

관심 영역에 기반한 저심도 이미지 후처리 구현 방법

*이유경, **허정환, 정제창

*한양대학교 융합전자공학부, **한양대학교 전자컴퓨터통신공학과

*iam422kr@naver.com, **hur881122@hanyang.ac.kr, jjeong@hanyang.ac.kr

Depth-of-Field Image Post-Processing Method Based on Visual Attention

*Lee Yu-Kyeong, **Heo Jeong-Hwan, Jeong Jechang

*Department of Electronic Engineering, Hanyang University

**Department of Electronics and Computer Engineering, Hanyang University

요약

본 논문에서는 단일 이미지의 관심 영역에 기반한 저심도 후처리 방법을 제안한다. 저심도 이미지란 사진에서 초점이 선명하게 포착되는 깊이의 범위가 좁은 이미지를 말한다. 기존의 광학적 특성을 이용한 저심도 이미지를 만드는 과정은 물리적인 구조 설계비용 문제가 존재한다. 또한, 이미지의 후처리 보정을 통한 방법은 이미지상의 사물 깊이 정보를 알기 어렵기 때문에 이미지의 심도를 후처리하기 어려웠다. 이에 따라 본 논문에서는 슈퍼 픽셀 군집화 방법을 통해 관심 영역을 찾고, 이에 기반하여 관심 영역이 부각될 수 있는 저심도 후처리 방법을 제안한다. 제안하는 후처리 방법은 슈퍼픽셀 군집화 방법을 통해 관심 영역을 설정하여 배경 영역을 분리하고 블러 과정을 수행한다. 관심 영역을 제외한 부분을 확장 한 뒤 배경 블러를 거치기 때문에 후광효과가 현저히 줄어든 저심도 효과가 적용된 이미지를 얻을 수 있었고 MSRA-1000 데이터 셋 이미지에서 우수한 주관적 화질 결과를 보였다.

1. 서론

최근 스마트폰과 같은 모바일 기기들의 급속한 발전으로 인해 더 작은 카메라로 눈으로 보는 것과 같은 저심도 사진을 촬영하는 방법에 대한 기술 수요가 증가하고 있다. 심도란, 초점이 선명하게 포착되는 깊이의 범위를 말한다. 즉, 고심도 영상이란 초점이 맞는 깊이의 범위가 넓은 영상을 말하고, 그 범위가 좁을 때는 저심도 영상을 말한다. 심도는 렌즈의 초점 거리와 렌즈 구경 그리고 카메라와 피사체의 거리에 의해 결정된다. 그러나 렌즈나 촬영 환경의 조건이 만족하지 않을 경우를 비롯해서 이러한 광학 특성을 이용한 심도 효과에는 물리적 비용적 한계가 존재한다. 또한, 단일 이미지 보정을 통한 후처리 방법은 이미지상의 사물 깊이 정보를 알기 어렵기 때문에 이미지의 심도를 후처리 하기는 어려웠다. 사람은 인지 기반 시각 특성 HVS(Human Visual System)[1]에 따라 복합적인 집중 과정을 통해 ROI(Region of Interest)를 판단한다. 이때, 심리적인 요인, 시선의 집중 상태 등에 따라 우리의 눈은 심도가 다른 이미지를 느끼게 된다. 또한, 심도 효과를 통해 집중을 유도할 수 있다. 이에 따라 본 논문에서는 관심 영역 (Region of Interest-ROI) 기반 심도 후처리를 통해 사람의 눈으로 바라보는 것과 가까운 심도 효과 후처리 기술을 제안한다.

