

다중스케일 서라운드 스위칭 맵을 이용한 다중 노출 영상 융합

고영호*¹, 이승환*², 이성학*³

Multi-Exposed Image Fusion using Multiscale-Surround Switching Map

Young-Ho Go*¹, Seung-Hwan Lee*², and Sung-Hak Lee*³

이 논문은 2021년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임
(NRF-2021R1I1A3049604)

요약

HDR(High Dynamic Range)영상 합성은 다양한 휘도 범위의 영상들을 합성을 통해 동적 범위를 확장하여 디스플레이 상에서 사람의 눈으로 관측한 장면을 재현하기 위한 방법이다. HDR 영상을 획득하기 위해서는 전문가용 카메라 장비를 사용해야하는 제한점이 있다. 이러한 촬영 한계를 극복하기 위해서 다양한 노출의 LDR(Low Dynamic Range) 영상을 이용한 HDR 영상 합성이 연구되고 있다. 본 논문은 Retinex 알고리즘을 기반으로하여 고노출, 저노출 두 영상에서 다중 스케일의 디테일 영상을 생성한다. 생성된 다중 영상에 다중 스케일 서라운드 스위칭 맵(Multiscale-surround switching map)을 적용하여 각 스케일에서 선택적으로 세부 정보를 가져온다. 제안된 방법은 각 영상의 스케일 별로 우수한 세부 정보 영역을 SMap으로 선택하여 합성을 진행하였고 기존방법과 비교해 영상의 노이즈와 헤일로 현상을 감소시키고 선명도의 증가 및 색 밸런스를 개선하였다.

Abstract

High Dynamic Range(HDR) image synthesis is a method of expanding the dynamic range by synthesizing images with various luminance ranges to reproduce scenes observed by the human eye on displays. There is a limitation that professional camera equipment must be used to obtain HDR images. To overcome these shooting limitations, research is being conducted on HDR image synthesis using various exposures of Low Dynamic Range(LDR) images. This paper generates multi-scale detail images from high-exposure and low-exposure images using the Retinex algorithm. The proposed method selectively retrieves detailed information from each scale using the Multiscale-surround Switching Map. The proposed method synthesizes excellent detail information areas for each scale of the image using SMap and compared to existing methods, it reduces noise and halo effect in the image, increases sharpness, and improves color balance.

Keywords

HDR, image fusion, multiscale switching map, MSR

* 경북대학교 IT대학 전자공학부(*³ 교신저자)
- ORCID¹: <https://orcid.org/0009-0007-0980-1161>
- ORCID²: <https://orcid.org/0009-0005-4430-4975>
- ORCID³: <https://orcid.org/0000-0002-1030-381X>

· Received: Mar. 18, 2024, Revised: Apr. 15, 2024, Accepted: Apr. 18, 2024
· Corresponding Author: Sung-Hak Lee
School of Electronics Engineering, Kyungpook National University, 80
Daehak-ro, Buk-gu, Daegu, 41566, Korea
Tel.: +82-53-950-7216, Email: shak2@ee.knu.ac.kr

I. 서 론

영상의 동적범위(Dynamic range)는 디지털 이미징과 컴퓨터 비전 분야에서 중요한 개념으로서 영상의 가장 어두운 영역과 가장 밝은 영역 사이의 밝기 레벨 차이를 의미한다. 현존하는 디스플레이 기술은 인간 눈의 넓은 동적 범위에 비해 한정된 동적 범위를 가지고 있어, 인간의 시야에 들어온 장면을 완벽하게 재현하는 것에 어려움을 겪고 있다. 이러한 시각적 현실감을 개선하기 위한 연구 방법 중 하나가 HDR(High Dynamic Range) 합성 기법이다[1].

HDR 합성 기법은 여러 다른 노출 수준에서 촬영된 LDR(Low Dynamic Range) 영상들을 효과적으로 합성하여 하나의 HDR 영상을 생성한다. Debevec 등은 각 노출 시간이 다른 LDR 영상들을 사용하여 영상 센서와 빛의 양 사이의 상호작용 법칙(Reciprocity law)을 활용한다. 이후 영상센서의 응답함수를 추정하여 단일 HDR 영상을 생성한다[2].

그러나, 생성된 HDR 영상은 기존의 LDR 디스플레이에서 표현하기 어렵다. 따라서 톤 매핑(Tone mapping) 알고리즘을 이용하여 HDR 영상의 정보를 영상 출력 장치의 동적범위에 맞게 LDR 영상으로 압축해야 한다. 이러한 LDR 디스플레이에 렌더링하기 위한 톤 매핑은 전역적 톤 매핑과 국부적 톤 매핑 알고리즘이 있다. 전역적 톤 매핑 알고리즘은 실시간 처리가 가능하여 속도면에서 좋은 성능을 보여주지만, 극단적으로 넓은 다이내믹 레인지의 영상의 경우에는 한계가 있다. 대표적으로는 Drago Logarithmic 매핑, Ward histogram adjustment 등이 있다. 국부적 톤 매핑은 HVS(Human Visual System)에서 빛의 절대적인 값보다 픽셀간의 주변 빛의 값 사이의 차이를 통해 빛의 세기를 인식하는 특성을 잘 반영한 결과를 보여준다. 그 예시로 iCAM, Retinex 알고리즘 등이 있다.

대표적인 Retinex 알고리즘은 영상재현 시에 HVS에서는 반사된 빛의 분포보다 물체의 반사도에 더 큰 영향을 받는다는 이론을 기반으로 한다. 물체의 반사도만큼 주변의 빛을 반사시켜 우리의 시야에 맏히게 되는 빛을 내보낸다[3]. 각 픽셀에 컨볼루션(Convolution) 마스크를 적용하여 주변 화소와 차이 값을 통한 상대적 변화율을 측정한다. Retinex 기반

알고리즘은 단일의 마스크를 사용하는 SSR(Single Scale Retinex)과 영상에 여러 마스크를 적용하여 다중 스케일(Multi scale)로 분해한 각 스케일에서의 결과를 조합하여 사용하는 MSR(Multi Scale Retinex)로 나누어진다[4]. 다양한 공간적 크기의 특징을 고려하여 영상의 더 풍부한 특징을 제공하지만 기존 MSR로 다중스케일의 정보를 통합하여 자연스러운 결과물을 만들어내는 것은 어렵다.

본 논문에서는 MSR을 이용하여 분산값에 따른 다중스케일 영상을 획득하고, 각 스케일들에 Multiscale surround switching map을 적용하여 스케일별 영상의 장점들을 효과적으로 합성한다. 해당 SMap(Switching map)들은 양방향 필터(Bilateral filter)를 적용한 서라운드맵(Surround map)으로 영상의 세부 정보를 살리고 적응적인 시그모이드 함수(Sigmoid function)를 사용하여 스위칭 성능을 향상시킨다.

이러한 다중스케일 SMap을 영상에 적용할 경우 기존의 MSR 기법보다 색감 및 채도 성분이 잘 보존되고 영상 경계 성분의 표현과 강한 경계 영역에서 헤일로 현상(Halo effect)을 줄임으로써 영상의 자연스러움을 증가시키는 결과를 얻을 수 있다.

