

Invisibility Cloak

1. Summary

과제명	Motion Edge Flow 예측과 Gradient 합성을 통한 Invisibility Cloak (투명 망토)의 구현
과제 분야	Computer Vision, AI, Deep Learning-based Image Inpainting
과제 기간	2021.08.27 ~
과제 요약	1. 영상에서 HSV 변환을 통해 특정 색상을 검출하고, 마스크 처리를 하여 해당 영역을 지운 손상된 비디오를 생성한다. 2. 손상된 비디오를 optical color flow field로 변환하고 motion edge를 추출하여 손상 영역의 motion flow edge를 예측하여 연결한다. 3. 완성된 flow edge와 비디오의 gradient domain의 합성으로 손상 영역에 대해 프레임별 image inpainting (이미지 복원)을 수행하고, 복원 이미지를 비디오에 적용하여 손상된 비디오를 복구하여 invisibility cloak (투명 망토)의 효과를 구현한다.

2. Result Overview



3. Plan

3.1. 개요

Image Inpainting (이미지 복원)은 마스크 영역의 주변을 인식하고 새로운 내용을 생성하여 마스크 영역을 재구성하고 복원하는 작업이다. 손상된 이미지의 복원, 특정 대상 제거, 워터마크 및 로고 제거 등 많은 부분에서 활용할 수 있으며, 생성된 내용은 자연스럽게 어색하지 않아야 한다.

이 과제에서는 특정 색으로 마스크된 비디오에서 [FGVC](#) 딥러닝 모델을 활용하여 motion flow edge를 예측하여 연결하고, gradient domain의 합성으로 주변 배경을 복원한 이미지를 생성하여 비디오에 적용하는 invisibility cloak video inpainting을 구현할 것이다.

3.2. 기존 기술의 문제점

기존의 inpainting 방법은 패치 (작은 이미지 영역) 기반 합성 기술을 사용하는데, 이 방법은 비디오에서 이미 존재하는 패치에 대해서만 작업할 수 있기 때문에 속도가 느리고, 합성 후 이미지 품질이 저하되는 경우가 많다. 특히, 비디오에 대해서는 각 프레임마다 inpainting 작업을 실행하기 때문에 복구 능력에 한이 많다.

이를 해결하기 위해 각 프레임별 motion edge에 대해서 complete flow edge connection을 계산하여 패치 기반 합성을 제거하고 수행 시간을 단축한다. 그리고 예측 결과와 비디오의 픽셀 정보를 담고 있는 gradient domain에 대하여 weighted average of color gradients (색상에 대한 가중치 평균 계산)을 수행하고, 새로 복원된 이미지 영역을 원본 비디오에 적용하여 이미지 품질의 향상을 기대할 수 있다.

3.4. 구현 방법

1. 입력된 비디오에서 특정 색상의 검출을 위해서 HSV 채널로 변환하고 마스크 영역을 지정한다.
2. 마스크 영역이 포함된 비디오를 color video와 binary masking video로 나눈다.
3. FlowNet 2.0 기반 [RAFT](#) 모델을 활용하여 color 비디오를 optical color flow field로 변환하고, Canny Edge Detector를 활용하여 인접 프레임과 주변 배경에 대한 motion edge flow를 추출한다.
4. 추출된 motion edge flow에서 마스크된 영역에 대해 [EdgeConnect](#) 모델을 활용하여 motion edge flow를 예측한다.
5. Color video에서 gradient domain을 추출하고, complete motion edge flow connection 채널과 합성한다.
6. 합성된 부분을 기존 마스크 영역에 적용하여 이미지 프레임별로 image inpainting을 수행하고, 해당 프레임을 비디오에 적용하여 video inpainting을 완성한다.

3.5. 효과

1. Optical color flow field 기반으로 edge flow를 예측하기 때문에 패치 기반 합성보다 image inpainting 결과에 대해 **높은 품질**을 얻고, 계산 시간을 단축할 수 있다.
2. Optical color flow field에서 추출된 edge flow의 결과를 예측하기 때문에 실제 비디오 환경의 **자연스러운 edge flow connection** 상태를 얻을 수 있다.
3. 픽셀 정보가 있는 gradient domain을 최종 합성에 사용하기 때문에 **원활한 블렌딩 결과**를 얻을 수 있다.
4. 프레임 단위로 image inpainting을 수행하기 때문에 **비디오에 적용하기 수월**하다.

4. Method

4.1. Overview

1. HSV Conversion

- Color 비디오를 HSV 채널로 변환하고 필요한 부분을 마스크 처리한다.

