

특집논문 (Special Paper)

방송공학회논문지 제25권 제6호, 2020년 11월 (JBE Vol. 25, No. 6, November 2020)

<https://doi.org/10.5909/JBE.2020.25.6.882>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭

이 성 배^{a)}, 강 전 호^{a)}, 김 규 현^{a)†}

Image Stitching focused on Priority Object using Deep Learning based Object Detection

Seongbae Rhee^{a)}, Jeonho Kang^{a)}, and Kyuheon Kim^{a)†}

요 약

최근 Panorama와 360° 영상이 대표되는 몰입형 미디어 콘텐츠의 활용이 증가하고 있다. 일반적인 카메라 한 대를 통해서 해당 콘텐츠를 생성하기에는 시야각이 제한되기 때문에, 다수의 카메라로 촬영한 영상을 넓은 시야각을 갖는 하나의 영상으로 합성하는 영상 스티칭이 주로 사용되고 있다. 그러나 촬영하는 카메라 간의 시차(Parallax)가 크다면 스티칭 영상에서 시차 왜곡이 발생할 수 있고, 이는 사용자의 콘텐츠 몰입을 제한하기 때문에 시차 왜곡을 극복할 수 있는 영상 스티칭 기술이 필요하다. 시차 왜곡을 극복하기 위한 기존의 Seam Optimization 기반 영상 스티칭 방법은 사물의 위치 정보를 반영하기 위하여 에너지 함수나 객체 세그먼트 정보를 활용하고 있지만, 초기 Seam 생성 위치, 배경 정보, 사물 검출기의 성능 그리고 사물의 배치 등의 제한 사항으로 인해 기술의 적용이 제한될 수 있다. 이에 본 논문에서는 딥러닝 기반 사물 검출을 활용하여 사물의 종류에 따라 다르게 설정한 가중치 값을 시각적 인지 에너지 값에 더함으로써, 기존 기술의 제한 사항을 극복할 수 있는 영상 스티칭 방법을 제안하고자 한다.

Abstract

Recently, the use of immersive media contents representing Panorama and 360° video is increasing. Since the viewing angle is limited to generate the content through a general camera, image stitching is mainly used to combine images taken with multiple cameras into one image having a wide field of view. However, if the parallax between the cameras is large, parallax distortion may occur in the stitched image, which disturbs the user's content immersion, thus an image stitching overcoming parallax distortion is required. The existing Seam Optimization based image stitching method to overcome parallax distortion uses energy function or object segment information to reflect the location information of objects, but the initial seam generation location, background information, performance of the object detector, and placement of objects may limit application. Therefore, in this paper, we propose an image stitching method that can overcome the limitations of the existing method by adding a weight value set differently according to the type of object to the energy value using object detection based on deep learning.

Keyword : Image Stitching, Parallax Distortion, Seam Optimization, Object Detection, Priority Object

1. 서론

최근 파노라마와 360도 영상으로 대표되는 몰입형(Immersive) 미디어 콘텐츠는 게임, 스포츠, 쇼핑 등 다양한 산업에서 활용되고 있으며, 교육, 군사, 의료, 제조 등 다양한 전문 분야로의 확산까지 전망되고 있다. 몰입형 미디어 콘텐츠는 기존 영상 콘텐츠와 달리 사람의 육안보다 더 넓은 시야각(Field of view)을 갖는 초고화질 영상으로써, 좁은 시야각을 갖는 전통적인 구조의 카메라 한 대를 통해서만 해당 콘텐츠의 생성이 제한되기 때문에 주로 광각 렌즈가 부착된 카메라를 통해 촬영되거나, 다수의 일반 카메라로 촬영한 영상을 하나의 영상으로 합성하는 영상 스티칭 기술을 통해 생성된다. 광각 카메라 렌즈를 통해 촬영한 영상은 렌즈의 곡률로 인해 실제 모습과 현저히 다르게 나타나는 방사 왜곡(Radial distortion)이 발생하므로, 사물의 모습이 중요한 몰입형 미디어 콘텐츠는 주로 영상 스티칭 기술을 통해 생성되고 있다^[1].

영상 스티칭 기술은 특징점(Keypoint) 추출 및 매칭, 공통 영역 설정, 호모그래피(Homography) 계산, 왜곡(Warping) 그리고 합성까지의 모든 과정을 지칭하며^[2], 영상 스티칭을 통해 생성된 몰입형 미디어 콘텐츠는 광각 렌즈를 통해 생성한 영상보다 방사 왜곡이 적게 나타난다. 그러나 촬영하는 카메라 간에 시차(Parallax)가 크다면 영상 내 사물이 사라지거나 중복되어 나타나는 시차 왜곡(Parallax distortion)이 발생할 수 있으며, 이는 사용자의 콘텐츠 몰입을 방해할 수 있다^[3].

따라서, 사용자의 콘텐츠 몰입을 저해하는 시차 왜곡을 줄이기 위해 많은 연구들이 진행되었으며, 그중에서 다중 호모그래피 기반의 영상 스티칭 방법과 Seam Optimization

기반 영상 스티칭 방법이 대표적이다^[4]. 다중 호모그래피 기반의 영상 스티칭 방법은 영상을 여러 개로 분할하고, 각각 분할된 영상별로 호모그래피를 추정하고 왜곡하여 영상 간의 차이를 최소화함으로써^[5], 스티칭 영상의 시차 왜곡을 줄이는 방법이다. 그러나 분할 경계 부분에서 Local distortion이 발생할 수 있고, 해당 왜곡을 교정하는 것은 복잡도가 높다는 제한 사항이 있다^[4].

반면, Seam Optimization 기반 영상 스티칭 방법은 두 입력 영상의 중복 영역에서 시각적 인지 에너지 함수나 객체 세그먼트 정보를 통해 Seam 생성 행렬을 구성하고, 행렬의 누적이 최소인 경로를 따라 Seam을 생성하여 합성하는 방법이다^[6]. 이때 Seam 생성 행렬의 사물 영역에는 높은 가중치가 할당되어 Seam으로 선택되지 않는 특징을 활용하여, 사물에서 시차 왜곡이 발생하는 것을 방지하는 기술이다. 하지만 Seam Optimization 기반 영상 스티칭에서는 Seam 생성 초기 위치, 사물 검출기의 성능, 사물의 배치 등에 따라서 올바른 Seam을 생성하는 것이 제한될 수 있고, 이로 인해 시차 왜곡이 발생할 수 있다^[4].

