

영상 분할을 이용한 블록 기반의 스테레오 정합

김 지 수*, 박 철 형*, 이 덕 우°

Block-Based Stereo Matching Using Image Segmentation

Ji-su Kim*, Cheol-hyeong Park*, Deok-woo Lee°

요 약

본 논문에서는 영상 분할(image segmentation)을 이용한 블록(block) 기반의 스테레오 정합(stereo matching) 기법을 제안한다. 블록 기반의 매칭 방식은 스테레오 정합 알고리즘 중 국부적 방법의 가장 대표적인 알고리즘이다. 그러나 스테레오 정합 문제에 적용될 때 가려진 영역(occlusion)이나 텍스처(texture) 정보가 부족한 영역에서 오정합 발생 확률이 매우 높아 만족할 만한 성능을 제공하지 못한다. 본 논문에서 제안하는 알고리즘은 영상을 분할하여 정확한 경계정보를 획득한 다음 한 영상의 작은 영역에 대하여 상관성이 가장 높은 영역이나 차이값이 가장 낮은 영역을 참조 영상에서 탐색하여 기존의 알고리즘보다 더 높은 정확도를 제공한다. Middlebury dataset에서 제공된 영상들을 활용하여 스테레오 영상에 적용한 실험 결과들은 제안된 알고리즘이 기존의 블록 기반 알고리즘보다 더 좋은 성능을 보여주는 것을 입증해준다.

Key Words : Image Segmentation, Block matching, Stereo Matching, Stereo camera, Occlusion

ABSTRACT

In this paper, we propose the approach to stereo matching based on block-based method and using image segmentation. Block based matching, local method, has been popularly and widely employed for stereo matching. In practice, however, it sometimes does not provide sufficiently accurate matching results if textureless region or occlusion is inherent in a scene of interest. This paper suggests the algorithm that provides higher matching accuracy than the existed ones by segmenting the sub-region of an image to acquire boundary regions followed by searching the matched region in the reference image. Searching process is carried out based on high correlation and the lowest difference between matching areas. Middlebury dataset is used in the experiments and the proposed approach outperforms the existed block-matching algorithm.

I. 서 론

인간이 3차원 입체감을 느끼게 하는 방법은 두 눈이 한 장면을 바라볼 때 생기는 양안 시차(binocular parallax)를 이용하는 것이다. 이러한 원리를 활용하여 스테레오 방식으로 기계가 3차원 영상을 인식할 수

있다. 즉, 좌, 우 두 눈에 서로 다른 영상이 보이게 함으로써 우리가 실생활에서 보는 것과 같은 입체감을 느낄 수 있게 하는 것이다. 이러한 스테레오 방식을 사람이 아닌 기계가 계산할 수 있도록 하는 기술이 컴퓨터 비전이다. 컴퓨터 비전에서 스테레오 정합(stereo matching)은 기존 영상과 정합이 이루어지는 영상 사

※ 본 연구는 2018년도 계명대학교 연구기금으로 이루어졌음

♦ First Author : Keimyung University Department of Computer Engineering, koo5679@gmail.com, 학생회원

° Corresponding Author : Keimyung University Department of Computer Engineering, dwoolee@kmu.ac.kr, 정회원

* Keimyung University Department of Computer Engineering, parkcheolhyeong@gmail.com

논문번호 : 201903-022-D-RN, Received March 18, 2019; Revised May 9, 2019; Accepted May 10, 2019

