

일반논문 (Regular Paper)

방송공학회논문지 제25권 제6호, 2020년 11월 (JBE Vol. 25, No. 6, November 2020)

<https://doi.org/10.5909/JBE.2020.25.6.1004>

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

실시간 고해상도 동영상 스티칭을 위한 고속 콘텐츠 보존 시접선 추정 방법

김 태 하^{a)}, 양 성 엽^{a)}, 강 병 근^{b)}, 이 희 경^{c)}, 서 정 일^{c)}, 이 의 진^{a)*}

Fast Content-preserving Seam Estimation for Real-time High-resolution Video Stitching

Taeha Kim^{a)}, Seongyeop Yang^{a)}, Byeongkeun Kang^{b)}, Hee Kyung Lee^{c)}, Jeongil Seo^{c)}, and Yeejin Lee^{a)*}

요 약

본 논문은 실시간 고해상도 비디오 스티칭을 위한 새로운 콘텐츠 보존 시접선 추정 알고리즘을 제안한다. 시접선 추정은 영상 스티칭 후 중첩 영역에서의 시각적 왜곡을 최소화하기 위한 요소 기술 중 하나이다. 기존 시접선 추정 알고리즘들은 요구되는 연산량과 메모리 사용량이 높은 최적화 알고리즘에 기반을 두고 있음에도 불구하고, 추정된 시접선이 객체를 피하지 못해 객체를 자르거나 반복하는 현상을 유발한다. 또한, 프레임 간의 추정된 시접선의 시간적 일관성이 부족하여 불필요한 잦은 변동이 발생한다. 따라서, 본 논문에서는 직선의 시접선을 활용하여 효율적이고 시간적 일관성이 있으며, 심층신경망 기반 객체 세그멘테이션 알고리즘을 활용하여 객체를 피하여 시접선을 형성하는 시접선 추정 알고리즘을 제안하고자 한다. 고해상도 360° 다중 시점 동영상을 사용한 실험을 통해 제안하는 알고리즘이 기존 알고리즘보다 짧은 시간에 시각적으로 유사한 360VR 동영상을 생성하는 시접선을 추정함을 확인하였다.

Abstract

We present a novel content-preserving seam estimation algorithm for real-time high-resolution video stitching. Seam estimation is one of the fundamental steps in image/video stitching. It is to minimize visual artifacts in the transition areas between images. Typical seam estimation algorithms are based on optimization methods that demand intensive computations and large memory. The algorithms, however, often fail to avoid objects and results in cropped or duplicated objects. They also lack temporal consistency and induce flickering between frames. Hence, we propose an efficient and temporally-consistent seam estimation algorithm that utilizes a straight line. The proposed method also uses convolutional neural network-based instance segmentation to locate seam at out-of-objects. Experimental results demonstrate that the proposed method produces visually plausible stitched videos with minimal visual artifacts in real-time.

Keyword : Image stitching, video stitching, seam estimation, instance segmentation

I. 서론

최근 가상현실(virtual reality), 증강현실(augmented reality)과 같은 실감 미디어 서비스에 대한 관심이 증가하고 있으며, 이에 따라 보다 나은 몰입감과 현장감을 제공하기 위한 넓은 시야각(field of view, FoV)을 가진 고해상도, 고화질의 콘텐츠 생성 수요도 증가하고 있다. 그러나 이러한 고해상도의 콘텐츠 생성을 위해서는 고가의 전문가용 장비가 필요하여 영상 콘텐츠 수급에 어려움을 겪고 있다. 따라서, 비교적 저렴한 다수의 고화질 영상을 활용하여 고해상도의 초고화질 영상을 생성하는 방식으로 콘텐츠 부족 현상을 극복하고자 하는 노력이 진행되고 있다.

다수의 카메라를 이용하여 넓은 시야각을 가진 고해상도의 영상을 획득하는 방법은 일정한 시야각을 가진 카메라들을 이용하여 중첩된 영역을 가진 여러 장의 영상을 획득한 후, 서로 중첩된 영역을 가진 영상들의 위치 관계를 파악하여 정렬(registration)하고, 정렬된 영상의 중첩된 영역을 하나로 결정하고 해당 영역의 시각적 질을 향상시키는 과정을 통해서 이루어진다. 이와 같은 스티칭 기법 및 후처리에 관한 연구는 현재까지 주로 정지 영상에 대하여 이루어져 왔으나, 최근에는 기술의 발달과 사용자의 초고해상도 콘텐츠에 대한 요구 증가로 동영상 기반의 스티칭 기술에 대한 수요가 증가하고 있다. 그러나 기존의 정지 영상을 위한 스티칭 기술은 연산량이 많고 처리 시간이 길어 실시간 360° 동영상 제작 등과 같이 실시간으로 고해상도의 영상 처리가 필요한 경우에는 적용이 불가능하다는 한계점이 존재한다. 또한, 연속적으로 촬영된 동영상의 경우, 프레임 간의 조도

차이, 카메라 흔들림, 움직이는 객체 등에도 불구하고 프레임 간 시각적 일관성을 유지해야 하는 어려움이 있다.

