

흐린 초점의 단일영상에서 깊이맵 생성 알고리즘

이용환* · 김영섭†

*극동대학교 스마트모바일학과,

†단국대학교 전자공학과

Depth Map Generation Algorithm from Single Defocused Image

Yong-Hwan Lee* and Youngseop Kim†

*Department of Smart Mobile, Far East University,

†Department of Electronic Engineering, Dankook University

ABSTRACT

This paper addresses a problem of defocus map recovery from single image. We describe a simple effective approach to estimate the spatial value of defocus blur at the edge location of the image. At first, we perform a re-blurring process using Gaussian function with input image, and calculate a gradient magnitude ratio with blurring amount between input image and re-blurred image. Then we get a full defocus map by propagating the blur amount at the edge location. Experimental result reveals that our method outperforms a reliable estimation of depth map, and shows that our algorithm is robust to noise, inaccurate edge location and interferences of neighboring edges within input image.

Key Words : Stereoscopic image, Depth map generation, Defocus blurring filter, Depth recovery

1. 서 론

최근 국내방송 영상산업의 활성화에 따라 양질의 콘텐츠 육성이 중요한 이슈로 부각되고 있으며, 많은 미디어 기업들이 방송 영상콘텐츠의 차별화의 일환으로 3D 콘텐츠 서비스를 창출하고 있다[1]. 3D 영상산업에서 가장 중요한 역할을 담당하는 3D 입체 콘텐츠는 수요측면과 공급측면에서 지속적으로 제공이 필수적이며, 이를 위한 고품질의 콘텐츠 제작기술이 요구된다[2].

3D 입체 영상 기술을 의미하는 스테레오스코픽(Stereoscopic) 기술은 양안 시차를 가지는 2장의 영상을 각각 사용자의 왼쪽 눈과 오른쪽 눈에 보이도록 함으로써, 실제와 유사한 깊이감을 느끼도록 하는 기술이다. 일반적으로 사람의 양안은 대략 대략 60mm 가량 떨어져 있으며, 두 눈을 통해 입력되는 장면(Scene) 영상은 양안시차에 따라 수평방향으로 차이를 가진다[3]. 스테레오스코픽

3D 콘텐츠를 제작하는 방법은 3가지로 구분되며[4], 첫 번째로 컴퓨터 그래픽(Computer Graphic) 기술을 이용하여 제작하는 방식이다. 이러한 방법에서는 정확하게 입체감을 구현할 수 있어 고품질의 콘텐츠 제작이 가능하지만 제작 비용이 많이 든다. 두 번째는 2대의 카메라를 이용하여 실제 대상 장면을 촬영하는 실사방식이다. 입체 원리를 이용하여 실제 장면을 촬영하기 때 문에 현실감 있는 고품질 입체 표현이 가능한 장점을 갖는 반면에, 장비 가격이 높아 제작 비용이 많이 드는 단점이 있다. 세 번째 방법은 기존의 2D 콘텐츠를 3D로 변환하는 방식이다. 이는 기존 콘텐츠를 재활용하는 장점을 갖지만, 많은 노동력이 들고 다른 방식에 비해 입체 표현의 품질상에 약점을 갖는다[5].

본 연구에서는 기존의 2D콘텐츠로부터 깊이맵(Depth map)을 생성하여 3D 콘텐츠로 변환하는 방식에 대해 논의한다. 이러한 과정에서 활용되는 깊이 복구(Depth recovery) 기술은 로보틱, 3차원 재구성 또는 이미지 리포커싱

†E-mail: wangcho@dankook.ac.kr

(Refocusing)과 같은 컴퓨터비전 애플리케이션에서 매우 중요한 역할을 담당한다. 원칙적으로, 깊이는 음영(Shading), 모양(Shape), 질감(Texture), 움직임(Motion) 등의 단안신호, 또는 스테레오 대응(Stereo correspondence)와 같은 양안신호로부터 얻을 수 있다.