이에 실제계의 동일한 한 점을 나타내는 픽셀의 위치 차를 구하는 과정인 디스패리티(disparity) 계산에서 반드시 먼저 수행되어야 할 과정으로, 이 정보를 이용하여 2차원의 영상으로부터 3차원의 기하학적 구조를 복원할 수 있다. 두 픽셀간의 위치 차를 디스패리티 또는 변이(disparity)라 부르며, 전체 영상에 대해서 구한 변이 값들을 나타낸 것을 변이 지도(disparity map)라 한다. 스테레오 정합은 잡음, 물체 간 불연속(depth discontinuity), 폐색 영역(occlusion region) 등으로 인한 모호성으로 인해 완전한 정확도를 제공하는 정합 알고리즘은 존재하지 않는다^[1]. 때문에 많은 정합 방법들이 영상의 환경에 최적화하는 방향으로 제시되고 있다. 그동안 제안된 정합 방법은 전역적 정합 방법(global matching method), 국부적 정합 방법(local matching method)으로 나눌 수 있다.^[2] 전역적 방법은 텍스처 정보가 부족한 영역(textureless region)이나 폐색영역에서의 오정합(matching error) 발생을 줄이기 위해 비국부적 제약조건을 활용하는 방법이다. 이러한 전역적 방법은 정확한 변이를 추정할 수 있는 반면 계산 비용이 크다는 단점이 있으며, 전역적 방법을 활용하는 대표적인 알고리즘으로는 그래프 컷(graph cut), 신뢰확산(belief propagation), 다이내믹 프로그래밍(dynamic programming) 기법 등이 있다.^[3] 그래프 컷은 컴퓨터 네트워크 흐름제어에서 출발하였으며 최대-흐름/최소-컷(max-flow/min-cut) 알고리즘을 이용하여 변이를 할당하는 방식으로 변이 추정에 가장 좋은 성능을 내는 방법 중 하나이다. 신뢰확산은 메시지 전달에 기반을 둔 반복 알고리즘으로 성능은 우수한 반면 완전히 수렴된 메시지를 얻기 위해서 많은 반복이 필요해 실행 시간이 길다는 단점이 있다.^[4] 다이내믹 프로그래밍은 본래 어떠한 문제를 해결하는데 있어 최적화 문제에 대한 계산상의 복잡성을 줄이기 위해 제안된 수학적 방법이다. 다이내믹 프로그래밍 기반 변이 추정 방식이란 기준영상과 참조영상 사이의 관계를 이용하여 생성한 변이 공간 영상(DSI : disparity space image)으로부터 최적의 변이 경로를 찾아가는 방식을 말한다.^[5] 이에 반해 국부적 방법은 비교적 연산량이 적고 구현이 쉬운 반면 가려진 영역(occlusion)이나 텍스처 정보가 부족한 영역에서 오정합 발생 확률이 매우 높다. 또한, 정합창 사용으로 인한 경계 영역에서의 부정확한 변이가 할당되는 단점으로 인해 정확성 보다는 실시간이나 고속의 프로세스가 요구되는 분야에 사용되고 있다.^[6] 국부적 방법은 특징 기반(feature-based) 정합 방법과 지역 기반(area-based) 정합 방법으로 나눌 수 있다. 특징 기반

매칭 방식은 단순히 픽셀의 밝기 값을 이용하지 않고 모서리(edge), 구석(corner), 골(cave) 등 영상이 가지는 특징을 이용한다.^[7] 특징 기반 정합 방법에는 SIFT(Scale Invariant Feature Transform)방법과 SURF(Speed-Up Robust Feature)방법 등이 있다^[11,12]. SIFT 알고리즘은 영상 회전, 스케일 변화, 유사성을 가진 변형, 관점 변화, 잡음, 조명 변화에 매우 강인성을 가진 특징 추출 알고리즘이다. SURF 알고리즘은 SIFT의 단점인 연산의 높은 복잡도를 해결하기 위한 방법으로 제안되었으며, 다중-스케일 공간 정리(Multi-Scale Space Theory)에 기반하며, 특징 추출은 성능과 정확성에서 우수한 헤시안 행렬(Hessian Matrix)을 기반으로 수행된다.^[8] 국부적 방법 중 지역 기반 정합 방법의 가장 대표적 알고리즘은 한 영상의 작은 영역에 대하여 상관성이 가장 높은 영역이나, 영역 간의 특징값의 차이가 가장 낮은 영역을 참조 영상에서 탐색하는 블록 매칭 알고리즘이다. 블록 매칭 알고리즘은 전역적 정합 방법에 비해 상대적으로 부정확한 깊이 영상을 출력하지만, 전역적 정합 방법에 비해 알고리즘이 단순하고 빠르게 계산이 가능하다는 장점이 있다.

본 논문에서는 영상 분할을 이용한 블록 기반의 스테레오 정합 방법을 제시한다. 기존의 블록 기반의 스테레오 정합 방법의 결과로부터 나온 정확도를 개선하기 위하여 영상 분할을 이용한 이미지를 블록 기반의 스테레오 정합 방법에 적용하였다. 제안하는 방법은 일반적인 블록 기반의 스테레오 정합 방법의 장점을 유지하면서 더 높은 정확도를 제공하며 기존의 블록 기반의 스테레오 정합 알고리즘보다 우수한 결과를 보여준다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 방식의 블록 기반 스테레오 정합 방법에 대해서 알아보고, 3장에서는 제안된 영상 분할을 이용한 블록 기반의 스테레오 정합 방법에 대하여 기술한다. 3장에서 제안한 알고리즘을 구현하여 그 결과를 4장에서 제시 및 분석하고 5장에서 본 논문을 끝맺는다.

