



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2024년03월14일
(11) 등록번호 10-2648270
(24) 등록일자 2024년03월12일

(51) 국제특허분류(Int. Cl.)
G06N 3/08 (2023.01) G06N 3/04 (2023.01)
G06N 5/02 (2023.01)
(52) CPC특허분류
G06N 3/08 (2023.01)
G06N 3/04 (2023.01)
(21) 출원번호 10-2020-0185765
(22) 출원일자 2020년12월29일
심사청구일자 2020년12월29일
(65) 공개번호 10-2022-0094502
(43) 공개일자 2022년07월06일
(56) 선행기술조사문헌
KR1020180080081 A*
Yihui He et al., "Bounding Box Regression with Uncertainty for Accurate Object Detection," arXiv:1809.08545v3 [cs.CV] 16 Apr 2019 (2019.04.16.)*
Y. Tu et al., "An Accurate and Real-time Surface Defects Detection Method for Sawn Lumber", IEEE Transactions on Instrumentation and Measurement (2020.09.17.)*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
포항공과대학교 산학협력단
경상북도 포항시 남구 청암로 77 (지곡동)
(72) 발명자
원주철
경상북도 포항시 북구 양학로9번길 7, 203호
조성현
경상북도 포항시 남구 지곡로 278, 207동 102호
(74) 대리인
특허법인이상

전체 청구항 수 : 총 10 항

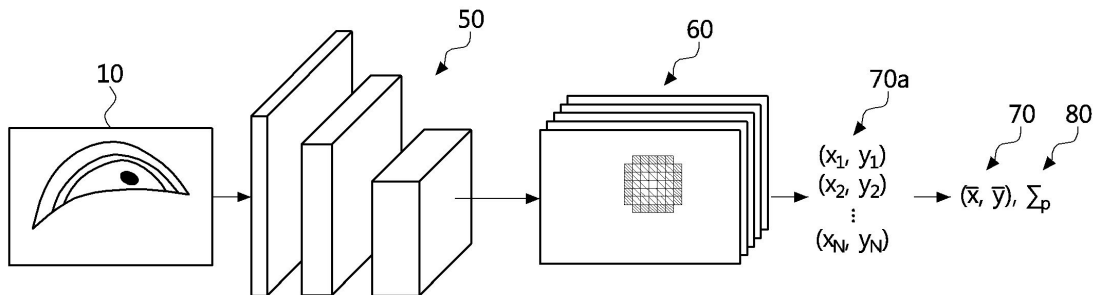
심사관 : 송근배

(54) 발명의 명칭 이미지에서의 좌표 및 불확실성 추정 시스템 및 방법

(57) 요약

컨벌루션 신경망으로 이미지에서 좌표 추정을 할 때 히트맵으로부터 각 좌표의 단일 값이 아닌 각 좌표의 정규 분포를 추정하고 추정한 분포에 기반하여 좌표 및 각 좌표에 대응하는 불확실성을 함께 추정하는 시스템 및 방법이 개시된다. 이미지에서의 좌표 및 불확실성 추정 방법은, 주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 방법으로서, 컨벌루션 신경망으로 복수의 초기 좌표들을 획득하는 단계, 초기 좌표들에 대한 손실 함수를 정의하는 단계 및 초기 좌표들로부터 최종 좌표와 최종 좌표의 불확실성을 획득하는 단계를 포함한다.

대표도



(52) CPC특허분류
G06N 5/02 (2023.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711116260
과제번호	2017M3C4A7066316
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	차세대정보·컴퓨팅기술개발(R&D)
연구과제명	초실감 원격가상 인터랙션을 위한 사용자 복원 기술 개발
기여율	1/1
과제수행기관명	고려대학교
연구기간	2020.04.01 ~ 2020.12.31

명세서

청구범위

청구항 1

주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 방법으로서,

프로세서에 의해, 입력 이미지에서 컨벌루션 신경망을 통해 복수의 히트맵들을 생성하는 단계;

상기 프로세서에 의해, 상기 컨벌루션 신경망으로 상기 복수의 히트맵들의 개수와 동일한 개수로 상기 복수의 히트맵들에서 복수의 초기 좌표들을 추정하는 단계-상기 컨벌루션 신경망은 추정하는 복수의 초기 좌표들이 정규 분포를 따를 수 있도록 정규 분포의 확률 밀도 함수인 목적 함수를 최대화하는 방향으로 학습되고, 상기 컨벌루션 신경망의 손실 함수로서 상기 목적 함수의 음의 로그를 사용함-;

상기 프로세서에 의해, 상기 초기 좌표들의 대표값으로 최종 좌표를 추정하는 단계; 및

상기 프로세서에 의해, 상기 최종 좌표와 상기 복수의 초기 좌표들의 표준 편차를 토대로 상기 최종 좌표의 불확실성을 계산하는 단계;

를 포함하는 이미지에서의 좌표 및 불확실성 추정 방법.

청구항 2

청구항 1에 있어서,

상기 복수의 초기 좌표들의 개수는 상기 최종 좌표의 개수의 N배이고, 여기서 상기 N은 2 이상의 자연수인, 이미지에서의 좌표 및 불확실성 추정 방법.