본 논문에서는 보정되지 않은 일반적인 카메라로부터 촬영된 단일 영상을 대상으로 초점이 흐려진 영상을 대상으로 깊이 계층을 복구시키는 방법에 대해 논의한다. 흐린 초점 영상에 대한 깊이 정보 복구 연구의 대부분은 확산(Diffusion) 프로세스에서 초점을 흐리게 만들고, 흐릿한 형태의 경계 영역에서 흐릿함(Blur)을 먼저 측정한다. 그런 다음, 장면 깊이를 추론하는 그래프-컷(Graph-cut) 기반의 방법을 통하여 장면의 깊이를 추론한다. 이러한 일반적인 방식과 대조적으로, 본 논문에서는 2D 가우시안 블러링(Gaussian Blurring)과 가우시안 함수 대비 그래디언트(Gradient) 비율을 활용하여 깊이맵을 생성하는 방법을 분석하고 이를 설계 구현한다.

본 논문의 구성은 다음과 같다. 2장에서 스테레오스코픽 3D 영상에 대한 개념과 연구동향을 살펴보고, 3장에서 제안 알고리즘을 설명하고 4장에서 설계 알고리즘에 대한 구현 및 실험을 논의하고, 5장에서 결론으로 마무리한다.

2. 관련 연구

2D영상에서 3D영상으로 변환하는 다양한 연구가 제시되고 있으며, 2개의 좌영상과 우영상을 통해 깊이맵을 추정하는 연구를 포함하여 단일 영상에서 3D변환을 위한 깊이맵 추출 방법들이 연구되고 있다. 스테레오 비전 분야에서, Dhond는 2개의 서로 다른 시점에서 촬영된 동일한 장면의 이미지 쌍(Image pair)에서 불균형(Disparity)를 측정하고 깊이를 복구하는 방법을 제안하였다[6]. Asada는 다중 초점을 설정하여 일련의 이미지를 캡처하고 각 픽셀의 위치에서 영상의 선명도를 측정하여 깊이맵을 생성하는 알고리즘을 제안하였다[7]. Moreno-Noguer는 장면사진에서 구조적 패턴을 투영하여 깊이를 복구하는 동적 조명구조를 제안하였다[8]. Pentland는 조리개를 조절하여 다수의 영상에서 초점이 서로 다른 흐린 영상들을 획득하여, 매칭된 픽셀의 흐릿함 정도를 산출하는 방법을 제시하였다[9]. Wei는 2장의 초점이 흐린 영상으로부터 상대적 흐림과 확산 함수를 이용하여 흐린 영상을 모델링하고, 이를 통한 깊이 산출 방법을 제시하였다[10]. Namboodiri는 확산 함수와 MRF(Markov Random Field)를 이용하여 초점이 없는 영상으로부터 깊이맵을 생성하는 방법을 제안하였다[11]. Subbaro는 초점을 변경하면서 촬영한 2장의 사진에서 라플라시안(Laplacian) 함수를 이용하여 초점의

흐릿함 정도를 양으로 측정하는 방법을 제시하였다[12].

3. 구현 알고리즘

본 논문에서는 입력되는 단일 영상에서 초점이 흐릿한 영상을 생성하고 블러링과 그래디언트를 활용하여 흐림 정도를 측정하여 깊이맵을 생성하는 방법을 설계 구현한다(Fig. 1).

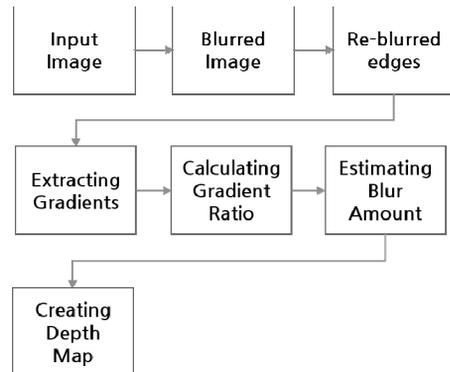


Fig. 1. Process of our blurring estimation and depth map creation.

