

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2020-0061478
(43) 공개일자 2020년06월03일

(51) 국제특허분류(Int. Cl.)

G06T 5/50 (2006.01) G06T 3/40 (2006.01)

G06T 3/60 (2006.01) G06T 5/20 (2006.01)

(52) CPC특허분류

G06T 5/50 (2013.01)

G06T 3/40 (2013.01)

(21) 출원번호 10-2018-0146693

(22) 출원일자 2018년11월23일

심사청구일자 2018년11월23일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

이상훈

서울특별시 서대문구 연세로 50, 제3공학관 715호(신촌동)

허정우

서울특별시 서대문구 연세로 50 제3공학관 703호(신촌동)

(74) 대리인

민영준

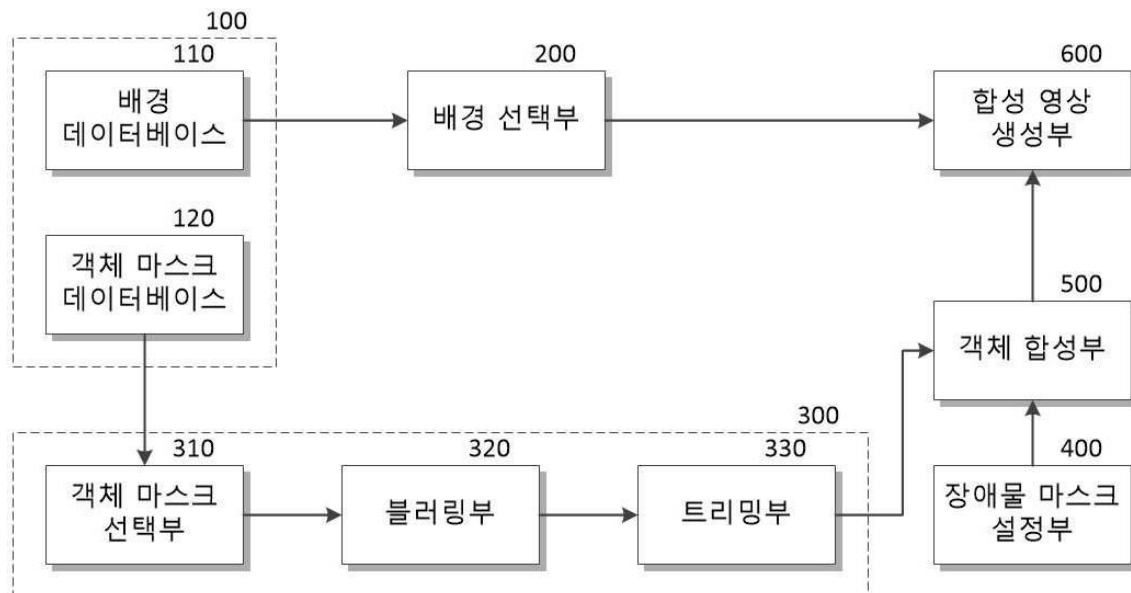
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 실시간 객체 탐지 시스템 학습을 위한 합성 데이터 생성 장치 및 방법

(57) 요약

본 발명은 다수의 배경 영상 및 다수의 객체 마스크가 저장된 데이터베이스부, 데이터베이스부에서 배경 영상을 선택하여 획득하는 배경 선택부, 데이터베이스부에서 적어도 하나의 객체 마스크를 선택하고, 선택된 적어도 하나의 객체 마스크에 대해 블러링, 크기 조절 및 회전각 조절 중 적어도 하나를 적용하여 변형하는 객체 마스크 설정부 및 배경 선택부에서 선택된 배경 영상에 변형된 적어도 하나의 객체 마스크로부터 획득되는 적어도 하나의 최종 객체 마스크를 합성하여 합성 영상을 생성하는 합성 영상 생성부를 포함하는 합성 데이터 생성 장치 및 방법을 제공할 수 있다.

대표도



(52) CPC특허분류

G06T 3/60 (2013.01)

G06T 5/20 (2013.01)

G06T 2207/20221 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2016-0-00406

부처명 과학기술정보통신부

연구관리전문기관 정보통신기술진흥센터(NIPA산하)

연구사업명 정보통신방송연구개발사업

연구과제명 [이지바로][주관/숙명여자대학교] (기반SW-창조씨앗2단계)SIAT형 CCTV 클라우드 플랫폼 기

술 개발 (2/5)

기 여 율 1/1

주관기관 숙명여자대학교 산학협력단

연구기간 2018.01.01 ~ 2018.12.31

명세서

청구범위

청구항 1

다수의 배경 영상 및 다수의 객체 마스크가 저장된 데이터베이스부;

상기 데이터베이스부에서 배경 영상을 선택하여 획득하는 배경 선택부;

상기 데이터베이스부에서 적어도 하나의 객체 마스크를 선택하고, 선택된 적어도 하나의 객체 마스크에 대해 블러링, 크기 조절 및 회전각 조절 중 적어도 하나를 적용하여 변형하는 객체 마스크 설정부; 및

상기 배경 선택부에서 선택된 상기 배경 영상에 변형된 적어도 하나의 객체 마스크로부터 획득되는 적어도 하나의 최종 객체 마스크를 합성하여 합성 영상을 생성하는 합성 영상 생성부; 를 포함하는 합성 데이터 생성 장치.

청구항 2

제1 항에 있어서, 상기 객체 마스크 설정부는

상기 데이터베이스부에 저장된 다수의 객체 마스크 중 적어도 하나의 객체 마스크를 선택하는 객체 마스크 선택부;

선택된 적어도 하나의 객체 마스크에 대해 기지정된 블러링 확률값에 따라 선택적으로 블러링하는 블러링부; 및

적어도 하나의 객체 마스크의 크기 및 회전각을 조절하여 변형된 적어도 하나의 객체 마스크를 출력하는 트리밍부; 를 포함하는 합성 데이터 생성 장치.

청구항 3

제2 항에 있어서, 상기 블러링부는

블러링 크기 및 블러링 방향각에 따라 상기 객체 마스크를 블러링하는 블러링 커널을 이용하여 객체 마스크를 블러링하는 합성 데이터 생성 장치.

청구항 4

제1 항에 있어서, 상기 합성 데이터 생성 장치는

기지정된 장애물 확률값에 따라 적어도 하나의 장애물 마스크를 획득하고, 획득된 장애물 마스크를 상기 객체 마스크 설정부에서 출력되는 객체 마스크와 합성하여, 상기 최종 객체 마스크를 상기 합성 영상 생성부로 전달하는 객체 합성부; 를 더 포함하는 합성 데이터 생성 장치.

청구항 5

제4 항에 있어서, 상기 객체 합성부는

적어도 하나의 장애물 마스크의 크기 및 회전각을 조절하여 변형하고, 변형된 장애물 마스크를 상기 객체 마스크와 합성하는 합성 데이터 생성 장치.

청구항 6

제1 항에 있어서, 상기 합성 영상 생성부는

상기 배경 영상에 적어도 하나의 최종 객체 마스크를 배치하여 합성하고, 최종 객체 마스크의 개수가 다수개이면, 배치된 다수개의 최종 객체 마스크 사이의 IOU(intersection of union)가 기지정된 허용값 미만인 되도록 배치 위치를 가변하며,

상기 합성 영상에 합성된 적어도 하나의 최종 객체 마스크의 각각에 대한 검증 레이블을 생성하는 합성 데이터 생성 장치.