청구항 3

청구항 1에 있어서,

상기 목적 함수는 수학식 2와 같이 정의되고,

[수학식 2]

$$f(p) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_p|}} \exp\left(-\frac{1}{2} (p_{\text{mean}} - p_{\text{gt}})^T \Sigma_p^{-1} (p_{\text{mean}} - p_{\text{gt}})\right)$$

상기 수학식 2에서, P_{mean} 은 초기 좌표들(P)의 평균을, P_{gt} 는 정답 좌표를 각각 나타내며, 불확실성을 나타내는 분산(Σ_p)은 초기 좌표들(P)의 분산을 따르는, 이미지에서의 좌표 및 불확실성 추정 방법.

청구항 4

청구항 1에 있어서,

상기 손실 함수는 수학식 3과 같이 정의되고,

[수학식 3]

$$\mathcal{L} = \log(\Sigma_p) + \frac{(p_{\text{mean}} - p_{\text{gt}})^2}{\Sigma_p}$$

상기 수학식 3에서, Σ_p 는 불확실성을 나타내는 분산을, P는 초기 좌표들을, P_{mean} 은 초기 좌표들(P)의 평균을, P_{gt} 는 정답 좌표를 각각 나타내는, 이미지에서의 좌표 및 불확실성 추정 방법.

청구항 5

청구항 1에 있어서,

상기 프로세서에 의해, 상기 손실 함수를 최소화하도록 상기 컨벌루션 신경망을 학습하는 단계를 더 포함하는, 이미지에서의 좌표 및 불확실성 추정 방법.

청구항 6

주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 프로세서를 포함하는 추정 시스템으로서, 상기 프로세서에 의해,

입력 이미지에서 컨벌루션 신경망을 통해 복수의 히트맵들을 생성하는 단계;

상기 컨벌루션 신경망으로 상기 복수의 히트맵들의 개수와 동일한 개수로 상기 복수의 히트맵들에서 복수의 초기 좌표들을 추정하는 단계-상기 컨벌루션 신경망은 추정하는 복수의 초기 좌표들이 정규 분포를 따를 수 있도록 정규 분포의 확률 밀도 함수인 목적 함수를 최대화하는 방향으로 학습되고, 상기 컨벌루션 신경망의 손실 함수로서 상기 목적 함수의 음의 로그를 사용함-;

상기 복수의 초기 좌표들의 대표값으로 최종 좌표를 추정하는 단계; 및

상기 최종 좌표와 상기 복수의 초기 좌표들의 표준 편차를 토대로 상기 최종 좌표의 불확실성을 계산하는 단계;를 수행하는,

이미지에서의 좌표 및 불확실성 추정 시스템.

청구항 7

청구항 6에 있어서,

상기 복수의 초기 좌표들의 개수는 상기 최종 좌표의 개수의 N배이고, 여기서 상기 N은 2 이상의 자연수인, 이미지에서의 좌표 및 불확실성 추정 시스템.

청구항 8

청구항 6에 있어서,

상기 목적 함수는 수학식 2와 같이 정의되고,

[수학식 2]

$$f(p) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_p|}} \exp\left(-\frac{1}{2} (p_{\text{mean}} - p_{\text{gt}})^T \Sigma_p^{-1} (p_{\text{mean}} - p_{\text{gt}})\right)$$

상기 수학식 2에서, P_{mean} 은 초기 좌표들(P)의 평균을, P_{gt} 는 정답 좌표를 각각 나타내며, 불확실성을 나타내는 분산(Σ_p)은 초기 좌표들(P)의 분산을 따르는, 이미지에서의 좌표 및 불확실성 추정 시스템.

청구항 9

청구항 6에 있어서,

상기 손실 함수는 수학식 3과 같이 정의되고,

[수학식 3]

$$\mathcal{L} = \log(\Sigma_p) + \frac{(p_{\text{mean}} - p_{\text{gt}})^2}{\Sigma_p}$$

상기 수학식 3에서, Σ_p 는 불확실성을 나타내는 분산을, P는 초기 좌표들을, P_{mean} 은 초기 좌표들(P)의 평균을, P_{gt} 는 정답 좌표를 각각 나타내는, 이미지에서의 좌표 및 불확실성 추정 시스템.

청구항 10

청구항 6에 있어서,

상기 프로세서에 의해, 상기 손실 함수를 최소화하도록 상기 컨벌루션 신경망을 학습하는 단계를 더 수행하는, 이미지에서의 좌표 및 불확실성 추정 시스템.

발명의 설명

기술 분야

[0001] 본 발명은 이미지에서 좌표 및 불확실성을 추정하는 기술에 관한 것으로, 보다 상세하게는 컨벌루션 신경망(Convolutional Neural Networks)으로 이미지에서 좌표 추정을 할 때 히트맵(heatmap)으로부터 각 좌표의 단일 값이 아닌 각 좌표의 정규 분포를 추정하고 추정된 분포에 기반하여 좌표 및 각 좌표에 대응하는 불확실성(uncertainty)을 함께 추정하는 시스템 및 방법에 관한 것이다.

