

전방위 영상 생성을 위한 SIFT 연산량 감소

*문원준 *이재은 *서영호 *김동욱

*광운대학교 전자재료공학과

wonjun94@kw.ac.kr wodms6364@kw.ac.kr yhseo@kw.ac.kr dwkim@kw.ac.kr

Reducing the SIFT Calculation for Omni-directional Image Generation

*Moon, Won-Jun *Lee, Jae-Eun *Seo, Young-Ho *Kim, Dong-Wook

*Department of Electronic Materials Engineering, Kwangwoon University

요약

본 논문에서는 시장규모가 점진적으로 커지고 있는 VR의 콘텐츠 제작에 있어서 전방위 영상 생성을 위해 촬영한 영상들을 스티칭하는 데 소요되는 연산량을 감소시키는 방법을 제안한다. 제안하는 방법은 널리 사용되는 SIFT의 알고리즘에서 가장 많은 연산을 차지하는 가우시안 피라미드 생성 부분을 파라미터를 조절하여 감소시키고, 이를 통해 얻은 점들을 스티칭하는데 사용하여 영상을 합성하는 방법이다. 영상에 대한 평가는 변환에 의한 점들의 평균오차를 계산하여 유효성을 판단한다.

1. 서론

가상현실(Virtual Reality, VR)에 대한 시장규모가 확대됨에 따라 콘텐츠 시장 역시 대규모로 성장하고 있다[1]. VR 콘텐츠는 360도의 정보를 모두 포함해야 하는데, 이를 만족하는 영상을 전방위 영상이라고 한다. 전방위 영상을 만들기 위해서는 스티칭(stitching), 즉 획득한 영상을 붙이는 작업을 수행해야 하는데, 스티칭을 위해서는 영상에서 동일한 부분을 찾아 정합하는 과정이 필요하다.

스티칭에는 주로 SIFT 알고리즘이 사용되고 있는데, 이를 통해 특징점을 추출하고 각 영상간의 특징점을 비교하여 정합 여부를 판단하여 스티칭을 수행한다[2,3].

SIFT에서는 영상에 분산을 키워가면서 가우시안 필터를 반복 적용하여 피라미드를 만드는 과정이 포함되는데, 이 연산에 많은 시간이 소요되기 때문에 성능은 떨어지지만 속도를 향상시킨 SURF와 같은 다른 알고리즘도 연구되었다[4,5].

본 논문에서는 연산시간 문제를 해결하기 위해 가우시안 필터에서의 초기 분산과 피라미드의 한 옥타브 내에서의 레이어 수를 조절한다. 이에 따라 SIFT 알고리즘에서의 연산량은 감소되지만, 스티칭의 결과는 근사하게 나오는 방법을 제안한다.

2. 연산량 감소 방법

2.1. 필터의 특성

본 논문에서는 SIFT 알고리즘을 사용하는데, SIFT는 크게 특징점을 찾는 과정과 특징점의 특징을 추출하는

과정으로 나뉜다. 특징점을 찾는 과정은 영상을 가우시안 필터로 반복 처리하게 된다.

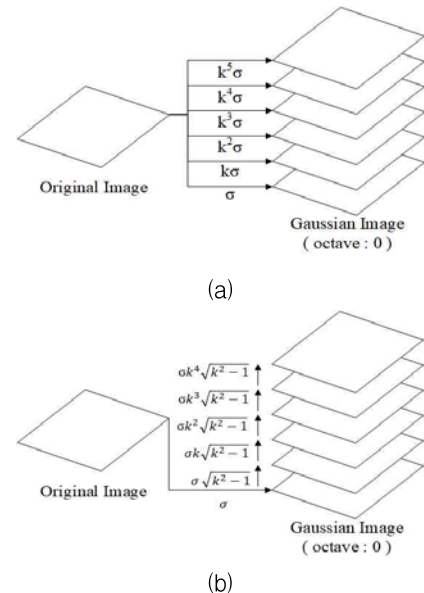


그림 1. 가우시안 피라미드 구성 방법: (a) 기본적인 방법, (b) 캐스케이드 특성을 통한 계산량 감소 방법.

그림 1(a)은 원본 영상에 분산 값을 증가시키어 가우시안 필터링을 수행하여 피라미드를 만드는 과정인데, 그림 1(b)와 같이 가우시안 필터의 캐스케이드 특성을 이용하여 원본 영상이 아닌 이전에 처리된 영상에 누적적으로 필터링하는 방식을 통해 연산량을 감소시킬 수 있다[6].

영상의 해상도가 4,000×3,000인 영상에 대해 8 옥타브까지 SIFT 특징점 추출을 위한 필터링 연산을 했을 때, 사용되는 MAC(multiplier-accumulator)의 연산 횟수가 (a) 방식은 7.4×10^{10} 번의 연산이 필요하고, (b)는 3.2×10^{10} 번의 연산이 필요하여 (a)의 연산량이 약 2.3배 많았다.

