

특집논문 (Special Paper)

방송공학회논문지 제24권 제5호, 2019년 9월 (JBE Vol. 24, No. 5, September 2019)

<https://doi.org/10.5909/JBE.2019.24.5.775>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

영상 스티칭의 지역 차분 픽셀 평가 방법

이성배^{a)}, 강전호^{a)}, 김규현^{a)†}

Local Differential Pixel Assessment Method for Image Stitching

Seongbae Rhee^{a)}, Jeonho Kang^{a)}, and Kyuheon Kim^{a)†}

요 약

영상 스티칭은 다수의 영상을 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 영상 스티칭 기술의 필요성이 커지고 있다. 또한, 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 PSNR과 SSIM 측정 방식의 문제점을 밝히고, 해당 방법의 한계점을 극복하여 기하적 유사성과 광도 측정 정보를 포괄하는 지역 차분 픽셀 평가(LDPM: Local differential pixel mean)방법을 제안한다. 또한, 본 논문에서 제안하는 LDPM(Local differential pixel mean) 평가 방식을 테스트 영상을 통해 증명하고 SSIM과 비교를 통해 해당 평가 방법의 이점을 밝힌다.

Abstract

Image stitching is a technique for solving the problem of narrow field of view of a camera by composing multiple images. Recently, as the use of content such as Panorama, Super Resolution, and 360 VR increases, the need for faster and more accurate image stitching technology is increasing. So far, many algorithms have been proposed to satisfy the required performance, but the objective evaluation method for measuring the accuracy has not been standardized. In this paper, we present the problems of PSNR and SSIM(Structural similarity index method) measurement methods and propose a Local Differential Pixel Mean method. The LDPM evaluation method that includes geometric similarity and brightness measurement information is proved through a test, and the advantages of the evaluation method are revealed through comparison with SSIM.

Keyword : Image Stitching, Assessment Method, PSNR, SSIM, LDPM

a) 경희대학교 전자공학과(Department of Electronic Engineering, KyungHee University)

† Corresponding Author : 김규현 (kyuheon Kim)

E-mail: kyuheonkim@khu.ac.kr

Tel: +82-31-201-2442

ORCID: <https://orcid.org/0000-0003-1553-936X>

※ 이 논문의 연구 결과 중 일부는 한국방송·미디어공학회 “2019년 하계학술대회”에서 발표한 바 있음.

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

※ This research was supported by Korea Electric Power Corporation. (Grant number:R18XA02).

· Manuscript received July 29, 2019; Revised September 23, 2019; Accepted September 23, 2019.

Copyright © 2016 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. 서론

현대 사회는 ‘초고화질 시대’로 향하고 있다. 2017년 5월 31일에 한국 지상파 3사가 세계 최초로 UHD(Ultra High Definition) 지상파 방송을 시작하였고, 가전업계에서는 UHD, 4K, 8K의 초고해상도를 지원하는 초고화질 초대형 TV를 경쟁적으로 출시하고 있다^[1]. 2018년에는 8K TV시장이 대증화될 것으로 업계는 예상하고 있다^[1,2].

그러나 초고해상도 영상 콘텐츠의 부족으로 인해 초고화질 시장의 생태계는 어려움을 겪고 있다^[1]. 콘텐츠 부족의 원인 중 하나로는 초고화질 장비의 가격 제한이다. 이를 극복하기 위해 비교적 저렴한 다수의 고화질 영상을 스티칭하여 초고화질 영상을 생성하는 기술을 활용함으로써 초고화질 콘텐츠 부족 현상을 극복하고자 하는 노력이 진행되고 있다^[3].

영상 스티칭 기술은 다수의 영상을 하나의 영상으로 합성하는 기술로서, 각각 다른 시점에서 촬영된 영상을 합성하여 기존 카메라의 좁은 시야각 문제를 해결하고, 인간에게 직관적인 공간 정보를 제공한다^[3]. 영상 스티칭 기술 초기에는 정지 영상에 대한 처리가 주된 관심사였지만, 최근에는 기술의 발달과 대중의 초고해상도 콘텐츠 요구로 동영상 기반 영상 스티칭이 주를 이루고 있다.

360 VR, Super Resolution, Panorama 등 다양한 콘텐츠에서 사용되는 동영상 스티칭 기술은 기존 정지 영상 기반의 기술보다 빠르고 정확함을 요구한다. 동영상 스티칭 기술의 속도는 프로그램의 시작과 종료까지의 시간이라는 명확한 측정 기준이 존재하지만, 스티칭 기술의 ‘정확도’에 대해서는 아직 표준화된 것이 없다. 그렇기에 동영상 스티칭 기술 초기에는 성능 평가를 육안으로 수행하는 ‘주관적 평가’가 주를 이루었다가^[4], 최근에는 PSNR과 SSIM^[5] 수치를 제시하는 것으로 변화하였다.