앞서 설명한 바와 같이, 다중 호모그래피 추정 방법과 Seam Optimization 방법 모두 스티칭 영상에서 원치 않는 시차 왜곡이 발생할 수 있다. 다중 호모그래피 추정 방법은 영상을 분할한 경계선에서 Local distortion이 발생할 수 있기 때문에 실질적인 활용에 어려움이 존재하지만, Seam Optimization 방법의 제한 사항은 올바른 Seam 생성 행렬을 새로 정의함으로써 극복할 수 있다. 이에 본 논문에서는 Seam Optimization 기반의 영상 스티칭을 바탕으로, 딥러닝 기반의 사물 검출을 활용하여 사물의 종류에 따라 다른 가중치를 설정하고, 이를 Seam 생성 행렬에 적용함으로써 시차 왜곡을 극복할 수 있는 영상 스티칭 방법을 제안하고자 한다.

이에 본 논문의 2장에서는 기존 Seam Optimization 기반의 영상 스티칭에서 사물의 위치 정보를 Seam 생성 행렬에 반영하는 방법을 간단하게 살펴보고, 이때 Seam Optimization 기반의 영상 스티칭이 갖는 제한 사항을 분석하고자 한다. 이어지는 3장에서는 기존 Seam Optimization 기반 영상 스티칭 방법의 제한 사항을 극복하기 위하여 딥러닝 기반의 사물 검출을 활용한 우선순위 사물 중심 영상 스티칭 방법을 제안하며, 4장에서는 테스트 영상으로 실험을 진행

a) 경희대학교 전자공학과(Department of Electronic Engineering, Kyung Hee University)

‡ Corresponding Author : 김규현(kyuheon Kim)

E-mail: kyuheonkim@khu.ac.kr

Tel: +82-31-201-3810

ORCID: <http://orcid.org/0000-0003-1553-936X>

※ 이 논문의 연구결과 중 일부는 “2020년 한국방송·미디어공학회 하계학술대회”에서 발표한 바 있음.

※ This research was supported by Korea Electric Power Corporation. (Grant number:R18XA02)

· Manuscript received September 14, 2020; Revised October 21, 2020; Accepted October 27, 2020.

하여 제안 기술의 효용성을 검증하고자 한다. 마지막으로 5장에서는 결론을 지으며 마무리한다.

II. 배경 기술

앞서 설명한 바와 같이 일반적인 영상 스티칭 과정은 여러 장의 입력 영상에 대하여 영상의 크기, 회전, 밝기 등에 강인한 특징점을 추출(Feature extraction)하는 것으로 시작된다. 각 입력 영상에서 검출된 다수의 특징점 중에서 동일한 쌍을 찾는 매칭 작업(Feature matching)을 바탕으로 호모그래피(Homography)를 계산하고, 호모그래피에 맞지 않은 지점(Outlier)은 RANSAC을 통해 제거된다. 이때 보존된 특징점을 모두 포함할 수 있는 영역을 공통 영역(Overlap Region)으로 설정하고 와핑(Warping)을 진행한다. 이에 동일한 평면으로 정렬된 다수의 공통 영역 영상은 가중치 연산이나 Seam을 바탕으로 합성(Compositing)되며, 입력 영상 간의 밝기 차이를 보정하는 Blending 작업이 수행되기도 한다²⁾. 또한, 최근에 연구되고 있는 Seam Optimi-

zation 기반의 영상 스티칭 방법은 시차 왜곡을 극복하기 위하여 배경을 중심으로 합성하는 것이 아닌, 영상 내 객체를 중심으로 스티칭하고 있으며, 이때 영상 내 객체 정보를 활용할 수 있도록 Object detection 모듈을 추가하여 그림 1과 같이 나타난다⁴⁾.

본 논문에서 활용하고자 하는 Seam Optimization 기반의 영상 스티칭 방법은 영상을 합성할 때 공통 영역 간의 픽셀 차이 값을 기준으로 시각적 인지 에너지 함수^[7,8,9]나 객체 세그먼트^[10,11]의 사물 반영 기술을 Seam 생성 행렬에 적용하고, Seam 생성 행렬 누적이 최소인 경로를 Seam으로 설정하여 영상을 합성하는 방법이다. 이때 Seam 생성 행렬의 사물 영역에는 높은 가중치가 설정되어 Seam으로 선택되지 않기에 시차 왜곡이 발생할 확률을 최소화한다. 하지만, Seam Optimization 기반의 스티칭 영상에서는 여전히 시차 왜곡이 발생하고 있으며, 시각적 인지 에너지 함수를 활용한 방법과 객체 세그먼트 방법 각각에서 시차 왜곡이 발생하는 이유는 다음과 같다.

먼저, 시각적 인지 에너지 함수를 활용하는 방법은 식 (1)과 같이 나타난 에너지 값을 Seam 생성 행렬에 적용하고

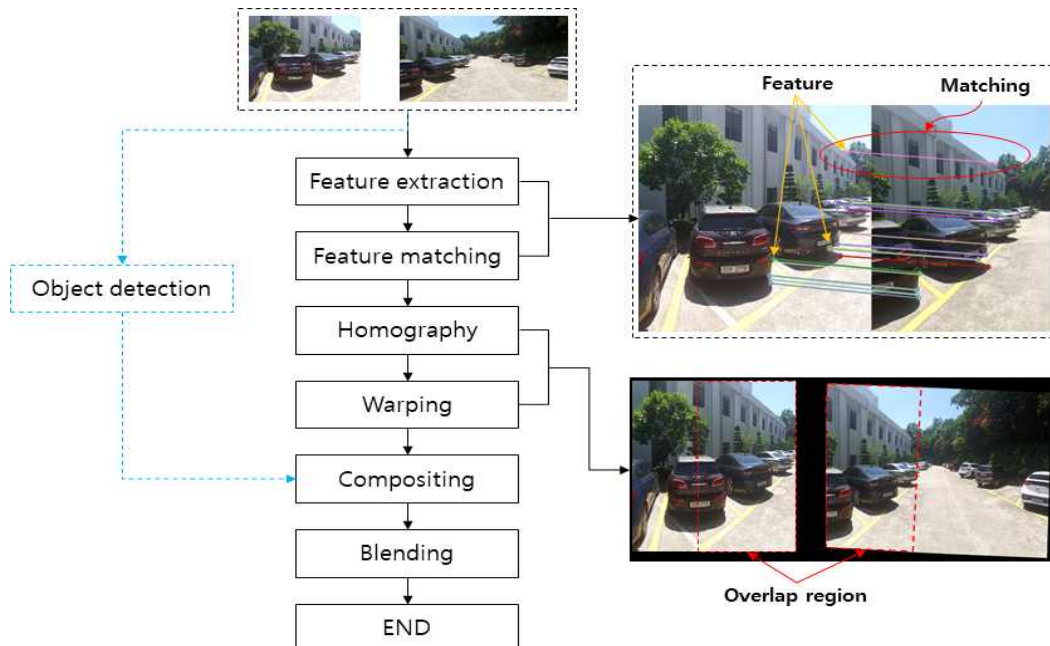


그림 1. Seam optimization 기반 영상 스티칭 과정
Fig. 1. Process of seam optimization based image stitching

