

Mask R-CNN을 활용한 Priority Object 영상 스티칭

*이성배 ** 김규현

경희대학교

*rhee@khu.ac.kr **kyuheonkim@khu.ac.kr

Mask R-CNN based Priority Object Image Stitching

*Rhee, Seong Bae **Kim, Kyuheon

Kyunghee University

요약

최근 Panorama와 360도 영상이 대표되는 몰입형(Immersive) 미디어 콘텐츠의 사용이 증가하고 있다. 몰입형 영상 콘텐츠는 사용자에게 현장감을 제공해야 하지만, 촬영 카메라 간의 시차(Parallax)로 인해 영상 콘텐츠에서 시차 왜곡이 발생할 수 있고, 이는 사용자의 콘텐츠 몰입을 제한하기 때문에 해당 영상 콘텐츠의 제작 기술인 영상 스티칭의 높은 정확도가 요구되고 있다. 지금까지 스티칭 영상의 시차 왜곡을 줄이기 위하여 다중 호모그래피 추정 방법과 Seam Optimization 방법이 제안되었지만, 영상 내 사물 배치에 따라 기술 적용이 제한될 수 있다. 이에 본 논문에서는 Mask R-CNN을 활용하여 사물을 세그먼트화하고, 사물의 종류에 따라 각각 다른 가중치 적용을 통해 시차 왜곡을 방지하며, 영상 내 사물의 배치에 따라 시차 왜곡이 발생할 상황에 서는 사용자의 인지 중요도가 낮은 사물로 시차 왜곡을 유도하는 영상 스티칭 방법을 제안한다.

1. 서론

최근 Panorama, 360도 영상이 대표되는 몰입형(Immersive) 미디어 콘텐츠의 사용이 증가하고 있다. 이러한 몰입형 영상 콘텐츠는 사용자에게 현장감을 제공해야 하지만, 촬영 카메라 간의 시차(Parallax)로 인해 영상 콘텐츠에서 사물이 중복해서 나타나거나 사라지는 시차 왜곡이 발생할 수 있고[1], 이러한 실제와 다른 모습은 사용자의 콘텐츠 몰입을 제한하므로 해당 콘텐츠의 제작 기술인 영상 스티칭 기술의 높은 정확도가 요구되고 있다.

영상 스티칭 기술은 기존 카메라 한 대를 통해서는 파노라마 및 360도 영상 콘텐츠를 생성함에 있어 카메라의 시야각(Field of View)이 제한되므로, 다수의 카메라를 통해 촬영한 영상을 넓은 시야각을 가지는 하나의 영상으로 합성하는 기술이다[2]. 이때 촬영하는 카메라 간의 시차로 인해 왜곡 현상이 발생할 수 있고, 영상 스티칭 분야에서는 시차 왜곡을 극복하기 위하여 다중 호모그래피를 추정하고 다중 와핑을 진행하는 방법[3]과 사물을 피해 Seam을 생성하는 Seam Optimization[4] 방법이 제안되고 있다.

다중 호모그래피 추정 및 다중 와핑을 통해 시차 왜곡을 극복하는 방법은 영상을 패치 단위로 분할하고 각각의 패치에 대하여 변환 행렬을 계산하므로, 하나의 호모그래피를 계산하는 방법보다 입력 영상 간의 차이를 줄일 수 있고, 이와 같은 효과를 통해 시차 왜곡을 줄이는 방법이다. 그러나 패치 경계 부분에서 Local Distortion[5]이 발생할 수 있고, 해당 왜곡을 교정하는 것은 복잡도가 높다는 단점으로 활용에 제한이 있다.

한편, Seam Optimization 방법은 두 입력 영상의 중복 영역에서 픽셀의 차이 값[6]이나 시각적 인지 에너지 함수[7] 등으로 생성한

Seam 생성 행렬을 바탕으로 합계가 최소인 경로로 Seam을 생성한다. 이때 Seam 생성 행렬에서 사물의 위치에는 큰 값이 할당되므로 Seam은 사물 영역을 피해 생성되고, 결국 시차 왜곡이 발생할 확률은 낮아진다. 그러나 영상 내 사물의 크기, 모양, 배치 등에 따라서 올바른 Seam을 생성하지 못하는 상황이 발생할 수 있고, 이로 인해 시차 왜곡이 발생할 수 있다는 단점이 존재한다.