PSNR과 SSIM을 이용하여 성능을 판단하는 방법은 PSNR, SSIM 각각의 측정값이 높게 나온 영상을 좋은 성능의 합성 영상이라고 판단한다. 일반적으로 육안으로 한 주관적 평가와 PSNR, SSIM을 통한 평가가 일치하는 결과를 제공하고 있으나, PSNR과 SSIM으로 평가하는 것과 육안으로 평가한 결과가 일치하지 않는 경우도 발생한다. 따라서, 본 논문에서는 PSNR과 SSIM 기반 영상 스티칭 결과

평가 방법의 문제점을 밝히고 영상 스티칭 평가의 새로운 방법을 제시하고 추가적인 이점 또한 밝히고자 한다.

II. 배경 기술

1. 영상 유사도 측정을 위한 PSNR 및 SSIM

PSNR(Peak signal-to-noise ratio)이란 신호가 가질 수 있는 최대 신호 크기 대 잡음의 비로 정의된다. 대표적으로 PSNR이 사용되는 부분은 영상의 압축과 전송 시에 영상 평가 및 화질 측정에 사용되는 경우이다. 영상 처리 시에 잡음은 원본 영상과 비교 영상의 픽셀 정보 차이이며, PSNR이 큰 영상을 잡음이 적은 좋은 품질의 영상이라고 정의한다. PSNR은 인간의 시각 능력을 반영하여 최대 신호 크기 대 잡음 비를 dB 단위로 표현한다. 일례로, 8비트의 양자화 레벨을 갖는 $M \times N$ 크기의 흑백 원본 영상 f 와 손실 영상 g 에 대한 PSNR의 계산식은 다음과 같다.

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

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

상기 식 (1)의 분자는 영상이 가질 수 있는 최대 신호의 크기이기에, 8비트 양자화 레벨의 최댓값인 255로 설정하였다. PSNR은 영상의 평균 자승 오차(Mean Squared Error)분의 최대 양자화 레벨의 제곱을 로그 스케일로 정의하며, 식 (1)에 나타난 MSE는 식 (2)와 같이 두 영상의 같은 위치에 대한 픽셀 정보의 분산을 의미한다.

PSNR 평가 방법은 PSNR이 높은 영상을 좋은 품질의 영상이라고 정의하지만, PSNR이 높은 영상이 좋은 품질의 영상으로 평가되기 위해서는 원본 영상이 최상의 품질이라는 전제조건이 필요하다. 전제조건을 만족하지 않은 상황에서 PSNR은 아래 그림 1과 같은 잘못된 평가를 내릴 수 있다.

그림 1에는 PSNR 수치가 비슷한 (a), (b)의 사진 두 장이 있다. PSNR의 수치는 영상 (a)가 더 높지만, 인간의 주관적 화질 평가로 판단한 경우, (b)가 해당 사진의 물체 형상



그림 1. PSNR과 영상
 Fig. 1. PSNR and Image

을 보다 구체적으로 제공하고 있기에, (b)를 상대적으로 화질이 더 좋은 영상이라고 평가를 내린다. 이를 통해서 확인할 수 있는 사항은 PSNR이 높다는 의미는 영상의 절대적인 품질이 좋다는 의미보다는 원본 영상과 비교 영상의 픽셀 정보가 유사하다는 것을 확인할 수 있다.

또한, 영상 스티칭의 결과를 측정하는 방법으로 사용되는 SSIM(Structural similarity index method)은 다른 영상 품질 평가 방법들과 달리 인간이 영상을 평가하는 시스템을 기반으로 고안한 방법이다. SSIM은 인간의 시각 체계에 큰 역할을 하는 휘도(luminance), 명암비(contrast), 구조(structure)의 값을 연립하여 전체 영상의 구조적 유사도를 파악하고 이를 통해 영상의 품질을 평가하는 방법이다. 평가 결과는 0부터 1 사이의 값을 가지며, 1에 가까울수록 좋은 품질의 영상으로 평가된다. 일례로, $M \times N$ 크기의 원본 영상 f 와 손실 영상 g 를 가정했을 때, SSIM의 계산식은 다음 식으로 정의된다.

$$SSIM(f, g) = l(f, g)c(f, g)s(f, g) \quad (3)$$

$$\begin{cases} l(f, g) = \frac{2\mu_f\mu_g + c_1}{\mu_f^2 + \mu_g^2 + c_1} \\ c(f, g) = \frac{2\sigma_f\sigma_g + c_2}{\sigma_f^2 + \sigma_g^2 + c_2} \\ s(f, g) = \frac{2\sigma_{fg} + c_3}{\sigma_f\sigma_g + c_3} \end{cases} \quad (4)$$

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

SSIM은 식 (3)에서 보듯이 휘도(l), 명암비(c), 구조(s) 정보의 곱 연산으로 표현된다. 영상의 휘도, 명암비, 구조 각각의 함수는 식 (4)와 같이 정의되며, 식 (4)에서 μ_f 와 μ_g 는 각각 신호 f 와 g 의 평균을 의미하고, σ_f 와 σ_g 는 신호 f 와 g 의 표준편차를, σ_{fg} 는 신호 f 와 g 의 공분산을 의미한다. C_1, C_2, C_3 는 분모가 0이 되는 경우를 방지하기 위한 안정화 변수이다. 최종 평가 지수를 구하기 위해 전체 영상을 크기의 작은 윈도우(window) 단위로 분할하고, 각 윈도우 단위로 계산한 SSIM의 평균인 MSSIM(Mean SSIM)을 구하여 0부터 1 사이의 평가 지수로 제시한다^[5].

