

이미지 스티칭의 지역 차분 픽셀 평가 방법

*이성배 ** 김규현
경희대학교

*rhee@khu.ac.kr **kyuheonkim@khu.ac.kr

Local Differential Pixel Assessment Method for Image Stitching

*Rhee, Seong Bae **Kim, kyuheon
Kyunghee University

요약

이미지 스티칭은 다수의 이미지를 합성하여 카메라의 좁은 시야각(Field of View) 문제를 해결하는 기술이다. 최근 동영상 기반 Panorama, Super Resolution, 360 VR (Virtual Reality) 등의 콘텐츠 사용이 증가함에 따라, 보다 빠르고 정확한 이미지 스티칭 기술의 필요성이 커지고 있다. 지금까지 필요 성능을 만족시키기 위해 많은 알고리즘이 제안되고 있지만, 정확성을 측정하는 객관적 평가 방법은 표준화되지 않고 있다. 최근에서야 PSNR (Peak Signal-to-Noise Ratio) 과 SSIM (Structural Similarity index method) 측정값을 제시하는 방법이 주를 이루고 있지만, 본 논문에서는 PSNR 과 SSIM 측정 방식의 문제점을 밝히고 지역 차분 픽셀 평가 방법을 제안한다. 기하적 유사성과 광도 측정 정보를 포괄하는 LDPM(Local Differential Pixel Mean) 평가 방식을 테스트 이미지를 통해 증명하고 SSIM 과 비교를 통해 해당 평가 방법의 이점을 밝힌다.

1. 서론

이미지 스티칭 기술은 다수의 이미지를 하나의 이미지로 합성하는 기술이다. 각각 다른 시점에서 촬영한 이미지를 합성하여, 카메라의 좁은 시야각 문제를 해결한다. 기술 초기에는 정지 영상에 대한 처리가 주된 관심사였지만, 최근에는 동영상 기반 기술이 주를 이루고 있다.

360 VR, Super Resolution, Panorama 에서 사용되는 스티칭 기술은 정지 영상 기반에서 보다 빠르고 정확함을 요구한다. 스티칭 기술의 속도는 프로그램의 시작과 종료 시간이라는 명확한 측정 방법이 존재하지만, ‘정확도’에 대해서는 표준화된 측정 방법이 없다. 스티칭 성능을 판단하는 방법은 기술 초기에는 육안으로 결과를 판단하는 주관적 평가가 주를 이루었다가[1], 최근에는 PSNR 과 SSIM[2] 수치를 제시하는 것이 대부분이다.

PSNR 과 SSIM 을 이용하는 방법은 각각의 측정 결과가 높은 것을 좋은 성능의 알고리즘이라고 여긴다. 육안으로 한 주관적 평가와 PSNR, SSIM 을 통한 평가가 대부분 일치하므로 타당한 평가 방법이라고 생각할 수 있다. 하지만, 이미지 스티칭 과정에서 생기는 문제점과 PSNR·SSIM 의 평가 특성이 합쳐져 잘못된 성능을 . 본 논문에서는 PSNR·SSIM 평가의 문제점들을 밝히고 이 이미지 스티칭 평가의 새로운 방법을 제시하고자 한다.

2. 기존 평가 방법 분석

2.1 PSNR 정의

PSNR(Peak Signal-to-noise ratio) 은 gray-level (8 bits) MxN 사이즈 원본 이미지 f, 비교 이미지 g에 대해

서 다음과 같이 정의된다.

$$PSNR(f, g) = 10 \log_{10} \left(\frac{255}{MSE(f, g)} \right)^2 \quad (1)$$

$$MSE(f, g) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f_{ij} - g_{ij})^2 \quad (2)$$

식 (1) 의 PSNR 은 영상의 평균 자승 오차(Mean Squared Error) 분의 최대 양자화 레벨의 로그 스케일로 정의된다. 식 (2) 의 MSE 는 두 영상의 픽셀 단위 차이이다.



(a) 30.3653 (dB)



(b) 29.5176 (dB)

그림 1. PSNR 수치와 이미지

그림 1 에 PSNR 수치가 비슷한 (a), (b) 두 장의 사진이 있다. PSNR 의 수치는 (a) 사진이 더 높지만 사람의 평가는 (b) 이미지가 더 좋은 이미지라고 평가를 내릴 것이다. PSNR 이 높다는 의미는 원본과 픽셀 정보가 유사하다는 것에 불과하다.

