

특집논문 (Special Paper)

방송공학회논문지 제24권 제6호, 2019년 11월 (JBE Vol. 24, No. 6, November 2019)

<https://doi.org/10.5909/JBE.2019.24.6.992>

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

이동 객체를 보존하는 시접선 추정 기술

곽 문 성^{a)}, 이 찬 혁^{a)}, 이 희 경^{b)}, 정 원 식^{b)}, 양 승 준^{a)†}

Moving Object Preserving Seamline Estimation

Moonsung Gwak^{a)}, Chanhyuk Lee^{a)}, HeeKyung Lee^{b)}, Won-Sik Cheong^{b)}, and Seungjoon Yang^{a)†}

요 약

다수의 카메라를 이용하여 취득된 영상을 정합 하여 하나의 넓은 시야각을 가지는 영상을 합성하는 기술이 많은 영역에서 사용되고 있다. 영상의 정합 영역 내에 이동 객체가 등장한 경우 객체의 왜곡 없이 영상을 정합 하기 위하여 움직임 정보를 이용하는 시접선 추정 방법을 제안한다. 기존 시접선 추정 기반 기술들은 영상 외곽선 정보와 시차 등을 이용하여 에너지 함수를 형성하고 시접선을 추정 하는 방법을 사용하여 왔으나, 본 논문에서는 프레임 별 외곽선 정보를 이용하여 움직임 정보를 계산하고, 이를 에너지 함수에 반영하여 영상 정합 시 이동 객체의 왜곡을 방지하는 시접선 추정 기술을 제안한다. 또한, 정합 영상의 왜곡 정도를 정량화 할 수 있는 지표를 제안하여 기존 기술과 제안된 방법의 성능 차이를 검증한다.

Abstract

In many applications, images acquired from multiple cameras are stitched to form an image with a wide viewing angle. We propose a method of estimating a seam line using motion information to stitch multiple images without distortion of the moving object. Existing seam estimation techniques usually utilize an energy function based on image gradient information and parallax. In this paper, we propose a seam estimation technique that prevents distortion of moving object by adding temporal motion information, which is calculated from the gradient information of each frame. We also propose a measure to quantify the distortion level of stitched images and to verify the performance differences between the existing and proposed methods.

Keyword : image stitching, moving object, seam estimation, distortion quantitative measurement, temporal motion information

a) 울산과학기술원 전기전자컴퓨터공학부(Ulsan Institute of Science and Technology, UNIST)

b) 한국전자통신연구원 실감미디어연구실(Electronics and Telecommunications Research Institute, ETRI)

† Corresponding Author : 양승준(Seungjoon Yang)

E-mail: syang@unist.ac.kr

Tel: +82-52-217-2100

ORCID: <https://orcid.org/0000-0001-9109-1582>

※ This work was supported by Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government (MSIT) (No. 2018-0-00207, Immersive Media Research Laboratory)

· Manuscript received September 23, 2019; Revised November 20, 2019; Accepted November 20, 2019.

I. 서론

최근 VR 시장의 급속한 성장과 더불어, 360도 영상 등 넓은 시야각(viewing angle)을 갖는 영상에 대한 수요가 증가하였다. 이러한 넓은 시야각의 영상은 일정한 시야각을 갖는 다수의 카메라를 이용하여, 중첩된 공간 영역을 촬영한 다수의 영상을 취득하고, 서로 중첩된 부분의 영상들을 정합 하는 과정을 통하여 얻어지게 된다. 이 과정에서, 취득된 영상들의 시차 등에 의해 정합 된 영상에 오류가 발생할 수 있다. 이러한 정합 오류를 수정하기 위하여, 특징점 기반의 공간 변환을 사용한 영상 정합과, 와핑(warping) 기반의 영상 정합 및 시점선(seamline) 기반의 영상 정합이 사용되고 있다¹⁻⁸⁾.

특징점 기반의 공간 변환 영상 정합 기술은, 영상 간 특징점을 이용하여 특징점 쌍을 형성하고, 이를 기반으로 공간 변환 행렬을 형성하여 영상을 정합 하는 기술이다¹⁻³⁾. 와핑 기술은 주로 광역적, 지역적 특징점에 의해 형성된 위치 변환 함수를 이용하여, 기하학적 투영 변환을 사용해 영상을 변환하고 이를 정합에 사용하는 방식으로 이루어진다⁴⁾⁵⁾. 그러나 객체와 배경이 카메라로부터 서로 다른 거리에 존재할 경우 하나의 변환 행렬 또는 변환 함수를 이용하여 완벽하게 정합 된 영상을 만들 수 없고, 이로 인해 정합 결과 영상이 실제로 의도된 촬영 결과와 다소 달라 질 수 있다.

시점선 기반의 영상 정합 기술은, 영상의 중첩 영역을 이루는 두 영상의 차이 등을 이용하여 에너지 함수를 구성하고 이를 최소화하는 위치를 최적 시점선으로 추정하며⁶⁾, 추정된 시점선을 따라 두 영상의 중첩 영역을 이어 붙여 왜곡을 방지하는 방식을 사용한다⁷⁾⁸⁾. 이때, 두 영상에 존재하는 강한 윤곽선의 위치에서 높은 값을 갖도록 에너지 함수를 구성하여 정합된 영상에서 윤곽선이 끊어지는 것을 방지한다. 그러나 기존의 방식들은 중첩 영역내의 에너지를 계산할 때, 이동 객체와 배경을 선택적으로 사용할 수 있는 에너지를 고려하지 않았기 때문에 이동 객체를 관통하는 시점선이 추정되는 경우가 발생하였다. 높은 시각적 중요도를 가지는 움직이는 객체의 왜곡은 정합 영상 콘텐츠 사용자로 하여금 심각한 불편을 초래할 수 있기 때문에⁹⁾, 이동 객체의 왜곡 없이 영상을 정합 할 수 있는 기술이

필요하다.

