

딥러닝 기반 특징점 추출 및 매칭 기술을 활용한 밝기 적응형 영상 스티칭

김민영, 김건호, 이민석, 이성배 *김규현

경희대학교

klasyoung97@khu.ac.kr, kimgh2836@khu.ac.kr, qsibmini@khu.ac.kr,

rhee@khu.ac.kr *kyuhyeonkim@khu.ac.kr

Luminance-adaptive Image Stitching by Using Deep learning based Feature Detection and Matching

Minyoung Kim, Geonho Kim, Minseok Lee, Seongbae Rhee *Kyuhyeon Kim

Kyunghee University

요약

최근 가상 현실(Virtual Reality), 파노라마(Panorama) 영상 등에 관한 관심과 수요가 증가함에 따라 고해상도 영상을 얻기 위한 영상 스티칭(Image Stitching)에 관한 연구가 활발하게 진행되고 있다. 영상 스티칭은 다수의 영상을 하나의 영상으로 합성해 카메라의 좁은 시야각 문제를 해결함으로써 사용자에게 몰입감과 현장감을 제공할 수 있는 기술이다. 영상 스티칭에 있어 특징점 추출 및 매칭 과정의 정확도는 스티칭 영상의 품질을 결정짓는 핵심적인 요소이지만, 기존의 특징점 추출 및 매칭 방법은 밝기가 어둡고 선명도가 낮은 영상의 스티칭에서 정확도가 저하될 수 있고 생성된 스티칭 영상의 품질 또한 저하될 수 있다는 제한 사항이 있다. 이에 본 논문에서는 앞선 제한적 영상에 대하여 특징점 추출 및 매칭의 정확도를 높여 스티칭 영상의 품질을 높이기 위하여 SuperPoint와 SuperGLUE를 활용한 입력 영상의 밝기 적응형 영상 스티칭 방법을 제안하고자 한다.

1. 서론

최근 영상의 특징점(Keypoint)을 활용하는 다양한 연구가 진행되고 있으며, 그중 영상 스티칭(Image Stitching) 기술이 대표적이다[1]. 영상 스티칭은 크게 특징점 추출, 특징점 매칭, 호모그래피(Homography) 추정, 영상 왜곡(Warping), 합성까지의 일련의 과정을 지칭하며[2], 이때 추출된 특징점은 기술자(Descriptor)를 통해 특징점 간 유사도 계산을 통해 매칭되며, 매칭된 특징점은 영상 간의 중첩 영역(Overlap Region)의 설정과 호모그래피 추정에 활용된다[3]. 이에 스티칭 영상의 품질을 높이기 위해서는 정확한 특징점 추출 및 매칭 과정이 필수적이다[4].

대표적인 특징점 추출 및 매칭 알고리즘은 ORB(Oriented FAST and Rotated BRIEF)[5]를 활용한 방법으로써, 이미지 피라미드를 활용한 FAST(Features from Accelerated Segment Test)[6] 알고리즘을 사용해 특징점을 추출하고, 행렬의 회전변환을 활용한 BRIEF(Binary Robust Independent Elementary Features)[7] 기술을 활용하는 기술이다. 이때 ORB 기술은 FLANN(Fast Library for Approximate Nearest Neighbors)[8]과 같은 개선된 매칭 알고리즘과 결합하여 영상을 스티칭할 수 있다. ORB 기술은 특징점 추출 및 매칭에 있어서 높은 정확도를 갖기에 다양한 스티칭 기술에서 활용되지만, ORB 기술의 기저 알고리즘인 FAST는 영상의 단위 패치 내의 픽셀 패턴을 활용한 기술로써, 영상이 어둡거나 선명도가 낮은

경우에는 추출되는 특징점의 수와 매칭 정확도가 저하된다는 제한 사항이 있다. 이에, 본 논문에서는 어둡거나 선명도가 떨어지는 영상을 효과적으로 스티칭할 수 있는 방법을 이어지는 2장에서 제안하고자 하며, 3장에서는 테스트 영상을 통해서 제안 기술의 효용성을 검증하고자 한다. 마지막 4장에서는 결론을 지으며 마무리하고자 한다.

