

조명 변화에 강인한 상호 정보량 기반 스테레오 정합 기법

허용석[°]

An Illumination-Insensitive Stereo Matching Scheme Based on Weighted Mutual Information

Yong Seok Heo[°]

요약

본 논문에서는 조명 변화에 강인한 상호 정보량 기반의 스테레오 정합 기법을 제안한다. 일반적으로 다양한 조명 조건에서 취득한 스테레오 영상은 좌우 영상 간에 컬러의 변화가 발생하기 때문에 정확한 스테레오 정합점을 찾는 것이 쉽지 않다. 이 경우, 컬러를 보정하는 작업을 우선 수행하는 것이 일반적이다. 그러나, 좌우 스테레오 영상에 대해서는 컬러 값을 동일하게 보정하는 작업도 시점 차이로 인한 좌우 영상의 변화로 인해서 좌우 영상에 대한 정합 정보가 요구되므로 쉽지 않다. 본 논문에서는 다양한 조명 조건에서 취득한 영상에 강인한 스테레오 정합 기법을 제안한다. 이를 위해서 선형적인 관계를 갖는 로그색도(log-chromaticity) 컬러 공간으로 변형을 수행하였고, 이 컬러 공간에서 상호 정보량에 기반한 새로운 스테레오 정합 비용(cost)을 제안하였다. 제안하는 비용은 가중치가 적용된 상호 정보량과 SIFT(Scale Invariant Feature Transform) 묘사 벡터의 정보를 화소(pixel)마다 적응적으로 결합한다. 또한, 보다 정확한 변위 지도 예측을 위해서 세그먼트 기반의 평면 제한 조건도 제안하는 비용에 포함되었다. 다양한 실험 데이터에 대해서 테스트한 결과, 제안하는 방법이 기존의 방법들에 비해서 보다 정확한 변위 지도 결과를 얻는 것을 확인하였다.

Key Words : Stereo Matching, Mutual Information, Illumination Variation, Camera Variation

ABSTRACT

In this paper, we propose a method which infers an accurate disparity map for radiometrically varying stereo images. For this end, firstly, we transform the input color images to the log-chromaticity color space from which a linear relationship can be established during constructing a joint pdf between input stereo images. Based on this linear property, we present a new stereo matching cost by combining weighted mutual information and the SIFT(Scale Invariant Feature Transform) descriptor with segment-based plane-fitting constraints to robustly find correspondences for stereo image pairs which undergo radiometric variations. Experimental results show that our method outperforms previous methods and produces accurate disparity maps even for stereo images with severe radiometric differences.

* 본 연구는 2015년도 아주대학교 정착연구비 지원에 의하여 연구되었음.

[°] First and Corresponding Author : Department of Electrical and Computer Engineering, Ajou University, ysheo@ajou.ac.kr, 정회원
논문번호 : KICS2015-07-243, Received July 31, 2015; Revised October 4, 2015; Accepted November 2, 2015

I. 서 론

스테레오 정합 (stereo matching)은 두 장 이상의, 시점 (view-point)이 다른 영상들로부터 정합점 (correspondence)을 계산하여, 3차원 정보를 포함하고 있는 변위 (disparity)를 추출하는 것을 목표로 한다^[32]. 이때, 대부분의 알고리즘들이 영상의 컬러 (color) 또는 밝기 (intensity) 정보의 값을 정합점을 찾기 위한 정보로 활용하고 있다^[4,6,7,32]. 일반적으로 이러한 컬러 정보는 다양한 요인들이 결합되어 나타난다^[9,38]. 즉, 영상의 컬러 값은 조명 조건, 물체의 반사도 (reflectance), 그리고 촬영 시의 카메라 세팅 조건 등의 영향을 받는다^[2].

이러한 요인들은 영상에 변화를 일으키는 범위에 따라서 전역적인 요인과 지역적인 요인으로 구분할 수 있다^[17]. 전역적인 요인으로는 카메라의 노출 시간, 카메라 증폭 (gain), 감마 보정 (gamma correction) 등과 같은 요인이 있다. 한편, 지역적인 요인으로는 광원 (source of light)의 위치, 비네팅 현상 (vignetting effect) 등이 있다. 따라서, 정합점을 요구하는 대부분의 컴퓨터 비전 알고리즘에서는 이러한 영상의 변화 요인들에 강인한 알고리즘을 설계하는 것이 중요한 이슈 중의 하나이다^[1-4,17].

대부분의 스테레오 정합 알고리즘들은 입력 영상들이 다양한 조명 조건에 대해서 컬러 값이 보정 (radiometrically calibrated)이 되어 있다고 가정한다. 이렇게 보정된 조건에서는 정합되는 영상의 밝기 값들이 정확히 같거나 매우 비슷한 값을 가지고 있기 때문에, 화소의 절대 값 차이 (absolute difference)와 같은 단순한 측정 비용 (cost)을 사용해서도 매우 정확한 결과를 얻을 수 있다. 그러나, 위에서 언급한 바와 같이 보다 일반적인 상황에서는 좀 더 강인한 정합 알고리즘을 요구하고 있다. 예를 들면, 항공 영상을 이용하여 3차원 복원을 하는 경우와 다양한 다시점 영상을 활용하여 3차원 복원을 하는 경우가 있다. 항공 영상은 영상 촬영 시에 넓은 지역을 일정한 시간 간격을 두고 촬영을 수행하고 있으므로, 시간차에 의한 조명 조건의 변화에 따른 영상의 변화가 발생하게 된다. 또한, 최근 인터넷에 있는 영상들을 이용하여 3차원 복원을 하는 연구가 발표되었다^[5,33]. 이 경우는 다양한 조명과 카메라로 취득한 영상들을 이용하기 때문에 영상들 사이에 불가피하게 위에서 언급한 다양한 전역적인 또는 지역적인 변화가 발생하게 되어, 임의의 영상들 사이에 정확한 정합점을 찾는 작업은 매우 어려운 일이라고 할 수 있다. 따라서, 이러한 변화 요인

들이 있는 경우, 보다 강인한 스테레오 정합 알고리즘이 요구되고, 최근 이를 해결하기 위한 몇몇 알고리즘들이 제안되었다^[13-15,28,34,36].

일반적으로 정합하려는 이미지들 사이에 조명 및 카메라 노출 변화에 대해서는 NCC (Normalized Cross-Correlation), MI (Mutual Information)^[16,22], Census^[41]와 같은 정합 비용들이 널리 쓰이고 있다^[17]. NCC는 카메라 노출 변화와 같은 전역적인 밝기 변화가 있는 영상들의 정합에 가장 널리 쓰이는 기법 중 하나이다. 이는 정합하려는 각 화소를 중심으로 특정 윈도우 범위 안에 있는 밝기 값들을 하나의 벡터로 표현했을 때, 정합하려는 두 벡터 사이의 코사인 (cosine) 각도를 측정하는 방법이다. 따라서 NCC는 비교하려는 두 벡터가 선형적인 변화 (affine transformation)가 있다고 가정할 때, SAD (Sum of Absolute Difference)나 SSD (Sum of Squared Difference)와 같이 밝기의 절대적인 차이 값을 이용하는 다른 윈도우 정합 기법들에 비해 매우 효과적인 방법이다. 그러나, NCC는 기존의 다른 윈도우 정합 방법들과 마찬가지로 정합에 사용되는 윈도우 안에서 서로 다른 깊이 정보를 갖는 화소들이 존재할 때, 정합 오류가 발생할 수 있다는 단점이 있고, 윈도우 크기를 증가시킬수록 윈도우의 경계부근에서 이러한 오류가 증가하게 된다. 또한, 밝기 변화가 선형적으로 변화한다는 가정에서 벗어나는, 보다 일반적인 경우에도 오류가 많이 발생하게 된다. 이러한 단점을 극복하기 위해 ANCC (Adaptive Normalized Cross-Correlation) 기법이 제안되었다^[13]. ANCC는 윈도우 크기의 증가에도 윈도우의 경계 지역에서 정확한 값을 예측할 뿐만 아니라, 밝기 값이 비선형적으로 발생하는 경우도 비교적 정확한 값을 예측하지만, 윈도우 크기가 커짐에 따라서 계산량이 급격히 증가하고, 카메라의 극심한 노출 변화에는 여전히 오류가 발생하기 쉬운 단점이 있다.