청구항 7

미리 획득된 다수의 배경 영상 및 다수의 객체 마스크 중 하나의 배경 영상과 적어도 하나의 객체 마스크를 선택하는 단계;

선택된 적어도 하나의 객체 마스크에 대해 블러링, 크기 조절 및 회전각 조절 중 적어도 하나를 적용하여 변형하는 단계; 및

선택된 상기 배경 영상에 변형된 적어도 하나의 객체 마스크로부터 획득되는 적어도 하나의 최종 객체 마스크를 합성하여 합성 영상을 생성하는 단계; 를 포함하는 합성 데이터 생성 방법.

청구항 8

제7 항에 있어서, 상기 변형하는 단계는

상기 다수의 객체 마스크 중 적어도 하나의 객체 마스크를 선택하는 단계;

선택된 적어도 하나의 객체 마스크에 대해 기지정된 블러링 확률값에 따라 선택적으로 블러링하는 단계; 및

적어도 하나의 객체 마스크의 크기 및 회전각을 조절하여 변형된 적어도 하나의 객체 마스크를 출력하는 단계; 를 포함하는 합성 데이터 생성 방법.

청구항 9

제8 항에 있어서, 상기 블러링하는 단계는

블러링 크기 및 블러링 방향각에 따라 상기 객체 마스크를 블러링하는 블러링 커널을 이용하여 객체 마스크를 블러링하는 합성 데이터 생성 방법.

청구항 10

제8 항에 있어서, 상기 변형하는 단계는

기지정된 장애물 확률값에 따라 적어도 하나의 장애물 마스크를 획득하는 단계; 및

획득된 장애물 마스크를 변형된 객체 마스크와 합성하여 상기 최종 객체 마스크를 생성하는 단계; 를 포함하는 합성 데이터 생성 방법.

청구항 11

제10 항에 있어서, 상기 변형하는 단계는

획득된 상기 적어도 하나의 장애물 마스크의 크기 및 회전각을 조절하여 변형하는 단계; 를 더 포함하는 합성 데이터 생성 방법.

청구항 12

제7 항에 있어서, 상기 합성 영상을 생성하는 단계는

상기 배경 영상에 적어도 하나의 최종 객체 마스크를 배치하는 단계;

상기 최종 객체 마스크의 개수가 다수개이면, 배치된 다수개의 최종 객체 마스크 사이의 IOU(intersection of union)가 기지정된 허용값 미만인 되도록 배치 위치를 조절하는 단계; 및

상기 합성 영상에 합성된 적어도 하나의 최종 객체 마스크의 각각에 대한 검증 레이블을 생성하는 단계; 를 포함하는 합성 데이터 생성 방법.

발명의 설명

기술 분야

본 발명은 합성 데이터 생성 장치 및 방법에 관한 것으로, 실시간 객체 탐지 시스템 학습을 위한 합성 데이터 생성 장치 및 방법에 관한 것이다.

[0001]

배경 기술

- [0002] 최근 증강 현실(Augmented Reality: 이하 AR)이나 보안 시스템과 같은 다양한 응용 분야에서 객체를 식별하는 객체 탐지 기술이 요구되고 있다. 특히 AR 기술에서는 보다 현실적인 경험을 제공하기 위해 높은 수준의 사용자 상호 작용을 추구하고 있으며, 이에 실시간으로 주변 환경을 인식하는 것이 매우 중요하다.
- [0003] 현재 영상에서 어떠한 객체가 포함되어 있는지 인지할 수 있다면 주변 환경에 대한 충분한 정보를 제공할 수 있기 때문에, 실시간으로 주변 환경을 인식하기 위해서는 객체 탐지 기술 또한 실시간으로 수행될 수 있어야 한다.
- [0004] 이에 최근에는 실시간 객체 탐지 기술에 딥 러닝(Deep learning) 기법으로 학습된 인공 신경망(Artificial Neural Network)을 객체 탐지 시스템에 이용함으로써, 그 성능을 획기적으로 향상시키기 위한 연구가 계속되고 있다. 인공 신경망이 신뢰성 있는 성능을 나타내기 위해서는 다양한 환경에서 획득된 다수의 학습 데이터를 이용하여 학습시켜야만 한다.
- [0005] 일부 공용 데이터베이스들은 몇몇 종류의 객체에 대해 여러 영상들을 제공하므로 학습용 데이터를 획득할 수 있다. 그러나 탐지하고자 하는 객체가 해당 데이터베이스들에 존재하지 않는다면, 직접 학습 데이터를 제작해야 하는 상황이 발생한다.
- [0006] 객체 탐지를 위한 학습 데이터 제작은 영상에 존재하는 모든 목표 객체들에 대해 테두리를 수동으로 설정하여야 하며, 인공 신경망 모델을 학습시키기 위해 통상적으로 몇 만장 이상의 방대한 학습 데이터가 필요하기 때문에 실질적으로 매우 어려운 작업이다.
- [0007] 뿐만 아니라 실시간 객체 탐지 시스템을 학습시키고자 하는 경우, 해결되어야 하는 또 다른 문제가 있다.
- [0008] 이는 객체와 환경 사이의 복잡한 상호 작용에 의한 시각적 열화(visual degradation)이다. 시각적 열화는 일반적으로 모션 블러(motion blur)와 폐색(occlusion)에 의해 유발된다.
- [0009] 도1 은 시각적 열화가 유발된 영상의 일례를 나타낸다.
- [0010] 도1 에서 (a)는 모션 블러에 의한 시각적 열화가 발생된 영상을 나타내고, (b)는 폐색에 의한 시각적 열화가 발생된 영상을 나타낸다.
- [0011] 모션 블러는 (a)에 도시된 바와 같이, 객체(여기서는 야구 배트)의 움직임에 따라 영상에 객체가 흐리게 나타나는 현상을 의미하며, 이는 저품질의 카메라에서 주로 발생한다. 이러한 모션 블러에 의해 객체의 윤곽이 변형됨에 따라, 개별 객체에 각각에 대해 객체 탐지 시스템이 이미 학습되었더라도, 객체를 탐지하지 못하는 결과를 초래한다.
- [0012] 폐색은 (b)에 도시된 바와 같이, 장애물(distractor)(여기서는 사람의 손)에 의해 영상에서 객체(여기서는 나이프)의 일부가 가려지는 현상을 의미하며, 이 또한 객체 탐지 시스템이 객체의 윤곽을 판별할 수 없도록 함으로써, 객체에 대한 학습 여부에 무관하게 객체를 탐지하지 못하는 결과를 초래한다.
- [0013] 만일 객체 탐지 시스템이 시각적 열화에도 불구하고 객체를 탐지할 수 있도록 학습 시키기 위해서는 각 객체에 대한 다양한 형태의 시각적 열화가 반영된 학습 데이터를 이용하여야 하므로, 요구되는 학습 데이터의 양이 기하급수적으로 증가하게 되는 문제를 초래한다.
- [0014] 따라서 객체 탐지 시스템이 효율적으로 제작되더라도, 학습 데이터의 부족으로 인해 객체 탐지 장치가 신뢰성 있게 객체를 탐지하지 못하는 문제가 발생한다.

