

다중 조리개 영상으로부터 선명한 영상의 복원 기법

박태장, 박인규^o

인하대학교 정보통신공학과

taejang.park@gmail.com, pik@inha.ac.kr

요약

본 논문에서는 다른 조리개 크기로 취득한 영상을 이용하여 영상의 상대적인 깊이 정보를 추정하고 이를 이용하여 blur 를 제거한 디블러링 영상을 취득하는 기법을 제안한다. 즉, 카메라 조리개의 크기 변화로 취득된 영상의 피사계 심도를 이용하여 취득된 장면의 상대적인 깊이 정보를 추정한다. 먼저 조리개의 크기에 의하여 노출이 다른 두 영상간의 밝기를 보정한 후 두 영상간의 정확한 blur 비율을 측정하기 위하여 에지 위치에서 blur 를 측정한다. 그 후 측정된 blur 비율을 영상전체로 전파시켜 상대적인 깊이 정보인 defocus map 을 추정한다. 마지막으로 추정된 defocus map 을 공간변화적인 PSF 로 이용하여 blur 를 제거한 디블러링 영상을 생성한다.

1. 서론

최근 영상만을 이용하여 장면의 3 차원 깊이 정보를 추정하는 연구가 활발히 진행되고 있다. 카메라만을 이용한 간단한 기법으로 기하학적 방법을 이용한 스테레오 기법과 카메라 인자를 이용한 depth from defocus 기법이 존재한다.

Depth from defocus[1][2]는 초점영역이 다른 두 장의 영상을 이용하여 취득된 장면의 깊이 정보를 추정하는 기법으로서, 정확한 3 차원 정보를 취득하기 어렵지만 카메라 인자의 설정만으로 깊이 정보를 추정할 수 있는 장점을 가지고 있다. 최근에는 특정 코드에 의한 마스크로 조리개를 변형하고, 이후 디코딩을 통해 장면의 깊이 정보를 추정하는 coded aperture 기법의 연구도 활발히 진행되고 있다[3]. 또한 한 장의 영상만을 이용하여 장면의 깊이 정보를 추정하는 기법도 연구가 진행되고 있다[4][5].

본 논문에서는 조리개의 크기를 변화시켜 취득된 노출이 다른 영상으로부터 장면의 defocus map 을 추정한 후 이를 이용하여 blur 가 제거된 영상을 취득하는 기법을 제안한다.

2. 취득된 장면의 Defocus Map 추정

본 장에서는 노출 및 피사계 심도가 다른 영상으로부터 취득된 장면의 defocus map 을 추정하는 기법을 제시한다.

2.1 밝기 변환 함수

조리개의 크기 변화에 의하여 노출이 다른 영상간의 blur 를 측정하기 위하여 먼저 영상간의 밝기를 동일하게 보정하여야 한다. 이를 해결하기 위하여 밝은 영상 I_1 의 각 화소에 대응되는 어두운 영상 I_2 의 화소들을 $K \times K$ 배열에 저장하여 joint histogram 을 생성한다. 일반적으로 8 비트 밝기 영상에서 K 는 256 이다. 본 논문에서는 영상의 잡음과 포화상태에 대한 잡음을 줄이기 위하여 화소값의 범위를 [10, 245]로 제한하였다. I_1 의 각 화소 집합 Y 에 대응하는 I_2 의 최적의 화소 집합 X 을 추정하기 위하여 MMSE (minimum mean squared error) 를 적용하였으며 식 (1)과 같이 표현된다.

$$x_i = \frac{\sum_{x=0}^{K-1} xp(x, y_i)}{\sum_{x=0}^{K-1} p(x, y_i)} \quad (1)$$

여기서 $p(x, y)$ 는 joint probability density function 이다. 식 (1)에 의하여 추정된 X , Y 에 LSF (least square fitting) 알고리즘을 적용하여 두 영상간의 모든 화소에 대응하는 BTF (brightness transfer function) 곡선을 추정한다. 그림 1 은 노출이 다른 입력 영상과 BTF 에 의하여 변환된 영상, 그리고 각 채널에 대하여 추정된 BTF 를 보여주고 있다.