II. 블록 기반 스테레오 정합

블록 기반 스테레오 정합 방식은 지역 정합 방법 중에서도 영역 기반의 정합 방식이다. 일정 영역인 정합창(window)의 픽셀 정보들을 이용하여 정합점을 찾는다. 정합창의 크기는 매우 중요한데, 창 크기가 너무 작으면 잡음에 민감하지만 매칭의 정확도가 높아지는 경향이 있고, 정합창의 크기가 너무 크면 일중

의 저역 필터(low-pass filter)의 역할을 하게되어 깊이 불연속점을 쉽게 놓치게 된다. 정합창의 크기가 커질수록 변이 지도는 세밀하지만 연산 속도가 증가하고, 정합창의 크기가 작아질수록 변이 지도는 세밀하지 못하지만 연산 속도가 감소하므로 적당한 정합창을 설정하는 것이 중요하다.

그림 1은 기본적인 블록 기반 스테레오 정합 알고리즘을 나타낸다. 영상의 각 라인에 대하여 변이 탐색 범위(disparity search range)내에서 정합 비용(matching cost)을 계산한 후, 변이 지도를 생성한다. 변이 탐색 범위는 변이 값이 최대가 될 수 있는 값에서부터 변이 값이 최소가 될 수 있는 값까지의 범위에 해당한다. 스테레오 정합이 이루어지는 스테레오 카메라 환경에서는, 에피폴라 기하학(epipolar geometry)에 의하여 최대 변이 값이 될 수 있는 지점과, 최소 변이 값이 될 수 있는 지점까지의 범위가 있는데 그 범위만큼 이동하며 기준 영상과 목표 영상을 정합창을 이용해 비교한다. 정합 비용(matching cost)은 기준 영상과 목표 영상 간의 변이 탐색 범위 내 정합창을 이용하여 비교를 한 후 계산하여 나온 비교 결과값으로 정합 비용을 정한다. 계산은 SAD(Sum of Absolute Difference) 방법을 사용하는데, SAD는 기준 영상의 한 점에 대한 정합창 크기의 픽셀값과 목표 영상의 변이 탐색 범위 안의 한 점에 대한 정합창 크기의 픽셀값 사이의 차이값에 절대값을 취한 후, 합산한 것으로서 이 결과값이 정합 비용(matching cost)이 된다.

계산하여 나온 정합 비용들 중에서 가장 작은 값을 그 픽셀의 변이로 선택한다. 그렇게 각 픽셀당 계산된

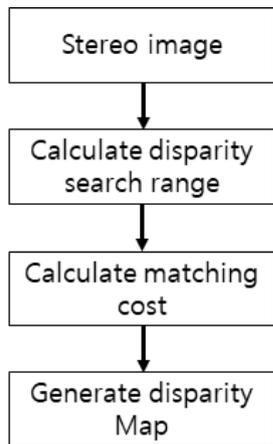


그림 1. 기존 방식의 블록 기반 스테레오 정합 알고리즘
Fig. 1. Conventional block-based stereo matching algorithm

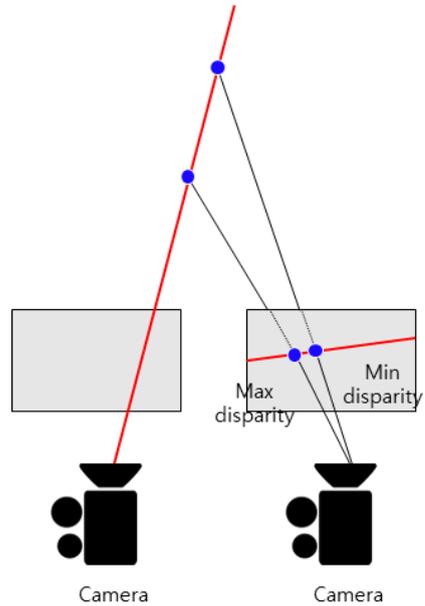


그림 2. 정합을 하기위한 탐색범위 측정
Fig. 2. Estimation of search range for stereo matching

변이값으로 변이 지도를 그린다.