따라서, 본 논문에서는 프레임 간의 일관성을 유지하면서도 시각적으로 만족스러운 360VR 동영상을 생성하기 위하여, 영상 정렬 후 스티칭 된 영역의 왜곡을 최소화하기 위한 효율적인 시접선 추정 알고리즘을 제안하고자 한다.

II. 관련 연구

중첩된 영역을 가진 영상의 스티칭은 일반적으로 시점이 다른 영상마다 특징점에서의 특징 기술자(feature descriptor)를 추출하고, 추출된 특징 기술자 간의 대응 쌍(correspondences)을 생성한 후 카메라 간의 자세 정보를 예측하여, 다른 시점 카메라의 위치 좌표를 기준 카메라의 위치 좌표에 투영하는 호모그래피 추정(homography estimation) 방식으로 수행된다^{[1][2]}. 이러한 특징점 기반의 영상 스티칭 성능은 호모그래피의 정확도에 의해 크게 영향을 받음으로 보다 정확한 호모그래피를 추정하기 위한 기법들이 제안되었다^{[3][4][5]}. 예를 들어, 추정된 하나의 호모그래피를 영상 내 작은 영역에 가중치가 적용된 특징점에 부합하도록 변경하는 방식이 제안되었다^[4]. 또한, 여러 개의 호모그래피를 예측한 뒤 영역별로 가장 적합한 호모그래피를 선택하여 영상을 스티칭하는 방식도 제안되었다^[5].

이와 같이 보다 정교한 방식으로 호모그래피를 추정하는 방식들이 제안되었음에도 불구하고, 시차(parallax), 거리(depth) 차, 카메라 간 조도 차에 의해 장면 내 서로 다른 거리에 위치한 물체들의 정확한 호모그래피를 추정하는 것은 매우 어려우며, 이에 따라 스티칭 된 영상의 중첩된 영역에서는 합성에 의한 왜곡이 자주 발생한다. 이와 같은 문제점을 해결하고자 영상 정렬(registration) 후 스티칭 된 영역의 왜곡을 최소화하기 위한 시접선(seam)을 활용한 합성 방식^{[6][7]}과 영상 블렌딩 방법들도 연구되어 왔다^{[1][2][8][9]}. 이중 그래프 컷(graph-cut) 알고리즘은 화소들의 유사도를 최대화하는 방식으로 시접선을 추정하며, 이후 제안된 많은 시접선을 추정하는 알고리즘의 바탕이 되었다^[6]. Zhang 등은 정렬 오류와 영상 간의 색깔 차이로 구성된 에너지 함수

a) 서울과학기술대학교 전기정보공학부(Department of Electrical and Information Engineering)

b) 서울과학기술대학교 전자IT미디어공학과(Department of Electronic and IT Media Engineering)

c) 한국전자통신연구원 실감미디어연구실(Immersive Media Research Laboratory)

‡ Corresponding Author : 이의진(Yeejin Lee)

E-mail: yeejinlee@seoultech.ac.kr

Tel: +82-2-970-6443

ORCID: <https://orcid.org/0000-0002-3439-5042>

* This work was supported by Electronics and Telecommunications Research Institute(ETRI) grant funded by ICT R&D program of MSIT/IITP[2018-0-00207, Immersive Media Research Laboratory].

• Manuscript received October 15, 2020; Revised November 16, 2020; Accepted November 16, 2020.

를 그래프 컷 알고리즘을 활용하여 최적화한 시점선 추정 기법을 제안하였다^[7].

연속적인 영상으로 구성된 동영상 스티칭의 경우 장면의 움직임, 객체 출현 등에 의해 영역별 시차 차이가 크며, 이에 따라 객체가 겹쳐 보이거나, 일그러지거나, 잘려 나가는 등 스티칭에 의한 시각적 결함 문제가 가중된다. 그러나 현재까지 동영상 스티칭에 관한 연구는 활발히 이루어지지 않고 있으며, 영상 스티칭 방법을 동영상 스티칭에 적용하는 방식이 사용되거나^[10], 그래프 컷 최적화 방식을 공간 영역과 시간 영역에 동시에 적용하는 방식이 제안되었다^[11].

III. 전경 객체 보존을 위한 시점선 추정 알고리즘

II절에서 기술된 바와 같이 영상 및 동영상 스티칭을 위한 다양한 방법이 제시되었음에도 불구하고, 그림 1에서 보는 바와 같이 중첩 영역에 있는 전경 객체가 잘려 나가거나 겹쳐 보이는 현상은 여전히 완화하기 어려운 문제이다. 또한, 영상 정렬 후, 영역별 왜곡 보안을 위하여 보다 정교한 알고리즘을 사용하면 긴 처리 시간이 요구되어 실시간 고해상도 동영상 처리에 적용하기는 어렵다. 따라서, 본 절에서는 위에 기술된 문제점들을 해결하고자 전경 객체를 보존하면서도 고해상도 동영상에서도 빠른 속도로 적용 가능한 효율적인 시점선 추정 알고리즘을 설계한다. 제안하는 알고리즘은 전경 객체의 정확한 위치 및 영역 정보를 활용하여 시점선을 추정하고, 추정된 시점선을 따라 중첩 영역의 화소 정보를 선택하여 스티칭된 영역에서의 객체의 왜곡을 최소화한다.