한편, 영상들 사이에 선형적인 변화를 포함한 임의의 전역적인 변화가 발생할 때에는 MI에 기반한 기법이 일반적으로 널리 쓰이고 있다. 이는 특히 MR (Magnetic Resonance)과 CT (Computed Tomography)와 같은 서로 다른 특성을 갖는 영상들을 정합할 때 많이 쓰이는 기법 중 하나이다^[29,30,31]. 스테레오 정합을 위해서는 Egnal^[11]이 지역적인 윈도우에 대해서 MI를 사용하는 기법을 제안하였다. 그러나 이 기법도 텍스처 (texture)가 없는 지역에서는 제한된 정보량으로 인하여 기존의 윈도우 정합 기법이 갖는 단점들을 그대로 가지고 있다. 이러한 단점을 극

복하기 위해 Kim et al.^[22]은 화소에 기반한 MI 기법을 전역적인 에너지 최소화 구조에 융합하는 알고리즘을 제안하였다. 이러한 기법을 더욱 개선하여 Hirschmuller^[16]는 가려지는 (occluded) 지역의 정보를 고려하여 보다 정확한 화소 기반의 MI 값을 예측하는 기법을 제안하였다. Zitnick et al.^[42]은 세그먼트(segment) 기반의 근사적인 MI 기법을 제안하여 노이즈에 강인한 스테레오 정합 기법을 제안하였다. 이러한 기존의 MI에 기반한 기법들은 임의의 전역적인 변화 함수를 사전에 모른다고 하더라도 동시 확률 분포(joint probability)의 측정을 통해서 전역적인 변화 함수를 예측하여 정확하게 정합을 수행할 수 있다는 장점이 있다. 그러나, MI 기반의 방법들은 조명의 방향 변화 등으로 인한 지역적인 변화에 대해서는 전역적인 함수를 예측하는 것이 불가능하다는 단점이 있다^[17].

또한, 조명 변화에 널리 쓰이는 정합 방법 중의 하나로 Census^[41] 기반의 방법이 있다. 이는 정합하려는 화소 주변의 윈도우 안에 있는 화소들에 대해서 중심 화소와의 밝기 값의 크기를 비교하여 크면 0, 작으면 1의 이진수로 변화하여 얻은 값들을 벡터(vector)화 하여, 정합하려는 두 벡터 사이의 해밍 거리(Hamming distance)를 이용하여 유사도를 측정하는 방법이다. 그 밖에도, Montesinos et al.^[27]은 밝기의 일차 미분 값을 이용하여 특징 점에 대해서만 강인하게 정합하는 방법을 제안하였다. Ogale와 Aloimonos^[28]는 영상으로부터 얻은 여러 대역의 주파수(frequency) 값들을 이용하여 지역적인 정합을 수행하는 방법을 제안하였다. Hirschmuller와 Scharstein^[17]은 위에서 언급한 대부분의 정합 비용들을 여러 가지 환경에서 다양한 실험을 통하여 성능 평가를 수행하였다. Hirschmuller와 Scharstein^[17]은 대부분의 알고리즘들이 전역적인 환경에서는 비교적 잘 동작하지만, 조명 변화 등으로 인한 지역적인 변화에 대해서는 잘 동작하지 않음을 확인하였다. 이는 이러한 대부분의 기존 기법들은 단순히 영상으로부터 관찰된 밝기 값을 그대로 입력 값으로 쓰고 있기 때문에, 실제 환경에서 일어날 수 있는 다양한 조명 및 카메라 조건에 의해서 발생하는 변화를 다루지 못한다는 점도 오류의 원인중 하나로 분석되고 있다^[13].

따라서 본 논문에서는 이러한 다양한 조명 변화에 강인한 스테레오 정합 알고리즘을 제안한다^[14,15]. 본 논문에서는 다양한 조명 변화 환경에서 변위 지도를 효과적으로 측정하도록 최적화된 비용 함수들을 결합하여 최적화 과정(optimization)을 수행하였다. 그림

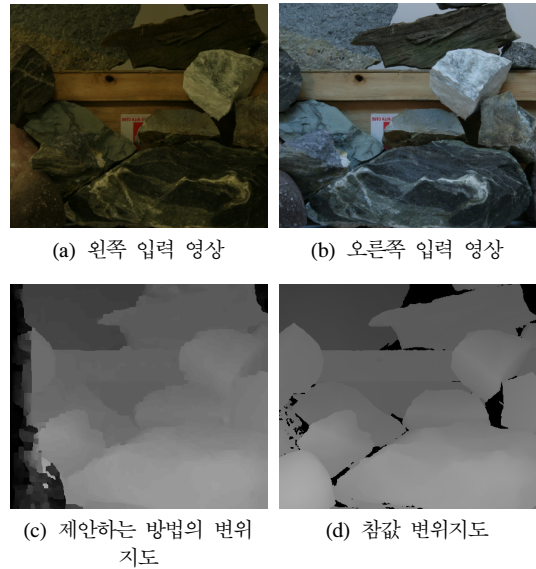


그림 1. 제안하는 방법의 결과 예시
Fig. 1. An example of the result of the proposed method

1은 제안하는 방법을 사용한 결과를 보여주고 있다. 그림 1(a)와 그림 1(b)는 각각 서로 다른 조명 및 카메라 조건에서 취득한 스테레오 좌우 영상이다. 이때, 제안하는 방법을 사용하여 그림 1(c)와 같은 변위 지도를 얻을 수 있다. 그림 1(d)의 참값과 비교했을 때, 비교적 정확한 결과를 얻는 것을 확인할 수 있다.

본 논문의 구성은 다음과 같다. 2장에서는 제안하는 스테레오 정합 기법에 대해서 기술한다. 2장 1절에서 비선형적인 컬러 영상으로부터 선형성을 보장하는 로그-색도 컬러 공간으로의 변환 과정을 설명하고, 2장 2절에서는 일반적인 상호 정보량 추출 과정과 제안하는 상호 정보량 예측 기법에 대해서 기술한다. 2장 3절에서는 여러 가지 정보가 융합된 스테레오 정합 비용을 제안한다. 3장에서는 다양한 실험 영상에 대해서 제안하는 알고리즘과 기존 기법들과의 비교 실험을 수행한다. 마지막으로, 4장 결론에서는 제안하는 방법의 장단점과 향후 연구의 필요성에 대해서 요약하고 논문을 마무리한다.

II. 가중치가 적용된 상호 정보량에 기반한 스테레오 정합 기법

제안하는 방법은 일반적인 조명 변화에 강인한 스테레오 정합을 통해서 변위 지도를 예측하기 위해서 다양한 과정을 거친다^[13,14,15]. 그림 2는 제안하는 방법의 개요도를 나타내고 있다. 각 모듈은 다음의 각 절

에서 자세히 설명하기로 한다.