최근의 시점선 추정 및 정합 방법에서는 이동 객체를 보존하며 영상을 정합하는 기술들이 시도되고 있다. 시점선 추정 시 사용되는 외곽선(gradient)과 시차 정보 외에 객체의 왜곡을 세가지 분류(나타남, 잘림, 복제)로 정의하고 이를 중첩 영역 내의 에너지 계산 시 활용하거나¹⁰⁾, 인간의 시각적 중요도를 나타내는 항목인 세일리언시(Saliency)를 이용하여 에너지 계산시에 객체 별 가중치를 부여하는 등의 방법이 연구되었다¹¹⁾. 하지만, 이러한 방법들은 실시간 환경에서 사용되기에는 계산량이 많다는 문제가 있다.

본 논문에서는 다수의 카메라를 사용하여 취득하고 공간 변환을 적용한 영상에서 기존 시점선 추정 시 이용되는 영상의 외곽선 정보, 시차 정보와 더불어, 이동 객체를 고려한 에너지 정보를 추가하여 시점선을 추정하고 추정된 시점선을 이용하여 정합한 결과를 서술한다. 또한, 영상에서 중첩 영역 내 객체의 왜곡 정도를 정량화 할 수 있는 지표를 제안하고, 이를 이용하여 기존의 시점선 추정 및 정합 방법과의 성능을 비교 검증한다.

본 논문은 다음과 같이 구성된다. 섹션 II에서는 제안한 알고리즘에 대한 설명이며, 섹션 III에서는 제안한 알고리즘을 평가하기 위한 방법을 제안한다. 섹션 IV에서 제안 알고리즘과 기존 알고리즘의 성능 비교 실험을 서술하며, 섹션 V에서 결론을 제시한다.

II. 객체의 움직임을 이용한 시점선 추정 및 영상 정합

2장에서 설명에 사용되는 영상들은 영상 취득을 위해 다수의 렌즈 부착이 가능한 고정된 카메라에서 어안 렌즈를 이용해 취득한 영상들을 등장방형도법(equirectangular projection, ERP) 기법으로 공간 변환한 영상들이라 가정한다. 영상들을 이용하여 넓은 시야각을 가지는 영상을 합성하기 위해, 이웃한 두 카메라에서 촬영한 영상을 정합 하는 방법을 설명한다. 이 방법을 연속적으로 이웃한 카메라에 수행하면, 최대 360도의 시야각을 가지는 영상을 합성할 수 있다.

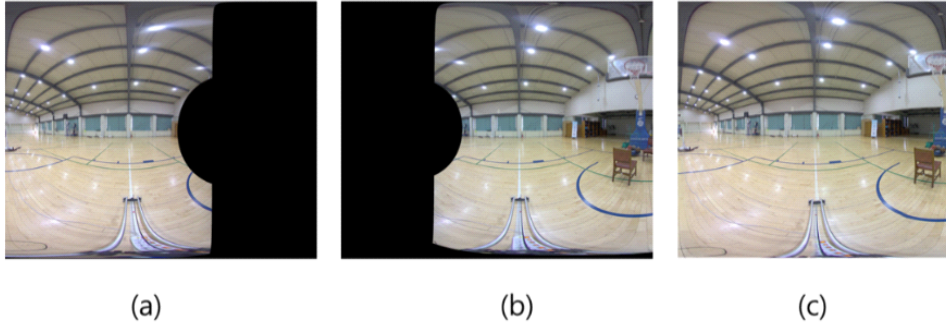


그림 1. (a), (b): ERP 변환 영상의 일부, (c) 정합 영상

Fig. 1. (a), (b): parts of ERP converted images, (c) stitching image therefrom

1. 시접선 추정

시접선 추정은 정합 하고자 하는 두 영상에 포함된 객체 및 배경의 구조를 나타내는 에너지를 구한 후 에너지가 작은 부분을 통과하는 시접선을 추정한다. 추정된 시접선 위치에서 두 영상을 정합 하면 큰 에너지를 갖는 객체 및 배경의 구조가 어긋나지 않은 정합 영상을 만들 수 있다.