에너지 값이 작은 저주파수 영역을 따라 Seam을 생성하는 방법으로써^[7], 초기 Seam 생성 위치에 따라서 시차 왜곡이 발생할 수 있다. 식 (1)에서 나타난 바와 같이 수평 변화의 크기와 수직 변화의 크기를 더한 에너지 값은 영상 내 사물의 윤곽선(Edge)과 비슷한 모습으로 나타나며, 그림 2의 (a)를 예시로 에너지 함수를 구한다면 그림 2의 (b)와 같이 나타난다. 이때 그림 2의 (b)에서 나타난 빨간색 점에서 Seam이 처음 생성된다면, Seam은 사물의 내부 영역을 따라 생성되기에 시차 왜곡이 발생할 수 있다. 또한, 에너지 값은 픽셀의 변화율로 정의되기 때문에 배경과 비슷한 픽셀 값을 가지는 사물 영역에서는 가중치가 제대로 반영되지 않는다는 제한 사항이 있고, 이는 그림 2의 (b)에서 노란색으로 표시된 영역에서 확인할 수 있다.

$$E(x,y) = \left| \frac{\partial}{\partial x} I(x,y) \right| + \left| \frac{\partial}{\partial y} I(x,y) \right| \quad (1)$$

반면, 객체 세그먼트를 활용하는 방법은 영상 내 사물의 윤곽선을 획득하고 이를 그룹화하여 사물의 모양 단위로 Seam 생성 행렬에 높은 가중치를 반영하는 방법으로써^[11], 앞선 그림 2의 (a)에 대하여 객체 세그먼트를 적용한 예시는 그림 3과 같이 나타난다. 그림 3에서 나타난 바와 같이 객체 세그먼트를 활용하는 방법은 그림 2-(b)의 예시로 나타나는 시각적 인지 에너지 함수를 활용한 방법과 달리 사물 내부의 영역 전체에 대하여 높은 가중치를 적용하기 때문에 초기 Seam 생성 지점은 사물 내부 영역이 아닌 다른

지점에서 결정된다. 이에 따라서, 객체 세그먼트를 활용하는 방법은 시각적 인지 에너지 함수를 활용하는 방법보다 초기 Seam 생성 위치로 인해 시차 왜곡이 발생하는 경우가 적다는 장점이 있다.



그림 3. 객체 세그먼트 예시
 Fig. 3. Example of object segmentation

그러나 객체 세그먼트를 활용하는 방법은 사물 검출기의 검출 성능에 따라 시차 왜곡이 발생할 수 있다. 그림 4에서 나타난 바와 같이 실제 사물 영역이지만 검출하지 못하는 영역이 있을 수 있고, 검출하지 못한 사물 영역 위로 Seam이 생성된다면 시차 왜곡이 나타날 수 있다^[12]. 또한, 객체 세그먼트를 통한 방법은 영상 내 객체를 추출하는 것이 목적인 기술로써, 바닥과 벽면 등의 배경 정보를 제대로 Seam 생성 행렬에 반영하지 못하기 때문에 스티칭 영상의 배경 부분에서 시차 왜곡이 발생할 수 있다^[11].



(a)



(b)

그림 2. 에너지 함수 예시
 Fig. 2. Example of energy function

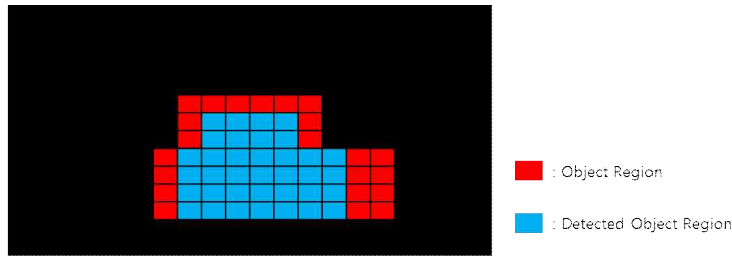


그림 4. 객체 세그먼트의 오검출
Fig. 4. Misdetection of object segment

또한, 앞서 설명한 에너지 함수를 활용한 방법과 객체 세그먼트를 활용한 방법은 공통적으로 다음 두 가지 이유로 시차 왜곡이 발생할 수 있다. 첫 번째로, 영상 합성의 기준선이 되는 Seam은 그림 5에서 나타나는 바와 같이 공통 영역을 두 영역으로 분할할 수 있는 연속(Continuous)의 곡선 - 그림 5의 주황색 픽셀로 구성된 선- 이라는 것이고, 두 번째로는 사물의 종류와 무관한 가중치 값을 Seam 생성 행렬에 반영한다는 것이다. 사물이 수평으로 넓게 배치된 영상을 스티칭하는 경우, 해당 영상에서 사물 영역에는 동일한 값의 가중치가 설정되기 때문에 해당 사물 영역을 피해 연속된 Seam을 생성하는 것이 제한된다.

결론적으로, 에너지 함수 또는 객체 세그먼트 방법으로 검출한 사물 영역이 수평으로 넓게 형성된 상황에서는, 사물 영역을 우회할 수 있고, 공통 영역을 수직 방향으로 분할 가능하며, 연속인 Seam을 생성할 수 없기 때문에 시차 왜곡이 발생할 수 있다. 일례로, 그림 2의 (a)를 에너지 함수와 객체 세그먼트 방법을 각각 적용하였을 때 모습은 그림 6의 (a)와 (b)로 나타나며, 이때 흰색 영역으로 나타난 사물 영

역을 지나지 않고 연속의 Seam을 생성하는 것은 불가능하다. 이와 같이 Seam이 사물 위로 생성될 수밖에 없는 상황에서는 Seam이 상대적으로 중요하지 않은 차선이나 가드레일 위치에서 생성되는 것이 차량 위치에서 Seam이 생성되는 것보다 만족스러운 스티칭 영상을 생성하겠지만, 기존 방법은 사물의 종류에 따라 다른 가중치를 적용하지 않기 때문에 임의의 지점에서 시차 왜곡이 발생하게 된다.