2.1 로그-색도 (log-chromaticity) 컬러 공간

제안하는 방법에서는 보다 일반적인 상황에서 다양한 조명 변화에 강한 정합 비용을 정의하기 위해서 물체 표면이 난반사 표면 (diffuse reflection)이고, 영상의 컬러 변화가 아래와 같은 비선형(non-linear) 조건식을 만족한다고 가정한다^[13-15,24].

$$I_{L,k}(p) = \sigma_L(p)a_{L,k}(I_{L,k}^o(p))^{\gamma_L}, k \in \{R, G, B\}, \quad (1)$$

여기서, 식 (1)은 왼쪽 영상 I_L 에 대한 비선형 조건식을 나타낸다. 오른쪽 영상 I_R 에 대한 비선형 조건식도 식 (1)과 유사하게 표현할 수 있다. 식 (1)에서 $\sigma_L(p)$ 는 I_L 에서 화소 p 의 밝기를 결정하는 요인에 대한 함수이고, $a_{L,k}$ 는 조명과 관련된 요인, $I_{L,k}^o(p)$ 는 왼쪽 영상의 k 채널에 대해서 화소 p 의 영상 고유 밝기 값을 의미한다. 그리고, 카메라 내에서 감마 보정 요인을 변수 γ_L 에 의해서 결정된다고 가정한다. 일반적으로 영상이 주어지면, 위의 각 요인들을 모두 예측하는 것은 매우 어려운 작업이다. 대부분 이러한 상황에서는 물체 고유의 영상 값을 얻는 과정인 컬러 일반화 (color constancy) 과정^[9,10,20,21,23,27,39,40]이 적용될 수 있다. 그러나, 컬러 일반화 과정은 밝기 값의 구분성 (discriminability)을 잃게 만들 수 있으므로, 컬러 일반화 특성과 영상 구분성은 서로 상충 관계 (trade-off)에 있다고 할 수 있다. 또한, 스테레오 입력 영상들의 경우는 시점 차이로 인해서 정확한 컬러 일

반화를 수행하기는 쉽지 않다. 따라서, 본 논문에서는 밝기 값의 구분성을 최대한 보존하면서, 컬러 일반화가 적용될 수 있도록, 컬러 공간을 적절히 변형하는 방법을 사용한다. 본 논문에 사용하는 컬러 공간을 로그-색도 (log-chromaticity) 컬러 공간이라고 하고, 아래 과정과 같이 유도할 수 있다.

우선, 식 (1)에 로그 (log)를 취하면 $\log I_{L,k}(p) = \log(\sigma_L(p)) + \log(a_{L,k}) + \gamma_L \log(I_{L,k}^o(p))$ 와 같다. 이때, 화소 p 에서 각 컬러 채널의 평균값 $I_L^m(p)$ 은 아래와 같다.

$$I_L^m(p) = \log(\sigma_L(p)) + \frac{\log(a_{L,R}a_{L,G}a_{L,B})}{3} + \frac{\gamma_L \log(I_{L,R}^o(p)I_{L,G}^o(p)I_{L,B}^o(p))}{3}. \quad (2)$$

한편, 식 (1)과 식 (2)로부터 로그-색도 컬러 값 $I'_{L,k}(p)$ 은 아래와 같이 정의할 수 있다.

$$I'_{L,k}(p) = \log(I_{L,k}(p)) - I_L^m(p) = \log \frac{a_{L,k}}{\sqrt[3]{a_{L,R}a_{L,G}a_{L,B}}} + \gamma_L \log \frac{I_{L,k}^o(p)}{\sqrt[3]{I_{L,R}^o(p)I_{L,G}^o(p)I_{L,B}^o(p)}} = A_{L,k} + \gamma_L B_{L,k}(p), \quad (3)$$

여기서, $A_{L,k}$ 는 각 채널에 관련된 상수이고, $B_{L,k}(p)$ 는 원본 컬러 값 $I_{L,k}^o(p)$ 으로 이루어져 있는 채널 k 와 화소 p 에 대한 함수이다. 마찬가지로, 오른쪽 영상 I_R 에 대해서도 식 (3)과 같은 과정을 수행하면, 아래 식과 같이 표현할 수 있다.

$$I'_{R,k}(p + f_p) = A_{R,k} + \gamma_R B_{R,k}(p + f_p), \quad (4)$$

이때, 왼쪽 영상에서 화소 p 의 오른쪽 영상에서의 대응점 위치는 $p + f_p$ 로 표현하였다. 따라서, 정합되는 최우 스테레오 영상 화소들에 대해서는 아래와 같은 식이 성립한다.

$$B_{L,k}(p) = B_{R,k}(p + f_p) = \log \frac{I_{L,k}^o(p)}{\sqrt[3]{I_{L,R}^o(p)I_{L,G}^o(p)I_{L,B}^o(p)}} \equiv B_k, \quad (5)$$

따라서, 대응되는 두 화소 $I'_{L,k}(p)$ 와

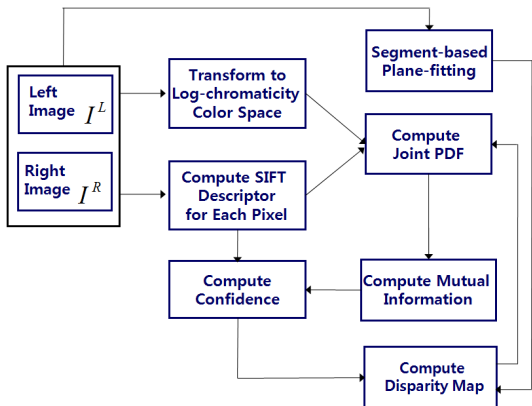


그림 2. 제안하는 알고리즘
Fig. 2. An overview of the proposed method

$I'_{R,k}(p+f_p)$ 사이에는 다음과 같이 선형적인 관계가 성립함을 알 수 있다.

$$\begin{aligned} I'_{R,k}(p+f_p) &= A_{R,k} + \gamma_R B_k, \\ I'_{L,k}(p) &= A_{L,k} + \gamma_L B_k \\ &= A_{L,k} + \gamma_L \left(\frac{I'_{R,k}(p+f_p) - A_{R,k}}{\gamma_R} \right) \\ &= A_{L,k} - \frac{\gamma_L A_{R,k}}{\gamma_R} + \frac{\gamma_L}{\gamma_R} I'_{R,k}(p+f_p) \\ &= M_k I'_{R,k}(p+f_p) + N_k, \end{aligned} \quad (6)$$

이때, M_k 과 N_k 은 각각

$$M_k = \frac{\gamma_L}{\gamma_R}, N_k = A_{L,k} - \frac{\gamma_L}{\gamma_R} A_{R,k}$$

와 같이 표현되는 상수이고, 이는 각 컬러 채널 k 에 대해서 동일한 값을 갖는 것을 알 수 있다.

한편, 다음 절에서 설명할 상호 정보량의 계산을 위해서 $I'_k(p)$ 의 컬러 공간을 양자화 (quantization) 할 필요가 있고, 양자화된 값 $J_k(p)$ 은 아래식과 같이 정의할 수 있다.

$$J_k(p) = \lfloor C I'_k(p) \rfloor, \quad (7)$$

이때 C 는 양자화 상수이고, 본 논문에서는 상수 C 의 값을 실험적으로 결정하였고, 모든 실험에서 $C=500$ 으로 고정하였다.

2.2 가중치가 적용된 상호 정보량 측정

변위 지도 f 를 구하기 위하여 MAP (Maximum A Posteriori)-MRF (Markov Random Field) 구조를 가정하였고, 이때 변위 지도 f 는 다음과 같은 전체 에너지 $E(f)$ 를 최소로 하는 f 로 결정을 한다^[7,35].

$$E(f) = E_d(f) + E_s(f), \quad (8)$$

이때, 전체 데이터 비용(total data cost) $E_d(f)$ 와 전체 평탄화 비용 (total smoothness cost) $E_s(f)$ 는 각각 다음 식과 같이 정의할 수 있다.

$$\begin{aligned} E_d(f) &= \sum_p D_p(f_p), \\ E_s(f) &= \sum_p \sum_{q \in N(p)} V_{pq}(f_p, f_q), \end{aligned} \quad (9)$$

여기서, $N(p)$ 는 화소 p 의 이웃 화소들을 의미하고, $D_p(f_p)$ 와 $V_{pq}(f_p, f_q)$ 는 각각 화소 p 에서의 변위지도 값이 f_p 를 가질 때, 데이터 비용 (data cost)과 평탄화 비용 (smoothness cost)을 의미한다. 한편, 데이터 비용으로 사용하기 위한 상호 정보량 (mutual information)은 아래 식과 같이 정의될 수 있다^[11,16,22,29,30,31,37,42].

$$\begin{aligned} E_d(f) &= -MI(J_L, J_R, f) \\ &= -(H(J_L) + H(J_R) - H_f(J_L, J_R)), \end{aligned} \quad (10)$$