선행기술문헌

특허문헌

- [0015] (특허문헌 0001) 한국 공개 특허 제10-2017-0037024호 (2017.03.23 공개)

발명의 내용

해결하려는 과제

[0016] 본 발명의 목적은 실시간 객체 탐지 시스템을 학습 시키기 위한 다수의 학습용 데이터를 생성할 수 있는 합성 데이터 생성 장치 및 방법을 제공하는데 있다.

[0017] 본 발명의 다른 목적은 모션 블러 및 페색과 같은 시각 열화 현상에서도 객체 탐지 시스템이 용이하게 객체를 탐지할 수 있도록 하는 합성 데이터 생성 장치 및 방법을 제공하는데 있다.

과제의 해결 수단

[0018] 상기 목적을 달성하기 위한 본 발명의 일 실시예에 따른 합성 데이터 생성 장치는 다수의 배경 영상 및 다수의 객체 마스크가 저장된 데이터베이스부; 상기 데이터베이스부에서 배경 영상을 선택하여 획득하는 배경 선택부; 상기 데이터베이스부에서 적어도 하나의 객체 마스크를 선택하고, 선택된 적어도 하나의 객체 마스크에 대해 블러링, 크기 조절 및 회전각 조절 중 적어도 하나를 적용하여 변형하는 객체 마스크 설정부; 및 상기 배경 선택부에서 선택된 상기 배경 영상에 변형된 적어도 하나의 객체 마스크로부터 획득되는 적어도 하나의 최종 객체 마스크를 합성하여 합성 영상을 생성하는 합성 영상 생성부; 를 포함한다.

[0019] 상기 객체 마스크 설정부는 상기 데이터베이스부에 저장된 다수의 객체 마스크 중 적어도 하나의 객체 마스크를 선택하는 객체 마스크 선택부; 선택된 적어도 하나의 객체 마스크에 대해 기지정된 블러링 확률값에 따라 선택적으로 블러링하는 블러링부; 및 적어도 하나의 객체 마스크의 크기 및 회전각을 조절하여 변형된 적어도 하나의 객체 마스크를 출력하는 트리밍부; 를 포함할 수 있다.

[0020] 상기 블러링부는 블러링 크기 및 블러링 방향각에 따라 상기 객체 마스크를 블러링하는 블러링 커널을 이용하여 객체 마스크를 블러링할 수 있다.

[0021] 상기 합성 데이터 생성 장치는 기지정된 장애물 확률값에 따라 적어도 하나의 장애물 마스크를 획득하고, 획득된 장애물 마스크를 상기 객체 마스크 설정부에서 출력되는 객체 마스크와 합성하여, 상기 최종 객체 마스크를 상기 합성 영상 생성부로 전달하는 객체 합성부; 를 더 포함할 수 있다.

[0022] 상기 객체 합성부는 적어도 하나의 장애물 마스크의 크기 및 회전각을 조절하여 변형하고, 변형된 장애물 마스크를 상기 객체 마스크와 합성할 수 있다.

[0023] 상기 합성 영상 생성부는 상기 배경 영상에 적어도 하나의 최종 객체 마스크를 배치하여 합성하고, 최종 객체 마스크의 개수가 다수개이면, 배치된 다수개의 최종 객체 마스크 사이의 IOU(intersection of union)가 기지정된 허용값 미만인 되도록 배치 위치를 가변하며, 상기 합성 영상에 합성된 적어도 하나의 최종 객체 마스크의 각각에 대한 검증 레이블을 생성할 수 있다.

[0024] 상기 목적을 달성하기 위한 본 발명의 다른 실시예에 따른 합성 데이터 생성 방법은 미리 획득된 다수의 배경 영상 및 다수의 객체 마스크 중 하나의 배경 영상과 적어도 하나의 객체 마스크를 선택하는 단계; 선택된 적어도 하나의 객체 마스크에 대해 블러링, 크기 조절 및 회전각 조절 중 적어도 하나를 적용하여 변형하는 단계; 및 선택된 상기 배경 영상에 변형된 적어도 하나의 객체 마스크로부터 획득되는 적어도 하나의 최종 객체 마스크를 합성하여 합성 영상을 생성하는 단계; 를 포함한다.

발명의 효과

[0025] 따라서, 본 발명의 실시예에 따른 합성 데이터 생성 장치 및 방법은 다수의 배경과 다수의 객체를 별도로 획득하여 합성함으로써, 객체 탐지 시스템이 다양한 환경에서의 객체 탐색을 수행할 수 있도록 하는 학습 데이터를 생성할 수 있다. 또한 객체에 블러링 처리, 크기 및 방향 조절과 같은 시각적 변화를 적용하여 배경과 합성함으로써, 객체 탐지 시스템이 다양한 상태의 객체를 탐지할 수 있도록 하며, 장애물을 추가로 합성함으로써 페색이 발생된 영상에서도 객체를 신뢰도 있게 탐지할 수 있도록 하는 학습 데이터를 생성하여 제공할 수 있다.

도면의 간단한 설명

[0026] 도1 은 시각적 열화가 유발된 영상의 일예를 나타낸다.

도2 는 본 발명의 일 실시예에 따른 학습 데이터 생성 장치의 개략적 구조를 나타낸다.

도3 은 도2 의 학습 데이터 생성 장치의 각 구성 요소의 동작을 설명하기 위한 도면이다.

도4 는 알파 채널 블러링의 수행 여부에 따른 합성 데이터의 차이를 비교한 도면이다.

도5 는 본 발명의 일 실시예에 따른 학습 데이터 생성 방법을 나타낸다.

도6 은 본 발명의 일 실시예에 따른 학습 데이터 생성 방법을 테스트하기 위한 예제 프레임을 나타낸다.