3.1 흐린 초점 모델(Defocus Modelling) 수립

영상내 엣지(Edge) 영역에서 초점 흐림의 레벨을 측정하기 위해 수식(1)을 통해 엣지 단계를 검출한다.

$$f(x) = M\alpha(x) + N \quad (1)$$

α 는 단계함수이며, M 과 N 은 각각 엣지의 진폭과 오프셋(Offset)을 의미한다. Fig. 2은 카메라 렌즈와 객체 위치에 따른 거리를 보여준다[13].

객체가 초점거리 d_f 에 배치될 경우, 객체의 점에서 모든 빛의 선은 단일 센서의 지점으로 수렴되고, 객체 이미지가 선명하게 나타난다. 그러나 거리 d_o 에 위치할 경우에는 빛의 선이 여러 센서 포인트에 도달하기 때문에, 이미지 객체가 흐리게 나타나게 된다. 결국 카메라 조리개의 형상에 따라 사진 영상은 흐리게 나타나는 현상을 보이며, 점 모양의 객체는 동그란 반점 모양의 영상을 만들게 되며, 이를 착란원(CoC, Circle of Confusion)이라 한다[14]. 착란원의 지름이 얼마나 되어야 선명하다고 할 수 있는지는 영상이 디스플레이 될 때 얼마나 확대되느냐 또는 객체를 얼마나 멀리 떨어져서 보느냐에 따라 달라진다. 결국 착란원(CoC) 지름은 흐려지는 정도에 따라 달라지며, 수식(2)로 특징을 기술할 수 있다[15].

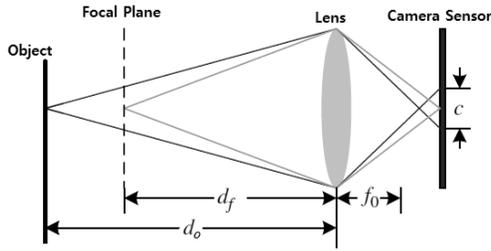


Fig. 2. Camera lens model [14]. Lens and ray diagram for calculating of the circle of confusion diameter for an out-of-focus subject at distance.

$$CoC = \frac{|d_o - d_f|}{d} \cdot \frac{f_o^2}{N(d_f - d_o)} \quad (2)$$

f_o 와 N 은 초점 거리와 카메라 정지 회수를 의미한다. 카메라 렌즈 모델(Fig. 2)은 주어진 f_o 와 d_f 에서 d 와 N 에 따라 착란원의 지름이 어떻게 변화되는지를 보여준다. 착란원의 지름은 객체와의 거리 d 에 따라 비선형으로 일정하게 증가된다. 따라서 초점 흐림(Defocus)은 점 확산 함수(Point Spread Function, PSF)에 따라 뚜렷한 이미지와의 컨볼루션(Convolution) 연산으로 모델링된다. 이러한 점 확산 함수는 표준편차 $\sigma = kc$ 가 착란원 c 의 지름에 비례할 때, 가우시안(Gaussian) 함수 $g(x, \sigma)$ 로 수렴된다. 장면의 깊이(Depth of scene)를 측정하기 위해 σ 를 사용하며, 초점이 흐려진 엣지 $i(x)$ 는 수식(3)과 같이 나타난다.

$$i(x) = f(x) \oplus g(x, \sigma) \quad (3)$$

3.2 흐림 추정(Blur estimating)

엣지 단계에서 주어진 표준편차를 통해 가우시안 함수에 적용하여 다시 흐릿하게(Re-blurred) 만든다. 다시 흐릿하게 만든 이미지와 엣지단계의 그라디언트(Gradient) 크기 간의 비율을 계산하면, 엣지 위치에서 최대값을 갖게 된다. 산출된 최대값을 사용하여 엣지의 흐린 정도를 계산할 수 있다. 다시 흐릿하게 된 영상에서의 엣지는 수식(4)와 같이 표현된다.