배경 기술

[0002] 동공의 중심 좌표를 추정할 수 있다면 시선의 방향 또는 위치를 추적할 수 있다. 또한, 얼굴 랜드마크의 좌표를 추정할 수 있다면 표정의 변화 또는 얼굴의 움직임 등을 파악할 수 있다. 이처럼 이미지에서 관심이 있는 부분의 좌표를 정확하게 추정할 수 있으면 많은 것들을 할 수 있다.

[0003] 신경망을 이용하면 이미지에서 좌표 추정을 효과적으로 수행할 수 있는데 그 경우 신경망은 입력 이미지로부터 사전에 설정된 개수만큼의 좌표를 추정하도록 설계된다.

[0004] 그러나 기본적인 구조의 신경망은 추정된 좌표가 얼마나 정확한 값인지에 대한 정보를 전혀 알지 못한다. 신경망이 추정한 결과에 대해 예상되는 오차의 범위 또는 확실하지 않은 정도를 불확실성이라고 한다.

[0005] 불확실성을 획득할 수 있으면 사용자가 신경망의 추정 결과를 선택적으로 사용할 수 있으며 불확실성에 기반한 후처리를 통해 추정 결과의 보정도 가능하다. 예를 들어, 실시간 시선 추적을 위해 동공의 중심 좌표를 추정하는 시스템에서, 추정된 동공 중심 좌표의 불확실성이 높다면 불확실성이 낮은 인접 프레임의 추정 좌표를 이용해서 더 정확한 시선 추적이 가능하다.

[0006] 한편, 신경망의 추정값에 대한 불확실성을 추정하는 방법들이 기존에 연구되고 있었다. 한 연구 결과의 [참고문헌 1]에서는 드랍아웃(dropout)을 활용하여 불확실성을 추정할 수 있는 베이지안 신경망(Bayesian Neural Networks)을 근사화하는 방법을 제공한다. 그러나 이 종래 방법은 베이지안 신경망을 구성하기 위해 추가적으로 요구되는 학습 파라미터(trainable parameters)와 메모리(memory)가 많고 구현이 어렵다.

[0007] 또 다른 연구 결과의 [참고문헌 2]에서는 신경망이 추정값에 대한 불확실성을 동시에 직접적으로 추정하도록 설계한다. 이 종래 방법은 구현이 쉽지만 좌표 추정에서 좋은 성능을 보이는 히트맵 생성 기반의 컨벌루션 신경망에 적용하기에는 적합하지 않다.

[0008] 도 1은 비교예에 따른 이미지에서의 좌표 추정 방법을 설명하기 위한 개략도이다. 도 2는 또 다른 비교예에 따른 이미지에서의 좌표 추정 방법을 설명하기 위한 개략도이다.

[0009] 디지털 이미지에서 좌표를 추정하는 방법으로 컨벌루션 신경망이 좋은 성능을 보이고 있다. 컨벌루션 신경망을 이용해 좌표를 추정하는 방법은 대표적으로 두 가지 방법이 있다.

[0010] 첫 번째 방법은 도 1에 도시한 바와 같이 완전연결(fully-connected) 레이어(22)를 가진 컨벌루션 신경망(20)를 이용한 회귀 기반으로 이미지(10)에서 추정 좌표(30)를 직접 생성하는 방법이고, 두 번째 방법은 도 2에 도시한 바와 같이 입력 이미지(10)에서 컨벌루션 신경망(20a)을 통해 히트맵(25)을 생성하고 히트맵(25)의 최대값 또는 soft-argmax 등의 활성화 함수를 이용하여 추정 좌표(30)를 결정하는 방법이다.

[0011] 전자의 회귀 기반의 방법은 가장 단순하지만 컨벌루션 신경망(20)에서 막대한 양의 학습 파라미터를 요구하고 신경망이 깊어질수록 공간적 정보의 손실이 커져서 정확도가 상대적으로 낮다.

[0012] 반면, 후자의 히트맵 기반의 방법은 컨벌루션 신경망(20a)이 도 1의 컨벌루션 신경망(20)만큼 깊어지기 전에 히트맵(25)을 추정하는 구조이기 때문에 적은 양의 학습 파라미터를 요구하며 공간적 정보의 손실이 적고 정확도가 높다.

- [0013] 이처럼 좌표 추정에서는 히트맵 기반의 컨벌루션 신경망을 사용하는 것이 적합하고 좋은 성능을 가진다.
- [0014] 신경망의 추정값에 대한 불확실성을 추정하는 방법은 베이지안(Bayesian) 신경망을 사용하는 것으로부터 시작되었다. 그러나 베이지안 기반의 방법은 복잡하고 구현이 어려워서 적용하기 어렵기 때문에 비 베이지안(non-Bayesian) 기반의 방법들이 제안되고 있다.
- [0015] 도 3은 또 다른 비교예에 따른 이미지에서의 불확실성 추정 방법을 설명하기 위한 개략도이다.
- [0016] 도 3을 참조하면, 비 베이지안 기반으로 불확실성을 추정하는 대표적인 방법으로서, 이 방법은 컨벌루션 신경망(20b)이 추정값과 추정값의 불확실성을 동시에 직접적으로 추정하는 방법으로 쉽고 간단하다. 그러나 완전연결 레이어(22)를 이용해 불확실성을 직접적으로 추정하는 방법이기 때문에 좌표 추정에 적합한 히트맵 기반의 컨벌루션 신경망에 적용하기에는 적합하지 않다.
- [0017] 이와 같이 이미지에서의 좌표 추정에서 성능이 좋은 히트맵 기반의 컨벌루션 신경망에서 기존의 불확실성 추정 방법들은 적합하게 적용되기가 어려워 이에 대한 해결 방안이 요구되고 있는 실정이다.