표 1. 분산 값에 따른 필터사이즈 (레이어는 6개로 고정)

Sigma	Basic method					Cascade method						
	I0	I1	I2	I3	I4	I5	I0	I1	I2	I3	I4	I5
0.6	7	7	9	11	13	17	7	5	7	7	9	11
0.8	7	9	11	15	17	21	7	7	7	9	11	13
1	9	11	15	17	21	27	9	7	9	11	13	17
1.2	11	13	17	21	25	31	11	9	11	13	17	21
1.4	13	15	19	23	29	37	13	11	13	15	19	23
1.6	15	17	21	27	33	43	15	11	13	17	21	27
1.8	15	19	25	31	37	47	15	13	15	19	23	29
2	17	21	27	33	41	53	17	13	17	21	27	33

표 2. 레이어 개수에 따른 필터사이즈 (분산은 1.6 고정)

Layer	Basic method					Cascade method				
	L4	L5	L6	L7	L8	L4	L5	L6	L7	L8
10	15	15	15	15	15	15	15	15	15	15
11	27	19	17	17	17	23	15	11	9	9
12	53	27	21	19	19	45	19	13	11	9
13	103	37	27	23	21	91	27	17	13	11
14	▪	53	33	27	23	▪	37	21	15	13
15	▪	▪	43	31	27	▪	▪	27	17	15
16	▪	▪	▪	37	31	▪	▪	▪	21	15
17	▪	▪	▪	▪	35	▪	▪	▪	▪	19

2.2 파라미터 변화

가우시안필터는 분산에 따라 필터의 사이즈가 달라지는데, SIFT에서의 분산은 초기로 설정해주는 분산 값과, 레이어(Ln)의 개수가 영향을 미친다. 분산 값과 레이어 개수에 따른 필터의 사이즈는 표 1, 2 와 같다.

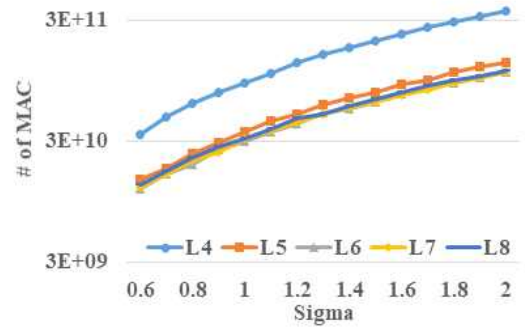
3. 실험 및 결과

3.1 실험환경 및 조건

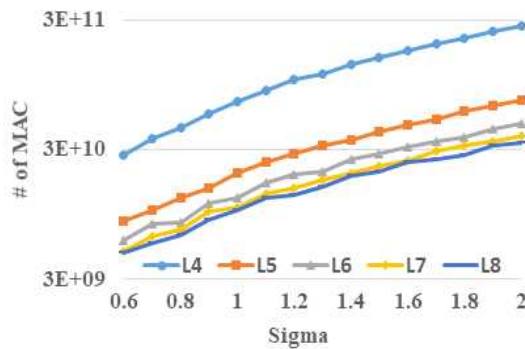
본 실험에서는 검치는 영역이 존재하는 2개의 영상을 한 세트로 하여 총 10세트에 대해 분산과 레이어의 개수를 조절해가며 스티칭을 수행하였다. 생성된 영상에 대한 평가는 SIFT 알고리즘을 통해 매칭이 된다고 정해진 점과 RANSAC을 통한 호모그래피(homography)를 구하여 실제 변환을 통해 얻어지는 점의 유클리드 거리를 계산하여 평균을 내는 방식으로 오차 값을 계산하였다.

3.2 연산량의 감소

초기 분산은 0.6~2, 레이어는 4~8의 범위에 대해서 각 파라미터의 변화에 따른 그림 1(a)와 (b) 방식에 대한 MAC의 연산수를 그림 2에서 나타내었다. 영상의 해상도는 2.1절과 마찬가지로 4,000×3,000에 대해 수행하였다.



(a)



(b)

그림 2. 분산과 레이어의 개수에 따른 MAC연산 횟수: (a) 기본적인 피라미드, (b) 캐스케이드 특성을 이용한 피라미드.

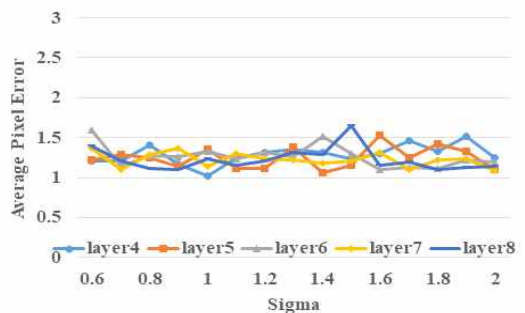


그림 3. 분산과 레이어의 개수에 따른 스티칭 오차

3.3 초기 분산과 레이어의 개수에 따른 결과 비교

그림 3에서는 파라미터에 상관없이 화소의 평균 제곱 오차가 1~1.5픽셀 사이에 위치하는 것을 확인할 수 있는데, 이 오차는 사람이 구분할 수 없을 정도로 미세하다.

따라서 초기 분산과 레이어 선택에 따른 스티칭 결과는 비슷하다고 볼 수 있으므로, 2.2절에서의 가장 적은 MAC의 수를 가지는 값이 가장 효율적이다.

가장 적은 MAC의 수를 가지는 8개의 레이어를 가질 때, 분산에 따른 스티칭 결과를 그림 4에서 나타내었다.

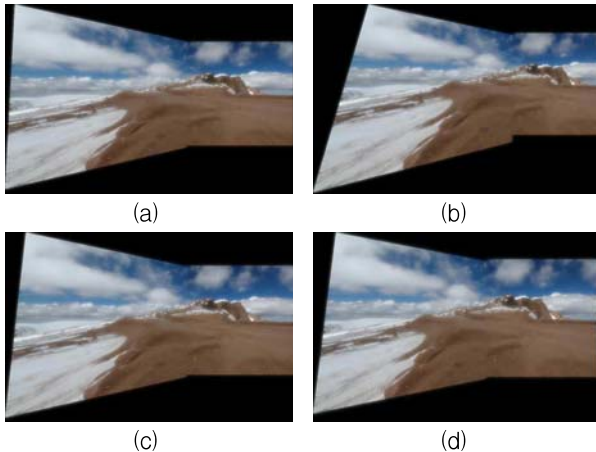


그림 4. 레이어 8개일 때의 분산의 변화에 따른 스티칭 결과 (a) 0.6, (b) 1.0, (c) 1.6, (d) 2.0

4. 결론

본 논문에서는 전방위 영상 생성과정에서의 연산량 감소를 위해 초기 분산과 레이어의 개수를 선정하는 방법을 제시하였다. 가장 연산량이 적은 초기분산 0.6과 레이어 8개에 대해 비교했을 때, 전방위 영상의 결과는 비슷하지만 계산상에서 약 42.9배의 이득을 얻는다. 두 장의 영상에 대해서만 결과를 비교한 것이기에, 실제 여러 장의 영상을 합성하여 만드는 전방위 영상에서는 더 큰 효율을 가질 것이라 기대한다.