이와 같은 제한 사항으로 인해 시각적 인지 에너지 함수를 활용하는 방법과 객체 세그먼트를 활용하는 방법 모두 시차 왜곡이 발생할 수 있다. 객체 세그먼트를 활용하는 방법은 사물 검출기의 성능에 따라서 사물에 시차 왜곡이 발생할 수 있고, 배경 정보가 충분히 고려되지 못하여 배경에서 시차 왜곡이 발생할 수 있다는 단점이 존재했다. 반면, 시각적 인지 에너지 함수를 활용하는 방법은 사물의 내부 영역이 높은 가중치 값으로 설정되지 않기 때문에 사물의 내부에 Seam이 생성될 수 있다는 단점이 있지만, 객체 세그먼트를 활용하는 방법과 달리 배경 정보를 Seam 생성 행렬에 제대로 반영할 수 있기에, 사물의 내부 영역을 검출하

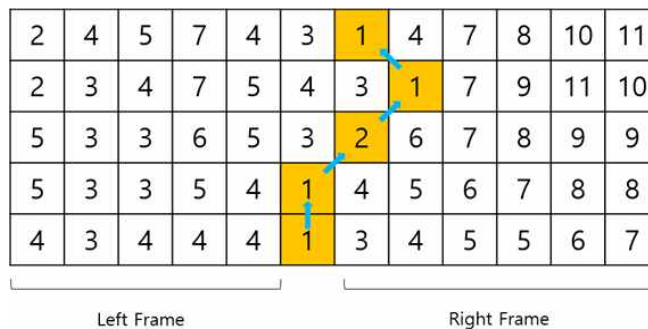


그림 5. Seam 생성 예시
Fig. 5. Example of seam generation



그림 6. 에너지 함수 (a)와 객체 세그먼트(b)에서 Seam 생성 제한 예시
 Fig. 6. Example of seam generation restriction in energy function (a) and object segment (b)

고 사물의 종류에 따라서 다른 가중치를 설정하는 것을 새로 추가함으로써, 기존의 시각적 인지 에너지 함수를 활용하는 방법과 객체 세그먼트를 활용하는 방법이 지닌 제한 사항을 극복할 수 있다는 장점이 있다.

이에 본 논문에서는 딥러닝 기반의 사물 검출을 활용하여 사물의 종류에 따라 다른 가중치 값을 설정하고, 이를 에너지 함수 기반 Seam 생성 행렬에 추가함으로써 기존 사물 반영 기술이 가진 제한 사항을 극복하는 영상 스티칭 방법을 이어지는 3장에서 제안하고자 한다.

III. 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭

앞서 설명한 바와 같이 본 논문의 3장에서는 딥러닝 기반의 사물 검출을 활용하여 사물의 종류에 따라 다른 가중치를 설정하고, 이를 Seam 생성 행렬에 반영하여 기존 사물 반영 기술이 가진 제한 사항을 극복하는 영상 스티칭 방법을 제안하고자 한다. 이에 따라서 본 장은 딥러닝 기반 사물 검출 기술을 간략하게 설명하며, 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭 방법을 설명한다.

딥러닝 기반 사물 검출은 딥러닝 네트워크를 통하여 영상 내 사물을 검출하는 기술로써, 사물이 존재하는 영역을 박스 형태로 출력하는 Object detection 기술과 사물의 위치 정보를 사물의 모양과 동일한 세그먼트로 제공하는 Instance segmentation 기술로 구분된다. 본 논문에서 정의하는 딥러닝 기반 사물 검출은 Mask R-CNN(Mask Region-

based Convolutional Neural Network)^[13]과 YOLACT (You Only Look At CoefficientTs)^[14]가 대표되는 Instance segmentation 기술을 의미하며, 일반 영상에 딥러닝 기반 사물 검출을 적용한 모습은 그림 7과 같이 나타난다.



그림 7. Instance segmentation 예시
 Fig. 7. Example of instance segmentation

딥러닝 기반의 사물 검출기는 딥러닝을 활용하지 않은 사물 검출기보다 뛰어난 검출 성능을 보이지만, 그림 8 (a)의 예시 영상에서 나타난 하단 좌측의 노란 박스에서 볼 수 있듯이 딥러닝 기반 사물 검출기를 통해 검출한 사물 영역은 실제 사물의 모든 영역을 나타내지 못할 수 있다. 이와 같이 완벽하지 않은 검출기의 성능을 보완하기 위하여 검출한 사물 영역의 박스 정보를 바탕으로 윤곽선 검출을 진행하여 그림 8의 (b)를 생성하고, 윤곽선 정보와 세그먼트 정보를 더해 그림 8의 (c)와 같이 생성한다. 마지막으

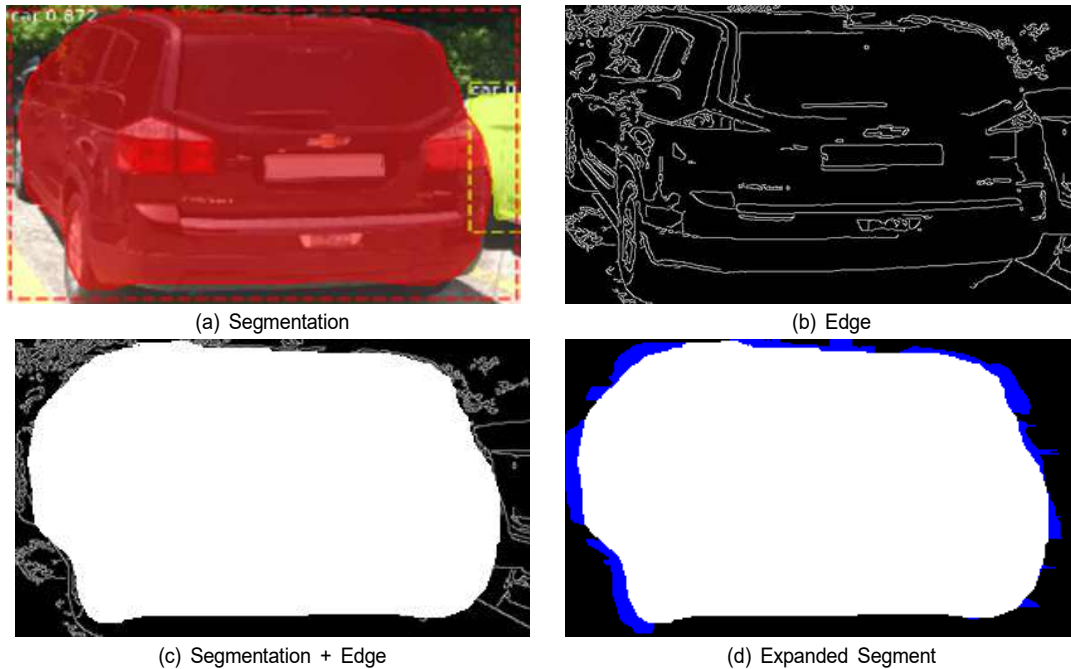


그림 8. 세그먼트 영역 확장 예시
 Fig. 8. Example of expanding the segment region

로, 그림 8의 (c)에서 나타난 세그먼트 정보를 중심으로 일정 거리 이내의 윤곽선까지 세그먼트 범위를 확장하여 그림 8의 (d)를 생성하고, 이를 Seam 생성 행렬에 적용하고자 한다. 이때 확장된 영역은 그림 8의 (d)에서 파란색 영역으로 나타난다.

