

## MegaDepth Network를 활용한 깊이 기반 영상 스티칭

김가현, 장혜민, 최유진, 이성배, \*김규헌  
경희대학교

kahyunapril@khu.ac.kr, wkdgpals9@khu.ac.kr, zenith0410@khu.ac.kr,  
rhee@khu.ac.kr, \*kyuheonkim@khu.ac.kr

## Depth-based Image Stitching Using MegaDepth Network

Kahyun Kim, Hyemin Jang, Yujin Choi, Seongbae Rhee, \*Kyuheon Kim  
Kyunghee University

## 요약

영상 스티칭은 다수의 영상을 넓은 시야각을 갖는 하나의 영상으로 합성하여 사용자들에게 몰입감과 현장감을 제공하는 기술이다. 그러나 영상에 시차(Parallax)가 존재하는 경우 스티칭된 영상에서 왜곡이 발생할 수 있는데 이는 사용자의 몰입을 방해할 수 있다. 따라서 스티칭 영상의 다양한 활용을 위해서는 시차로 인한 왜곡을 최소화하여 자연스러운 스티칭 영상을 만드는 것이 중요하다. 기존 호모그래피 추정 방법으로 발생할 수 있는 고스트 현상을 최소화하기 위해서 seam 기반 스티칭 방법이 사용되었지만, 단순히 작은 특징값을 따라 생성된 seam은 사물 영역 정보가 반영되지 않아 seam이 특징이 있는 부분을 지나가면서 시차 왜곡이 발생할 수 있다. 이에 본 논문에서는 딥러닝 기반의 MegaDepth를 활용한 depth 예측 정보를 에너지 함수 기반의 seam 생성 행렬의 가중치로 사용하여 seam이 사물을 피해 생성되면서 시차가 작은 영역으로 유도되도록 하는 seam optimization 기법을 제안한다.

## 1. 서론

영상 스티칭 기술이란 다수의 영상을 넓은 시야각을 갖는 하나의 영상으로 합성하는 기술이다[1]. 스티칭된 영상은 넓은 시야각을 통해 사용자들에게 몰입감과 현장감을 제공할 수 있기에, VR(Virtual Reality), 파노라마 영상 등 다양한 몰입형 미디어 콘텐츠에 활용된다. 그러나 다수의 영상을 합성할 때 영상에 시차(Parallax)가 존재하는 경우 스티칭 영상에서 왜곡이 나타날 수 있으며[2], 이는 사용자의 몰입을 방해할 수 있다. 따라서 스티칭 영상의 다양한 활용을 위해서는 시차로 인한 왜곡을 최소화하여 자연스러운 스티칭 영상을 만드는 것이 중요하다.

일반적으로 영상 스티칭은 특징점 추출, 특징점 매칭, 호모그래피 추정, 왜핑, 합성, 블렌딩까지를 지칭한다[1]. 매칭된 특징점이 가장 많은 영역의 시차를 기준으로 호모그래피를 추정하기 때문에 여러 시차가 존재하는 영상의 경우에는 고스트 현상[3]이 발생할 수 있다. 이러한 왜곡 현상을 최소화하기 위해서 seam을 기반으로 한 스티칭 방법들이 제안되었다.

Seam 기반 스티칭 방법은 공통 영역 내 특징이 있는 부분을 피해 seam을 생성하여 영상을 스티칭하는 방법으로서, 공통 영역 내 특징값이 반영된 행렬의 최소 누적 합을 따라 seam을 생성한다. 단순히 작은 특징 값을 따라 생성된 seam은 사물 영역 정보가 반영되지 않아 사물이 있는 부분을 지나가면서 시차 왜곡이 발생할 수 있다[4]. Seam 생성 시 사물 영역을 지나가지 않기 위해 다양한 seam optimization 기술[4,5,6]이 제안되었지만, seam이 특징이 있는 부분을 지나갈 수밖에 없는 상황[7]에서는 시차 왜곡이 발생한다는 한계점이 있다.

Seam optimization 기술에는 seam 생성 행렬에 사물 영역 정보

를 반영하여 사물이 있는 부분을 최대한 피해 seam을 생성하는 방법이 있다. 사물 영역을 반영하기 위해서 object detection[8]을 활용하기도 하지만, 사물 검출기의 성능에 따라 시차 왜곡이 발생할 수 있다는 제한 사항이 있다. 또한, depth 정보를 활용하여 사물 영역을 반영하는 방법도 있다[1]. Depth 정보는 LiDAR, TcF, 구조광에 기반한 센서를 활용하거나, 스테레오 비전을 통해서 생성된다. 고가의 장비나 오랜 연산 시간이 요구된다는 제한 사항이 있지만, depth 정보를 이용하면 시차가 작은 영역을 영상에서 구분할 수 있다는 장점이 있다.