감사의 글

이 논문은 2016 년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2016R1D1A1B03930691)

참고문헌

- [1] Institute for Information & Communications Technology Promotion: “Technology Development Trend and Market Forecast of VR/AR,” Weekly Technology Trends, Vol. 1803, pp. 48–55, July, 2017.
- [2] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” International Journal of Computer Vision, Vol. 60, No. 2, pp. 91–110, Jan. 2004.

- [3] D. G. Lowe, “Automatic Panoramic Image Stitching using Invariant Features,” International Journal of Computer Vision, Vol. 74, No. 1, pp. 59–73, Aug. 2007.
- [4] F. C. Huang, S. Y. Huang, J. W. Ker Y. C. Chen “High-Performance SIFT Hardware Accelerator for Real-Time Image Feature Extraction,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, No. 3, pp. 340–351, Mar. 2012.
- [5] H. Bay, T. Tuytelaars, L. V. Gool “SURF: Speeded Up Robust Features,” European Conference on Computer Vision, pp. 404–417, 2006.
- [6] R. Hess, “An Open-Source SIFT Library,” ACM Multimedia, pp. 1493–1496, Oct. 2010.

실외 영상에서 반사율의 유사 패치를 이용한 복합 광원 추정

이상호, 김종욱

고려대학교

franky_@korea.ac.kr, jokim@korea.ac.kr

Multiple Light Sources Estimation Using Similar Patches of Reflectance in Outdoor images

Sang-Ho Lee Jong-Ok Kim

Korea University

요 약

본 논문에서는 실외 영상에서의 새로운 접근 방식의 복합 조명 알고리즘을 제안한다. 기존의 복합 조명 알고리즘들이 동시에 두 조명을 추정하는 것에 비해 제안 알고리즘은 먼저 단일 조명 기법을 적용하여 첫번째 광원의 색을 추정한 후에 각 영역에서 유사 패치 쌍을 찾아 두번째 광원의 색을 추정하는 방식이다. 일반적인 복합 조명 환경에서는 적용하기 힘들지만 환경을 실외로 제한하여 실외의 광원인 햇빛과 그늘 사이의 관계를 이용하여 효과적으로 유사 패치를 찾아 두 광원의 색을 추정한다. 따라서 실외 환경을 촬영하여 얻은 raw 파일 영상에 제안 알고리즘을 적용하여 효과적으로 광원들의 영향을 제거할 수 있다.

1. 서론

사람은 같은 객체에 대하여 조명의 색에 관계없이 동일한 색으로 지각하는 시지각 능력을 가지고있고 이것을 color constancy 라고 한다. 카메라는 사람의 눈과 다르게 color constancy 를 가지고 있지 않기 때문에 white balance 가 필요하다. White balance 는 조명의 색을 추정하여 조명의 효과를 제거하는 후처리 과정이다.

기존 카메라에는 단일 조명에 대한 화이트 밸런스 기법이 내제되어 있다. 따라서 복합 조명 환경에서 촬영할 때, 색 재현 정확도가 낮은 사진을 얻게 된다. 복합 조명 환경에 대응하는 기법들이 최근에 제안되고 있으나 단일 조명 환경일 때에 비해 복잡도가 높고 성능이 좋지 않다. [1]~ [3]

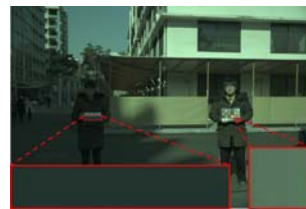
단일 조명 환경에서는 뛰어난 성능을 가진 white balance 방법들이 있지만 복합 조명 환경에서 사람이 실제로 인지하는 색을 복원하기는 어렵다. 복합 조명 환경은 크게 실내, 실외로 나눌 수 있고, 본 논문에서는 실외 환경의 복합 조명 환경이 다뤄진다. 실외는 햇빛과 그늘의 복합 조명 환경이라 할 수 있으며, 따라서 햇빛과 그늘 조명 사이의 관계를 이용할 수 있다.

본 논문에서는, 복합 조명 환경에 대응하는 새로운 접근 방식의 알고리즘을 제안한다. 제안하는 기법은 먼저 기존의 단일 조명 기법으로 햇빛 조명을 측정하여 화이트 밸런스 한 다음 그늘 조명의 영역을 보정한다. 이와 같은 방식으로 실외 영상에 집중하여 기존 복합 조명 기법들의

문제점을 보완하고 결과 영상의 색 재현 정확도가 높다.

본 논문의 구성은 다음과 같다. 2 절에서는 제안 알고리즘에 대해 설명하고, 3 절에서는 제안 알고리즘을 적용 후 실험 결과에 대해 토의한다. 마지막으로 4 절에서는 본 논문에 대한 결론을 맺고 있다.

