



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

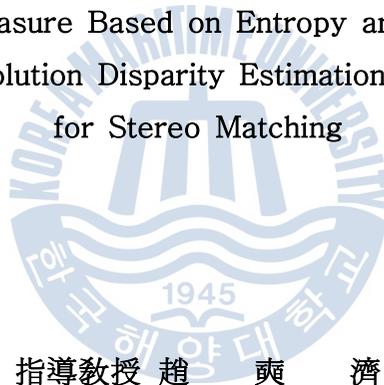
이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

工學博士 學位論文

스테레오 정합을 위한 엔트로피 및  
Census 기반의 유사도 척도와  
다해상도 변위 추정 기법

Similarity Measure Based on Entropy and Census and  
Multi-Resolution Disparity Estimation Technique  
for Stereo Matching



指導教授 趙 爽 濟

2013年 8月

韓國海洋大學校 大學院

制 御 計 測 工 學 科

洪 錫 根

본 논문을 홍석근의 공학박사 학위논문으로 인준함.

위원장 공학박사 김 중 화



위원 공학박사 류길수



위원 공학박사 하윤수



위원 공학박사 김정엽



위원 공학박사 조석제



2013년 6월 19일

한국해양대학교 대학원

# 목 차

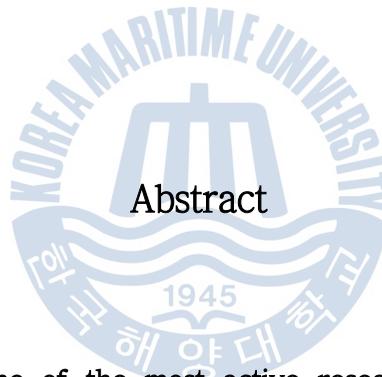
제 1 장 서 론 .....	1
1.1 연구 목적 및 배경.....	1
1.2 관련 연구.....	3
1.3 연구 내용.....	6
1.4 논문의 구성.....	7
제 2 장 스테레오 시각과 스테레오 정합.....	8
2.1 스테레오 시각.....	8
2.2 스테레오 정합.....	10
2.2.1 유사도 척도.....	10
2.2.2 최적화 방법.....	15
2.3 환경 변화에 강인한 유사도 척도 .....	18
2.3.1 특징 기반 유사도 척도 .....	18
2.3.2 명암도 기반 유사도 척도.....	19
2.3.3 비모수 유사도 척도.....	22
제 3 장 엔트로피 및 Census 기반의 유사도 척도 .....	24
3.1 엔트로피 기반의 유사도 척도 .....	26
3.1.1 엔트로피 .....	26
3.1.2 엔트로피를 이용한 MI 유사도 척도 .....	27
3.2 제안한 Haar 웨이블릿 특징을 결합한 엔트로피 유사도 척도 .....	28
3.2.1 화소 단위 엔트로피 .....	28
3.2.2 Haar 웨이블릿 특징을 결합한 엔트로피 .....	35
3.3 제안한 Census 변환 기반의 유사도 척도 .....	46

3.3.1 Census 변환.....	46
3.3.2 제안한 Census 변환을 이용한 유사도 척도.....	49
<b>제 4 장 8방향 동적 계획법을 이용한 변위 추정.....</b>	<b>53</b>
4.1 동적 계획법.....	53
4.2 제안한 8방향 동적 계획법.....	57
<b>제 5 장 다해상도 기반의 스테레오 정합.....</b>	<b>67</b>
5.1 가우시안 영상 피라미드.....	67
5.2 제안한 다해상도 기반 스테레오 정합.....	71
<b>제 6 장 실험 및 고찰.....</b>	<b>77</b>
6.1 정합 성능 평가 방법.....	77
6.2 스테레오 정합 실험.....	79
6.2.1 RDS 영상 실험.....	79
6.2.2 환경 변화가 없는 표준 영상 실험.....	84
6.2.3 환경 변화가 발생한 표준 영상 실험.....	92
6.2.4 실제 영상 실험.....	110
6.3 계산 속도.....	118
6.4 제안한 방법의 정합 성능에 대한 고찰.....	121
<b>제 7 장 결    론.....</b>	<b>123</b>
<b>참고 문헌.....</b>	<b>126</b>

Similarity Measure Based on Entropy and Census and  
Multi-Resolution Disparity Estimation Technique  
for Stereo Matching

*Seok-Keun, Hong*

*Department of Control and Instrumentation Engineering,  
Graduate School of Korea Maritime University*



Abstract

Stereo matching is one of the most active research areas in the field of computer vision. Stereo matching aims to obtain 3D information by extracting correct correspondence between two images captured from different point of views. There are two research parts in stereo matching: similarity measure between correspondence points and optimization technique for dense disparity estimation.

The crux of stereo matching problem in similarity measure perspective is how to deal with the inferent points ambiguity that results from the ambiguous local appearances of image points. Similarity measures in stereo matching are classified as feature-based, intensity-based or non-parametric

measure. And most similarity measures in the literatures are based on pixel intensity comparison. When images are taken at different illumination conditions or different sensors used, it is very unlikely that the corresponding pixels would have the same intensity creating false correspondences if it is only based on intensity matching functions alone. Especially illumination variations between input images can cause serious degrade in the performance of stereo matching algorithms. In this situation, mutual information-based method is powerful. However, it is still ambiguous or erroneous in considering local illumination variations between images. Therefore, similarity measure to these radiometric variations are demanded and become inevitable for stereo matching.

Optimization method in stereo matching can be classified into two categories: local and global optimization methods, and most state-of-the-art algorithms fall into global optimization method. Global optimization methods can greatly suppress the matching ambiguities caused by various factors such as occluded and textureless regions. However, They are usually computationally expensive due to the slow-converging optimization process.

In this paper, it was proposed that a stereo matching similarity measure based on entropy and census transform and an optimization technique using dynamic programming to estimate disparity efficiently based on multi-resolution method. Proposed similarity measure is composed of entropy, Haar wavelet feature vector, and modified Census transform. In general, mutual information similarity measure based on entropy about stereo images and disparity map is a popular and powerful similarity measure which is robust to complex intensity transformation. However, it is still ambiguous or erroneous with local radiometric variations, since it only accounts for global variation between images, and does not contain spatial information. Haar

wavelet response can express frequency properties of image regions and is robust to various intensity changes and bias. Therefore, entropy was utilized with Haar wavelet feature vector as geometric measure. Modified Census transform was used as another spatial similarity measure. Census transform is a well-known non-parametric measure. And it is powerful to textureless and disparity discontinuity region and robust to noisy environment. A combination of entropy with Haar wavelet feature vector and modified Census transform as similarity measure was proposed to find correspondence. It is invariant to local radiometric variations and global illumination changes, so it can be applied to find correspondence for images which undergo local as well as global radiometric variations.

Proposed optimization method is a new disparity estimation technique based on dynamic programming. A method using dynamic programming with 8-direction energy aggregation to estimate accurate disparity map was applied. Using 8-direction energy aggregation, accurate disparities can be found at disparity discontinuous region and suppress a streaking phenomenon in disparity map.

Finally, the multi-resolution scheme was proposed to increase efficiency while processing and disparity estimation method. A Gaussian pyramid which prevent the ailing at low-resolution image pyramid levels was used. And the multi-resolution scheme was proposed to perform matching at every levels to find accurate disparity. This method can perform matching efficiently and make accurate disparity map.

And proposed method was validated with experimental results on stereo images.

# 제 1 장 서 론

## 1.1 연구 목적 및 배경

최근 컴퓨터 산업과 로봇 산업이 발달함에 따라 카메라와 컴퓨터를 이용하여 인간의 시각 기능을 구현하는 기술에 대한 관심이 크다. 컴퓨터나 로봇의 시각 기능을 구현하는 분야를 컴퓨터 시각(computer vision)이라 한다. 컴퓨터 시각은 영상을 분석하여 특정한 물체를 인식(recognition)하고 장면을 이해(understanding)하는 기술로 관련된 연구 분야로는 영상 내 물체 분류 및 추적, 형상 인식, 로봇의 자율 주행, 3차원 모델링 및 거리 인식 등이 있다[1-4]. 여기서 3차원 영상 응용 분야는 스테레오 시각(stereo vision)이라는 별도의 영역에서 다룬다. 스테레오 시각은 인간이 두 눈과 뇌로 장면의 공간을 인식하는 것처럼 컴퓨터가 두 장 이상의 2차원 평면 영상을 처리하여 3차원 정보를 추출하는 기술이다[5-7].

스테레오 시각으로 얻은 3차원 정보는 물체 인식 및 거리 인식, 로봇의 자율 주행을 위한 시각 기능 구현, 장면이나 물체의 3차원 재구성 등에 응용되고 있다. 이와 같은 3차원 기술의 구현을 위하여 시차가 다른 두 영상에서 같은 물체의 지점, 즉 대응점(corresponding point)을 찾는 과정을 스테레오 정합(stereo matching)이라 한다. 스테레오 정합은 스테레오 시각에서 가장 중요하고 어려운 과정이다. 정확한 정합을 어렵게 하는 대표적인 요인들은 두 영상의 시차로 인해 발생하는 가려지는 폐색 영역(occluding region), 질감(texture)이 약한 영역, 물체의 경계선(edge), 두 영상의 밝기 차이 및 잡음 등이 있다. 그래서 대부분의 스테레오 시각에 대한 연구는 이러한 요인들로 인해 발생하는 오정합을 줄이고 정합점을 정확히 찾는 데 초점이 맞춰져 있다[8-10].

스테레오 정합은 정합 요소에 따라 특징기반(feature-based) 정합과 영역기반(area-based) 정합으로 분류된다[5]. 정합 요소는 두 영상의 대응점을 정량적으

로 표현하고 유사도를 계산하는데 쓰인다. 그리고 유사도는 변위(disparity)를 추정하는데 쓰이는 에너지 함수(energy function) 또는 비용 함수(cost function)의 중요한 요소이다. 특징기반 정합은 중간단계(mid-level) 영상처리를 거쳐 얻을 수 있는 경계선이나 분할 영역(segmentation region) 등 검출된 기하학적 특징에 대해 변위를 얻기 때문에 처리속도가 빠르고 정교한 변위를 계산할 수 있으나 특징으로 검출되지 않은 화소의 변위는 정합이 아닌 보간(interpolation)으로 얻어야 한다[11-13]. 영역기반 정합은 화소의 명암도나 컬러 정보를 그대로 이용하기 때문에 적용하기 쉽고 모든 화소를 정합하기 때문에 전체적으로 조밀한 변위도를 얻을 수 있지만 영상의 공간 정보(spatial information)를 표현하기 어렵기 때문에 질감이 약하거나 폐색이 발생하는 물체의 경계선 근처에서 오정합을 많이 발생시킨다. 이와 같이 공간 정보의 부족으로 인해 발생하는 영역기반 정합의 단점을 보완하기 위해 폐색 영역이나 경계선 영역을 검출하여 처리할 수 있는 변위 추정 최적화 방법들에 대한 연구들이 활발히 진행 중이다 [14-21].

스테레오 정합의 변위 추정 최적화 방법은 지역적 최적화 방법(local optimization method)[8,15]과 전역적 최적화 방법(global optimization method)[12,13,16-23]으로 분류할 수 있다. 지역적 최적화 방법은 극상선(epipolar line) 상의 점들 중에서 현재의 기준이 되는 점과 가장 일치하는 점을 주어진 변위 탐색 범위 내에서 찾는 방법이다. 대표적인 방법은 일정한 크기의 윈도우 영역에 대한 유사도를 비교하는 지역적 WTA(Winner-Take-All) 방식이다. 지역적 최적화 방법은 유사도 척도와 에너지 함수의 형태가 단순하여 적용하기 쉽고 질감이 풍부한 영역에서 정확한 변위를 얻을 수 있다. 그러나 질감이 없는 영역이나 경계선 영역에서 정확한 변위를 추정하기 어렵다.

전역적 최적화 방법은 스테레오 정합의 유일성(uniqueness), 연속성(smoothness) 등 다양한 제약조건들(constraints)을 이용하여 에너지 모델을 정의하고 영상 전체의 에너지 함수를 최소화시키며 정합문제를 해결하는 방법이다. 여러 연구들의 결과를 통해 전역적 최적화 방법으로 얻은 변위도가 지역적 최

적화 방법으로 얻은 변위도보다 더 정확하다는 사실을 알 수 있다[14-23]. 전역적 최적화 방법의 성능은 수학적 최적화 도구와 밀접한 관련이 있다. 이와 관련하여 마르코프 랜덤 필드(MRF: Markov Random Field)[7], 신뢰확산법(BP: Belief Propagation)[17,18], 그래프 컷(GC: Graph-Cut)[19], 신경회로망(Neural Network)[23,24], 동적 계획법(DP : Dynamic Programming)[25,26] 및 유전 알고리즘(GA: Genetic Algorithm)[20,27-29]을 이용하여 변위를 탐색하는 방법들이 연구되고 있다. 전역적 최적화 방법은 스테레오 영상과 변위도에 대한 에너지 함수를 정의하여 이를 최소화 시키는 해, 즉 전체 변위도를 반복 계산을 통해 얻기 때문에 계산이 복잡하다. 그래서 계산이 복잡하지 않으면서 정확한 변위를 추정할 수 있는 최적화 방법에 대한 연구가 필요하다.