선행기술문헌

비특허문헌

- [0018] (비특허문헌 0001) [참고문헌 1] Gal et al., Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, ICML 2016.
- (비특허문헌 0002) [참고문헌 2] Lakshminarayanan et al., Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles, NIPS 2017.

발명의 내용

해결하려는 과제

- [0019] 본 발명은 종래 기술의 문제점을 해결하기 위해 도출된 것으로, 본 발명의 목적은 히트맵 기반의 컨벌루션 신경망을 사용하여 좌표를 정확하게 추정할 수 있고 히트맵 기반의 신경망 구조에서도 적합하게 적용 가능한 비 베이지안(non-Bayesian) 기반의 불확실성 추정 방법을 제공하는데 있다.
- [0020] 본 발명의 또 다른 목적은 컨벌루션 신경망(Convolutional Neural Networks)으로 이미지에서 좌표 추정을 할 때 히트맵(heatmap)으로부터 각 좌표의 단일 값이 아닌 각 좌표의 정규 분포를 추정하고 추정한 분포에 기반하여 좌표 및 각 좌표에 대응하는 불확실성(uncertainty)을 함께 추정하는 시스템 및 방법을 제공하는데 있다.

과제의 해결 수단

- [0021] 상기 기술적 과제를 해결하기 위한 본 발명의 일 측면에 따른 이미지에서의 좌표 및 불확실성 추정 방법은, 주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 방법으로서, 컨벌루션 신경망으로 복수의 초기 좌표들을 획득하는 단계; 상기 초기 좌표들에 대한 손실 함수를 정의하는 단계; 및 상기 초기 좌표들로부터 최종 좌표와 상기 최종 좌표의 불확실성을 획득하는 단계를 포함한다.
- [0022] 일실시예에서, 상기 복수의 초기 좌표들을 획득하는 단계는, 상기 컨벌루션 신경망으로 복수의 히트맵을 생성하고, 상기 복수의 히트맵에서 상기 복수의 초기 좌표들을 획득한다.
- [0023] 일실시예에서, 상기 손실 함수를 정의하는 단계는, 상기 초기 좌표들이 구성하는 정규 분포의 평균이 정답 좌표와 같아지도록 할 때 상기 정규 분포의 확률 밀도 함수의 음의 로그를 손실 함수로 사용한다.
- [0024] 일실시예에서, 상기 불확실성을 획득하는 단계는, 상기 초기 좌표들의 평균이나 중앙값에 대응하는 대표값을 최종 좌표로 사용하고 상기 초기 좌표들의 표준 편차를 상기 최종 좌표에 대응하는 불확실성으로 사용한다.
- [0025] 상기 기술적 과제를 해결하기 위한 본 발명의 다른 측면에 따른 이미지에서의 좌표 및 불확실성 추정 시스템은 주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 시스템으로서, 컨벌루션 신경망으로 복수의 초기 좌표들을 생성하는 초기 좌표 획득부; 상기 초기 좌

표들에 대한 손실 함수를 정의하는 손실 함수 정의부; 및 상기 초기 좌표들로부터 최종 좌표와 상기 최종 좌표의 불확실성을 생성하는 불확실성 획득부를 포함한다.

- [0026] 일실시예에서, 상기 초기 좌표 획득부는, 상기 컨벌루션 신경망으로 복수의 히트맵을 생성하고 상기 복수의 히트맵에서 상기 복수의 초기 좌표들을 획득한다.
- [0027] 일실시예에서, 상기 손실 함수 정의부는, 상기 초기 좌표들이 구성하는 정규 분포의 평균이 정답 좌표와 같아지도록 할 때 상기 정규 분포의 확률 밀도 함수의 음의 로그를 손실 함수로 사용한다.
- [0028] 일실시예에서, 상기 불확실성 획득부는, 상기 초기 좌표들의 평균이나 중앙값에 대응하는 대표값을 최종 좌표로 사용하고 상기 초기 좌표들의 표준 편차를 상기 최종 좌표에 대응하는 불확실성으로 사용한다.