2. 제안 알고리즘



영역	그늘	햇빛
R:G:B	0.41:1:0.75	0.51:1:0.60

그림 1. Color checker 를 이용한 그늘과 햇빛 광원의 RGB 비율 비교

일반적인 복합 조명 환경에서는 두 조명 간의 상관관계가 적으므로 조명들을 추정하기 어렵다. 하지만 햇빛과 그늘 조명의 실외 영상의 경우에는 그늘 영역이 상대적으로 밝기 값이 낮으므로 조명에 의한 색 차이와 밝기 간의 상관관계가 높다. 따라서 해당 영상의 경우에는 영상의 밝기 값을 이용하여 두 조명의 영역을 어느 정도 구분할 수 있다. 그리고 그림 1. 에서 color checker 를

통하여 G 값을 기준으로 광원의 색을 비교했을 때 그들이 햇빛에 비해 상대적으로 R 이 낮고 B 가 높은 것을 확인할 수 있다.

제안 알고리즘에 대해 설명하면, 먼저 영상 최대 밝기 값의 일정 퍼센트 값(10%)을 기준으로 햇빛 영역과 그늘 영역을 구분한다. 그 후에 단일 조명 알고리즘을 통해 햇빛 조명을 추정하여 영상 전체에 대해 색을 재생성 한다. 그늘 조명은 햇빛 조명보다 상대적으로 색온도가 높으므로 영상에서 그늘 영역 부분이 원래의 색보다 파랗게 표현되므로 그늘 영역에 대한 보정이 필요하다. 그늘 영역에 대한 보정을 하기 위해 햇빛 영역과 그늘 영역에서 각각 유사 패치를 찾는다. 햇빛 영역을 10 등분하여 각각의 영역에서 랜덤으로 후보 패치를 선출한다. 각 후보 패치와 3 가지의 조건을 만족하는 타겟 패치를 그늘 영역에서 찾아 유사도를 비교한다. 첫번째 조건은 패치 내부의 픽셀들의 색이다. 그늘 광원에 비해 햇빛 광원의 색의 비율에서 G 를 기준으로 R 이 높고 B 가 낮기 때문에 조건에 해당하는 패치들 중에서만 검색을 한다. 두번째는 각 유사 패치의 중심 픽셀 간의 거리이다. 반사율이 같은 물체는 거리상으로 멀리 떨어져있지 않다고 가정한다. 마지막으로 그늘 영역의 타겟 패치의 주변 일정 반경 이내가 모두 그늘 영역일 때만 유사도를 비교한다. 이 세가지 조건을 만족하는 타겟 패치만 검색하여 보다 정확하게 찾을 수 있고 시간도 단축시킬 수 있다.

$$I = L \times R \quad (1)$$

위 식과 같이 영상에서 픽셀 값은 광원의 조명 값과 반사율의 곱으로 나타낼 수 있다. I 는 영상에 나타나는 픽셀 값이고 L 은 광원의 색에 해당하는 조명 값, R 은 물체의 반사율이다. I, L, R 은 모두 RGB 로 3 채널로 이루어져 있다. 반사율이 비슷한 패치 쌍을 찾는게 목적이므로 패치 내부에서 각각의 픽셀들을 모든 픽셀의 중간 값으로 나누어 준다. 패치 내부의 픽셀들은 같은 조명 값을 가지고 비슷한 반사율을 가지기 때문에 중간 값으로 나누면 반사율의 RGB 비율과 유사한 비율을 얻을 수 있다.

후보 패치와 타겟 패치의 반사율의 유사도를 비교하기 위해 Peak signal-to-noise ratio (PSNR)을 이용한다. PSNR 은 주로 영상의 손실 정보를 평가할 때 쓰이는데 두 영상이 유사도를 나타낼 때 사용되기도 한다. 제안 알고리즘에서는 각 패치를 중간 값으로 나눈 뒤 얻은 반사율의 유사도를 측정하는데 PSNR 을 사용하였다. PSNR 의 수식은 다음과 같다.

$$PSNR = 10 \times \log \left(\frac{255^2}{MSE} \right) \quad (2)$$

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N [I(i, j) - I'(i, j)]^2$$

제안 알고리즘으로 추정한 그늘 영역 광원의 색을 그늘 영역의 모든 픽셀에 대해 동일하게 적용하면 밝기 값을 기준으로 햇빛, 그늘 영역을 정확하게 나누지 못하기 때문에 경계 부근에 색 아티팩트가 나타나게 된다. 따라서 영역을 구분하는 영상 최대 밝기 값의 일정 퍼센트 값(10%)과 유사한 값(9-11%)에서는 햇빛과 그늘 광원의 중간에 해당하는 색으로 보정하였다.

3. 실험결과



(a) raw 파일 영상 (b) 결과 영상
그림 2. 제안 알고리즘 적용 전후 영상

시간에 따라 바뀌는 실외 환경의 특성상 실험에 사용할 사진을 직접 촬영하였다. 사진은 Canon EOS 5D Mark III 카메라로 촬영하여 2880x1920 크기의 raw 파일로 얻었다. 그림 2 의 보정 전, 후의 사진을 비교하면 색 재현 정확도가 높아졌음을 알 수 있다. 또한 눈으로 봤을 때 경계 부근에서도 아티팩트가 나타나지 않는것을 확인할 수 있다.

4. 결론

본 논문에서는 햇빛과 그늘 광원의 밝기, 색 차이 등을 이용하여 실외에서 효과적으로 광원의 색을 추정할 수 있는 알고리즘을 제안한다. 제안 알고리즘은 단일 조명 알고리즘으로 두 광원의 색 중에서 하나만 잘 찾는다면 어떠한 단일 조명 알고리즘과도 결합할 수 있는 장점이 있다. 유사 패치 쌍을 잘 찾을 수 있도록 유사도 측정 방식을 바꾸는 것도 좋은 결과를 얻을 수 있을 것이다.