도7 은 본 발명의 일 실시예에 따른 학습 데이터를 이용하여 학습된 객체 탐지 시스템의 객체 탐지 성능을 실험한 결과를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0027] 본 발명과 본 발명의 동작상의 이점 및 본 발명의 실시에 의하여 달성되는 목적을 충분히 이해하기 위해서는 본 발명의 바람직한 실시예를 예시하는 첨부 도면 및 첨부 도면에 기재된 내용을 참조하여야만 한다.
- [0028] 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.
- [0029] 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.
- [0030] 도2 는 본 발명의 일 실시예에 따른 학습 데이터 생성 장치의 개략적 구조를 나타내고, 도3 은 도2 의 학습 데이터 생성 장치의 각 구성 요소의 동작을 설명하기 위한 도면이다.
- [0031] 도2 및 도3 을 참조하여, 본 실시예에 따른 학습 데이터 생성 장치를 설명하면, 학습 데이터 생성 장치는 데이터베이스부(100), 배경 선택부(200), 객체 마스크 설정부(300), 장애물 마스크 설정부(400), 객체 합성부(500) 및 합성 영상 생성부(600)를 포함한다.
- [0032] 데이터베이스부(100)는 객체 탐지 시스템 학습을 위한 합성 영상을 생성하기 위한 소스 영상인 배경 영상과 객체 영상을 저장한다. 데이터베이스부(100)는 배경 영상을 저장하는 배경 데이터베이스(110)와 객체 영상을 저장하는 객체 마스크 데이터베이스(120)를 각각 구비할 수 있다.
- [0033] 배경 데이터베이스(110)는 합성 데이터가 복잡한 실제 환경과 유사하게 생성될 수 있도록 다양한 배경 영상이 수집되어 저장될 수 있다.
- [0034] 그리고 객체 마스크 데이터베이스(120)는 탐색 대상이 될 수 있는 다양한 객체에 대한 영상이 저장된다. 여기서 객체 영상은 객체의 윤곽에 따라 칠해진 영상으로, 이후 합성 영상을 생성할 때 배경 영상에 중첩되는 일종의 마스크로 기능한다. 즉 본 실시예에서 객체 영상은 객체 마스크로서 이용된다.
- [0035] 객체 마스크 데이터베이스(120)에 저장되는 객체 영상은 다양한 방식으로 수집될 수 있으나, 일례로 별도의 보조 세그멘테이션 신경망을 통해서도 획득될 수 있다.
- [0036] 보조 세그멘테이션 신경망 또한 인공 신경망의 일종으로 학습을 요구하므로, 많은 경우에 효율적이지 않다. 그러나 여기서는 지정된 심플 배경에서 객체가 포함된 영상을 획득하고, 획득된 영상에 대해 단순 이미지 마스킹 알고리즘을 이용하여 배경을 제거함으로써 보조 세그멘테이션 신경망을 학습시키기 위한 객체 영상을 수집할 수 있으며, 학습된 보조 세그멘테이션 신경망은 다양한 객체 영상을 획득할 수 있다.
- [0037] 배경 선택부(200)는 데이터베이스부(100)에 저장된 다수의 배경 영상 중 합성하고자 하는 배경 영상을 선택하여 합성 영상 생성부(600)로 전달한다. 배경 선택부(200)는 합성 영상에 객체가 다양한 환경에 배치된 상황으로 합성될 수 있도록 랜덤하게 배경을 선택하거나, 기지정된 순서로 선택할 수 있다.
- [0038] 객체 마스크 설정부(300)는 데이터베이스부(100)에 저장된 다수의 객체 마스크 중 배경 선택부(200)에서 선택된 배경 영상에 배치되어 합성 영상으로 합성될 N(여기서 N은 자연수)개의 객체 마스크를 선택하고, 선택된 N개의 객체 마스크에 대해 모션 블러링, 크기 조절 및 회전 중 적어도 하나를 적용함으로써 시각적으로 변형시켜 객체 합성부(500)로 전달한다. 즉 객체 마스크 설정부(300)는 합성하고자 하는 N개의 객체 마스크를 선택하고 변형 시킴으로써, 선택된 객체 마스크에 다양한 상호 작용에 의해 발생할 수 있는 시각적 변형(Visual Variation) 효과를 부여한다.

- [0039] 객체 마스크 설정부(300)는 객체 마스크 선택부(310), 블러링부(320) 및 트리밍부(330)를 포함할 수 있다.
- [0040] 객체 마스크 선택부(310)는 데이터베이스부(100)에 저장된 다수의 객체 마스크 중 합성하고자 하는 N개의 객체 마스크를 선택한다.
- [0041] 블러링부(320)는 N개의 객체 마스크에 대해 모션 블러 효과를 추가한다. 블러링부(320)는 모션 블러 합성(motion blur synthesis) 방식으로 모션 블러 효과를 부과한다. 블러링부(320)는 객체 마스크의 블러링 크기(w_{mb})와 x축 방향에 대한 블러링 방향각(direction angle)(θ_{mb})을 조절하는 모션 블러 커널을 이용할 수 있다. 모션 블러 커널은 선형 특성을 가지며, 블러링 크기(w_{mb})는 모션 블러 커널의 중심에 대해 블러링 방향각(θ_{mb})을 통한 평균 픽셀이다.
- [0042] 블러링부(320)는 모션 블러 커널을 이용하여 인가된 객체 마스크에 대해 도3 에 도시된 바와 같이 블러링 효과를 추가할 수 있다. 이때 블러링부(320)는 배경 영상과의 자연스러운 융합을 위해 객체 마스크의 알파 채널에 대해서도 블러링 효과를 적용할 수 있다.
- [0043] 그리고 블러링부(320)는 객체 마스크에 대해 선택적으로 블러링 효과를 적용할 수 있다. 예를 들면, 블러링부(320)는 기지정된 블러링 확률값에 따라 인가되는 N개의 객체 마스크 중 일부 객체 마스크에만 블러링 효과를 적용하고 나머지 객체 마스크에는 블러링 효과를 적용하지 않을 수 있다. 또한 블러링 효과를 적용하는 객체 마스크에 대해서도 블러링 크기(w_{mb})와 블러링 방향각(θ_{mb})을 서로 다르게 적용할 수도 있다.
- [0044] 한편, 트리밍부(330)는 블러링부(320)로부터 객체 마스크를 인가받아 크기 조절 및 회전하여 출력한다.
- [0045] 트리밍부(330)는 다양한 상황에 배치된 객체를 연출하기 위해, 객체 마스크의 크기를 스케일링 파라미터(s_{obj})에 따라 조절하고, 객체 마스크의 x 축에 대한 회전각을 회전 파라미터(θ_{obj})에 따라 조절한다. 여기서 스케일링 파라미터(s_{obj})는 배경 영상의 크기에 대한 비율로 설정될 수 있다. 일례로 스케일링 파라미터(s_{obj})가 0.5로 설정되면, 객체 마스크는 배경 영상의 절반 크기로 조절될 수 있다. 이때 객체 마스크가 배경 영상의 크기를 초과하지 않도록, 스케일링 파라미터(s_{obj})는 배경 영상의 긴 변의 길이에 대한 상대값으로 설정될 수 있다.
- [0046] 그리고 스케일링 파라미터(s_{obj})와 회전 파라미터(θ_{obj})는 인가되는 다수의 객체 마스크 각각에 대해 서로 다른 값으로 설정될 수 있다.
- [0047] 한편, 트리밍부(330)는 크기 조절 및 회전된 객체 마스크의 크기 및 형태에 대응하여, 객체 마스크의 바운딩 박스(bounding box)를 트리밍하여 객체 합성부(500)로 출력한다.
- [0048] 장애물 마스크 설정부(400)는 도3 에 도시된 바와 같이, 객체에 대한 폐색을 유발하는 장애물(distractor) 마스크를 생성하여 폐색 합성부로 전달한다. 이때 장애물 마스크 설정부(400)는 모든 객체 마스크에 대해 장애물 마스크를 생성하여 전달하는 것이 아니라, 기지정된 장애물 확률값에 따라 선택적으로 장애물 마스크를 생성할 수 있다.
- [0049] 그리고 경우에 따라서 데이터베이스부(100)는 장애물 마스크를 저장하기 위한 장애물 마스크 데이터베이스를 더 포함할 수도 있으며, 장애물 마스크 설정부(400)는 데이터베이스부(100)에 저장된 장애물 마스크를 선택적으로 인가받을 수 있다.
- [0050] 장애물 마스크 설정부(400)는 객체 마스크 설정부(400)와 유사하게 장애물 마스크의 크기와 회전 각도를 조절하여 출력할 수 있다. 이때, 장애물 마스크의 크기와 회전 각도는 장애물 스케일링 파라미터(s_{dis})와 장애물 회전 파라미터(θ_{dis})에 의해 조절되며, 장애물 마스크가 객체 마스크를 오버레이하지 않도록 장애물 스케일링 파라미터(s_{dis})와 장애물 회전 파라미터(θ_{dis})는 스케일링 파라미터(s_{obj})와 회전 파라미터(θ_{obj})를 고려하여 결정될 수 있다.
- [0051] 뿐만 아니라 경우에 따라서는 장애물 마스크 설정부(400)는 장애물 마스크에 대해 블러링 효과를 적용할 수도 있다. 즉 장애물 마스크 설정부(400)는 객체 마스크 설정부(400)와 유사하게 구성될 수 있다.
- [0052] 객체 합성부(500)는 객체 마스크 설정부(400)에서 인가되는 변형된 객체 마스크와 장애물 마스크 설정부(400)에서 인가되는 장애물 마스크를 합성하여 최종 객체 마스크를 생성한다. 여기서 장애물 마스크는 객체 마스크 상에 오버랩되되, 상기한 바와 같이 장애물 마스크에 의해 객체 마스크가 완전히 가려지지 않도록 합성된다. 도3

에서는 일예로 객체 마스크인 가방과 장애물 마스크인 사람의 팔이 합성된 최종 객체 마스크를 도시하였다.