한편 대부분의 스테레오 정합 방법들은 스테레오 영상 획득시 조명 조건과 카메라의 노출 정도가 일정하다는 가정을 토대로 적용된다. 이러한 전제조건에서 기존의 유사도 척도들은 그 각각의 특성에 따라 결과의 차이를 보일 수 있으나 대체적으로 성능이 양호하다. 그러나 두 영상의 광학적인 조건이 일치한다는 전제조건은 엄격한 실험 환경이 아닌 이상 실현하기 어렵고, 실제로 조명의 방향, 조명 색의 변화, 카메라 노출 정도의 차이 및 잡음의 분포가 두 영상에서 일치하게 않는 경우 기존의 스테레오 정합 기법의 성능은 크게 떨어진다[30-35]. 그래서 이러한 환경의 영향에 강인한 유사도 척도와 최적화 방법에 관련된 연구는 스테레오 정합에서 매우 중요하다.

## 1.2 관련 연구

환경 변화의 영향에도 성능이 좋은 스테레오 정합에 대한 연구는 기존의 유사도 척도를 개선하는 방법과 스테레오 영상에서 서로 다른 환경 요인을 제거하는 방법으로 나뉜다. 스테레오 정합에서 발생할 수 있는 대표적인 환경 요인은 잡음, 전역적인 밝기 변화, 지역적인 밝기 변화 등이 있다. 여기서 잡음은

전처리 과정으로 충분히 해결할 수 있으므로 관련 연구 대부분은 밝기 변화에 초점을 맞추고 진행되고 있다[32,35-47].

영상의 특징 정보를 정합요소로 사용하는 특징 기반 유사도 척도는 지역적인 공간 정보를 다양하게 반영할 수 있는 장점이 있다. 이는 영상의 기하학적 특징을 계산하여 얻기 때문에 밝기 변화에 민감하지 않고 좋은 성능을 보인다. 대표적인 특징 기반 유사도 척도는 SIFT(Scale-Invariant Feature Transform) [41,42], SURF(Speed-Up Robust Feature)[43], DAISY[44] 등이 있다. 이들은 보정(calibration) 과정을 거치지 않은 스테레오 영상에 대해서도 우수한 성능을 보이지만 한 화소를 표현하는 특징 벡터의 차원수가 100개 이상이기 때문에 계산량이 많고 검출된 특징점에 대해서만 정합을 수행하므로 전체적으로 조밀한 변위도를 얻기 힘들다.

명암도 기반 유사도 척도는 절대 차의 합(SAD : Sum of Absolute Differences)[2], NCC(Normalized Cross-Correlation)[16,17,19,25], BT (Birchfield and Tomasi)[45,46], AWSM(Adaptive Weight Stereo Matching)[34] 및 상호정보(MI : Mutual Information)[35-38] 등이 있다. SAD와 NCC는 환경적 제약이 엄격할 경우 성능이 비슷하지만 전역적인 밝기 변화에 대해 NCC가 SAD보다 강한 성능을 보인다. NCC는 가장 일반적인 비용 척도로써 대비가 변한 영상에 대해서는 성능이 우수하지만 지역적인 조명의 변화에 강인하지 못하다. BT는 대응점과 주변 화소와의 선형관계를 이용한 유사도 척도이다. 그리고 영상의 아웃라이어에 대한 영향을 줄이기 위해 LoG(Laplacian of Gaussian) 필터를 적용한 BT[45], 평균값 필터를 적용한 BT, Rank 변환을 이용한 BT[46] 등의 방법들이 소개되었다. 이들은 전역적인 밝기 변화에서 좋은 성능을 보이지만 지역적인 밝기 변화에서는 두 영상의 선형 관계가 변하기 때문에 성능이 좋지 않다. MI는 두 영상의 명암도 조합에 대한 확률분포를 이용하는 방법이다. MI는 명암도가 반전되는 등 복잡한 전역적인 변화가 발생한 영상에 대해 NCC나 BT 등 다른 기법들보다 더 좋은 성능을 보이지만 지역적인 변화가 발생한 영상은 다루기 어렵다. AWSM은 임의의 한 화소를 중심으로 형성된 정합 윈도우

(support window)내에서 중심 화소와 주변 화소의 명암도차와 거리를 이용한 가중치를 NCC와 결합하여 지역적인 밝기 변화에 대해 다른 유사도 척도들보다 좋은 성능을 보이지만 경계선 영역의 오정합 문제를 해결하지 못했고 창의 크기가 기존의 NCC보다 커야 성능이 보장되기 때문에 계산 복잡도가 높다.

지역적인 밝기 변화에 의한 오정합을 줄이기 위해 비모수(non-parametric) 특성을 가지는 Rank 변환과 Census 변환을 이용한 유사도 척도들이 소개되었다 [20,47-51]. 이들은 대응점과 그 주변 화소와의 관계를 순위 또는 대소 관계로 표현하여 지역적 구조가 비슷한 영역에 대한 모호함을 줄일 수 있다. 이러한 특성은 조명 조건이 서로 다른 스테레오 영상에서 공간 정보를 반영하는데 이 점 잘 반영할 수 있게 하지만, 상세한 명암도 정보를 사용하지 않기 때문에 지역적으로 구조가 비슷한 영역에서 오정합이 많이 발생한다. 특히 Census 변환의 경우 경계선 영역에서 좋은 성능을 보이지만 질감이 반복적으로 나타나는 영역에서 오정합이 많고 해밍 거리(hamming distance)를 계산하는 과정에서 많은 계산량과 메모리를 요구한다.

스테레오 영상에서 환경 변화의 요인을 찾아 그 영향을 최소화하는 방법들도 연구되고 있다. LTC(Light Transport Constancy)는 지역적인 밝기 변화를 유발한 요인을 극복하는 방법이다[30]. 그러나 각각 다른 환경에서 얻은 스테레오 영상을 두 세트 이상 필요로 한다. 다중 주파수 채널 기법은 채널별 주파수 변환을 이용하여 대비(contrast)의 변화에 강인하지만 지역적인 밝기 변화에서 좋은 성능을 보이지 못한다[39]. 그리고 RGB 컬러 영상에서 두 채널에 대한 로그 크로메티시티(log chromaticity) 변환을 이용하는 방법[40]과 RGB 각 채널에 대해 로그 크로메티시티를 수행하는 방법[32]이 소개되었다. 로그 크로메티시티를 이용한 방법은 영상에서 조명의 영향을 제거하는 방법으로, 지역적인 조명의 영향을 받은 영상에서 기존의 방법들보다 좋은 성능을 보이지만 저대비 영상이나 전역적인 변화가 발생한 영상에서는 정합 성능이 떨어지고, 흑백 영상에는 적용할 수 없는 문제가 있다.

이렇듯 유사도 척도를 개선하는 방법들은 전역적인 밝기 변화에 강인하지만

지역적인 밝기 변화에 대해 좋은 성능을 보이지 못하고, 스테레오 영상에서 환경 변화 요인을 제거하는 방법은 환경 변화가 없는 영상에 대해 적용하기 어렵거나 성능이 떨어지는 문제를 가지고 있다.

### 1.3 연구 내용

본 논문에서는 스테레오 정합을 위해 엔트로피와 Census 기반의 유사도 척도와 다해상도 변위 추정 방법을 제안하였다. 제안한 유사도 척도는 전역적인 밝기 변화에 강인한 엔트로피에 공간적인 정보를 반영하기 위해 밝기 변화에도 특성이 변하지 않는 Haar 웨이블릿 특징 벡터와 Census 변환을 스테레오 정합에 맞게 변형하여 결합하였다. 그리고 변위를 정확히 추정하고 계산 시간의 단축을 위해 8방향 동적 계획법을 이용한 다해상도 변위 추정 기법 제안하였다.

제안한 유사도 척도는 환경 변화에도 대응점의 유사도를 정확히 반영하도록 대응점과 그 주변의 공간적 특성을 반영하는 특징 벡터를 엔트로피와 결합한 형태이다. 특징 벡터는 Haar 웨이블릿 필터 응답을 구하고 주변 화소와의 관계에 대한 대표성을 부여하여 다차원 특징 벡터의 형태로 얻었다. 또한 질감이 없는 영역이나 물체의 경계선 영역 등 특징 벡터를 결합한 엔트로피로 정확히 표현하기 어려운 질감이 없는 영역에서의 정합 성능을 높이기 위해 변형된 Census 변환을 이용하였다. 제안한 Census 변환은 변환 영역을 블럭 단위로 나누고 각 블럭의 분산과 평균값을 이용하여 만들었다.

다해상도 기반의 스테레오 정합 기법은 변위도를 반복적으로 개선하는 과정에서 계산의 효율성을 높이기 위해 사용하였다. 다운샘플링 과정에서 정보의 손실을 줄이기 위해 가우시안 영상 피라미드를 이용한 다해상도 기법을 적용하였고 조밀한 변위도를 얻기 위해 상위 계층에서 얻은 변위도의 정보를 이용하여 각 계층에서 변위 추정 최적화 방법을 적용하였다. 변위 추정 최적화 방법은 8개의 탐색 방향에 대한 누적 에너지를 통합(agggregation)하여 동적 계획법

을 적용하였다.

제안한 유사도 척도는 변위가 급격히 변하는 경계선 영역과 질감이 없는 평탄한 영역에서 정합 성능이 우수하였고 광학적인 조건이 다른 스테레오 영상에 대해서도 대응점간의 유사도를 정확히 반영하였다. 제안한 최적화 방법은 에너지 함수 최적화 과정에서 8방향의 누적 에너지를 통합함으로써 동적 계획법 탐색 방향과 어긋나는 경계선 영역의 변위와 점진적으로 변하는 변위를 잘 추정하였고 기존의 동적 계획법에서 발생하는 스트리킹(streaking) 현상을 개선하였다. 또한 영상 피라미드의 모든 계층에 대해 정합을 수행하여 최상위 계층에서만 정합을 수행하는 기존의 다해상도 스테레오 정합 기법보다 조밀한 변위도를 얻을 수 있었다.

## 1.4 논문의 구성

본 논문의 구성은 다음과 같다. 제 2 장에서는 스테레오 시각과 스테레오 정합에 대한 이론과 기존의 스테레오 정합 기법을 설명하고, 환경 변화에 강인한 스테레오 정합 연구 사례에 대해 기술한다. 제 3 장에서는 Haar 웨이블릿 지역 특징 결합 엔트로피와 변형된 Census 변환을 이용한 유사도 척도에 대해 설명한다. 제 4 장에서는 기존의 동적 계획법과 제안한 8방향 동적계획법에 대해 설명하고 제 5 장에서 제안한 다해상도 기법을 설명하고, 3장과 4장의 내용을 포함한 제안한 다해상도 스테레오 정합 기법을 설명한다. 제 6 장에서 제안한 방법의 결과를 기존의 방법들의 결과와 정량적, 정성적으로 비교하고 성능을 분석하여 고찰하고, 제 7 장에서 전체 내용에 대해 결론을 내린다.

## 제 2 장 스테레오 시각과 스테레오 정합

### 2.1 스테레오 시각

스테레오 시각은 시차가 다른 두 장의 영상에서 대응점 관계에 있는 화소들의 상대 거리를 계산하여 영상의 3차원 정보를 추출하는 과정이다. 카메라를 이용하여 2차원의 영상 평면에 장면을 투영시키면 공간상의 여러 점들이 영상의 한 점에 대응되기 때문에 3차원 깊이(depth) 정보를 잃어버리게 된다. 2차원 영상은 3차원 정보를 표현할 수 없기 때문에 3차원 정보를 복원하기 위해서 같은 장면을 다른 시점에서 얻은 영상이 추가적으로 필요하다. 스테레오 시각은 인간이 두 눈에서 발생하는 시차를 이용해서 원근감을 느끼는 원리를 이용하여 인간의 눈 역할을 하는 두 대의 카메라와 컴퓨터를 이용해 3차원 깊이 인지 능력을 구현하는 기술이다.

스테레오 시각을 이용해 3차원 정보를 얻는 과정은 그림 2.1과 같다. 획득된 스테레오 영상은 카메라 보정과정을 거쳐 저장된다. 보정이 완료된 좌영상(left image)과 우영상(right image)에서 정합 요소를 추출하여 대응점 여부를 판별하고 대응점 간의 변위(disparity)를 추정한다. 변위는 유사도 척도 계산 및 변위 추정 최적화 방법을 이용하여 추정하고, 추정된 변위와 삼각측량법을 이용하여 3차원 깊이 정보를 얻는다. 여기서 대응점 여부 판별과 변위 추정 과정을 스테레오 정합이라 한다.