이때, $H(\cdot)$ 과 $H(\cdot, \cdot)$ 는 각각 영상의 엔트로피 (entropy)와 동시 엔트로피 (joint entropy)를 의미하고, 다음 식과 같이 정의할 수 있다.

$$\begin{aligned} H(J) &= - \int P(j) \log(P(j)) dj, \\ H_f(J_L, J_R) &= - \iint P_f(j_L, j_R) \log(P_f(j_L, j_R)) dj_L dj_R, \end{aligned} \quad (11)$$

여기서, $P(j)$ 는 밝기 값 j 의 주변 확률 (marginal probability)를 의미하고, $P_f(j_L, j_R)$ 은 좌우 영상에서 대응되는 두 화소 값 j_L 과 j_R 에 대한 동시 확률 분포 (joint probability)를 의미한다. 이때, 동시 확률 분포를 위해서 본 논문에서는 아래와 같이 정의한다.

$$\begin{aligned} P_f^{SIFT}(j_{L,k}, j_{R,k}) &= \frac{1}{Z} \sum_{p \in J_{L,k}} \rho_k(p, p+f_L(p)) \cdot \\ T[(j_{L,k}, j_{R,k}) &= (J_{L,k}(p), J_{R,k}(p+f_L(p))), \end{aligned} \quad (12)$$

여기서 Z 는 정규화 상수이고, \mathcal{J} 는 식 (7)에서 정의한 로그-색도 컬러 영상을 의미하고, $T[\cdot]$ 는 괄호안의 조건식이 만족되면 1을 그 외에는 0을 반환하는 함수이다. 한편, $\rho_k(\cdot, \cdot)$ 는 SIFT^[25,26] 묘사 벡터를 가중치로 하는 가중치 성분을 의미하고, 다음 식과 같이 정의한다.

$$\rho_k(p, q) = \exp\left(-\frac{\|\phi_{L,k}(p) - \phi_{R,k}(q)\|}{l}\right), \quad (13)$$

여기서, $\phi_{L,k}(p)$ 와 $\phi_{R,k}(p)$ 는 k 채널 왼쪽과 오른쪽 영상의 화소 p 와 q 에 대한 SIFT^[26] 묘사 벡터

(descriptor)를 의미한다. 이때, 본 논문에서는 상수 l 의 값을 실험적으로 $l = 10$ 와 같이 두었다. SIFT 묘사 벡터는 선형 변화에 매우 강인한 특징이 있으므로, 제안하는 방법에서 가장 적합한 묘사 벡터중의 하나이다. 그림 3은 제안하는 가중치가 적용된 동시 확률 분포를 구하는 예를 보여주고 있다. 따라서, 본 논문에서 제안하는 데이터 비용은 위의 식 (12)에서 정의된 동시 확률 분포를 이용하여 정의된 상호 정보량을 통해서 정의된다. 즉, 제안하는 데이터 비용은 아래 식과 같이 정의한다.

$$E_d(f) = -MI(J_L, J_R, f) \approx -\sum_p m(J_L, J_R, f). \quad (14)$$

이때, 화소별 동시 엔트로피 $h(\cdot, \cdot)$ 는 아래식과 같이 정의할 수 있다^{16,22)}.

$$\begin{aligned} H_f(J_L, J_R) &\approx \sum_p h(J_L(p), J_R(p + f_L(p))) \\ &= \sum_p -\frac{1}{N} \log \left\{ P_f^{SIFT}(J_L(p), J_R(p + f_L(p))) \right. \\ &\quad \left. \otimes G(J_L(p), J_R(p + f_L(p))) \right\} \end{aligned} \quad (15)$$

여기서, \otimes 는 컨볼루션 (convolution) 연산을 의미하고, N 은 대응하는 화소의 개수를 의미한다. 또한, $G(\cdot, \cdot)$ 는 2차원 가우시안 (Gaussian) 함수를 의미한다. 마찬가지로, 화소별 주변 엔트로피 (marginal entropy)는 아래와 같이 정의할 수 있다.

$$h(J(p)) = -\frac{1}{N} \log \{ P(J(p)) \otimes G(J(p)) \} \quad (16)$$

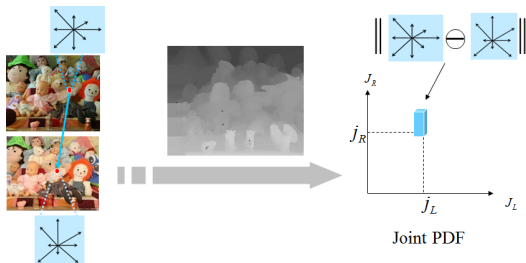


그림 3. 가중치가 적용된 동시 확률 분포
Fig. 3. An example of the weighted joint probability construction

따라서, 제안하는 화소별 상호 정보량은 아래 식과 같이 화소별 엔트로피를 이용하여 정의할 수 있다.

$$\begin{aligned} m(J_L(p), J_R(p + f_L(p))) \\ = h(J_L(p)) + h(J_R(p + f_L(p))) - h(J_L(p), J_R(p + f_L(p))). \end{aligned} \quad (17)$$

2.3 제안하는 데이터 비용의 계산

한편, 제안하는 데이터 비용은 다양한 조건에서 강인함을 유지하기 위하여 식 (14)와 식 (17)에서 정의된 화소별 상호 정보량과 SIFT 묘사 벡터 정보, 그리고 세그먼트 기반의 평면 제한 조건을 적응적으로 결합하여 다음 식 (18)과 같이 정의한다.

$$D_p(f(p)) = \lambda_p m_p(f(p)) + (1 - \lambda_p) \xi_p(f(p)) + \omega \pi_p(f(p)), \quad (18)$$

여기서, $m_p(f(p))$ 와 $\xi_p(f(p))$ 는 각각 상호정보량과 SIFT 묘사 벡터의 화소별 거리에 관한 비용을 의미하고, $\pi_p(f(p))$ 는 세그먼트 기반의 평면 제한 조건 비용을 의미한다. 이때, 각 항목에 대한 가중치는 λ_p , $1 - \lambda_p$, 그리고 ω 로 각각 표현된다. 우선, $m_p(f(p))$ 와 $\xi_p(f(p))$ 는 다음과 같이 정의할 수 있다.

$$m_p(f(p)) = \frac{1}{3} \sum_k \{-m(J_{L,k}(p), J_{R,k}(p + f(p)))\}, \quad (19)$$

$$\xi_p(f(p)) = \frac{1}{3} \sum_k \left\{ \frac{\| \phi_{L,k}(p) - \phi_{R,k}(p + f(p)) \|}{l} \right\}, \quad (20)$$

$k \in \{R, G, B\}$,

여기서, $m_p(f(p))$ 은 상호 정보량에 관한 식이고, $\xi_p(f(p))$ 는 SIFT 묘사 벡터에 관한 식이다. 이 두 가지 비용은 각각 서로 다른 특징이 있으므로, 이 둘 사이의 가중치를 적응적으로 결정하기 위해서 가중치 비중 λ_p 을 도입하였다. 이때, λ_p 는 각 비용의 신뢰도 (confidence) 비율에 의해서 결정이 되고, 다음 식과 같이, 계산한 비용 공간 (cost volume)의 비용 커브 (cost curve)에 의해서 결정이 된다¹⁹⁾. 즉, 각 화소마다 변위에 따른 비용 함수의 최소값과 두 번째 최소값의 비율에 의해서 신뢰도를 정의한다. 일반적으로, 이러한 최소값과 두 번째 최소값의 비율이 클수록 신뢰

도가 높다고 판단할 수 있다. 따라서, 각 비용의 신뢰도는 다음 식과 같이 정의한다.

$$\Lambda_m(p) = \frac{m_p^{c2}}{m_p^{c1}}, \Lambda_\xi(p) = \frac{\xi_p^{c2}}{\xi_p^{c1}}, \quad (21)$$