- [0053] 여기서는 설명의 편의를 위해 장애물 마스크 설정부(400)와 객체 합성부(500)를 구분하여 도시하였으나, 장애물 마스크 설정부(400)는 객체 합성부(500)에 포함될 수 있다.
- [0054] 합성 영상 생성부(600)는 배경 선택부(200)에서 전달되는 배경 영상 위에 객체 합성부(500)에서 전달되는 N개의 최종 객체 마스크를 오버랩하여 합성 영상을 생성한다. 즉 도3에 도시된 바와 같이 합성 영상 생성부(600)는 하나의 배경 영상에 다수의 최종 객체 마스크를 배치할 수 있다.
- [0055] 이때 합성 영상 생성부(600)는 이전 배경 영상 위에 배치된 최종 객체 마스크와 중첩되어 배치되지 않도록 조절할 수 있다. 이를 위해, 합성 영상 생성부(600)는 N개의 최종 객체 마스크를 순차적으로 랜덤하게 배치한다. 그리고 이전 배치된 최종 객체 마스크와 현재 배치되는 최종 객체 마스크 사이의 IOU(intersection of union)를 측정하고, 측정된 IOU가 기지정된 허용값(ε_{iou}) 이상이면 현재 배치되는 최종 객체 마스크의 위치를 재설정할 수 있다. 즉 IOU가 허용값(ε_{iou}) 미만이 될 때까지 반복적으로 최종 객체 마스크의 위치를 가변할 수 있다.
- [0056] 합성 영상 생성부(600)는 다양한 환경에서 다양한 객체가 배치된 상황에 대한 합성 영상을 생성할 수 있으며, 생성된 합성 영상은 객체 탐지 시스템을 학습 시키기 위한 학습 데이터로 이용될 수 있다. 특히 합성 영상 생성부(600)는 배경 영상에 최종 객체 마스크를 직접 위치 시킴에 따라, 합성 영상 생성 시에 각 객체에 대한 검증(ground truth) 레이블을 자동으로 생성한다.
- [0057] 결과적으로 본 발명의 실시예에 따른 합성 데이터 생성 장치는 적은 수의 배경 영상과 객체 영상을 이용하여 여러 환경 및 조건과 다양한 객체에 대한 대량의 영상을 획득할 수 있도록 한다. 따라서 객체 탐지 시스템을 학습 시키기 위해 대량의 학습 데이터를 용이하게 획득할 수 있도록 한다.
- [0058] 도4는 알파 채널 블러링의 수행 여부에 따른 합성 데이터의 차이를 비교한 도면이다.
- [0059] 도4에서 (a)는 객체 마스크 설정부(300)의 블러링부(320)가 객체 마스크의 알파 채널에 모션 블러링을 적용하여 합성 영상을 생성한 경우를 나타내고, (b)는 알파 채널에 모션 블러링을 적용하지 않고 합성 영상을 생성한 경우를 나타낸다.
- [0060] 알파 채널은 그래픽상의 한 픽셀의 색이 다른 픽셀의 색과 겹쳐서 나타날 때 두 색을 효과적으로 융합하기 위한 채널로서, 만일 객체 마스크에서 알파 채널을 제외하고 모션 블러링을 적용하는 경우, 즉 알파 채널에 대해 모션 블러링을 적용하지 않는 경우에는 (b)의 확대 영상에 도시된 바와 같이, 합성 영상에 아티팩트(artifact)가 발생하는 문제가 있다. 따라서 신뢰성 있는 학습 데이터를 생성하기 위해서는 객체 마스크에 대해 모션 블러링 효과를 적용할 때, 알파 채널도 함께 적용되는 것이 바람직하다.
- [0061] 도5는 본 발명의 일 실시예에 따른 학습 데이터 생성 방법을 나타낸다.
- [0062] 도2 및 도3을 참조하여, 도5의 학습 데이터 생성 방법을 설명하면, 우선 배경 선택부(200)가 데이터베이스부(100)에 저장된 다수의 배경 영상 중 하나의 배경 영상을 선택하고, 객체 마스크 설정부(300)가 다수의 객체 마스크 중 N개의 객체 마스크를 선택한다(S10).
- [0063] 그리고 객체 마스크 설정부(300)는 선택된 N개의 객체 마스크 각각에 대해 모션 블러링을 수행할 지 여부를 기지정된 블러링 확률값에 따라 결정한다(S20). 만일 모션 블러링을 수행할 것으로 판별되면, 객체 마스크 설정부(300)는 블러링 크기(w_{mb})와 블러링 방향각(θ_{mb})에 따라 객체 마스크에 모션 블러링하여 모션 블러 효과를 추가한다(S30). 이때 블러링 크기(w_{mb})와 블러링 방향각(θ_{mb})은 각각의 객체 마스크에 대해 서로 상이하게 조절될 수 있다.
- [0064] 또한 객체 마스크 설정부(300)는 N개의 객체 마스크 각각의 크기를 스케일링 파라미터(s_{obj})에 따라 조절하고, 객체 마스크의 x 축에 대한 회전각을 회전 파라미터(θ_{obj})에 따라 조절한다(S40). 이때 객체 마스크 설정부(300)는 배경 영상의 크기를 초과하지 않도록, 객체 마스크의 크기를 조절할 수 있다. 또한 객체 마스크 설정부(300)는 크기 조절 및 회전된 객체 마스크의 크기 및 형태에 대응하여, 객체 마스크의 바운딩 박스를 트리밍할 수 있다.
- [0065] 그리고 장애물 마스크 설정부(300)는 기지정된 장애물 확률값에 따라 객체 마스크에 장애물 마스크를 추가할지 여부를 결정한다(S50). 장애물 마스크 설정부(300)는 장애물 마스크를 추가할 것으로 결정되면, 추가할 장애물을 선택하고, 객체 마스크 설정부(300)는 장애물 마스크 설정부(300)에서 선택된 장애물 마스크를 객체 마스크

설정부(300)에서 변형되어 객체 마스크에 합성하여 최종 객체 마스크를 생성한다(S60).

[0066] 이때 장애물 마스크 설정부(300)는 객체 마스크 설정부(400)와 유사하게 장애물 마스크의 크기와 회전 각도를 조절할 수 있으며, 경우에 따라서는 블러링 효과를 적용할 수도 있다.