2. Edge-guided Flow Completion

- **Flow computation** : FlowNet 2.0을 기반으로 하는 RAFT 모델을 활용하여 color 비디오를 optical color flow field로 변환한다.
- **Flow edge completion** : Canny Edge Detector를 활용하여 optical flow에서 마스킹 된 영역에 대해 EdgeConnect 모델을 활용하여 motion edge flow를 완성한다.
- **Flow completion** : 완성된 motion edge flow를 기존의 optical flow에 합성한다.

3. Non-local Temporal Neighbors Completion


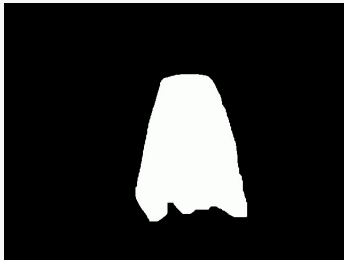

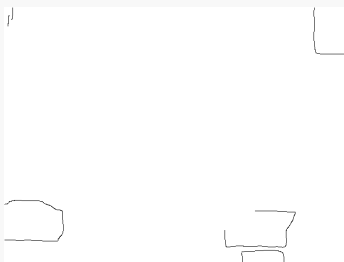
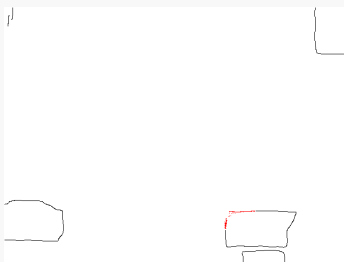


- Non-local 프레임3개를 추가로 생성하고 원래 프레임에서 missing region 를 non-local 프레임을 통해 계산한다.
- `get_flowNN.py` & `get_flowNN_gradient.py` 에서 연산한다.

4. Fusing Temporal Neighbors Completion

- **Gradient-domain processing** : 색상값을 그대로 사용하면 빛, 그림자 등의 문제로 visible seam이 발생하기 때문에 color gradients를 계산하고, 최종 이미지 합성에 Poisson blending을 활용한다.
- `--seamless` 옵션에 따라 `get_flowNN.py` & `get_flowNN_gradient.py` 의 모듈 사용을 결정한다.

5. Result

5.1. Result Image

1. color	2. mask	3. optical flow
		
4. canny edge	5. edge completion	6. flow completion
		
7. result		
		

5.2. Improvement

1. 수행 시간 단축

이미지 프레임에서 전폐적인 edge flow를 분석함으로써 이미지의 모든 영역에 대한 픽셀 검사를 하지 않는다. 즉, 복구 영역에 대해서만 계산하기 때문에 효율적으로 계산 처리를 할 수 있고, image inpainting을 위한 수행 시간을 단축할 수 있다.

2. 이미지 품질 향상

패치 기반 합성 작업이 불필요하기 때문에, 복원된 이미지 영역과 원본 영역을 자연스럽게 연결할 (seamless) 수 있다. 이를 위해 색상 영역과 gradient domain을 활용하여 품질 높은 결과를 도출하는 점이 기존 과제들과 차별화된다.

3. 높은 접근성

기존의 방법에서는 색상 값을 직접 다루기 때문에 조명이나 그림자 등의 환경에 따라 image inpainting의 결과 품질이 저하될 때가 있다. 이 과제에서는 optical color flow field와 edge flow를 함께 합성하여 image inpainting의 품질을 높이고, HSV 변환으로 실생활에서 쉽게 발견할 수 있는 손상 비디오 처리에 대한 해결 방안을 제시할 수 있다.

6. References

[1]	Chen Gao, Ayush Saraf, Jia-Bin Huang, Johannes Kopf. Flow-edge Guided Video Completion. ECCV Sep. 2020.
[2]	Dahun Kim, Sanghyun Woo, Joon-Young Lee, In So Kweon. Deep Video Inpainting. CVPR May 2019.
[3]	Eddy Ilg, Nikolaus Mayer, Tonmoy Saikia, Margret Keuper, Alexey Dosovitskiy, Thomas Brox. FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks. CVPR 2017
[4]	Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, Thomas Huang. Free-Form Image Inpainting with Gated Convolution. ICCV Oct. 2019.
[5]	M. Bertalmio, A.L. Bertozzi, G. Sapiro. Navier-stokes, fluid dynamics, and image and video inpainting. CVPR Dec. 2001.
[6]	RAFT: Recurrent All-Pairs Field Transforms for Optical Flow Zachary Teed, Jia Deng. ECCV 2020
[7]	Zili Yi, Qiang Tang, Shekoofeh Azizi, Daesik Jang, Zhan Xu. Contextual Residual Aggregation for Ultra High-Resolution Image Inpainting. CVPR May 2020.