1. 시점선 추정을 위한 에너지 함수 설계

중첩 영역에서 시점선이 주요 객체를 벗어나는 영역으로 형성되도록 시점선을 추정함으로써 주요 객체를 보존할 수 있다. 따라서, 시점선을 객체 보존을 위한 스티칭 기법에 활용할 수 있으며, 중첩 영역에서 시점선을 추정하여 영상을 이어 붙인다는 것은 서로 다른 시점에서 촬영된 영상의 화소 중 어떤 시점의 화소를 중첩 영역에 보존할지를 결정하는 문제로 정의할 수 있다.

시점선 추정을 위해서는 중첩 영역 내에 중요도가 낮은 화소를 결정하기 위한 기준이 필요하며, 이는 인접 화소 간의 기울기(gradient) 차이 등으로 표현될 수 있다^{[8],[9],[10]}. 그러나 화소 간의 기울기 차이만으로는 전경 객체와 배경 분리 한계가 있으므로, 제안하는 시점선 추정 알고리즘에서는 전경 객체만을 완벽히 분리해내는 영상 분할(segmentation) 결과를 에너지 함수로 사용한다. 이에 따라, 서로 다른 시점(view point)에서 촬영된 두 동영상의 I_A , I_B 의 t 번째 프레임의 높이가 M , 넓이가 N 인 중첩 영역 내의 화소 위치 (y, x) 에서의 영상 분할 결과 $O^t(y, x)$ 는 다음과 같이 표현할 수 있다.

$$O^t(y, x) = O_A^t(y, x) \parallel O_B^t(y, x) \quad (1)$$

위의 식에서 $O_A^t(y, x)$ 와 $O_B^t(y, x)$ 는 동영상 A, B의 t 번째 프레임의 중첩 영역에서의 영상 분할 결과이며, 객체에 속하는 화소의 값은 1, 속하지 않는 화소의 값은 0인 이진 영상으로 표현된다. 연산 \parallel 은 OR 연산으로 동영상 A, B의 t 번째 프레임에서 객체에 속하는 모든 픽셀의 합집합을 의미한다.

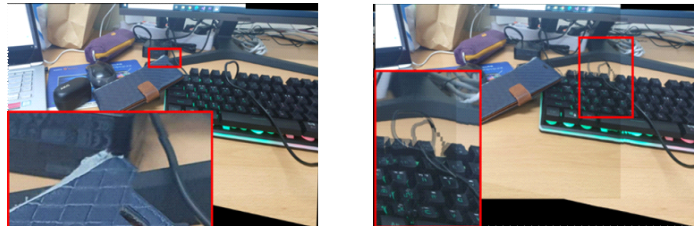


그림 1. 스티칭 오류 예시: 객체가 어긋나는 현상(좌)과 객체가 겹쳐 보이는 현상(우)

Fig. 1. Examples of visual artifacts on objects. object deformation (left) and duplication (right)

그러나 개별 프레임 기반의 영상 분할 결과 O^t 를 동영상에 활용할 경우 객체의 움직임에 따라 시접선의 잦은 변화가 중첩 영역에서 나타날 수 있으므로 시간 영역(temporal domain)에서 영상 분할 결과를 누적 활용한 에너지 E^t 를 시접선 추정을 위한 t 번째 프레임에서의 최종 에너지로 사용하고 아래와 같이 정의한다.

$$E^t(y, x) = \sum_{i=0}^k w_i O^{t-i}(y, x) \quad (2)$$

위의 식에서 O^{t-i} 는 $(t-i)$ 번째 프레임의 영상 분할 결과이며, w_i 는 k 개의 프레임 결과를 누적할 때, 각 프레임의 결과를 E^t 에 어느 정도 반영할지 결정하는 가중치이다. 즉, 식 (2)의 에너지 맵은 이전 프레임들의 가중합(weighted sum)이며, 이는 프레임 간 영상 분할 결과의 갑작스러운 변화를 완화하여 스티칭된 동영상의 연속성 향상에 도움이 된다. 또한, 그림 2에서도 볼 수 있듯이, 공간 영역(spatial domain)에서도 근접한 프레임의 영상 분할 정보를 결합하여 개별 프레임에 존재할 수 있는 불완전한 영상 분할 정보 보완도 가능하게 한다.