II. 관련 연구

2.1 다중 노출 영상 합성

다중 노출 영상의 합성(MEF, Multi-exposed image fusion)은 HDR 영상을 생성하는 중요한 기술로서, 여러 다른 노출 수준의 LDR영상을 적절하게 합성하여 영상의 밝은 영역과 어두운 영역의 세부 정보를 모두 표현하여 결과적으로 동적 범위의 확장이 가능하다[5]. 이러한 합성 결과 영상은 디스플레이에서 현실적인 영상을 구현한다. MEF의 일반적인 방법으로 영상 블렌딩(Image blending)이 있고 합성된 단일 HDR 영상을 디스플레이에 렌더링하기 위해 톤 매핑 기법(Tone mapping operator)을 사용한다.

다중 노출 영상의 블렌딩은 동일한 장면을 서로 다른 노출 시간으로 촬영한 여러 장의 LDR 영상들을 부드럽게 혼합하여 자연스러운 전환 효과를 내며 HDR 영상을 획득한다.

이때, 다중 노출 입력 영상으로 사용되는 LDR 영상은 서로 다른 노출 시간에 따라 다른 광학 특성을 가지며, 노출시간에 따라 나타난 특성은 생성될 HDR 영상에 영향을 미친다.

노출 시간이 길어지면 어두운 영역의 세부적인 정보가 더 많이 드러나지만, 밝은 영역에서는 과포화(Overexposure)현상이 발생한다. 반면에, 노출 시간이 짧으면 밝은 영역의 세부 정보는 더 잘 보존되지만 어두운 영역은 정보가 소실되고 채도가 낮아진다. 다중 노출의 영상을 적절한 노출 시간 비율로 혼합함으로써 단일 노출 영상보다 휘도 범위가 넓은 HDR 영상의 합성이 가능하다. 그림 1(a), (b)에서 표시된 영역은 노출도가 다른 두 영상의 주요 세부 영역을 나타낸다. 이러한 다중 노출 영상 블렌딩 기술은 HDR 영상의 생성과정에서 중요한 역할을 하며, 다양한 노출 시간의 영상들을 조합하여 풍부한 휘도 정보를 효과적으로 표현하는데 활용된다. 다만, 합성되는 영상들의 노출 수준에 따라 노이즈가 심해지거나 포화현상의 발생으로 세부 정보의 손실이 발생한다.

TMO(Tone Mapping Operator)는 동적범위의 한계를 가진 LDR 디스플레이에서 HDR 영상의 표현을 개선하기 위해 사용하는 방법으로 전역적 톤매핑과 국부적 톤매핑으로 나뉜다.

전역적 톤매핑은 영상의 전체 영역에 한가지 톤매핑 함수를 적용하는 방법이다. 영상의 모든 영역을 동일한 함수로 조절하게 되어 영상의 밝기 및 색상을 일괄적으로 조절이 가능하다. 전역적 톤매핑의 예시로, R. Fattal et al.[6]은 사진 노출 블렌딩 기술에서 정보의 손실을 최소화 하기위해 로그 도메인 엔트로피를 사용한다. 로그 도메인에서 데이터를 처리하면서 이상치와 같은 극단적 값에 대한 영향을 줄여 원본 사진의 세부 정보를 보존한다[7].

E. Reinhard et al.[8]은 인간의 photoreceptor 생리학을 기반으로 광수용체의 동적 범위 메커니즘을 모델로 한다. 조명 수준에 대해서 photoreceptor가 자동 조절하는 HVS의 메커니즘에서 착안하여 전역적 연산자를 제공한다. 다만, 동일한 함수를 전체 영상에 적용하기 때문에 특정 부분에서 세부정보 손실이 발생한다.

국부적 톤매핑은 영상의 작은 영역 또는 패치로 분할하여 각 패치에 대해 톤 매핑 함수를 적용한다. 각 패치를 개별적 조정으로 영상의 세부 정보를 보존하는데 중점을 둔다. 예시로, F. Durand and J. Dorsey[9]은 결합형 양방향 필터를 사용하여 세부 정보들을 넓은 강도 변이(Broader intensity variation)로부터 분리한다. 해당 방법으로 영상을 부드럽게 처리하면서 경계 정보들을 손상시키지 않는 방법을 제안한다.

E. H. Land and J. J. McCann[10]도 HVS의 작동 원리를 기반으로 영상의 조명 성분과 반사 성분을 추정하기 위해 필터를 적용하고 이후 분리된 스케일을 조합하여 영상의 품질을 향상시킨다. 국부적 톤 매핑 연산자들은 각각의 방식으로 HDR 영상을 효과적으로 처리한다. 다만, 국부적 톤매핑의 경우 계산에 있어서 많은 시간이 소요되고 연산이 복잡하다. 또한, 밝거나 어두운 부분에서 정보왜곡, 노이즈 등 문제가 발생한다.



(a) (b)
그림 1. 다중 노출 영상 (휘도 채널):
(a) 저노출 영상, (b) 고노출 영상

Fig. 1. Input multi-exposure images (Luminance channel):
(a) Low-exposure image and (b) High-exposure image

2.2 딥러닝 기반 영상 합성

최근에는 딥러닝 기술이 컴퓨터 비전 및 영상처리에 많이 활용된다. 그 중 다중 노출 영상 합성에서도 성능을 향상시키기 위해 딥러닝을 기반으로 하는 다중 노출 영상 합성 방법이 제안되었다.

딥러닝 기반 영상 합성은 end-to-end 학습이 가능하여 전체 시스템을 하나의 모델로 설정함으로써, 간소화된 단계로 다양한 입력데이터에 대해서 좋은 성능을 보인다[11].

MEF에서 딥러닝 기반 영상 합성은 지도 학습과 비지도 학습으로 나뉜다. 지도 학습으로는 학습을 위한 레이블이 있는 훈련 데이터를 사용한다. 학습한 관계를 기반으로 다양한 조명 조건이나 환경에서도 개선된 합성 효과를 보여준다. P. Debevec et al.[12]은 3개의 LDR 영상 세트를 기반으로, 컨볼루션 신경망(CNN)을 사용하여 고스트가 줄어든 개선된 HDR 영상을 획득하였다. 지도학습에서는 ground truth 영상이 충분하지 않더라도 일정량 이상의 영상이 필요하다. 다만 생성된 ground truth 영상은 실제가 아니기 때문에, 사용할 경우에는 합성 성능이 감소할 수 있다. 따라서, 많은 연구에서는 비지도 방식으로 네트워크를 학습을 시도한다.

MEF를 위한 비지도 학습은 입력 영상에서 특징을 추출하기 위해 네트워크 구조를 수정한다. L. Qu et al.[13]은 자가지도 다중 학습을 사용하는 Transformer를 기반으로 퓨전 프레임 워크인 TransMEF를 제안하였다. 해당 프레임 워크는 내추럴 영상 데이터셋에서 훈련될 수 있는 인코더-디코더 네트워크를 기반으로 하기때문에 ground truth 영상을 필요로 하지 않는다. CNN 모듈과 Transformer 모듈이 결합된 인코더를 기반으로 하여 로컬 및 전역 정보에 모두 집중할 수 있다. 우수한 합성 성능을 보이고 좋은 품질의 합성 영상을 생성할 수 있지만, 복잡한 네트워크의 아키텍처는 아직까지 많은 데이터셋이 필요하다. 이는 더딘 훈련과 증가된 메모리 사용으로 이어진다.