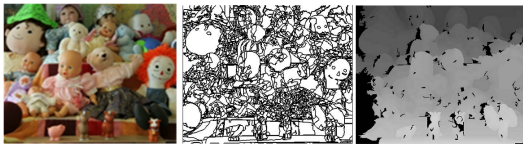
이때, m_p^{c1} 과 m_p^{c2} 는 각각 화소 p 에서 상호 정보량의 비용 커브에서 최소값과, 두 번째 최소값을 의미한다. 마찬가지로, ξ_p^{c1} 과 ξ_p^{c2} 도 SIFT 묘사 벡터의 화소 p 에서의 비용 커브에서 최소값과 두 번째 최소값을 의미한다. 위의 식에서 정의된 Λ_m 과 Λ_ξ 를 이용하여 가중치 비중 λ_p 는 다음과 같이 정의된다.

$$\lambda_p = \frac{\Lambda_m(p)}{\Lambda_m(p) + \Lambda_\xi(p)}. \quad (22)$$

한편, 보다 정확한 변위 지도의 예측을 위하여, 제안하는 데이터 비용은 추가적으로 세그먼트 기반의 평면 제한 조건^[18]도 포함이 되었다. 이는 식 (18)에서 $\pi_p(f(p))$ 로 표현이 되어 있다. 일반적으로, 평면은 다음과 같은 3개의 파라미터로 결정할 수가 있다.

$$f = ax + by + c, \quad (23)$$

여기서 $\{a, b, c\}$ 는 평면을 결정하는 파라미터이고, $\{x, y\}$ 는 화소의 위치에 대한 파라미터이다. 각 세그먼트 내에서 위의 평면의 파라미터 $\{a, b, c\}$ 를 구하기 위해서, 2 단계에 거치는 작업을 수행하였다. 이를 위해서 첫 단계에서는 세그먼트로 영상을 분할을 수행하였다. 이를 위하여, mean-shift^[8] 기반의 영상 분할 알고리즘을 수행하였다. 그 다음, 각 세그먼트에 대해서 최소 자승법 (least-square)을 이용하여 평면에 대한 최적의 파라미터 $\{a, b, c\}$ 를 예측하였다. 이를 위하여 특정 세그먼트 s 에 포함된 화소 값들을 이용



(a) (b) (c)

그림 4. 세그먼트 기반 평면 예측의 예 (a) 입력 영상 (b) mean-shift 세그먼트 영상 (c) 세그먼트 기반 평면 예측 결과
Fig. 4. An example of the segment-based plane fitting

하여 다음과 같은 선형 시스템 식을 전개할 수 있다.

$$Av = b \Leftrightarrow \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \dots & \dots & \dots \\ x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} a_s \\ b_s \\ c_s \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}, \quad (24)$$

이때, 각 세그먼트 s 에서의 평면 파라미터는 다음 식과 같이 결정이 된다.

$$\begin{bmatrix} a_s \\ b_s \\ c_s \end{bmatrix} = v = (A^T A)^{-1} A^T b, \quad (25)$$

여기서, n 은 세그먼트 s 의 화소 개수를 의미한다. 따라서, $\pi_p(f(p))$ 는 아래 식과 같이 정의한다.

$$\pi_p(f(p)) = |a_s x_p + b_s y_p + c_s - f(p)|. \quad (26)$$

위의 식과 같이 평면 제한 조건식 (26)은 세그먼트를 통해서 구한 최적의 평면 식과 구하려는 변위 값의 절대값 차이에 의해서 결정이 된다. 그림 4는 세그먼트의 예와 제안하는 세그먼트 기반 평면 예측의 예를 보여주고 있다.

2.4 전역적인 비용 최적화를 통한 변위 지도 예측

한편, 위에서 정의한 MAP-MRF 비용 최적화를 위해서 평탄화 비용 $V_{pq}(f_p, f_q)$ 은 아래 선형 모델 (truncated linear model)과 같이 정의한다.

$$V_{pq}(f_p, f_q) = \eta \cdot \min(|f_p - f_q|, V_{\max}), \quad (27)$$

여기서, 상수는 $\eta = 1$, $V_{\max} = 5.0$ 과 같이 설정하였다. 식 (14)와 식 (27)을 결합하여 식 (8)의 전체 에너지는 TRW-S (Sequential Tree-ReWeighted message passing)^[43,44] 기법을 사용하여 최적화를 통해서 최적의 변위 지도를 예측하였다. 일반적으로 TRW-S 기법은 다양한 비선형 비용 함수 최적화에 사용할 수 있고, 다른 최적화 기법들보다 최적의 해를 찾는 데 있어서 성능이 가장 우수한 것으로 알려져 있다^[43,44].

III. 실험 결과

제안하는 알고리즘의 성능을 평가하기 위해서 다양한 데이터에 대해서 실험을 수행하였다. 실험의 다양

성을 위해서 미들베리 (Middlebury) 사이트^[4]에 있는 다양한 데이터들과 조명 변화가 발생하는 항공 영상들에 대해서 제안하는 알고리즘의 성능을 테스트하였다.

3.1 미들베리 데이터에 대한 실험

우선, 미들베리 사이트에 있는 다양한 영상 데이터에 대해서 실험을 수행하였다. 미들베리 데이터의 경우, 각 데이터 셋은 3가지 종류의 조명의 변화와 3가지 종류의 카메라의 노출 변화를 수행했을 때의 영상 데이터를 포함하고 있다. 또한, 각 데이터에 대해서 참값 변위 지도를 공개하여서 정량적인 성능 평가를 할 수가 있다. 제안하는 알고리즘의 성능 평가를 위해서 기존의 다양한 스테레오 알고리즘들과 비교 및 분석을 수행하였다. 이때 비교 실험을 위해서 사용한 기존의 여러 강인한 스테레오 정합 기법들은 Census^[41], NCC^[17], ANCC^[13], MI^[16,22] 기법들을 포함하고 있다. 각 기법들의 변위 지도의 오류를 측정하기 위하여, 제공된 참값 변위 지도와 알고리즘들의 결과로 얻은 변위 지도를 비교하였고, 변위 값의 차이가 2 이상인 경우를 오류로 평가하였다^[32].

이러한 기법들을 다양한 실험 데이터 영상들 Dolls, Laundry, Moebius, Rocks1에 대해서 비교 분석하였다. 그림 5와 그림 6은 각각 Dolls 데이터에 대해서 조명 변화와 카메라 노출을 변화시켰을 때 각각 다양한 기법에 대해서 정성적으로 비교한 결과를 나타낸다. 그림들에서 확인할 수 있듯이 제안하는 방법이 기존의 기법들에 비해서 보다 정확한 변위 지도를 예측하는 것을 알 수 있다.

그림 7은 조명 변화에 대하여 다양한 데이터에 대한 제안하는 방법의 정성적 실험 결과를 보여주고 있다. 그림에서 확인할 수 있듯이 다양한 실험 데이터에

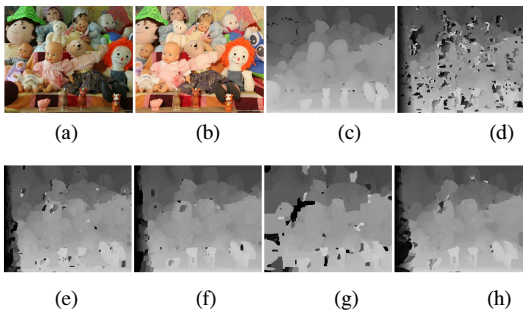


그림 5. Dolls 실험 데이터에 대한 조명 변화에 따른 비교 실험 결과 (a) 왼쪽 영상 (b) 오른쪽 영상 (c) 참 값 (d) Census 결과 (e) NCC 결과 (f) ANCC 결과 (g) MI 결과 (h) 제안하는 방법의 결과
Fig. 5. Comparative Results for different light configurations of Dolls dataset