2. 시접선 추정 알고리즘

개별 프레임별로 시접선을 추정할 경우, 장면의 변화가 없는 경우라도 조도 차이, 카메라의 흔들림 등에 의해 프레임마다 다른 위치에서 시접선이 추정될 수 있다. 이와 같은 시접선 위치 변화는 스티칭된 동영상에서 장면 전환이 크

지 않음에도 불구하고, 장면이 전환되는 것과 같은 깜빡거림(flickering) 현상을 유발할 수 있다. 기존의 에너지가 큰 객체의 경계 영역을 따라 시접선을 추정하는 알고리즘들은 화소 값의 작은 변화에도 민감하게 반응하여 객체 영역을 침범하거나 좁은 영역 내에서 시접선의 위치가 자주 변하는 지그재그(zigzag)와 같은 패턴을 가진 시접선을 추정할 가능성이 높다. 또한, 이러한 알고리즘들은 최적화 문제를 풀기 위하여 많은 연산량과 긴 수행 시간이 필요하므로 실시간으로 고해상도 동영상을 처리해야 하는 시스템에 사용하기에는 적합하지 않다.

따라서, 위와 같은 문제점을 해결하기 위해서는 프레임별 시접선의 불필요한 변화를 최소화하고, 고해상도 동영상에 실시간 적용이 가능하면서도 주요한 전경 객체 영역을 피해 시접선을 추정하는 방식이 필요하다. 일반적으로 시접선은 중첩 영역의 최상단 행으로부터 최하단 행까지 에너지 맵을 검사하여, 행 방향으로 누적 에너지를 계산하고 계산된 누적 에너지의 최하단에서부터 다시 최솟값을 검색하는 방식으로 결정된다. 그러나 모든 행에서 최솟값을 검색하는 방식은 연산 시간이 오래 걸리므로, (2)와 같이 객체만의 정보를 가진 에너지를 활용하여 최단 경로 탐색이 필요 없는 시접선 추정 방식을 사용한다. 이와 같이 객체 정보만을 활용하는 경우에는 추정된 시접선이 주요 객체가 아닌 다른 요소에 영향을 받지 않으므로, 최단 경로 탐색이 필요 없으며 대상 객체에서 벗어난 가장 가까운 경계 영역을 결정하여 수직인 시접선을 추정하여 시접선 패턴 변화에 의한 깜빡거림 현상을 시각적으로 완화한다.

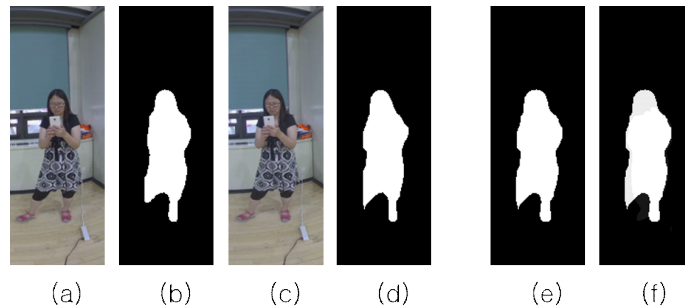


그림 2. 객체 에너지 결과 예시 (a) 입력 영상 I_A^t (b) I_A^t 의 객체 분할 결과 O_A^t (c) 입력 영상 I_B^t (d) I_B^t 의 객체 분할 결과 O_B^t (e) O_A^t 와 O_B^t 를 활용하여 연산한 O^t (f) 시간 영역의 객체 분할 결과를 누적한 객체 에너지 E^t

Fig. 2. Examples of energy maps. (a) I_A^t (b) O_A^t (c) I_B^t (d) O_B^t (e) O^t (f) E^t

대상 객체에서 벗어난 가장 가까운 경계 영역을 결정하기 위해서는 중첩 영역을 구성하는 열 중 객체에 속하지 않는 열을 찾아야 하며, 객체에 속하는지 속하지 않는 열인지는 아래 정의된 $E^t(c)$ 에 의해 결정된다.

$$E^t(c) = \sum_{y=0}^{M-1} E^t(y, c) \quad (3)$$

만약 $E^t(c)$ 이 0이라면 c 번째 열을 구성하는 모든 화소가 객체에 속하지 않음을 의미하며, 해당 열에 시접선을 위치시킬 경우 주요 객체를 피해 중첩 영역이 스티칭 될 수 있음을 알 수 있다. 그러나, $E^t(c)$ 이 0이 아니라면 c 번째 열의 화소 중 일부는 객체에 속함을 의미하며, 해당 열을 시접선 위치로 선택할 경우 객체를 관통하는 시접선이 형성될 가능성이 있다.

위의 (3)에 의하면, 중첩 영역 내의 주요 객체의 크기에 따라 객체에 속하지 않는 열이 다수 존재하거나, 모든 열이 객체에 속하는 경우가 발생할 수 있으므로 시접선 추정 규칙을 추가하여 시접선의 위치를 결정한다. 시접선 위치 결정 방식은 $E^t(c)$ 를 이용하여 첫 번째 열부터 마지막 열까지 차례대로 객체에 속하는 열인지를 검사하며, 시각적으로 연속적인 스티칭 동영상 생성을 위하여 이전 프레임의 시접선 위치를 고려한 현재 프레임의 시접선 위치를 결정한다. 만약 객체에 포함되지 않는 열이 다수 존재한다면, 현재 프레임에서의 시접선 위치 c^t 는 $E^t(c)$ 가 0인 n 개의 열

$\{c_1^t, c_2^t, \dots, c_n^t\}$ 에 대하여 이전 프레임의 시접선 위치 c^{t-1} 에서 가장 가까운 위치로 수식 (4)와 같이 결정된다.