2. 딥러닝 기반 특징점 추출 및 매칭 기술을 활용한 밝기 적응형 영상 스티칭

본 논문에서는 입력 영상이 어둡거나 선명도가 떨어져 기존 특징점 추출 및 매칭 기술을 통한 영상 스티칭이 제한되는 것을 극복하기 위하여 딥러닝 기반 특징점 추출 및 매칭 기술을 활용한 밝기 적응형 영상 스티칭 방법을 제안하고 있다. 이를 지원하기 위하여 제안 기술의 구조는 그림 1과 같이, 두 장의 입력 영상을 입력으로 받아 해당 영상의 밝기를 판단하는 휘도 측정 모듈, 2D 필터 모듈, Feature Detection 모듈, Feature Matching 모듈, Homography 추정 모듈, 영상 합성 모듈로 구성된다.

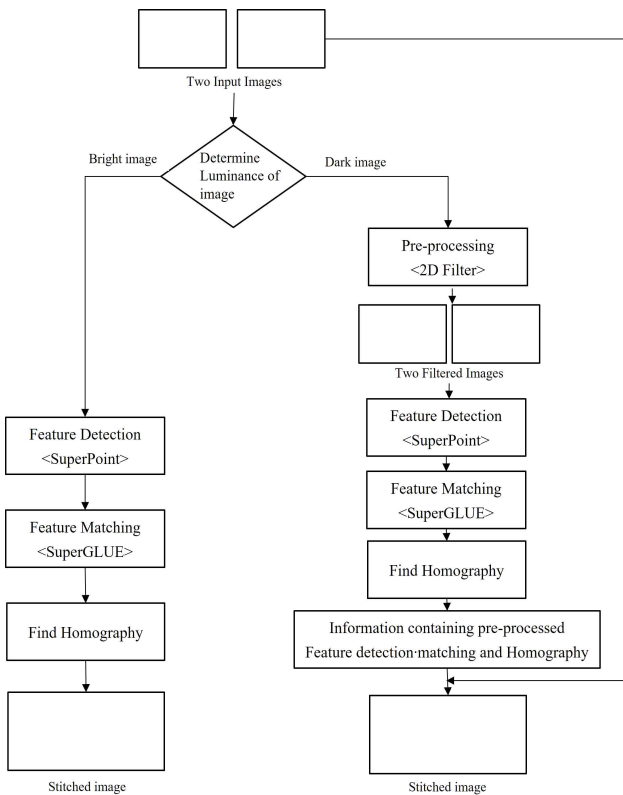


그림 1. 제안 기술 구조도

2.1 영상의 휘도 측정

영상의 휘도(Luminance) 측정 모듈은 입력된 영상이 어두운 영상인지를 판단하는 모듈로써, 영상의 휘도 정보를 활용하며, 이때의 휘도 값(Y)은 입력 영상의 채널 정보인 R, G, B를 활용한다. 본 논문에서 사용된 변환은 ITU-R BT.601-7[11]에 의해 정의된 방식을 따르며, 아래 식 1과 같이 계산된다.

$$Y = (0.257 \times R) + (0.504 \times G) + (0.098 \times B) + 16 \quad (1)$$

계산된 Y는 휘도 임계값(Y_{th})과의 비교를 위하여 식 2와 같은 백분율(P_Y)로 변환된다. 이때 식 2에서 나타나는 M과 N은 입력 영상의 가로와 세로 크기이며, Y_{max} 는 휘도 값이 가질 수 있는 최댓값으로 정의된다. 일례로, 입력 영상이 RGB 채널 각각에 1바이트씩 할당되는 영상에서 Y_{max} 는 235로 나타난다. 또한, Y_{th} 는 실험을 통해서 정의된 임계값으로써, Y_{th} 보다 작은 P_Y 의 영상에 대해서는 2D 필터 활용 모듈이 진행된다. 본 논문의 실험에서는 Y_{th} 를 25로 설정하여 실험하였다.