시접선 추정에 사용하는 에너지 E 는 세가지 에너지의 가중치 합으로

$$E(x, y, t) = c_d E_d(x, y, t) + c_g E_g(x, y, t) + c_o E_o(x, y, t) \quad (1)$$

로 구성된다. 여기서 E_d , E_g , E_o 는 세가지 에너지 함수이고, c_d , c_g , c_o 는 각각의 에너지 항에 곱해지는 가중치이다. 각각의 에너지 함수를 설명하면 다음과 같다. 이웃하는 두 영상의 중첩 영역 영상을 각각 $I^1(x, y, t)$ 와 $I^2(x, y, t)$ 라고 한다. 영상의 사이즈를 $n \times m \times r$ 라 하였을 때, (x, y) 는 픽셀의 위치를 나타내는 벡터이고 t 는 동영상의 프레임 시간을 나타낸다. 3차원 공간 상의 한 점이 두 카메라의 시차에 의해 두 영상의 서로 다른 픽셀 위치에서 취득되는 것을 반영한 에너지는

$$E_d(x, y, t) = |I^1(x, y, t) - I^2(x, y, t)| \quad (2)$$

로 주어진다. 이 에너지는 큰 시차를 가지는 동일한 객체 및 배경이, 서로 어긋나지 않도록 시접선을 형성하게 해 준

다. 두 영상의 외곽선이 서로 일치하지 않는 것을 반영하는 에너지는

$$E_g(x, y, t) = \left| \frac{\partial}{\partial x} I^1(x, y, t) - \frac{\partial}{\partial x} I^2(x, y, t) \right| + \left| \frac{\partial}{\partial y} I^1(x, y, t) - \frac{\partial}{\partial y} I^2(x, y, t) \right| \quad (3)$$

로 주어진다. 이 에너지는 영상 정합 시 외곽선이 존재하는 부분에 시접선이 위치하지 않도록 해 준다. 객체의 움직임 을 반영하는 에너지는

$$E_o(x, y, t) = |E_g(x, y, t) - E_g(x, y, t-1)| + |E_g(x, y, t-1) - E_g(x, y, t-2)| \quad (4)$$

로 주어진다. 이 에너지는 객체의 움직임에 따라 픽셀 값의 시간적 변화가 큰 부분에 시접선이 위치하지 않도록 해 준다.

두 영상이 서로 유사하여 눈에 띄는 왜곡 없이 서로 붙일 수 있는 위치를 나타내는 시접선을 \vec{s} 라 할 때, 이를

$$\vec{s} = \{s_i, i = 1, 2, \dots, n\} = \{(x^i, i), i = 1, 2, \dots, n\} \quad (5)$$

로 나타낼 수 있다. 여기서 x^i 는 1 부터 m 사이의 상수 값을 가지는 상수이고, i 는 위에서 아래로 연속되는 픽셀 y 좌표값이다. 시접선의 에너지 $E(\vec{s}, t)$ 는 시접선을 이루는 각 픽셀 에너지의 총합인 $\sum_{i=1}^n E(x^i, i, t)$ 로 표현될 수 있다. 최

적의 시접선을 \vec{s}^* 이라 할 경우, 이는 아래 수식에 의해 정의된다.

$$\vec{s}^* = \min_{\vec{s}} E(\vec{s}, t) = \min_{\vec{s}} \sum_{i=1}^n E(x^i, i, t) \quad (6)$$

최적의 시접선은 동적 계획법에 의해 찾아 질 수 있다. 전체 에너지 E 와 같은 사이즈의 누적 최소 에너지 M 이 있다고 가정하고, 첫 번째 행에는 전체 에너지 E 의 첫 번째 행 값과 동일한 값을 부여한다. 그리고 다음 과정을 통해 누적 최소 에너지의 두 번째 행부터 마지막 행 까지를 연산해가며 모든 가능한 시접선의 에너지를 계산한다^[6].

두번째 행부터 마지막 행까지, 행을 먼저 고정하고 열을 이동시키며 식 (7)을 수행하고, 마지막 열에 도달한 경우 행을 바꾸어 다시 첫 열부터 마지막 열까지 식 (7)을 수행한다. 연산이 완료되었을 때, 마지막 행에서 최소값을 가지는

픽셀 위치가 최적 시접선의 마지막 픽셀 위치이다. 그 위치로부터 위 행의 근접한 3픽셀을 살펴며 최소값을 가지는 위치를 기록하며 전진하여 최적 시접선 \vec{s}^* 를 연산한다^[6]. 구해진 누적 최소 에너지 M 을 활용하면 영상에서 중요도가 높은 영역을 시각적으로 확인할 수 있다.

2. 시접선 업데이트

시접선은 수식 (7)을 이용한 일련의 과정으로 매 프레임 추정된다. 모든 프레임에서 추정된 시접선을 이용하여 영상을 정합 하면, 정합 영역에서 잦은 시접선 위치 변화에 따른 영상 왜곡이 영상 관측자에게 인지될 수 있다. 이를 방지하기 위해, 주어진 프레임에서 추정된 시접선과 이전에 추정된 시접선의 에너지 차가 큰 경우에만 시접선을 업데이트한다.

그림 2는 시접선 업데이트가 불필요한 경우와 필요한 경우를 보여준다. t_0 와 t_1 시간에서 추정된 시접선이 각각 빨

$$M(x, y, t) = E(x, y, t) + \min(M(x-1, y-1, t), M(x, y-1, t), M(x+1, y-1, t)) \quad (7)$$