2.2 이미지 스티칭과 PSNR

동영상 스티칭의 경우 영상에서 움직이는 물체, 각 카

메라 간의 시차(parallax), 부정확한 이미지 변환 관계식(Homography) 사용 등의 이유로 그림 2 (a)와 같이 특정 물체가 여러 위치에서 나타나는 ‘고스트 (Ghost) 현상’이 생길 수 있다. 또한 촬영 지점의 조명 환경 차이로 인해 그림 2 (b)와 같은 ‘솔기 (Seam)’ 에러가 생길 수 있다[3]. 이러한 인공물(Artifacts)을 제거하기 위해 블렌딩(Blending) 작업을 이미지 스티칭 기술 말단에서 수행한다. 시차로 인한 ‘고스트 현상’에 대해서는 이미지 가중치 설정을 통한 모션 보상(Motion compensation) 등을 진행하고, 조도 차이로 인한 ‘솔기’ 부분에는 Multiband blending, 톤맵(Tone map) 블렌딩 작업 등이 수행된다[3,4,5]. 더 나은 성과물을 위해서 블렌딩을 적용한 이미지는 원본 영상과 픽셀 정보 차이가 더 커지기 때문에 오히려 PSNR 은 더 낮아진다. 이에 PSNR 은 적절한 측정 방법이 될 수 없다.



그림 2. 인공물 예시 (a) Ghost Effect (b) Seam

2.3 SSIM 과 이미지 스티칭

SSIM 은 gray - level (8 bits) MxN 사이즈 원본 이미지 f, 비교 이미지 g 에 대해서 다음과 같이 정의된다.

$$SSIM(f,g) = I(f,g)c(f,g)s(f,g) \tag{3}$$

$$\begin{cases} I(f,g) = \frac{2\mu_f\mu_g+C_1}{\mu_f^2+\mu_g^2+C_1} \\ c(f,g) = \frac{2\sigma_f\sigma_g+C_2}{\sigma_f^2+\sigma_g^2+C_2} \\ s(f,g) = \frac{2\sigma_{fg}+C_3}{\sigma_f\sigma_g+C_3} \end{cases} \tag{4}$$

- μ_x : x 의 평균
- σ_x^2 : x 의 분산
- σ_{xy} : x, y 의 공분산
- C_1, C_2, C_3 : 안정화 변수

SSIM 은 식 (3)에서 보듯이 휘도(I), 대비(c), 구조(s) 정보의 곱으로 표현된다. 이미지를 작은 윈도우 단위로 픽셀들을 묶어서 각각의 SSIM 을 구한 후, 그 값들의 평균을 통해 MSSIM (Mean SSIM) 을 구하여 [0, 1] 사이의 값을 평가 지수로 제시한다[2].

이미지 스티칭의 성능을 가장 잘 나타내는 방법은 단연 SSIM 이라고 할 수 있다. 하지만 SSIM 의 휘도, 대비, 구조는 식 (4)처럼 밝기 정보의 평균, 분산, 공분산 들로 구성된다. 따라서 MSSIM 은 구조 유사도 평균들의 평균 값이므로 픽셀 단위 정보가 사라지는 문제가 있다. 이는 PSNR 또한 마찬가지이다.

픽셀 단위의 정보는 더 나은 스티칭 구현을 위한 파라미터로 활용될 수 있다. 이런 픽셀 단위 정보가 사라지는 문제를 해결하기 위해 본 논문에서는 픽셀 단위 기반

스티칭 이미지 평가 방법을 3 장에서 제시한다.

3. LDPM

3.1 측정 공간 정의

이미지 스티칭의 성능을 평가할 공간은 추출된 특징점을 매칭하여 구한 이미지들 간 공통 영역(Overlap Region)으로 설정한다[1]. 본 논문에서 이미지 전체가 아닌 중첩 영역에 대해서만 평가를 적용하는 이유는 호모그래피 추정 및 블렌딩 작업이 공통 영역을 기준으로 계산되고 적용되기 때문이다[3-5]. 중첩되지 않은 영역을 평가에 추가하는 것은 비교 공간의 크기를 늘려 오류의 민감도를 낮추게 된다.