이에 본 논문에서는 depth 정보 생성의 제한 사항을 극복하기 위하여 딥러닝 기반의 depth 생성 네트워크를 활용하고, seam 생성 행렬에 가중치 값으로 적용함으로써 인지적 시차 왜곡을 최소화할 수 있는 seam optimization 방법을 제안하고자 한다. 본 논문 2장에서는 제안 기술의 배경 기술을 간단하게 분석하고, 3장에서는 딥러닝을 통해 추출한 depth 정보로 seam 생성 행렬에 가중치를 더하는 방법을 설명한다. 4장에서는 테스트 영상을 통해 기존 기술의 결과와 제안 기술의 결과를 비교하여 제안 기술의 효용성을 검증하고, 5장에서는 결론을 내린다.

## 2. 배경 기술 분석

앞서 설명한 바와 같이 영상 스티칭은 특징점 추출, 특징점 매칭, 호모그래피 추정, 왜핑, 합성, 블렌딩까지의 과정을 지칭한다[1]. 먼저 영상에서 이동, 회전, 확대 및 축소 등의 변환 과정에도 동일한 지점으로 나타나는 강인한 특징점(Keypoint)을 추출하고, 각각의 영상에서 동일한 특징점 쌍을 찾는 매칭 작업을 한다. 그리고 매칭된 특징점들을 통해서 호모그래피를 추정하며 왜핑을 진행한다. 왜핑을 통해 동일한 평면으

로 정렬된 다수의 영상에서 영상 간 차이가 최소인 경로(Seam)를 따라서 하나의 영상으로 합성하고, 마지막으로 블렌딩을 통해서 영상의 밝기나 모양을 보정한다.

이때 영상을 합성하기 위해 사용되는 seam은 합성할 영상 간 차이가 최소인 경로로서, 주로 두 영상 간 픽셀 차이나 시각적 인지 에너지 함수[9,10]를 통해서 정의된 seam 생성 행렬에서 누적 합이 최소인 경로로 정의된다. 그러나 이와 같은 방법들은 사물 영역을 제대로 반영하지 못하기 때문에 시차 왜곡이 발생할 수 있다. 일례로, 시각적 인지 에너지 함수의 경우 영상의 X축과 Y축에 대하여 좌표 (x, y)에서 식 1과 같이 나타나는데 사물과 배경의 저주파 성분에서는 작은 에너지로 표현되고, 윤곽 부분과 같은 고주파 성분에서 큰 에너지로 표현되기 때문에 사물 내부를 반영하는데 한계점이 있다.

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

이와 같이 사물 내부 영역을 제대로 반영하지 못하는 방법으로 seam 생성 행렬을 구성하게 된다면, seam이 사물 내부를 지나가며 생성되어 시차 왜곡이 발생할 수 있다[4]. 따라서 이를 방지하기 위해서는 seam 생성 행렬에 사물 영역을 가중치로 반영할 필요가 있다. 이에 본 논문에서는 사물 영역을 반영하기 위하여 딥러닝 네트워크를 통해서 depth 정보를 추출하고, depth 정보를 통하여 사물 영역을 반영하여 인지적 시차 왜곡을 줄일 수 있는 seam optimization 기술을 이어지는 3장에서 설명하고자 한다.

### 3. MegaDepth Network기반 Depth 가중치를 활용한 Seam Optimization

앞서 설명한 바와 같이 사물 내부 영역을 반영하지 않은 seam 생성 행렬을 통해 영상을 합성한다면, seam이 사물 내부 영역을 따라서 생성될 수 있고 이때 시차 왜곡이 발생할 수 있다는 제한 사항이 있었다. 또한, 영상 내 사물이 수평으로 넓게 펼쳐진 경우라면 seam이 사물을 지나갈 수밖에 없기 때문에 시차 왜곡이 불가피하다. 이에 본 논문에서는 딥러닝을 통해 추출한 depth 정보를 에너지 함수 기반의 seam 생성 행렬의 가중치로 사용함으로써, seam이 사물을 피해 생성되면서 시차가 작은 영역으로 유도되도록 하는 seam optimization 기법을 제안한다.

이에 제안 기술의 구조도는 그림 1과 같이 두 개의 영상 입력, 딥러닝 기반 depth 예측 정보 추출, 공통 영역 계산 모듈, 시각적 인지 에너지 함수 기반의 seam 생성 행렬 추출, depth 정보 가중치를 이용한 seam optimization 모듈, seam 생성 및 영상 합성으로 구성된다.

