

4 차원 Light Field 영상에서의 학습 기반 초해상도 알고리즘

이승재, 박인규
 인하대학교 정보통신공학과
 {creative.sjlee@gmail.com, pik@inha.ac.kr}

Learning-Based Superresolution for 4D Light Field Images

Seung-Jae Lee and In Kyu Park
 Department of Information and Communication Engineering, Inha University

요 약

영상을 취득한 후 다양한 응용프로그램으로 확장이 가능한 4 차원 light field 영상은 일반적인 2 차원 공간 (spatial) 영역과 추가적인 2 차원 각 (angular) 영역으로 구성된다. 그러나 이러한 4 차원 light field 영상을 2 차원 CMOS 센서로 취득하므로 이에 따른 해상도 제약이 존재한다. 본 논문에서는 이러한 4 차원 light field 영상이 가지는 해상도 제약 조건을 해결하기 위하여, 4 차원 light field 영상에 적합한 학습 기반 (learning-based) 초해상도 (superresolution) 알고리즘을 제안한다. 제안하는 알고리즘은 공간영역 해상도 그리고 각영역의 해상도를 각각 2 배 향상시킨다. 실험에 사용되는 영상은 상용 light field 카메라인 Lytro 에서 취득하며, 기존의 선형 보간 기법인 bicubic 기법과의 비교를 통해 제안하는 기법의 우수성을 검증한다.

1. 서론

최근 3D TV 등 최근 영상 산업분야에서 4 차원 light field 영상은 많은 관심을 받고 있다. Light field 영상은 재초점 영상 (refocusing), 시점 변환 (viewpoint change), 그리고 깊이 지도 복원 (depth map reconstruction) 등 다양한 응용이 가능하다. 상용 light field 카메라인 Lytro 카메라는 공간영역 해상도 380×380 , 각영역 해상도로 9×9 의 11 메가 광선(mega-rays)을 CMOS 센서로 취득한다. 이러한 장비는 4 차원 light field 영상을 2 차원 센서로 취득하기에 공간, 각영역에서 해상도 제약이 존재한다.

이러한 문제를 해결하기 위하여 4 차원 light field 영상에 적용이 가능한 해상도 향상 기법이 필요하다. 이를 위하여 본 논문에서는 4 차원 light field 영상에 적합한 패치를 구성하고 이를 바탕으로 해상도 향상을 도출하는 초해상도 알고리즘을 제안한다.

제안하는 알고리즘은 두 가지 단계로 구성된다. 첫째, 4 차원 light field 영상에서 저해상도 패치와 고해상도 패치간의 짝으로 이뤄진 대규모의 4 차원 패치를 추출하고 K 평균 군집 (K-means clustering) 기법을 이용하여 딕셔너리 (dictionary)를 구성하는 단계이다. 위와 같은 딕셔너리 학습 단계는 오프라인 (offline) 과정으로 수행한다. 둘째, 입력 영상으로부터 추출한 저해상도 light field 패치와 대규모 패치를 이용하여 학습 과정을 거친 딕셔너리를 바탕으로 4 차원 light field 영상에서 공간, 각영역으로 해상도를 향상시키는 단계이다. 위와 같은 복원 단계의 과정은 온라인 (online) 과정으로 수행한다.

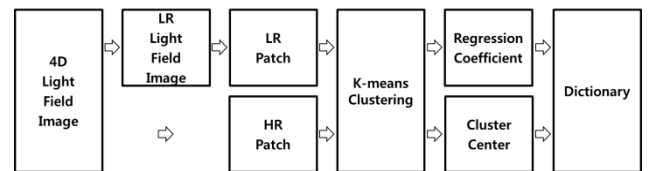


그림 1. 딕셔너리 훈련 과정

2. 제안하는 알고리즘

2.1 딕셔너리 훈련 단계

제안하는 알고리즘에서는 배열 렌즈 (lenslet) 기반 light field 카메라를 사용하여 4 차원 light field 영상을 취득한다. 다양한 장면에 대해 취득한 영상을 기준으로 공간영역, 각영역으로 절반씩 해상도를 낮추어 만든 저해상도 영상을 이용하여 훈련 데이터로 사용한다. 훈련 단계에서 사용하는 4 차원 light field 패치는 각영역 영상들의 동일한 위치에서 공간영역의 2 차원 패치들의 합으로 구성되어 공간영역과 각영역 정보를 모두 포함한다.

제안하는 알고리즘의 군집화 단계에서는, K 평균 군집 기법을 이용하여 패치들을 효율적으로 군집화 하고, 각 군집에서 식(1)을 이용하여 대표 값(cluster center)의 회귀 계수(regression coefficient)를 추정하고 이를 바탕으로 딕셔너리를 구성한다.

$$C^* = \underset{c}{\operatorname{argmin}} \left\| \mathbf{H} - \mathbf{C} \begin{pmatrix} \mathbf{L} \\ 1 \end{pmatrix} \right\|^2 \quad (1)$$



그림 2. Lytro 카메라로 취득한 영상

이때, \mathbf{C} 는 $n \times m$ 크기의 회귀 계수 행렬을 의미하며, n 과 m 은 고해상도, 저해상도 패치의 픽셀의 개수를 의미한다. 또한 \mathbf{H} 와 \mathbf{L} 은 동일한 군집 내에 존재하는 고해상도, 저해상도 패치들로 구성된 행렬을 의미한다. 그림 1 에서 딥러닝 훈련과정의 개요를 도시하였다.