단일 이미지의 심도 후처리 방법의 구현에는 먼저 이미지 안에서 관심 영역을 추출해 내는 과정이 필요하다. 이미지의 두드러진 정도 (Saliency)를 파악하여 관심 영역 추출 방법이 사용된다. 이를 구현하는 데는 크게 3가지 종류의 방법이 사용된다. 첫째로 이미지상의 공간적 고주파 영역을 검출하는 전통적인 방법[2]가 있다. 이 방법은 주파

수 변환의 특성상 배경에 복잡한 패턴의 무늬가 있을 경우에 취약하다. 둘째로 슈퍼픽셀에 기반한 군집화 방법[3][4][5]이 있다. 이 방법은 계산 속도가 빠르고 비선형적인 연산 특성에 의해 관심 영역을 분리해 내는데 높은 성능을 보인다. 셋째로 최근에는 CNN을 통한 딥네트워크의 지도학습을 활용하여 사람의 관심 영역을 네트워크가 학습하는 방법[6]을 사용한다. 훈련된 딥러닝 네트워크는 주요한 물체를 인식해 낼 뿐만 아니라 비선형적인 pooling 연산으로 이미지의 상황을 인식하여 관심 영역을 추출해내는 성능이 뛰어나지만 연산량이 매우 높다는 단점이 있다. 본 논문에서는 계산 복잡도가 낮으면서도 분류 성능이 뛰어난 슈퍼픽셀에 기반한 군집화 방법을 통해 Saliency를 구하는 방법을 사용하였다.

본 논문에서는 관심 영역 기반의 심도처리를 통해 눈으로 보는 것과 유사한 이미지를 만들어 내었고, 제안하는 방법은 MSRA-1000[2] 데이터 셋에서 주관적 화질평가 결과 눈으로 보는 것과 같은 자연스러운 이미지를 얻을 수 있음을 보였다.

2장에서는 기존의 Saliency 기반의 관심 영역 추출 알고리즘과 3장에서는 제안하는 블러 이미지 알고리즘, 4장에서는 실험 결과 및 분석, 마지막으로 5장에서는 향후 연구 방향을 기술한다.

2. 기존의 알고리즘

2.1 슈퍼픽셀 기반의 군집화

본 논문에서는 관심 영역을 설정하기 위하여 슈퍼픽셀에 기반한

군집화 과정을 사용하였다. 관심 영역을 설정하기에 앞서 군집화 과정을 수행하면 고해상도 이미지의 해상도를 줄이는 효과를 얻을 수 있기 때문에 관심 영역을 구하는 과정에서의 연산량을 줄일 수 있으며, 이웃한 군집의 특성에 따라 관심 영역을 결정하기에도 유리하다. 군집화 알고리즘에는 여러 가지 종류가 있지만 simple linear iterative clustering (SLIC)[7]을 통한 방법은 연산 속도 및 결과 면에서 모두 우수하므로 본 논문의 아웃포커스 구현에도 적절한 결과물을 얻을 수 있었다. SLIC의 주요 연산 과정은 다음과 같다. 초기에 설정된 군집의 대푯값과 대상 픽셀 간의 거리 D 를 다음과 같은 수식으로 계산하여 군집의 대푯값을 다시 갱신한다.

$$\begin{aligned}
 d_{color} &= \sqrt{(R_k - R_i)^2 + (G_k - G_i)^2 + (B_k - B_i)^2} \\
 d_{spatial} &= \sqrt{(x_c - x)^2 + (y_c - y)^2} \\
 D &= d_{color} + d_{spatial}
 \end{aligned}
 \tag{1}$$

이 과정을 반복 계산하여 수렴함으로써 인접한 특성의 픽셀들이 군집으로 형성된다. 군집의 이동이 일정 수치 이하일 때까지 반복하면 최종적으로 그림 1과 같은 결과물을 얻을 수 있다.



그림 1. SLIC 알고리즘의 군집화 과정

2.2 매니폴드 랭킹을 사용한 saliency map

사람의 눈으로 사물을 인식할 때 두드러진 정도를 나타내는 saliency map은 매니폴드 랭킹[4] 알고리즘을 사용하여 구할 수 있다. 매니폴드 랭킹 알고리즘은 구글의 웹페이지 중요도 순위 계산 알고리즘으로 사용되는 PageRank[8]로도 잘 알려져 있다. 앞서 구한 군집화 결과를 그림 2와 같이 그래프 구조로 나타낸 뒤 군집의 순위를 바탕으로 saliency map을 구한다.