먼저 두 입력 영상의 depth 예측 정보를 MegaDepth[11] network를 통해 추출하고, 공통 영역 계산 모듈에서는 두 입력 영상의 특징점 추출 및 매칭 작업을 통해 공통영역을 찾는다. 구해진 공통영역에 대해 시각적 에너지 함수식을 통해 seam 생성 행렬을 만든다. 다음으로 seam optimization 모듈에서는 depth 정보를 활용하여 seam 생성 행렬에 가중치를 반영한다. 이때 시차는 촬영 지점으로부터 사물까지의 거리가 멀어질수록 작아지기 때문에, 촬영 지점으로부터 가까운 지점에는 큰 가중치를 적용하고 먼 지점에는 작은 가중치를 적용함으로써, 시차가 작은 영역에서 seam이 생성되도록 유도한다. 따라서 본 논문에서는

Deep/Shallow Depth Region Separation Submodule과 Shallow Depth Region Quantization Submodule을 구성하였으며, 각각의 설명은 3.1절과 3.2절에서 자세히 설명한다. 최종적으로 가중치가 더해진 seam 생성 행렬에서 누적 합이 최소인 경로를 따라 영상을 합성한다.

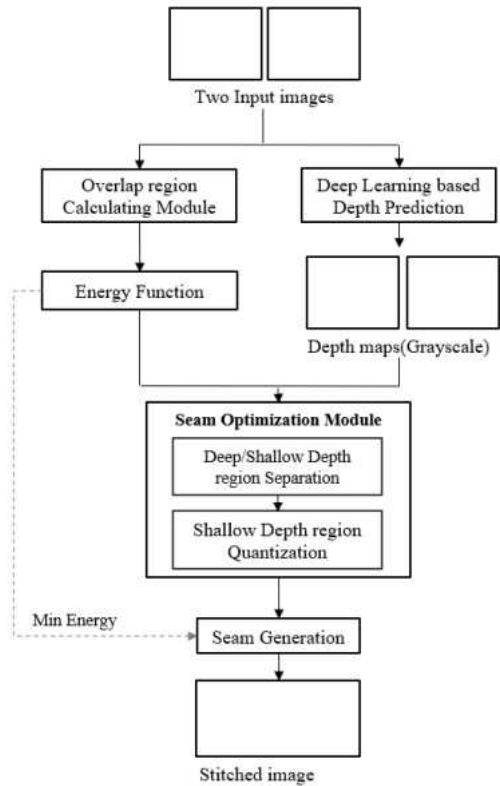


그림 1. Depth 가중치를 활용한 Seam Optimization 구조도

#### 3.1 Deep/Shallow Depth Region Separation Submodule

Deep/Shallow Depth Region Separation Submodule은 depth를 깊은 영역과 얇은 영역으로 구분하는 기능을 한다. 그림 2와 같이 depth가 가장 깊은 최솟값부터 depth가 가장 얇은 최댓값을 백분율로 나타내고, P% 지점을 구분 경계로 설정하며 기준 이하인 지점을 deep region, 기준 이상인 지점을 shallow region으로 설정한다.

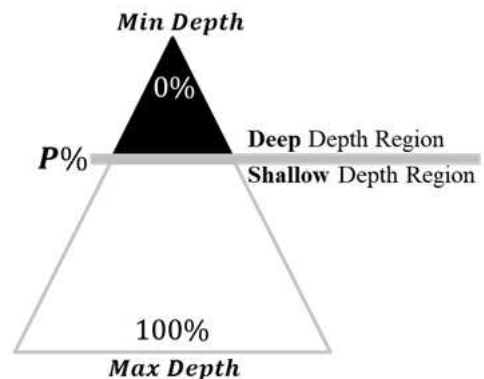


그림 2. Deep/Shallow Depth Region Separation 도식

이에, 시차가 큰 shallow region에서 seam이 생성되는 것을 방지하기 위하여 해당 지점의 에너지 값은 식 2와 같이 정의하여 가중치를 둔다. 이때  $E_{max}$ 는 공통 영역에서 계산한 시각적 인지 에너지 함수의 최댓값에 해당하며,  $E$ 는 현재 픽셀에서의 시각적 인지 에너지 함수값이다.