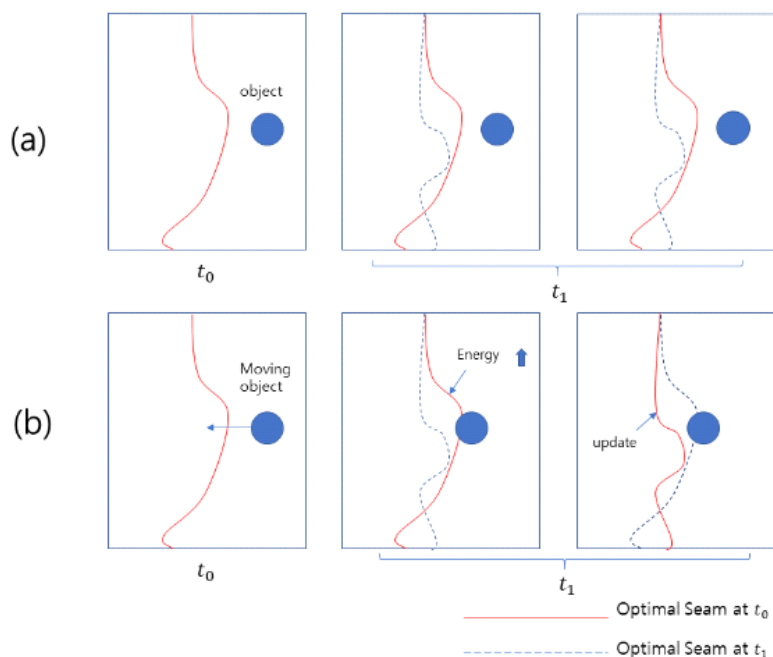


그림 2. (a) 객체의 출현이 없어 시접선 변화가 없는 경우, (b) 객체가 출현하여 시접선이 업데이트 되는 경우

Fig. 2. (a) No change in seam line due to no appearance of object, (b) When seam line is updated due to appearance of object

간 선과 파란 점선으로 표시 되어있다. 그림 2의 (a)의 경우에는 움직이는 객체가 존재하지 않아, t_0 시간의 영상과 t_1 시간의 영상이 비슷하고 따라서 두 시간에서 각각 추정된 시접선의 에너지 차가 적다. 따라서 시접선 업데이트가 일어나지 않는다. 그림 2의 (b)의 경우에는 객체의 이동에 의해 t_0 와 t_1 시간에서 추정된 시접선 위치의 영상이 달라져 두 시접선 위치에서 에너지 차이가 발생한다. 그 결과 시접선 업데이트가 일어난다.

현재 프레임에서 추정된 시접선과 이전 프레임에서 추정된 시접선이 통과하는 위치의 에너지차는

$$\delta(t) = |E(\vec{s}^*(t), t) - E(\vec{s}^*(t-1), t)| \quad (8)$$

로 구해진다. 이를 이용하여 주어진 프레임에서 사용할 시접선은

$$\hat{s}^*(t) = \begin{cases} \vec{s}^*(t) & \text{if } \delta(t) > \epsilon \\ \vec{s}^*(t-1), & \text{otherwise} \end{cases} \quad (9)$$

에 의해 결정된다.

3. 영상 정합

추정된 시접선을 이용하여 이웃한 두 영상을 정합 하여 하나의 영상으로 합성한다. 시접선 기준으로 좌우 영상을 그

대로 이용하면 시접선의 위치가 그대로 들어나 영상 품질을 저하시킬 수 있기 때문에, 시접선을 기준으로 마스크를 연산하고, 연산된 마스크의 경계를 기준으로 좌우 선형 블렌딩을 수행한 후 두 영상을 합하여 정합 영상을 합성한다.

그림 3은 선형 블렌딩을 이용한 영상 정합의 예시를 나타낸다. 그림 3(a) (b)는 정합에 이용되는 좌우 영상 $I^1(x, y, t)$, $I^2(x, y, t)$ 의 일부이고 그림 3(c)는 선형 블렌딩에 이용되는 마스크 $w(x, y, t)$ 이다. 마스크 $w(x, y, t)$ 는, 추정된 시접선의 왼쪽 픽셀이 1을 가지도록 하고, 오른쪽에 있는 픽셀들이 0을 가지도록 한 후, 자연스러운 정합 결과를 얻기 위하여 가우시안 필터 등을 이용하여 필터링을 수행하였다. 이를 이용하여

$$I^{1,2}(x, y, t) = w(x, y, t) * I^1(x, y, t) + (1 + w(x, y, t)) * I^2(x, y, t) \quad (10)$$

로 정합 영상을 합성한다.

k 개의 카메라로부터 취득된 k 개의 ERP 공간 변환 영상 $I^1(x, y, t)$, $I^2(x, y, t)$, ..., $I^k(x, y, t)$ 을 중첩 영역에서 추정된 개의 시접선을 이용하여 연속적으로 정합 하여 최대 360도 시야각의 정합 영상을 합성할 수 있다. 이 사실을 적용하여 수식 (10)을 변형하면

$$I^{1,\dots,k}(x, y, t) = w^t(x, y, t) * I^{1,\dots,k-1}(x, y, t) + (1 - w^k(x, y, t)) * I^k(x, y, t) \quad (11)$$