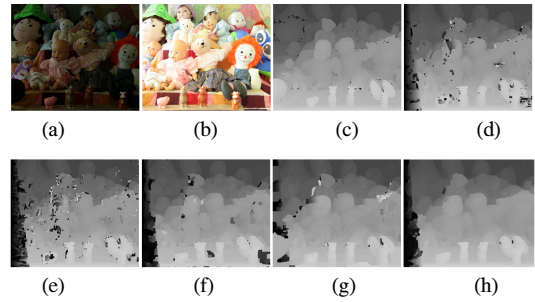


그림 6. Dolls 데이터에 대한 카메라 노출 변화에 따른 비교 실험 결과 (a) 왼쪽 영상 (b) 오른쪽 영상 (c) 참 값 (d) Census 결과 (e) NCC 결과 (f) ANCC 결과 (g) MI 결과 (h) 제안하는 방법의 결과
Fig. 6. Comparative Results for different camera exposures of Dolls dataset

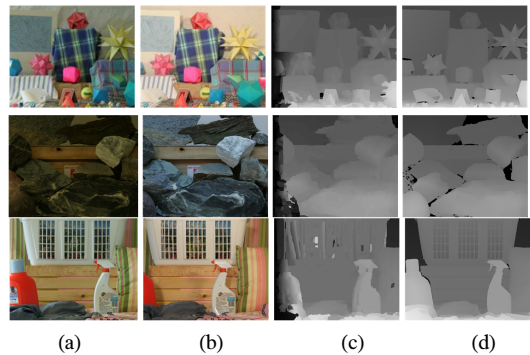


그림 7. 미들베리 데이터에 대한 조명 변화에 따른 실험 결과 (a) 왼쪽 영상 (b) 오른쪽 영상 (c) 제안하는 방법의 결과 (d) 참 값
Fig. 7. Qualitative Results for different light configurations

대해서 제안하는 방법은 안정적으로 정확한 결과를 내는 것을 확인할 수 있다. 그림 8은 조명 변화에 대해서 미들베리 데이터에 대한 정량적 비교 결과를 보여주고 있다. 여기서, x축은 미들베리 영상에 대해서 조명 변화에 대한 번호 (index)를 표시하고, y축은 가려지지 않은 지역 (unoccluded region)에 대한 오류 화소의 비율을 의미한다. x축의 경우, 예를 들면, “1/3”은 왼쪽 영상의 번호가 “1”이고, 오른쪽 영상의 번호가 “3”임을 의미한다. 각각의 번호는 각각 다른 조명 방향을 의미한다. 따라서, “1/1” 또는 “2/2”인 경우에 비해서 “2/3” 또는 “1/3”과 같이 좌우 영상의 번호 차이가 많이 날수록 좌우 영상의 조명 차이가 크게 발생한다. 이러한 조명 변화로 인해서 발생하는 영향은 주로 그림자 변화 등으로 인해서 지역적으로 밝기 값이 변화하는 것이다.

이러한 조명 변화로 인한 지역적인 영상의 변화는

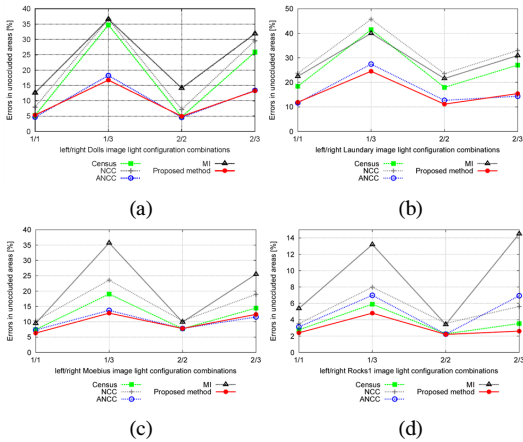


그림 8. 미들베리 데이터에 대한 조명 변화에 따른 정량적인 실험 결과 (a) Dolls (b) Laundry (c) Moebius (d) Rocks1
 Fig. 8. Quantitative Results for light configuration combinations

영상을 보정하기도 매우 어렵고, 대부분의 기법들이 전역적인 변화를 가정하고 있기 때문에 이러한 지역적인 변화에 대해서는 매우 불안정한 결과를 내는 것을 알 수 있다. 이에 반해서 제안하는 방법은 기존의 방법들에 비해서 정량적으로도 가장 안정적이고 우수한 성능을 나타내는 것을 알 수 있다. 그러나, “1/3”과 같은 극심한 조명 변화가 있는 데이터에 대해서는 그림자 변화가 크게 발생하거나,

물체 표면이 매끈한 경우 입사된 빛의 일부가 전반사 (specular reflection)되어 밝기 값이 매우 큰 경우가 발생한다. 제안하는 알고리즘은 물체 표면이 난반사 (diffuse reflection)임을 가정하고 있기 때문에, 이러한 경우들에 대해서는 오류 화소가 크게 증가하게 된다.

한편, 그림 9는 카메라 노출 변화에 대하여 다양한 데이터에 대해서 제안하는 방법의 정성적 실험 결과를 보여주고 있다. 카메라 노출 변화는 스테레오 영상들 사이에 전역적인 변화를 발생시킨다. 카메라 노출을 짧은 시간 동안 설정하면, 대체로 어두운 영상이 획득이 되고, 노출 시간을 길게 하면 밝은 영상이 획득되는 것을 알 수 있다. 그림 9에서

확인할 수 있듯이, 이러한 전역적인 변화에 대해서도 다양한 실험 데이터에 대해서 제안하는 방법이란 정적인 정확한 결과를 내는 것을 확인할 수 있다.

그림 10은 카메라 노출 변화에 대해서 미들베리 데이터들에 대한 정량적 비교 결과를 보여주고 있다. 그림 8의 경우와 마찬가지로, x축은 미들베리 영상 데이터에 대해서 카메라의 노출 시간 변화에 대한 번호를

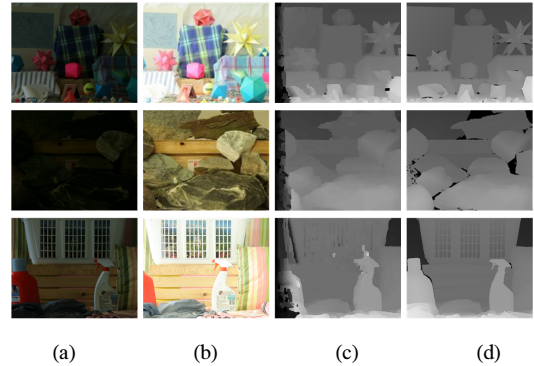


그림 9. 미들베리 데이터에 대한 카메라 노출 변화에 따른 실험 결과 (a) 왼쪽 영상 (b) 오른쪽 영상 (c) 제안하는 방법의 결과 (d) 참 값
 Fig. 9. Qualitative Results for different camera exposures

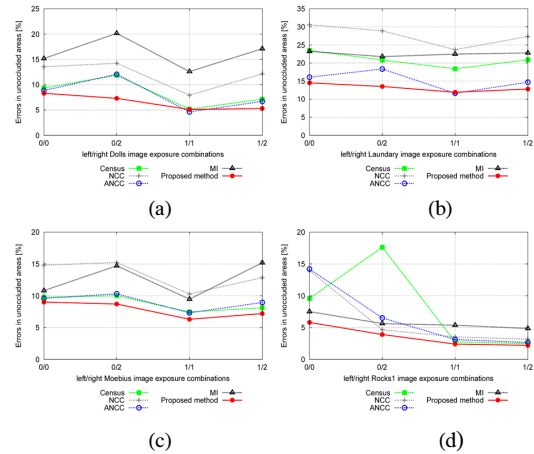


그림 10. 미들베리 데이터에 대한 카메라 노출 변화에 따른 정량적인 실험 결과 (a) Dolls (b) Laundry (c) Moebius (d) Rocks1
 Fig. 10. Quantitative Results for light configuration combinations

표시하고, y축은 가려지지 않은 지역 (unoccluded region)에 대한 오류 화소의 비율을 의미한다. 예를 들면, x축에서 “0/2”는 왼쪽 영상의 노출 시간의 번호가 “0”이고, 오른쪽 영상의 노출 시간의 번호가 “2”임을 의미한다. 각각의 번호는 각각 다른 노출 시간을 의미하고, 좌우 영상의 번호의 차이가 많이 날수록 좌우 영상의 카메라 노출 변화의 차이가 크다. 카메라 노출 값의 변화는 전역적인 밝기 변화에 영향을 미치고, 주로 영상에 밝기 값이 포화된 (saturated) 화소를 생성할 수 있다. 즉, 번호 “0”에 해당하는 짧은 노출 시간은 밝기 값이 “0”에 가까운 어두운 화소들을 생성할 수 있고, 긴 노출 시간에 해당하는 번호 “2”는 밝기 값이 “255”에 가까운 매우 밝은 화소들을 생성할 수

있다.