$$E_p = E + 2E_{max} \quad (2)$$

### 3.2 Shallow Depth Region Quantization Submodule

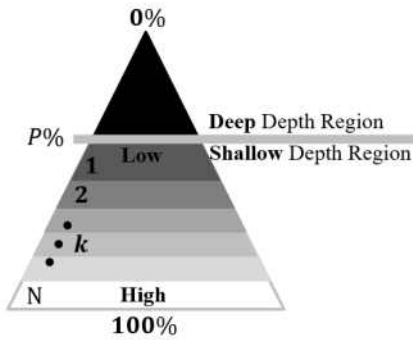


그림 3. Shallow Depth Region Quantization 도식

Shallow Depth Region Quantization Submodule은 3.1절에서 정의한 shallow region에 대하여 depth에 따른 차등적인 가중치를 설정하는 기능을 한다. 차등적인 가중치는 그림 3에서 나타나는 바와 같이 shallow region을 depth 백분율 값을 기준으로 N개로 양자화하고 영역의 순서에 따라 가중치를 차등 부여한다. 이때, k 번째 영역에 부여되는 가중치 값은 아래 식 3과 같이 나타난다.

$$E_k = E_p + \frac{E_{max}}{N}k \quad (\text{for } k = 1, 2, 3, \dots, N) \quad (3)$$

이어지는 4장에서는 본 논문의 3장에서 제안한 MegaDepth Network 기반 Depth 가중치를 활용한 seam optimization 방법을 테스트 영상에 적용하여 기존 기술과 비교함으로써, 제안 기술의 효용성을 검증하고자 한다.

## 4. 실험 결과

본 논문에서는 딥러닝 기반의 depth 정보를 seam 생성 행렬의 가중치로 사용하는 seam optimization 기법을 제안한다. 본 논문의 4장에서는 그림 4와 같이 시차가 존재하는 두 장의 입력 영상을 스티칭하는 실험을 통해 기존 스티칭 방법과 제안 하는 스티칭 방법의 성능을 비교하고자 한다. 성능 평가를 위한 실험은 Windows 10, i7-8565U CPU와 Visual Studio, Google Colaboratory를 통해 설계되었다.

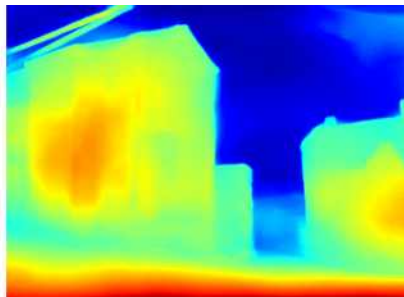


그림 4. 두 장의 실험 입력 영상

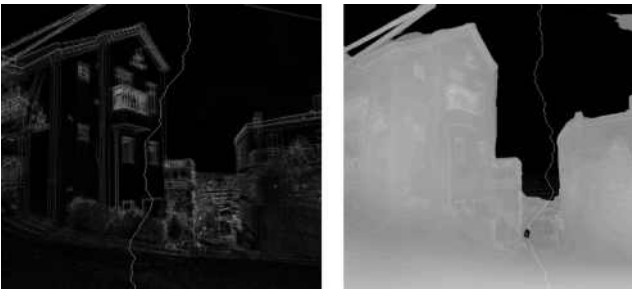
그림 4와 같이 시차가 존재하는 두 장의 입력 영상을 스티칭하기 위하여 각각의 영상에서 특징점 추출 및 매칭을 통해 공통영역을 설정하고 공통 영역에 대하여 에너지 함수를 계산하였으며, 결과는 그림 5의 (a)와 같이 나타난다. MegaDepth를 이용한 depth 예측 정보는 그림 5의 (b)와 같이 추출되며, seam optimization 모듈을 통해 가중치를 반영한 seam 생성 행렬은 그림 5의 (c)와 같이 나타난다.

시각적 에너지 함수만을 활용한 기존 방법[10]의 seam 생성 결과는 그림 6의 (a)와 같이 나타나며, 영상 스티칭 결과는 그림 7의 (a)와 같이 시차 왜곡이 발생한 것을 확인할 수 있다. 반면, 제안 기술의 경우 seam 생성 결과는 그림 6의 (b)와 같이 나타나며, 영상 스티칭 결과는 그림 7의 (b)와 같이 시차 왜곡이 발생하지 않은 것을 확인할 수 있다.