[0067] 합성 영상 생성부(600)는 배경 선택부(200)에서 선택된 배경 영상에 객체 합성부(500)에서 전달되는 N개의 최종 객체 마스크를 오버랩하여 합성 영상을 생성한다(S70). 합성 영상 생성부(600)는 IOU가 기지정된 허용값(ϵ_{iou}) 이상인지 여부를 판별하여, 이전 배경 영상 위에 배치된 최종 객체 마스크와 중첩되어 배치되지 않도록 조절할 수 있다. 이때 합성 영상 생성부(600)는 합성 영상에서 각 객체에 대한 검증(ground truth) 레이블을 자동으로 생성한다.

[0068] 이하에서는 본 실시예의 학습 데이터 생성 장치 및 방법에 따라 생성되는 학습 데이터의 학습 성능을 테스트한 결과를 나타낸다.

[0069] 여기서는 일례로 표1에 나타난 바와 같이, 12개의 객체 마스크 카테고리를 선택하였으며, 각 카테고리에 대한 객체 마스크의 개수는 상이할 수 있다.

표 1

Object Category	Rifle	Bat	Bag	Bottle	Knife	Handgun	Laptop	Umbrella	Book	Phone	Broom	Chair	Total
# of Object Patches	17	24	15	18	9	13	39	17	39	18	8	9	226

[0070]

[0071] 학습 데이터 생성 시에 블러링 크기(w_{mb})는 {20, 40} 중에서 선택되도록 설정되었으며, 블러링 방향각(θ_{mb})은 {-45, 0, 45, 90} 중에서 선택되도록 설정되었다. 그리고 스케일링 파라미터(s_{obj})와 장애물 스케일링 파라미터(s_{dis})는 {0.2, 0.3, 0.4}에서 선택되고, 회전 파라미터(θ_{obj})와 장애물 회전 파라미터(θ_{dis})는 {-45, 0, 45, 90}에서 선택되도록 설정되었다. 또한 IOU에 대한 허용값(ϵ_{iou})은 0.1로 설정되었다.

[0072] 표2는 본 발명의 실시예에 따른 합성 영상을 학습 데이터로 이용하여 객체 탐지 시스템을 학습한 경우의 객체 탐지 성능을 나타내며, 평균 정밀도(average precision: AP) 메트릭으로 객체 탐지 시스템의 성능을 평가하였다.

표 2

Visual Variations	Rifle	Bat	Bag	Bottle	Knife	Handgun	Laptop	Umbrella	Book	Phone	Broom	Chair	mAP
<i>SynDB₁</i> w/ none	48.4	33.6	31.5	64.3	27.5	32.5	44.2	41.5	52.5	29.6	16.7	15.6	36.5
<i>SynDB₁</i> w/ occl	68.5	76.8	21.0	74.4	31.0	31.8	58.0	64.4	63.9	57.4	31.3	24.9	50.3
<i>SynDB₁</i> w/ mb	64.5	85.6	55.3	90.8	25.6	64.4	47.4	59.2	73.0	77.3	41.0	14.8	58.3
<i>SynDB₁</i> w/ mb and occl	89.7	78.8	67.7	87.0	46.8	69.7	70.6	82.7	87.3	81.9	77.2	87.7	77.3

[0073]

[0074] 표2에서는 합성 데이터를 이용한 객체 탐지 시스템의 학습 성능을 분석할 때, 시각적 열화에 대한 성능을 비교할 수 있도록 학습 데이터 집합을 4가지로 분류하여 테스트한 결과를 나타낸다.

[0075] 이하에서 $SynDB_n$ 은 n(n은 자연수)개의 영상에 포함된 객체의 개수를 의미한다. 즉 표2의 $SynDB_1$ 은 영상에 1개의 객체가 포함되어 있음을 의미한다.

[0076] 그리고 w/none는 모션 블러와 페색이 포함되지 않은 학습 데이터 집합을 의미하고, w/occl은 페색된 객체가 포함된 학습 데이터 집합, w/mb는 모션 블러가 용된 객체 포함된 학습 데이터 집합을 의미하며, w/mb and occl은 모션 블러와 페색이 모두 포함된 학습 데이터 집합을 의미한다.

[0077] 표2를 살펴보면, 본 실시예에 따라 생성된 학습 데이터를 이용하여 학습된 객체 탐지 시스템은 시각적 열화가 발생된 영상에 대해 매우 우수한 성능을 나타냄을 알 수 있다.

- [0078] 도6 은 본 발명의 일실시예에 따라 생성된 학습 데이터를 테스트하기 위한 예제 영상을 나타낸다.
- [0079] 도6 에서는 (a)와 (b)는 합성 데이터를 이용한 객체 탐지 시스템의 학습 성능을 비교하여 분석하기 위해 기존 방식으로 획득된 학습 데이터의 일례를 나타내며, 미리 수작업을 통해 검증 레이블이 설정되어 있는 것으로 가정한다. 그리고 (c)는 본 실시예에 따라 생성된 학습 데이터의 일례를 나타내며, (c)는 배경 영상과 객체 마스크가 합성된 합성 영상이므로, 검증 레이블 자동으로 생성된다.
- [0080] (a)는 표1 에 나타난 영상에 객체 카테고리의 객체들 중 하나의 객체가 포함된 영상($N = 1$)으로 상단 영상에는 총이 포함되어 있으며, 하단 영상에는 빗자루가 포함되어 있다. 그리고 (b)는 영상에 객체 카테고리의 객체들 중 다수개의 객체가 포함된 영상으로 상단 영상에는 가방과 배트 및 의자 포함되어 있으며, 하단 영상에는 병, 나이프, 총 및 폰이 포함되어 있다.
- [0081] 한편 (c)는 배경 영상과 객체 마스크가 합성된 합성 영상으로 다수의 객체가 포함되어 있다.
- [0082] 표3 은 도6 과 같이 기존의 방식으로 획득된 학습 데이터와 본 실시예에 따라 생성된 학습데이터를 이용하여 객체 탐지 시스템을 학습시킨 결과를 비교하였다.
- [0083] 표3 에서는 영상에 포함된 객체의 개수를 1개에서 8개까지 가변($n = 1 \sim 8$)하면서 테스트를 수행한 결과이며, 좌측은 검증 레이블이 수작업으로 획득된 실제 영상을 의미하며, 우측은 본 실시예에 따라 합성된 합성 영상을 의미하며, 모든 영상에 객체의 시각적 열화가 포함된 경우를 테스트하였다.

표 3

Dataset Type	Multiple Object Interaction	Synthetic Multiple Object
<i>SynDB₁</i>	54.3	63.0
<i>SynDB₂</i>	56.4	80.4
<i>SynDB₃</i>	59.6	84.7
<i>SynDB₄</i>	57.4	83.3
<i>SynDB₅</i>	57.7	85.3
<i>SynDB₆</i>	60.4	84.3
<i>SynDB₇</i>	55.9	85.0
<i>SynDB₈</i>	60.9	85.0