이때, 이전 시접선의 위치로부터 가장 가까운 위치에 해당하는 열이 이전 시접선으로부터 왼쪽(c_l^t), 오른쪽(c_r^t)에 모두 존재한다면($|c_l^t - c^{t-1}| = |c_r^t - c^{t-1}|$), 주요 객체의 움직임 경향을 저장하고 있는 m^{t-1} 를 활용하여 움직임 방향을 추정한다. 움직임 방향을 예측하도록 도와주는 움직임 지시자(motion indicator)는 다음과 같이 정의한다.

$$m^{t-1} = \begin{cases} -1, & \text{if } c^{t-1} \leq c^{t-2} \\ 1, & \text{otherwise} \end{cases} \quad (5)$$

위의 $m^{t-1} = -1$ 인 경우는 객체가 왼쪽으로 이동하고 있음을 의미하고, 따라서, 현재 프레임의 시접선 위치 c^t 는 c_l^t 로 추정되며, 1인 경우에는 객체가 오른쪽으로 이동하고 있음을 의미하므로 c^t 를 c_r^t 로 추정한다.

$$c^t = \begin{cases} c_l^t & \text{if } m^{t-1} = -1 \\ c_r^t, & \text{otherwise} \end{cases} \quad (6)$$

또한, 모든 열이 객체에 속한다면 현재 프레임에서의 시접선 위치 c^t 는 에너지가 최소인 위치로 (7)과 같이 결정된다.

$$c^t = \min_{C_i} \{E^t(c_1), E^t(c_2), \dots, E^t(c_N)\}, \quad i = 1, 2, \dots, N \quad (7)$$

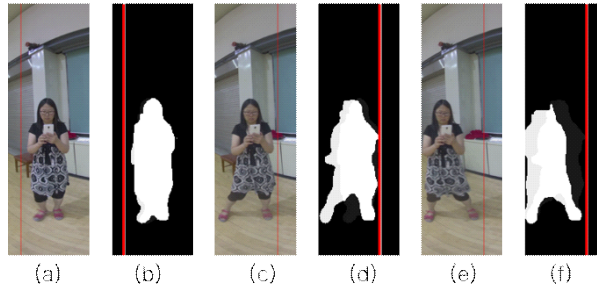


그림 3. 시접선 변화 예시 (a) 입력 영상 I_A^{t-1} (b) 에너지 맵 E^{t-1} (c) 입력 영상 I_A^t (d) 에너지 맵 E^t (e) 입력 영상 I_A^{t+1} (f) 에너지 맵 E^{t+1}
Fig. 3. Examples of seam changes. (a) I_A^{t-1} (b) E^{t-1} (c) I_A^t (d) E^t (e) I_A^{t+1} (f) E^{t+1}

$$c^t = \min_{C_i} \{|c_1^t - c^{t-1}|, |c_2^t - c^{t-1}|, \dots, |c_n^t - c^{t-1}|\}, i = 1, 2, \dots, n \quad (4)$$

그림 3은 연속적인 프레임에서의 시접선 변화 예시이며, 예시에서도 볼 수 있듯이 연속적인 객체의 이동으로 인하여, 이동 방향 쪽 중첩 영역에서 시접선이 존재할 열이 없을 때 반대 방향의 중첩 영역으로 시접선이 이동하였다. 이후 예측된 위치에서 객체가 감지되지 않으면 시접선이 고정되어 불필요한 시접선 변경을 방지하는 것을 확인할 수 있다.

IV. 실험 결과

본 장에서는 제안한 동영상 스티칭 알고리즘의 성능 검증 위한 실험 결과를 논의한다. 실험을 위한 영상은 6개

의 카메라로 구성된 360VR 카메라인 Insta360 Pro 2와 KanDao-Obsidian R로 획득하였다. 알고리즘을 통해 생성된 동영상의 해상도는 7680×3840 이며, 초기 영상은 호모그래피 방식을 사용하여 정렬하였다. 그림 4는 실험에서 사용한 정렬된 입력 영상의 예시로 6개의 영상을 스티칭하여 하나의 영상을 생성한다.

제안하는 알고리즘에서는 너무 많은 수의 프레임 정보를 누적할 경우, 객체의 움직임을 정확하게 예측하기 어렵고, 객체가 없는 열이 줄어들거나 없어지므로 두 프레임의 에너지를 누적하여 사용하였으며, (2)에서 정의된 최종 에너지 E^t 는 $k = 1, w_0 = 0.9, w_1 = 0.09$ 를 사용하여 계산하였다. 객체 정보 추출을 위한 영상 분할에는 다양한 알고리