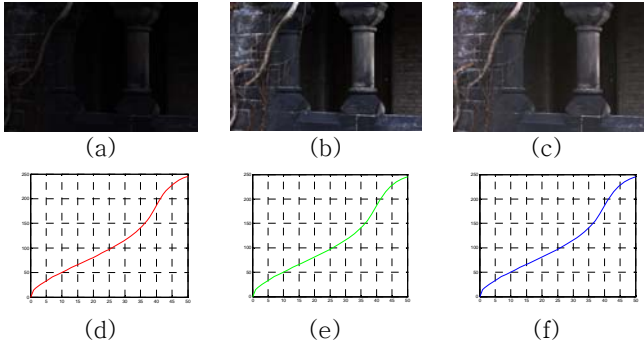


그림 1. 입력 영상 및 BTF 변환 영상. (a) 조리개 크기 F8로 취득한 영상. (b) 조리개 크기 F4로 취득한 영상. (c) (a)의 BTF 적용 영상. (d) R 채널에 대하여 추정된 BTF. (e) G 채널에 대하여 추정된 BTF, (f) B 채널에 대하여 추정된 BTF.

2.2 에지 기반의 Blur 비율 추정

본 논문에서는 Canny 알고리즘으로 추정된 에지 위치에서 blur 비율을 추정한다. 여기서 에지는 계단 함수로 모델링 하였으며 영상에 존재하는 blur 는 저주파 필터링된 성분을 가지므로 Gaussian 커널에 의하여 발생한다고 가정한다.

식 (2)는 1 차원 에지 위치 ($x=0$)에서의 gradient 를 나타내고 있다.

$$\begin{aligned} \nabla I(x) &= \nabla \{ (Au(x) + B) \otimes g(x, \sigma) \} \\ &= \frac{A}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right) \end{aligned} \quad (2)$$

여기서 σ 는 Gaussian 커널의 표준편차이다. 식 (2)로부터 blur 가 다른 동일한 에지 위치에서의 gradient 크기의 비율은 식 (3)과 같다.

$$BR = \frac{|\nabla I_1(0)|}{|\nabla I_2(0)|} = \frac{\sigma_2}{\sigma_1} \quad (3)$$

Blur 의 차이가 클수록 Gaussian 의 크기 차이 또한 커지므로 식 (3)의 BR 의 크기로부터 blur 의 비율 을 추정한다. 두 영상의 동일한 에지 위치 ($x=0, y=0$)에서의 blur 비율은 식 (4)로부터 얻어진 Gaussian 크기의 비율로 추정한다.

$$\|\nabla i(x, y)\| = \sqrt{\nabla G_x^2 + \nabla G_y^2} \quad (4)$$

여기서 $\nabla G_x, \nabla G_y$ 는 x, y 방향으로의 gradient 이다. 에지 위치에서 추정된 blur 는 카메라 센서에서 발생한 잡음, 양자화 잡음 그리고 BTF 에 의하여 변환된 영상 I_m 과 I_1 의 밝기 차이에 의한 잡음 등으로 정확한 blur 의 비율을 추정하기 어렵다. 따라서 식 (5)의 JBF (joint bilateral filtering)을 적용하여 잡음의 영향을 줄였으며 참조영상은 I_1 을 사용하였다.

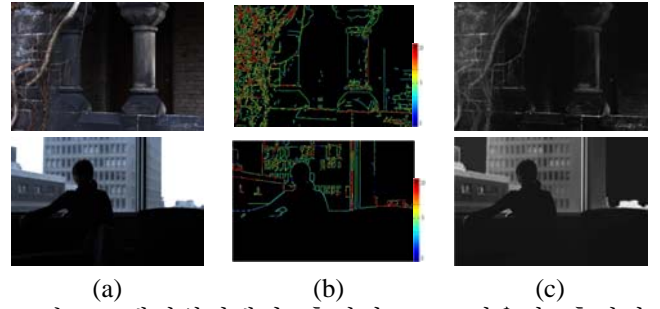


그림 2. 에지위치에서 추정된 blur 비율과 추정된 defocus map. (a) 입력영상 (b) 에지에서 추정된 blur 비율. (c) (b)의 blur 정보에 기반하여 추정된 defocus map.

$$\begin{aligned} JBF(BM)_p &= \\ \frac{1}{W(p)} \sum_{q \in N(p)} G_{\sigma_s}(\|p-q\|) G_{\sigma_r}(\|C_i(p) - C_i(q)\|) BM_q \end{aligned} \quad (5)$$

여기서 σ_s 는 공간 영역의 표준편차, σ_r 는 밝기 영역의 표준편차, $W(p)$ 는 정규화 계수, $N(p)$ 는 중심이 p 인 7×7 커널 그리고 C_i 는 각 채널의 밝기 값이다.