발명의 효과

- [0029] 전술한 이미지에서의 좌표 및 불확실성 추정 시스템 및 방법을 사용하는 경우에는, 히트맵 기반의 컨벌루션 신경망(Convolutional Neural Networks)을 사용하여 좌표를 정확하게 추정할 수 있고 히트맵 기반의 신경망 구조에서도 적합하게 적용 가능한 비 베이지안(non-Bayesian) 기반의 불확실성 추정 방법을 제공할 수 있다.
- [0030] 또한, 본 발명에 의하면, 컨벌루션 신경망으로 이미지에서 좌표 추정을 할 때 히트맵(heatmap)으로부터 각 좌표의 단일 값이 아닌 각 좌표의 정규 분포를 추정하고 추정된 분포에 기반하여 좌표 및 각 좌표에 대응하는 불확실성(uncertainty)을 함께 추정하는 시스템 및 방법을 제공할 수 있다.
- [0031] 또한, 본 발명에 의하면, 기존에 히트맵 기반의 컨벌루션 신경망으로 좌표를 추정하는 시스템이 있다면 그 구조를 그대로 유지하면서 불확실성까지 쉽게 추정이 가능한 시스템으로 손쉽게 변경할 수 있는, 이미지에서의 좌표 및 불확실성 추정 시스템을 제공할 수 있다.

도면의 간단한 설명

- [0032] 도 1은 비교예에 따른 이미지에서의 좌표 추정 방법을 설명하기 위한 개략도이다.
- 도 2는 다른 비교예에 따른 이미지에서의 좌표 추정 방법을 설명하기 위한 개략도이다.
- 도 3은 또 다른 비교예에 따른 이미지에서의 불확실성 추정 방법을 설명하기 위한 개략도이다.
- 도 4는 본 발명의 일실시예에 따른 이미지에서의 좌표 및 불확실성 추정 방법을 설명하기 위한 개략도이다.
- 도 5는 본 발명의 다른 실시예에 따른 이미지에서의 좌표 및 불확실성 추정 시스템에 대한 개략적인 블록도이다.
- 도 6은 도 5의 이미지에서의 좌표 및 불확실성 추정 시스템의 작동 원리를 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0033] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0034] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0035] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0036] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조

합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

- [0037] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가진 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0038] 이하, 첨부한 도면들을 참조하여, 본 발명의 바람직한 실시예를 보다 상세하게 설명하고자 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면상의 동일한 구성요소에 대해서는 동일한 참조부호를 사용하고 동일한 구성요소에 대해서 중복된 설명은 생략한다.
- [0039] 도 4는 본 발명의 실시시에 따른 이미지에서의 좌표 및 불확실성 추정 방법을 설명하기 위한 개략도이다.
- [0040] 도 4를 참조하면, 본 실시예에 따른 이미지에서의 좌표 및 불확실성 추정 방법(이하 간략히 '추정 방법'이라고 한다)은 입력 이미지(10)로부터 최종 좌표(70) 및 최종 좌표값의 불확실성(80)을 추정하기까지의 일련의 과정을 포함한다.
- [0041] 즉, 본 실시예의 추정 방법은 입력 이미지(10)에서 컨벌루션 신경망(50)으로 히트맵(60)을 생성하고 생성된 히트맵(60)으로부터 초기 좌표들(70a)을 추정한다. 초기 좌표들(70a)은 정규 분포를 따르도록 추정되는데 이를 위한 손실 함수를 제공한다.
- [0042] 마지막으로 초기 좌표들(70a)로부터 최종 좌표(70) 및 최종 좌표값의 불확실성(80)을 추정한다.
- [0043] 보다 상세히, 본 실시예의 추정 방법은 초기 좌표들(70a)을 정확하게 추정할 수 있도록 컨벌루션 신경망(50)으로 히트맵(60)을 생성한다. 히트맵(60)이 생성되면 히트맵(60)으로부터 좌표를 획득할 수 있는데 획득 가능한 초기 좌표들(70a)의 개수는 히트맵(60)의 개수와 동일하다. 1개의 초기 좌표를 추정하고 싶다면 1개의 히트맵을 생성하고 5개의 초기 좌표를 추정하고 싶다면 5개의 히트맵을 생성하면 된다.
- [0044] 본 실시예의 추정 방법은 최종 좌표(70)를 추정하기 전에 앞서 초기 좌표들(70a)을 먼저 추정하는데 초기 좌표들(70a)의 개수는 최종 좌표(70)의 개수의 N배의 개수로 결정한다. 즉, 1개의 최종 좌표를 추정하고 싶다면 N개의 초기 좌표들을 추정하도록 N개의 히트맵들(60)을 생성한다. N은 임의의 2 이상의 자연수이다.
- [0045] 추정된 N개의 초기 좌표들(70a)은 [수학식 1]과 같이 정의할 수 있다.
- [0046] [수학식 1]
- [0047]
$$p = [p_1 p_2 \dots p_N]^T, \quad p_n = [x_n y_n]$$
- [0048] 위의 [수학식 1]에서, x_n 과 y_n 은 각각 n번째 초기 좌표의 x 좌표와 y 좌표를, p_n 은 n번째 초기 좌표로 x_n 과 y_n 을 포함하는 2차원 벡터를, p 는 N개의 초기 좌표들을 표현하는 Nx2 차원의 벡터를 의미한다.
- [0049] 1개의 최종 좌표(70)를 추정하기 위해 N개의 초기 좌표들을 추정하는 이유는 N개의 초기 좌표들의 분포에 기반하여 불확실성을 추정할 수 있도록 하기 위함이다.
- [0050] 간략하게 부연하면, N개의 초기 좌표들이 비슷한 위치에 집중되어 있다면 높은 확률로 정답에 가까운 좌표를 추정할 수 있을 것이고, N개의 초기 좌표들이 넓게 퍼져 있다면 추정 범위가 넓기 때문에 정답 좌표를 추정할 수 있을 확률이 낮으며, 이 경우 불확실성이 높다고 판단할 수 있다.
- [0051] 본 실시예에서는 N개의 초기 좌표가 정답 좌표를 평균으로 하는 정규 분포를 따르도록 설계한다. 그리고 신경망이 추정하는 초기 좌표들이 정규 분포를 따를 수 있도록 하기 위해서 정규 분포의 확률 밀도 함수를 목적 함수로 정의한다.
- [0052] 목적 함수는 [수학식 2]와 같이 정의된다.