3차원 공간상의 한 점이 두 영상에 투영되면 두 영상에 존재하는 두 정합점 간의 거리를 변위라고 한다. 변위는 영상 평면에서 물체의 표면까지의 거리인 깊이를 알아내는데 중요한 단서이다. 사람의 두 눈으로 본 시각의 차이를 시차라고 하며, 두 눈에 맺히는 상의 위치 차이인 변위로 해석된다. 사람은 변위로 깊이 정보를 인식할 수 있는데 좌,우 시각 간의 변위가 큰 지점은 가까운 거리에 있는 것이고, 변위가 작은 지점은 먼 거리에 있는 것으로 인식한다.

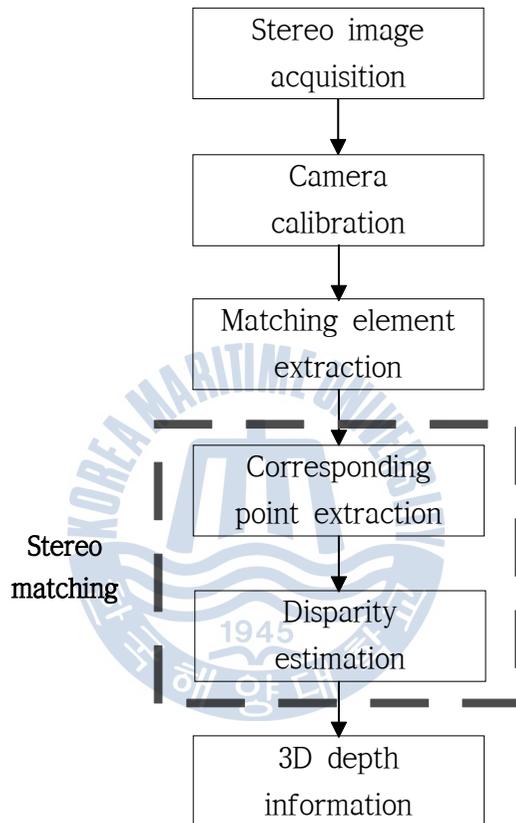


그림 2.1. 3차원 깊이 정보 계산을 위한 스테레오 시각의 순서도

Fig. 2.1. A flowgraph of stereo vision for 3D depth information

스테레오 영상에서 대응점의 변위와 실제 거리와의 관계는 그림 2.2와 같다. 두 개의 카메라로부터 두 영상이 얻어질 때 3차원 공간상의 한 점  $P(x', y', z')$ 가 좌우 영상에  $P_\ell, P_r$ 로 맺히고 그 두 점 간의 변위가  $d$ 이다. 이 때 변위  $d$ 와 3차원 깊이  $z'$ 는 서로 반비례 관계이다.

$C_\ell$ 에서  $C_r$ 를 잇는 선을 기준선(base-line)이라고 하고 그 길이를  $B$ 라고 했을 때, 점  $P$ 의 깊이  $z'$ 는 각 영상의 수평좌표  $x_\ell$ 과  $x_r$ 의 차이인 변위  $d$ 를 이용하여 식 (2.1)과 같이 표현할 수 있다.

$$z' = \frac{Bf}{x_\ell - x_r} = \frac{Bf}{d} \quad (2.1)$$

여기서  $f$ 는 카메라 초점 거리이며, 초점거리와 기준선의 길이를 알고 있다면 변위를 계산함으로써 물체의 깊이 정보인  $z$ 좌표를 구할 수 있다.

식 (2.1)이 성립하기 위해서는 영상을 획득하는 두 렌즈의 광학적, 기하학적 특성이 동일하여 스테레오 영상 간의 기하학적 왜곡이 없고, 극상선(epipolar line)이 수평으로 일치하도록 두 카메라의 좌표계가 정렬되어 있어야 한다.

## 2.2 스테레오 정합

### 2.2.1 유사도 척도

스테레오 영상에서 각 화소들에 대한 대응성 여부를 판별하기 위해서 각 화소를 대표하는 정량화된 정합 요소와 이들의 유사성을 측정하기 위한 척도를 선택하는 것이 중요하다. 정합 요소로 사용할 수 있는 것은 화소의 명암도나 컬러값, 영상에서 계산한 경계나 분할영역 등이다. 두 영상의 화소를 표현하는 정합 요소와 변위의 관계를 이용하여 유사도를 수치화한 것을 유사도 척도

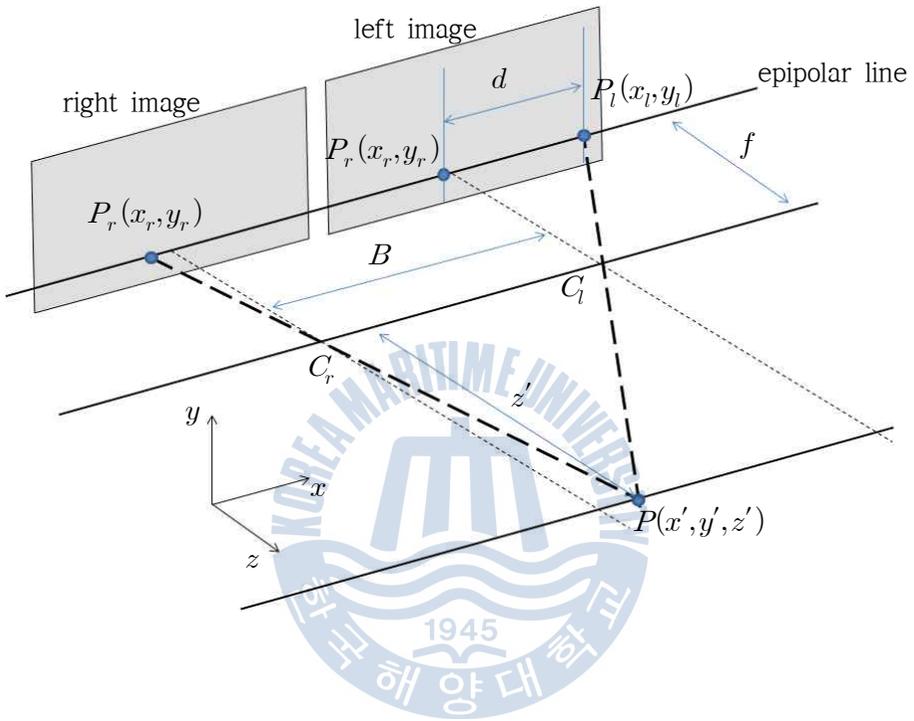


그림 2.2. 스테레오 시각의 카메라 기하학

Fig. 2.2. Camera geometry of stereo vision

(similarity measure)라고 한다.

스테레오 정합을 위해 유사도를 계산하여 깊이 정보를 얻는 과정은 그림 2.3과 같다. 좌영상의 임의의 화소  $p$ 에 대한 대응점을 우영상에서 찾기 위해 좌영상의 화소  $p$ 를 중심으로 하는 윈도우 영역과 크기가 같은 우영상의 여러 후보 영역간의 유사도를 계산하여 오차가 가장 적은 영역의 변위  $D_p$ 를 구한다. 이와 같은 방법으로 모든 화소에 대해 변위를 구하여 변위도를 생성한다.

스테레오 정합은 정합 요소에 따라 특징기반 정합과 영역기반 정합으로 나눌 수 있다. 특징기반 정합은 추출한 영상의 특징이 분포된 공간에서 특징들의 정합이 정확히 이뤄지기 때문에 정확한 변위를 구할 수 있고 조명의 영향이나 잡음에 강인하다. 특징기반 정합은 영상에서 추출된 특징에 대해서만 정합을 수행하므로 정합 속도가 빠른 장점이 있지만 전체적으로 조밀한 변위도를 얻기 어렵다. 이러한 단점은 특징 집합간의 인접관계를 사용하여 보완할 수 있지만 경계나 영역분할 결과의 영향을 많이 받을 뿐만 아니라 특징이 필요 이상으로 많이 검출되는 복잡한 영상에 대해서는 정합이 제대로 이뤄지지 않는다. 따라서 전체적으로 정확한 변위를 구하기 위해 각 화소의 명암도 또는 컬러 정보와 그 주변 화소와의 관계를 반영할 수 있는 영역기반 정합이 주로 사용된다.

유사도 척도는 두 카메라로부터 얻은 좌우 영상의 한 점  $p=(x,y)$ 에 대한 명암도  $I(p)$ 와 변위도  $D_p$ 에 대해 식 (2.2)와 같은 조건을 만족해야 한다.

$$I_L(p) - I_R(p - D_p) = 0 \quad (2.2)$$

여기서  $I_L(p)$ 와  $I_R(p)$ 는 화소  $p$ 에 대한 좌영상과 우영상 좌표의 명암도 또는 컬러값이다.

유사도 척도의 가장 기본적인 형태는 제곱차(SD : Squared Difference), 절대차(AD : Absolute Difference)와 같이 화소단위로 표현된다. 제곱차는 식 (2.3)이고, 절대차는 식 (2.4)와 같다.

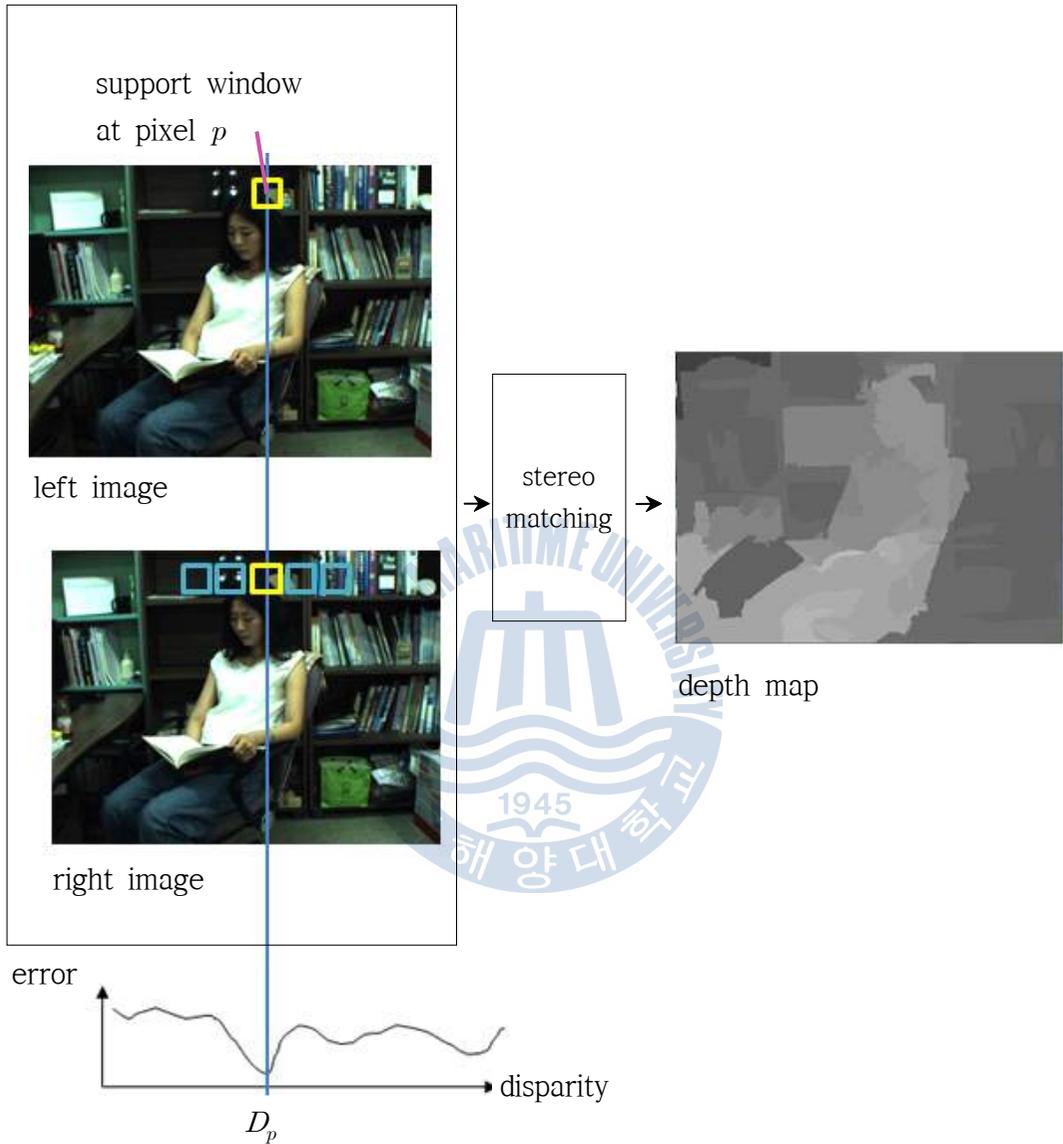


그림 2.3. 스테레오 정합을 위한 유사도 계산과 깊이 정보

Fig. 2.3. Calculating similarity measure and depth map for stereo matching

$$C_{SD}(p, D_p) = [I_L(p) - I_R(p - D_p)]^2 \quad (2.3)$$

$$C_{AD}(p, D_p) = |I_L(p) - I_R(p - D_p)| \quad (2.4)$$

영역기반 정합은 주로 블록 또는 윈도우 단위에 적용된다. 이는 유사도 척도가 그 이웃 화소의 정보를 포함하기 때문이다. 그러므로 제곱차와 절대차에 주변 화소 정보를 포함한 제곱차합(SSD : Sum of Squared Difference)과 절대차합(SAD : Sum of Absolute Difference)이 사용되며 각각 식 (2.5)와 식 (2.6)과 같다.

$$C_{SSD}(p, D_p) = \sum_{q \in N_p} [I_L(q) - I_R(q - D_p)]^2 \quad (2.5)$$

$$C_{SAD}(p, D_p) = \sum_{q \in N_p} |I_L(q) - I_R(q - D_p)| \quad (2.6)$$

여기서  $N_p$ 는 유사도를 구하려는 화소  $p$ 를 포함한 윈도우 내의 주변 화소들이고  $q$ 는 영역  $N_p$ 에 포함되는 화소이다.