로 정의된다. 여기서 k 는 1 보다 큰 양수이며, w^k 는 $k-1$ 와

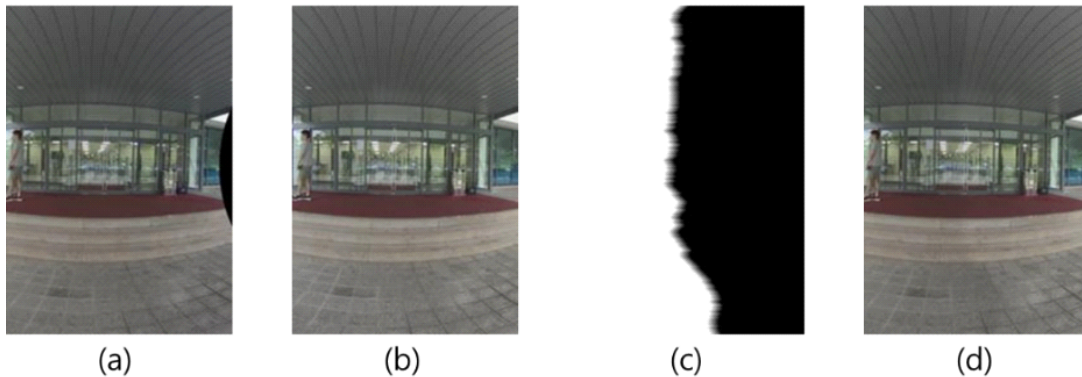


그림 3. 선형 블렌딩을 이용한 영상 정합 예시. (a) 좌측 영상, (b) 우측영상, (c) 선형 블렌딩 마스크, (d) 정합 영상

Fig. 3. Example of image stitching using linear blending. (a) left image, (b) right image, (c) linear blending mask, (d) stitching image

k 번째 영상의 시점선 추정 결과로부터 얻어진 선형 블렌딩 마스크이다. 이때, w^k 는 합성 영상 $I^{1,\dots,k}(x,y,t)$ 와 같은 해상도를 가진다.

III. 정합 영상의 왜곡 측정을 위한 정량적 지표

이웃한 카메라로부터 취득된 두 영상에는 같은 공간을 촬영한 중첩 영역이 있다. 중첩 영역의 객체 및 배경은 시점선을 기준으로 좌측 또는 우측 영상에 왜곡 또는 깨짐 없이 존재한다. 따라서 좌우 영상과 정합된 영상을 비교하여 객체 및 배경의 왜곡 및 깨짐에 따른 영상 품질을 측정할 수 있다.

이웃하는 두 카메라로부터 취득한 영상 $I^1(x,y,t)$, $I^2(x,y,t)$ 와 정합된 영상 $I^{1,2}(x,y,t)$ 을 이용하여

$$ErrMap^{1,2}(x,y,t) = \min((I^1(x,y,t) - I^{1,2}(x,y,t))^2, (I^2(x,y,t) - I^{1,2}(x,y,t))^2) \quad (12)$$

를 연산하고, 연산의 결과로 정합된 영상과 좌측 또는 우측 영상의 차이가 발생하는 곳의 에러 맵을 구성할 수 있다. 이를 확장하면, k 개의 ERP 영상을 이용하여 정합된 영상 $I^{1,\dots,k}(x,y,t)$ 의 에러 맵은 식 (13)으로 나타내어진다. 이로부터 왜곡 정량 지표는 식 (14)를 이용하여 연산한다.

$$DQM(t) = \sum_{(x,y)} E^{1,2,\dots,k}(x,y,t) \quad (14)$$

IV. 실험

1. 영상 취득 시스템

실험에 사용한 영상들은 어안렌즈가 장착된 카메라가 60도 간격으로 이웃하여 6개 존재하고, 각 렌즈의 화각이 185도인 360도 카메라에서 취득한 동영상들을 프레임별로 ERP

공간 변환하여 생성하였다. 영상들은 다양한 배경 환경에서 객체의 이동이 있는 상태에서 촬영되었으며, 이를 이용하여 제안 방법이 움직이는 객체를 성공적으로 보존하는 시점선을 형성하는지 확인할 수 있었다. 제안한 알고리즘 구현과 테스트는 운영체제 WINDOWS 10에서 MATLAB R2019a를 사용하였다. 알고리즘 테스트에 사용된 동영상은 총 4종이며, 영상 별로 약 300 ~ 900 프레임을 포함하고 있다. 4종의 영상을 이용하여 제안하는 시점선 추정 및 업데이트 기술과 인용 논문 [8]의 시점선 추정 및 업데이트 방법을 비교하며, 정합 결과 영상에 정합 영상 왜곡 측정 지표를 적용하여 비교 실험을 수행한다.

ERP 변환 영상에서 시점선 추정을 위해 사용되는 영역은 이웃한 두 영상에서 같은 공간을 포함하는 중첩 영역이다. 6대의 어안렌즈 카메라를 60도 간격으로 배치한 카메라에서 취득한 영상을 정합하기 위하여 총 6개의 중첩 영역이 사용되며, 중첩 영역들의 좌우 영상은 실험적으로 확인된 고정 좌표에서 취득한다.

2. 시점선 추정 비교 실험

인용 논문[8]에서는 시점선을 추정하기 위하여, 본 논문의 2장에서 설명한 세 에너지 항목 중 외곽선과 시차 에너지 항목만을 이용한다. 시점선 업데이트를 수행하기 위해서는, 새로운 프레임에서 시점선을 추정하고, 기존 시점선과의 픽셀 별 외곽선 에너지 차이를 연산한다. 연산된 외곽선 에너지 차이가 일정 값 이상을 넘어가는 픽셀의 개수를 기록하여, 시점선 전체 픽셀 개수의 0.3배가 넘어가면 업데이트를 수행한다.