III. 제안하는 스테레오 정합 기법

스테레오 정합에서는 여러가지 방법 중에서 지역 기반 정합 방법으로 알고리즘이 단순하고 빠른 장점을 가진 블록 기반 스테레오 정합 방법을 많이 적용하고 있다. 다른 정합 방법들과 비교했을 때 전역적 방법에 비해 텍스처가 많은 지역에서는 비교적 정확한 시차를 얻을 수 있으나 텍스처가 적은 지역에서는 부정확한 변이를 만들어내고, 변이가 불연속한 경계(boundary) 근처에서는 변이가 뭉개지며(blur) 폐색(occlusion) 영역에서는 변이값이 부정확하게 된다.^[9] 본 논문에서는 부정확한 변이를 만들어 내는 블록 기반 스테레오 정합 방법에서 정확도가 개선된 알고리즘을 제안하였다. 이전에도 신뢰성을 검증하기 위한 방법은 제안되었으나, 본 논문에서는 보다 효율성을 높이기 위하여 영상의 특징 정보를 이용하여 더 정확한 변이를 얻을 수 있도록 하였다.^[10] 그림 3에서는 제안하는 스테레오 정합 과정의 전체 구조를 나타낸다. 영상은 일반적으로 모서리(corner), 경계(edge) 등에서 특징 점이 존재하며, 이러한 특징 점을 추출하기 위한 방법은 다양하다. 본 논문에서는 Mean-Shift 알고리즘을 이용하여 영상을 분할하고 경계 정보를 획득하며 군집화(clustering)로부터 나온 영상을 이용하

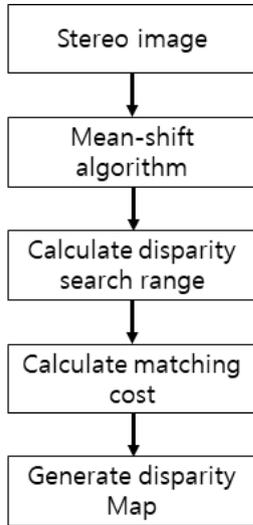


그림 3. 제안하는 스테레오 정합 알고리즘
Fig. 3. Proposed algorithm for stereo matching

여 매칭 알고리즘을 순서대로 진행한다. Mean-Shift 알고리즘은 탐욕 알고리즘(Greedy algorithm)에 기반한 알고리즘으로써, 어느 공간상에 존재하는 점들의 밀도가 높은 곳에서 군집화가 이뤄지도록 하는 방법이다¹³⁾. 즉, 점들의 경사도(gradient)가 큰 쪽으로 이동을 해 가면서 국부최대가 되는 지점에서 군집화를 이루게 되는 것이다. Mean-Shift의 탐색 반경(search radius)의 설정은 mean-shift 알고리즘에서 가장 중요한 요소중 하나이다. 탐색 반경이 너무 크면 정확한 최대(peak) 위치 추정에 오류가 발생할 수 있고, 너무 작은 경우 탐색 결과가 전역 극댓값(global maximum)이 아닌, 국부 극댓값(local maximum)으로 수렴하기 쉬워진다. 식 (1)에서는 mean-shift 알고리즘을 활용하여 밀집도가 가장 높은 지역, 즉 업데이트 되는 새로운 평균 위치를 계산하는 방법을 나타낸다.

$$m_h(x) = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{x-x_i}{h}\right\|^2\right)} - x \quad (1)$$

$m_h(x)$ 는 이동한 위치 x (여기서 x 는 탐색 윈도우의 중심점이기도 하다)에서의 새로운 평균값을 나타낸다. n 은 탐색 반경 내에서 존재하는 점들의 개수이며, x_1 부터 x_n 으로 표현된다. h 는 탐색 윈도우의 반경이고 $g(\cdot)$ 는 커널함수(kernel function)을 나타



그림 4. (a) 원영상 (b) Mean-Shift Algorithm을 적용한 결과 영상
Fig. 4. (a) image (b) Result image using Mean-Shift Algorithm

낸다. 식 (1)에 따르면 $m_h(x)$ 는 항상 밀도가 최대로 증가하는 방향을 가리키고 있으며 국부 최대 밀도에 수렴하도록 정의되어 있다. 그림 4(a)의 영상에 Mean-Shift 알고리즘을 적용한 결과가 그림 4(b)에 제시되어 있으며 영상의 경계는 보존하면서 공간 영역에서의 화소값을 균일하게 하여 군집화를 정확하게 하는 결과를 제공한다. K-means 군집화(clustering) 알고리즘은 빠른 분류 성능을 나타내지만 클러스터의 초기 값에 따라 사용자가 무작위로 지정하는 클러스터 수와 분류 성능에 크게 달라지는 단점이 있다. 이에 비해, Mean-Shift 알고리즘은 클러스터 수를 지정할 필요가 없으며 클러스터의 초기 값을 고려하지 않아도 된다는 장점이 있다. 또한, 각 화소 색상 차이의 최대값을 반복적으로 계산해나가는 알고리즘을 사용하므로 영상에 포함된 객체 영역의 경계를 효율적으로 찾아낼 수 있다. 제시된 그림 4(b)에서처럼, 영상의 경계는 보존하면서 공간 영역에서의 화소값을 균일하게 하여 군집화를 정확하게 하는 결과를 제공한다.