정규화된 상호상관(NCC: Normalized Cross-Correlation)은 SAD와 더불어 널리 쓰이는 유사도 척도이다. NCC를 이용하여 구한 화소  $p$ 의 위치에서 변위  $D_p$ 에 대한 유사도 척도는 식 (2.7)과 같다.

$$C_{NCC}(p, D_p) = \frac{\sum_{q \in N_p} I_L(q) I_R(q - D_p)}{\sqrt{\sum_{q \in N_p} I_L(q)^2 \sum_{q \in N_p} I_R(q - D_p)^2}} \quad (2.7)$$

NCC는 전역적인 밝기 변화를 보상할 수 있고 가우시안 잡음에 강인하다. 그러나 변위 불연속 영역에서 변위가 흐려지고, 일정 범위를 벗어난 비정상적인 화소값(outlier)이 존재하는 영역에서 큰 오차를 유발한다. 이는 NCC 뿐만 아니

라 SAD 등 다른 윈도우 기반 기법들을 사용할 때 흔히 나타나는 문제들이다. 이러한 문제를 해결하기 위해 영역의 평균값을 빼주는 ZSAD(Zero-mean Sum of Absolute Difference)와 ZNCC(Zero-mean Normalized Cross-Correlation)를 사용한다. ZSAD와 ZNCC는 각각 식 (2.8)와 식 (2.9)와 같다.

$$C_{ZSAD}(p, D_p) = \sum_{q \in N_p} \left| \{I_L(q) - \bar{I}_L(p)\} - \{I_R(q - D_p) - \bar{I}_R(p - D_p)\} \right| \quad (2.8)$$

$$C_{ZNCC}(p, D_p) = \frac{\sum_{q \in N_p} [I_L(q) - \bar{I}_L(p)] [I_R(q - D_p) - \bar{I}_R(p - D_p)]}{\sqrt{\sum_{q \in N_p} [I_L(q) - \bar{I}_L(p)]^2 \sum_{q \in N_p} [I_R(q - D_p) - \bar{I}_R(p - D_p)]^2}} \quad (2.9)$$

여기서  $\bar{I}_L$ 와  $\bar{I}_R$ 은 좌영상과 우영상 각각의 영역에 대한 평균명암도이고 식 (2.10)과 같다.

$$\bar{I}(p) = \frac{1}{N_W} \sum_{q \in N_p} I(p) \quad (2.10)$$

여기서  $N_W$ 는 화소  $p$ 를 포함한 주변영역  $N_p$ 의 화소수이다. ZSAD와 ZNCC는 영역  $N_p$ 에서 큰 오프셋(offset)이 발생했을 때 그에 대한 오차를 감소시킬 수 있고, ZNCC는 영상에 이득(gain)차가 발생했을 때 그에 따른 차이를 보상해주는 장점이 있다. 그러나 지역적인 밝기 변화에서 성능이 좋지 않고, 경계선 영역이 흐려지는 문제점은 여전히 존재한다.

## 2.2.2 최적화 방법

스테레오 정합의 최적화 방법은 지역적 최적화 방법과 전역적 최적화 방법으로 분류할 수 있다. 지역적 최적화 방법은 일정한 크기의 정합 윈도우 영역 내

에서 유사도를 구하여 비교하는 지역적 WTA 방식이 대표적이다. 지역적 최적화 방법은 질감이 높은 영역에서 정확한 변위를 구할 수 있으나 넓은 영역에 걸쳐 질감이 없으면 정확한 변위를 구하기 어렵다. 이러한 점을 개선하기 위해 적응적으로 윈도우의 크기와 형태를 결정하여 정합하는 방법이 소개되기도 하였다. 전역적 최적화 방법은 변위도의 정확성을 높이기 위해 스테레오 정합의 유일성, 연속성 등의 제약조건들을 포함한 에너지 모델을 정의하고, 영상 전체의 에너지 함수를 최소화시키는 변위도를 구하는 방법이다.

전역적 최적화 방법의 성능은 최적화 도구와 밀접한 관련이 있다. 그래서 동적 계획법, 신경 회로망, 그래프 컷, 신뢰 확산법 및 유전 알고리즘 등을 이용한 사례가 많이 있다. 전역적 최적화 방법에는 신뢰 확산법과 그래프 컷과 같이 베이시안 추론을 적용하여 유도된 에너지 함수를 이용하는 방법과 유전 알고리즘과 동적 계획법 등 정합 문제를 최소화 문제와 정규화 문제로 표현하여 반복법으로 변위를 구하는 방법이 있다.

전역적 최적화 방법은 지역적 최적화 방법에서 나타나는 정합 윈도우로 인한 오정합이 발생하지 않고 비교적 정확한 변위도를 얻을 수 있다. 하지만 전역적 최적화 방법 자체에 폐색 영역을 다루거나 추정할 수 있는 과정이 없으므로 별도의 폐색 영역을 추정하는 알고리즘이 필요하다. 이 과정은 영상 전체에 걸쳐 에너지를 최소화하는 과정과 더불어 변위 추정 단계에서 많은 계산량을 요구한다.

전역적 최적화 방법에서 최적화하는 에너지 함수는 전체 변위도에 대한 유사도 척도 에너지항과 변위 연속성 에너지항의 합으로 구성되며 식 (2.11)과 같다.

$$E(D) = E_{data}(D) + E_{smooth}(D) \quad (2.11)$$

여기서  $D$ 는 변위도이다.  $E_{data}(D)$ 와  $E_{smooth}(D)$ 는 각각 변위도  $D$ 에 대해 계산

되는 유사도 척도 에너지항과 변위 연속성 제약조건에 의한 에너지항이다. 전역적 최적화 방법의 목표는 에너지 함수  $E(D)$ 를 최소로 만드는 변위도  $D$ 를 구하는 것이다.

신뢰 확산법은 2차원 그래프 모델에서 각 노드들이 이웃하는 노드들과 확률적인 정보를 반복적으로 상호 교환하여 최적화된 상태를 추정하는 방법이다. 한 화소에 할당된 변위가 그 화소의 변위가 될 확률은 그 변위가 이웃 화소들의 변위가 될 확률들을 고려하여 결정한다. 현재 반복 횟수  $t$ 에서 현재 화소  $p$ 의 변위가  $D_p$ 일 때, 이웃 화소  $q$ 와 상호 교환하는 확률 정보를 신뢰도  $m_{pq}^t(D_p)$ 라고 하며 식 (2.12)와 같다.

$$m_{pq}^t(D_p) = \min \left[ C_{data}(D_p) + V(D_p, D_q) + \sum_{s \in N_p} m_{sq}^{t-1}(D_p) \right] \quad (2.12)$$

여기서  $C_{data}(D_p)$ 는 변위  $D_p$ 에 대한 유사도 척도 에너지항이고,  $V(D_p, D_q)$ 는 연속성을 나타내는 에너지항이다. 이 때 주로 반복적인 확산을 이용하는데 확산이 일어나면서 각 변위들의 확률이 다른 화소로 전달된다.

신뢰 확산법을 이용하는 스테레오 정합은 영상 전체의 화소를 이용하기 때문에 폐색 영역에 대한 오정합이 낮은 편이다. 그러나 여러 단계에 걸쳐 반복적으로 정합을 수행해야하고 모든 화소에 대해 주어진 변위 탐색 범위 내의 변위들의 확률들을 모두 고려해야 하므로 계산 복잡도가 높고 메모리 점유율이 증가하여 과부하(overhead)를 많이 발생시킨다. 그리고 무질감 영역에 있는 화소들이 주변 화소에 영향을 받아 잘못된 정보를 전달하면 최종적으로 오정합이 발생할 수 있다.

그래프 컷은 마르코프 랜덤 필드(MRF: Markov Random Field)와 그래프 이론을 이용한 최적화 방법이다. 스테레오 영상의 모든 화소를 변위를 이용해 라벨링(labeling)하여 화소들의 라벨, 즉 변위가 같은지 여부를 확률로 판별해 에너지 함수를 최소화하는 라벨들을 연결하는 방식이다. 이 때 라벨이 다른 확률이

높은 화소는 끊어내어(cut) 해당 화소의 라벨은 다음 단계 처리에서 고려하지 않는다. 그래프 컷으로 얻은 변위도의 특징은 경계 화소를 기준으로 변위가 달라지기 때문에 일반적으로 영상 분할(segmentation) 기법과 함께 적용한다. 그래프 컷을 이용한 방법은 경계선에서 변위의 정확도가 높으나 물체의 경계가 아닌 단순한 에지 화소에서도 변위값을 다르게 구하기 쉽기 때문에 물체의 세세한 부분에서 정합 성능이 좋지 않고, 경계 영역에서도 연속성 제약으로 인해 뭉개지는 현상이 나타날 수 있다.

동적 계획법은 최적화 문제에 많이 적용되는 방법이다. 동적 계획법은 이전에 얻은 해를 반복적으로 참고하는 최적화 문제를 해결하는데 효율성이 입증된 방법이다. 특히 변위의 유일성과 순서성 제약조건을 이용해 적은 계산량으로 폐색 영역 추정을 동시에 할 수 있기 때문에 스테레오 정합에 많이 적용된다. 그러나 동적 계획법은 탐색 방향과 어긋나는 경계선 영역에서 줄무늬 형태의 스트리킹 현상이 발생하는 단점이 있다.

## 2.3 환경 변화에 강인한 유사도 척도

### 2.3.1 특징 기반 유사도 척도

특징 기반 스테레오 정합은 밝기 변화에 강인한 기하학적 특징을 정합요소로 이용하는 방법이다[52-54]. 이러한 정합 요소는 특징점과 그 주변의 공간적 특징을 서술하기 때문에 특징점에 대한 정확한 정합이 가능하다는 장점이 있다. 최근에는 변위가 커지면서 각 카메라에 대해 기하학적인 왜곡이 발생한 기준선이 긴 스테레오 시각에서 주로 특징 기반 기법을 사용한다. 특히 이와 같은 경우 기존의 영역기반 정합으로는 유사도를 정확히 반영하기 어렵기 때문에 대응점을 크기와 방향 성분을 가지는 벡터의 형태로 표현하는 지역 서술자(local descriptor)를 주로 사용한다[41-44].

지역 서술자는 영상의 크기 변화 또는 회전과 같은 기하학적인 변화에도 동일한 영역에 대해 동일한 서술자를 생성하며, 지역적인 서술에 기반하고 있다는 특성 때문에 시점의 변화가 발생한 영상 또는 가려짐이 발생한 영상에서도 강인한 정합을 가능케 한다. 이런 장점을 이유로 지역 서술자는 패턴 인식(pattern recognition), 컴퓨터 시각, 컴퓨터 그래픽스 분야에서 폭넓게 활용되고 있으며, 컴퓨터 시각의 세부 분야 중 물체 인식(object recognition), 영상 검색(image retrieval) 및 스테레오 정합에 이용되고 있다. 스테레오 정합에 이용되는 대표적인 지역 서술자는 SIFT, DAISY, SURF가 있다. 이들 중 Lowe에 의해 제안된 SIFT가 크기, 회전 및 시점 변화와 같은 기하학적 변화에 대해서 다른 서술자들에 비해서 성능이 상대적으로 안정된 것으로 알려져 있다.

SIFT는 영상에서 DoG(Difference of Gaussian) 필터를 이용하여 키포인트를 찾아 그 점을 중심으로  $8 \times 8$ 영역에 대해 각각의 화소에 대해 기울기를 계산하여 기울기의 방향, 즉 각도에 대한 히스토그램을 생성한다. 생성된 히스토그램을 통해 방향 정규화 과정을 거쳐 단위 벡터를 계산한다. 기울기를 구했던 영역을  $4 \times 4$ 크기의 부영상(sub-image) 블록으로 분리하여 각 영역에 대해 기울기 성분을 계산한다. 이와 같은 방법으로  $4 \times 4 \times 8$ 크기의 SIFT 지역 서술자를 생성된다. SIFT는 크기 변환에 대한 요소들을 포함하기 때문에 영상 피라미드를 만든 후 각 계층마다 특징 벡터를 추출한다.