이에 본 논문에서는 Seam Optimization 기반 영상 스티칭 기술이 가진 제한 사항을 분석하고, 이를 해결하기 위해 영상 내 사물의 종류에 따라 다른 가중치 적용을 추가한 Seam 생성 행렬을 통하여, 시차 왜곡에 강인한 영상 스티칭 방법을 제안하고자 한다. 본 논문의 2장에서는 기존 Seam Optimization 방법에서 사물의 위치 정보를 반영하는 방법을 간단하게 살펴보고, 3장에서는 기존 방법의 제한 사항을 해결하기 위한 딥러닝 기반의 사물 검출기를 활용한 Priority Object 영상 스티칭 방법을 설명한다. 또한, 4장에서는 테스트 영상을 통해 제안 방법을 검증하며 마지막 5장에서 결론을 내린다.

2. 기존 방법 분석

앞서 설명한 바와 같이 Seam Optimization 기반의 영상 스티칭 방법은 입력 영상 간의 중복 영역에 대하여 픽셀의 차이 값이나 에너지 함수 등으로 생성한 Seam 생성 행렬을 바탕으로 합계가 최소인 경로로 Seam을 생성한다. 이때 사물 영역에는 큰 값의 가중치가 설정되어 Seam으로 선택되는 것이 방지되는데, 이에 기존 기술에서는 사물 영역을 반영하기 위하여 사물의 윤곽선과 같은 고주파수 지점을 활용하거나 [7] 사물 영역 전체를 세그먼트화하여 반영하고 있다[6].

영상의 주파수 성분을 통해 사물 영역을 반영하는 방법은 영상 내 고주파 위치를 피해 Seam을 생성하는 방법으로써, 저주파 위치를 따라 Seam을 생성하여 사물을 우회하는 방법이다. 그러나 초기 Seam이 생성된 위치에 따라서, 사물 단면의 저주파 성분을 따라 Seam이 생성되어 시차 왜곡이 발생할 수 있다.

반면, 사물 영역을 세그먼트화하여 반영하는 방법은 사물의 윤곽 선을 획득하고 이를 그룹화하여 사물의 모양 단위로 반영하는 방법으로써, 사물 영역 자체를 Seam 생성 행렬에 가중치로 반영하므로 사물 영역 내에서 Seam이 생성될 가능성은 앞선 방법보다 낮은 편이다. 그러나 사물을 세그먼트화하여 반영하는 방법은 세그먼트 기술의 성능에 따라서 사물 영역을 제대로 나타내지 못할 수 있고, 검출하지 못한 사물 영역 위로 Seam이 생성된다면 시차 왜곡이 발생할 수 있다.

또한, 앞선 두 가지 방법 모두 영상 내 사물이 그림 1의 (a)와 같이 나타난 경우, 올바른 Seam 생성에 제한될 수 있다. 영상 내 사물의 모양을 직관적으로 확인하기 위하여 그림 1의 (b)와 같이 세그먼트화한다면, 수평으로 넓게 퍼진 사물 영역이 Seam 생성 행렬에 반영되었기 때문에 사물 영역을 피해 공통 영역을 분할할 수 있는 Seam을 생성하는 것은 불가능하고, 결국 Seam은 사물 영역을 지나면서 생성되게 된다. 이와 같이 Seam이 사물 위로 생성될 수밖에 없는 상황에서는 Seam이 차선이나 가드레일 위치를 지나면서 생성되는 것이 차량 위로 생성되는 것보다 좋은 영상 결과를 만들겠지만, 사물의 중요도를 구분할 방법이 없다.