인간의 시각 체계를 모방한 SSIM은 식 (3)에서 보는 바와 같이, 휘도, 명암비, 구조 정보에서만 동작하기 때문에 영상 간 색 왜곡에는 취약하다는 한계가 존재한다. 일례로, 그림 2 (a)에는 원본 영상을 그림 2 (b)에는 컬러 톤 왜곡 영상을 제공하였으며, 인간의 주관적 판단으로 두 영상의 컬러 차이를 분명하게 인지할 수 있지만, 두 영상에 대한 MSSIM를 계산한 결과값은 0.9900으로서 매우 높은 수치를 제공하기에 MSSIM을 통해서 두 영상 간 차이가 거의 없다고 평가한다.

이와 같이, 현재 영상 스티칭 결과에 대한 화질 측정으로 활용되고 있는 PSNR 및 SSIM 평가 방법은 절대적 품질 평가와 색 왜곡 측면에서 각각 그 한계점을 포함하고 있는 것을 확인할 수 있다.



(a)



(b)

그림 2. 원본 영상과 컬러 톤 왜곡 영상
Fig. 2. Original image and color distorted image

2. 영상 스티칭 결과 측정을 위한 PSNR 및 SSIM

동영상 기반의 스티칭에서는 영상 속 움직이는 물체 또는 각 카메라 간의 시간 동기화 차이 등으로 발생하는 시차(Parallax) 그리고 부정확한 영상 변환 관계식(Homography) 사용 등의 이유로 그림 3 (a)와 같이 특정 물체가 여러 위치에서 나타나는 ‘고스트(Ghost) 현상’이 생길 수 있다. 또한 촬영 지점의 조명 환경 차이로 인해 그림 3 (b)와 같은 ‘솔기(Seam)’ 에러가 생길 수 있다⁶⁾.

‘고스트 현상’과 ‘솔기’ 같은 인공물(artifacts)을 제거하기 위해, 영상 스티칭 기술의 후반부에서는 블렌딩(Blending) 작업을 일반적으로 수행한다. 시차로 인한 ‘고스트 현상’에 대해서는 영상 가중치 설정을 통한 모션 보상(Motion compensation) 등을 진행하고, 조도 차이로 인한 ‘솔기’ 부

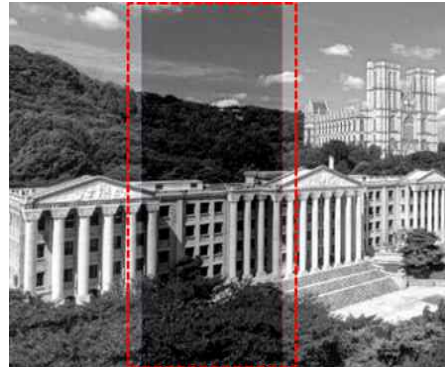
분에는 Multiband blending, 톤맵(Tone map) 블렌딩 방법 등이 수행된다^{6,7,8)}.

앞 1절에서 설명한 바와 같이, PSNR은 원본 영상과의 픽셀 정보 유사성을 측정하는 방법이다. 동영상 스티칭에서는 상기의 고스트 및 솔기 에러를 제거하기 위하여 블렌딩 작업을 일반적으로 수행하며, 이를 적용한 영상은 원본과의 픽셀 정보가 달라지기 때문에, 오히려 PSNR이 더 낮게 나타나는 경향이 있다. 그러므로, Billah 등의 논문⁹⁾ 같이 PSNR 기반의 영상 스티칭 성능 평가는 적절한 방법이 될 수 없다.

반면, SSIM의 경우, 동영상 스티칭 작업에서 수행되는 블렌딩 작업이 색채 값의 변화를 가지고 오지만, 색채 보정에는 SSIM이 둔감하기에 영향을 주지 않으며, 반면에 블렌딩을 통해 영상 간 구조적 유사도를 증가시키고 있기에,



(a)



(b)

그림 3. 인공물 예시 (a) Ghost Effect (b) Seam
Fig. 3. Examples of artifacts (a) Ghost Effect (b) Seam

SSIM 평가에서 활용되는 영상의 휘도, 명암비, 구조 함수의 값을 증가시킴으로써, 결과적으로 스티칭 영상 평가에 보다 정확한 결과를 나타낸다^[5,10,11]. 이러한 이유로 영상 스티칭 품질 평가에서 SSIM이 많이 활용되고 있으며, SSIM을 기반으로 한 IW-SSIM^[12], FSIM^[13], GSM^[14] 등과 같은 평가 방법들이 제안되었다.