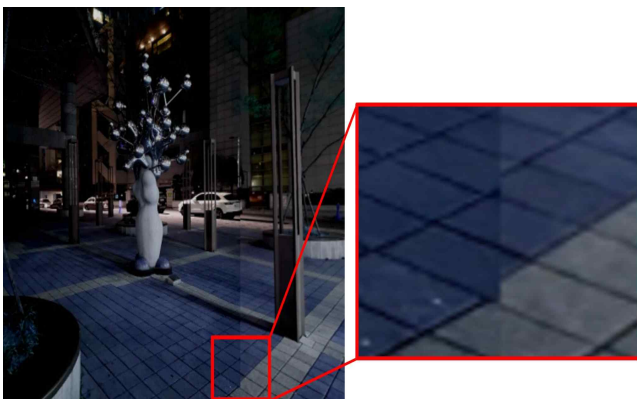
$$P_Y = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M \frac{Y_{i,j}}{Y_{max}} \times 100 \quad (2)$$

2.2 2D 필터

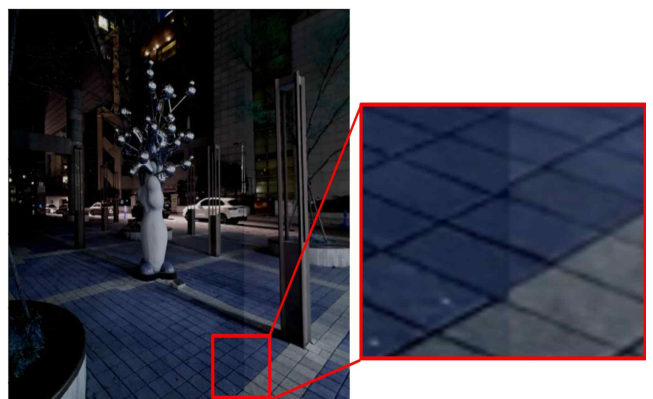
2D 필터 모듈은 앞선 휘도 측정 모듈에서 계산된 휘도 백분율(P_Y)이 임계값(Y_{th})보다 작은 경우에 대하여 수행된다. 일반적으로 2D 필터는 특정한 커널(Kernel)을 이미지에 콘벌루션(Convolution)하여 블러(Blur) 처리하거나 경계선(Edge)을 추출하는 데 활용되지만, 본 논문의 2D 필터 모듈은 입력 영상의 선명도를 높이기 위하여 그림 2와 같이 설계된 커널을 활용한다. 일례로, 그림 3의 (a)와 같이 스티칭 되는 입력 영상에 대하여 2D 필터 모듈을 적용하는 경우, 입력 영상의 선명도가 향상되어 특징점 추출 및 매칭 정확도가 높아지고, 이에 따라서 그림 3의 (b)에서 나타나는 바와 같이 스티칭 결과가 개선된 것을 확인할 수 있다.