이러한 전역적인 변화에 대해서는, 대부분의 알고리즘들이 안정적인 결과를 보여 주고 있고, 제안하는 방법이 기존의 방법들에 비해서 가장 우수한 성능을 나타내는 것을 알 수 있다. 그러나, 일부 포화된 화소 값들은 정합을 하기 위한 정보량이 거의 없으므로, 정합을 어렵게 만드는 카메라 변화 요인 중 하나이고, 제안하는 기법도 다른 기법들과 마찬가지로 이러한 포화 화소의 정도에 따라서 성능의 영향을 받는 것을 알 수 있다.

그림 11에서는 제안하는 방법에 대해서 원본 컬러를 사용했을 때와 로그-색도 컬러를 사용했을 때의 결과를 비교하였다. 비교를 위해서 “1/3”과 같은 극심한 조명 변화가 있는 경우에 대한 실험을 수행하였다. 이러한 경우에 심한 조명 변화가 발생하는 부분에 대해서 로그-색도 컬러 공간을 사용하는 방법이 원본 컬러를 사용하는 경우보다 적은 오류를 갖는 변위 지도를 생성하는 것을 확인할 수 있다.

한편, 각 정합 기법들의 복잡도를 비교하기 위하여 그림 12에서는 각 기법을 얻는데 필요한 수행 시간을 비교하였다. 이를 위하여 Intel Core i7 CPU@3.4GHz PC 환경에서 실험을 수행하였고, 실험 영상은 427x370 화소 크기의 영상에 대해서 변위 범위 (disparity range)가 [0,70]인 경우로 설정을 하였다. 이때, Census, NCC, ANCC의 경우 모두 윈도우 기반의 정합 기법이므로 동일한 7x7 윈도우 크기로 설정하였다. MI의 경우는 화소 기반의 반복적인 방법이므로, 수렴하는 경우까지 3회 반복을 수행하였다. 제안

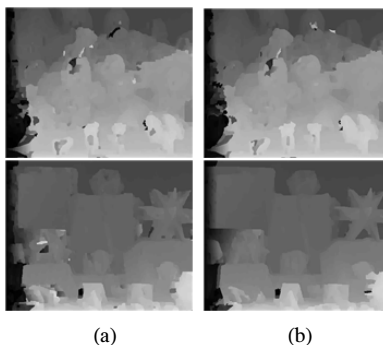


그림 11. 원본 컬러와 로그-색도를 이용한 제안하는 알고리즘의 결과 비교 (a) 원본 컬러를 이용한 제안하는 알고리즘 결과. Dolls(변위 지도 오류: 18.9%), Moebius(변위 지도 오류: 16.1%) (b) 로그-색도 컬러를 이용한 제안하는 알고리즘 결과. Dolls(변위 지도 오류: 16.7%), Moebius(변위 지도 오류: 12.8%)
Fig. 11. Comparative Results for the original color and log-chrom color for the proposed method

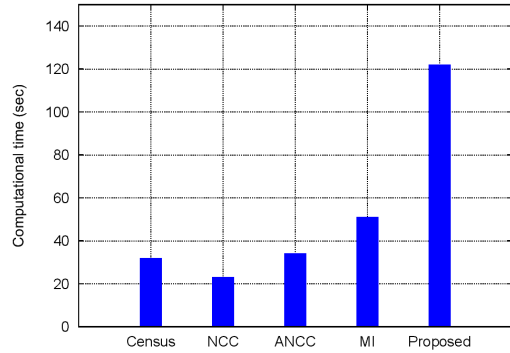


그림 12. 각 정합 비용의 수행 시간 비교
Fig. 12. Comparison of computational time for each method

하는 기법은 여러 가지 정보를 융합하여 사용하므로, 다른 기법들에 비하여 비교적 시간이 많이 걸리는 단점이 있지만, GPGPU와 같은 병렬 고속 연산이 가능한 하드웨어를 사용하여 비교적 계산량이 많은 SIFT 묘사 벡터 계산 부분이나 각 세그먼트의 평면 파라미터를 계산하는 부분을 고속화를 할 수 있을 것으로 기대하고 있다.

3.2 항공 영상에 대한 실험

이번 절에서는 항공 영상에 대해서 실험을 수행하였다. 일반적으로 항공 영상은 넓은 지역을 일정한 시간 간격을 두고 촬영을 하므로, 시간차에 따른 다양한 조명 변화가 좌우 스테레오 영상에 매우 다르게 나타날 수 있다. 아래 그림 13은 이러한 항공영상에 대해서 실험한 결과를 나타낸다. 항공 영상의 경우 취득 시간의 차이로 인해서 다양한 조명 변화가 발생한다. 이러한 특유의 다양한 조명 변화에 대해서도 제안하는 방법이 비교적 효율적으로 변위 지도를 예측하는 것을 알 수 있다. 그러나, 물체의 경계 부근에

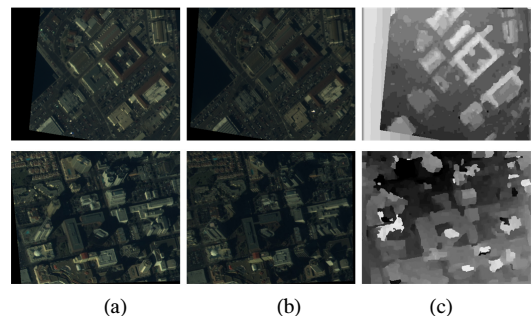


그림 13. 항공 영상에 대한 실험 결과 (a) 왼쪽 영상 (b) 오른쪽 영상 (c) 제안하는 방법의 결과
Fig. 13. Experiments for aerial images

서는 극심한 조명 변화로 인하여 그림자의 변화가 크게 발생하는데, 이러한 경우는 제안하는 알고리즘이 오류를 발생하는 것을 확인할 수 있다.

IV. 결 론

본 논문에서는 가중치가 적용된 상호 정보량과 SIFT 묘사 벡터, 그리고 세그먼트 기반의 평면 제한 조건을 적응적으로 결합하여 최적화를 통하여 다양한 조명 및 카메라 변화에 강인한 스테레오 정합 기법을 제안하였다. 다양한 실험 데이터에 대해서 기존의 알고리즘들과 비교 실험을 수행하여 제안하는 알고리즘이 기존 알고리즘들에 비하여 강인하게 동작하는 것을 확인할 수 있었다. 특히, 조명 변화가 극심한 경우는 대부분의 기존 기법들의 오류가 크게 증가하고 있는 것에 비해서 제안하는 알고리즘이 가장 적은 오류를 보이는 것을 확인하였다. 그러나, 조명의 변화로 인하여 그림자의 변화가 크게 발생하거나, 전반사와 같이 이상적인 램버시안 (Lambertian) 물체 표면의 가정에서 벗어나는 화소들의 경우는 제안하는 알고리즘도 오류 화소가 크게 증가하는 것을 확인할 수 있다. 또한, 카메라 노출 변화가 극심한 경우에 대해서는 노출 변화가 없는 경우와 비교했을 때, 대부분의 알고리즘들이 오류의 차이가 크지 않아 안정적으로 동작함을 확인할 수 있었다. 이러한 카메라 노출 변화의 경우에는 대부분의 방법들이 밝기 값의 포화 화소로 인한 정보량의 손실로 인해서 오류가 주로 발생하고, 이러한 문제는 향후 연구에서 지속적으로 연구되어 해결될 것으로 기대하고 있다.

