인터페이스(26)는 통신 버스(18)에 연결된다. 입출력 장치(24)는 입출력 인터페이스(22)를 통해 컴퓨팅 장치(12)의 다른 컴포넌트들에 연결될 수 있다. 예시적인 입출력 장치(24)는 포인팅 장치(마우스 또는 트랙패드 등), 키보드, 터치 입력 장치(터치패드 또는 터치스크린 등), 음성 또는 소리 입력 장치, 다양한 종류의 센서 장치 및/또는 촬영 장치와 같은 입력 장치, 및/또는 디스플레이 장치, 프린터, 스피커 및/또는 네트워크 카드와 같은 출력 장치를 포함할 수 있다. 예시적인 입출력 장치(24)는 컴퓨팅 장치(12)를 구성하는 일 컴포넌트로서 컴퓨팅 장치(12)의 내부에 포함될 수도 있고, 컴퓨팅 장치(12)와는 구별되는 별개의 장치로 컴퓨팅 장치(12)와 연결될 수도 있다.

[0144] 이상에서 대표적인 실시예를 통하여 본 발명에 대하여 상세하게 설명하였으나, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자는 전술한 실시예에 대하여 본 발명의 범주에서 벗어나지 않는 한도 내에서 다양한 변형이 가능함을 이해할 것이다. 그러므로 본 발명의 권리범위는 설명된 실시예에 국한되어 정해져서는 안 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

부호의 설명

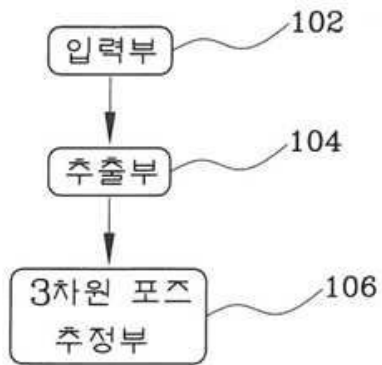
[0146]

- 10 : 컴퓨팅 환경
- 12 : 컴퓨팅 장치
- 14 : 프로세서
- 16 : 컴퓨터 판독 가능 저장 매체
- 18 : 통신 버스
- 20 : 프로그램
- 22 : 입출력 인터페이스
- 24 : 입출력 장치
- 26 : 네트워크 통신 인터페이스
- 100 : 3차원 포즈 추정 시스템
- 102 : 입력부
- 104 : 추출부
- 106 : 3차원 포즈 추정부