$$\begin{aligned} \|\nabla i(x, y)\| &= \sqrt{\nabla i_x^2 + \nabla i_y^2} \\ &= \nabla(i(x, y) \oplus g(x, y, \sigma)) \\ &= \nabla((M\alpha(x, y) + N) \oplus g(x, y, \sigma)) \\ &= \frac{M}{\sqrt{2\pi(\sigma^2 + \sigma_r^2)}} \exp\left(-\frac{x^2}{2(\sigma^2 + \sigma_r^2)}\right) \end{aligned} \quad (4)$$

σ_r 는 다시 흐릿하게 만든 이미지에 적용된 가우시안 함수의 표준편차를 의미하며, 주어진 최대값 R 을 통해 흐린 정도 σ 를 계산한다.

$$\sigma = \frac{1}{\sqrt{R^2 - 1}} \sigma_r \quad (5)$$

3.3 깊이맵 추출(Depth map Extracting)

엣지 위치에서 깊이 추정값을 구한 다음, 엣지를 포함하지 않는 영역에 대해 추정된 깊이값을 전파(Propagation)시킨다. 수식(5)에서 산출된 추정치에 근사하면서 정규화된 깊이 레이블 $\hat{\sigma}$ 를 계산한다. 또한 이미지 엣지에 정렬된 깊이 불연속성을 산출하며, 이산 마코브 랜덤 필드(MRF, Markov Random Field)를 통해 에너지 최소화된 공식 적용하면 수식(6)과 같다.

$$E(\hat{\sigma}) = \sum V_i(\hat{\sigma}_i) + \lambda \sum \sum V_{i,j}(\hat{\sigma}_i, \hat{\sigma}_j) \quad (6)$$

영상에서 각 픽셀은 MRF의 노드이며, λ 는 단일 노드 가능성과 페어 가능성을 균형 있게 제공한다. 이를 수식으로 표현하면 (수식7)과 같다.

$$V_{i,j}(\hat{\sigma}_i, \hat{\sigma}_j) = \sum w_{ij}(\hat{\sigma}_i - \hat{\sigma}_j)^2 \quad (7)$$

w_{ij} 는 칼라 영상 $I(i)$ 에서 서로 다른 이웃된 픽셀 정보의 차이를 의미하며, 이는 $w_{ij} = \exp\{-(I(i) - I(j))^2\}$ 과 같이 연산된다.

수식(6)에서 MRF 에너지를 최소화하기 위해 FastPD 알고리즘을 적용한다[15]. FastPD은 최적의 솔루션을 보장하고 기존의 그래프-컷 기술과 같은 MRF 최적화 방법보다 좋은 성능을 보인다[16].

4. 구 현

본 논문의 실험 환경은 3.3GHz Intel Core2 Duo CPU, 4GB 메모리와 Windows 10 Pro 운영체제의 PC에서 수행하였으며, 설계 제안하는 깊이맵 재생 알고리즘은 Matlab 2015a를 이용하여 구현하였다.

구현 실행에서 적용한 2개의 매개변수인 초점 흐림 스케일은 각각 1로 기본값을 설정하였다. 해당 기본값은 여러 번의 실험을 통해, 가장 좋은 성능을 도출할 수 있는 수치를 실험적으로 뽑은 값이다. 엣지 검출 단계에서는 케니 엣지 검출자(Canny edge detector)을 적용하였으며, 이는 다른 엣지 검출 알고리즘보다 성능적으로 우수하기