III. 제안 방법

본 논문에서는 다중스케일 스위칭 맵을 이용한 HDR 영상 합성 알고리즘을 제안한다. 그림 2는 제안한 알고리즘의 플로우차트를 확인 할 수 있다. 먼저 노출도가 다른 저노출, 고노출의 입력영상을 사용한다. 두 영상을 색 공간 변환(CSC, Color Space Conversion)을 통해 CIELAB 채널에서 휘도 성분을 합성 대상으로 사용하기 위해 분리한다. CIELAB

색 공간의 영상은 색채 채널(a,b)과 휘도(L)채널로 분리된다. 영상의 색 정보와 휘도 정보를 독립적으로 처리하여 채도 및 색상 정보를 보존하면서 영상 합성이 가능하다는 장점이 있다[14]. 다음으로 휘도 채널은 다른 스케일의 SSR(Single Scale Retinex)영상으로 변환된다. 톤 압축 영상 생성을 위해 각 입력 영상의 국부 배경 성분을 제거하는데, 배경 성분은 영상에 가우시안 필터(Gaussian filter)를 적용해서 얻을 수 있다[15]. SSR 출력 영상은 다음식으로부터 얻어진다.

$$R_{SSR}(x,y;v) = \log(I(x,y)) - \log(I(x,y) * G(x,y;v)) \quad (1)$$

여기서 $I(x,y)$ 는 입력 영상이다. $G(x,y;v)$ 는 가우시안 필터이며 v 는 필터에 적용되는 분산 값이다. SSR의 출력에 로그를 취함으로써 반사성분의 dynamic range를 압축하는 효과를 가져온다.

가우시안 필터의 분산 값을 다르게 적용하여 서로 다른 특징을 가진 3가지 배경 분산 스케일을 가지는 영상을 생성하며 다음과 같이 표기한다.

$$\begin{aligned} SSR_1 &: R_{SSR}(x,y;v_1) \\ SSR_2 &: R_{SSR}(x,y;v_2) \\ SSR_3 &: R_{SSR}(x,y;v_3) \end{aligned} \quad (2)$$

본 논문에서 각 스케일별로 분산값 v 에 따라 다른 Large, Middle, Small 스케일들을 식 (2)에서 $v_1=20$, $v_2=40$, $v_3=70$ 순서대로 SSR_1 , SSR_2 , SSR_3 이라고 지칭한다. 해당 Retinex 알고리즘 적용 과정은 FFT(Fast Fourier Transform)로 주파수 영역에서 진행하여 정보 처리의 속도를 향상 시킨다[16].

고노출과 저노출의 영상에서 각각 원하는 세부영역의 정보를 선택하기 위해 스위칭 맵(SMap)을 생성한다. 해당 SMap은 식 (3)에서처럼 두 영상의 휘도 성분에 양방향 필터(Bilateral filter)를 적용하여 서라운드 맵을 생성한다. 이후 생성된 서라운드 맵에 시그모이드 함수를 적용하여 SMap의 강한 경계영역 선택 능력을 향상시킨다. 서라운드 맵을 생성할 때, 양방향 필터의 공간 범위(Spatial range)의 값을 조정하여 다른 특성을 가진 다중 스케일 $Blur_i(x,y;\sigma_s)$ 들을 생성한다.

$$Blur_i(x, y; \sigma_s) = L_{LE_i}(x, y) * BF(x, y; \sigma_s) \quad (3)$$

여기서 $L_{\leq L}(x, y)$ 은 저노출 영상의 L채널 영상이다. $BF(x, y; \sigma_s)$ 는 양방향 필터 함수이며 σ_s 는 양방향 필터의 공간 범위의 값이다. 공간 범위는 양방향 필터에서 좌표공간에 적용되는 가우시안 필터의 분산 값을 의미한다. 주변 픽셀의 강도가 비슷한 경우, 즉 강한 경계 부분이 아닐 때 값이 커질수록 주변 픽셀에 영향을 크게 미친다. $Blur_i(x, y; \sigma_s)$ 는 σ_s 값에 따라 $Blur_{1,2,3}$ 으로 분류된 3개의 SMap을 생성한다.

$$Sig(\alpha, x) = \frac{1}{1 + e^{-\alpha x}} \quad (4)$$

$$SMap_i = Sig(\alpha, Blur_i(x, y; \sigma_{s_i})) \quad (5)$$

식 (4)는 시그모이드 함수이다. α 는 시그모이드 함수의 기울기를 조절하는 값이며 커질수록 함수의 기울기가 커진다. $Blur_i(x, y; \sigma_{s_i})$ 에 $Sig(\alpha, x)$ 을 적용하여 $SMap_{1,2,3}$ 을 생성한다.

SSR_1, SSR_2, SSR_3 의 각 장점을 살릴 수 있는 스케일별 SMap들을 영상에 각각 적용한다. 저노출 영상에는 다중스케일 $SMap_{1,2,3}$ 을, 고노출 영상에는 그와 반대되는 영역을 선택하도록 $\in v_SMap_{1,2,3}$ 을 적용한다. 식 (6)에서 식 (8)을 통해 각 저노출(LE), 고노출(HE) SSR 영상에 대해 SMap이 적용된 선택영상이 생성된다.

$$Inv_SMap_i = (1 - SMap_i) \quad (6)$$

$$S_{LE_i} = SSR_LE_i \times SMap_i \quad (7)$$

$$S_{HE_i} = SSR_HE_i \times Inv_SMap_i \quad (8)$$

식 (7), (8)에서 노출도에 적합한 SMap들을 SSR 출력 영상에 적용해준다. 이후 두 입력 영상의 다중스케일 영상들을 식 (9)에서 각 스케일별로 저노출 영상과 고노출 영상을 합성한다.

$$L_{Fused_i} = S_{LE_i} + S_{HE_i} \quad (9)$$

$$L_{Fused} = L_{Fused_1} + L_{Fused_2} + L_{Fused_3} \quad (10)$$

여기서 L_{Fused} 는 각 스케일들의 영상을 합성한 출력 영상이다.

색 성분 채널에서는 식 (11), (12)와 같이 저노출 영상의 a, b 채널에는 $SMap_1$ 을 고노출 영상의 a, b 채널에는 $\in v_SMap_1$ 을 적용한다. 각 스위칭 맵들로 보존하고자 하는 색 성분의 영역들을 선택한다.

$$a_{SLE}, b_{SLE} = (a_{LE}, b_{LE}) \times SMap_1 \quad (11)$$

$$a_{SHE}, b_{SHE} = (a_{HE}, b_{HE}) \times Inv_SMap_1 \quad (12)$$

$$a_{Fused}, b_{Fused} = (a_{SLE} + a_{SHE}, b_{SLE} + b_{SHE}) \quad (13)$$

생성된 각 $L_{Fused}, a_{Fused}, b_{Fused}$ 채널을 RGB채널로 색 공간 변환 후 합성을 통해 결과 영상을 생성한다.