그림 1. Seam 생성이 제한되는 예시

3. Priority Object based Image Stitching

앞서 설명한 바와 같이, Seam 생성 행렬에 사물 영역을 반영하기 위하여 주파수 성분을 이용하는 방법은 Seam의 초기 생성 위치에 따라서, 사물 내부의 저주파 성분을 따라 Seam이 생성될 수 있는 문제가 있었고, 사물의 세그먼트를 이용한 방법은 세그먼트 기술의 성능에 따라서 검출하지 못한 사물 영역 위로 Seam이 생성될 수 있는 문제가 있었다. 또한, 두 가지 방법 모두 영상 내 사물의 모양, 배치 등에 따라서 올바른 Seam 생성이 제한되는 문제가 있었기에, 본 논문에서는 영상 내 주파수 성분을 통해 Seam 생성 행렬을 생성하는 방법을 바탕으로 Mask R-CNN[8]을 통해 검출한 사물의 종류에 따라서 다른 가중치 값을 Seam 생성 행렬에 더하여 Seam이 사물을 피해 생성하고, 앞선 그림 1과 같이 Seam이 사물 영역을 피하지 못해 시차 왜곡이 불가피한 상황에서는 사용자의 영상 인식에서 중요도가 가장 낮은 사물로 시차 왜곡을 유도하는 방법을 제안하고자 한다.

제안 기술은 Mask R-CNN의 사물 세그먼트 기술을 활용하여, 검출된 사물의 종류에 따라서 각각 다른 중요도를 설정하고, 이에 중요도가 높은 사물의 위치에는 큰 가중치 값을 Seam 생성 행렬에 할당하

고, 중요도가 낮은 사물의 위치에는 작은 가중치 값을 설정한다. 그림 2의 (a)와 같이 사물이 수평으로 넓게 퍼져 Seam 생성이 어려운 영상을 Mask R-CNN을 통해서 영상 내 사물을 검출하고, 중요도에 따라 다른 가중치를 적용한 모습은 그림 2의 (b)와 같이 나타난다. 이때 그림 2의 (b)에서 밝기는 상대적인 가중치 값을 표현하였으며, 흰색에 가까울수록 높은 가중치가 설정된 영역이다.



그림 2. Priority Object 가중치 설정 예시

그림 2의 (b)와 같이 사물의 중요도에 따라서 다른 가중치 값으로 표현된 영상 정보는 기존 주파수를 활용한 Seam 생성 행렬에 더해지므로, 기존 주파수를 활용한 Seam 생성 행렬이 가지고 있는 초기 Seam 위치에 따라 사물 내부의 저주파 성분으로 Seam이 생성되는 문제를 해결할 수 있으며, 사물의 모양, 배치에 따라서 올바른 Seam을 생성할 수 없어 시차 왜곡이 불가피한 상황에서 영상 내 사용자의 중요도가 낮은 사물로 시차 왜곡을 유도할 수 있다.

4. 실험 결과

본 논문에서는 전통적인 Seam Optimization 기술의 주파수 성분을 활용한 Seam 생성 행렬을 바탕으로, Mask R-CNN을 통해 검출한 사물의 종류에 따라 다른 가중치를 Seam 생성 행렬에 더함으로써, 기존 기술이 가진 제한 사항을 극복하는 Mask R-CNN 기반의 Priority Object 영상 스티칭 방법을 제안하고 있다. 이에 본 논문의 4장에서는 아래 그림 3과 같이 시차가 큰 세 개의 입력 영상을 스티칭하는 실험을 통해 제안 기술을 검증하고자 한다. 기술 검증 및 성능 평가의 실험 환경은 Windows 10, i7-6700의 CPU, RTX-2080의 GPU와 VS Code를 통해 설계되었다.



그림 3. 시차가 큰 테스트 입력 영상

그림 3과 같이 시차가 큰 세 개의 입력 영상을 스티칭하기 위하여 각각의 영상에서 특징점[9]을 추출하고, 공통 영역을 설정하여 호모그래피를 추정하기 위해 사용된 기술자는 ORB[10]로 그림 4와 같이 나타난다.



그림 4. ORB 매칭 적용 예시

한편, Mask R-CNN을 통하여 사물 영역을 세그먼트화한 것은 그림 5와 같이 나타나며, 본 실험에서는 사물을 검출한 위치와 대응되는 공통 영역의 Seam 생성 행렬 위치에는 Seam 생성 행렬 값 중에서 가장 큰 값의 10배에 해당하는 값을 가중치로 설정하였다.