0	-1	0
-1	5	-1
0	-1	0

그림 2. 선명도 향상을 위한 2D 필터 커널

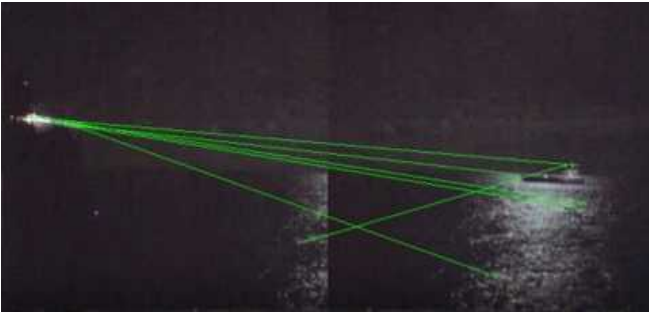


(a) 원본 영상 스티칭 결과

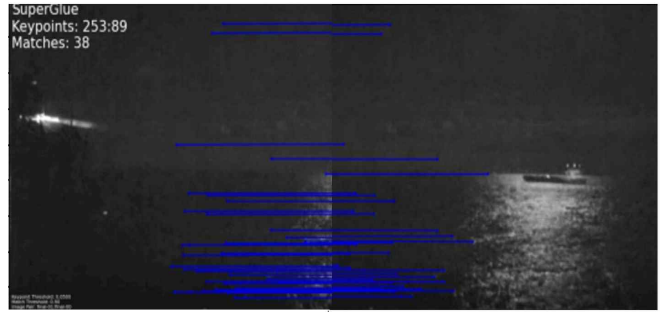


(b) 필터를 사용한 전처리 후 영상 스티칭 결과

그림 3. 2D 필터를 적용한 영상 스티칭 결과 예시



(a) ORB를 이용한 특징점 추출 및 매칭 결과



(b) SuperPoint 기반 특징점 추출 및 매칭 결과

그림 4. ORB와 SuperPoint의 특징점 추출 및 매칭 결과 비교

2.3 SuperPoint 및 SuperGLUE 기반 특징점 추출 및 매칭

본 논문에서는 기존 ORB 알고리즘이 갖는 제한 사항을 극복하기 위하여 딥러닝 기반의 특징점 추출 및 매칭 기술을 활용하고자 한다. 이때 딥러닝 기반 특징점 추출 네트워크로는 SuperPoint[9]를 사용하여, 딥러닝 기반 특징점 매칭 네트워크로는 SuperGLUE[10]를 사용한다. SuperPoint는 영상을 패치 단위로 나누어 패치 내 어떤 픽셀이 특징점인지를 분류(Classification)하는 네트워크로써, 이는 영상에 가해지는 노이즈 평행 이동, 회전, 확대 및 축소, 시점 변환 그리고 다양한 호모그래피 변환 등에 강인한 특징점을 추출하는 알고리즘이기 때문에 영상 스티칭에 적합한 특징점 추출 네트워크이다. 그림 4의 (a)와 (b)에서 나타나는 바와 같이, SuperPoint는 기존 ORB에 비하여 어두운 영상에서 특징점 추출 및 매칭 정확도가 높게 나타나는 것을 확인할 수 있다.

SuperPoint 네트워크는 특징점 기술자를 포함하기에 특징점 매칭까지 수행할 수 있지만, 더 높은 매칭 정확도를 획득할 수 있도록 SuperGLUE를 특징점 매칭에 활용하였다. SuperGLUE는 GNN(Graph Neural Network)을 기반으로 설계된 딥러닝 네트워크로써, 특징점의 지역적 특성과 구조적 특성을 모두 고려하여 특징점을 매칭하므로 Nearest neighbor 방식의 지역적 특성만을 활용하는 기존 매칭 기술보다 높은 매칭 정확도를 얻을 수 있다.

본 논문에서는 입력 영상의 휘도 정보에 따라 2D 필터를 적용하여 영상의 선명도를 높이고, SuperPoint와 SuperGLUE를 활용하여 높은 특징점 추출 및 매칭 정확도를 얻음으로써 어두운 영상을 높은 품질로 스티칭할 수 있는 기술을 제안하였다. 이에 다음 장에서는 테스트 영상을 통해서 제안 기술의 효용성을 검증하고자 한다.

3. 실험 결과

본 논문에서는 2D 필터를 활용하여 입력 영상의 선명도를 높이고, SuperPoint와 SuperGLUE를 활용하여 특징점 추출 및 매칭 정확도를 높여 어둡고 선명도가 낮은 영상에 대하여 영상 스티칭 기술을 적용할 수 있는 방법을 제안하고 있다. 이에 본 논문의 3장에서는 그림 5와 같이 나타나는 어둡고 선명도가 떨어지는 테스트 영상을 통해서 제안 기술의 효용성을 검증하고자 한다. 이때 실험 환경으로는 Windows 10의 운영체제에서 i7-6700 HQ CPU와 Visual Studio Code를 통해 설계되었다.