한편, 그림 8과 같이 나타난 Instance segmentation 정보를 Seam 생성 행렬에 그대로 적용하는 것은 기존 객체 세그먼트 방법과 동일하기 때문에, 바닥과 벽면 등 배경 정보를 제대로 반영하지 못해 합성 영상의 배경 부분에서 시차 왜곡이 발생할 수 있다. 이를 극복하기 위하여 본 논문에서는 배경 정보를 반영할 수 있는 시각적 인지 에너지 함수 기반의 Seam 생성 행렬에 대하여 사전에 정의된 사물의 우선순위에 따라 다른 가중치를 더한 새로운 Seam 생성 행렬 구성 방법을 제안한다.

본 논문에서 제안하는 우선순위 사물 중심의 Seam 생성 행렬(Priority Object based Seam Matrix, POSM)은 식 (2)에서 나타나는 바와 같이 영상의 픽셀 좌표 (x,y) 에 대한 시각적 인지 에너지 함수 $E(x,y)$ 를 이용하기 때문에 바닥과 벽면 등의 배경 정보를 반영할 수 있으며, Instance seg-

mentation 정보를 사물의 우선순위(k_p)와 최대 에너지 값 (E_{max})에 따라 설정된 가중치를 더하므로 기존 에너지 함수 기반의 방법에서 초기 Seam 생성 위치가 사물 내부일 때 시차 왜곡이 발생할 수 있는 제한 사항을 해결할 수 있다.

$$POSM(x,y) = E(x,y) + (1+k_p)E_{max} \quad (2)$$

식 (2)를 구성하는 사물의 우선순위 k_p 는 공통 영역의 입력 영상에서 검출된 사물의 중요도에 따라 설정되는 상대 변수로써, 중요도가 가장 낮은 사물에는 0이 설정되고 중요도 순위에 따라 순차적으로 큰 정수 값이 설정된다. 이때 사물의 중요도는 지도 학습 기반 딥러닝 사물 검출에서 사전에 정의된 클래스들을 시차 왜곡으로부터 보존하고 싶은 순서에 따라 차등적으로 수치화한 것을 의미한다. 또한, 사물 영역 위로 Seam이 생성되는 것을 방지하기 위하여, 사물 영역의 가중치 값은 항상 배경의 에너지 값보다 크게 설정되도록 식 (3)과 같이 정의된 해당 영상의 최대 에너지 값(E_{max})을 식 (2)에서 나타난 사물의 우선순위를 고려한 에너지 값을 더해 POSM을 생성한다. 일례로, 중요도가 설

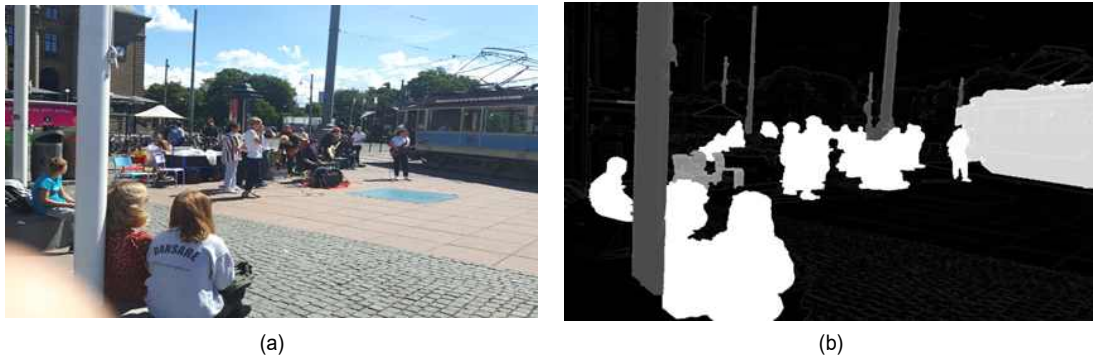


그림 9. 우선순위 사물 중심의 Seam Matrix 생성 예시
 Fig. 9. Example of priority object based seam matrix generation

정된 102개의 클래스를 학습한 딥러닝 기반 사물 검출을 그림 9의 (a)에 적용했을 때 해당 영상에서 검출된 사물과 중요도는 pole(중요도:21), chair(중요도:61), train(중요도:90), person(중요도:102)이었으며, 해당 영상에서 사물의 우선순위 k_p 는 pole(우선순위:0), chair(우선순위:1), train(우선순위:2), person(우선순위:3)으로 설정된다. 이를 식 (2)에 따라 POSM을 생성한 예시는 그림 9의 (b)와 같이 나타나며, 영상에서 흰색에 가까울수록 높은 가중치가 설정되었다.

분의 누적이 최소인 경로를 Seam으로 설정하고, Seam을 기준으로 영상을 합성함으로써, 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭 방법은 마무리된다.

이어지는 4장에서는 본 논문에서 제안하는 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심 영상 스티칭 방법을 테스트 영상에 적용하여 결과를 보임으로써, 제안 기술의 성능을 확인하고자 한다.

$$E_{\max} = \max(E(x,y)) \{x,y|x \in X, y \in Y\} \quad (3)$$

상기의 방법으로 시각적 인지 에너지 함수와 우선순위 사물 가중치를 더하여 생성한 POSM을 바탕으로 행렬 성

IV. 실험 결과

본 논문에서는 전통적인 Seam Optimization 기반 영상 스티칭에서 에너지 함수가 적용된 Seam 생성 행렬을 바탕



그림 10. 시차가 있는 테스트 영상
 Fig. 10. Test images with parallax



그림 11. 특징점 추출 및 매칭
Fig. 11. Feature extraction and matching

으로 딥러닝 기반 사물 검출 기술을 활용하여 사물의 종류에 따라 다른 가중치를 Seam 생성 행렬에 더해 POSM을 구성함으로써, Seam Optimization 기반의 영상 스티칭 기술이 가진 제한 사항을 극복하는 영상 스티칭 방법을 제안하고 있다. 이에 본 논문의 4장에서는 그림 10과 같이 시차가 크게 나타난 세 장의 입력 영상을 스티칭하는 실험을 통하여 제안 기술을 성능을 검증하고자 하며, 본 실험에서는 작은 사물을 세그먼트화할 수 있는 Mask R-CNN을 활용하였다. 또한, 기술 검증 및 성능 평가의 실험 환경은 Windows 10, i7-6700의 CPU, RTX-2080의 GPU와 Python 3.4, TensorFlow 1.3, Keras 2.0.8을 통해 설계되었다.

그림 10과 같이 시차가 큰 세 장의 입력 영상을 스티칭하기 위하여 각각의 영상에서 특징점을 추출하고 매칭하는 과정을 통하여 공통 영역을 설정하고 호모그래피를 추정하

였다. 이때 사용된 기술자는 ORB(Oriented FAST and Rotated BRIEF)^[14]이며, 해당 매칭 과정은 그림 11과 같이 나타난다.

