

비지도 딥 호모그래피 성능 향상을 위한 이미지 마스킹의 실험적 연구

이현용, 김낙우, 이준기, 이병탁
한국전자통신연구원

{hyunyonglee, nwkim, jungi, bytelee}@etri.re.kr

An Empirical Study on Image Masking for Improving Unsupervised Deep Homography

HyunYong Lee, Nac-Woo Kim, Jun-Gi Lee, and Byung-Tak Lee
Electronics Telecommunications Research Institute (ETRI)

요약

본 논문은 이미지 스티칭을 위한 딥 호모그래피의 성능 향상을 위한 방법에 대해서 연구한다. 특히, 입력 이미지의 불필요한 부분을 마스킹처리하여 호모그래피 계산을 위해 사용되는 인공신경망에서 불필요한 특징이 추출되지 않게 함으로써 성능 향상을 꾀하고자 한다. COCO 데이터셋에 기반하여 진행한 실험 결과를 통해 본 연구의 타당성을 검증하며 향후 연구 방향에 대해서 논의한다.

I. 서론

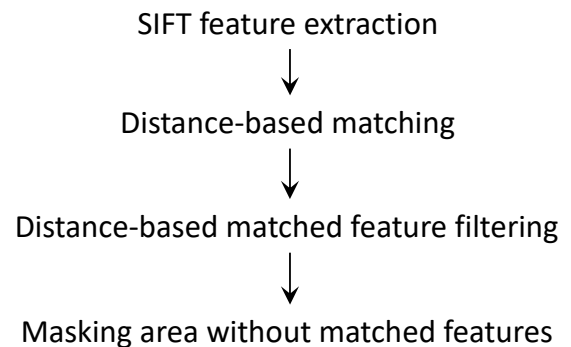
이미지 스티칭(Image Stitching)은 서로 중복된 영역을 포함하는 두 개 이상의 이미지들을 정합하여 하나의 보다 큰 이미지를 생성하는 기술이다. 이미지 스티칭을 위한 전통적인 방법은 특징 기반 방법이다[1]. 특징 기반 이미지 스티칭 방법에서는, 먼저 Scale Invariant Feature Transform(SIFT) 등과 같이 잘 알려진 기법에 기반하여 주어진 이미지들에서 특징 추출을 진행한다. 그리고 특징 매칭을 수행한 후에 매칭된 특징에 기반하여 이미지 정렬을 위한 매트릭스, 즉 호모그래피(homography)를 계산한다. 호모그래피에 기반하여 필요한 이미지들을 변형하고, 변형된 이미지들을 픽셀 기반으로 정합함으로써 이미지 스티칭을 구현한다.

딥 호모그래피는 Convolutional Neural Network(CNN)등과 같은 인공신경망을 활용하여 호모그래피를 계산하는 기법을 말한다[2]. 딥 호모그래피 구현에 있어 중요한 것은 인공신경망을 통해 호모그래피 계산에 사용될 특징을 잘 추출하는 것이다.

본 논문에서는 딥 호모그래피의 성능 향상을 위한 방법을 실험적으로 연구한다. 특히, 입력 이미지의 마스킹을 통해 인공신경망을 통한 불필요한 특징 추출을 억제함으로써 성능 향상을 꾀하고자 한다. 잘 알려진 COCO 데이터셋에 기반하여 제안 방법의 성능을 검증하고 향후 개선을 위한 연구 방향을 논의하고자 한다.

II. 본론

딥 호모그래피 구조. 본 연구에서는 딥 호모그래피 구현을 위해 (그림 1)에 도식된 구조를 사용하였다[3]. 본 연구에서는 두 개의 이미지를 하나의 이미지로 정합하는 것을 고려하였다. 딥 호모그래피의 첫번째 단계는 CNN 을 통해 특징을 추출하는 것이다. 이 때, 정합 대상이 되는 두 개의 이미지에 적용하는 CNN 은 weight 를 공유한다. 두번째 단계는, 두 개의 이미지로부터 추출된 특징에 기반하여 cost volume 을 생성하는 것이다. Cost volume 은 두 개의 이미지로부터