[0053] [수학식 2]

$$f(\mathbf{p}) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma_p|}} \exp\left(-\frac{1}{2}(\mathbf{p}_{\text{mean}} - \mathbf{p}_{\text{gt}})^T \Sigma_p^{-1} (\mathbf{p}_{\text{mean}} - \mathbf{p}_{\text{gt}})\right)$$

[0054]

[0055] 위의 [수학식 2]에서, \mathbf{p}_{mean} 은 초기 좌표들(\mathbf{p})의 평균을, \mathbf{p}_{gt} 는 정답 좌표를 각각 의미하며, 불확실성을 나타내는 분산(Σ_p)은 초기 좌표들(\mathbf{p})의 분산을 따른다.

[0056] 초기 좌표들의 평균이 정답 좌표와 가까워질수록, 초기 좌표들의 분산이 낮을수록 목적 함수는 커지게 된다.

[0057] N개의 초기 좌표들이 모두 정답 좌표와 같을 때 목적 함수는 최대값이 되며 이 경우 초기 좌표들의 평균은 정답 좌표와 같고 초기 좌표들의 분산은 0이다.

[0058] 본 실시예의 추정 방법은 신경망이 상기의 목적 함수를 최대화하는 방향으로 학습할 수 있도록 목적 함수에 음의 로그를 취하는 것으로 손실 함수를 정의한다.

[0059] 이러한 손실 함수는 [수학식 3]과 같이 정의할 수 있다.

[0060] [수학식 3]

$$\mathcal{L} = \log(\Sigma_p) + \frac{(\mathbf{p}_{\text{mean}} - \mathbf{p}_{\text{gt}})^2}{\Sigma_p}$$

[0061]

[0062] 위의 [수학식 3]을 손실 함수로 정의하고 손실 함수를 최소화하도록 신경망을 학습하면, 신경망은 목적 함수가 최대화되는 방향으로 학습된다. 궁극적으로 신경망은 최종 좌표와 불확실성을 추정할 수 있도록 정규 분포를 따르는 초기 좌표들을 추정할 수 있다.

[0063] 앞서 설명한 바와 같이, 손실 함수를 최소화하도록 잘 학습된 신경망으로 추정한 초기 좌표들이 주어졌을 때, 초기 좌표들의 평균 또는 중앙값 등의 대표값으로 최종 좌표를 추정할 수 있으며 초기 좌표들의 표준 편차로 최종 좌표의 불확실성을 계산할 수 있다.

[0064] 예를 들어, 추정된 초기 좌표들이 모두 같은 값을 가진다면 표준 편차가 0이고 추정된 최종 좌표의 불확실성이 0이 된다. 반대로 추정된 초기 좌표들의 값들이 넓게 퍼져 있어 표준 편차가 높다면 추정된 최종 좌표의 불확실성은 높을 것이며 이 경우 최종 좌표가 가지는 오차가 클 확률이 높다고 추정할 수 있다.

[0065] 이와 같이 신경망을 이용한 좌표 추정에서 불확실성을 추정할 수 있다면 사용자 입장에서 더 높은 신뢰도를 가지고 시스템을 사용할 수 있다. 불확실성을 추정하지 못한다면 큰 오차가 발생할 수 있더라도 항상 같은 신뢰도를 가지고 추정값을 사용해야 한다. 반면에 신경망이 추정값과 불확실성을 함께 제공한다면 불확실성이 낮을 때에는 높은 신뢰도를 가지고 추정값을 사용할 수 있고, 불확실성이 높다면 오차가 발생할 확률이 높다는 것을 사전에 인지할 수 있기 때문에 다른 조치를 취할 수 있다.

[0066] 또한, 불확실성에 기반한 후처리를 통해 좌표 추정 시스템의 성능을 향상시킬 수도 있다. 예를 들어, 실시간 시선 추적 시스템에서 빛 반사 또는 다른 객체로 인한 가려짐 등으로 인해 정확한 추정이 어려운 경우가 발생하기 쉬운데, 연속적인 프레임 형태의 양 쪽 눈의 이미지가 입력으로 주어질 때, 한 쪽 눈에서의 동공 중심 좌표 추정 결과의 불확실성이 높다면 불확실성이 상대적으로 낮은 반대쪽 눈에서의 추정 결과를 사용할 수 있고, 양 쪽 눈에서의 불확실성이 모두 높다면 불확실성이 낮은 이전 프레임에서의 추정 정보를 활용하는 것이 가능하다.