그림 4는 비교 실험 수행 결과 중 일부이다. 움직이는 객체에 의한 업데이트가 일어나는 연속 프레임이고 (a)와 (b)는 인용 논문[8]의 알고리즘을, (c)와 (d)는 제안하는 방법을 적용한 경우이다. [8]의 알고리즘에 의한 결과물은, 객체에 의한 왜곡이 다소 많이 발생한 이후 업데이트 동작을 수행하는 것으로 보이며, 이는 업데이트 시, 일정 개수의 픽셀 이상에서 외곽선 에너지 값 차이가 존재하여야 하기 때문으로 판단된다. 반면 제안 알고리즘에 의한 결과는, 시

$$ErrMap^{1,\dots,k}(x,y,t) = \min((I^1(x,y,t) - I^{1,\dots,k}(x,y,t))^2, \dots, (I^k(x,y,t) - I^{1,\dots,k}(x,y,t))^2) \quad (13)$$

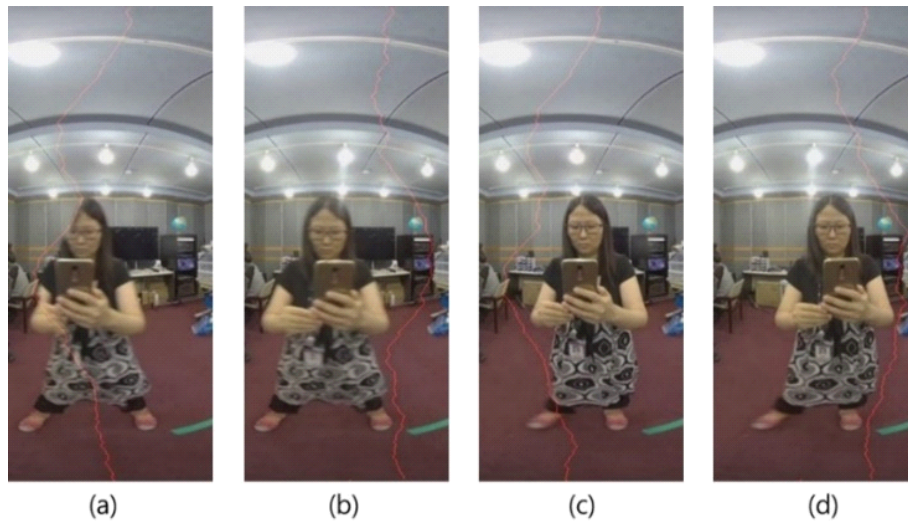


그림 4. (a) 인용 논문 [8] 알고리즘 시접선 변경 전, (b) 인용 논문 [8] 알고리즘 시접선 변경 후, (c) 제안 알고리즘 시접선 변경 전, (d) 제안 알고리즘 시접선 변경 후

Fig. 4. (a) algorithm of the [8]: before changing the seam line (b) algorithm of the [8]: after changing the seam line (c) proposed algorithm: before changing the seam of the proposed paper, and (d) proposed algorithm: after changing the seam of the proposed paper

접선 에너지 총합의 변화량이 일정량이 넘는 경우 업데이트 되도록 제안되었고, 움직이는 객체에 의한 에너지가 기존 알고리즘에 비해 높기 때문에, 이동 객체에 의한 업데이

트가 빠르게 나타나는 것으로 판단된다.

그림 5의 (a)는 에너지 맵, (b)는 누적 최소 에너지 맵이며 각 열은 그림 4에 대응한다. 영상 별로 최대값 정규화가 적

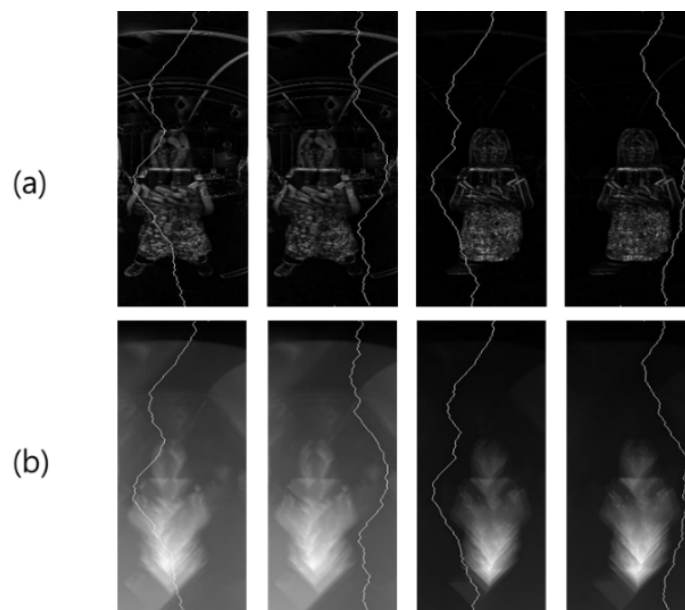


그림 5. (a): 에너지 맵, (b): 누적 최소 에너지 맵, 각 열은 그림 4의 (a) (b) (c) (d)에 대응

Fig. 5. (a): energy map, (b): cumulative minimum energy map, each column corresponding to (a) (b) (c) (d) in Fig 4

용되어, 높은 값을 가지는 영역을 강하게 보여주고 있다. 이를 통해 제안한 알고리즘에서 움직이는 객체가 배경에 비해 높은 에너지와 중요도를 가지는 것을 확인할 수 있다.