본 논문의 3장에서 제안한 우선순위 사물 중심의 스티칭을 수행하기 위하여, 본 논문에서는 Mask R-CNN을 통하여 그림 12와 같이 사물 영역을 세그먼트화하고, 이와 같은 세그먼트 정보를 기반으로 세그먼트된 사물의 위치에 식 2와 3에서 정의한 바와 같이 최대 에너지 함수 값 E_{max} 을 추가하여 POSM을 생성하였다. 또한, 영상의 와핑 작업 이후에 사물을 검출하는 것은 검출 성능이 저하되어 제대로 된 세그먼트 정보를 추출하지 못할 수 있으므로, 와핑을 진행하기 전에 사물을 검출하고 POSM을 생성한 뒤에 와핑을 적용하였다. 1112×1080 크기의 공통 영역 영상에 대하여 Mask R-CNN을 통해 사물 검출을 진행하고 POSM을



그림 12. Instance segmentation 적용 예시
Fig. 12. Example of applying instance segmentation

생성하는데 걸리는 시간은 471ms이었으며, YOLACT를 통해 사물 검출을 진행하고 POSM을 생성하는데 걸리는 시간은 171ms이었다.

최종적으로 기존 시각적 인지 에너지 함수만을 활용하는 Seam Optimization 기반의 영상 스티칭 결과와 제안 기술을 적용한 영상 스티칭 결과는 그림 13과 같이 나타난다. 기존 기술^[8]은 어두운 영역에서 변화율로 정의되는 에너지 함수가 작은 값으로 나타나기 때문에, 검은색 차량 위로 Seam이 생성되어 시차 왜곡이 발생한 것을 그림 13-(a)의 표시된 영역에서 확인할 수 있다. 반면, 제안 기술에서 차량의 위치에는 큰 가중치가 추가로 더해지기 때문에, 스티칭 영상 내 어떤 차량에서도 시차 왜곡이 발생하지 않은 것을 그림 13의 (b)에서 확인할 수 있다.

앞선 그림 13을 통한 실험에서 픽셀 변화율로 정의되는 시각적 인지 에너지 함수만으로 Seam 생성 행렬을 구성하는 것은 영상 내 어두운 영역에서 가중치 값이 낮게 설정되기 때문에 시차 왜곡이 발생할 수 있음을 확인했다. 최근에는 이와 같은 시각적 인지 에너지 함수의 제한 사항을 극복하고자 영상의 평균 밝기를 계산하고, 평균 밝기와 일정 크기 이상의 차이를 갖는 픽셀의 인접 단면에 큰 가중치를 할당한 Prominent map을 시각적 인지 에너지 함수에 더하는 방법^[9]이 제안되었기에, 해당 방법과 제안 기술을 테스트 영상을 통해서 비교 실험하고자 한다.

그림 14의 (a)와 (b)는 공통 영역의 입력 영상이며, 입력 영상의 촬영 지점 간 거리 차이가 크기 때문에 빨간색 영역과 초록색 영역의 영상이 크게 다른 것을 확인할 수 있다. 먼저, 시각적 인지 에너지 함수에 Prominent map을 적용한 기술^[9]의 스티칭 영상에서는 그림 14의 (c)에서 볼 수 있듯이 표지판이 사라지게 합성되었지만, 평균 밝기와 비슷한 픽셀에 대하여 가중치가 적용되지 않는다는 특성으로 인해, 평균 밝기와 비슷한 그림 14-(b)의 흰색 영역으로 표시된 기둥 부분이 매우 작은 가중치로 설정되었고, 이에 따라서 Seam이 해당 지점을 지나면서 생성되어 시차 왜곡이 발생했다. 반면, 제안 기술을 적용한 스티칭 영상에서는 사물의 종류에 따라서 개별적인 가중치가 설정되었기 때문에 시차 왜곡이 발생하지 않은 것을 그림 14의 (d)에서 확인할 수 있다.

그림 15의 (a)와 (b)는 비교 방법 [8], [9]와 제안 방법의 영상 스티칭 결과를 비교하기 위하여 설정된 테스트 영상이다. 해당 테스트 영상은 각 방법의 결과 영상에서 시차 왜곡이 발생할 확률을 높이기 위하여 촬영 시간과 촬영 시점이 모두 다르게 촬영되었으며, 전체 영상 크기에 비하여 현저하게 작은 공통 영역이 설정되었고, 공통 영역의 중앙에는 기둥이 위치하여 Seam이 사물을 우회하며 생성되는 것을 어렵게 하였다. 이와 같이 시차 왜곡이 발생하기 쉬운 입력 영상을 바탕으로 비교 방법 [8], [9]와 제안 기술의 영상 스티칭 결과는 다음과 같다.

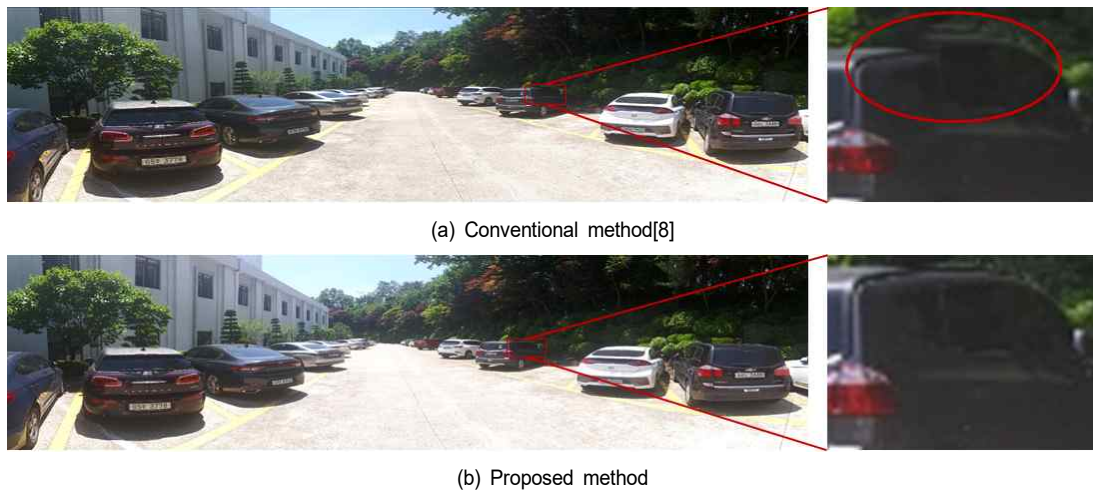


그림 13. 기존 기술과 제안 기술 비교
 Fig. 13. Comparison of conventional method and proposed method

