

잡음 제거 여부에 따른 비디오 인페인팅의 성능 변화에 대한 사례 연구

서정운, 백한결, 박상효*

경북대학교

s.park@knu.ac.kr*

A Case Study on the Video Inpainting Performance by Denoising Technique

Jeong-yun Seo, Han-gyul Baek, Sang-hyo Park*

Kyungpook National University

요약

본 논문에서는 영상의 누락된 영역을 채우는 기술인 비디오 인페인팅에서 노이즈가 미치는 영향을 연구하였다. 현실 세계의 영상에는 노이즈가 포함되어 있으므로 인페인팅시 영상의 품질에 영향을 줄 수 있다. 따라서 본 논문에서는 노이즈가 포함된 영상을 디노이징 후 인페인팅을 적용한 결과와 인페인팅을 적용한 후 디노이징한 영상을 비교하였으며, 이에 따라 노이즈가 인페인팅 기술을 적용하는 비디오의 품질에 끼치는 영향에 대해 관찰하고 비디오의 품질을 개선할 수 있는 방법을 제시한다.

I. 서론

현재 컴퓨터 비전 분야는 빠른 속도로 성장하고 있다. 관련 분야 중 인페인팅(inpainting) 분야는 크게 이미지 인페인팅(image inpainting)과 비디오 인페인팅(video inpainting)의 두 개 분야로 나누어진다. 현재까지도 완벽한 인페인팅 알고리즘을 구현하기에는 어려움이 존재하여 지속적으로 많은 연구가 이루어지고 있는 추세이다. 더 나은 이미지 인페인팅을 위해 초기 구멍 문제(initial hole problem)로 인해 모델이 부자연스러운 이미지를 생성하는 문제에 대해 부분적 합성곱 층(partial convolutional layer)의 계산을 통하여 손상된 영역으로부터 좋은 결과를 얻는 연구가 발표된 바 있다[1]. 하지만 해당 기법의 한계점이 존재하는데, 부분적 합성곱은 layer를 지나는 순간마다 feature가 점점 줄어들게 되고 결국 모든 마스크 값들이 1로 변환된다. 또한 사용자가 입력한 마스크(user-guided mask)를 사용하지 못하여 사용자가 원하는 형태로 이미지를 생성할 수 없다는 문제를 참고문헌 [2]에서 밝히며, 해당 문제를 동적 특성 선택 매커니즘(dynamic feature selection mechanism)이라는 방법을 통하여 새로운 합성곱인 gated convolution으로 문제를 해결하였다[2].

비디오 인페인팅은 이미지 인페인팅과 마찬가지로 누락된 영역을 채우는 기술로, 특정 객체를 제거하고 배경을 완성하거나 손상된 영역을 복원하는데 사용한다는 점은 유사하다. 하지만 비디오 인페인팅은 비디오의 기본 단위인 프레임(frame)이라하는 연속된 이미지에서 프레임 간의 시공간적 일관성 유지가 필요하다는 차이점이 있다[3]. 시공간적 일관성을 보존하는 문제를 해결하는 연구는 활발히 진행되고 있지만, 현실 세계의 노이즈가 포함된 영상에 인페인팅을 적용할 때 나타나는 문제점은 거의 다루어지지 않았다. 그러나 현실 세계의 영상, 특히 어두운 곳에서 촬영된 영상에는 노이즈가 불가피하게 포함되어있고 이것은 인페인팅을 적용한 영상의 품질에 영향을 줄 가능성이 있다. 따라서 본 논문에서는 비디오 인페인팅 기술을 사용할 때 노이즈가 미치는 영향에 관해 연구하여 비디오의 품질을 개선할 수 있는 방법을 제안한다.

II. 본론

1. 배경 기술

본 논문에서는 비디오 인페인팅 실험에 Deep Flow-Guided Video Inpainting 모델[4]을 사용했다. 이 모델은 각 프레임의 누락된 영역에 RGB 픽셀을 채우는 대신 픽셀의 전파를 통해 프레임을 완성한다. 그림 1은 모델의 구조로, 3개의 subnetwork(DFC-S)가 스택 구조로 Deep Flow Completion Network(DFCNet)을 구성하고 있다. 이 DFCNet 모델이 각 DFC-S에서 추출한 flow를 기반으로 영상의 flow-field를 완성하고, 완성된 flow-field의 양방향 전파를 통해 누락된 영역을 채운다. 원본 모델의 Flow 추출에는 FlowNet2.0 모델이 사용되었지만, 본 실험에서는 전체 모델의 경량화를 위해 LiteFlowNet2.0 모델을 사용했다.