3.1 저노출 스위칭 맵 생성

저노출(LE, Low-Exposure) 영상과 고노출(HE, High-Exposure) 영상의 경우 가지고 있는 유효한 정보가 다르기 때문에 가중치를 가진 SMap으로 원하는 영역을 매핑하여 노출별 영상의 국부영역을 선택한다. 그림 2의 입력 영상처럼 LE 영상에서 유효한 정보가 많은 창틀 안의 실외 영역을 선택하기 위해 SMap을 적용하고, HE 영상에서 정보가 포화되지 않은 실내 영역을 선택하기 위해 SMap의 반전 영상을 적용한다. 이때, 사용되는 SMap은 LE, HE 영상의 휘도 채널에 양방향 필터를 적용하여 블러된(Blurred) HE, LE 서라운드 맵을 생성한다.

그림 3(a)와 같이 블러된 LE 영상에서는 어두운 배경 영역의 구분이 명확히 되고 상대 노출의 디테일 영역의 선택 특성이 뛰어나다. 다만, 밝은 포화 부분의 디테일 영역의 블러링이 잘 이루어지지 않을 경우 부분적인 영역의 low level 포화레벨의 간섭으로 유효하지 않은 HE 영상의 정보가 선택될 수 있다.

반면에 그림 3(b)는 HE 영상의 밝은 배경의 포화영역에서 플레어 현상과 빛 번짐으로 인해 경계영역이 뚜렷하게 구분되지 않는 블러링 현상이 일어난다.

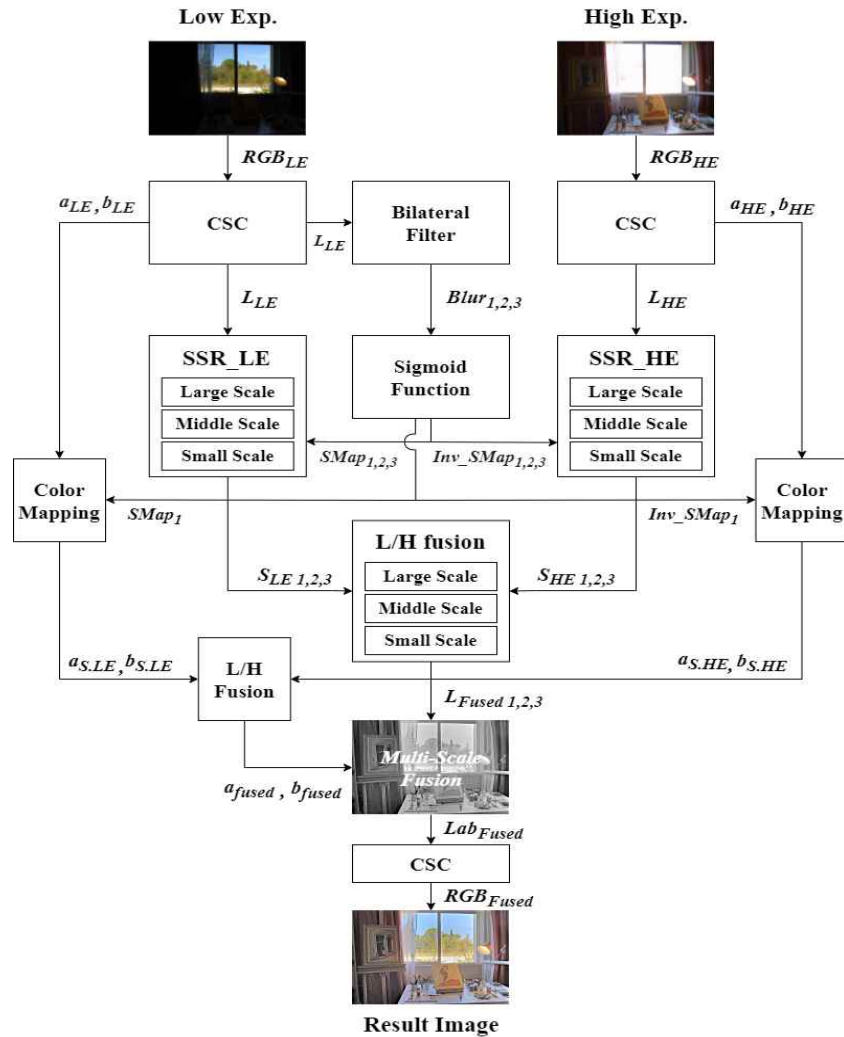


그림 2. 제안된 알고리즘
Fig. 2. Proposed algorithm

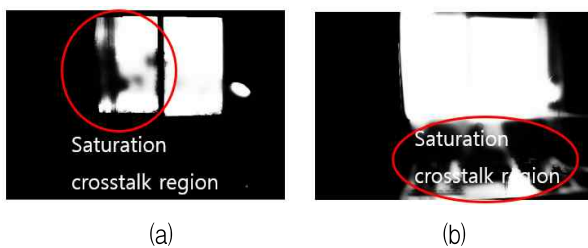


그림 3. SMap에 사용된 블러된 LE, HE 영상:
(a) 블러된 LE 영상, (b) 블러된 HE 영상
Fig. 3. Blurred LE, HE images for SMap:
(a) Blurred LE image and (b) Blurred HE image

따라서 SMap을 생성할 때, 각 블러된 HE 영상과 LE 영상을 합성하여 생성된 최종 SMap을 SSR가 적용된 각 스케일 영상에 매핑하면 포화 영역에서 LE 영상의 유효한 정보가 선택되지 않고 HE 영상의 포화레벨의 간섭으로 유효하지 않은 정보가 선

택된다. 따라서, 본 논문에서는 각 HE, LE 영상의 비포화 영역의 효과적인 선택(Selection)과 합성(Fusion)을 위해 블러된 LE 영상만을 활용하여 경계 성능이 개선된 합성 스위칭 맵을 생성하였다.

3.2 다중 스케일 서라운드 스위칭 맵 생성

SSR 알고리즘을 적용할 때, LE, HE 입력영상에 적용된 가우시안 필터의 분산값을 다르게 조정하면 각 스케일별로 반사성분의 다른 특징이 존재한다. 큰 스케일에서는 넓게 블러링(Wide blurring)된 배경성분이 제거된다. 그림 4(a)와 같이 노이즈가 감소하지만 세부 성분이 부드럽게 처리되며 전역 톤 압축이 된다.

작은 스케일에서는 좁게 블러링(Narrow blurring)된 필터를 적용한 배경성분을 제거한다. 그 결과 그림 4(b)와 같이 세부사항들이 더욱 보존되지만 노이즈가 부각되는 부분들이 존재한다.