참고문헌

- [1] A. Gijsenij, L. Rui, T. Gevers. "Color constancy for multiple light sources" IEEE Transactions on Image Processing 21.2 (2012): 697-707.
- [2] M. Ebner, "Color constancy using local color shifts" European Conference on Computer Vision. Springer Berlin Heidelberg, 2004.
- [3] Funt, B. V., Ciurea, F., & McCann, J. J. (2004). Retinex in matlab tm. Journal of electronic imaging, 13(1), 48-58.

다중 잡음 제거 영상을 이용한 Non-convex Low Rank 최소화 기법 기반 영상 잡음 제거 기법

유준상, *김종옥
고려대학교

look2017@korea.ac.kr, *jokim@korea.ac.kr

Image Denoising via Non-convex Low Rank Minimization Using Multi- denoised image

Jun-Sang Yoo, *Jong-Ok Kim
Korea University

요 약

행렬의 rank 최소화 기법은 영상 잡음 제거, 행렬 완성(completion), low rank 행렬 복원 등 다양한 영상처리 분야에서 효과적으로 이용되어 왔다. 특히 nuclear norm 을 이용한 low rank 최소화 기법은 convex optimization 을 통하여 대상 행렬의 특이값(singular value)을 thresholding 함으로써 간단하게 low rank 행렬을 얻을 수 있다. 하지만, nuclear norm 을 이용한 low rank 최소화 방법은 행렬의 rank 값을 정확하게 근사하지 못하기 때문에 잡음 제거가 효과적으로 이루어지지 못한다. 본 논문에서는 영상의 잡음을 제거 하기 위해 다중 잡음 제거 영상을 이용하여 유사도가 높은 유사 패치 행렬을 구성하고, 유사 패치 행렬의 rank 를 non-convex function 을 이용하여 최소화시키는 방법을 통해 잡음을 제거하는 방법을 제안한다.

1. 서론

최근 영상을 이용한 신호 처리 기술과 컴퓨터 비전 기술이 급격히 발달되었으며, 다양한 분야에 활용되고 있다. 하지만 영상은 취득과 송신 과정에서 발생하는 잡음에 의하여 본래의 정보가 불가피하게 왜곡되며, 이러한 영상 신호 왜곡은 후 처리로 적용되는 영상 처리 및 컴퓨터 비전 기술의 정확도에 치명적인 영향을 미치기 때문에, 전처리 기술으로써 영상 잡음 제거 기법이 필수적으로 적용되어야 한다.

영상 잡음 제거 기법은 관찰된 잡음 영상으로부터 잡음이 없는 원래의 영상을 얻는 것을 목표로 한다. 하지만 본래의 영상 신호를 잡음과 완벽하게 분리 하는 것은 불가능하기 때문에 [1], 잡음과 함께 영상이 원래 갖고 있는 세부(detail) 정보들을 잃어버리게 된다. BM3D [2], LPGPCA [3], CSR [4] 등의 기법들은 기존 다른 기법들에 비하여 뛰어난 성능을 갖고 있지만, 마찬가지로 texture 복원 성능은 아직 부족하다.

자연 영상은 스스로 높은 자기-유사도를 갖고 있으며, 이러한 자기 유사 성질은 많은 잡음 제거 기법에 이용되고 있다. 이러한 반복적인 패턴을 갖는 자기-유사도가 높은 영상들은 rank 가 낮으며, 따라서 잡음 제거를 위한 사전 정보(prior information)로 이용될 수 있다. Nuclear norm penalty 는 low rank 행렬을 얻는 가장 대표적인 convex optimization 방법이며, 행렬의 특이값(singular value)을 이용하여 쉽게 low rank 행렬을 얻을 수 있다. 하지만 nuclear norm 은 loose 한 rank 의 근사 방법으로 정확도가 낮기 때문에, low rank 행렬이 정확하게 추정되지 못하며, 잡음 제거가 효과적으로 이루어지지 못한다.

본 논문에서는 다중 잡음 제거 영상을 이용하여 non-convex rank 최소화 기법을 기반으로 한 잡음 제거 방법을 제안한다. 구체적으로, 서로 다른 level 로 잡음 제거된 영상의 동일한 위치의 패치를 통해 유사 패치 행렬을 구성하여 유사도를 높이며, 해당 유사 행렬의 rank 를 non-convex function 을 이용하여 최소화한다.

본 논문의 구성은 다음과 같다. 2 절에서는 non-convex low rank 행렬의 최소화 기법을 소개하며, 3 절에서는 다중 잡음 제거 영상을 이용하여 2 절에서 소개한 low rank 최소화 기법을 통해 잡음을 제거하는 방법을 소개한다. 4 절에서는 제안한 기법의 성능을 실험을 통해서 확인한다. 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

2. Non-convex Low rank 최소화 기법

영상의 rank 값은 실제로 계산이 어렵고 복잡하기 때문에, nuclear norm 을 rank 의 convex approximation 으로서 많이 사용한다. 하지만 rank 의 manifold 는 실제로 non-convex 하기 때문에, nuclear norm 을 이용한 근사법은 정확도가 떨어진다. 본 논문에서는 다중 잡음 제거 영상의 유사 패치 행렬을 이용하여 효과적으로 잡음을 제거하기 위해 non-convex function 을 이용하여, 더욱 정확하게 low rank 행렬을 얻는다.

$$\hat{X} = \arg \min_X \sum_{i=1}^m g(\sigma_i(X)) + \frac{1}{2} \|X - Y\|_F^2$$

Y 는 잡음 유사 패치 행렬이며, X 는 잡음이 제거된 low rank 유사 패치 행렬, g 는 non-convex function 이다.