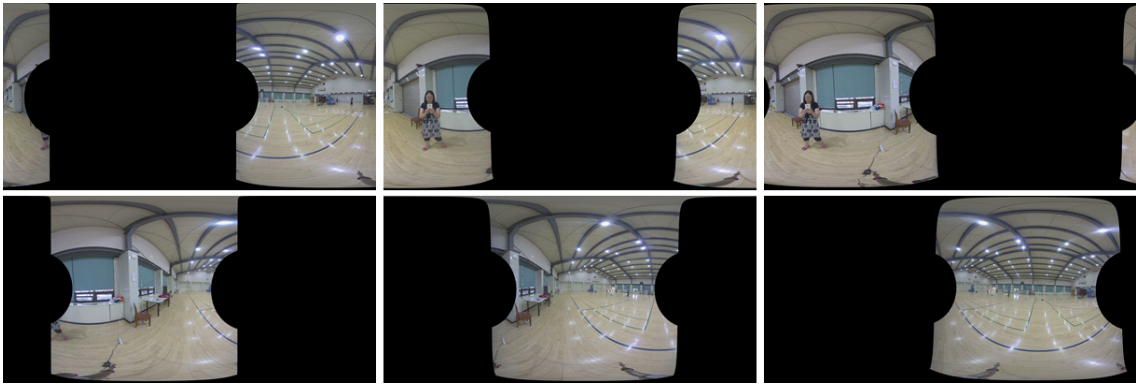


그림 4. KanDao-Obsidian R 카메라로 획득한 영상을 등장방형도법으로 투영한 영상 예시

Fig. 4. Examples of equirectangular projection of images collected using KanDao-Obsidian R

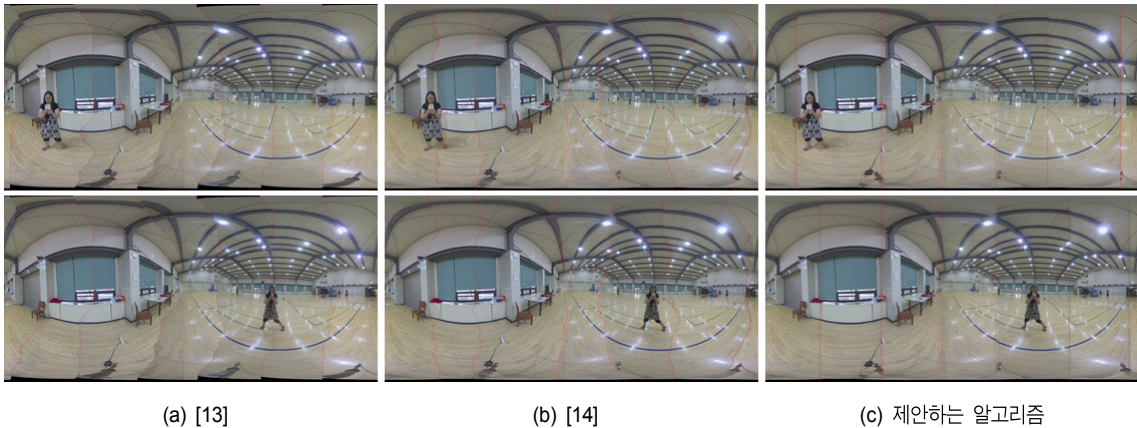


그림 5. 스티칭 결과 예시. (a) [13]의 알고리즘이 생성한 시접선 및 스티칭된 영상 (b) [14]의 알고리즘 결과 (c) 제안하는 알고리즘의 결과
Fig. 5. Examples of stitched frames. (a) [13] (b) [14] (c) Proposed method

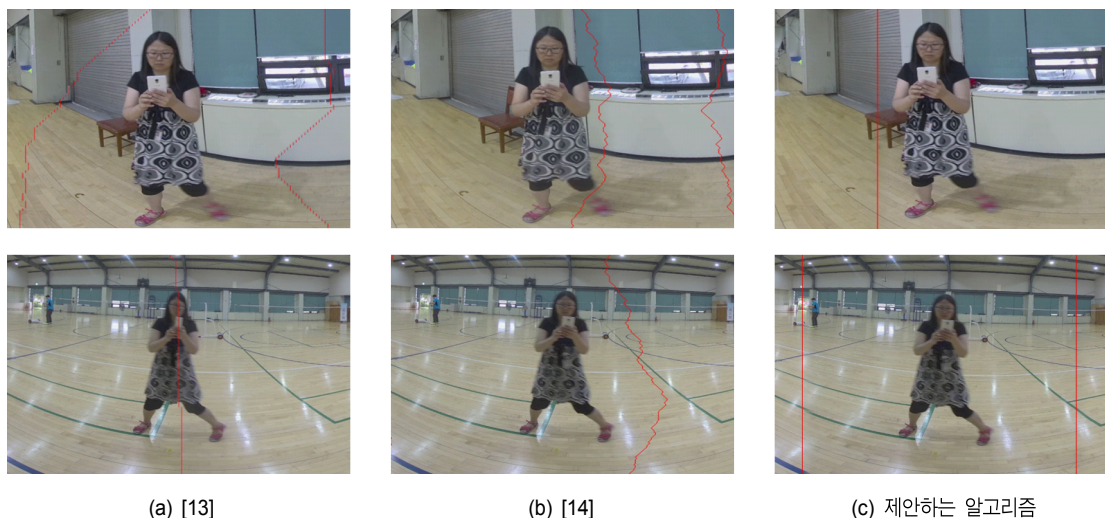


그림 6. 확대한 스티칭 결과 예시. (a) [13]의 알고리즘이 생성한 시점선 및 스티칭된 영상 (b) [14]의 알고리즘 결과 (c) 제안하는 알고리즘의 결과
Fig. 6. Examples of stitched frames (zoomed-in). (a) [13] (b) [14] (c) Proposed method