Mean-Shift 알고리즘을 적용한 영상에서 정합점을 찾는데 있어서 정합창의 크기를 결정하는 절차는 정합 결과에 영향을 준다. 주로 3×3 , 5×5 , 7×7 , 9×9 의 사이즈를 가지는 창을 사용하며 창 크기가 너무 작으면 잡음에 민감하지만 매칭의 결과 정확도는 높아지는 경향이 있는 반면, 정합창이 너무 크면 저역 필터(low-pass filter)의 역할을 하게 되어 깊이 불연속 점 등을 쉽게 놓치는 오류를 범할 수 있다. 정합창의 크기가 커질수록 변이 지도는 세밀하지만 연산 속도가 증가하고, 정합창의 크기가 작아질수록 변이 지도는 세밀하지 않지만 연산 속도가 감소하므로 정합창 크기의 최적화 절차가 필요하다. 본 연구에서는 3×3 과 9×9 의 정합창을 사용하였다 (그림 5 참조).

정합창에 존재하는 영상들 사이의 비유사도(dissimilarity) 값을 계산하기 위해 SAD, SSD(Sum

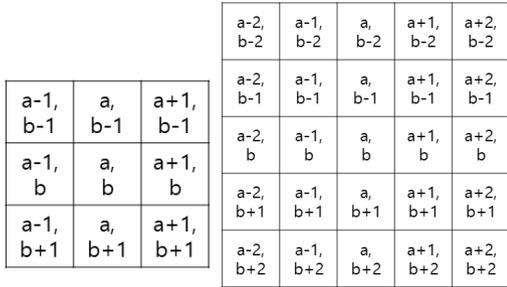


그림 5. 3×3, 9×9 정합창
Fig. 5. 3×3, 9×9 Matching window

of Squared Difference), Census transform, Rank transform 등을 사용하는데, 본 논문에서는 SAD방법을 사용한다. 제한한 정합창을 사용한 SAD의 비용 함수 (C_{SAD}) 는 식 (2)와 같이 표현된다.

$$C_{SAD}(x_l, x_r) = \sum_{m=-W}^W \sum_{n=-W}^W |I_R(x_l + m, s + n) - I_R(x_r + m, s + n)|, (x_l, s) \in R \quad (2)$$

x_l 와 x_r 은 각각 좌측과 우측 정합창 내에 존재하는 영상의 화소 위치의 x 좌표 (또는 수평 위치 좌표)를 나타낸다. W 는 $W \times W$ 사이즈의 정합창의 크기를 나타낸다. I_R 은 분할된 이미지 영역을 나타내고, s 는 수직위치 좌표를 나타내며, 정합절차를 수행하기 전에 수직 위치 좌표는 동일하다. SAD 방식은 기준 영상에서의 한 지점을 지정한 후 목표 영상의 변이 탐색 범위(disparity search range) 내 존재하는 모든 윈도우에 대해 이동(shift)하며 모두 비교한다. 변이 탐색 범위는 여러 번 실험을 거듭한 결과 가장 최적의 결과가 나오는 100으로 지정하여 실험을 진행하였다. SAD 방식으로 계산하는 방법은, 기준 영상의 한 점에 대한 정합창 크기의 픽셀값과 목표 영상의 변이 탐색 범위 안의 한 점에 대한 정합창 크기의 픽셀값을 각각의 차이에 절대값을 취한 후, 합산한 것이 정합 비용(Matching cost)이 된다. 계산하여 나온 정합 비용들 중에서 가장 작은 값을 그 픽셀의 변이로 선택한다. 그림 6에서 좌우 영상 간의 정합을 수행하면서 계산되는 정합비용을 제시하였다. 왼쪽 위 영상을 기준 영상으로 보았을 때 한 픽셀을 지정한다. 지정된 변이 탐색 범위 만큼 이동하며 모든 정합 비용을 계산한 후 가장 비유사도가 낮은, 즉, 가장 유사한 픽셀을 변이로 정한다.