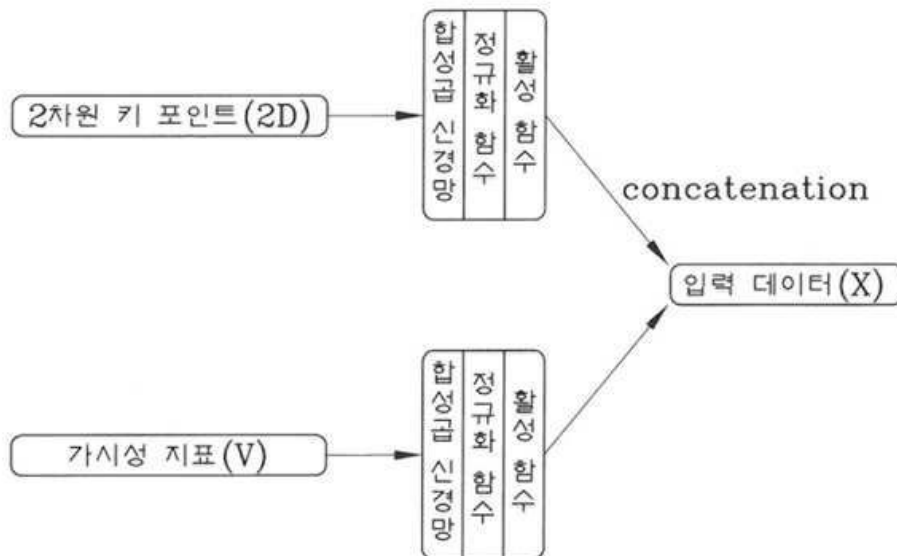
도면

도면1

100



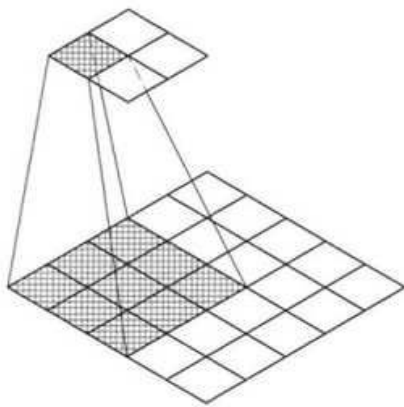
도면2



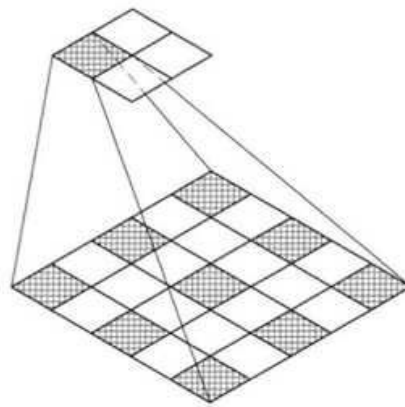
도면3



도면4

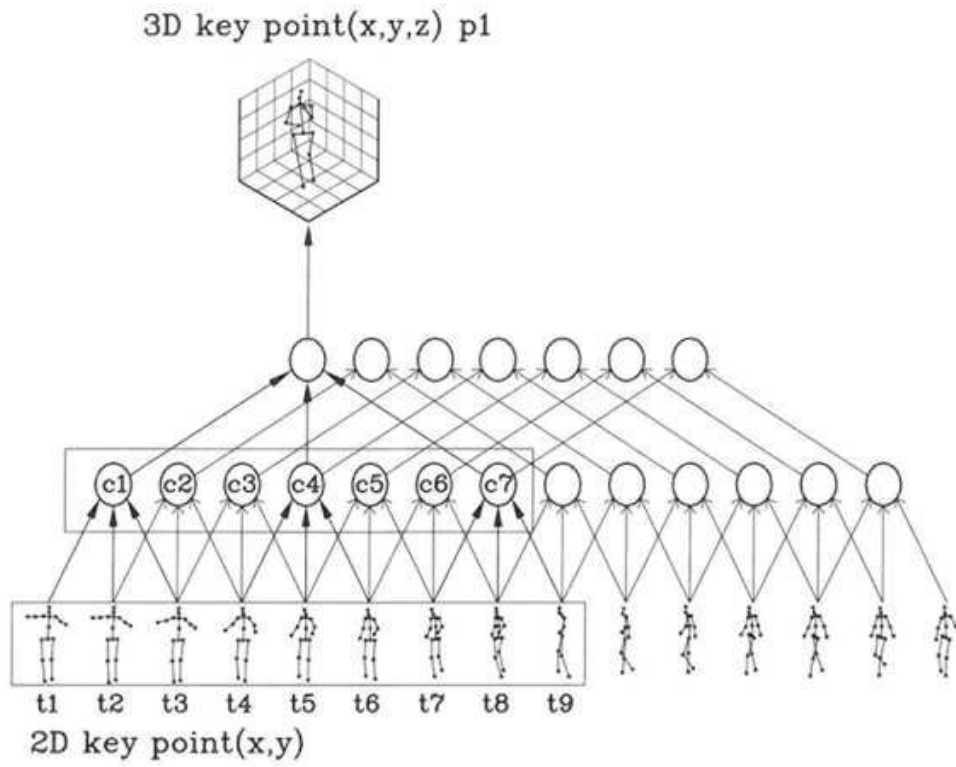


(a)