즘이 사용 가능하며, 본 논문에서는 심층신경회로망 기반의 객체 세그멘테이션 알고리즘 중 처리 속도가 빠른 YOLACT (You Only Look At CoefficientTs)^[12] 알고리즘을 사용하였다. 알고리즘의 실시간 구현을 위해서 객체 영역 추출은 GPU를 활용하고 시점선 추정 등 그 외 연산은 CPU를 활용하여 처리하도록 구현하였다. 시점선 추정 성능 검증을 위해서 기존의 [13]의 알고리즘, [14]의 알고리즘을 활용한 스티칭 결과와 제안하는 알고리즘의 스티칭 결과를 비교하였다. 비교 결과는 그림 5와 같으며, 빨간 선은 추정된 시점선이며, 그림 5의 객체 부분을 확대한 영상은 그림 6에 나타내었다. 그림 5와 6의 예시에서 볼 수 있듯이, [13]과 [14] 알고리즘은 영상의 로우 레벨(low-level) 특징만을 사용하여 시점선을 추정하므로 추정된 시점선이 객체를 피하지 못하는 경우가 발생하였으며, 시간 영역에서의 일관성을 고려하지 않아, 프레임 간 잦은 시점선 변화가 관찰되었다. 이에 반해, 제안하는 알고리즘은 시간적 일관성을 유지하면서도 전경 객체를 피해 시점선을 형성할 수 있음을 확인하였다.

또한, 제안하는 시점선 추정 방식이 에너지를 누적하여 최솟값을 찾는 방식에 비하여 처리 시간이 단축되면서도 전경 객체를 피해 시점선을 추정할 수 있음을 확인하기 위하여, 시점선 추정 시간을 측정하였다. 표 1에서 보는 바와 같이 제안한 알고리즘은 비교 알고리즘들 보다 약 3배

이상 시간이 단축되는 것을 확인하였다. 스티칭 된 동영상의 품질을 객관적으로 평가하기 위해서 Ground-truth 영상 없이 영상의 품질을 평가하는 지표인 BRISQUE(Blind/Referenceless Image Spatial Quality Evaluator)^[15]를 사용하였다. BRISQUE는 영상의 자연스러움을 평가하는 지표로 점수가 낮을수록 더 나은 시각적 품질을 나타낸다. 시험 동영상의 모든 프레임에 대한 평균 BRISQUE 점수는 표 2와 같으며, 제안하는 알고리즘이 생성한 동영상은 기존 방식과 유사하거나 높은 품질을 가지는 것을 확인하였다.

표 1. 프레임 당 평균 처리 시간 비교 (ms/frame)
Table 1. Comparison on processing time (ms/frame)

	[13]	[14]	Proposed method
Video sequence 1	6.89	4.48	1.12
Video sequence 2	15.63	5.33	1.89

표 2. 평균 BRISQUE 점수 비교
Table 2. Comparison on BRISQUE Score

	[13]	[14]	Proposed method
Video sequence 1	54.1101	53.6017	53.7667
Video sequence 2	40.7437	40.4309	40.3415

V. 결 론

본 논문에서는 동영상 스티칭의 중첩 영역 내 객체 왜곡 문제를 해결하기 위한 세그먼테이션 기반 시접선 추정 알고리즘을 제안하였다. 최소 에너지 경로를 찾기 위해 최적화 문제를 푸는 일반적인 시접선 추정 알고리즘의 처리 속도 문제를 해결하기 위하여, 제안하는 알고리즘에서는 주요 객체를 피해 시접선을 추정할 수 있도록 객체 에너지를 사용하였으며, 효율적으로 프레임 간 일관성을 유지하기 위해 직선의 시접선을 제안하였다. 실험 결과 제안하는 알고리즘을 사용하여 시접선을 추정할 경우, 비교 알고리즘보다 빠른 처리 속도로 시각적으로 유사한 360VR 동영상을 생성할 수 있음을 확인하였다.