3. Low rank 기반 잡음 제거 기법

영상내의 자기 유사도(self similarity)를 기반으로 한 잡음 제거 방법은 얻어진 유사 행렬의 유사도가 높을수록 높은 성능을 보이며, 유사도가 낮은 패치로 유사 패치 행렬을 구성할 경우 잡음 제거 성능을 저하시킨다. 따라서 유사도가 높은 유사 패치 행렬을 구성하기 위하여, 본 논문에서는 서로 다른 level 로 잡음 제거된 잡음 제거 영상을 사용한다. 구체적으로, 서로 다른 잡음 제거 영상에서 동일 위치의 패치들을 이용하여 유사 패치 행렬을 구성한다. 유사 패치들은 목표(target) patch 에 대하여 비지역적으로 검색하여, Euclidean 거리를 기반으로 잡음 level 과 동일하게 잡음 제거된 영상에서 구한다.

Euclidean 거리를 기반으로 얻은 잡음 유사 패치 행렬(Y)을 non-convex low rank 최소화 기법을 통해 잡음 제거 유사 패치 행렬 X 를 얻는다. 유사도가 높은 서로 다른 잡음 제거 영상의 패치들을 이용하였기 때문에 low rank 최소화 기법으로 인하여 영상의 잡음이 효과적으로 제거되며, 영상의 세부 정보들을 보존시킬 수 있다.

4. 실험 결과

본 연구의 실험은 잡음 제거 연구에 많이 사용되는 5 개의 영상을 사용하였으며, $\sigma = 20$ 인 잡음을 원본 영상에 더해 주어 CSR [2], BM3D [3], LPGPCA [4] 의 3 가지 비교 기법과 제안 기법을 이용해 잡음을 제거하여 성능을 평가하였다. 제안 기법에서는 BM3D 기법을 이용하여 $\sigma = 20$ 인 잡음 영상을 서로 다른 3 가지 잡음 레벨 ($\sigma = 15, 17, 20$)로 잡음 제거 하였다. 또한 Non-convex function 은 l_p norm 을 사용하였다.

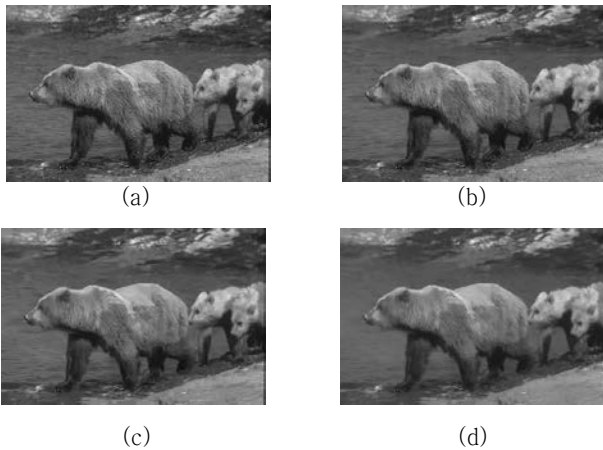


그림 1. 주관화질 비교 평가 (a) 원본 영상 (b) CSR (c) BM3D (d) LPG-PCA (e) 제안 기법

표 1. PSNR 비교 평가

	CSR	BM3D	LPGPCA	Proposed
Monkey	26.23	26.27	26.05	26.41
Bear	29.61	29.57	29.27	29.70
Horse	27.62	27.63	27.20	27.80
Bridge	27.15	27.22	26.87	27.41
Seal	27.14	27.16	26.91	27.31

5. 결론

본 논문에서는 다중 잡음 제거 영상을 이용한 non-convex low rank 최소화 기법 기반 잡음 제거 기법을 제안하였다. 기존의 비지역적 유사도 기반의 잡음 제거 연구들은 유사 패치의 유사도가 낮아 잡음 제거의 정확도가 높지 않았다. 본 연구에서는 다중 잡음 제거 영상을 이용하여 유사도가 높은 유사 패치 행렬을 구성하여 non-convex low rank 최소화 기법을 통해서 효과적으로 잡음을 제거하였다.

참고 문헌

[1] Levin, Anat, and Boaz Nadler. "Natural image denoising: Optimality and inherent bounds." *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.

[2] Dong, Weisheng, et al. "Sparsity-based image denoising via dictionary learning and structural clustering." *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011.

[3] Dabov, Kostadin, et al. "Image denoising by sparse 3-D transform-domain collaborative filtering." *IEEE Transactions on image processing* 16.8 (2007): 2080-2095.

[4] Zhang, Lei, et al. "Two-stage image denoising by principal component analysis with local pixel grouping." *Pattern Recognition* 43.4 (2010): 1531-1549.

그래프 간 정합을 이용한 포인트 클라우드 시퀀스 압축

이선호, 김지수, 이세호, 김창수
고려대학교

{seonholee, jisukim, seholee}@mcl.korea.ac.kr, changsukim@korea.ac.kr

Point Cloud Sequence Compression by Matching between Graphs

Seonho Lee, Ji-Su Kim, Se-Ho Lee, Chang-Su Kim
Korea University

요 약

본 논문에서는 그래프 간 정합을 이용한 포인트 클라우드 시퀀스 압축 기법을 제안한다. 우선, 그래프를 활용하여 포인트 클라우드 시퀀스의 시변하는 기하학적 구조를 표현하고, 그래프로부터 웨이블릿 변환을 사용하여 추출한 특징 벡터를 매칭하는 방법으로 인접 프레임 간 움직임 예측을 수행한다. 그리고 움직임 예측을 통해 얻은 움직임 벡터 중 정합 점수가 높은 소수의 움직임 벡터를 보간하여 프레임 전체의 움직임 필드를 얻는다. 최종적으로 움직임 정보를 활용하여 얻은 예측 프레임과 타겟 프레임의 차이를 선택적 엔트로피 부호화 방식으로 코딩하여 포인트 클라우드 시퀀스 압축을 수행한다. 실험 결과 제안하는 기법이 3D 포인트 클라우드 시퀀스를 효과적으로 압축함을 확인할 수 있다.