그림 4. 다른 분산값의 고노출 다중스케일 SSR 영상:
(a) 큰 스케일 영상, (b) 작은 스케일 영상
Fig. 4. Multiscale High Exposure SSR images with different variance values: (a) Large scale image and (b) Small scale image

그림 5와 그림 6은 각각 양방향 필터의 공간 범위 변화에 따른 SMap과 각 SMap이 적용된 SSR L 채널 영상을 보여준다. 양방향 필터의 Intensity

range는 타겟 픽셀과 주변 픽셀과의 강도 유사도에 따라 가중치를 부여한다. 그림 5에서 공간 범위가 작은 스케일에서는 그림 5(a)처럼 하이라이트 영역의 블러링이 부족하여 포화영역 레벨의 간섭이 심해지고 합성 영상에서 detail loss가 발생한다. 반면 큰 스케일에서는 그림 5(c)처럼 블러 효과가 좋지만 스위칭이 필요한 강한 경계부분에서까지 블러링 현상이 동일하게 발생하여 헤일로 현상을 초래한다. 그림 6은 양방향 필터의 Intensity range는 100, 시그모이드 함수의 slope를 10으로 설정하였을 때의 그림 5의 각 SMap이 적용된 SSR L 채널 영상결과를 보여준다. 결과에서 그림 6(a)처럼 공간 범위가 작은 영상에 세부 정보 손실이 나타나고, 그림 6(c)처럼 공간 범위가 큰 SMap을 적용한 영상에서는 헤일로 현상이 더 심하게 나타난다.

본 논문에서는 이러한 공간 범위에 따른 SMap의 특성을 반영하여 Small scale, Middle scale, 및 Large scale 영상에 대해 공간 범위의 값을 각각 20, 40, 70으로 적용하였다.

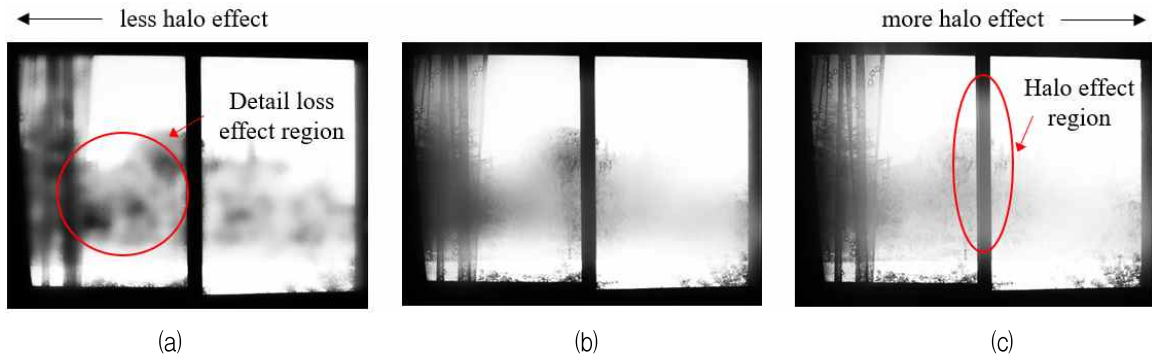


그림 5. 공간 범위 값에 따른 SMap 영상: (a) 공간 범위 20, (b) 공간 범위 40 (c) 공간 범위 70
Fig. 5. SMap images by spatial range: (a) Spatial range 20, (b) Spatial range 40 (c) Spatial range 70

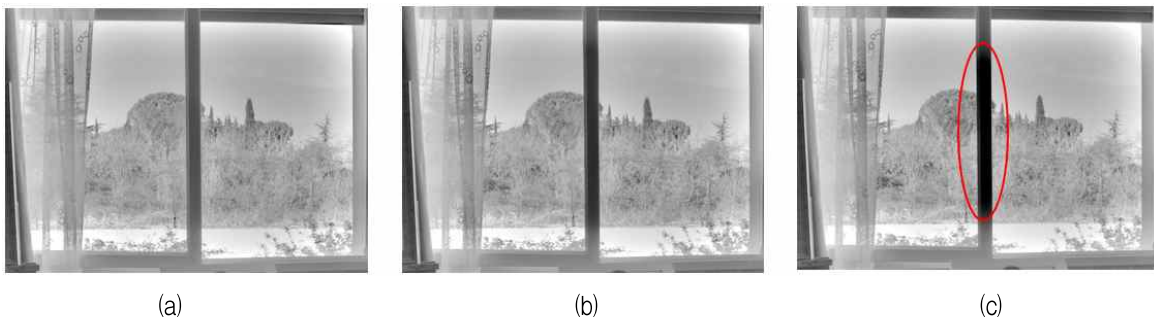


그림 6. 다중스케일 SMap이 적용된 SSR 영상: (a) 공간 범위 20, (b) 공간 범위 40 (c) 공간 범위 70
Fig. 6. SSR images with multi-scale SMap: (a) Spatial range 20, (b) Spatial range 40 (c) Spatial range 70

IV. 실험 결과

제안 방법의 다중 노출 영상 합성 성능 검증을 위해 기존의 MSR[17], EF[18], FMMR[19], DCT-based[20], GRW[21]의 방법들과 비교 평가 실험을 진행하였다.

그림 7은 입력영상 그림 7(a), (b)에 대한 제안방법과 기존의 합성방법을 적용한 비교영상들이다. 위 그림 7(c)에서 SMap없이 L채널에 대해서는 기존 MSR 알고리즘을 적용하고 색상채널에서는 LE, HE 영상의 색성분을 단순 합성하여 나온 결과이다.

그림 7(d) EF의 결과에는 밝은 부분의 채도가 더 좋아 보이지만 어두운 곳에서의 세부사항들이 식별하기 어렵다. 그림 7(e) FMMR의 결과에서는 로컬 명암대비, 밝기, 색상 등 영상의 특징을 고려하여

가중치를 추정한다. 다만, 헤일로 현상 및 노이즈가 심하다. 그림 7(f) DCT-based 방법의 결과는 영상이 자연스럽게 나오지만 어두운부분에서의 세부정보가 잘 드러나지 않는다. 그림 7(g) GRW방법의 결과는 헤일로 현상이 심하며 밝은 부분에서의 채도가 낮아진다. 그림 7(h)에서는 제안된 방법을 사용하였으며 다른 방법들에 비해 어두운 부분의 세부 정보를 잘 살리고 헤일로 효과가 적게 나타나 영상의 자연스러움이 증가한다.

그림 8(a)에서는 입력 영상의 HE 영상에서 노출의 부족으로 인해 다른 방법들은 전반적으로 어두운 실내가 잘 보이지 않는다. 그림 8(g)의 제안방법의 결과 영상은 SSR로 톤매핑을 이용한 합성이기 때문에 확연히 어두운 실내 부분의 세부정보 및 색채가 잘 드러난다.