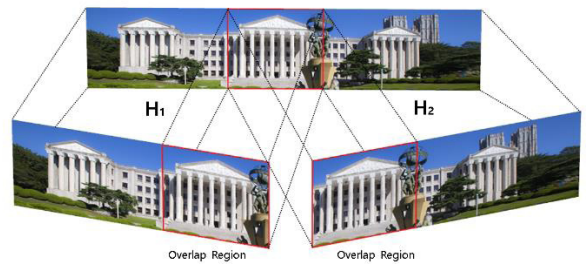


그림 3. Homography 와 Overlap Region

3.2 왜곡 측정 방법

공통 영역의 gray - scale (8 bits) 크기 MxN 원본 이미지 f, 스티칭 이미지 g 에 대하여 픽셀 좌표 (x,y) 의 gray - scale 값 $I_f(x,y), I_g(x,y)$, span = {(x, y) | x, y 는 공통 영역 좌표 공간}에 대하여 다음 사항들을 정의한다.

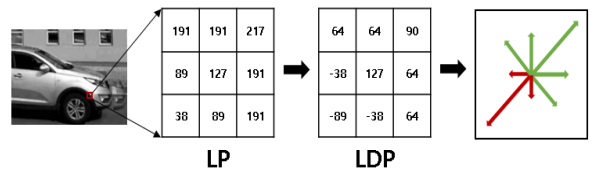


그림 4. LP 와 LDP

$$LP(I(x,y)) = \begin{bmatrix} I(x-1,y-1) & I(x,y-1) & I(x+1,y-1) \\ I(x-1,y) & I(x,y) & I(x+1,y) \\ I(x-1,y+1) & I(x,y+1) & I(x+1,y+1) \end{bmatrix} \tag{5}$$

이미지 픽셀 좌표 (x, y) 에 대하여 (x-1, y-1) ~ (x+1, y+1) 좌표까지의 주변 픽셀을 묶어 식 (5)의 3x3 의 그림 4 의 LP 구조를 생성한다. 각 LP 구조에는 9 개의 gray - scale 값이 지정된다.

$$LDP(I(x,y)) = \begin{bmatrix} D_1 & D_2 & D_3 \\ D_4 & I(x,y) & D_5 \\ D_6 & D_7 & D_8 \end{bmatrix} \tag{6}$$

- $D_1: I(x-1,y-1) - I(x,y)$, $D_2: I(x,y-1) - I(x,y)$
- $D_3: I(x+1,y-1) - I(x,y)$, $D_4: I(x-1,y) - I(x,y)$
- $D_5: I(x+1,y) - I(x,y)$, $D_6: I(x-1,y+1) - I(x,y)$
- $D_7: I(x,y+1) - I(x,y)$, $D_8: I(x+1,y+1) - I(x,y)$

LP 구조의 외각 픽셀 각각에서 중앙 픽셀 값을 빼 식 (6)의 LDP 구조를 생성한다. LP 구조의 가장자리 값에서 중앙값의 차이 값인 $D_1 \sim D_8$ 은 두 픽셀 간의 밝기 차이만 저장되므로 LDP 구조 비교는 블렌딩 효과로부터 영향이 적어지게 된다.

$$LDPM(f, g) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \text{matc} \boxtimes LDP(I_f(i, j)), LDP(I_g(i, j)) \quad (7)$$

LDP 구조는 그림 4에서 보듯이 방향이 고정된 벡터와 같다. 식 (7)은 이미지 전체에 대해서 두 공통 영역 이미지의 동일 좌표에서의 LDP 벡터 구조를 비교하기 때문에 이미지 간 구조적 유사성을 비교하는 평가 방법이다.

$$\text{matc} \boxtimes LDP(I_f(i, j)), LDP(I_g(i, j)) = \begin{cases} 1 & \text{if } e < 8 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$e = \sum_{k=1}^8 cnt, cnt = \begin{cases} 1 & \text{if } |D_{fk} - D_{gk}| < \theta_{t\Delta} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