1. 서론

3D 포인트 클라우드 시퀀스는 애니메이션, 게임, 가상 현실, 몰입형 통신 등 다양한 분야에서 점점 널리 쓰이고 있다. 다양한 분야에서의 응용을 위해서는 3D 포인트 클라우드 시퀀스의 효과적인 압축이 필수적이다. 그러나 효과적인 압축에 필요한 움직임 벡터 추측은 여전히 어려운 과제로 남아있다. 이는 포인트 클라우드 시퀀스를 생성하는 과정에서 RGBD 카메라에 의해 순간적으로 포착된 점의 개수가 매 순간 다르기 때문에, 각 포인트 클라우드 프레임을 구성하는 점의 개수가 계속해서 변하기 때문이다.

본 논문에서는 기존의 포인트 클라우드 프레임 간 움직임 필드 추측 기법과 인접한 두 프레임의 팔진트리 기반 압축 방법을 결합한 3D 포인트 클라우드 시퀀스 압축 방법을 제안한다.

2. 제안 기법

2.1 팔진 트리의 그래프 표현

3D 포인트 클라우드로부터 만들어진 팔진트리의 잎사귀 노드는 공간 상의 복셀과 대응된다[1]. 점을 하나 이상 포함하는 점유된 복셀은 불규칙하게 분포하는데 이를 효과적으로 표현하기 위해서 그래프로 표현한다. [2]와 같이 모든 점유된 복셀을 방향이 없는 가중 그래프 $G = (V, E, W)$ 로 나타낸다. V 와 E 는 G 의 꼭짓점과 에지를 의미한다. 꼭짓점은 팔진트리의 각 점유된 잎사귀 노드에 대응하고, 에지의 연결성은 k -최근접 이웃으로 정한다. 본 논문에서는 $k = 26$ 을

사용한다. G 의 가중치 행렬 $W(i, j)$ 는 꼭짓점 i 와 j 를 잇는 에지의 가중치로 구성되고, 이 가중치 값은 두 꼭짓점 간의 유클리드 거리에 반비례한다.

2.2 움직임 예측

A. 특징 벡터 추출

먼저, 그래프 G 상의 i 번째 꼭짓점에 대하여 [3]과 같이 팔분면 지시 함수 $o_{k,i}, \forall k = [1, 2, \dots, 8]$ 를 정의한다. 예를 들어, 1 팔분면에 대한 지시 함수는 다음과 같다.

$$o_{1,i}(j) = 1_{\{x(j) \geq x(i), y(j) \geq y(i), z(j) \geq z(i)\}}(j) \quad (1)$$

V 에 속하는 꼭짓점 j 의 좌표가 꼭짓점 i 의 좌표보다 큰 경우에만 $o_{1,i}(j)$ 는 0이 아닌 값을 갖는다. 팔분면 지시함수는 각 꼭짓점이 팔진트리에서 꼭짓점 i 에 대해 어떤 방향에 위치하는지를 나타낸다.

다음으로, 위치와 색상 정보에 기초한 그래프 스펙트럴 특징을 계산한다. V 에 속하는 모든 꼭짓점 i 에서 [4]와 같이 스펙트럴 그래프 웨이블릿 $\psi_{s,i}$ 를 구한다. 꼭짓점 i 에서 지시 방향 k 와 위치 및 색상 특징 정보 $f \in \{x, y, z, r, g, b\}$ 를 독립적으로 고려하여 그래프 웨이블릿 특징 벡터를 다음과 같이 계산한다.

$$\phi_{i,s,o_{k,i},f} = \langle f \cdot o_{k,i}, \psi_{s,i} \rangle \quad (2)$$

$s \in S = \{s_1, \dots, s_{\max}\}$ 는 웨이블릿의 스케일을 의미하고 \cdot 은 벡터의 단순곱을 의미한다. 그래프 웨이블릿을 계산할 때 사용하는 스케일링 함수 h 를 고려한 특징 벡터는 다음과 같이 정의된다.

$$\phi_{i,h,o_{k,i,f}} = (f \cdot o_{k,i}, h(L)\delta_i) \quad (3)$$

최종적으로, 특징 벡터 $\phi_i = [\phi_{i,s,o_{k,i,f}}, \phi_{i,h,o_{k,i,f}}]$ 를 얻는다.

B. 움직임 필드 계산

참조 프레임 I_t 와 목표 프레임 I_{t+1} 내 그래프들의 꼭짓점들을 정합을 수행하여 움직임 정보를 추정한다. 움직임 정보를 추정하기 위해 다음과 같이 특징 벡터 간의 정합 점수를 계산한다.

$$\sigma(m, n) = (\phi_m - \phi_n)^T P (\phi_m - \phi_n), \forall m \in \mathcal{V}_t, n \in \mathcal{V}_{t+1} \quad (4)$$

꼭짓점 n 은 정합 점수가 가장 작을 때의 꼭짓점 m 과 매칭되고, 두 꼭짓점의 좌표의 차이가 이들의 움직임 벡터를 나타낸다.