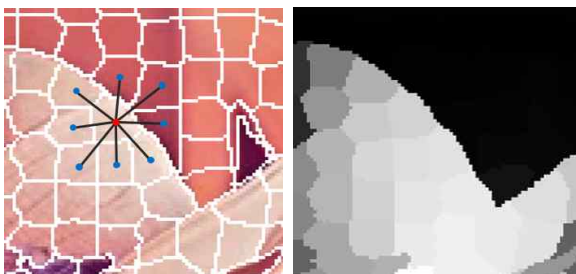


그림 2. 인접 군집간의 그래프 연결 방법과 saliency map

saliency map은 (0, 255)의 정수 범위로 출력되지만 이 출력을 관심 영역과 배경 영역을 구분 짓는 마스크로 사용할 경우, 관심 영역에서 조금만 벗어나더라도 이미지가 흐려지게 되므로 주관적 화질 저하를 일으키게 된다. 따라서 saliency map의 값을 두 배로 곱하여 수치를 벗어나는 값은 버리는 trim을 수행한 후 (0, 1)의 실수 범위로 바꾼 값을 관심 영역으로 사용한다. 이와 같은 과정을 통해 관심 영역에 해당하는 물체는 선명하게 유지할 수 있다.

3. 제안하는 알고리즘

이번 장에서는 2장에서 설명한 관심 영역을 통해 관심 영역 이외의 배경 영역은 흐리게 하여 이목이 집중될 수 있는 방법을 제안한다. 심도 후처리 과정을 구현하는데 생각해 볼 수 있는 가장 간단한 방법은 원본 이미지와 블러된 이미지를 관심 영역 지도 (Region of Interest map)의 값에 따라 적절히 섞는 방법이다. 하지만 원본과 블러 이미지를 섞는 단순 구현을 할 경우 그림 3의 (c)처럼 관심 영역과 블러된 배경 영역 경계면이 번져 나오는 듯한 후광 효과가 발생할 수 있다. 이는 관심 영역의 주요 이미지가 배경과의 경계면에서 가우시안 블러 연산을 통해 범위가 확장되어 배경 영역까지 영향을 미쳤기 때문이다. 이는 관심 영역을 제외한 후 배경 영역을 padding하는 방법을 사용하면 후광 효과를 줄일 수 있다.

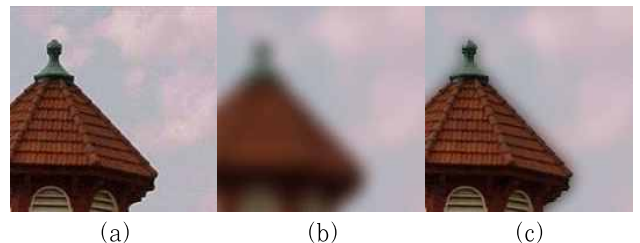


그림 3. 관심영역과 배경영역의 경계면의 후광효과 발생과정 (a) 원본이미지 (b) 블러 이미지 (c) 관심영역 지도에 따라 (a)와 (b)를 더해 후광효과가 발생한 이미지

제안하는 알고리즘의 전체 흐름도는 다음 그림 4와 같으며 후광 효과 제거를 위한 배경 이미지 패딩 과정과 최종 저심도 효과를 적용한 이미지를 위한 배경 이미지 블러 과정의 두 부분으로 나누어 설명한다.

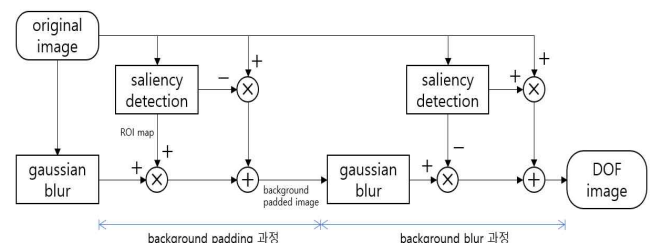


그림 4. 제안하는 알고리즘의 흐름도

A. 배경 이미지 패딩 과정

그림 4의 흐름도에 따라, 입력된 원본 이미지에서 블러된 이미지를 생성한다. 초점이 흐려진 이미지는 점 확산 함수(Point-Spread Function, PSF)[9] 모델의 연산 결과로 생각 할 수 있다. 따라서 블러 이미지를 만드는데 이상적인 가우시안 함수를 통해 심도를 조절효과를 구현하였다. 2차원 입력 이미지의 좌표를 x, y 라고 할 때, 가우시안 이미지 $G(x, y)$ 는