3. 정합 영상 품질 비교 실험

그림 4의 연속 프레임 결과에 정합 영상의 왜곡 정량 지표 값을 적용하여 보면 테이블 1의 결과와 같다. (13) 번 수식을 이용하여 에러 맵을 연산하고 결과물에 색 코딩을 적용하면 그림 6과 같이 왜곡이 발생한 영역을 시각화 할 수 있다. 작은 왜곡 정량 지표값을 가지는 위치는 파란색으로 시각화 되었으며 왜곡이 거의 없음을 의미한다. 반대로 붉

표 1. 그림 4 (a) (c) 정합 영상 왜곡 정량 지표 값 비교 (단위 :1e-5)
Table 1. Figure 4 (a) (c) stitching image distortion quantitative measurement comparison (Unit: 1e-5)

	[8]	proposed
distortion quantitative measurement	4.0715	0.7320

은색으로 시각화 된 구역은, 큰 왜곡 정량 지표값을 가지고 있으며 왜곡이 크게 발생하였음을 의미한다.

정합 영상의 왜곡에 대한 정량적 지표를 이용하여 동영상의 전반적인 왜곡 정량 지표 값을 계산할 수 있다. 객체가 중첩 영역을 지나감에 따라, 왜곡 정량 지표 값이 일정 프레임 내에서 크게 변동하며, 객체를 피해 형성된 마지막 시접

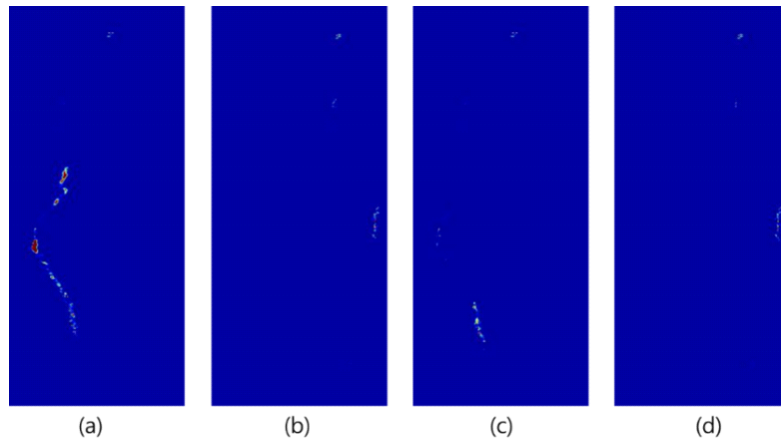


그림 6. (a), (b), (c), (d): 그림 4 (a), (b), (c), (d)의 에러 맵 연산과 색 코딩을 이용한 왜곡 시각화
Fig. 6. (a), (b), (c), (d): Distortion visualization using error map calculation and color coding of Fig. 4 (a), (b), (c) and (d)

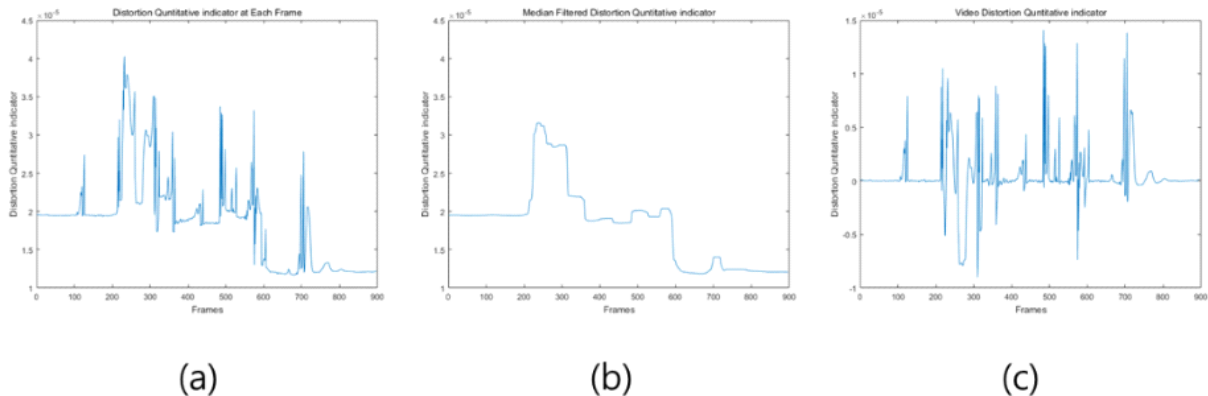


그림 7. (a): 프레임 별 왜곡 정량 지표 값, (b): (a)에 50프레임 중간값 필터 적용 결과 (c): (a)와 (b)의 차연산 수행 결과
Fig. 7. (a): distortion quantitative measurement for each frame, (b): result of applying 50 frame median filter to (a), (c): Result of difference calculation of (a) and (b)

선 위치에 따라 왜곡 정량 지표 값이 업데이트 전과 달라진다. (그림 7의 (a)). 이동 객체에 의한 업데이트는 대체로 50프레임 안에서 동작하기 때문에 객체 이동에 의한 왜곡 정량 지표의 영향만을 고려하기 위하여, 50 프레임 안에서 구해진 중간 값과 정량 지표 값을 이용하여 차연산을 수행한다(그림 7의 (b), (c)). 이를 절대 값을 취한 후 합 연산하여 동영상의 중간값 보정 왜곡 지표를 계산한다.