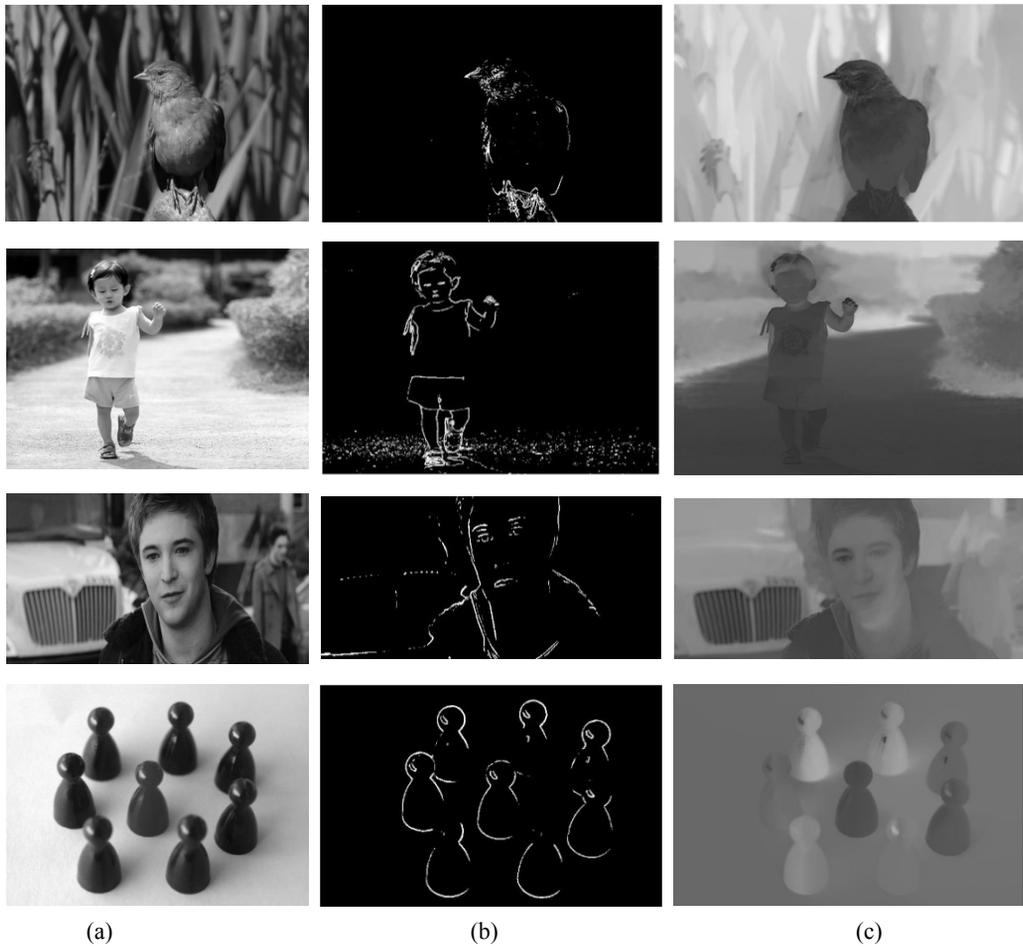


Fig. 3. Example of results, Input image in (a), Sparse blue map of edge detected image in (b), and Result image of the depth map in (c).

때문이다[17]. 픽셀 위치에서 원하는 에지 검출 결과를 도출하기 위해 매개변수를 조정하였으며, 실제 측정되는 값들을 기반으로 깊이맵을 재생하였다. Fig 3는 입력 원본영상과 실험을 통해 얻어진 깊이맵 영상에 대한 예제이다.

5. 결 론

본 논문에서는 단일 화면 영상(Single Scene)을 대상으로, 초점이 흐릿한 영상을 생성하고 이를 통해 영상의 깊이 정보를 분석하여 깊이맵을 생성하는 알고리즘을 설계 구현하였다. 가우시안 블러링 함수와 그라디언트 비율을 기반으로 에지 영역에서 흐려짐 정도를 분석하는 방법을

적용하였다. 에지 검출을 위한 Canny detector와 MRF 최적화를 위한 FastPD 알고리즘은 기존에 연구된 내용을 적용적으로 활용하였으며, 다수의 실험을 통해 좋은 성능을 도출하는 매개변수 값을 적용하였다.