· LDPM (f, g): 이미지 f, g 구조적 차이

- e: 에러 픽셀 개수 cnt: LDP 매칭 횟수
- D_{fk} : 이미지 f의 LDP $D_1 \sim D_8$ D_{gk} : 이미지 g의 LDP $D_1 \sim D_8$
- $\theta_{t\Delta}$: 에러 임계값

식(8)의 두 LDP를 비교하는 방법은 각각의 $D_1 \sim D_8$ 가 같은 값인지를 비교하는 것이다. 하나라도 다르다면 에러 픽셀로 여겨 1을, 모두 같으면 에러가 없는 픽셀이므로 0으로 설정한다. 식 (9)에서 두 LDP의 $D_1 \sim D_8$ 를 비교하는 과정에서 $\theta_{t\Delta}$ 의 임계 값을 설정하는 이유는 이미지의 색공간 이동 및 호모그래피 변환 과정에서 달라질 수 있는 값을 보상하기 위함이다. 본 논문의 실험에서는 임계값을 0으로 설정하였다.

LDP 구조를 스티칭 이미지 평가에서 스티칭 이미지를 f, 좌, 우 이미지를 m, g라고 한다면, 호모그래피 값을 이용하여 이미지 f, g, m의 공통 영역의 구역을 일치시키는 것을 시작으로 한다. 공통 영역에서 LPDM(f, m)과 LPDM(f, g)의 평균을 구해 이미지 스티칭에서 발생한 오류 픽셀의 비율을 구해 이미지 스티칭을 평가하는 방법을 제안한다.

4. 실험 결과

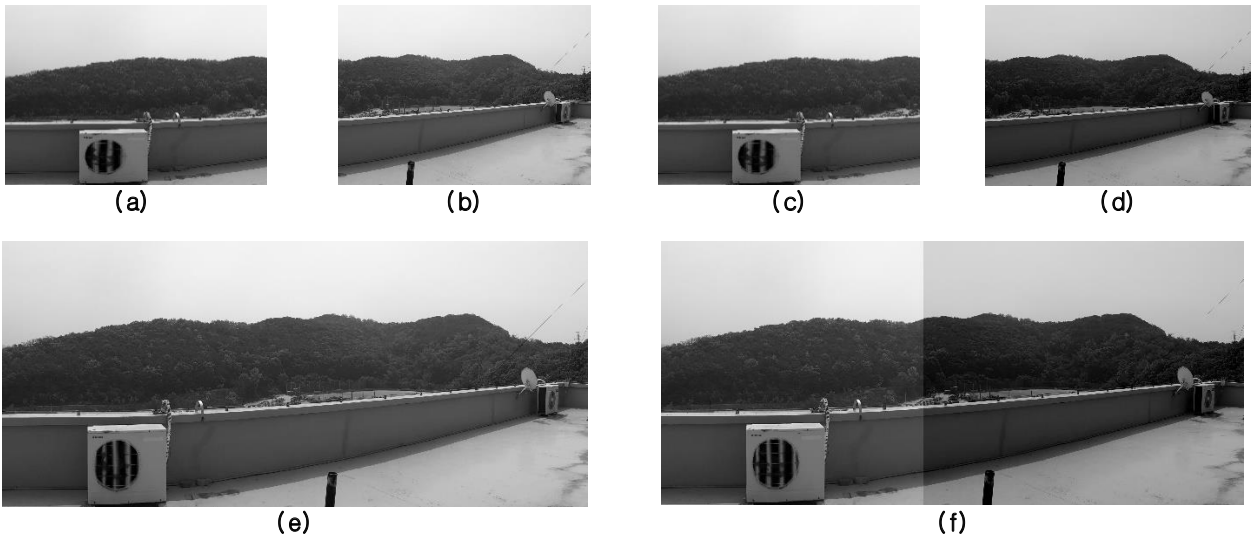


그림 5. 테스트 이미지

Test(Overlap)	SSIM	Proposed(Error)
(a), (b) and (e)	1.0000	0 %
(c), (d) and (e)	0.9402	0.9315 %
(c), (d) and (f)	0.9335	1.7605 %

표 1. 이미지 스티칭 평가

그림 5의 (a), (b), (c), (d)는 이미지 스티칭을 진행할 원본 영상에 해당하고, (e), (f)는 스티칭된 이미지이다.