참 고 문 헌 (References)

- [1] R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2010.
- [2] M. Brown and D. Lowe., "Automatic Panoramic Image Stitching Using Invariant Features," *International Journal of Computer Vision*, Vol.74, No.1, pp. 59-73, Dec. 2007.
- [3] J. Gao, S. J. Kim, and M. S. Brown. "Constructing Image Panoramas Using Dual-Homography Warping," *Proceedings of Conference on Computer Vision and Pattern Recognition*, CO, USA. pp. 49 - 56, 2011.
- [4] J. Zaragoza, T. J. Chin, M. S. Brown, D. Suter, "As-Projective-As-Possible Image Stitching with Moving DLT," *Proceedings of Conference on Computer Vision and Pattern Recognition*, OR, USA, pp. 2339 - 2346, 2013.
- [5] C. Herrmann, C. Wang, R.S. Bowen, E. Keyder, M. Krainin, C. Liu, R. Zabih, "Robust Image Stitching Using Multiple Registrations," *Proceedings of the European Conference on Computer Vision*, Munich, Germany, pp.53-69, 2018.
- [6] V. Kwatra, A. Schodl, I. Essa, G. Turk, A. Bobick, "Graphcut Textures: Image and Video Synthesis Using Graph Cuts," *ACM Transactions on Graphics*, Vol.22, No.3, pp.277-286, Jul. 2003.
- [7] G. Zhang, Y. He, W. Chen, J. Jia, and H. Bao, "Multi-Viewpoint Panorama Construction with Wide-Baseline Images," *IEEE Transactions on Image Processing*, Vol.25, No.7, pp.3099 - 3111, Jul. 2016.
- [8] P. Perez, M. Gangnet, A. Blake, "Poisson Image Editing," *ACM Transactions on Graphics*, Vol.22, No.3, pp. 313 - 318, Jul. 2003.
- [9] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless Image Stitching in the Gradient Domain," *Proceedings of the European Conference on Computer Vision*, Prague, Czech Republic, pp.377-389, 2006.
- [10] M. El-Saban, M. Izz, and A. Kaheel, "Fast Stitching of Videos Captured from Freely Moving Devices by Exploiting Temporal Redundancy," *Proceedings of International Conference on Image Processing*, Hong Kong, pp. 1193 - 1196, 2010.
- [11] W. Jiang and J. Gu, "Video Stitching with Spatial-Temporal Content Preserving Warping," *Proceedings of Conference on Computer Vision and Pattern Recognition Workshops*, MA, USA, pp. 42 - 48, 2015.
- [12] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-Time Instance Segmentation," *Proceedings of International Conference on Computer Vision*, Seoul, Korea, pp. 9157 - 9166, 2019.
- [13] OpenCV SeamFinder, https://docs.opencv.org/3.4/d0/dd5/classcv_1_1detail_1_1_DpSeamFinder.html
- [14] A. Efros, and T. William. "Image Quilting for Texture Synthesis and Transfer," *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, CA, USA, pp. 341-346. 2001.
- [15] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Transactions on Image Processing*, Vol.21, No.12, pp. 4695-4708, Dec. 2012.

저 자 소 개



김 태 하

- 2020년 2월 : 서울과학기술대학교 전기정보공학과 공학사
- 2020년 2월 ~ 현재 : 서울과학기술대학교 전기정보공학과 석사과정
- 주관심분야 : 컴퓨터 비전, 영상 처리, 기계 학습

저 자 소 개



양 성 엽

- 2020년 8월 : 서울과학기술대학교 전기정보공학과 공학사
- 2020년 8월 ~ 현재 : 서울과학기술대학교 전기정보공학과 석사과정
- 주관심분야 : 컴퓨터 비전, 영상 처리, 기계 학습



강 병 근

- 2018년 12월 : 미국 University of California San Diego 전기컴퓨터공학과 공학박사
- 2018년 ~ 2019년 : 미국 Carnegie Mellon University Postdoctoral Fellow
- 2019년 ~ 현재 : 서울과학기술대학교 조교수
- ORCID : <https://orcid.org/0000-0003-2537-7720>
- 주관심분야 : 컴퓨터 비전, 영상 처리, 기계 학습



이 희 경

- 1999년 2월 : 영남대학교 공과대학 컴퓨터공학과 공학사
- 2002년 2월 : KAIST-ICC 정보통신공학부 공학석사
- 2002년 ~ 현재 : 한국전자통신연구원 실감미디어연구실 책임연구원
- ORCID : <https://orcid.org/0000-0002-1502-561X>
- 주관심분야 : 디지털방송 HCI, Gaze Tracking, VR/AR/MR



서 정 일

- 1994년 2월 : 경북대학교 전자공학과 (공학사)
- 1996년 2월 : 경북대학교 대학원 전자공학과 (공학석사)
- 2005년 8월 : 경북대학교 대학원 전자공학과 (공학박사)
- 1998년 2월 ~ 2000년10월 : LG반도체 주임연구원
- 2010년 8월 ~ 2011년7월 : 영국 Southampton University, ISVR 방문연구원
- 2000년 11월 ~ 현재 : 한국전자통신연구원 실감미디어연구실 실장
- ORCID : <http://orcid.org/0000-0001-5131-0939>
- 주관심분야 : 오디오 신호처리, 실감음향, 디지털방송, 멀티미디어 표준화



이 의 진

- 2017년 9월 : 미국 University of California, San Diego 전기컴퓨터공학과 공학박사
- 2017년 ~ 2018년 : 미국 University of California, Los Angeles Postdoctoral Fellow
- 2018년 ~ 현재 : 서울과학기술대학교 조교수
- ORCID : <https://orcid.org/0000-0002-3439-5042>
- 주관심분야 : 컴퓨터 비전, 영상 처리, 기계 학습