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

로 표현할 수 있다. 여기서 σ 는 표준편차를 의미하며 블러의 정도를 조절하는 역할을 한다. 본 논문에서는 이미지 패딩을 위한 블러로는 15를, 최종 배경 이미지 블러를 위해서는 7의 값을 사용하였다. 15의 시그마 값으로 가우시안 함수로 생성한 블러 이미지와 원본 이미지를 관심 영역 지도를 통해 알파 블렌딩으로 섞는다. 블러된 이미지를 I_{blur} , 원본 이미지를 $I_{original}$ 라 할 때, 알파 블렌딩 결과물인 패딩된 배경이미지는 다음과 같이 표현할 수 있다.

$$I_{background_padded} = (1 - \alpha) \circ I_{original} + \alpha \circ I_{blur} \quad (3)$$

여기서 \circ 는 element wise product이고, α 는 관심 영역 지도의 출력 값으로 두 이미지간의 혼합 비율을 결정하는 역할을 한다. α 는 관심 영역일 때는 1, 관심 영역이 아닐 때는 0의 실수 값을 가진다. 알파 블렌딩을 이용하면 자연스럽게 블러된 패딩된 배경 이미지를 얻을 수 있다. 이렇게 관심 영역과 배경 영역의 블러 이미지를 섞기 전에 배경 영역을 먼저 블러해주는 이유는 앞서 설명하였듯이 그림 3의 (c)에서 보였던 후광 효과를 방지해주기 위함이다. 원본과 블러된 이미지만을 이용해 알파 블렌딩 하는 경우, 관심 영역의 외곽선이 번져 보이는데 이러한 효과를 후광효과라 한다. 후광효과는 배경 이미지를 먼저 블러시켜 관심 영역 안으로 먼저 들어가게 만듦으로써 이 효과를 억제할 수 있다. 여기까지가 이를 위한 알고리즘 흐름도 왼쪽의 과정이었다.

saliency detection 블록은 2장에서 구한 saliency map을 계산하는 과정이다. 이후 구해진 saliency map값을 두 배로 곱하고 범위를 벗어나는 값은 버리는 방법으로 trim을 하였다. Saliency map은 (0, 255) 범위의 값을 가지기 때문에 관심 영역에서 벗어날수록 점점 흐려지도록 만들 경우 관심 영역 인근의 화질 저하를 불러 올 수 있으므로 이와 같은 trim 과정을 통해 128 이상의 Saliency 값에 대해서는 원본을 보여주기 위함이다.

B. 배경 이미지 블러 과정

패딩 된 배경 이미지는 심도 효과를 위해 다시 한 번 가우시안 블러를 거친다. 패딩된 배경 이미지를 $I_{background_padded}$ 라 하자. 가우시안 블러를 거쳐 흐려진 이미지를 베이스 이미지 I_{base} 라 할 때 최종 후처리 결과는 다음과 같다.

$$I_{DOF_image} = \alpha \circ I_{original} + (1 - \alpha) \circ I_{base} \quad (4)$$

이와 같은 배경 패딩을 통한 블러 과정을 수행하면, 관심 영역의 주요 물체의 외곽선이 번져 보이는 후광효과와 발생을 억제한 저심도 효과 이미지를 얻을 수 있다.

4. 실험 결과 및 분석

제안하는 방법을 평가하기 위해 MSRA-1000 데이터 셋을 사용하여 관심 영역 분류를 수행하고 제안하는 알고리즘을 적용하여 주관적 화질 비교 실험을 수행하였다. 실험 환경은 Intel Pentium CPU에서 MATLAB R2019a 버전을 사용하였다.