그러나, SSIM에서 사용되는 휘도, 명암비, 구조는 각각 영상 윈도우의 평균, 분산, 공분산으로 구성된 상수로써 스티칭 영상의 픽셀 단위 정보를 활용하지 않는 문제점이 존재한다. 또한, MSSIM은 여러 SSIM의 평균이기 때문에 SSIM과 마찬가지로 픽셀 정보를 따로 사용할 수 없는 한계점이 존재한다. 이는 SSIM을 기반으로 한 IW-SSIM, FSIM, GSM 등 또한 마찬가지이다. 픽셀 단위 정보는 보다 정확한 스티칭 구현에 필수적인 데이터이기에, 본 논문에서는 픽셀 단위 정보와 기하적 유사도를 결합한 영상 스티칭 평가 방법을 III장에서 제시한다.

III. 지역 차분 픽셀 평가 방법 (Local Differential Pixel Assessment Method)

스티칭 영상 평가에서 PSNR은 픽셀 단위의 측정 방법이지만 블렌딩한 스티칭 영상이 갖고 있는 색 오차에 따라 스티칭 영상 평가에 적합하지 않으며, SSIM은 구조적 유사도를 기반으로 평가를 진행하기에 블렌딩한 스티칭 결과에 상대적으로 정확한 평가를 나타냈지만 픽셀 단위의 정보가

사라지는 문제로 인해 정확도에 한계를 나타냄을 II장에서 설명했다. 이에 III장에서는 픽셀 단위 정보를 활용할 수 있으며, 구조적 유사도를 측정하는 새로운 스티칭 영상 평가 방법인 지역 차분 픽셀 평가 방법 (LDPM: Local Differential Pixel Assessment Method)을 제시하고자 한다. 본 논문에서 제시하는 LDPM은 크게 측정 공간을 정의하는 부분과 스티칭에서 발생하는 왜곡을 측정하는 방법으로 구성되어 있다.

1. 측정 공간 정의

본 논문에서는 스티칭 평가의 비교 공간을 영상 간 공통 영역(Overlap region)으로 정의한다^[3]. 일례로, 그림 4에서 처럼 좌측 영상과 우측 영상을 스티칭하는 상황을 가정한다면, 좌측 영상과 우측 영상에서 공통으로 존재하는 영상 구역이 공통 영역이다. 비교 공간을 영상 전체가 아닌 공통 영역으로 설정하는 이유는 영상 스티칭에서 호모그래피 추정 및 블렌딩 작업이 주로 공통 영역을 기준으로 계산되고 적용되기 때문이다^[6,7,8]. 또한, 두 영상 간 중첩되지 않은 영역을 비교 공간에 추가하는 것은 측정 계산식의 분모의 크기를 늘려 오류의 민감도를 낮추게 된다.

2. 왜곡 측정 방법

$M \times N$ 크기의 그레이 스케일(8 bits) 원본 영상 f 와 스티칭된 영상 g 를 가정했을 때, 두 영상의 공통 영역의 픽셀

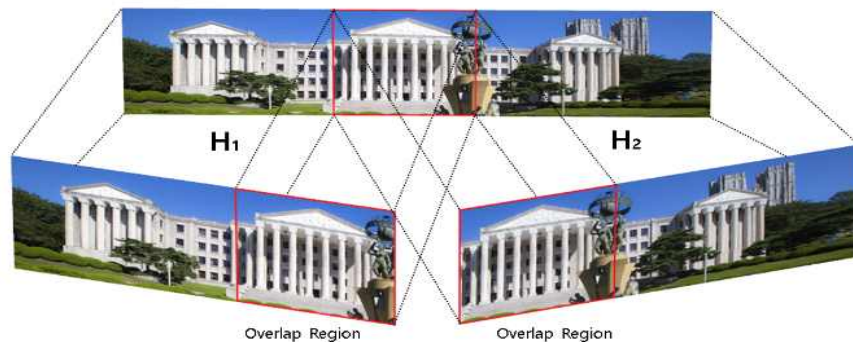


그림 4. Homography와 Overlap Region
 Fig. 4. Homography and Overlap Region

좌표 (x,y) 에 해당하는 픽셀 값을 $I_f(x,y)$, $I_g(x,y)$ 으로 각각 정의한다. 식 (5)처럼 영상의 픽셀 좌표 (x,y) 에 대해서 좌표 $(x-1,y-1)$ 부터 $(x+1,y+1)$ 까지에 해당하는 인접 픽셀을 묶어 3×3 크기의 LP(Local Pixel) 구조를 생성하며, 9개의 그레이 스케일 픽셀 값이 설정된다.

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

LP 구조의 각각의 외각 픽셀에서 중앙 픽셀 값을 빼 식 (6)의 LDP 구조를 생성한다.

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

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

LDP 구조는 그림 5에서와 같이 방향이 고정된 벡터로써, 영상 스티칭의 블렌딩 작업으로 달라지는 픽셀 값을 α 라 가정할 때, 식 (6)의 D_1 부터 D_8 의 값은 두 픽셀 간의 밝기 차이만 저장되므로 Blending 효과에 해당하는 α 값이 제거되는 것을 그림 6에서 확인할 수 있다. 그러므로 LDP를 통해 두 영상을 비교하는 방법은 블렌딩 효과로부터 영향이 적다.