- [0084]
- [0085] 표3 에 나타난 바와 같이 본 실시예에 따른 학습 데이터 생성 방법에 의해 생성된 합성 영상을 이용하여 학습된 객체 탐지 시스템은 객체의 개수에 상관없이 더 나은 객체 탐지 성능을 나타냄을 확인할 수 있다.
- [0086] 도7 은 본 발명의 일실시예에 따른 학습 데이터를 이용하여 학습된 객체 탐지 시스템의 객체 탐지 성능을 실험한 결과를 나타낸다.
- [0087] 도7 에서 (a)는 폐색된 객체에 대한 검출 결과를 나타내고, (b)는 모션 블러된 객체에 대한 검출 결과를 나타낸다. 그리고 도7 에서 상단은 모션 블러와 폐색이 포함되지 않은 합성 영상을 이용하여 학습시킨 결과를 나타내고, 가운데단은 폐색이 포함된 합성 영상을 이용하여 학습시킨 결과를 나타내며, 하단은 모션 블러 및 폐색이 모두 포함된 합성 영상을 이용하여 학습시킨 결과를 나타낸다.
- [0088] 도7 을 살펴보면, 상단의 모션 블러와 폐색이 포함되지 않은 합성 영상을 이용하여 학습된 객체 검출 시스템은 (a)에서 폐색이 나타나지 않는 영역만을 검출하는데 반해, 모션 블러가 존재하는 (b)에서는 객체를 전혀 검출하지 못하였다.
- [0089] 그리고 가운데 단의 폐색이 포함된 합성 영상을 이용하여 학습된 객체 검출 시스템은 (a)에서 폐색이 발생하였음에도 객체의 대부분의 영역을 검출하는데 반해, 여전히 모션 블러가 존재하는 (b)에서는 객체를 전혀 검출하지 못하였다.
- [0090] 마지막으로 하단의 모션 블러 및 폐색이 모두 포함된 합성 영상을 이용하여 학습된 객체 검출 시스템은 폐색이

발생된 (a)와 모션 블러가 존재하는 (b) 모두에서 양호하게 객체를 검출할 수 있음을 알 수 있다.

[0091] 결과적으로 본 발명의 실시예에 따른 합성 데이터 생성 장치 및 방법에 의해 생성된 합성 영상은 배경에 다양한 객체를 삽입하여 합성 영상을 생성함으로써, 객체 검출 시스템이 다양한 환경에서 객체를 용이하게 검출할 수 있도록 한다. 뿐만 아니라 합성 영상에 포함되는 각 객체에 모션 블러 및 페색과 같은 시각적 열화 현상이 반영되도록 함으로써, 객체 검출 시스템이 시각적 열화 현상에도 불구하고 객체를 검출할 수 있도록 학습 시킬 수 있다. 따라서 객체 검출 시스템의 성능을 크게 향상 시킬 수 있다.

[0092] 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.

[0093] 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.

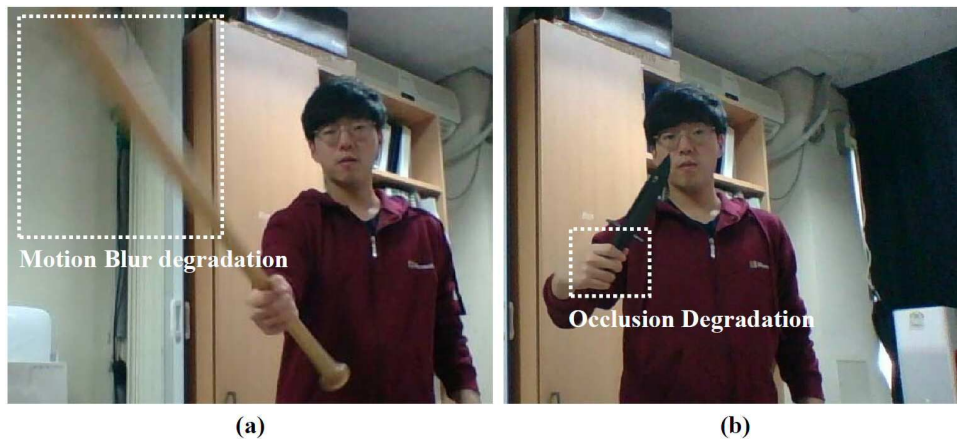
[0094] 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 청구범위의 기술적 사상에 의해 정해져야 할 것이다.

부호의 설명

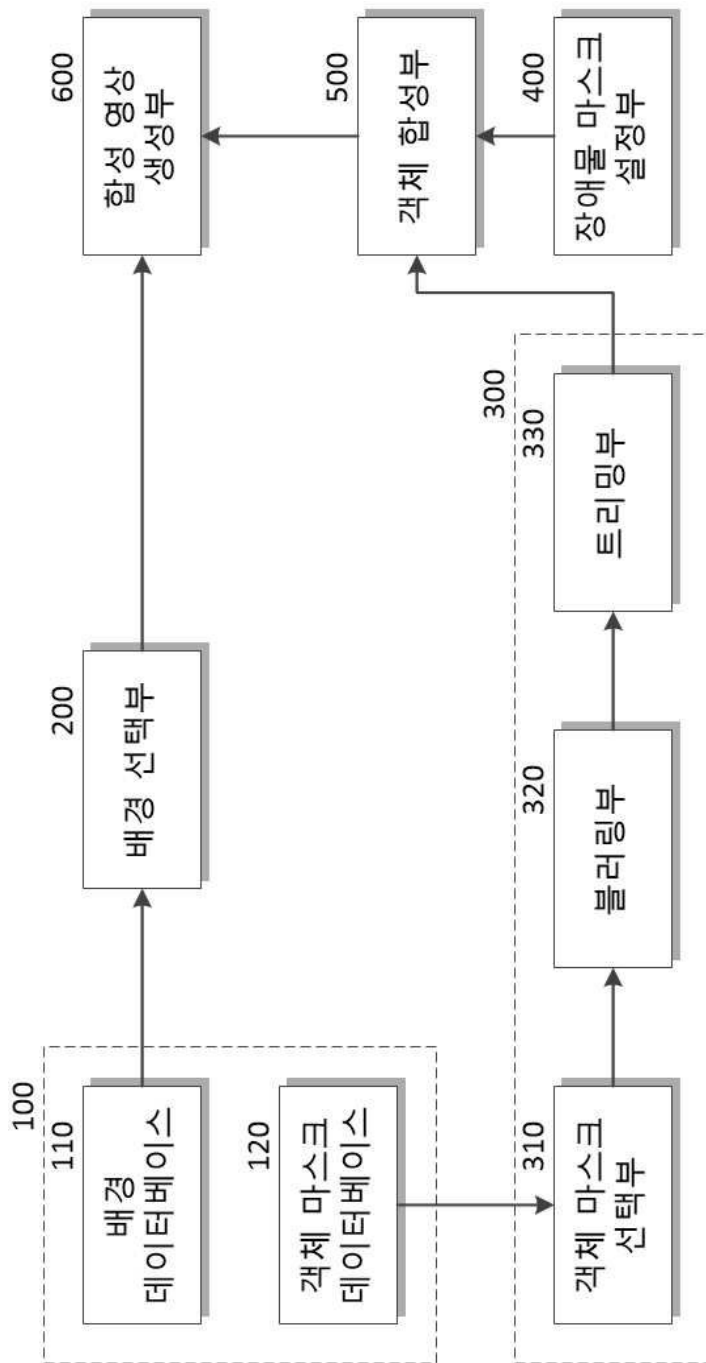
[0095] 100: 데이터베이스부 200: 배경 선택부
300: 객체 마스크 설정부 400: 장애물 마스크 설정부
500: 객체 합성부 600: 합성 영상 생성부