지역 서술자를 이용한 방법은 영상의 회전, 크기변화 및 조명의 변화에 강인하기 때문에 기준선이 긴 환경에서 좋은 성능을 보인다. 그러나 지역 서술자의 차원이 100차원 이상이기 때문에 영상의 모든 화소에 대해 적용할 경우 기존의 명암도 기반 스테레오 정합 기법에 비해 계산시간이 오래 걸리고, 검출된 특징 점에 대해서만 정합을 수행하므로 전체적으로 조밀한 변위도를 얻기 힘들다.

### 2.3.2 명암도 기반 유사도 척도

스테레오 영상의 밝기가 서로 다를 경우 명암도 기반의 스테레오 정합은 대

응점의 유사도를 잘 반영하지 못한다. 이를 해결하기 위해 기존의 유사도 척도에 공간 정보를 적용하거나 분포 확률이나 주변 화소와의 대소 비교를 통한 변환 등 다양한 형태의 유사도 척도들이 소개되었다.

BT(Birchfield and Tomasi)는 카메라 표본화에 강인한 유사도 척도이다. 이는 대응점의 이웃 화소들과의 선형보간이 가장 큰 값에 대해 절대차로 표현된다. BT를 이용한 유사도 척도는 두 영상의 대응화소와 주변 화소와의 선형관계를 각각 계산한 후 다른 영상의 대응 화소값에서 뺀 것 중 가장 작은 것을 택하여 얻는다. 이는 식 (2.13)과 같다.

$$C_{BT}(p, D_p) = \min(A, B) \quad (2.13)$$

여기서  $A$ 는 좌영상의 화소  $p$ 와 대응되는 우영상의 화소의 주변 선형관계에 대한 차이들 중 최대값을 의미하고,  $B$ 는 우영상의 화소와 좌영상 화소  $p$ 의 주변 선형관계에 대한 차이들 중 최대값을 나타낸다. 이를 식으로 표현하면 각각 식 (2.14)와 식 (2.15)와 같다.

$$A = \max[0, I_L(p) - I_R^{\max}(p - D_p), I_R^{\min}(p - D_p) - I_L(p)] \quad (2.14)$$

$$B = \max[0, I_R(p - D_p) - I_L^{\max}(p), I_L^{\min}(p) - I_R(p - D_p)] \quad (2.15)$$

$A$ 와  $B$ 를 얻기 위한 화소  $p$ 의 극값  $I^{\min}(p)$ 와  $I^{\max}(p)$ 는 화소  $p$ 의 주변 화소와의 평균값을 이용하여 구한다. 이는 각각 식 (2.16)과 식 (2.17)과 같다.

$$I^{\min}(p) = \min[I^-(p), I(p), I^+(p)] \quad (2.16)$$

$$I^{\max}(p) = \max[I^-(p), I(p), I^+(p)] \quad (2.17)$$

여기서 이전 화소와의 평균값  $I^-(p)$ 와 다음 화소와의 평균값  $I^+(p)$ 는 각각 다

음과 같다.

$$I^-(p) = \frac{[I(p-1) + I(p)]}{2} \quad (2.18)$$

$$I^+(p) = \frac{[I(p+1) + I(p)]}{2} \quad (2.19)$$

BT는 주변 화소와의 선형관계를 이용하므로 기존의 윈도우 기반의 유사도 척도보다 계산량이 적고 잡음과 전체적인 밝기 변화에 강인한 성능을 보인다. 그러나 정확하지 않은 이전 변위의 영향으로 변위도 전체의 정확도가 감소할 수 있고, 밝기 변화의 폭이 크거나 지역적인 밝기 변화가 발생했을 때 성능이 좋지 않다.

AWSM[34]는 기존의 NCC에 공간 정보를 적용한 유사도 척도이다. AWSM에서 사용하는 공간 정보는 NCC를 적용하는 윈도우의 위치에 따른 가중치로 표현된다. 현재 화소  $p$ 점을 중심으로 하는 윈도우에 포함된 화소  $q$ 점의 가중치는 두 화소의 거리와 명암도 또는 컬러값에 대해 계산되며 이는 식 (2.20)과 같다.

$$w(p, q) = \exp\left(-\frac{\|p - q\|^2}{2\sigma_d^2} - \frac{\|I(p) - I(q)\|^2}{2\sigma_s^2}\right) \quad (2.20)$$

여기서  $\sigma_d$ 와  $\sigma_s$ 는 각각 공간적 거리와 명암도 또는 컬러값의 가중치에 대한 표준편차이다. 가중치는 CIE\_Lab 색공간에 대해 계산되어 식 (2.21)의 AWSM에 이용된다.

$$C_{AWSM}(p, D_p) = \frac{\sum_{q \in N_p} w_L(q)w_R(q - D_p)I_L(q)I_R(q - D_p)}{\sqrt{\sum_{q \in N_p} [w_L(q)I_L(q)]^2 \sum_{q \in N_p} [w_R(q - D_p)I_R(q - D_p)]^2}} \quad (2.21)$$

AWSM은 화소의 명암도와 위치 정보를 혼합한 가중치를 사용하기 때문에 기존의 상관도 기반 유사도 척도에 비해 밝기 변화에 강인한 성능을 보인다. 그러나 윈도우의 영향으로 인한 경계선 영역에서 오정합이 발생한다.

MI는 영상의 명암도 분포에 대한 결합확률을 이용하여 두 영상의 상호 정보를 계산한다. MI는 전역적인 변형이 발생한 영상에서 대응점을 찾는 문제에 대해서 좋은 성능을 보이기 때문에 영상 정합, 스테레오 정합 등 다양한 컴퓨터 비전 문제를 해결하기 위해 사용되었다. Viola와 Wells[35]은 자기공명영상(MRI: Magnetic Resonance Image)과 컴퓨터 단층촬영(CT: Computed Tomography) 영상을 정합하는데 MI를 적용하였다. Egnal[36] 등은 지역적 정합 윈도우를 이용하여 변위를 찾는데 MI를 적용했으나, 불연속하거나 질감이 적은 영역에서 고정 크기의 윈도우가 갖는 문제를 해결하지 못했다. 이런 문제점을 개선하기 위해 Kim 등[37]은 전역적 최적화 기법에 MI를 화소단위(pixel-wise)로 근사화하는 방법을 제안하였다. 이들은 기존의 전체 영상에서 전역적으로 계산되는 MI를 Taylor 급수전개와 Parzen 윈도우 추정(Parzen window estimation) 기법을 이용하여 화소단위 MI의 합으로 근사화하였다. 그러나 MI는 대응점을 찾을 때 명암도의 통계적인 특성만을 활용하고 화소들의 공간 정보는 활용하지 않기 때문에 지역적인 밝기 변화가 발생할 경우 성능이 좋지 않다.

### 2.3.3 비모수 유사도 척도

Rank 변환과 Census 변환은 비모수(non-parametric) 특성을 가진다. Rank 변환은 대응점과 그 주변 화소와의 관계를 순위로 변환하고 좌영상과 우영상의 순위에 대한 차이를 유사도 계산에 사용하는 방법이다. Rank 변환은 명암도 분포에 대해 직접 적용할 수 있으나, 주로 SAD나 NCC 등의 다른 유사도 척도에 의한 결과에 적용하는 사례가 많다. Rank 변환은 데이터의 통계적 특성을 사용하기 때문에 아웃라이어나 잡음으로 인해 최빈값이 발생할 경우 강인하고 전체적인 명암도의 변화가 발생한 스테레오 영상에서 좋은 성능을 보인다. 그러나

대비가 급격히 축소된 영상이나 지역적인 밝기 변화에서는 성능이 좋지 않다.

Census 변환은 대응점과 그 주변 화소의 관계를 대소 비교를 통해 이진 비트로 변환하고 유사도를 계산하는 방법이다. 이러한 과정을 통해 생성된 이진 비트는 대응점 주변의 구조적 정보를 표현하는 특징이 있다. 그리고 정합 요소들은 0또는 1의 값으로만 표현되기 때문에 화소값을 그대로 이용해서 유사도를 계산할 때 발생하는 모호함을 방지할 수 있고 경계선 등 구조적인 특징이 있는 영역의 유사도를 잘 반영하는 장점이 있다.

Rank 변환과 Census 변환은 아웃라이어나 잡음과 같은 요인에 강인하고 변위가 급격히 변하는 영역에서 성능이 우수하다. 그러나 명암도의 상대적인 관계만을 사용하기 때문에 정보를 낭비하게 되고, 잡음의 영향으로 중심화소가 변하거나 지역적인 밝기 변화와 대비가 축소된 영상에 대해 성능이 좋지 않다.



### 제 3 장 엔트로피 및 Census 기반의 유사도 척도

본 장에서는 엔트로피 및 Census 기반의 유사도 척도를 제안한다. 스테레오 정합에서 대응점과 주변 화소의 관계를 공간적으로 표현하기 위해 Haar 웨이블릿 특징 벡터와 Census 변환을 이용한 유사도 척도를 확률기반 유사도 척도인 엔트로피에 결합하고자 한다. 엔트로피를 이용하는 기존의 확률 기반 유사도 척도인 MI는 공간 정보를 활용하지 않기 때문에 경계선 영역의 유사도를 정확히 반영하지 못하고 지역적으로 불균일한 밝기 변화가 발생했을 때 성능이 크게 떨어진다. Haar 웨이블릿 특징 벡터는 대응점의 공간 정보를 주파수 개념으로 표현하고, Census 변환은 대응점을 주변 화소와의 대소 관계로 표현하기 때문에 밝기가 변한 환경에서도 공간적인 특징을 잘 표현할 수 있다. 제안한 유사도 척도는 이러한 공간 정보들을 엔트로피 유사도 척도와 결합하여 두 영상의 환경 조건이 달라도 대응점의 유사도를 정확히 반영할 수 있다.

제안한 유사도 척도는 Haar 웨이블릿 특징 벡터와 엔트로피를 결합한 유사도 척도 및 Census 변환을 이용한 유사도 척도로 구성된다. 이 중 Haar 웨이블릿 특징 벡터와 결합한 엔트로피 유사도는 최초 변위도에 대한 확률 분포를 이용하여 화소 단위 엔트로피 유사도를 추정한 후 Haar 웨이블릿 필터 응답을 기반으로 얻은 특징 벡터 유사도와 곱하여 얻는다. Census 변환은 지역적으로 불균일한 밝기 변화에도 좋은 성능을 보이도록 변형하여 유사도를 얻고, Haar 웨이블릿 특징 벡터와 결합한 엔트로피 유사도와 가중합한다. 제안한 유사도 척도를 계산하는 과정은 그림 3.1과 같다. 화소 단위 엔트로피 유사도는 주어진 최초 변위도에 대한 스테레오 영상의 명암도 분포를 기반으로 추정한다. 이 때 최초 변위도는 무작위로 생성할 수 있다. 화소 단위 엔트로피 유사도를 계산한 후 모든 화소에 대해 Haar 웨이블릿 특징 벡터를 추출하여 유사도를 구한다. 그리고 변형된 Census 변환에 대한 유사도를 계산하여 다른 유사도 척도들과 결합함으로써 제안한 유사도 척도를 얻는다.