표 2. 정합 동영상의 중간값 보정 왜곡 정량 지표 값 (단위 :1e-4)
Table 2. Median Compensated Distortion quantitative measurement of video when distortion occurs (Unit: 1e-4)

	[8]	proposed
8k video 1	23.4847	9.4141
8k video 2	6.4450	4.7574
8k video 3	26.0437	11.7524
8k video 4	7.6989	4.4046

V. 결 론

영상 정합에 시접선을 이용한 추정 방법을 주로 사용하고 있으나, 객체와 배경의 에너지를 동일하게 취급하는 에너지 항목들 만으로는 추정된 시접선이 객체를 통과하여 영상에 왜곡이 발생하는 경우를 방지하기 어렵다. 이를 극복하기 위해 본 논문은 객체의 움직임을 반영하는 에너지를 시접선을 추정하기 위한 에너지 항목에 추가하여, 높은 시각적 중요도를 가지는 이동 객체에서의 정합 왜곡 현상을 방지하였다. 이를 증명하기 위하여 정합 영상의 왜곡을 정량적으로 나타내는 지표를 제시하고, 비교 실험을 수행하였으며, 왜곡 정량 지표 값을 통해 동영상의 전반적인 정합 품질이 향상되는 것을 확인하였다. 시접선 추정 기반의 영상 정합 기술들은, 배경 에너지와 분리될 수 있는 객체 에너지를 이용하여 객체의 실제 위치보다 몇 픽셀 상하좌

우의 에너지를 활용하여 시접선 업데이트가 더욱 빨리 일어날 수 있게 하거나, 느린 움직임을 보이는 객체 등에서 선별적으로 에너지를 강화하는 등의 다양한 응용이 가능할 것으로 기대 된다.

참 고 문 헌 (References)

- [1] L. Juan and O. Gwun, "SURF applied in panorama image stitching", IEEE 2010 2nd international conference on image processing theory, tools and applications, p. 495-499, July, 2010.
- [2] Y. Li, Y. Wang, W. Huang and Z. Zhang, "Automatic image stitching using SIFT", IEEE 2008 International Conference on Audio, Language and Image Processing, pp. 568-571, July, 2008.
- [3] W. Y. Lin, S. Liu, Y. Matsushita, T. T. Ng and L. F. Cheong, "Smoothly varying affine stitching", IEEE Conference on Computer Vision and Pattern Recognition 2011, pp. 345-352, June, 2011.
- [4] C.H. Chang, Y. Sato, and Y.Y. Chuang, "Shape-preserving half-projective warps for image stitching", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3254-3261, 2014.
- [5] J. Zaragoza, T. J. Julio, M.S. Brown and D. Suter, "As-projective-as-possible image stitching with moving DLT", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2339-2346, 2013.
- [6] S. Avidan and A. Shamir, "Seam carving for content-aware image resizing", ACM Transactions on graphics (TOG), Vol. 26, No. 3, p. 10, 2007.
- [7] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection", IEEE transactions on pattern analysis and machine intelligence, Vol. 34, No. 10, pp. 1915-1926, 2011.
- [8] B. He and S. Yu, "Parallax-robust surveillance video stitching", Sensors, Vol. 15, No. 7, 2015.
- [9] J. Gao, Y. Li, T. J. Chin and M. S. Brown, "Seam-Driven Image Stitching", In Eurographics (Short Papers), pp. 45-48, May, 2013.
- [10] C. Herrmann C. Wang, R. Strong Bowen, E. Keyder and R. Zabih, "Object-centered image stitching", Proceedings of the European Conference on Computer Vision (ECCV), pp. 821-835, 2018.
- [11] N. Li, T. Liao and C. Wang, "Perception-based seam cutting for image stitching", Signal, Image and Video Processing, Vol 12, No. 5, p967-974, 2018.

저 자 소 개



곽 문 성

- 2015년 8월 : 울산과학기술원 전기전자컴퓨터공학부 학사
- 2015년 9월 ~ 현재 : 울산과학기술원 전기전자컴퓨터공학부 석박통합과정
- ORCID : <https://orcid.org/0000-0003-2002-5961>
- 주관심분야 : 영상처리, 머신러닝



이 찬 혁

- 2018년 2월 : 울산과학기술원 전기전자컴퓨터공학부 학사
- 2018년 3월 ~ 현재 : 울산과학기술원 전기전자컴퓨터공학부 석사과정
- ORCID : <https://orcid.org/0000-0002-5749-0800>
- 주관심분야 : 영상처리, 머신러닝



이 희 경

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



정 원 식

- 1992년 : 경북대학교 공과대학 전자공학과 졸업 (공학사)
- 1994년 : 경북대학교 대학원 전자공학과 졸업 (공학석사)
- 2000년 : 경북대학교 대학원 전자공학과 졸업 (공학박사)
- 2000년 ~ 현재 : 한국전자통신연구원 실감미디어연구실 책임연구원
- 주관심분야 : 디지털방송, MPEG, VR/AR/MR



양 승 준

- 1990년 : 서울대학교 전기공학과 학사
- 1993년 : University of Wisconsin-Madison 석사
- 2000년 : University of Wisconsin-Madison 박사
- 2000년 ~ 2008년 : 삼성전자 수석연구원
- 2009년 ~ 현재 : 울산과학기술원 전기전자컴퓨터공학부 부교수
- ORCID : <https://orcid.org/0000-0001-9109-1582>
- 주관심분야 : 영상 처리