도면

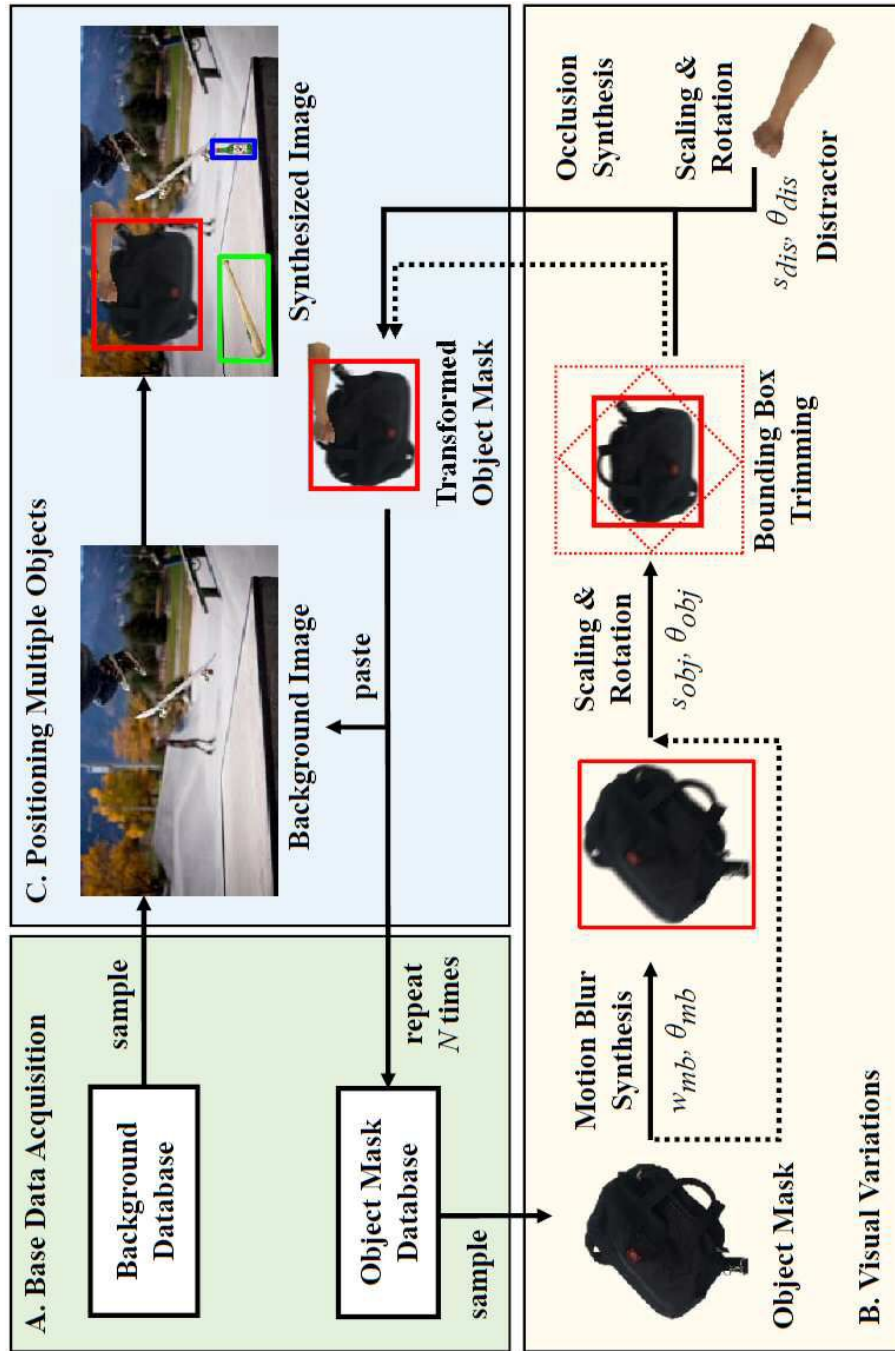
도면1



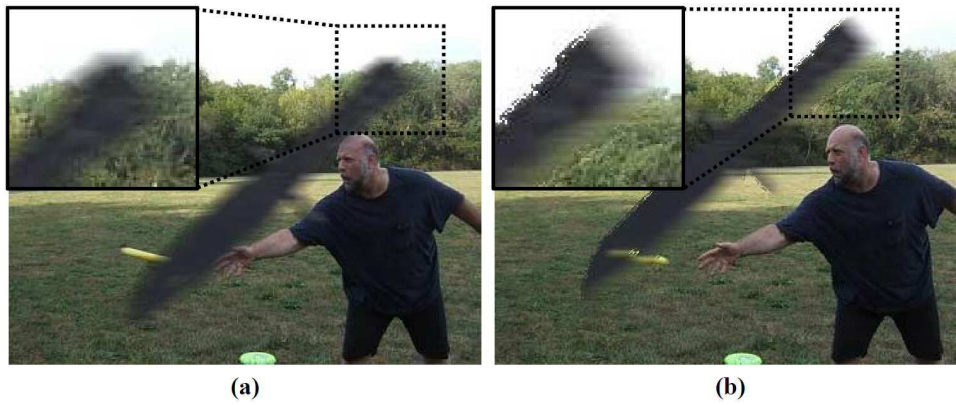
도면2



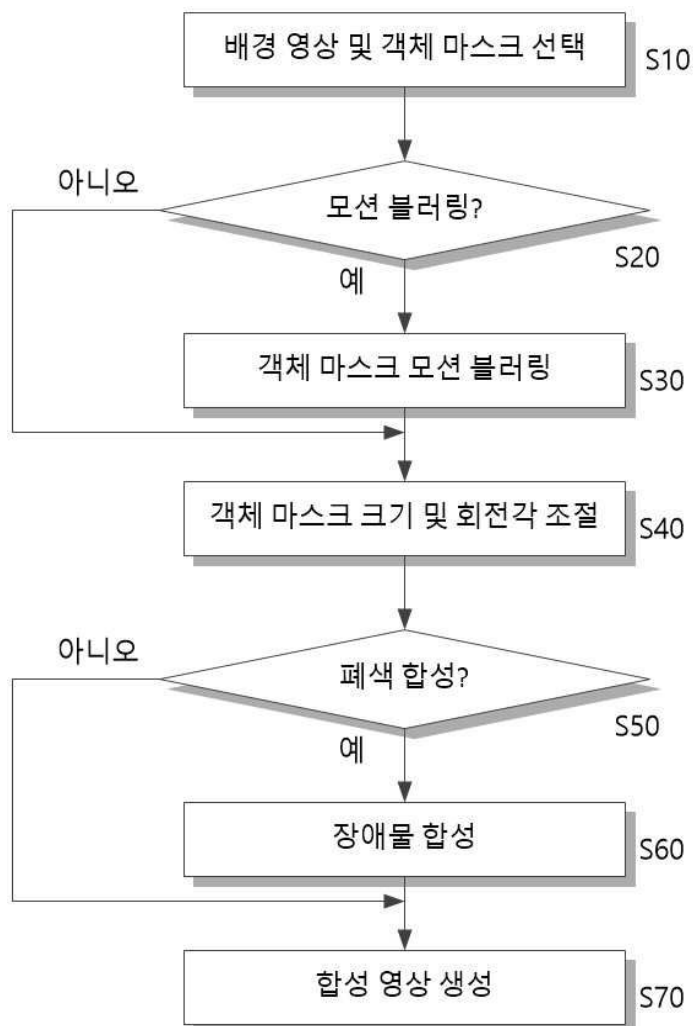
도면3



도면4



도면5



도면6



도면7