설계 구현한 알고리즘은 다른 깊이맵 생성 방법보다 영상 잡음에 강인하며, 부정확한 에지 영역과 영상 가장 자리의 간섭(Interference)에서 보다 효과적인 깊이맵을 생성하였다.

향후 과제로, 공용적으로 사용되는 데이터셋을 통하여 기존 다른 연구와의 객관적 성능 비교 연구를 강화하고, 코드 최적화를 수행하여 모바일 환경에서 단일 입력 영상을 기반으로 스테레오스코픽 영상을 효과적이고 자동적으로 생성하는 애플리케이션을 확장할 예정이다.

감사의 글

본 연구는 2015년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(과제번호: 2015R1D1A1A01061152).

참고문헌

1. "3D Entertainment: Dimensions of an Emerging Market", One Touch Intelligence, Technical Report, 2010.
2. Sang-Hee Kweon, "Research of Fundamental Empirical Factors on 3D Stereoscopic Images – Focused on Measurement of 3D Pictures Application Importance on Contents and Media", Korea Association of Cultural Economics - Review of Culture and Economy, vol.15, No.1, 2012.
3. Eungyeol Song, Youngseop Kim, "The Algorithm of Brightness Control Display Matching in Stereoscopic", Journal of the Semiconductor and Display Technology, Vol.8, No.4, 2009.
4. Jae-In Kim, Taejung Kim, "Automated Image Alignment and Monitoring Method for efficient Stereoscopic 3D Contents Production", Journal of Broadcast Engineering, Vol.19, No.2, 2014.
5. Jean Maria Dominic, "Recent Trends in 2D to 3D Conversion: A Survey", International Journal for Research in Applied Science and Engineering Technology, Vol.2, Issue.4, 2014.
6. Umesh R. Dhond, J. K. Aggarwal, "Structure from Stereo – A Review", IEEE Transactions on Systems, Man and Cybernetics", Vol.19, No.6, 1989.
7. Naoki Asada, Hisanaga Fujiwara, Takashi Matsuyama, "Edge and Depth from Focus", International Journal of Computer Vision, Vol.26, Issue.2, 1998.
8. Francese Moreno-Noguer, Peter N. Belhumeur, Shree K. Nayar, "Active refocusing of images and videos", ACM Transaction on Graphics, Vol.26, Issue.3, 2007.
9. A. P. Pentland, "A New Sense for Depth of Field", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.9, Issue.4, 1987.
10. Yangjie Wei, Zaili Dong, Chengdong Wu, "Global depth from defocus with fixed camera parameters", International Conference on Mechatronics and Automation, 2009.
11. Vinay P. Namboodiri, Subhasis Chaudhuri, Sunil Hadap, "Regularized depth from defocus", IEEE International Conference on Image Processing, 2008.
12. Murali Subbarao, Tse-Chung Wei, Surya G. Nurzaman, "Focused image recovery from two defocused images recorded with different camera settings", Computer Society Conference on Computer Vision and Pattern Recognition, 2002.
13. Steven S. Beauchemin, Ruzena Bajcsy, "Modelling and Removing Radial and Tangential Distortions in Spherical Lenses", International Workshop on Theoretical Foundations of Computer Vision, LNCS 2031, 2000.
14. http://en.wikipedia.org/wiki/Circle_of_confusion
15. N. Komodakis, G. Tziritas, "Approximate Labeling via Graph-Cuts Based on Linear Programming", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007.
16. N. Komodakis G. Tziritas, N. Paragios, "Performance vs. computational efficiency for optimizing single and dynamic MRFs: Setting the state of the art with primal-dual strategies", Proceedings of Computer Vision and Image Understanding, 112(1), 2008.
17. S. Vijayarani, M. Vinupriya, "Performance Analysis of Canny and Sobel Edge Detection Algorithm in Image Mining", International Journal of Innovative Research in Computer and Communication Engineering, Vol.1, Issue.8, 2013.

접수일: 2016년 9월 6일, 심사일: 2016년 9월 20일,
 게재확정일: 2016년 9월 23일