LDPM은 식 (7)과 같이 전체 영상 f 와 g 에서 동일 좌표의 LDP를 비교한다. 공통 영역 전체에 걸쳐 계산한 평가 값을 공통 영역의 영상 사이즈로 나눈 것이 최종적인 LDPM 값으로 정의된다. 본 논문에서는 오류 지점을 찾아내는 것이 주된 목적이기 때문에 LDP를 비교하는 것은 기존 PSNR과 SSIM과 달리 결과 수치가 0에 가까울수록 좋은 품질의 스티칭 영상으로 나타낸다.

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

식 (8)의 $match$ 는 좌표 (i,j) 에서 스티칭 오류를 확인하

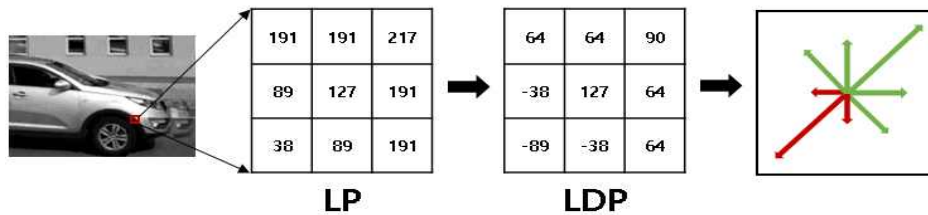


그림 5. LP와 LDP
Fig. 5. LP and LDP

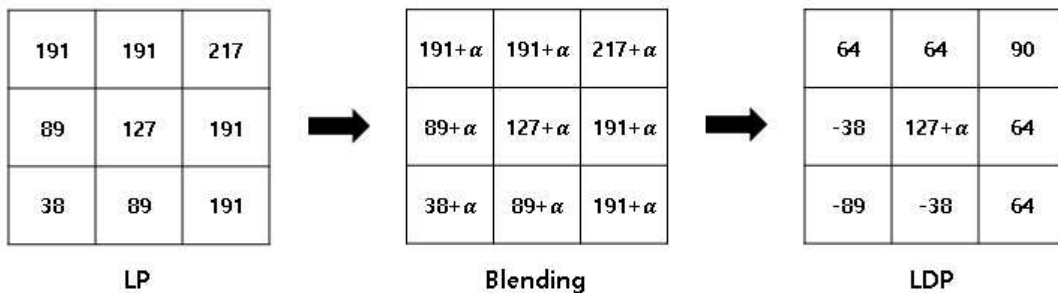


그림 6. Blending과 LDP
Fig. 6. Blending and LDP

는 함수로서, 해당 식(8)에서 사용되는 e 는 식 (9)에서 나타난 바와 같이 좌표 (i, j) 에서 cnt_k 합계이고, cnt_k 는 두 LDP 각각의 D_k 가 같은지 확인하여 일치할 시 '1'로 설정되고 다를 시에는 '0'으로 설정된다. D_1 부터 D_8 중에서 하나라도 값이 다르다면 e 가 8보다 작아 $match$ 를 '1'로 설정하고 모두 같으면 e 가 8이 되어 $match$ 를 '0'으로 설정한다. 식 (9) cnt_k 에서 두 LDP의 D_k 를 비교할 때 θ_{th} 의 임계값을 설정하는 이유는 영상의 색 공간 이동 및 호모그래피 변환 과정에서 달라질 수 있는 픽셀 값을 보상하기 위함이다. 본 논문의 실험에서는 평행하게 설치된 카메라를 통하여 촬영한 흑백 영상을 활용하여, 호모그래피 변환 과정과 색 공간 이동에 따른 픽셀 정보의 손실이 없으므로 임계값을 0으로 설정하였다. 또한 컬러 영상 실험의 경우 색 공간의 이동에 따른 픽셀 정보의 손실이 예상되므로 임계값을 2로 설정하였다.

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

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

- LDPM(f, g): 영상 f, g 구조적 차이
- e : 에러 픽셀 개수, cnt: LDP 매칭 횟수

- D_{fk} : 영상 f의 LDP $D_1 \sim D_8$
- D_{gk} : 영상 g의 LDP $D_1 \sim D_8$
- θ_{th} : 에러 임계값

IV. 실험 결과

본 논문에서 제안하는 LDPM은 LDP 구조의 중앙값에 픽셀 정보를 저장하며, LDP의 가장자리 픽셀 값과 중앙 픽셀 값의 차이 값을 통해 벡터 구조를 형성하므로 구조적 유사도 평가 방법이다. 본 IV장에서는 그림 7의 좌측 영상 (a), (c)와 우측 영상 (b), (d) 그리고 스티칭 영상 (e), (f)를 통해 LDPM을 검증하고자 한다. 좌측 영상과 우측 영상은 평행하게 설치된 카메라를 통해 촬영되었고, 스티칭 영상의 설계 시스템은 Visual Studio 2015에서 C++ 기반으로 Opencv의 AutoStitch를 통해 생성했다.