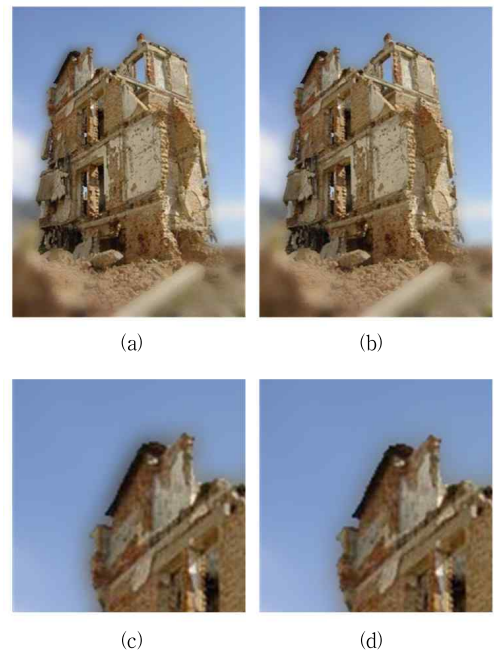


그림 5. 제안하는 알고리즘의 후광효과 화질 비교
(a) 단순 합으로 구현한 결과 (b) 제안하는 배경 패딩 방법을 사용 후 혼합 결과

그림 5는 원본과 블러 이미지를 섞는 단순 구현을 한 경우의 결과와 배경 이미지 패딩 후 원본과 블러 이미지를 섞어 구현한 결과이다. (a)는 원본과 블러 이미지를 섞는 단순 구현을 한 경우의 결과이고, (b)는 배경 이미지 패딩 후 원본과 블러 이미지를 섞어 구현한 결과이다. (a)의 모서리 부분을 확대한 (c)의 그림을 보면 관심 영역과 배경 영역의 경계 부분 색상이 타들어가는 것처럼 후광 효과가 발생하였음을 알 수 있다. 반면, 제안한 알고리즘을 이용하여, 배경 이미지를 패딩한 후 원본과 블러 이미지를 섞어, 구현한 (b)의 모서리 부분을 확대한 (d)의 그림은 후광 효과의 발생이 현저히 줄어들었음을 확인할 수 있다. 이렇듯 후광 효과를 억제하여 구현한 (b)의 결과가 (a)와 비교했을 때 훨씬 자연스럽게 깨끗한 저심도 효과가 적용된 사진을 얻을 수 있음을 볼 수 있다.