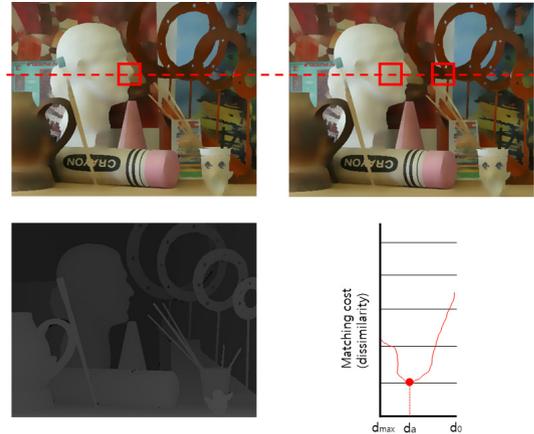


그림 6. (위) 기준 영상과 목표 영상의 스캔라인과 변이 탐색 범위, (아래) 변이 지도와 정합 비용 추정
Fig. 6. (Top) The scanline and disparity search range of the reference and target images, (Bottom) Disparity Map and matching cost cube

IV. 실험

스테레오 정합의 성능을 평가하기 위하여 본 실험에 사용된 데이터는 Middlebury stereo vision (<http://vision.middlebury.edu/stereo/>)의 8개의 실험 영상을 사용하였다^[14]. 실험에 사용된 파라미터 값들은 모든 영상에 공통적으로 적용하여 객관적인 결과가 도출되도록 하였다. 각 영상의 특성은 표 1과 같다.

표 2는 기존의 방식의 블록 기반 스테레오 정합을 하였을 때와 새롭게 제시한 스테레오 정합 기법을 사용하였을 때의 평균 제곱근 편차(Mean-square error)이다. 평균 제곱근 편차는 추정값 또는 모델이 예측한 값과 실제 환경에서 관찰되는 값의 차이를 다룰 때 사용되는 척도로써, 본 실험에서는 Middlebury dataset에서 제공되는 변이 지도와 비교하였다. 실험 영상에 대해서 3×3 정합창을 사용하였을 때와, 9×9 정합창

표 1. 실험 영상의 특성
Table 1. Characteristics of test images

Test Image	Size	Disparity Range
Cones	450x375	0-59
Teddy	450x375	0-59
Art	463x370	0-59
Books	463x370	0-59
Dolls	463x370	0-59
Laundry	447x370	0-59
Moebius	463x370	0-59
Reindeer	447x370	0-59

표 2. 변이 지도들의 평균 제공근 편차
Table 2. MSE of the disparity maps

Algorithm	Cones				Teddy				Art			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	11516	11259	11770	11000	7832	7442	8082	7573	10133	8812	10345	8921
Proposed	11259	10501	11571	10830	7748	7395	8077	7500	9945	8647	10232	8885
Improvement(%)	2.20	6.73	1.69	1.55	1.07	0.63	0.01	0.96	1.89	1.90	1.10	0.40

Algorithm	Books				Dolls				Laundry			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	9617	9871	9641	9987	10314	10097	10359	10287	8546	7114	8645	7213
Proposed	9370	9622	9310	9667	10207	9977	10350	10179	8322	6957	8570	7146
Improvement(%)	2.63	2.58	3.55	3.31	1.04	1.20	0.08	1.06	2.69	2.25	0.87	0.93

Algorithm	Moebius				Reindeer			
	3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right
Block-based Stereo Matching	6611	7106	6775	7233	8702	10953	8612	12204
Proposed	6430	6984	6766	7232	8335	10451	8544	12116
Improvement(%)	2.81	1.74	0.13	0.01	4.40	4.80	0.79	0.72

을 사용했을 때로 나누어 평균 제공근 편차 결과를 제시하였다. 표 2에서의 결과로부터 평균 제공근 편차를 비교하면, 전체적으로 평균 제공근 편차, 즉 정합오류값이 줄어든 것을 확인할 수 있다. 기존의 블록 기반의 스테레오 정합 기법은 속도는 빠르지만 비교적 부정확한 단점을 가지고 있다. 제안하는 기법은 블록 기반의 스테레오 정합 기법이 가진 장점을 살리면서 간단한 영상 분할을 통하여 기존의 블록 기반의 스테레오 정합 알고리즘들보다 우수한 결과를 보여준다. 현재 가장 우수하게 평가 받는 신뢰확산(Belief Propagation)에 기반한 방법들에 비해 유사하거나 영상의 특성에 따라 낮은 성능을 나타내는 경우도 있지만, 이러한 방법들은 복잡한 색 영역 분할 방법과 MRF(markov random fields)에 기반한 확률적 접근을 통해 정교한 변이를 획득한다. 또한, 이러한 방법들은 결과가 정교한 반면, 변이를 획득하는 과정의 복잡도가 높다는 단점이 있다. 이러한 신뢰확산에 기반한 방법들보다는 성능이 낮지만 복잡도가 낮은 장점이 있고, 기존의 블록 기반의 스테레오 정합 방법보다도 성능이 우수하다는 장점이 있다.