그림 7의 (a), (b), (e)의 테스트는 동일한 조명 환경을 갖는 좌측 영상과 우측 영상을 통해 인공물이 생성되지 않은 스티칭 영상을 가정하였다. 표 1에서 나타난 바와 같이 스티칭 영상 (e)에 대해서 MSSIM은 평가 지수 '1'을 기록하였고, LDPM 평가 방법은 '0'을 기록하여 두 평가 방법 모두 정확한 평가 결과를 나타냈다.

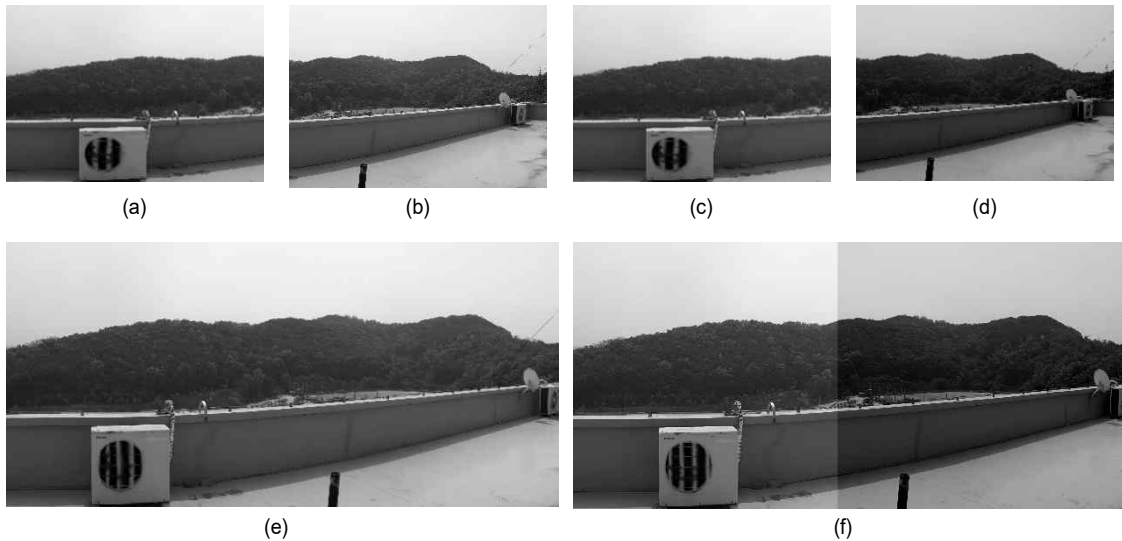


그림 7. 왼쪽 영상 (a) (c), 오른쪽 영상 (b) (d), 스티칭 결과 영상 (e) (f)
 Fig. 7. Left Image (a) (c), Right Image (b) (d), Stitching Image (e) (f)

그림 7의 (c), (d), (e)의 테스트는 서로 다른 조명 환경을 갖는 좌측 영상과 우측 영상을 스티칭하는 과정에서 스티칭 영상에 솔기가 없도록 블렌딩한 것을 가정하였다. 블렌딩을 통해 보정하였기 때문에 표 1에서 보는 바와 같이 스티칭 영상 (e)에 대해서 MSSIM은 0.9402로 나타났고, LDPM은 0.9315%로 나타났다.

마지막으로 그림 7의 (c), (d), (f)의 테스트는 서로 다른 조명 환경을 갖는 좌측 영상과 우측 영상을 이용하여 솔기가 있는 스티칭 영상을 생성한 것을 가정하였으며, 스티칭 영상 (f)에 대해서는 MSSIM은 0.9335로 나타났고, LDPM은 1.7605%로 나타났다. MSSIM과 LDPM 모두 블렌딩을 적용하지 않아 솔기가 존재하는 스티칭 영상을 가장 낮은 품질의 영상이라고 평가를 내렸다.

상기의 결과에서 본 바와 같이 MSSIM과 LDPM 모두 사람의 시각 인지와 같은 평가 결과를 나타낸다. 하지만 MSSIM의 평가 값을 살펴보면 (a), (b), (e)의 테스트에서 1, (c), (d), (e)의 테스트에서 0.9402, (c), (d), (f)의 테스트에서 0.9335로 기록되었다. 블렌딩을 적용한 스티칭 영상의 MSSIM이 1보다 낮은 것은 직관적 판단과 동일하지만, 블렌딩을 적용한 스티칭 영상과 블렌딩을 적용하지 않은 영상의 MSSIM 차이가 0.0067인 것은 스티칭 영상을 평가하기에는 오류 민감도가 낮은 것으로 판단된다.

반면 LDPM은 (a), (b), (e)의 테스트는 0%, (c), (d), (e)의

테스트는 0.9315%, (c), (d), (f)의 테스트는 1.7605%로 기록되었다. 블렌딩을 적용한 스티칭 영상과 블렌딩을 적용하지 않은 영상의 LDPM 차이가 약 두 배로 MSSIM보다는 오류 민감도가 크게 나타났으며, 또한, 동일한 평가 공간에서 평가를 진행했음에도 LDPM이 MSSIM보다 오류 민감도가 더 큰 것을 보였다.