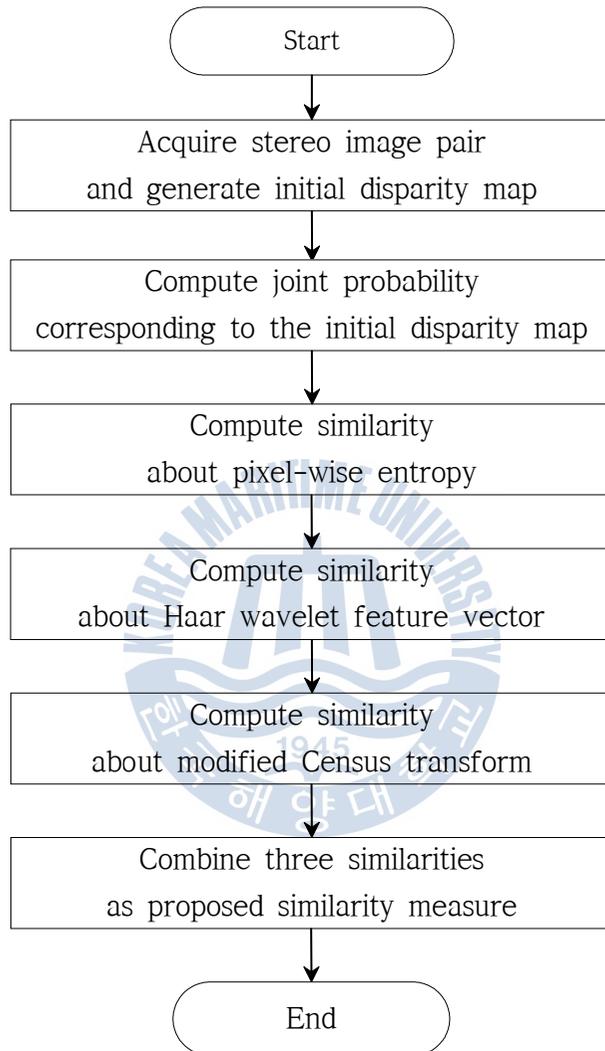


그림 3.1. 제안한 유사도 척도의 계산 순서

Fig. 3.1. A flowgraph of computing proposed similarity measure

### 3.1 엔트로피 기반의 유사도 척도

#### 3.1.1 엔트로피

어떤 사건  $a$ 의 발생확률이  $p(a)$ 일 때 해당 사건의 정보량(information content)은 식 (3.1)과 같다[55].

$$IC(a) = -\log p(a) \quad (3.1)$$

이산확률변수  $A$ 의 각 사건에 대한 정보량을 발생확률에 따라 가중 평균한 것을 평균정보량(average information content) 또는 엔트로피(entropy)라고 하며 식 (3.2)와 같다.

$$H(A) = -\sum_a p(a) \log p(a) \quad (3.2)$$

정보량의 경우 두 사건  $a$ 와  $b$ 가 동시에 발생할 확률과 독립적으로 발생할 확률의 비율을 상호정보량(mutual information)이라고 하며 식 (3.3)과 같다.

$$MI(a,b) = \log \frac{p(a,b)}{p(a)p(b)} \quad (3.3)$$

여기서  $p(a,b)$ 는 사건  $a$ 와  $b$ 에 대한 결합확률이다.

엔트로피에 대한 상호정보량은 식 (3.3)의 정보량에 대한 상호정보량을 동시 발생 사건의 확률을 곱한 평균값으로, 평균 상호정보량이라고 한다. 평균 상호정보량은 식 (3.4)와 같다.

$$MI(A, B) = \sum_a \sum_b p(a, b) \log \frac{p(a, b)}{p(a)p(b)} \quad (3.4)$$

평균 상호정보량은 log 연산을 합으로 풀면 식 (3.5)와 같다.

$$MI(A, B) = - \sum_a \sum_b p(a, b) \log p(a) - \sum_a \sum_b p(a, b) \log p(b) + \sum_a \sum_b p(a, b) \log p(a, b) \quad (3.5)$$

식 (3.5)의 우변의 항들은 식 (3.2)를 이용해 엔트로피로 나타낼 수 있다. 각 항을 순서대로 엔트로피로 나타내면 식 (3.6)과 같다.

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (3.6)$$

여기서  $H(A)$ 와  $H(B)$ 는 각각 확률변수  $A$ 와  $B$ 에 대한 엔트로피이고,  $H(A, B)$ 는 확률변수  $A$ 와  $B$ 에 대한 결합 엔트로피이다.

### 3.1.2 엔트로피를 이용한 MI 유사도 척도

스테레오 정합에서 평균 상호정보량(이하 MI)을 이용한 유사도 척도는 좌영상  $I_L$ 과 변위도  $D$ 에 의해 화소들의 좌표가 변형된  $I_R$ 의 명암도 분포에 대한 에너지 함수로 표현된다. 이는 식 (3.7)과 같다.

$$E_{data}(D) = -MI(I_L, I_R, D) = -H(I_L) - H(I_R) + H(I_L, I_R, D) \quad (3.7)$$

여기서 MI의 부호가 반대인 이유는 평균 상호정보량이 최대일 때 에너지가 최

소가 되어야하기 때문이다.  $H(I)$ 는 영상  $I$ 의 엔트로피이고  $H(I_L, I_R, D)$ 는 주어진 변위도  $D$ 에 대한 전체 스테레오 영상  $I_L$ 과  $I_R$ 의 엔트로피이다. 명암도 범위  $[0, L_{\max}]$ 에 대한 두 엔트로피는 각각 식(3.8)과 식(3.9)와 같다.

$$H(I) = - \int_0^{L_{\max}} P(i) \log(P(i)) di \quad (3.8)$$

$$H(I_L, I_R, D) = \int_0^{L_{\max}} \int_0^{L_{\max}} P(i_L, i_R, D) \log(P(i_L, i_R, D)) di_L di_R \quad (3.9)$$

$P(i)$ 는 명암도  $i$ 의 주변확률(marginal probability)이다. 결합확률  $P(i_L, i_R, D)$ 은 좌영상 화소의 명암도  $i_L$ 과 주어진 변위도  $D$ 에 대해 와핑(warping)된 우영상 화소의 명암도  $i_R$ 간의 결합확률을 의미한다.

## 3.2 제안한 Haar 웨이블릿 특징을 결합한 엔트로피 유사도 척도

### 3.2.1 화소 단위 엔트로피

MI를 이용한 유사도 척도는 좌영상과 우영상의 명암도쌍에 대한 도수를 구하고 전체 화소수에 대한 비율을 얻은 후 엔트로피를 계산한다. 이는 계산과정에서 많은 시간을 소요하기 때문에 제안한 방법은 화소 단위로 엔트로피를 추정하는 방법을 사용하였다[34].

두 영상의 MI는 식 (3.7)을 이용해 구할 수 있다. 식 (3.7)에서 좌영상의 주변 엔트로피  $H(I_L)$ 와 우영상의 주변엔트로피  $H(I_R)$ 는 두 영상의 시차가 크지 않다는 전제하에 변위도  $D$ 의 영향을 거의 받지 않는 일정한 상수를 갖는다고 볼 수 있다. 따라서 좌영상과 우영상에 대한 주변 엔트로피를 제외한 유사도 척도  $MI_{joint}(I_L, I_R, D)$ 는 식 (3.10)과 같이 두 영상과 변위도에 대한 결합엔트로피로 표

현할 수 있다.

$$MI_{joint}(I_L, I_R, D) = -H(I_L, I_R, D) \quad (3.10)$$

엔트로피  $H(I_L, I_R, D)$ 는 식 (3.9)와 같이 좌영상과 우영상의 명암도 쌍에 대한 결합확률을 구하여 얻을 수 있다. 여기서 결합확률  $P(i_L, i_R, D)$ 은 좌영상 화소의 명암도  $i_L$ 과 주어진 변위도에 대해 와핑된 우영상 화소의 명암도  $i_R$ 간의 대응하는 화소들의 수를 카운트하여 계산한다. 예를 들어 현재 좌영상의 화소  $p$ 의 명암도가 100이고 이 화소의 변위  $D_p$ 만큼 이동한 우영상의 화소 위치의 명암도가 110일 경우, 명암도쌍  $(i_L, i_R) = (100, 110)$ 의 도수가 1만큼 증가한다. 이와 같은 과정을 전체 영상의 모든 화소에 대해 수행하여 얻은 명암도쌍  $(i_L, i_R)$ 의 결합확률은 식(3.11)과 같다.

$$P(i_L, i_R, D) = \frac{1}{n_c} \sum_p T[(i_L, i_R) = (I_L(p), I_R(p - D_p))] \quad (3.11)$$

여기서 연산자  $T[\cdot]$ 는 괄호안의 조건, 즉 주어진 명암도쌍이 좌영상과 우영상의 명암도쌍과 일치할 경우 1을, 아닐 경우 0을 갖는다.  $n_c$ 는 스테레오 영상의 총 대응화소수이다.

엔트로피  $H(I_L, I_R, D)$ 는 변위도  $D$ 가 변할 때마다 전체 영상에 대해 계산해야 하므로 계산복잡도가 높다. 이는 화소단위로 엔트로피를 추정하는 방법을 통해 해결할 수 있다. 스테레오 영상의 엔트로피는 식 (3.12)와 같이 영상의 각 화소에서 계산한 화소단위 엔트로피의 합으로 근사화할 수 있다.

$$H(I_L, I_R, D) \approx \sum_p h\{I_L(p), I_R(p - D_p)\} \quad (3.12)$$

여기서  $D_p$ 는 좌영상의 화소  $I_L(p)$ 의 변위이다. 엔트로피를 유사도 척도로 사용한다면 화소  $p$ 의 화소단위 에너지 비용  $C(p, D_p)$ 는 식(3.13)과 같이 화소  $p$ 의 화소단위 엔트로피로 표현할 수 있다.

$$C(p, D_p) = h\{I_L(p), I_R(p - D_p)\} \quad (3.13)$$

여기서  $h\{I_L(p), I_R(p - D_p)\}$ 는 화소단위 엔트로피이고 이전 반복순번에서 얻은 변위도  $D_0$ 에 대한 확률분포를 기초로 Taylor 급수 전개와 Parzen 추정 기법을 이용하여 추정할 수 있으며 식(3.14)와 같다.

$$\begin{aligned} & h\{I_L(p), I_R(p - D_p)\} \quad (3.14) \\ & = -\frac{1}{n_c} \log P(I_L(p), I_R(p - D_p), D_0) * G\{I_L(p), I_R(p - D_p)\} * G\{I_L(p), I_R(p - D_p)\} \end{aligned}$$

여기서  $G\{I_L(p), I_R(p - D_p)\}$ 는 중심을  $(I_L(p), I_R(p - D_p))$ 에 둔 2차원 가우시안 함수이고  $P(I_L(p), I_R(p - D_p), D_0)$ 는 이전 반복순번에서 얻은 변위도  $D_0$ 를 이용하여 얻은 확률분포를 토대로 계산한 현재 할당된 변위  $D_p$ 에 대한 결합확률이다. 그리고  $*$ 은 컨볼루션 연산자이다.

앞에서 설명한 스테레오 영상의 엔트로피는 좌영상의 명암도 범위와 우영상의 명암도 범위를 축으로 하는 명암도 평면에서 표현할 수 있다. 조명 조건 등 환경 조건이 동일하고 시차가 없는 영상 두 개에 대한 결합확률분포는 명암도 평면에서 대각선 방향으로 일직선으로 나타날 수 있다. 그리고 시차가 발생한 스테레오 영상에서 각 화소의 변위가 정확할 경우 이상적인 일직선은 아니더라도 대각선 방향으로 분포가 집중된다. 변위도에 따라 변하는 스테레오 영상의 결합확률분포의 예는 그림 3.2와 같다. 그림 상단의 스테레오 영상에 서로 다른 변위도  $D_a$ 와  $D_b$ 에 대해 얻을 수 있는 확률분포는 각각  $P(I_L, I_R, D_a)$ 과  $P(I_L, I_R, D_b)$ 와 같이 나타난다. 변위도  $D_a$ 는 상대적으로  $D_b$ 에 비해 정확도가 떨

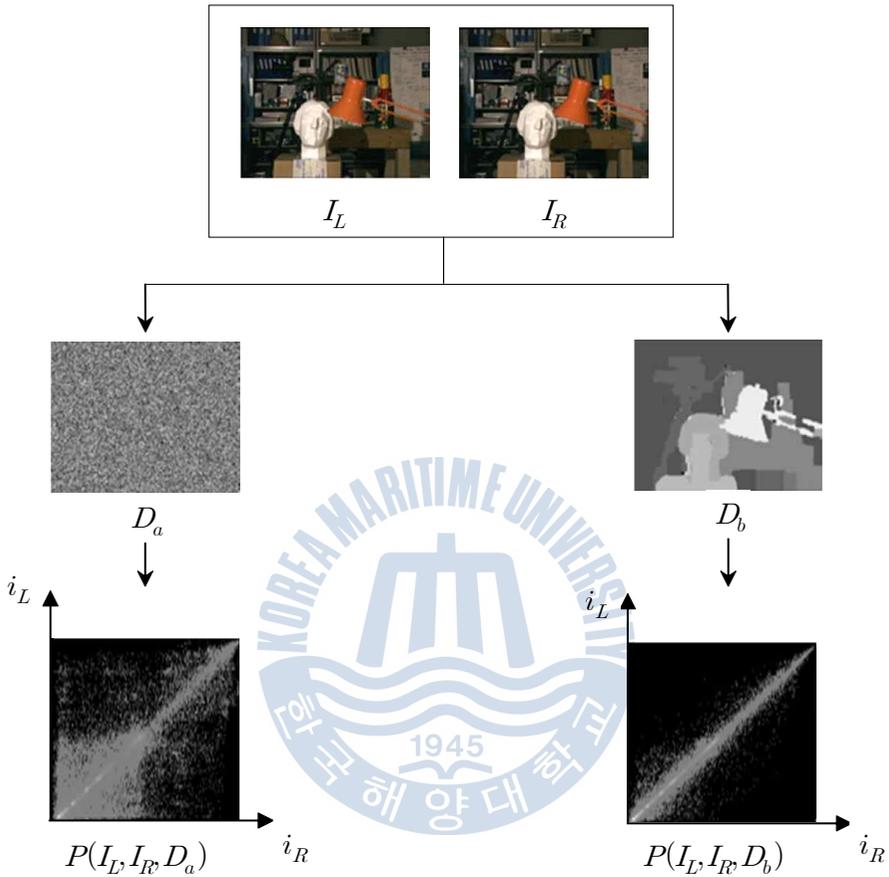


그림 3.2. 두 변위도  $D_a$ 와  $D_b$ 에 대한 스테레오 영상의 결합확률분포

( $I_L$ 과  $I_R$ 은 Middlebury Stereo Vision Lab에서 공개)

Fig. 3.2. Joint probability distribution of stereo image pair corresponding to two disparity maps,  $D_a$  and  $D_b$

(Image courtesy of  $I_L$  and  $I_R$  Middlebury Stereo Vision Lab)

어진다.  $D_a$ 에 대한 스테레오 영상의 명암도 분포  $P(I_L, I_R, D_a)$ 는 그림 3.2의 좌측하단과 같이 퍼져있는 분포 특성을 갖는다. 이러한 분포에서 엔트로피는 높게 계산된다.  $D_b$ 에 대해 얻은 명암도 분포  $P(I_L, I_R, D_b)$ 는 대각선 방향으로 분포가 집중되어 있어 낮은 엔트로피로 나타난다. 따라서 주어진 변위도의 정확도에 따라 스테레오 영상의 명암도 분포의 형태가 달라짐을 알 수 있다.