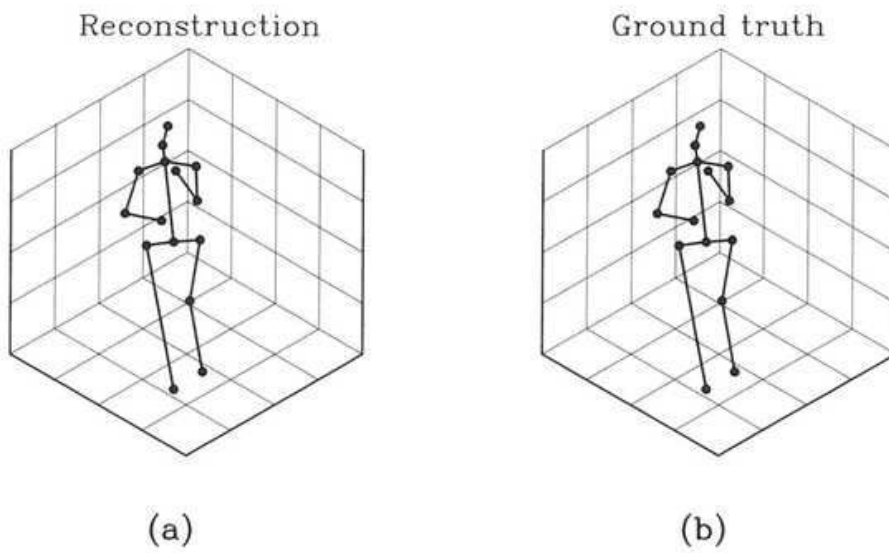


(b)

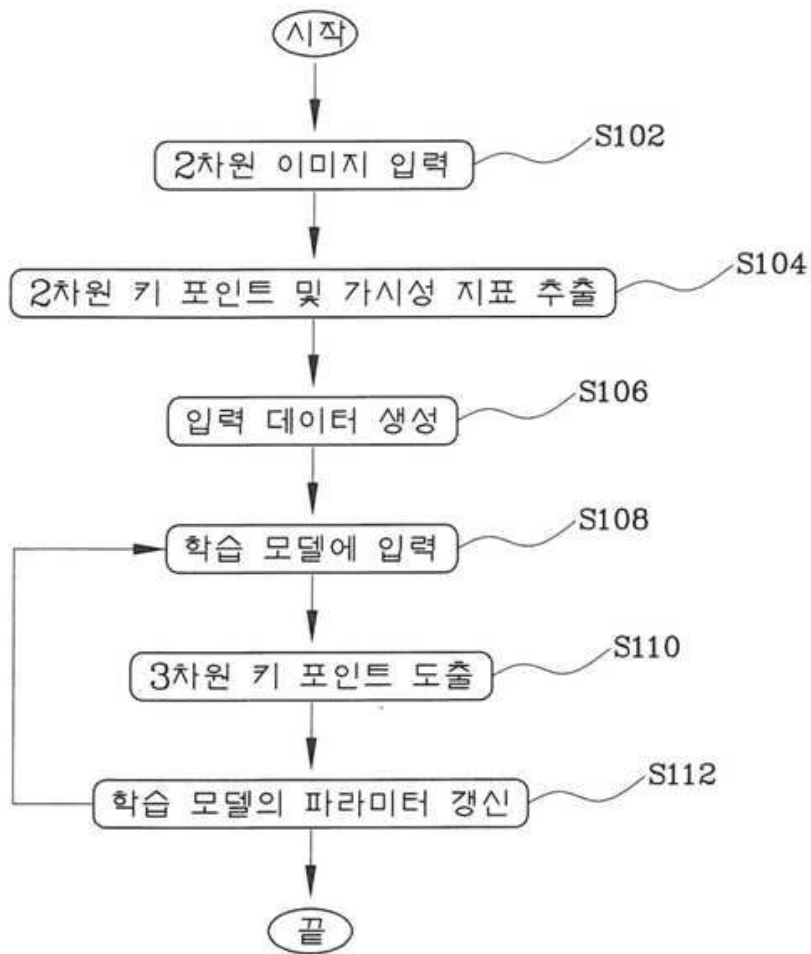
도면5



도면6



도면7



도면8

