

이미지 인페인팅을 위한 효과적인 제거 영역 선택 기법

윤상빈, 정순기
 경북대학교 컴퓨터학부
 kor135@knu.ac.kr, skjung@knu.ac.kr

Efficient Removal Region Selection for Image Inpainting

Sang Bin Yun, Soon Ki Jung

School of Computer Science and Engineering, Kyungpook National University

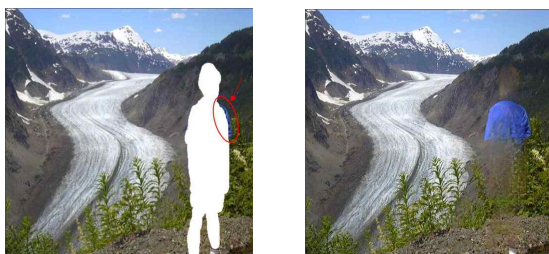
요 약

이미지 인페인팅은 이미지에서 특정 영역을 제거하고 매끄럽게 채우는 작업을 의미한다. 이미지 인페인팅에서 물체 제거 대상 영역을 설정하는 방법은 사람이 수작업을 하거나 사각형 박스로 설정한다. 수작업의 경우 많은 시간과 비용이 소모되고 사각형 박스의 경우 물체만 제거하기 위해 필요한 영역 이외에 사각형 박스 안의 나머지 부분도 새로 생성해내기 때문에 불필요한 정보의 손실이 발생한다. 따라서 본 논문에서는 물체를 제거하는 영상 작업에서 의미론적 영상분할 기술을 이용하여 물체 제거에 불필요한 영역 설정을 최소화하여 수작업의 방법에 비해서는 시간과 비용적 문제를 최소화하였고 사각형 박스보다는 자연스러운 이미지 생성하는 결과를 보였다.

1. 서 론

본 논문은 이미지 인페인팅[1] 분야 중 이미지에서 물체를 제거하고 물체를 지운 영역을 주변 배경과 자연스럽게 채우는 것에 중점을 두어, 지우고자 하는 대상 물체의 설정 과정에서 수작업을 최소화함과 동시에 불필요한 정보손실을 최소화함을 목표로 한다. 현재 대부분의 이미지 인페인팅 모델들은 목표 물체를 제거하기 위한 영역 설정을 위해 사각형 박스(bounding box) 또는 자유롭게 채색(freeform)하는 것으로 이루어져 있다. 사각형 박스의 경우 관심 물체 이외의 영역을 포함하여 제거하고 새로운 이미지를 생성하기 때문에 필요 이상의 정보가 손실되는 문제가 있다. 자유롭게 채색하는 경우 수작업으로 진행해야 하기 때문에 시간, 비용적 요소가 소모된다. 또한 지우고자 하는 물체를 모두 포함하지 않은 경우 해당 영역이 영향을 주게 되어 아래 (그림 1)과 같이 자연스러운 배경 생성을 하는 과정에 좋지 못한 영향을 주게 된다.

따라서 본 논문에서는 이러한 문제를 해결하기 위해 의미론적 영상분할(semantic segmentation)을 통해 지우고자 하는 물체의 영역을 정확하게 구별하여 물체 이외의 정보 손실을 최소화하여 영상 영역을 제거하고 자연스럽게 배경 영상을 생성한다.



관심 영역

결과 이미지

(그림 1)

2. 관련 연구

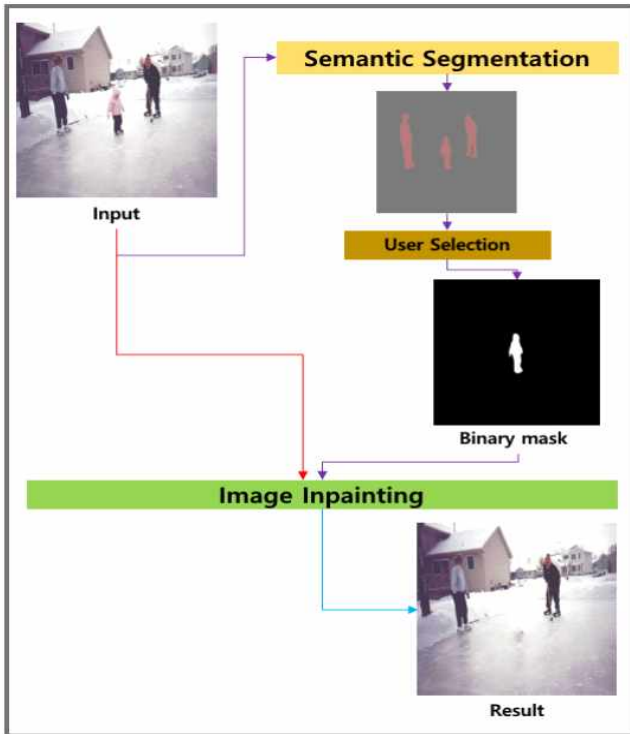
2.1 이미지 인페인팅(Image Inpainting)

이미지 인페인팅은 이미지에서 지워진 영역을 채우는 작업을 의미한다. 딥러닝 이전의 이미지 인페인팅 방식은 확산 기반(diffusion based) 방법[1]과 패치 기반(patch based) 방법[2]으로 해당 영역을 채우는 것으로 해당 문제를 해결해왔다. 확산 기반 방법[1]은 경계 주변의 픽셀 값의 미분방정식을 이용해 채우기 때문에 채워야 할 영역이 클 경우 블러링(blurring)되는 문제가 있고 패치 기반 방법[2]은 이미지 내의 패치만을 이용해서 채우기 때문에 새로운 이미지를 생성해 내야 하는 작업에서는 부자연스러운 느낌이 드는 이미지를 생성해낸다. 딥러닝 이후의 이미지 인페인팅 방식[3]은 특히 적대적 생성 모델[4] 이후 이미지 인페인팅의 성능이 크게 향상되었다. 기존의 방식은 현재 가지고 있는 영상의 정보만을 이용해서 빈 공간을 채우는 방식이기 때문에 새로운 이미지를 만들어내는 것에 한계가 있었지만 적대적 생성 모델은 딥페이크, 워터마크 제거, 물체 제거 등 다양한 응용 분야에 이용되고 있다.