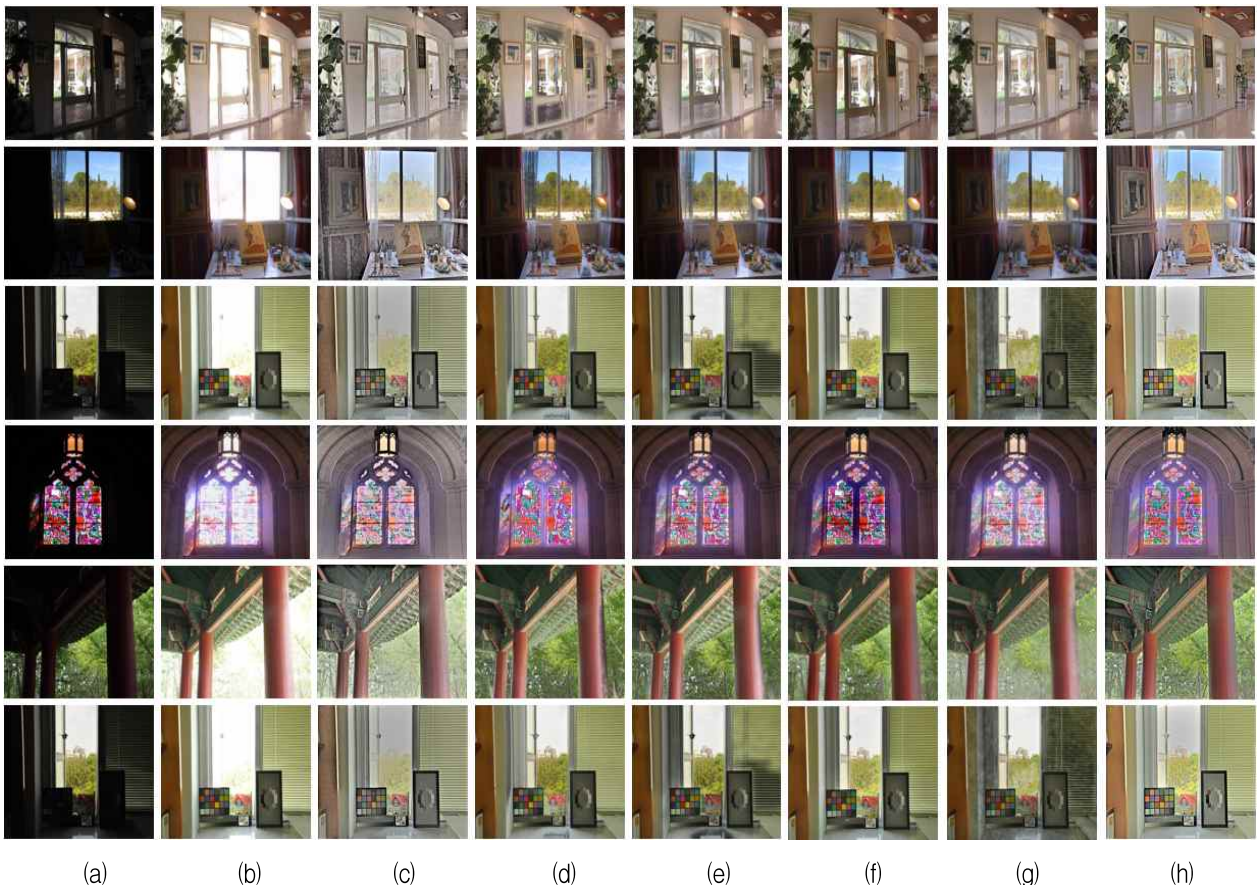


그림 7. SMap 적용 및 기존 방법 결과 비교: (a) 저노출 영상, (b) 고노출 영상, (c) 기존 MSR, (d) EF (e) FMMR (f) DCT-based (g) GRW (h) 제안된 방법

Fig. 7. Comparisons of the results of Smap application and other methods: (a) Low exposure, (b) High exposure, (c) Conventional MSR, (d) EF (e) FMMR (f) DCT-based (g) GRW (h) Proposed method

그림 9(a) 입력영상에서 어두운 처마 부분과 밝은 나무 영역을 잘라내 확대하였다. 해당 영역들에 대해 그림 9(b)의 기존의 MSR과 그림 9(e)의 GRW 방법의 결과 영상들은 나무의 색정보를 보존하지 못하는 것을 볼 수 있다. 그림 9의 EF, FMMR, DCT-based 방법의 결과 영상에서는 나무의 색채가 잘 보이지만 세부 정보가 손실되는 부분이 있고 처마부분에서 경계 정보가 잘 드러나지 않는다. 반면에 제안방법의 결과 영상에서는 처마부분의 경계부

분에 대한 세부정보가 명확히 보이고 나무 영역의 디테일 또한 잘 표현되어 있다.

그림 10에서는 실내와 실외의 입력영상을 통해 합성하였으며 컬러 체커(Color checker)와 회색 계조 패턴이 주요한 부분이다. 그림 10의 EF, FMMR, GRW 및 DCT-based 방법의 결과 영상에서보다 그림 10(g)의 제안방법의 결과에선 컬러 체커에서의 색채 및 회색 계조 패턴의 보상이 우수한 것을 확인할 수 있다.

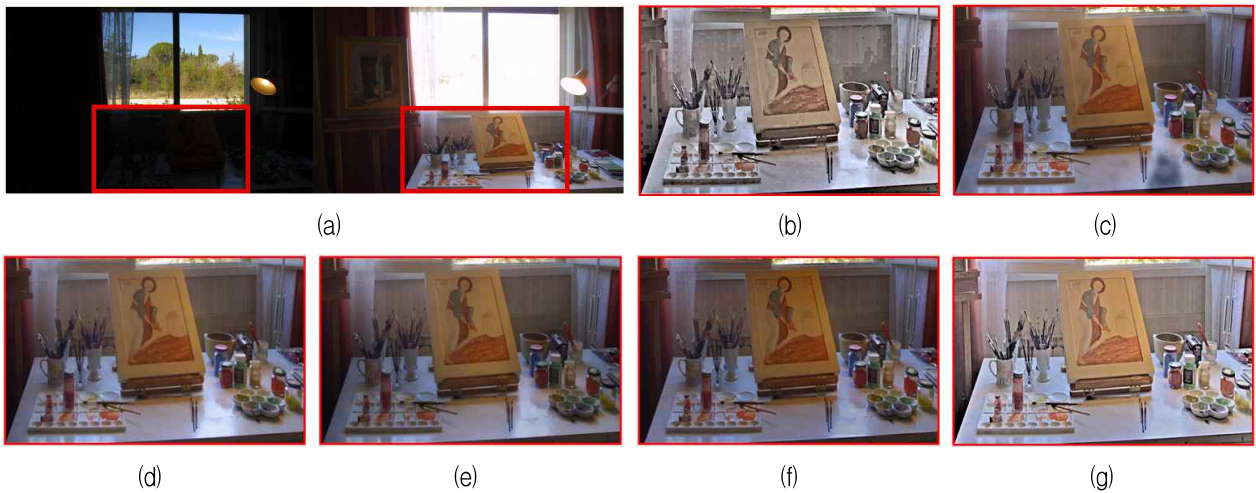


그림 8. 저노출 고노출 영상 합성 방법 비교: (a) 왼쪽: 저노출 영상, 오른쪽: 고노출 영상, (b) 기존 MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) 제안 방법