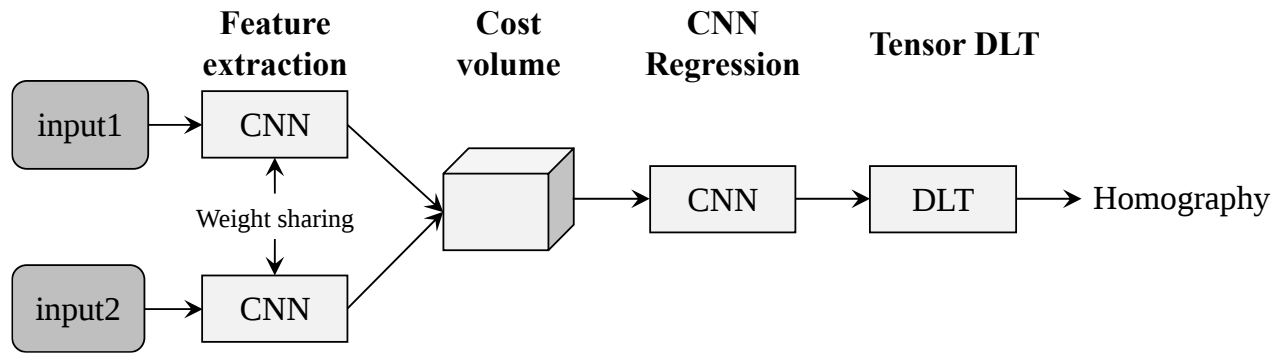


(그림 2) 이미지 마스킹 절차.

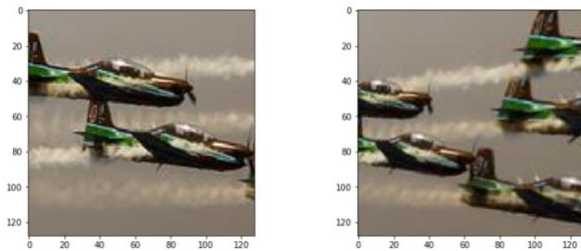
추출된 특징을 결합할 뿐 아니라 특징들 간의 거리를 계산하여 저장한다. 세번째 단계는 cost volume 을 입력으로 하여 CNN 기반 회귀를 진행하는 것이다. 이 단계의 출력은 호모그래피 계산에 사용되는 4 개의 좌표 정보이다. 마지막 단계는, Direct Linear Transformation(DLT)에 기반하여 4 개의 좌표로부터 호모그래피를 도출하는 것이다. 도출된 호모그래피에 기반하여 이미지를 정합하고 두 이미지의 중복되는 영역의 intensity 를 loss 로 간주하여 비지도 학습을 진행한다.

이미지 마스킹. 앞서 설명한 것처럼 딥 호모그래피는 첫단계인 CNN 을 통해 추출된 특징에 기반하여 다음 절차들이 진행된다. 따라서, 이 단계에서 어떤 특징이 추출되느냐에 따라 딥 호모그래피의 성능이 달라진다고 볼 수 있다. 본 연구에서는 딥 호모그래피 계산에 불필요한 특징의 추출되는 것을 막아 성능 향상을 꾀하기 위해서 입력 이미지를 마스킹하는 방법을 제안한다.

(그림 2)는 본 연구에서 사용하는 이미지 마스킹 절차를 보여준다. 먼저 SIFT 로 특징을 추출한 뒤에, 특징간 거리에 기반하여 특징들을 매칭한다. 계산 오류를 줄이기 위해서, 매칭된 특징들 중에서 일정 거리 이내의 특징들만 매칭된 특징으로 간주한다. 마지막으로, 매칭된 특징들이 없는 부분을 검정색으로 마스킹한다. 이 때



(그림 1) 딥 호모그래피 네트워크 구조.



(그림 3) 이미지 스티칭 대상 원본 이미지.

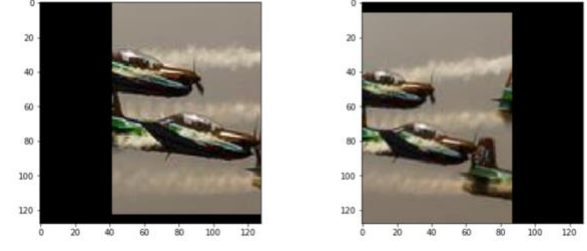
매칭된 특징들이 있는 부분을 사각형으로 인식하여 마스킹을 진행한다. (그림 3)은 마스킹 처리 하기 전의 원본 이미지를, (그림 4)는 마스킹 처리된 이미지의 예시를 보여준다.