그리고 타겟 프레임 I_{t+1} 을 K -평균 군집화한 후, 각 군집에서 가장 좋은 정합 점수를 갖는 군집 대표 꼭짓점 쌍을 찾는다. [3]과 같이 군집 대표 꼭짓점 쌍들의 움직임 벡터를 보간하는 방법으로 다른 모든 꼭짓점들의 움직임 벡터를 계산하여 움직임 필드를 얻는다.

2.3 포인트 클라우드 압축

A. 움직임 벡터 정보 압축

움직임 벡터 정보는 그래프 푸리에 도메인에서 효과적으로 압축 가능하다[5]. 따라서 [6]과 같이 움직임 벡터를 G 의 고유 벡터들과 내적하여 그래프 푸리에 변환을 수행한 후, 적절한 크기로 양자화하여 RLGR 엔트로피 코딩한다.

B. 기하 정보 압축

참조 프레임의 꼭짓점의 좌표 값에 움직임 필드를 더하여 움직임 보정 프레임을 얻는다. 움직임 보정 프레임과 타겟 프레임의 기하 정보는 유사하지만, 여전히 두 프레임의 꼭짓점의 개수가 다르다는 문제가 있다. 이를 해결하기 위해, 타겟 프레임과 움직임 보정 프레임으로부터 [7]의 방법을 사용하여 예측 프레임을 만든다.

본 논문에서는 예측 프레임과 타겟 프레임 사이의 차이를 효과적으로 압축하는 방법을 제안한다. 우선 예측 프레임과 타겟 프레임의 차이를 두 팔진트리의 XOR 연산을 통해 얻는다. 그 후에 프레임 간 차이를 Huffman 부호화와 산술 부호화 중 더 압축률이 좋은 방식을 선택하여 엔트로피 코딩한다. 이러한 방법을 적용하는 경우, 프레임 간 차이의 정도에 따라서 더 효율적인 압축이 가능하다.

C. 색상 정보 압축

움직임 보정 프레임의 꼭짓점 j 와 k -최근접 이웃($k = 3$)인 참조 프레임의 꼭짓점들의 색상 정보의 평균을 해당 꼭짓점의 색상 추정 값으로 사용한다. [2]와 같이 추정된 색상 값과 실제 타겟 프레임의 색상 정보 간 차이를 그래프 푸리에 변환 하여 엔트로피 코딩한다. 이 때도 2.3.B 에서와 같이 Huffman 부호화와 산술 부호화 중 더 효율적인 것을 선택한다. 복호화를 위해서는 프레임마다 사용한 엔트로피 부호화 종류를 함께 압축해야 한다.

3. 실험 결과

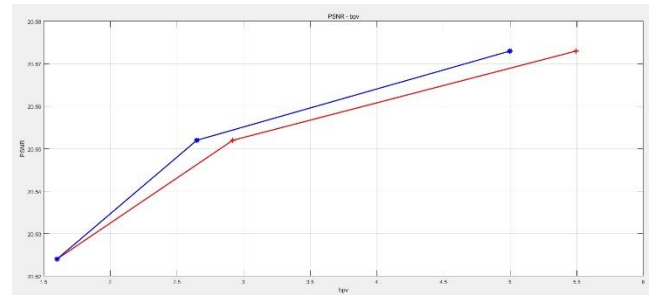


그림 1. 실험 결과.

그림 1 은 실험 결과를 보여준다. 파란 선은 제안 기법과 같이 선택적 엔트로피 부호화를 적용한 결과이다. 산술부호화만으로 압축했을 때보다 같은 PSNR 에서 더 높은 압축효율을 달성하고 있음을 확인할 수 있다..

4. 결론

본 논문은 그래프 간 정합을 이용한 포인트 클라우드 시퀀스 압축 기법을 제안하였다. 먼저 포인트 클라우드로부터 만들어진 그래프를 이용하여 특징 벡터를 추출하고, 이를 이용하여 움직임 필드를 추정한다. 그 후 프레임 간 차이를 선택적 엔트로피 부호화를 통하여 압축함으로써, 최종적으로 압축의 효율을 높였다. 실험 결과 제안 기법이 큰 왜곡 없이 포인트 클라우드 시퀀스 압축을 수행하는 것을 확인하였다.

감사의 글

이 논문은 2018 년도 정부(과학기술정보통신부)의 재원으로 '범부처 Giga KOREA 사업'의 지원을 받아 수행된 연구임 (No.GK18P0200, (4D 실감-총괄/1 세부) 4D 복원 및 동적 변형 거동 모델 기반의 초실감 서비스 기술 개발)

참고 문헌

- [1] C. Loop, C. Zhang, and Z. Zhang, "Real-time high-resolution sparse voxelization with application to image-based modeling," in *Proc. 5th High-Perform. Graph. Conf.*, 2013, pp. 73-79.
- [2] C. Zhang, D. Florencio, and C. Loop, "Point cloud attribute compression with graph transform," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2014, pp. 2066-2070.
- [3] D. Thanou, P. A. Chou, and P. Frossard, "Graph-based compression of dynamic 3D point cloud sequences," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1765-1778, Apr. 2016
- [4] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129-150, Mar. 2011.
- [5] C. Zhang, and D. Florencio, "Analyzing the optimality of predictive transform coding using graph-based models," *IEEE Signal Process. Lett.*, vol. 20, no. 1, pp. 106-109, Jan. 2013.
- [6] H. S. Malvar, "Adaptive run-length / golomb-rice encoding of quantized generalized gaussian sources with unknown statistics," in *Proc. Data Compression Conf.*, Mar. 2006.
- [7] J. Kermerl, N. Blodow, R. B. Rusu, S. Gedikli, M. Beetz, and E. Steinbach, "Real-time compression of point cloud streams," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2012, pp. 778-785.