(a), (b), (e)의 테스트에서는 오류가 없는 이미지 스

티칭을 가정한다. SSIM 평가와 제안된 평가 방법 모두 오류가 없는 결과가 나왔다.

(c), (d), (e)의 테스트는 다른 조명 환경을 가지는 촬영 시점에서 블렌딩 효과를 통해 솔기가 없는 스티칭 이미지 생성을 가정한다.

마지막으로 (c), (d), (f)의 테스트는 블렌딩 효과를 적용하지 않은 솔기가 있는 이미지 스티칭 결과를 가정한다.

(a), (b), (e)의 테스트, (c), (d), (e)의 테스트, (c), (d), (f)의 테스트의 순서로 이미지 스티칭 결과가 좋다는 것이 육안으로 실시한 주관적인 평가이다. 표 1에서 주관적인 평가 순서와 SSIM 평가, 그리고 제안된 평가가 모두 일치하므로 제안된 평가 방법은 스티칭 이미지를

올바르게 평가한다.

제한된 평가 방법은 픽셀 단위 정보를 가지고 있기 때문에 스티칭 이미지 평가 외에 추가적인 활용이 가능하다. 딥러닝 시스템에서 스스로 평가하고 에러로 여겨지는 픽셀의 좌표와 정보를 토대로 더 나은 스티칭 이미지를 생성하는 과정의 참조 파라미터로 사용될 수 있다.

5. 결론

이미지 스티칭 평가에서 SSIM 평가 지표는 납득할 만한 결론을 나타낸다. 하지만 SSIM 값은 평균 정보만을 담고 있기 때문에 픽셀 단위 정보는 사라지게 되는 문제가 있다. 이미지 스티칭을 진행하는 주체가 사람이라면 SSIM 으로 만족할 수 있다. 하지만 최근 딥러닝을 통한 이미지 스티칭 시도가 늘어나고 있다. 아직은 딥러닝을 통한 스티칭 이미지의 품질은 낮지만, 전처리 단계인 호모그래피 추정에는 어느 정도에 올라와있다[6]. 딥러닝 시스템의 스스로 판단하는 과정에서 스티칭 영상의 오류 픽셀의 정보가 필수적이기 때문에 픽셀 단위의 이미지 스티칭 평가 방법이 필요하다.

픽셀 단위 평가를 진행하면서 오류로 판단되는 픽셀을 버퍼에 저장하고, 오류 픽셀들의 좌표들을 묶어 라벨링(Labeling)을 진행하여 패턴들을 규정한다면 오류의 종류까지 분류가 가능해 딥러닝 기반 이미지 스티칭에서 추가적인 활용까지 기대된다.

본 논문에서는 LDPM 평가 방법이 이미지 스티칭 성능을 평가할 수 있는지의 가능성을 확인하고자 하였다. 스티칭 이미지 평가가 가능하다는 것을 확인한 시점에서 이제는 LDPM 의 블록 사이즈를 높이거나 SSIM 값들을 LDPM 로 묶는 방법 등으로 진행된다면 딥러닝 이미지 스티칭에 보다 적합한지 연구가 필요하다.

* 본 연구는 한국전력공사의 2018 년 착수 에너지 거점대학 클러스터 사업에 의해 지원되었음 (과제번호:R18XA02)

참고문헌

- [1] Matthew Brown and David G. Lowe. "Automatic panoramic image stitching using invariant features " International Journal of Computer Vision. Vol. 74, No.1, pp.55 - 73. 2007.
- [2] Wang, Zhou, et al. "Image Quality Assessment: from error visibility to structural similarity. " IEEE transactions on image processing, Vol.13, No.4, pp .600 - 612. 2004.
- [3] R. Szeliski, "Image Alignment and Stitching: A Tutorial. " Foundations and Trends in Computer Graphics and Computer Vision, Vol. 2, No.1, 2006.
- [4] A. Zomet, A. Levin, S. Peleg, Y. Weiss, "Seamless image stitching by minimizing false edges " IEEE Trans. Image Process, Vol. 15, No.4, pp.969 - 977, 2006.
- [5] Eden, Ashley, Matthew Uyttendaele and Richard Szeliski, "Seamless image stitching of scenes with large motions and exposure differences. ", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06), Vol. 2, 2006.
- [6] DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation. " arXiv preprint arXiv:1606.03798, 2016. .