데이터. 성능 검증을 위한 기본 데이터로 COCO 데이터셋트를 사용하였다. 이미지 정합을 위한 데이터셋 구성을 위해 하나의 COCO 원본 이미지에서 50%가 중복되는 두 개의 세그먼트를 추출하고 하나의 세그먼트에 임의의 변환을 적용하였다. 이와 같은 방식으로 50,000 개의 학습데이터와 1,106 개의 검증 데이터를 생성하였다.

실험 결과. 이미지 마스킹의 성능 검증을 위해 마스킹 처리되지 않은 이미지를 기반으로 학습, 검증을 진행한 경우와 마스킹 처리된 이미지를 기반으로 학습, 검증을 진행한 경우의 성능을 비교하였다. 본 연구에서 적용하는 비지도학습의 성능 검증을 위해서 두 개의 이미지가 정합되는 부분의 Peak Signal-to-Ratio(PSNR)과 Structural Similarity Index Map(SSIM)을 성능 메트릭으로 사용한다.

1,100 여개의 검증 데이터 중에서 입력 이미지의 마스킹을 통해 성능이 향상된 경우는 310 개이다. 이 경우 성능 향상 평균은 PSNR 기준 0.596 이며 SSIM 기준 0.023 이다. 반면, 오히려 성능이 저하된 경우는 795 개이다. 이 경우 성능 저하 평균은 PSNR 기준 1.26 이며 SSIM 기준 0.056 이다.

이 실험을 통해 일차적으로 알 수 있는 것은 이미지 마스킹이 모든 경우에 성능 향상을 도출하지 않는다는 것이다. 어떠한 경우에 성능이 향상되고 또는 저하되는지 파악하기 위해서 추가적인 분석을 진행하였다. 이 분석에서는 성능 변화 정도가 PSNR 기준 1 미만으로 미미한 경우는 제외하였다. 추가 분석 결과 성능 향상이 있는 경우의 마스킹 미적용 경우의 평균 PSNR 은 21.18, SSIM 은 0.688 이며, 성능 저하가 있는 경우의 마스킹 미적용 경우의 평균 PSNR 은 23.81, SSIM 은 0.799 임을 확인하였다. 이는 대체로 마스킹 미적용의 경우에 성능이 그리 좋지 않은 경우에, 마스킹을 적용하였을 경우 성능 향상이 된다는 것을 보여준다.



(그림 4) 이미지 스티칭 대상 마스킹 처리된 이미지.

향후 연구 방향. 위의 실험 결과에 기반하여 본 연구의 향후 방향에 대해서 간단히 논의한다. 첫번째로 적용할 수 있는 개선 방안은 이미지 마스킹 미적용시의 성능을 먼저 파악한 후에 개선 여지가 있는 경우에만 선택적으로 이미지 마스킹을 적용하는 것이다. 또 다른, 개선 방안은 입력 이미지에 적용하는 마스킹만으로는 한계가 있는 것으로 판단되어 (그림 1)에서 도식화된 구조의 개선이 함께 필요해보인다.

III. 결론

본 논문에서는 비지도학습 기반 딥 호모그래피 성능 개선을 위한 이미지 마스킹 방법의 타당성을 실험적으로 연구하였다. 또한, 실험 결과에 기반하여 향후 연구 방향을 설정하였다.

ACKNOWLEDGMENT

이 논문은 2022 년도 정부(산업통상자원부)의 재원으로 한국에너지기술 평가원(KETEP)의 신재생에너지핵심기술 개발사업으로 지원을 받아 수행한 연구 과제입니다. (No. 20223030020070)

참 고 문 헌

- [1] K.Y. Lee and J.Y. Sim, "Warping residual-based image stitching for large parallax," in Proc. of CVPR, 2020.
- [2] Nguyen, T. et al. "Unsupervised deep homography: A fast and robust homography estimation model," IEEE Robot. Autom. Lett., vol.3, pp.2346-2353, 2018.
- [3] Nie, L. Laio, K. Liu, S. and Zhao, Y. "Unsupervised deep image stitching: reconstructing stitched features to images," IEEE Transactions on Image Processing, vol.30, pp.6184-6197, 2021.