Fig. 8. Comparisons of LE and HE image fusion methods: (a) Left: low exposure, Right: high exposure, (b) Conventional MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) Proposed method

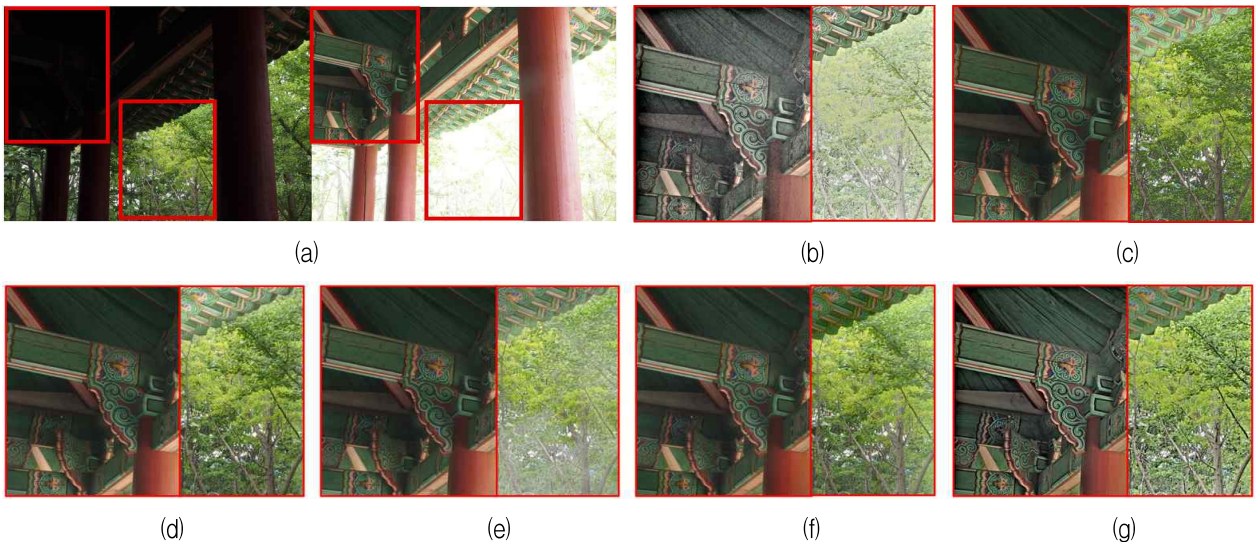


그림 9. 저노출 고노출 영상 합성 방법 비교: (a) 왼쪽: 저노출 영상, 오른쪽: 고노출 영상, (b) 기존 MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) 제안 방법

Fig. 9. Comparisons of LE and HE image fusion methods: (a) Left: low exposure, Right: high exposure, (b) Conventional MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) Proposed method

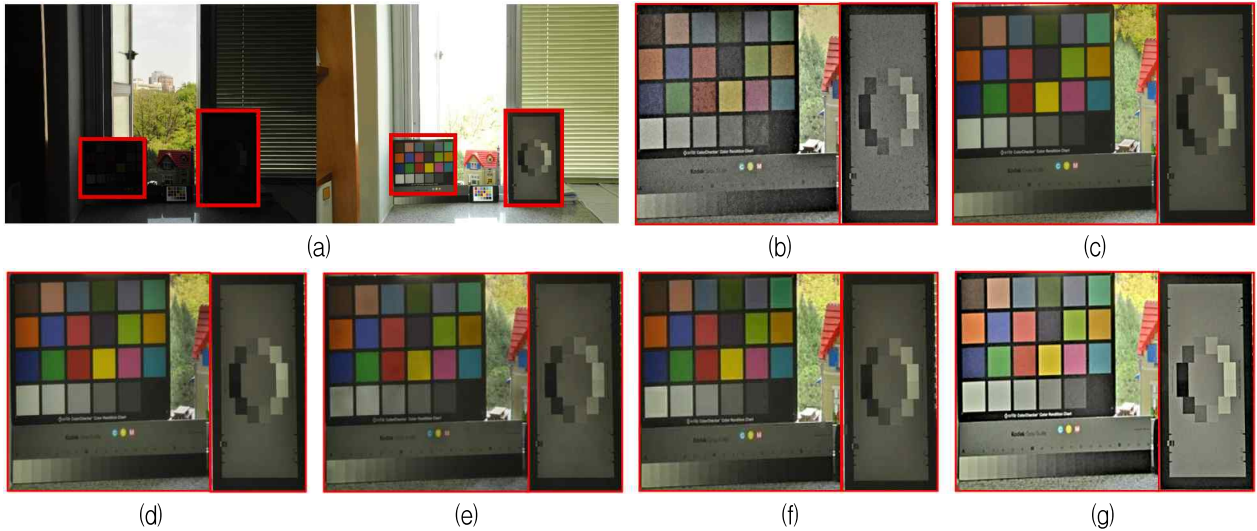


그림 10. 저노출 고노출 영상 합성 방법 비교: (a) 왼쪽: 저노출 영상, 오른쪽: 고노출 영상, (b) 기존 MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) 제안 방법

Fig. 10. Comparisons of LE and HE image fusion methods: (a) Left: low exposure, Right: high exposure, (b) Conventional MSR, (c) EF, (d) FMMR, (e) GRW, (f) DCT-based, (g) Proposed method

본 논문에서 제안한 방법은 특징이 다른 여러 SMap들을 생성하고 스케일별 영상에 각각 적용하여 합성한다. 따라서 전반적으로 영상의 노이즈와 헤일로 현상이 줄어들고 색상 보존 및 어두운 영역의 세부 정보의 보존이 개선된다.

V. 결론

본 논문에서는 다중 스케일 스위칭 맵을 이용한 다중 노출 영상 합성으로 HDR 영상을 생성하는 기법을 제안한다. 제안하는 방법에서는 분산 값에 따라 SSR이 적용된 다중 스케일의 각 특징에 맞는 적절한 스위칭 맵을 적용한다. 각 스케일들은 영상의 세부정보는 유지하지만 노이즈가 있거나, 자연스럽게 강한 헤일로가 존재하는 장단점이 있다. 선별되는 SMap들은 이러한 스케일 별 영상의 특징 중 장점 부분을 잘 합성한다. SMap은 스케일의 특성에 맞게 양방향 필터와 시그모이드 함수의 매개변수 값을 조정하여 경계 선택 능력이 향상되게 만든다. 실험 결과는 제안된 방법이 기존 모델보다 성능이 우수하다는 것을 보여준다. 향후 작업은 맵 스위칭 성능을 향상시켜 다중 노출 영상을 합성하였을 때 좀 더 나은 HDR 영상을 생성하는 것을 목표로 한다.