표 3에서는 실험 영상에 대해서 3×3 정합창을 사용하였을 때와, 9×9 정합창을 사용했을 때로 나누어 실행속도 결과를 제시하였다. 실험을 진행한 PC의 사양으로 CPU는 Intel Core i5 (7세대) 2.3 GHz 그리고 8GB RAM을 사용하였으며, 운영체제는 macOS Mojave를 사용하였다. 프로그래밍 언어로는 python을 사용하였으며 opencv 라이브러리를 함께 사용하였다. 표 2, 3의 결과를 종합해보면 정합오류값이 비슷하거나 줄어들었음에도 실행 속도는 큰 차이가 없다는 것을 알 수 있다. 표 4에서는 실험 영상에 대해서 3×3 정합창을 사용하였을 때와, 9×9 정합창을 사용했을 때로 나누어 불연속 영역들의 평균 제공근 편차 결과를 제시하였다. 표 4에서의 결과로부터 불연속 영역들의 정합오류값은 큰 차이가 없고 texture 영역에서의 정합 정확도가 상승하였음을 알 수 있다. 표 4에서는 실험 영상에 대해서 3×3 정합창을 사용하였을 때와, 9×9 정합창을 사용했을 때로 나누어 불연속 영역들의 평균 제공근 편차 결과를 제시하였다. 표 4에서의 결과로부터 불연속 영역들의 정합오류값은 큰 차이가 없고 texture 영역에서의 정합 정확도가 상

승하였음을 알 수 있다.

표 3. 실행속도 비교
Table 3. Execution speed comparison

Algorithm	Cones				Teddy				Art			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	17.16	15.97	109.27	108.54	14.61	14.16	95.18	94.59	14.78	14.53	97.05	96.70
Proposed	14.76	14.33	95.98	95.46	14.78	14.53	96.57	96.30	15.41	14.96	98.98	99.42

Algorithm	Books				Dolls				Laundry			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	15.61	15.59	103.63	102.09	15.02	14.48	96.72	96.54	15.52	15.41	105.31	104.93
Proposed	15.42	15.39	103.61	101.50	15.06	14.66	97.20	97.11	14.18	14.04	92.77	92.41

Algorithm	Moebius				Reindeer			
	3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right
Block-based Stereo Matching	15.09	14.60	108.73	97.43	14.53	14.13	94.11	93.60
Proposed	14.76	14.46	96.62	96.34	14.33	14.28	94.41	95.61

표 4. 불연속 영역들의 평균 제곱근 편차
Table 4. MSE of the discontinuous region

Algorithm	Cones				Teddy				Art			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	4376	4278	4472	4180	3289	3125	3394	3180	3749	3260	4138	3568
Proposed	4372	4276	4471	4172	3291	3122	3402	3182	3739	3248	4150	3577

Algorithm	Books				Dolls				Laundry			
	3×3		9×9		3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right	left	right	left	right
Block-based Stereo Matching	2692	2763	2699	2796	4084	3998	4102	4073	1880	1565	1901	1586
Proposed	2708	2779	2704	2812	4066	3993	4098	4059	1881	1567	1906	1593

Algorithm	Moebius				Reindeer			
	3×3		9×9		3×3		9×9	
	left	right	left	right	left	right	left	right
Block-based Stereo Matching	1943	2089	1991	2126	2810	2724	2928	4149
Proposed	1987	2134	2022	2151	2798	2712	2922	4146

V. 결 론

본 논문에서는 영상 분할을 이용하여 부정확한 경계 부분을 더욱 부각되도록 하였다. Mean-Shift 알고리즘을 적용한 균일화된 영상을 분할하여 경계정보를 획득하고, 변이 탐색 범위를 계산한 뒤, 정합창을 이용하여 변이 탐색 범위 안의 SAD 계산을 통해 정합 비용을 계산한 후 비유사도에 따라서 가장 최적의 변이를 할당하였다. Mean-Shift 알고리즘을 사용함으로써 기존의 비교적 부정확한 정확도를 향상시킨 결과를 얻게 되었고, 기존의 장점은 그대로 가질 수 있었다. 앞으로 이런 복합적인 기법에 대한 연구를 통해 좀 더 정확하면서 빠른 결과를 얻어낼 수 있는 방법에 대한 연구를 진행할 계획이다.