표 1. 흑백 영상 스티칭 평가

Table 1. Gray Color Image Stitching Assessment

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

그림 8의 (a), (b), (e)의 테스트는 동일한 조명 환경을 갖는 좌측 영상과 우측 영상을 통해 인공물이 생성되지 않는 스티칭 영상을 가정하였다. 흑백 영상 기반의 그림 7 (a), (b), (e) 테스트와 동일하게 MSSIM은 평가 지수 ‘1’을 기록하였고, LDPM 평가 방법은 ‘0’을 기록하여 컬러 영상에서도 정확한 평가 결과를 나타냈다.

그림 8의 (c), (d), (e) 테스트는 서로 다른 조명 환경을 갖는 좌측 영상과 우측 영상을 스티칭하는 과정에서 스티칭 영상에 솔기가 없도록 블렌딩한 것을 가정하였다. 블렌딩을 통해 보정하였기 때문에 표 2에서 보는 바와 같이 스티칭

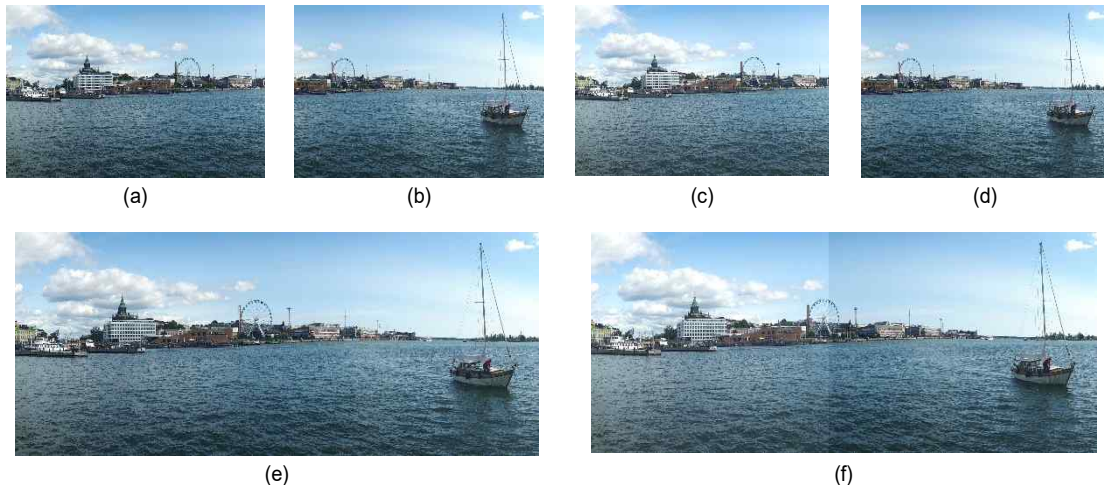


그림 8. 왼쪽 영상 (a) (c), 오른쪽 영상 (b) (d), 스티칭 결과 영상 (e) (f)
 Fig. 8. Left Image (a) (c), Right Image (b) (d), Stitching Image (e) (f)

티칭 영상 (e)에 대해서 MSSIM은 0.9942로 나타났고, LDPM은 0.1757%로 나타났다.

마지막으로 그림 8의 (c), (d), (f)의 테스트는 서로 다른 조명 환경을 갖는 좌측 영상과 우측 영상을 이용하여 솔기가 있는 스티칭 영상을 생성한 것을 가정하였으며, 스티칭 영상 (f)에 대해서는 MSSIM은 0.9935로 나타났고, LDPM은 0.3183%로 나타났다. 컬러 영상의 경우에도 MSSIM과 LDPM 모두 블렌딩을 적용하지 않아 솔기가 존재하는 스티칭 영상을 가장 낮은 품질의 영상이라고 평가를 내렸다.

앞선 그림 7의 흑백 영상 스티칭 평가에서는 좌측 영상과 우측 영상의 공통 영역 크기가 501×1760 인 반면, 컬러 영상 스티칭 평가에서는 공통 영역은 크기가 1400×1760 이다. LDPM의 계산식에서 공통 영역의 크기는 분모에 위치하기 때문에 컬러 영상의 평가에서 LDPM 측정값이 작아졌지만, SSIM의 평가보다 LDPM의 오류 민감도가 더 큰 것을 보였다.

표 2. 컬러 영상 스티칭 평가
 Table 2. Color Image Stitching Assessment

Test(Overlap)	MSSIM	Proposed LDPM (Error)
(a), (b) and (e)	1.0000	0 %
(c), (d) and (e)	0.9942	0.1757 %
(c), (d) and (f)	0.9935	0.3183 %

이와 같은 실험을 통해 본 논문에서 제안하는 LDPM 방법은 SSIM에서 제공하는 구조적 유사도 기반의 평가 측정을 제공하면서, 픽셀 기반의 정보를 활용함으로써, 스티칭 영상에 대한 품질에 대한 보다 정확한 평가 결과를 제공할 수 있음을 확인할 수 있다.