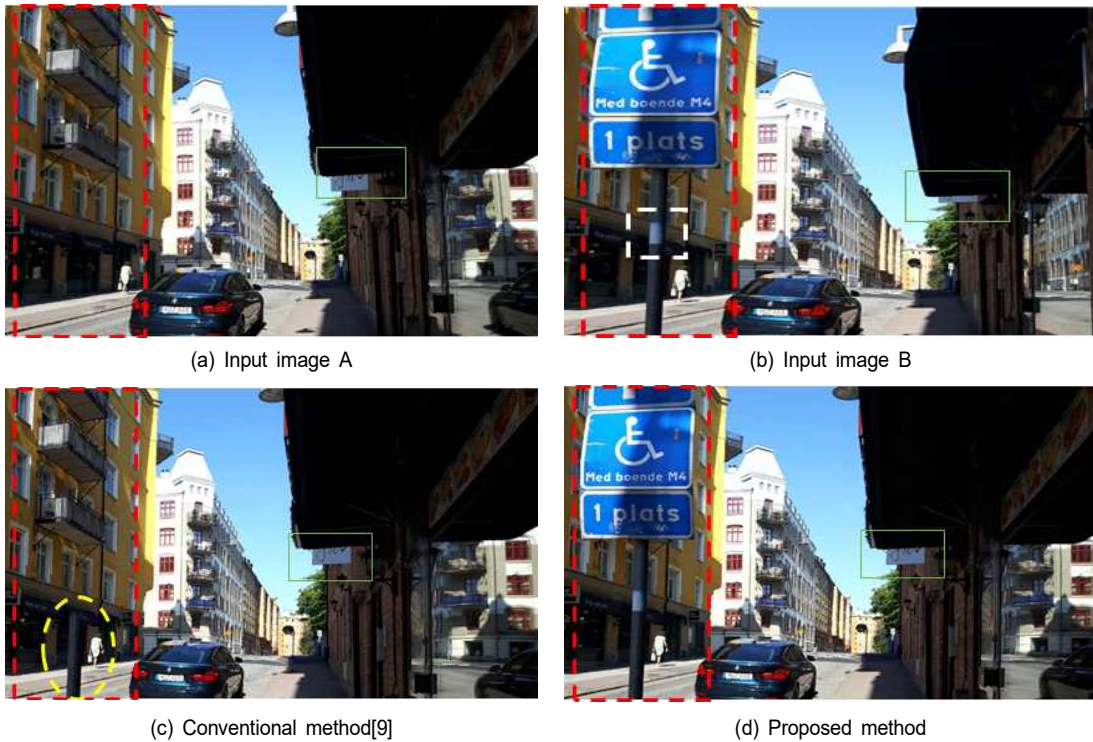


그림 14. 제안 기술 적용 예시
 Fig. 14. Example of applying the proposed method



그림 15. 영상 스티칭 비교 실험의 입력 데이터
 Fig. 15. Input data of image stitching comparison experiment

그림 16의 (a), (b), (c)는 앞선 그림 15의 (a)와 (b)를 통하여 비교 방법 [8], [9]과 제안 기술을 통해 스티칭한 각각의 결과 영상이며, 자세한 비교를 위해 공통 영역이 포함되는 ROI(Region of Interest) 영역만을 확대하여 나타내었다. 먼저, 시각적 인지 에너지 함수만을 사용한 비교 기술 [8]은

픽셀 변화율이 적은 부분에서 에너지 값이 작게 형성되기에, 주위 밝기와 비슷한 기둥 부분에서 에너지 값이 작게 형성되었고, 이에 따라서 기둥이 잘리게 나타나는 시차 왜곡이 발생한 것을 그림 16의 (a)에서 확인할 수 있다.

한편, 시각적 인지 에너지 함수에 평균 밝기와 일정 크

기 이상 차이나는 픽셀 값에 가중치를 적용한 비교 방법 [9]는 기둥 영역이 평균 밝기와 차이가 크기 때문에, 기둥 영역에 큰 가중치가 적용되었고, 이에 따라서 시각적 인지 에너지 함수만을 사용한 방법[8]과 달리 기둥에서 시차 왜곡이 발생하지 않은 것을 그림 16의 (b)에서 확인할 수 있다. 그러나 공통 영역 전체에 수평으로 형성된 건물들을 피해 Seam을 생성할 수 없기 때문에 건물의 합성 부분에서 시차 왜곡이 발생한 것을 확인할 수 있다.

마지막으로, 딥러닝을 통해 사물을 검출하고 사물의 종류에 따라서 개별적인 가중치를 설정하는 제안 기술의 경우에는 검출된 기둥에 높은 가중치를 적용하였기 때문에, 앞선 방법[9]와 마찬가지로 기둥에서 시차 왜곡이 발생하지 않은 것을 그림 16의 (c)에서 확인할 수 있다. 그러나 비교 방법 [9]와 마찬가지로 공통 영역 전체에 수평으로 형성된 건물들을 피해 Seam을 생성할 수 없기 때문에 건물의 합성 부분에서 시차 왜곡이 발생한 것을 확인할 수 있다.



그림 16. 영상 스티칭 기술 비교
 Fig. 16. Comparison of image stitching

한편, 비교 방법 [9]는 앞선 그림 14의 테스트에서 설명하였듯이, 공통 영역의 평균 밝기를 계산하고 평균 밝기와 일정 크기 이상 차이나는 픽셀에 대하여 높은 가중치를 적용하는 방법이므로, 평균 밝기와 유사한 픽셀 영역은 고려하지 못한다는 제한 사항이 있다. 그림 15의 (a)와 (b)에서 나타난 기둥을 공통 영역의 평균 밝기와 비슷하게 설정한 그림 17의 (a)와 (b)를 입력 영상으로, 비교 방법 [9]를 통해

스티칭한 결과는 기둥에 대하여 가중치 설정이 제한되기 때문에, 그림 17의 (c)에서 나타나는 바와 같이 기둥에서 시차 왜곡이 발생하는 것을 확인할 수 있다. 반면, 영상의 밝기 변화는 사물 검출에 영향이 적기 때문에 우선순위 가중치 설정에 영향이 없었고, 이에 따라서 제안 기술에서는 기둥에서 시차 왜곡이 발생하지 않은 것을 그림 17의 (d)에서 확인할 수 있다.



그림 17. 비교 방법 [9]의 제한 사항
 Fig. 17. Limitation of the comparison method [9]



그림 18. 제안 기술의 제한 사항
Fig. 18. Limitation of the proposed method

앞선 실험들을 통해서 제안 기술이 기존 시각적 인지 에너지 함수를 활용하는 방법에 적용되었을 때 보다 시차 왜곡이 적은 스티칭 영상을 생성할 수 있음을 확인할 수 있다. 그러나 제안 기술이 딥러닝 기반의 사물 검출 기술을 활용하기 때문에, 영상 내에서 사물을 검출하지 못하는 경우에는 시차 왜곡이 발생할 수 있다는 제한 사항이 있다. 이는 그림 15의 (a)와 (b)에 대하여 기둥에 대한 클래스를 제거하였을 때 스티칭 영상의 기둥 부분에서 시차 왜곡이 발생하는 것을 그림 18을 통해 확인할 수 있다.