그림 5. 두 장의 입력 영상

제안 기술과의 비교 실험을 위해 설정된 기존 스티칭 기술은 ORB를 활용하여 특징점을 추출하고 FLANN 기반 특징점 매칭 기술을 결합한 스티칭 기술이며, 그림 5의 테스트 영상을 통해서 스티칭한 결과는 각각 그림 6의 (a)와 (b)로 나타난다.



(a) 기존 기술 영상 스티칭 결과



(b) 제안 기술 영상 스티칭 결과

그림 6. 기존 기술과 제안 기술 영상 스티칭 결과 비교

그림 6의 (a)에서 나타나는 바와 같이 기존 스티칭 기술에서는 입력 영상이 어두우므로 특징점 추출 및 매칭의 정확도가 낮게 나타났고, 그에 따라서 호모그래피 추정 정확도가 낮아져 스티칭이 제대로 되지 않는 것을 확인할 수 있다. 반면, 그림 6의 (b)에서 나타나는 바와 같이 제안 기술을 통해서 스티칭한 결과는 올바르게 나타나는 것을 확인할 수 있다.

4. 결론

본 논문에서는 어둡고 선명도가 낮은 영상에 대하여 높은 품질의 스티칭 영상을 획득할 수 있는 딥러닝 기반 특징점 추출 및 매칭 기술을 활용한 밝기 적응형 영상 스티칭 방법을 제안하였다. 기존 ORB 기반의 특징점 추출 및 매칭을 활용한 영상 스티칭 기술에서는 어둡고 선명도가 낮은 영상을 스티칭함에 있어서 특징점 추출 및 매칭 정확도가 저하되어 낮은 품질의 스티칭 영상을 얻게 되지만, 제안 기술에서는 높은 품질의 스티칭 영상을 획득할 수 있음을 실험을 통해 검증하였다.

참고문헌

- [1] Sang-Burm Rhee, "Efficient Image Stitching Using Fast Feature Descriptor Extraction and Matching", KIPS Transactions on Software and Data Engineering, pp.65-70, 2013.
- [2] R. Szeliski, "Image Alignment and Stitching: A Tutorial," Foundations and Trends in Computer Graphics and Vision, Vol. 2, No.1, pp.1-104, 2006.
- [3] Jiheon Im, et al., "Images Grouping Technology based on Camera Sensors for Efficient Stitching of Multiple Images", in Journal of Broadcast Engineering, Vol 22, No.6, pp.713-722, 2017.
- [4] C. Y. Chen and R. Klette, "Image Stitching - Comparisons and New Techniques", Computer Analysis of Images and Patterns, pp.615-622, 1999.
- [5] Ethan Rublee, et al., "ORB: An efficient alternative to SIFT or SURF", International Conference on Computer Vision, 2011.
- [6] E. rosten and Tom Drummond, "Machine learning for high-speed corner detection", Cambridge University: Department of Engineering, 2006.
- [7] M. Calonder, et al. "BRIEF: Binary Robust Independent Elementary Features", European Conference on Computer Vision, pp.778-792, 2010.
- [8] Marius Muja and David G. Lowe, "Scalable Nearest Neighbor Algorithms for High Dimensional Data", Pattern Analysis and Machine Intelligence, Vol 36, 2014.
- [9] D. DeTone, et al., "SuperPoint: Self-Supervised Interest Point Detection and Description", IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.1-13, 2018.
- [10] Paul-Edouard Sarlin, et al., "SuperGlue: Learning Feature Matching with Graph Neural Networks", The IEEE / CVF Computer Vision and Pattern Recognition Conference, pp.1-18, 2020.
- [11] Recommendation ITU-R BT.601-7, Studio Encoding Parameters of Digital Televisions for Standard 4:3 and Wide-screen 16:9 aspect ratios, 2011.