데이터 디노이징(denoising)단계에는 Fast Deep Video Denoising Network(FastDVDNet) 모델[5]을 사용했다. FastDVDNet 모델은 컨볼루션 신경망 네트워크를 기반으로 프레임간의 flow 추출 없이 빠르게 노이즈를 제거할 수 있는 알고리즘이다. 연속된 5장의 프레임과 노이즈맵을 입력으로 사용하며 두 단계를 거쳐 노이즈를 제거한다.

2. 실험 과정

그림 3의 두 실험은 Ubuntu, python=3.7, pytorch=1.4.0, CUDA=11.3으로 동일한 환경에서 진행했다. 데이터셋은 DAVIS-dataset의 break-dance, rollerblade 두 개의 영상에 대해 각각 70장의 프레임에 사용한 결과를 나타냈다.

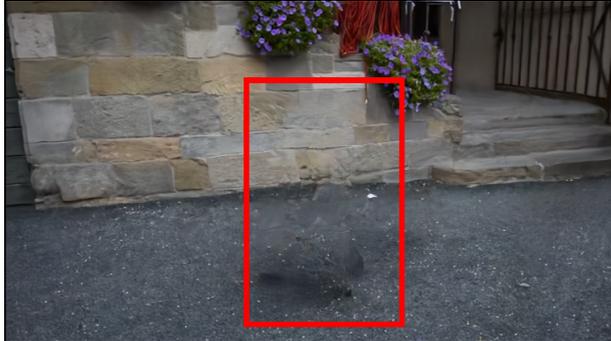
두 실험 모두 원본 비디오에 동일한 강도의 가우시안 노이즈를 임의로 추가하여 실험을 진행했다. 첫 번째 실험은 노이즈를 가진 각각의 프레임에 대해 먼저 FastDVDNet모델을 적용해 노이즈를 제거하고 결과 프레임의 객체를 인페인팅했다. 두 번째 실험은 노이즈를 가진 프레임의 객체를 인페인팅 모델로 먼저 제거한 후의 영상을 디노이징했다. 그림 1은 원본 프레임과 두 실험의 결과를 비교한 그림이다.



원본 break-dance 이미지



원본 rollerblade 이미지



실험1. break-dance(노이즈 제거 후 인페인팅 적용)



실험1. rollerblade(노이즈 제거 후 인페인팅 적용)



실험2. break-dance(인페인팅 적용 후 노이즈 제거)



실험2. rollerblade(인페인팅 적용 후 노이즈 제거)

그림 1. 실험 결과

III. 결론

앞서 진행한 두 실험을 비교해봤을 때, 노이즈가 포함된 영상을 인페인팅 후 노이즈를 제거한 영상은 노이즈가 완벽하게 제거되지 않았다. 반면에 노이즈를 제거한 후 인페인팅을 적용한 영상은 시각적으로 자연스러운 것을 확인했다. 실험 2의 경우 노이즈가 포함된 프레임의 픽셀이 전파되어 빈 영역을 채웠기 때문에 노이즈가 완벽히 제거되지 않은 것으로 추측된다. 이 실험 결과를 통해 디노이징 여부가 인페인팅 비디오의 품질에 영향을 미칠 수 있으며, 영상에 노이즈가 존재하는 경우 노이즈를 먼저 제거한 후 인페인팅을 진행하는 것이 비디오의 품질을 개선할 가능성이 있음을 알 수 있다. 본 실험을 실제 야간에 촬영된 영상이나 의도치 않은 노이즈가 발생한 영상에 적용한다면 실제 인페인팅 기술의 품질을 높이는 데 도움이 될 것으로 기대된다. 향후 본 실험을 확장시켜 여러 데이터에 적용해 보고, 인페인팅 모델을 다양하게 사용하여 추가 실험을 진행해볼 예정이다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2020R111A3072227).

참 고 문 헌

- [1] Liu, G., Reda, F. A., Shih, K. J., Wang, T. C., Tao, A., & Catanzaro, B. (2018). Image inpainting for irregular holes using partial convolutions. In Proceedings of the European conference on computer vision (ECCV) (pp. 85-100).
- [2] Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T. S. (2019). Free-form image inpainting with gated convolution. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 4471-4480).
- [3] 구윤희, 이한솔 이창화, 정민우, 하 선, 조운성, 정해성, 박한수, 김은서, 백승렬, 이미지 인페인팅을 위한 윤곽 유지 및 내부 문맥 개선법, 2022 제 34회 영상처리 및 이해에 관한 워크샵 (IPIU2022) (2022.2.9~11)
- [4] Xu, R., Li, X., Zhou, B., & Loy, C. C. (2019). Deep flow-guided video inpainting. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3723-3732).
- [5] Tassano, M., Delon, J., & Veit, T. (2020). Fastdvdnet: Towards real-time deep video denoising without flow estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1354-1363).