제안한 방법의 화소단위 엔트로피 유사도 척도는 스테레오 영상과 변위도가 주어졌을 때 화소 단위 결합확률과 엔트로피를 이용하여 계산되며 그림 3.3과 같다. 좌영상과 변위도  $D$ 에 의해 와핑된 우영상의 명암도쌍에 대한 분포를 구하고 식 (3.14)와 같이 가우시안 커널의 컨볼루션 연산을 통해 결합확률을 추정하여 엔트로피를 계산한다. 이 때 주어진 변위도  $D$ 의 정확도가 떨어질수록 그림 3.2의  $P(I_L, I_R, D_a)$ 와 같이 전체적으로 고르게 분포된 형태를 띤다.

기존의 MI는 조명 자체가 변하거나 지역적으로 밝기가 변한 환경에서 좋은 성능을 보이지 못한다. 그 이유는 MI는 전역적 변화가 발생했다는 전제로 결합확률밀도함수를 이용해서 대응점들을 찾기 때문이다. 카메라의 노출 시간의 변화로 인해 발생할 수 있는 전역적인 변화의 경우, 기존의 MI 기반의 비용 함수를 적용하면 비교적 정확한 변위도를 만들어낸다. 전역적인 변화가 발생한 경우, 동일한 결합확률에 대해 비선형적인 변화가 두 이미지 전역에 걸쳐 존재하기 때문이다.

지역적인 밝기 변화가 발생했을 때, 기존의 MI 기반 방법으로는 좋은 결과를 얻을 수 없다. 이러한 경우, 결합확률이 매우 고르게 퍼져서 분포하기 때문에 입력 스테레오 영상들의 전역적인 관계를 정확히 찾을 수 없다. 전역적인 밝기 변화가 발생한 Aloe 영상과 MI의 변위도는 그림 3.4와 같다. 카메라 노출 시간을 달리하여 얻은 좌영상과 우영상에 대해 기존의 MI를 이용하여 얻은 변위도가 그림 3.4의 (c)이다. 카메라 노출의 차이는 획득하는 영상의 전역적인 명암도의 변화를 의미한다. 이는 전 영역에 걸쳐서 화소별로 증가 혹은 감소하는 명암도가 거의 일정하다고 볼 수 있다. 이런 경우엔 그림 3.4의 (c)처럼 시각적으로 변위도가 양호하게 얻어짐을 확인할 수 있다.

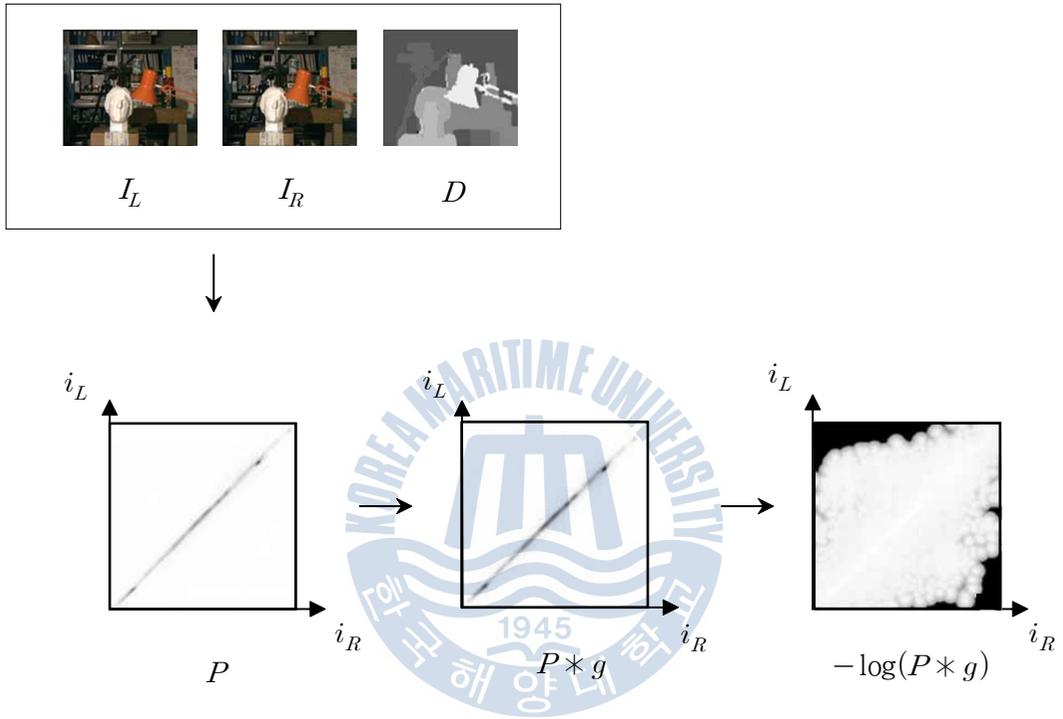


그림 3.3. 화소단위 결합확률과 엔트로피 계산 과정

( $I_L$ 과  $I_R$ 은 Middlebury Stereo Vision Lab에서 공개)

Fig. 3.3. Calculation process of a pixel-wise joint probability and entropy

(Image courtesy of  $I_L$  and  $I_R$  Middlebury Stereo Vision Lab)



(a)

(b)

(c)

**그림 3.4.** 카메라 노출 정도가 다른 Aloe 영상과 MI의 변위도

(a) 좌영상 (b) 우영상 (c) MI의 변위도

((a)와 (b)는 Middlebury Stereo Vision Lab에서 공개 )

**Fig. 3.4.** Image pair of Aloe with different camera exposure and disparity map of MI

(a) left image (b) right image (c) disparity map of MI

(Image courtesy of (a) and (b) Middlebury Stereo Vision Lab)

지역적인 밝기 변화가 발생한 Aloe 영상과 MI의 변위도는 그림3.5와 같다. 두 영상은 밝기의 차이가 약간 발생한 듯 보이나 자세히 관찰해보면 조명의 컬러가 다르고, 조명이 비추는 방향도 달라서 그림자가 우영상 중앙 벽에 위치해 있다. 이러한 경우 두 영상의 컬러 채널별 대비의 폭이 다르고 그림자의 영향 등 지역적인 명암도의 변형이 발생한 것이다. 명암도나 컬러를 기반으로 대응점을 찾는 기존의 유사도 척도를 이용하면 그림 3.5의 (c)처럼 오정합이 많이 발생한다.

### 3.2.2 Haar 웨이블릿 특징을 결합한 엔트로피

기존의 MI 유사도 척도는 지역적인 밝기의 변화가 발생했을 때 성능이 크게 떨어진다. 이는 결합확률분포의 형태가 지역적으로 불균일하게 밝기가 변한 경우 두 영상의 두 영상의 명암도 분포의 관계성이 모호해진다. 그리고 MI는 화소들의 공간 정보를 활용하지 않고 명암도쌍의 분포 특성만을 이용하기 때문에 주어진 변위도  $D$ 가 정확하지 않으면 명암도 평면에 골고루 퍼져 있는 결합확률을 얻게 되고, 엔트로피는 큰 값으로 계산된다. 이러한 문제점들은 주어진 영역의 공간적 특성을 반영할 수 있고 밝기 변화에 강인한 공간 정보를 이용하여 해결할 수 있다.

제안한 유사도 척도는 Haar 웨이블릿 필터를 이용하여 공간 정보를 얻는다. 대응점의 공간 정보는 대응점 중심의 윈도우 영역에 대해 Haar 웨이블릿 필터를 통과시켜 각 화소의 특징을 표현하는 4차원의 Haar 웨이블릿 응답 벡터를 얻고 주변 화소에 대해 대표성을 부여하여 정해진 크기의 특징 벡터를 얻을 수 있다. Haar 웨이블릿 특징 벡터를 얻는 과정은 그림 3.6과 같다. 특징 벡터를 얻기 위해 현재 화소  $p$ 를 중심으로  $m \times m$  크기의 윈도우를 형성한다. 여기에  $h \times h$  크기의 Haar 웨이블릿 필터를 통과시켜  $m \times m$  윈도우의 각 화소에 대해 4차원 Haar 웨이블릿 응답을 얻는다.  $m \times m$  영역에 대한 대표성을 부여한다. 이 때 대표성을 부여하는 화소의 갯수  $k$ 개에 대해 전체 벡터의 크기는  $4 \times k$ 개

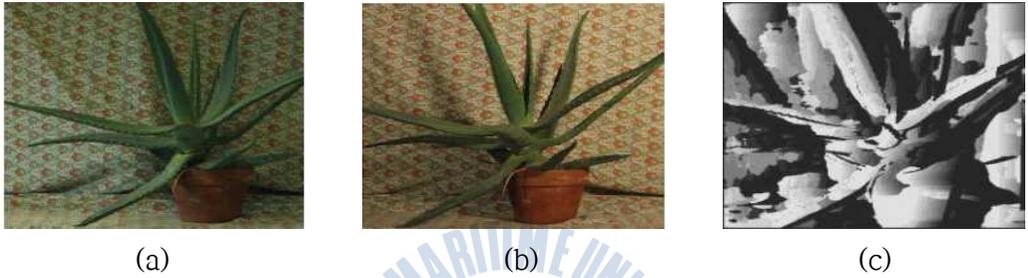


그림 3.5. 조명의 변화가 발생한 Aloe 영상과 MI의 변위도

(a) 좌영상 (b) 우영상 (c) MI의 변위도

((a)와 (b)는 Middlebury Stereo Vision Lab에서 공개 )

Fig. 3.5. Image pair of Aloe with illumination change and disparity map of MI

(a) left image (b) right image (c) disparity map of MI

(Image courtesy of (a) and (b) Middlebury Stereo Vision Lab)

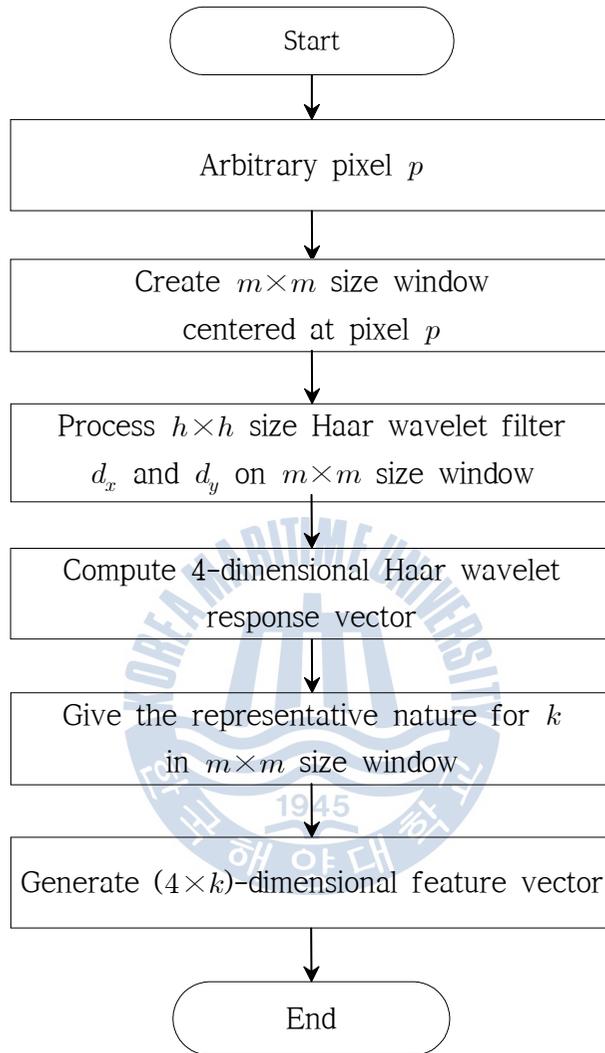


그림 3.6. 임의의 화소  $p$ 에서 제안한 Haar 웨이블릿 특징 벡터를 생성하는 순서

Fig. 3.6. A progress of generating proposed Haar wavelet feature vector at arbitrary pixel  $p$

로 얻을 수 있다.