2.2 복원 단계

복원 단계에서는 입력 저해상도 영상에서 4 차원 패치를 추출하고, 이를 기준으로 딥러닝에 저장되어 있는 중심 군집 값들과 비교를 통해 가장 근접한 중심 군집 값을 찾게 된다. 이후 식(2)와 같이 중심 군집 값에 저장 되어있는 회귀 계수인 \mathbf{C}^* 와 입력 저해상도 패치인 \mathbf{I} 의 연산을 통하여 고해상도 패치인 \mathbf{h} 를 복원해낸다.

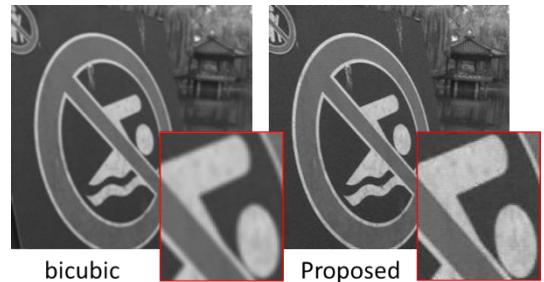
$$\mathbf{h} = \mathbf{C}^* \begin{pmatrix} \mathbf{I} \\ 1 \end{pmatrix} \quad (2)$$

해당 과정을 입력 영상의 모든 위치에서 추출한 모든 저해상도 패치에 반복적으로 적용함으로써 고해상도 영상을 복원해 낼 수 있다.

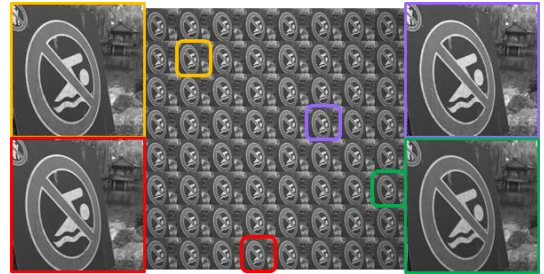
3. 실험 결과

본 논문의 실험은 Intel i7-3770K 3.5GHz CPU 와 16GB 메모리의 환경에서 진행한다. 또한 실험에 사용한 모든 영상은 $360 \times 360 \times 8 \times 8$ 의 크기로 Lytro 에서 취득하며 이를 바탕으로 저해상도 영상은 $180 \times 180 \times 4 \times 4$ 로 제작한다. 그리고 4 차원 고해상도, 저해상도 패치의 크기는 각각 $8 \times 8 \times 8 \times 8$, $4 \times 4 \times 4 \times 4$ 로 설정하고 이는 $4096 \times 1, 256 \times 1$ 의 벡터형태로 저장된다. 원활한 실험을 위하여 패치는 그림 2 의 light field 영상 40 장에서 200,000 개의 패치를 랜덤으로 추출한다. 또한 K 평균 군집에 사용할 K 는 512 개로 설정한다.

입력 영상을 제안하는 딥러닝 학습 기반 초해상도 알고리즘에 적용하여 두 배의 해상도 향상을 유도한다. 그림 3 은 bicubic 기법과 제안하는 알고리즘의 정량적, 정성적 비교이며, bicubic 기법은 공간영역의 해상도 향상만 도출하는 반면 제안하는 알고리즘은 공간영역, 각영역의 해상도 모두를 향상시키는 결과를 보여준다.



(a)



(b)

그림 3. 실험 결과 영상. (a) 공간영역에서 bicubic 보간법과의 시각적 비교. PSNR 기준으로 bicubic 기법은 31.15dB, 제안하는 알고리즘은 32.72dB. (b) 각영역에서의 결과 영상. 짝수 행과 열은 각영역에서 생성된 영상임.

4. 결론

본 논문에서는 4 차원 light field 영상에 적합한 딥러닝 학습 기반 초해상도 알고리즘을 제안하였다. 제안한 알고리즘을 사용하여 공간영역, 각영역으로 각각 두 배씩 해상도를 향상시켰고, 향상된 해상도를 가진 light field 영상을 정량적, 정성적으로 평가 및 검증하였다.

감사의 글

이 논문은 2013 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2013R1A2A2A01069181)

참고문헌

- [1] C.-Y. Yang and M.-H. Yang, "Fast direct super-resolution by simple functions," Proc. ICCV, pp. 561-568, December 2013.
- [2] D. G. Dansereau, O. Pizarro, and S. B. Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," Proc. CVPR, pp.1027-1034, June 2013.
- [3] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Computer Science Technical Report (CSTR), 2005-2, April 2005.
- [4] X. Huang and O. Cossairt, "Dictionary learning based color demosaicing for plenoptic cameras," Proc. CVPR Workshop, pp. 455-460, June 2014.