References

[1] W.-S. Jang, C. Lee, and Y.-S. Ho, "Efficient depth map generation for various stereo camera arrangements," *J. KICS*, vol. 37, no. 6, pp. 458-463, 2012.

[2] C. Lee, H. Song, B. Choi, and Y.-S. Ho, "Multi-view generation using high resolution stereoscopic cameras and a low resolution time-of-flight camera," *J. KICS*, vol. 37, no. 4, pp. 239-249, 2012.

[3] C. Song and J. Lee, "Detection of illegal u-turn vehicles by optical flow analysis," *J. KICS*, vol. 39, no. 10, pp. 948-956, 2014.

[4] <http://vision.middlebury.edu/stereo/>, 2015.

[5] <http://www.photomodeler.com/>, 2015.

[6] S. Birchfield and C. Tomasi, "A pixel dissimilarity measure that is insensitive to image sampling," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 20, no. 4, pp. 401-406, 1998.

[7] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 23, no. 11, pp. 1222-1239, 2001.

[8] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 24, no. 5, pp. 603-619, 2001.

[9] A. Chakrabarti, D. Scharstein, and T. Zickler, "An empirical camera model for internet color vision," in *Proc. British Machine Vision Conf.*, 2009.

[10] M. Ebner, *Color constancy*, ser. Wiley-IS&T Series in Imaging Sci. and Technol., John Wiley & Sons, 2007.

[11] G. Egnal, *Mutual Information as a Stereo Correspondence Measure*, in Technical Report MS-CIS-00-20, Computer and Information Science, Univ. of Pennsylvania, 2000.

[12] G. D. Finlayson, S. D. Hordley, G. Schaefer, and G. Y. Tian, "Illuminant and device invariant colour using histogram equalisation," *Pattern Recognition*, vol. 38, no. 2, pp. 179-190, 2005.

[13] Y. S. Heo, K. M. Lee, and S. U. Lee, "Robust stereo matching using adaptive normalized cross correlation," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 33, no. 4, pp. 807-822, 2011.

[14] Y. S. Heo, K. M. Lee, and S. U. Lee, "Mutual information-based stereo matching combined with SIFT descriptor in log-chromaticity color space," in *Proc. IEEE CVPR 2009*, pp. 445-452, Miami, FL, Jun. 2009.

[15] Y. S. Heo, K. M. Lee, and S. U. Lee, "Joint depth map and color consistency estimation for stereo images with different illuminations and cameras," *IEEE Trans. Pattern Anal. and*

- Machine Intell.*, vol. 35, no. 5, pp. 1094-1106, 2013.
- [16] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 30, no. 2, pp. 328-341, 2008.
- [17] H. Hirschmuller and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 31, no. 9, pp. 1582-1599, 2009.
- [18] L. Hong and G. Chen, "Segment-based stereo matching using graph cuts," in *Proc. IEEE CVPR*, vol. 1, pp. 74-81, 2004.
- [19] X. Hu and P. Mordohai, "Evaluation of stereo confidence indoors and outdoors," in *Proc. IEEE CVPR*, pp. 1466-1473, San Francisco, Jun. 2010.
- [20] H. Jeon, A. Basso, and P. F. Driessen, "A global correspondence for scale invariant matching using mutual information and the graph search," in *Proc. Int. Conf. Multimedia and Expo*, pp. 1745-1748, Toronto, Ont, Jul. 2006.
- [21] S. Kagarlitsky, Y. Moses, and Y. Hel-Or, "Piecewise-consistent color mappings of images acquired under various conditions," in *Proc. IEEE Int. Conf. Computer Vision*, pp. 2311-2318, Kyoto, Sept.-Oct. 2009.
- [22] J. Kim, V. Kolmogorov, and R. Zabih, "Visual correspondence using energy minimization and mutual information," in *Proc. IEEE Int'l Conf. Computer Vision*, vol. 2, pp. 1033-1040, Nice, France, Oct. 2003.
- [23] S. J. Kim and M. Pollefeys, "Robust radiometric calibration and vignetting correction," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 30, no. 4, pp. 562-576, Apr. 2008.
- [24] S. Lin, Y. Li, S. B. Kang, X. Tong, and H. Y. Shum, "Diffuse-specular separation and depth recovery from image sequences," in *Proc. Eur. Conf. Computer Vision*, 2002.
- [25] C. Liu, J. Yuen, A. Torralba, J. Sivic, and W. T. Freeman, "SIFT flow: Dense correspondence across different scenes," in *Proc. Eur. Conf. Computer Vision*, vol. 5034, pp. 28-42, 2008.
- [26] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int'l J. Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [27] P. Montesinos, V. Gouet, R. Deriche, and D. Pele, "Matching color uncalibrated images using differential invariants," *Image and Vision Computing*, vol. 18, no. 1, pp. 659-671, 2000.
- [28] A. S. Ogale and Y. Aloimonos, "Robust contrast invariant stereo correspondence," in *Proc. IEEE Int. Conf. Robot. and Automat.*, pp. 819-824, Apr. 2004.
- [29] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imaging*, vol. 19, no. 8, pp. 809-814, 2000.
- [30] J. P. W. Pluim, J. B. A. Maintz and M. A. Viergever, "Mutual-information-based registration of medical images : A survey," *IEEE Trans. Med. Imaging*, vol. 22, no. 8, pp. 986-1004, 2003.
- [31] D. B. Russakoff, C. Tomasi, T. Rohlfing, and C. R. Maurer Jr., "Image similarity using mutual information of regions," in *Proc. Eur. Conf. Computer Vision*, vol. 3023, pp. 596-607, 2004.
- [32] 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.
- [33] N. Snavely, S. M. Seitz, and R. Szeliski, "Modeling the world from internet photo collections," *Int. J. Computer Vision*, vol. 80, no. 2, pp. 189-210, 2008.
- [34] S. M. Seitz and S. Baker, "Filter flow," in *Proc. IEEE Int. Conf. Computer Vision*, 2009.
- [35] J. Sun, Y. Li, S. B. Kang, and H. Y. Shum, "Symmetric stereo matching for occlusion handling," in *Proc. CVPR*, vol. 2, pp. 399-406, Jun. 2005.
- [36] E. Tola, V. Lepetit, and P. Fua, "DAISY: An efficient dense descriptor applied to wide-

baseline stereo,” *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 32, no. 5, pp. 815-830, 2010.

- [37] P. Viola and W. M. Wells III, “Alignment by maximization of mutual information,” *Int. J. Computer Vision*, vol. 24, no. 2, pp. 137-154, 1997.
- [38] Y. Weiss, “Deriving intrinsic images from image sequences,” in *Proc. IEEE Int. Conf. Computer Vision*, vol. 2, pp. 68-75, Vancouver, 2001.
- [39] W. Xiong and B. Funt, “Stereo retinex,” in *Proc. Canadian Conf. Computer and Robot Vision*, Jun. 2006.
- [40] J. N. Yang and S. K. Shevell, “Stereo disparity improves color constancy,” *Vision Res.*, vol. 42, no. 1, pp. 1979-1989, 2002.
- [41] R. Zabih and J. Woodfill, “Non-parametric local transforms for computing visual correspondence,” in *Proc. Eur. Conf. Computer Vision*, vol. 801, pp. 151-158, Jun. 1994.
- [42] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, “High-quality video view interpolation using a layered representation,” in *Proc. SIGGRAPH*, pp. 600-608, Aug. 2004.
- [43] V. Kolmogorov, “Convergent tree-reweighted message passing for energy minimization,” *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 28, no. 10, pp. 1568-1583, 2006.
- [44] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, “A comparative study of energy minimization methods for markov random fields with smoothness-based priors,” *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 30, no. 6, pp. 1068-1080, Jun. 2008.

허 용 석 (Yong Seok Heo)



2005년 2월 : 서울대학교 전기공학부 졸업

2007년 2월 : 서울대학교 전기컴퓨터 공학부 석사

2012년 8월 : 서울대학교 전기컴퓨터 공학부 박사

2014년 9월~현재 : 아주대학교 전자공학과 조교수

<관심분야> 컴퓨터 비전, 패턴 인식