이와 같이 딥러닝 기반 사물 검출 기술에서 영상 내 사물을 검출하지 못하는 경우, 제안 기술에서 시차 왜곡이 발생할 수 있다는 제한 사항이 있지만, 이는 지도 학습을 통한 딥러닝 사물 검출에서 검출하지 못하는 사물의 클래스를 새로 정의하고, 해당 클래스에 대하여 학습함으로써 해결할 수 있다. 이에 충분한 클래스를 학습한 딥러닝 기반 사물 검출 기술을 활용할 경우, 본 논문에서 제안하는 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭 방법은 기존 시각적 인지 에너지 함수 기반의 Seam Optimization 기술에 적용될 때 보다 좋은 품질의 스티칭 영상을 획득할 수 있을 것으로 기대된다.

V. 결론

기존 영상 스티칭 기술이 가진 시차 왜곡 문제를 극복하기 위하여 Seam Optimization 기반 영상 스티칭 방법은 에

너지 함수를 이용하거나 객체 세그먼트를 적용하여 사물의 위치 정보를 Seam 생성 행렬에 반영하였다. 그러나 에너지 함수를 활용한 방법은 배경과 비슷한 사물에 대하여 가중치 설정이 제한되었으며, Seam 생성 초기 위치에 따라서 시차 왜곡이 발생할 수 있었고, 객체 세그먼트를 활용한 방법은 배경 정보를 제대로 반영하지 못하고 검출기의 성능에 따라 시차 왜곡이 발생할 수 있었다. 또한, 두 방법 모두 수평으로 넓게 사물 영역이 형성된 영상에서는 올바른 Seam 생성이 제한된다는 문제가 있었다.

이에 본 논문에서 제안하는 딥러닝 기반 사물 검출을 활용한 우선순위 사물 중심의 영상 스티칭에서는 기존 시각적 인지 에너지 함수 값을 바탕으로, 딥러닝 기반 Instance segmentation을 통해 검출한 사물의 종류에 따라 다르게 설정된 가중치를 더하여, 기존 Seam 생성 행렬의 제한 사항을 극복하였으며, 딥러닝 기반 Instance segmentation 기술의 부족한 검출 성능을 보완하기 위하여 검출 박스 영역 내에서 윤곽선 정보를 추출하고, 사물이 온전히 반영될 수 있도록 사물의 세그먼트 영역을 확장하였다. 또한, 제안 기술을 테스트 영상에 적용하여 시차 왜곡이 제거되는 것을 앞선 실험 결과를 통해 확인함으로써, 제안 기술이 기존 시각적 인지 에너지 함수를 활용하는 Seam Optimization 기술에 적용했을 때 보다 좋은 품질의 스티칭 영상을 획득할 수 있음을 검증하였다.

추가로, 딥러닝 기반 사물 검출을 적용하여 사물의 종류에 따라 다른 가중치를 설정하여도 올바른 Seam을 생성할 수 없는 상황, 즉, 영상 내 동일한 사물이 수평으로 형성되

어 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.

참 고 문 헌 (References)

- [1] Im, Jiheon, et al. "Images Grouping Technology based on Camera Sensors for Efficient Stitching of Multiple Images." *Journal of Broadcast Engineering*, Vol.22, No.6, pp.713-723, 2017.
- [2] R. Szeliski, "Image Alignment and Stitching: A Tutorial." *Foundations and Trends in Computer Graphics and Computer Vision*, Vol. 2, No.1, 2006.
- [3] Rhee, Seongbae, Jeonho Kang, and Kyuheon Kim. "Local Differential Pixel Assessment Method for Image Stitching." *Journal of Broadcast Engineering*, Vol.24, No.5, pp.775-784, 2019
- [4] 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.
- [5] Zhang, Guofeng, et al. "Multi-viewpoint panorama construction with wide-baseline images." *IEEE Transactions on Image Processing*, Vol.25, No.7, pp.3099-3111, 2016.
- [6] Lin, Kaimo, et al. "Seagull: Seam-guided local alignment for parallax-tolerant image stitching." *European conference on computer vision*. Springer, Cham, 2016.
- [7] Avidan, Shai, and Ariel Shamir. "Seam carving for content-aware image resizing." *ACM SIGGRAPH 2007 papers*. pp.10-es, 2007.
- [8] TANG, Yu; JIANG, Huiyan. "Highly efficient image stitching based on energy map". In: 2009 2nd International Congress on Image and Signal Processing IEEE, pp.1-5, 2009.
- [9] TANG, Yu; SHIN, Jungpil. "Image stitching with efficient brightness fusion and automatic content awareness.". *International Conference on Signal Processing and Multimedia Applications (SIGMAP)*, pp.60-66, 2014.
- [10] Abdukhlikov, Murodjon, and Taegkeun Whangbo. "Fast image stitching method for handling dynamic object problems in Panoramic Images." *KSII Transactions on Internet & Information Systems*, Vol.11, No.11, 2017.
- [11] Kang, Jeonho, et al. "Minimum Error Seam-Based Efficient Panorama Video Stitching Method Robust to Parallax." *IEEE Access* 7, pp. 167127-167140, 2019.
- [12] Rhee, Seongbae, Jeonho Kang, and Kyuheon Kim. "Parallax Distortion Detection and Correction Method for Video Stitching by using LDPM Image Assessment." *Journal of Broadcast Engineering*, Vol.25, No.5, pp.685-697, 2020.
- [13] He, Kaiming, et al. "Mask r-cnn." *Proceedings of the IEEE international conference on computer vision*, pp.2961-2969, 2017.
- [14] Bolya, Daniel, et al. "Yolact: Real-time instance segmentation." *Proceedings of the IEEE international conference on computer vision*, pp.9157-9166, 2019.
- [15] Rublee, Ethan, et al. "ORB: An efficient alternative to SIFT or SURF." *2011 International conference on computer vision*. IEEE, pp.2564-2571, 2011.

저 자 소 개



이 성 배

- 2019년 2월 : 경희대학교 전자공학과 공학사
- 2019년 3월 ~ 현재 : 경희대학교 전자공학과 석사과정
- ORCID : <https://orcid.org/0000-0002-7392-6870>
- 주관심분야 : 딥러닝, 영상처리, 멀티미디어 시스템



강 전 호

- 2010년 2월 : 경희대학교 전자공학과 공학사
- 2012년 2월 : 경희대학교 전자공학과 공학석사
- 2012년 3월 ~ 2015년 2월 : 주식회사 넥스트리밍 연구원
- 2015년 3월 ~ 현재 : 경희대학교 전자공학과 박사과정
- ORCID : <http://orcid.org/0000-0003-4405-3879>
- 주관심분야 : 영상처리, 멀티미디어 시스템, 디지털 대화형방송

저 자 소 개



김 규 현

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- ORCID : <http://orcid.org/0000-0003-1553-936X>
- 주관심분야 : 디지털 방송, 영상처리, 멀티미디어 통신, 디지털 대화형 방송