2.2 의미론적 영상분할(Semantic Segmentation)

의미론적 영상분할은 이미지 내에 있는 객체들을 의미 있는 단위로 분할하는 작업을 말한다. 즉 영상을 여러 개의 픽셀 집합으로 나누는 과정이며 단순히 이미지를 분류(classification)하고 객체를 탐지(detection)하는 것을 넘어서 객체의 경계 영역까지 정확하게 분류하는 문제이다. 의미론적 영상 분할 모델은 FCN[5] 이후로 딥-러닝 기반의 모델로 확장되어 다양한 응용분야에 이용되고 있다.

3. 의미론적 영상분할을 이용한 물체 제거



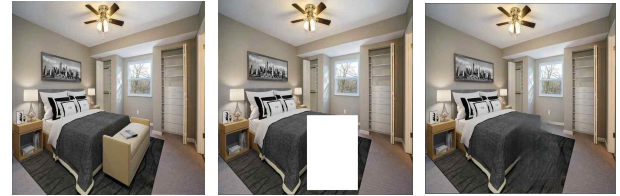
(그림 2) 제안된 물체제거 구조

위 그림 2의 구조에서 입력 이미지는 의미론적 영상분할 모델을 통과하여 물체 영역이 분할된다. 분할된 이미지에서 목표로 하는 물체를 유저 선택 모듈을 통해 원하는 객체를 이진 마스크 영상으로 만든 뒤 입력 이미지와 함께 이미지 인페인팅 모델의 입력으로 들어가 최종 이미지가 출력된다.

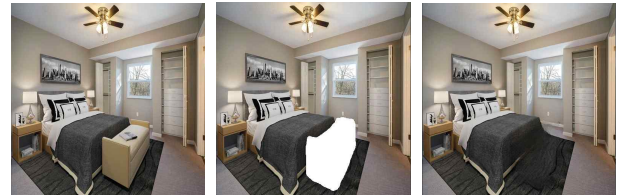
4. 실험

본 실험에서는 기존의 물체 제거 작업 중 사각형 박스를 이용한 영역 설정 방법과 제안된 모델을 비교하였다. 제안된 모델에서 의미론적 영상분할은 ADE20k 데이터셋을 기반으로 학습한 HRNetV2[6] 모델을 이용하였고 인페인팅 모델은 Places2 데이터셋으로부터 학습된 AOT-GAN[7]를 이용하여 실험을 진행하였다.

본 실험의 결과는 아래 (그림 3)에 나타나있다. (그림 3a)에서 사각형 박스의 경우 지우고자 하는 물체 이외의 정보가 함께 삭제되는 문제가 발생한다. 특히 물체의 모양이 사각형과 차이가 나거나 대상 물체와 다른 물체에 겹쳐있을 경우 (그림 3a)의 침대 모서리 부분처럼 불필요한 정보의 삭제로 인한 부자연스러움이 발생할 수 있다. 이 경우 제안된 모델을 활용할 경우 (그림 3b)의 결과와 같이 불필요한 정보가 삭제되는 것을 방지할 수 있고 해당 물체만을 삭제하여 주변 환경에 맞게 재생성이 가능하다.



(a) 사각형 박스



(b) 제안된 모델

(그림 3) 영역설정 에 따른 인페인팅 결과 이미지

5. 결론 및 향후 연구

본 논문에서는 이미지 생성 작업 중 물체 제거에서 기존의 물체 영역을 선택하는 방법의 문제를 해결하기 위해 의미론적 분할 모델을 이용하여 물체 영역을 선택할 수 있는 모델을 제안하였다. 제안된 모델을 활용하면 인적, 시간적 요소를 최소화하고 또한 불필요한 영역이 선택되는 문제를 최소화할 수 있다. 또한 제안된 모델은 비디오에서의 물체를 제거하는 응용 기술로 사용될 수 있다. 향후 물체의 그림자 영역을 포함한 의미론적 영상분할 모델로 확장하여 다양한 물체에 적용이 가능하도록 연구가 진행될 예정이다.

Acknowledgement

본 논문(저서)은 교육부 및 한국연구재단의 4단계 BK21 사업(경북대학교 컴퓨터학부 지능융합 소프트웨어 교육연구단)으로 지원된 연구임 (4199990214394)

참고 문헌

- [1]. M. Bertalmio, et al., "Image inpainting," in SIGGRAPH, 2000.
- [2]. A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," in TIP, 2004.
- [3]. Ian J. Goodfellow, et al., "Generative Adversarial Networks", in NIPS, 2014.
- [4]. A. Radford, et al. "Unsupervised representation learning with deep convolutional generative adversarial networks", in ICLR, 2016.
- [5]. J. Long, et al., "Fully convolutional networks for semantic segmentation", in CVPR, 2015.
- [6]. Ke Sun, et al., "High-Resolution Representations for Labeling Pixels and Regions", in arXiv:1511.06434, 2019.
- [7]. Yanhong Zeng, et al., "Aggregated Contextual Transformations for High-Resolution Image Inpainting", in arxiv.org/abs/2104.01431, 2021.