[0067] 도 5는 본 발명의 다른 실시예에 따른 이미지에서의 좌표 및 불확실성 추정 시스템(이하 간략히 '추정 시스템'이라 한다)에 대한 개략적인 블록도이다.

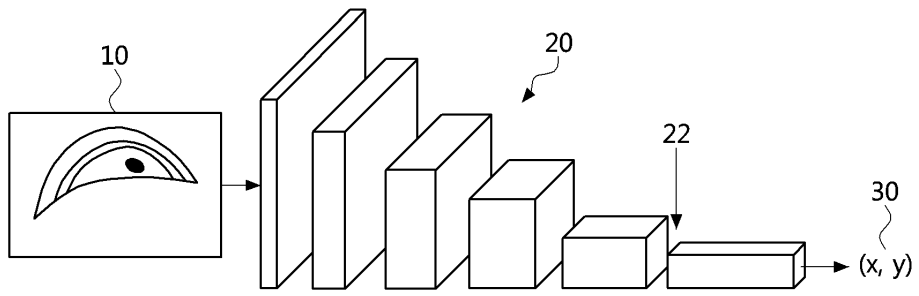
[0068] 도 5를 참조하면, 본 실시예에 따른 추정 시스템(100)은 컨벌루션 신경망(50)을 구비하는 초기 좌표 획득부(110), 손실 함수 정의부(120) 및 불확실성 획득부(130)를 포함하며, 주어진 입력 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정한다.

[0069] 이를 위해, 초기 좌표 획득부(110)는 컨벌루션 신경망(50)으로 복수의 초기 좌표들을 생성한다. 초기 좌표 획득부(110)는 컨벌루션 신경망(50)으로 복수의 히트맵을 생성하고 복수의 히트맵에서 복수의 초기 좌표들을 획득한다.

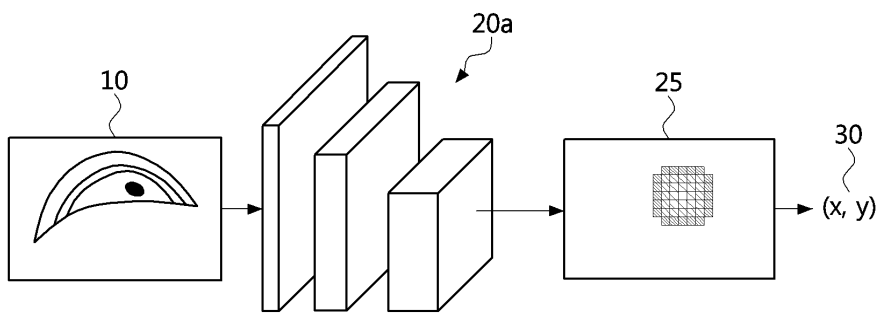
- [0070] 손실 함수 정의부(120)는 초기 좌표 획득부(110)로부터 받은 초기 좌표들에 대한 손실 함수를 정의한다. 손실 함수 정의부(120)는, 초기 좌표들이 구성하는 정규 분포의 평균이 미리 설정된 정답 좌표와 같아지도록 할 때 상기의 정규 분포의 확률 밀도 함수의 음의 로그를 손실 함수로 사용할 수 있다.
- [0071] 불확실성 획득부(130)는 초기 좌표들로부터 최종 좌표와 최종 좌표의 불확실성을 생성한다. 불확실성 획득부(130)는 초기 좌표들의 평균이나 중앙값에 대응하는 대표값을 최종 좌표로 사용하고 초기 좌표들의 표준 편차를 최종 좌표에 대응하는 불확실성으로 사용할 수 있다.
- [0072] 본 실시예에 따른 추정 시스템(100)은 전술한 컨벌루션 신경망(50), 초기 좌표 획득부(110), 손실 함수 정의부(120) 및 불확실성 획득부(130)를 소프트웨어 모듈 형태로 탑재한 프로세서로 구현될 수 있다. 이 경우, 프로세서는 메모리(200)에 연결되고, 메모리(200)는 컨벌루션 신경망(50), 초기 좌표 획득부(110), 손실 함수 정의부(120) 및 불확실성 획득부(130)와 그 기능을 수행하기 위한 소프트웨어 모듈이나 프로그램을 저장할 수 있다.
- [0073] 도 6은 도 5의 이미지에서의 좌표 및 불확실성 추정 시스템의 작동 원리를 설명하기 위한 흐름도이다.
- [0074] 도 6을 참조하면, 본 실시예에 따른 이미지에서의 좌표 및 불확실성 추정 방법은, 주어진 이미지에서 단일 또는 복수 개의 관심이 있는 부분의 좌표를 추정할 때 추정된 좌표의 불확실성을 함께 추정하는 방법으로서, 먼저 컨벌루션 신경망으로 복수의 초기 좌표들을 획득한다(S61). 본 단계(S61)는 컨벌루션 신경망으로 복수의 히트맵을 생성하고 복수의 히트맵에서 복수의 초기 좌표들을 획득하도록 구현될 수 있다.
- [0075] 다음, 초기 좌표들에 대한 손실 함수를 정의한다(S62). 본 단계(S62)는 초기 좌표들이 구성하는 정규 분포의 평균이 정답 좌표와 같아지도록 할 때 정규 분포의 확률 밀도 함수의 음의 로그를 손실 함수로 사용하도록 구현될 수 있다.
- [0076] 다음, 손실 함수가 최소화하도록 컨벌루션 신경망을 학습한다(S63). 본 단계(S63)는 추정 시스템의 학습(training) 모드에서만 실행되고, 테스트(test) 모드에서는 생략될 수 있다.
- [0077] 다음, 추정 시스템의 불확실성 획득부는 훈련된 컨벌루션 신경망을 통해 입력 이미지로부터 얻은 초기 좌표들로부터 최종 좌표와 최종 좌표의 불확실성을 획득한다(S64). 본 단계(S64)는 초기 좌표들의 평균이나 중앙값에 대응하는 대표값을 최종 좌표로 사용하고 초기 좌표들의 표준 편차를 최종 좌표에 대응하는 불확실성으로 사용할 수 있다.
- [0078] 한편, 전술한 실시예들을 통해 설명한 이미지에서의 좌표 및 불확실성 추정 방법은 다양한 컴퓨터 수단을 통해 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 컴퓨터 판독 가능 매체에 기록되는 프로그램 명령은 본 발명을 위해 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용가능한 것일 수 있다.
- [0079] 컴퓨터 판독 가능 매체의 예에는 롬(rom), 램(ram), 플래시 메모리(flash memory) 등과 같이 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러(compiler)에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터(interpreter) 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함한다. 상술한 하드웨어 장치는 본 실시예에 따른 이미지에서의 좌표 및 불확실성 추정 방법의 일련의 동작을 수행하기 위해 적어도 하나의 소프트웨어 모듈로 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0080] 이상 실시예를 참조하여 설명하였지만, 해당 기술 분야의 숙련된 당업자는 하기의 청구범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