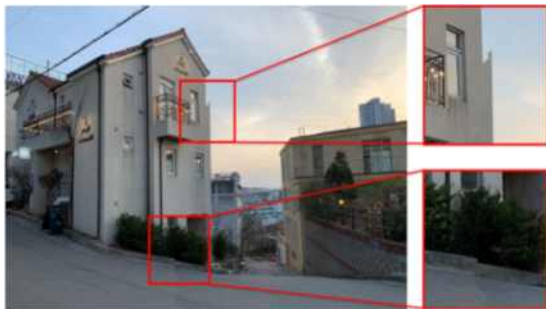
이와 같은 실험을 통해 본 논문에서 제안하는 depth 정보 가중치를 활용한 영상 스티칭 방법으로 시차 왜곡을 최소화한 스티칭 영상을 얻을 수 있음을 확인하였다.



(a) 시각적 에너지 함수 기반 seam 생성 행렬 (b) MegaDepth Network로 추출한 depth (c) depth 가중치가 적용된 seam 생성 행렬  
그림 5. 두 입력 영상의 공통 영역에 대한 결과 영상 예시



(a) 기존 기술 seam 생성 결과 (b) 제안 기술 seam 생성 결과  
 그림 6. 기존 기술과 제안 기술의 seam 생성 결과 비교



(a) 기존 기술 영상 스티칭 결과



(b) 제안 기술 영상 스티칭 결과

그림 7. 기존 기술과 제안 기술의 영상 스티칭 결과 비교

## 5. 결론

본 논문에서는 딥러닝 기반의 depth 정보를 seam 생성 행렬의 가중치로 사용하는 seam optimization 기법을 제안하였다. 기존 seam optimization 기반의 스티칭 방법은 시차 왜곡 문제를 해결하기 위하여 일반적으로 사물 영역을 우회하여 seam을 생성하였지만, seam의 초기 생성 위치, 사물이 수평으로 형성되어 seam이 물체를 통과할 수밖에 없는 상황 등으로 인해 올바른 seam을 생성하는 데 제한 사항이 있었다.

이러한 제한 사항을 극복하기 위하여, 본 논문에서는 MegaDepth로 depth 정보를 예측하였고, Deep/Shallow Depth Region Separation Submodule과 Shallow Depth Region Quantization Submodule을 통해 가중치를 생성하여 seam 생성 행렬에 반영함으로써, seam이 시차가 작은 깊은 영역으로 생성되어 시차 왜곡을 최소화할 수 있다. 이와 같이 시차가 작아 비교적 눈에 띄지 않는 깊은 영역으로 seam이 통과하게끔 함으로써 기존 seam optimization 기반의 영상 스티칭 기술이 가진 제한 사항을 극복할 수 있다.

추가적으로 시차가 작은 깊은 영역으로 seam을 유도할 뿐만 아니라, 사물의 경계에 가깝게 seam이 생성되도록 하고 Deep/shallow

depth region 경계를 구분하는 변수  $P$ 를 영상의 깊이 특성과 무관하게 설정할 수 있다면 보다 개선된 스티칭 기술을 제공할 수 있을 것으로 기대한다.

## 참고문헌

- [1] R. Szeliski. "Image Alignment and Stitching: A Tutorial." Foundations and Trends in Computer Graphics and Computer Vision, Vol.2, No.1, 2006.
- [2] Wei, L. Y. U., et al. "A survey on image and video stitching." in Virtual Reality & Intelligent Hardware, Vol.1, No.1, pp.55-83, 2019.
- [3] Fan Zhang, Feng Liu. "Parallax-tolerant Image Stitching" in IEEE conference on computer vision and pattern recognition, 2014.
- [4] Rhee Seongbae, et al. "Image Stitching focused on Priority Object using Deep Learning based Object Detection" in Journal of Broadcast Engineering, Vol.25, No.6, pp.882-897, 2020.
- [5] Kang Jeonho, et al. "Minimum Error Seam-Based Efficient Panorama Video Stitching Method Robust to Parallax." in IEEE Access 7 (2019): 167127-167140.
- [6] Lin Kaimo, et al. "Seagull: Seam-guided local alignment for parallax-tolerant image stitching." in European conference on computer vision, pp.370-385, 2016.
- [7] Rhee Seongbae, et al. "Parallax Distortion Detection and Correction Method for Video Stitching by using LDPM Image Assessment." in Journal of Broadcast Engineering, Vol.25, No.5, pp.685-697, 2020.
- [8] Charles Herrmann, et al. "Object-centered image stitching" in European conference on computer vision, pp.821-835, 2018.
- [9] Avidan, Shai, and Ariel Shamir. "Seam carving for content-aware image resizing." ACM SIGGRAPH 2007 papers. pp.10-es, 2007.
- [10] 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.
- [11] Zhengqi Li, Noah Snavely. "MegaDepth: Learning Single-View Depth Prediction from Internet Photos" in IEEE conference on computer vision and pattern recognition, 2018