V. 결론

영상 스티칭 평가에서 MSSIM 평가 지표는 사람의 시각 인지와 비슷한 결론을 나타낸다. 하지만 MSSIM 값은 평균 정보만을 담고 있기 때문에 픽셀 단위 정보는 사라지게 되는 문제가 있다. 최근 딥러닝을 통한 영상 스티칭의 시도가 늘어나고 있음에 따라, 딥러닝 시스템의 스스로 판단하는 과정에서 스티칭 영상의 결과에 대한 보다 객관적인 평가

및 판단 방법의 개발이 요구된다. 본 논문에서 제시하는 지역 차분 픽셀 평가 방법 (LDPM)은 오류 픽셀의 정보와 구조적 유사도 정보를 활용하기에, 픽셀 단위 평가를 진행하면서 오류로 판단되는 픽셀을 버퍼에 저장하고, 오류 픽셀들의 좌표들을 묶어 라벨링(labeling)을 통해 패턴을 규정함으로써 오류의 종류까지 분류가 가능해 딥러닝 기반의 영상 스티칭에서 활용할 것으로 기대된다. 본 논문에서는 기하적 평가를 가능하게 하는 최소 단위 크기 LP, LDP 블록을 통해 영상 스티칭 평가를 진행하지만, 추가로 LDPM의 LP, LDP의 블록 크기를 늘리거나 LDP 블록의 외각 성분을 또 다른 LDP로 묶어 활용하는 방법 등의 추가적인 연구를 진행함으로써 보다 만족스러운 평가 결과와 딥러닝 기반 스티칭 결과를 제공할 것으로 기대된다.

참고 문헌 (References)

- [1] 8K TV War which burned in the second half ... Korea, China, Japan, "Three Kingdoms", <http://news.heraldcorp.com/view.php?ud=20190621000013> (accessed June, 21, 2019).
- [2] Samsung Electronics Expands QLED 8K TV Market, http://www.naeil.com/news_view/?id_art=315419 (accessed June, 07, 2019).
- [3] Jeonho Kang, Junsik Kim, SangIL Kim, and Kyuheon Kim, "Method of Video Stitching based on Minimal Error Seam", The Korean Institute of Broadcast and Media Engineers, Vol.24, No.1, pp.142-152, January, 2019.
- [4] 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.
- [5] 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.
- [6] R. Szeliski, "Image Alignment and Stitching: A Tutorial." Foundations and Trends in Computer Graphics and Computer Vision, Vol. 2, No.1, 2006.
- [7] 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.
- [8] 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.
- [9] Meer Sadeq Billah, Heejune Ahn, "Objective Quality Assessment Method for Stitched Image", The Korean Institute of Broadcast and Media Engineers, Vol.23, No.2, pp.227-234, March, 2018.
- [10] Xu, Wei, and Jane Mulligan. "Performance evaluation of color correction approaches for automatic multi-view image and video stitching."

2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.263-270, 2010.

- [11] Qureshi, H. S., et al. "Quantitative quality assessment of stitched panoramic images." IET image processing, Vol.6, no.9, pp.1348-1358, 2012.
- [12] WANG, Zhou; LI, Qiang. Information content weighting for perceptual image quality assessment. IEEE Transactions on Image Processing, Vol.20, No.5, pp.1185-1198, 2010.

ssing, Vol.20, No.5, pp.1185-1198, 2010.

- [13] Zhang, Lin, et al. "FSIM: A feature similarity index for image quality assessment." IEEE transactions on Image Processing, Vol.20, No.8, pp.2378-2389, 2011.
- [14] Liu, Anmin, Weisi Lin, and Manish Narwaria. "Image quality assessment based on gradient similarity." IEEE Transactions on Image Processing, Vol.21, No.4, pp.1500-1512, 2011.

저 자 소 개



이 성 배

- 2019년 2월 : 경희대학교 전자공학과 공학사
- 2019년 3월 ~ 현재 : 경희대학교 전자공학과 석사과정
- ORCID : <https://orcid.org/0000-0002-7392-6870>
- 주관심분야 : 딥러닝, 영상처리, 멀티미디어 시스템



강 전 호

- 2010년 2월 : 경희대학교 전자공학과 공학사
- 2012년 2월 : 경희대학교 전자공학과 공학석사
- 2012년 3월 ~ 2015년 2월 : 주식회사 넥스트리밍 연구원
- 2015년 3월 ~ 현재 : 경희대학교 전자공학과 박사과정
- ORCID : <https://orcid.org/0000-0003-4405-3879>
- 주관심분야 : 영상처리, 멀티미디어 시스템, 디지털 대화형방송



김 규 현

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- ORCID : <https://orcid.org/0000-0003-1553-936X>
- 주관심분야 : 디지털 방송, 영상처리, 멀티미디어 통신, 디지털 대화형 방송