2.3 Defocus Map 추정

본 논문에서는 에지에서 추정된 blur 정보를 영상 전체로 전파시키기 위해서 주변 화소의 밝기 값 또는 색이 비슷하면 동일한 물체 또는 동일한 영역에 존재하며 blur 또한 비슷하다고 가정한다. 이 가정에 기반하여 화소값의 유사성을 이용한 matting Laplacian 알고리즘[6]을 사용하여 취득한 장면의 상대적인 깊이 정보인 defocus map 을 추정한다.

영상 전체에 대한 defocus map d 을 추정하는 비용함수는 식 (6)과 같이 표현된다.

$$o(d) = d^T L d + \lambda (d - \hat{d})^T D (d - \hat{d}) \quad (6)$$

여기서 \hat{d} 는 BR 의 벡터, D 는 에지 위치에서 1, 그렇지 않으면 0을 가지는 대각선 행렬, L 은 matting Laplacian로 식 (7)과 같이 정의되며 \hat{d}, D, L 은 sparse 행렬의 특성을 가진다.

$$L = \sum_{k(i,j) \in \omega_k} \left(\delta_{ij} - \frac{1}{|\omega_k|} \left(1 + (I_i - \mu_k)^T \left(\Sigma_k + \frac{\varepsilon}{|\omega_k|} E_3 \right)^{-1} (I_j - \mu_k) \right) \right) \quad (7)$$

여기서 ε 는 정규화 인자, $|w_k|$ 는 k 번째 마스크 내부의 총 화소 수, μ_k 와 Σ_k 는 평균과 공분산, δ_{ij} 와 E_k 는 Kronecker delta 함수와 3×3 단위 행렬이다. 식 (6)을 d 에 대하여 미분한 식 (8)로부터 defocus map 을 추정하였으며 결과는 그림 2 에 도시하였다.

$$(L + \lambda D)d = \lambda D \hat{d} \quad (8)$$

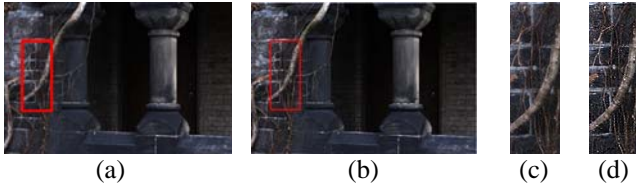


그림 3. Defocus map 을 PSF 로 이용하여 생성된 디블러링 영상. (a) 조리개 크기 F/4 로 취득한 영상. (b) (a)의 디블러링 영상. (c) (a)의 붉은 영역 영상. (d) (b)의 붉은 영역 영상.

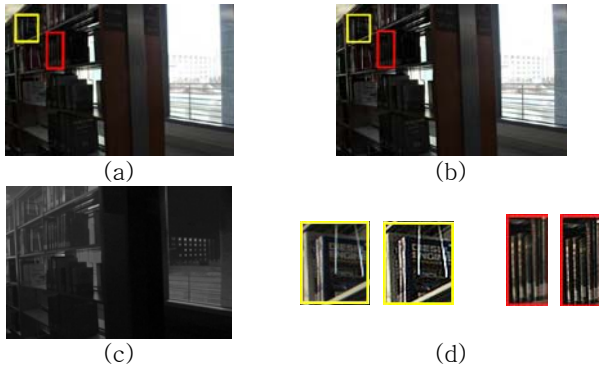


그림 4. Defocus map 추정 및 디블러링 영상 생성. (a) 조리개 크기 F/10 로 취득한 영상. (b) (a)의 디블러링 영상. (c) 추정된 defocus map. (d) (a)와 (b)의 표시된 영역의 확대 영상.