그림 5. Mask R-CNN을 통한 검출

최종적으로 기존 Seam Optimization 기반의 영상 스티칭 결과와 제안 기술을 적용한 영상 스티칭 결과는 그림 6과 같이 나타난다. 기존 기술에서는 검은색 차량에 Seam이 생성되어 시차 왜곡이 발생한 것을 확인할 수 있고, 제안 기술에서는 차량에 대한 가중치가 설정되어 어떤 차량에서도 시차 왜곡이 발생하지 않은 것을 확인할 수 있다. 이와 같은 실험을 통하여, 본 논문에서 제안하는 Mask R-CNN 기반 Priority Object 영상 스티칭 방법을 기존 Seam Optimization 기술에 적용할 때 보다 좋은 품질의 스티칭 영상을 획득할 수 있음을 확인하였다.

5. 결론

기존 스티칭 기술의 시차 왜곡 문제를 해결하기 위하여 Seam Optimization 기반의 영상 스티칭 기술은 사물 영역을 우회하여 Seam을 생성하지만, Seam의 초기 생성 위치, 사물 검출기의 성능, 영상 내 사물의 모양 및 배치 등으로 인해 올바른 Seam을 생성하는 데 제한 사항이 있었다.

이에 본 논문에서는 Mask R-CNN을 활용하여 사물을 검출하고 검출한 사물의 종류에 따라 다른 가중치를 설정하였고, 이를 테스트 영상을 통한 실험으로 확인함으로써, 기존 Seam Optimization 기반의 영상 스티칭 기술이 가진 제한 사항을 제안 기술을 추가함으로써 극복할 수 있음을 검증하였다.

추가적으로, Mask R-CNN을 통해 사물의 종류에 따라 다른 가중치 설정하더라도 Seam을 생성할 수 없는 상황에 대하여, 다시 말해, 영상 내 동일한 사물이 수평으로 형성되어 모두 같은 가중치를 지니게 되어, Seam을 생성할 수 없는 상황으로 인해 발생하는 시차 왜곡을 교정하기 위한 추가적인 연구를 진행함으로써, 보다 시차 왜곡을 극복한 개선된 스티칭 기술을 제공할 수 있을 것으로 기대한다.

* 본 연구는 한국전력공사의 2018년 착수 에너지 거점대학 클러스터 사업에 의해 지원되었음 (과제번호:R18XA02)



(a) 기존 기술 결과



(b) 제안 기술 적용 결과

그림 6. 기존 기술과 제안 기술 비교

참고문헌

- [1] Rhee, Seongbae, Jeonho Kang, and Kyuheon Kim. "Local Differential Pixel Assessment Method for Image Stitching." *Journal of Broadcast Engineering* pp.775-784, Vol.24, No.5, 2019.
- [2] R. Szeliski, "Image Alignment and Stitching: A Tutorial." *Foundations and Trends in Computer Graphics and Computer Vision*, Vol. 2, No.1, 2006.
- [3] Zaragoza, Julio, et al. "As-projective-as-possible image stitching with moving DLT." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013.
- [4] Lin, Kaimo, et al. "Seagull: Seam-guided local alignment for parallax-tolerant image stitching." *European conference on computer vision*. Springer, Cham, 2016.
- [5] Wei, L. Y. U., et al. "A survey on image and video stitching." *Virtual Reality & Intelligent Hardware* Vol. 1, No.1 pp.55-83, 2019.
- [6] Kang, Jeonho, et al. "Minimum Error Seam-Based Efficient Panorama Video Stitching Method Robust to Parallax." *IEEE Access* 7 (2019): 167127-167140.
- [7] Kwatra, Vivek, et al. "Graphcut textures: image and video synthesis using graph cuts." *ACM Transactions on Graphics (ToG)*, Vol.22, No.3, pp.277-286. 2003.
- [8] He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [9] Matthew Brown and David G. Lowe. "Automatic panoramic image stitching using invariant features" *International Journal of Computer Vision*. Vol. 74, No.1, pp.55-73. 2007.
- [10] Rublee, Ethan, et al. "ORB: An efficient alternative to SIFT or SURF." *2011 International conference on computer vision*. Ieee, 2011.