References

- [1] S. K. Nayar and T. Mitsunaga, "High dynamic range imaging: spatially varying pixel exposures", Proc. IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662), Hilton Head, SC, USA, pp. 472-479, Jun. 2000. <https://doi.org/10.1109/CVPR.2000.855857>.
- [2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs", Proc. of the 24th Annual Conference on Computer Graphics and Interactive Techniques, pp. 369-378, Aug. 1997. <https://doi.org/10.1145/258734.258884>.
- [3] M. E. Rudd, "Lightness computation by the human visual system", Journal of Electronic Imaging, Vol. 26, No. 3, pp. 31209, Jun. 2017. <https://doi.org/10.1117/1.JEI.26.3.031209>.
- [4] T. Watanabe, Y. Kuwahara, A. Kojima, and T. Kurosawa, "An Adaptive Multi-Scale Retinex Algorithm Realizing High Color Quality and High-Speed Processing", Journal of Imaging Science and Technology, Vol. 49, No. 5, pp. 486-497, Sep. 2005.

- [5] Y.-H. An, H.-J. Kwon, and S.-H. Lee, "Multi-Exposure Image Capturing and Synthesis Technique based on a Multi-Sensor Camera", *Journal of KIIT*, Vol. 19, No. 3, pp. 87-102, Mar. 2021. <https://doi.org/10.14801/jkiit.2021.19.3.87>.
- [6] R. Fattal, D. Lischinski, and M. Werman, "Gradient Domain High Dynamic Range Compression", *ACM Transactions on Graphics*, Vol. 21, No. 3, pp. 249-256, Jul. 2002. <https://doi.org/10.1145/566654.566573>.
- [7] N. D. B. Bruce, "ExpoBlend: Information preserving exposure blending based on normalized log-domain entropy", *Computers & Graphics*, Vol. 39, pp. 12-23, Apr. 2014. <https://doi.org/10.1016/j.cag.2013.10.001>.
- [8] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images", *ACM Transactions on Graphics*, Vol. 21, No. 3, pp. 267-276, Jul. 2002. <https://doi.org/10.1145/566654.566575>.
- [9] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images", *Proc. of the 29th annual conference on Computer graphics and interactive techniques*, San Antonio Texas, pp. 257-266, Jul. 2002. <https://doi.org/10.1145/566570.566574>.
- [10] E. H. Land and J. J. McCann, "Lightness and Retinex Theory", *Journal of the Optical Society of America*, Vol. 61, No. 1, pp. 1-11, Jan. 1971. <https://doi.org/10.1364/JOSA.61.000001>.
- [11] S. Wu, J. Xu, Y.-W. Tai, and C.-K. Tang, "Deep High Dynamic Range Imaging with Large Foreground Motions", *Proc. of the European Conference on Computer Vision (ECCV)*, Munich, Germany, Vol. 11206, pp. 120-135, Oct. 2018. https://doi.org/10.1007/978-3-030-01216-8_8.
- [12] P. Debevec, E. Reinhard, G. Ward, and S. Pattanaik, "High dynamic range imaging", *ACM SIGGRAPH 2004 Course Notes*, Los Angeles CA, pp. 14-es, Aug. 2004. <https://doi.org/10.1145/1103900.1103914>.
- [13] L. Qu, S. Liu, M. Wang, and Z. Song, "TransMEF: A Transformer-Based Multi-Exposure Image Fusion Framework Using Self-Supervised Multi-Task Learning", *Proceedings of the AAAI Conference on Artificial Intelligence*, Palo Alto, California USA, Vol. 36, No. 2, pp. 2126-2134, Jun. 2022. <https://doi.org/10.1609/aaai.v36i2.20109>.
- [14] C. Connolly and T. Fleiss, "A study of efficiency and accuracy in the transformation from RGB to CIELAB color space", *IEEE Transactions on Image Processing*, Vol. 6, No. 7, pp. 1046-1048, Jul. 1997. <https://doi.org/10.1109/83.597279>.
- [15] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex", *IEEE Transactions on Image Processing*, Vol. 6, No. 3, pp. 451-462, Mar. 1997. <https://doi.org/10.1109/83.557356>.
- [16] H. Buijs, A. Pomerleau, M. Fournier, and W. Tam, "Implementation of a fast Fourier transform (FFT) for image processing applications", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 22, No. 6, pp. 420-424, Dec. 1974. <https://doi.org/10.1109/TASSP.1974.1162620>.
- [17] Z. Rahman, D. J. Jobson, and G. A. Woodell, "Multi-scale retinex for color image enhancement", *Proc. of 3rd IEEE International Conference on Image Processing*, Lausanne, Switzerland, pp. 1003-1006, Sep. 1996. <https://doi.org/10.1109/ICIP.1996.560995>.
- [18] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure Fusion: A Simple and Practical Alternative to High Dynamic Range Photography", *Computer Graphics Forum*, Vol. 28, No. 1, pp. 161-171, Mar. 2009. <https://doi.org/10.1111/j.1467-8659.2008.01171.x>.
- [19] S. Li and X. Kang, "Fast multi-exposure image fusion with median filter and recursive filter", *IEEE Transactions on Consumer Electronics*, Vol. 58, No. 2, pp. 626-632, May 2012. <https://doi.org/10.1109/TCE.2012.6227469>.

[20] G.-Y. Lee, S.-H. Lee, and H.-J. Kwon, "DCT-Based HDR Exposure Fusion Using Multiexposed Image Sensors", *Journal of Sensors*, Vol. 2017, pp. 1-14, Dec. 2017. <https://doi.org/10.1155/2017/2837970>.

[21] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized Random Walks for Fusion of Multi-Exposure Images", *IEEE Transactions on Image Processing*, Vol. 20, No. 12, pp. 3634-3646, Dec. 2011. <https://doi.org/10.1109/TIP.2011.2150235>.

2009년 8월 ~ 2017년 7월 : 경북대학교 IT대학 전자공학부 연구초빙교수
 2018년 3월 ~ 현재 : 경북대학교 IT대학 전자공학부 교수
 관심분야 : Color Image Processing, Color Appearance Model, Color Management, HDR 영상처리, 영상융합, 인공지능영상처리

저자소개

고 영 호 (Young-Ho Go)



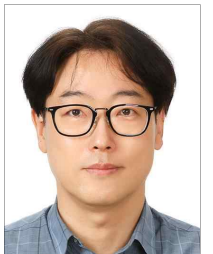
2024년 2월 : 경북대학교
 전자공학부(공학사)
 2024년 2월 ~ 현재 : 경북대학교
 대학원 전자전기공학부 석사과정
 관심분야 : HDR 영상처리, Multi exposure image fusion, Deep Learning, 컴퓨터 비전, Color Image Processing, 영상 융합

이 승 환 (Seung-Hwan Lee)



2016년 2월 : 한양대학교
 전자통신공학과(공학사)
 2018년 2월 : 한양대학교 전기
 전자 컴퓨터공학과(공학석사)
 2022년 9월 ~ 현재 : 경북대학교
 전자전기공학부 박사과정
 관심분야 : 컴퓨터 비전, 인공지능, 자율주행, Color Image Processing, 물체 인식

이 성 학 (Sung-Hak Lee)



1997년 2월 : 경북대학교
 전자공학과(공학사)
 1999년 2월 : 경북대학교
 전자공학과(공학석사)
 1999년 2월 ~ 2004년 6월 : LG
 전자 영상제품연구소 선임연구원
 2008년 2월 : 경북대학교
 전자공학과(공학박사)