References

[1] J. Xu, Q. Yang, and Z. Feng, "Occlusion-aware stereo matching," *Int. J. Computer Vision*, vol. 120, no. 3, pp. 256-271, 2016.

[2] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Trans. Magnetics, Pattern Anal. and Mach. Intell.*, vol. 25, no. 8, pp. 993-1008, Aug. 2003.

[3] Y. S. Seok and J. S. Yoo, "Dynamic programming-based stereo matching using image segmentation," *J. KICS*, vol. 35, no. 8, pp. 680-688, Jan. 2010.

[4] B. G. Kim and J. K. Eem "Stereo matching using belief propagation with line grouping," *The Inst. Electron. Eng. Korea*, vol. 42, no. 3, pp. 279-283, May 2005.

[5] P. H. S. Torr and A. Criminisi, *Dense stereo using pivoted dynamic programming*, Microsoft Res., MSR-TR-2002-51, 2002.

[6] I. J. Cox, S. L. Hingorani, S. B. Rao, and B. M. Maggs, "A maximum likelihood stereo algorithm," *Computer Vision and Image Understanding*, vol. 63, no. 3, pp. 542-567, May 1996.

[7] C. S. Kang, J. D. Kim, S. W. Lee, and K. G. Nam, "Stereo matching using dynamic programming with region partition," *The Inst. Electron. Eng. Korea*, vol. 20, no. 1, pp. 479-

482, Jul. 1997.

[8] Y. H. Lee, Y. S. Kim, and J. H. Park, "Comparative analysis of the performance of SIFT and SURF," *The Korean Soc. Semiconductor & Display Technol.*, vol. 12, no. 3, pp. 59-64, 2013.

[9] W. P. Dong and C. S. Jeong, "Stereo matching using dynamic programming and image segments," *National Digital Sci. Library*, vol. 32, no. 1, pp. 805-807, 2005.

[10] J. W. An and J. S. Yoo, "Stereo matching algorithm by using segmentation," in *Proc. Annu. Fall Conf. The Korean Inst. Broadcast and Media Eng.*, pp. 303-306, Seoul, South Korea, Nov. 2011.

[11] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. Seventh IEEE Int. Conf. Computer Vision*, pp. 1-8, Kerkyra, Greece, Sep. 1999.

[12] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 345-359, 2008.

[13] K. M. Lee and C. H. Lin, "An image segmentation and merge hierarchical region using mean-shift tracking algorithm," in *Proc. Annu. Conf. IEIE*, pp. 704-706, Incheon, South Korea, 2017.

[14] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Computer Vision*, vol. 47, no. 1, pp. 7-42, 2002.

김 지 수 (Ji-su Kim)



2019년 2월 : 계명대학교 컴퓨
터공학과 (공학사)
2019년 3월~현재 : 계명대학교
컴퓨터공학과 석사과정
<관심분야> 컴퓨터 비전, 영상
처리, 머신러닝

[ORCID:0000-0002-3836-2571]

박 철 형 (Cheol-hyeong Park)



2014년 3월~현재 : 계명대학교
컴퓨터공학과 학/석사 연계
과정
<관심분야> 컴퓨터 비전, 영상
및 신호처리, 머신러닝, 컴
퓨터그래픽스

[ORCID:0000-0002-2337-0018]

이 덕 우 (Deok-woo Lee)



2007년 2월 : 경북대학교 전자
전기 컴퓨터학부 (공학사)
2008년 12월 : North Carolina
State University (공학석사)
2012년 12월 : North Carolina
State University (공학박사)
2013년 1월~2013년 4월 :
Washington University in St.Louis, 박사후 연구
원
2013년 9월~2016년 8월 : 삼성전자 무선사업부 책임
연구원
2016년 9월~2018년 2월 : 영산대학교 가상현실콘텐
츠학과 조교수
2018년 3월~현재 : 계명대학교 컴퓨터공학전공 조교
수
<관심분야> 컴퓨터 비전, 영상 및 신호처리

[ORCID:0000-0002-0080-6886]