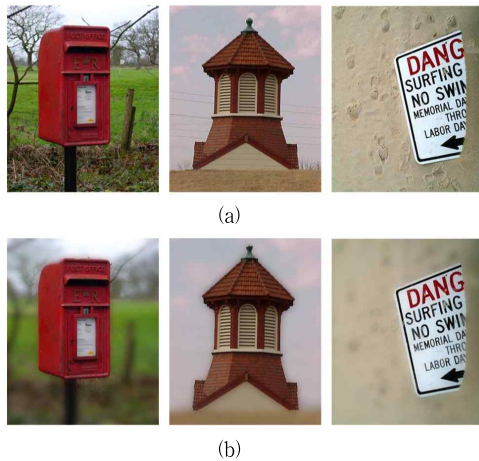


그림 6. 제안하는 알고리즘의 주관적 화질 비교

(a) 원본 MSRA-1000 데이터셋 (b) 제안하는 알고리즘의 저심도 후처리 효과 적용 결과

그림 6의 (b)는 각각의 단일 이미지 (a)에 제안하는 알고리즘의 저심도 후처리 효과를 적용한 결과의 이미지이다. 먼저, 그림에서 볼 수 있듯이 슈퍼 픽셀에 기반한 군집화 방법을 통한 Saliency 추출은 사람이 인지하는 관심 영역에 맞게 구해졌음을 확인할 수 있다. 그리고 이렇게 추출된 관심 영역에 따라 저심도 효과가 후광효과 등의 결함이 없이 적용되었음을 확인할 수 있다. 배경 이미지 페딩 과정을 이용하여 관심 영역과 배경 영역의 경계 부분이 서로 영향을 미치지 않아 자연스럽게 이어짐으로써 보다 높은 수준의 저심도 효과 이미지를 얻을 수 있었다. 그러나 몇몇 복잡한 패턴의 이미지에서 관심 영역과 관심 영역이 아닌 부분의 추출이 제대로 이루어지지 않아 저심도 후처리 성능이 떨어지는 경우도 있었다. 이는 관심영역을 설정하는 과정에서 생긴 문제이므로 추후 딥러닝 등을 활용한 관심영역 설정 방법을 활용하면 더 향상된 결과를 얻을 것으로 기대된다.

5. 결론

기존의 광학적 특성 등을 이용한 저심도 효과는 물리적 비용적 한계가 존재했고, 단일 이미지 보정을 통한 심도 후처리 방법은 이미지상의 사물 깊이 정보를 알기 어렵기 때문에 이미지의 심도를 후처리하기는 어려웠다. 이에 본 논문에서는 슈퍼 픽셀 군집화 방법을 통해 단일 이미지의 관심 영역을 찾고, 이에 기반한 저심도 후처리 방법을 제안했다. 제안하는 방법은 배경 이미지 페딩을 통해 효과적으로 후광효과를 억제하며 관심 영역에 저심도 효과가 적용된 이미지를 얻을 수 있었고 MSRA-1000 데이터 셋 이미지에서 대체로 우수한 주관적 화질 결과를 보였다. 그러나 몇 데이터 셋 이미지에서 관심 영역 추출이 제대로 이루어지지 않아 저심도 후처리 효과 이미지가 자연스럽게 않은 경우가 있었다. 이는 실험 환경에서 처리할 수 있는 연산량의 문제로 관심 영역 추출을 위해 정확성이 뛰어난 CNN을 통한 딥네트워크의 지도학습을 활용 방법 대신 슈퍼 픽셀 기반의 방법을 사용했기 때문으로 보인다. 본 연구의 이러한 점을 보완한 관심 영역 추출 방법을 다르게 한 추가 연구가 필요할 것으로 보인다.

참고 문헌

- [1] L. Itti, and C. Koch, "A Saliency-Based Search Mechanism for Overt and Covert Shifts of Visual Attention," *Vision Research*, vol. 40, pp. 10–12, Jun. 2000.
- [2] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned Salient Region Detection," *IEEE International Conference on Computer Vision and Pattern Recognition(CVRR)*, pp. 1597–1604, Jun. 2009.
- [3] Z. Liu, X. Zhang, S. Luo, and O. Le Meur, "Superpixel-Based Spatiotemporal Saliency Detection," *IEEE transactions on circuits and systems for video technology*, vol. 24, no. 9, pp. 1522–1540, Feb. 2014.
- [4] C. Yang, L. Zhang, H. Lu, X. Ruan, and M. Yang, "Saliency Detection via Graph-Based Manifold Ranking," *IEEE conference on computer vision and pattern recognition*, pp. 3166–3173, 2013.
- [5] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum, "Learning to Detect a Salient Object," *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 33, no. 2, pp. 353–367, 2011.
- [6] G. Li, and Y. Yu, "Visual Saliency Detection Based on Multiscale Deep CNN Features," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5012–5024, Aug. 2016.
- [7] K. Kim, D. Zhang, M. Kang, and S. Ko, "Improved Simple Linear Iterative Clustering Superpixels," *IEEE International Symposium on Consumer Electronics (ISCE)*, pp. 259–260, Jun. 2013.
- [8] W. Xing, and A. Ghorbani, "Weighted PageRank Algorithm," *In Proceedings. Second Annual Conference on Communication Networks and Services Research*, pp. 305–314, May 2004.
- [9] K. Rossmann, "Point Spread-Function, Line Spread-Function, and Modulation Transfer Function: Tools for the Study of Imaging Systems," *Radiology*, vol.93, no. 2, Aug. 1969.