도면

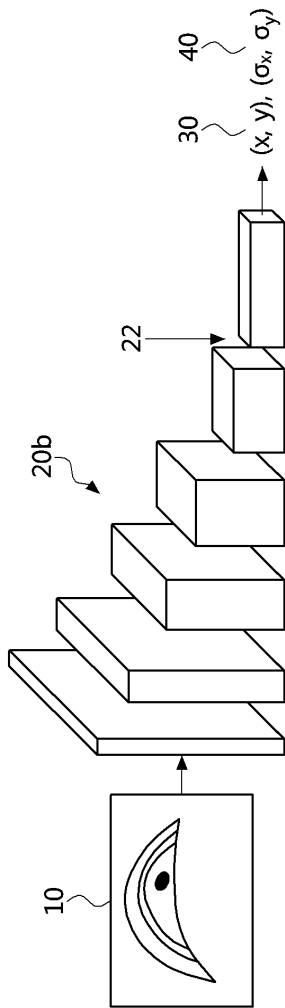
도면1



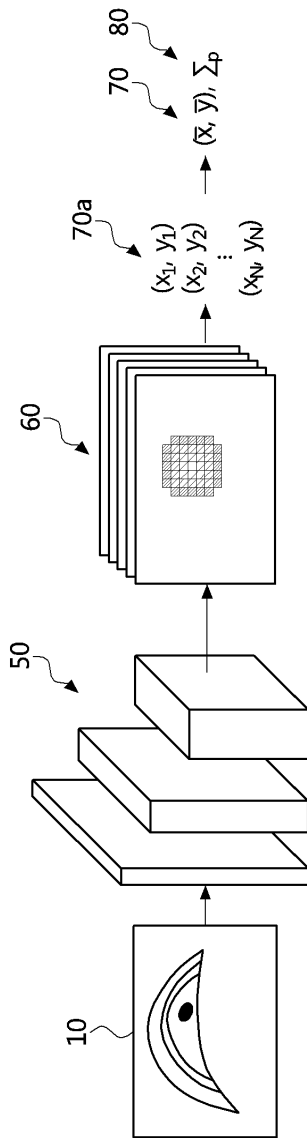
도면2



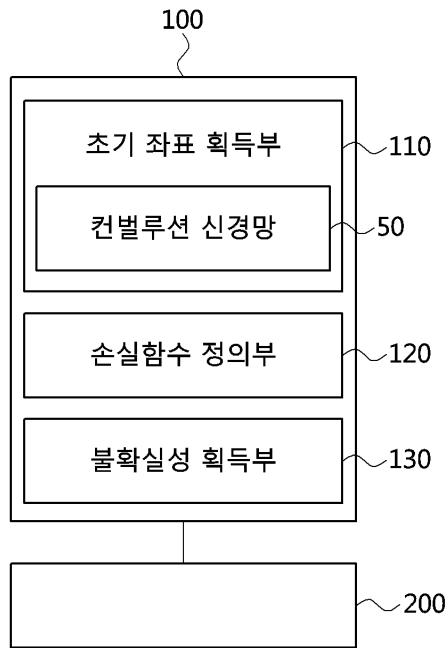
도면3



도면4



도면5



도면6