3. 디블러링 영상 생성

일반적으로 영상에 존재하는 blur 를 제거하기 위하여 RL (Richardson-Lucy) 알고리즘이 널리 사용되고 있다[7]. 본 논문에서는 추정된 defocus map 을 PSF 로 이용하였으며 deconvolution 과정에서 발생하는 잡음의 크기를 억제하기 위하여 식 (9)의 TV (total variation) 정규화를 적용하였다.

$$R_{TV}(I) = \int \sqrt{\|\nabla I(x)\|^2} dx \tag{9}$$

여기서 ∇I 는 수평, 수직 방향으로 일차 미분이다. 식 (9)의 TV 정규화를 RL 알고리즘에 적용하면 식 (10)와 같이 정의된다.

$$I_{t+1} = \frac{I_t}{1 - \lambda \nabla R_{TV}(I)} \times M \circ \frac{B}{I_t \otimes M} \tag{10}$$

여기서 \circ 는 correlation 연산자, \otimes 는 컨볼루션 연산자, B 는 blur 영상, I_t, I_{t+1} 은 $t, t+1$ 번째의 deconvolution 을 수행한 영상이며 M 는 Gaussian 커널이다. 그림 3 은 추정된 defocus map 을 RL 알고리즘의 PSF 로 이용하여 deconvolution 을 수행한 결과 영상을 보여주고 있다.

4. 실험 결과 및 분석

본 논문의 실험은 Intel[®] Core[™] i5 Quad 2.66 GHz CPU 에 2GB 의 메모리를 장착한 환경에서 수행되었다. 그림 4 는 동적 범위가 넓은 환경에서 취득된 영상으로부터 생성된 결과이다. 여기서 초점영역은 카메라와 가까운 책장 부분으로 설정하였으며 영상의 크기는 500×333 이다. 그림 4(c)는 추정된 defocus map 으로 초점영역으로부터 멀어질수록 더 밝은 값을 가진다. 그림 4(d)는 추정된 defocus map 을 이용하여 deconvolution 된 그림 4(b)와 입력영상인 그림 4(a)의 표시된 영역이며 보다 정확한 대비를 위하여 영상의 밝기를 증가시켰다. 그림 4(d)로부터 흐려진 영역의 정보가 보다 선명해 진 것을 알 수 있다.

5. 결론

본 논문은 초점 영역만 다른 영상을 사용한 기존 방법과 달리 조리개에 의하여 피사계 심도 및 노출이 다른 영상으로 defocus map 추정 및 선명한 영상을 생성하는 방법을 제안하였다. 조리개 변화로 인한 피사계 심도의 차이를 이용하여 취득된 장면의 상대적인 깊이 정보 추정이 가능하며 이 정보를 RL 알고리즘에 적용하여 흐려진 영역의 정보를 보다 정확히 인식할 수 있는 영상을 생성하였다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터지원사업의 연구결과로 수행되었음 (NIPA-2011-(C1090-1111-0003)).

참고문헌

- [1] A. P. Pentland, "A new sense for depth of field," *Proc. IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 523-531, JULY 1997.
- [2] H. Jin and P. Favaro, "A variational approach to shape from defocus," *Proc. European Conference on Computer Vision*, pp. 18-30, May 2002.
- [3] P. Green, W. Sun, W. Matusik, and F. Durand, "Multi-aperture photography," *ACM Trans. on Graphics*, vol. 26, no. 3, pp. 68, July 2007.
- [4] S. Bae and F. Durand, "Defocus magnification," *Computer Graphics Forum*, vol. 44, no. 9, pp. 571-579, September 2007.
- [5] S. Zhuo and T. Sim, "Defocus map estimation from a single image," *Pattern Recognition*, vol. 44, no. 9, pp. 1852-1858, September 2011.
- [6] A. Levin, D. Lischinski, and Y. Weiss, "A closed form solution to natural image matting," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228-242, February 2008.
- [7] Y. W. Tai, P. Tan, L. Gao, and M. S. Brown, "Richardson-Lucy deblurring for scenes under projective motion path," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1603-1618, August 2011.