Haar 웨이블릿 응답 영상을 얻는데 사용하는  $4 \times 4$  크기의 근사화된 Haar 웨이블릿 필터의 예는 그림 3.7과 같다.  $x$ 방향의 응답을 얻기 위해  $d_x$ 를 사용하고,  $y$ 방향의 응답을 얻기 위해  $d_y$ 를 사용한다.  $d_x$ 를 통해 얻는 응답  $\sum d_x$ 는  $x$ 방향으로 변하는 명암도 변화량의 총합으로, 정해진 영역의  $x$ 방향에 대한 명암도의 증가 혹은 감소의 정도를 의미한다.  $\sum |d_x|$ 는  $x$ 방향으로 명암도 변화의 발생 여부를 표현하며 고주파 성분이 포함되면 큰 값을 갖는다.  $y$ 방향에 대한  $d_y$ 의 응답 또한 동일한 특성을 갖는다. 이런 과정을 거쳐 식 (3.15)와 같은 4차원의 Haar 웨이블릿 응답 벡터  $V_{Haar}$ 를 얻는다.

$$V_{Haar} = [\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|] \quad (3.15)$$

Haar 웨이블릿 응답 벡터는 식 (3.15)의 4개의 성분으로 지역의 명암도 분포에 대한 주파수의 방향과 특성을 표현할 수 있다. 실제로 다양한 명암도 형태에 대한 Haar 웨이블릿 응답의 변화는 그림 3.8과 같다. 질감이 없는 이상적으로 평탄한 영역에서의 필터 응답은 그림 3.7의 (a)와 같이 특징 벡터의 성분 4개의 크기가 작다. 동일한 패턴이 반복적으로 나타나는 그림 3.7의 (b)의 경우 경계선 방향의 절대값 성분이 최대치를 갖고 나머지 성분은 0에 가까운 값을 갖는다. 마지막으로 명암도가 선형으로 변하는 그림 3.7의 (c)와 같은 경우 명암도가 변하는 방향의 성분 두개는 최대값을 갖고 나머지 값은 0이다. Haar 웨이블릿 응답은 지역의 명암도가 어느 방향으로 어떻게 변하는지 주파수 개념으로 서술할 수 있는 장점을 지니고 있다. 그리고 Haar 웨이블릿 응답은 실제 조명치의 추정 오차에 불변한 성질을 갖는다. 대비의 변화에 대해 불변한 특성을 지니고 있기 때문에 영상의 이득(gain)의 변화에 강인하다.

위에서 얻어진 4차원 벡터들에 대해 대표성을 부여하여 특징 벡터의 변별력을 높일 수 있다. 화소  $p$ 를 중심으로 하는  $13 \times 13$  크기의 윈도우 영역에 대해

1	1	-1	-1
1	1	-1	-1
1	1	-1	-1
1	1	-1	-1

$d_x$

(a)

1	1	1	1
1	1	1	1
-1	-1	-1	-1
-1	-1	-1	-1

$d_y$

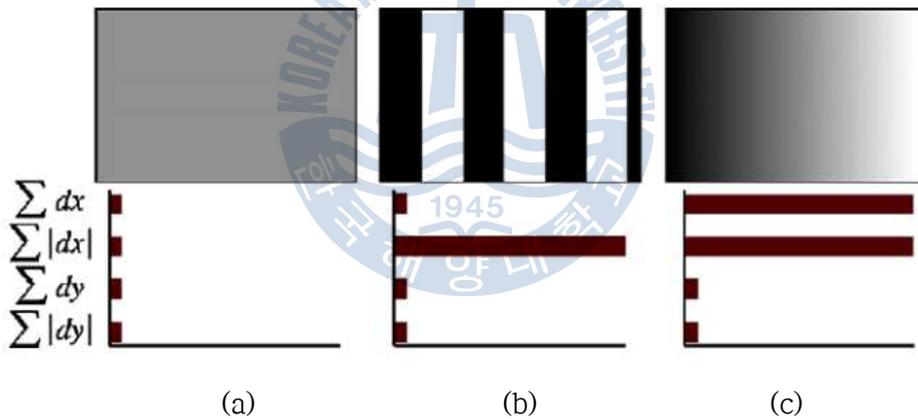
(b)

그림 3.7. 근사화된 Haar 웨이블릿 필터의 예 (4×4 크기)

(a)  $d_x$  (b)  $d_y$

Fig. 3.7. An example of approximated Haar wavelet filter (4×4 size)

(a)  $d_x$  (b)  $d_y$



(a)

(b)

(c)

그림 3.8. 3가지 명암도 패턴에 대한 Haar 웨이블릿 응답의 특성

(a) 균일 영역 (DC)

(b)  $x$ 방향으로 주파수들이 존재하는 영역

(c)  $x$ 방향으로 명암도가 점진적으로 변하는 영역

Fig. 3.8. Properties of Haar wavelet response about 3 intensity patterns

(a) homogeneous region (DC)

(b) region in presense of frequencies in  $x$  direction

(c) region in increasing intensity gradually in  $x$  direction

대표성을 부여하는 예는 그림 3.9와 같다. 특징 벡터 서술의 중심이 되는 화소  $p$ 를 기준으로 정의된 방향성을 계산하는 (a)의 형태나, (b)처럼 정해진 위치에서 벡터들의 평균을 계산하는 방법을 선택할 수 있다. 주변 화소와의 거리에 대한 가중치를 두기 위해 (c)와 같은 형태를 취할 수 있다. 이 때  $p$ 를 중심으로 반경  $R_2$ 는 7이 된다. 그리고 (d)와 같이 윈도우 영역의 정해진 위치에서 가중합을 구하여 특징 벡터를 구성할 수 있다. 이와 같은 경우 특징 벡터의 크기는 대표성을 띄는 9개의 화소가 4차원의 응답 벡터를 가지므로 36차원이 된다.

Haar 웨이블릿 특징 벡터 유사도 척도는 식(3.16)와 같이 두 벡터의 유클리드 거리를 이용하여 계산한다.

$$C_{local}(p, D_p) = \exp\left(\frac{\|V_L(p) - V_R(p - D_p)\|}{\gamma} - 1\right) \quad (3.16)$$

여기서,  $V_L(p)$ 와  $V_R(p - D_p)$ 는 각각 좌영상의 화소 위치  $p$ 와 우영상의 화소 위치  $(p - D_p)$ 에서 얻은 Haar 웨이블릿 특징 벡터이고  $D_p$ 는 화소  $p$ 에 할당된 변위이다.  $\gamma$ 는  $C_{local}(p, D_p)$ 의 최대치를 1로 정해주는 상수로써 스테레오 영상 정합 과정에서 계산된 Haar 웨이블릿 특징 벡터의 최대차이다.

제안한 Haar 웨이블릿 특징 벡터는 기존의 지역 서술자들에 비해 빠른 특징 추출 및 벡터 형성이 가능하다. 특징을 얻는 단계에서 Haar 웨이블릿 필터를 사용하기 때문에 영역의 명암도 분포를 주파수 형태로 표현할 수 있다. 이는 명암도 변화량 및 변하는 방향에 대한 표현력이 우수하고 지역적인 조명 특성에 강인하기 때문에 엔트로피와 결합하면 오정합을 줄일 수 있다.

앞에서 언급한 Haar 웨이블릿 특징 벡터를 이용하여, 제안한 방법은 엔트로피에 구조 정보를 결합한 형태의 유사도 척도를 구성한다. 화소 단위 유사도 척도는 화소 단위 엔트로피와 Haar 웨이블릿 특징 벡터가 결합된 항과 Census 변환을 이용하여 계산한 유사도 항의 합으로 표현된다. 현재 화소  $p$ 에서 변위

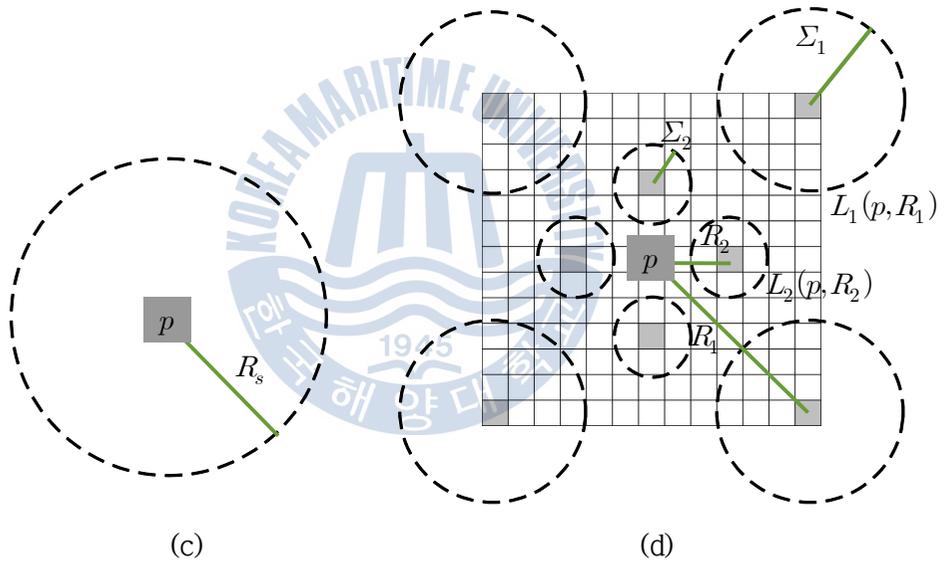
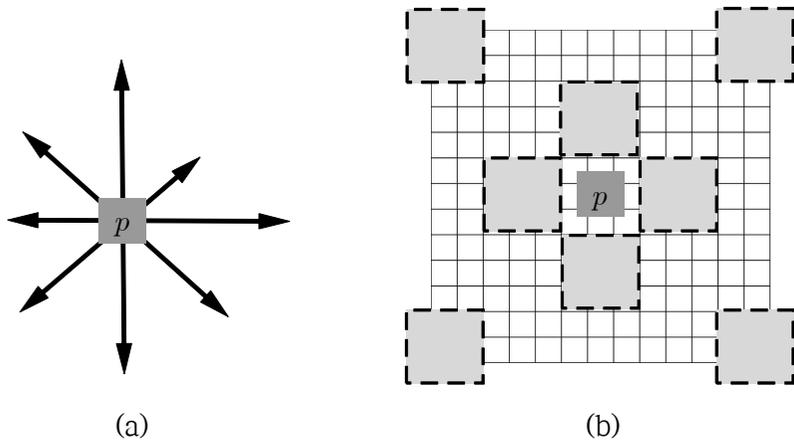


그림 3.9. 4가지 Haar 웨이블릿 특징 벡터의 형태

(a) 방향성 벡터 (b) 평균 벡터 (c) 가중합 벡터 (d) 고차원 벡터

Fig. 3.9. 4 patterns of Haar wavelet feature vector

(a) direction vector (b) mean value vector

(c) weighted sum vector (d) high dimension vector

$D_p$ 에 대해 계산된 화소 단위 유사도 척도는 식(3.17)과 같다.

$$C_{data}(p, D_p) = C_p(p, D_p) + C_{census}(p, D_p) \quad (3.17)$$

여기서  $C_p(p, D_p)$ 는 화소  $p$ 에 할당된 변위  $D_p$ 에 의해 계산되는 엔트로피 유사도  $C_{entropy}(p, D_p)$ 와 Haar 웨이블릿 특징 벡터를 이용한 유사도항  $C_{local}(p, D_p)$ 가 결합된 형태이다. 그리고  $C_{census}(p, D_p)$ 는 다음 절에 설명할 Census 변환을 이용한 유사도 척도이다. 화소 단위 엔트로피  $C_{entropy}(p, D_p)$ 은 식(3.18)과 같다.

$$C_{entropy}(p, D_p) = h\{I_L(p), I_R(p - D_p)\} \quad (3.18)$$

제안한 Haar 웨이블릿 특징 벡터를 결합한 엔트로피 유사도 척도는 엔트로피 유사도 척도와 Haar 웨이블릿 특징 벡터 유사도 척도를 곱한 형태이다. 밝기의 변화가 전역적으로 발생한 스테레오 영상에 대한 명암도 쌍의 확률분포는 선형적으로 나타나기 때문에 엔트로피 유사도 척도가 좋은 성능을 보이지만 지역적인 밝기 변화가 발생한 스테레오 영상에서 엔트로피 유사도 척도는 높은 값으로 계산되어 변위도의 정확도가 떨어진다. 이는 두 영상의 명암도 분포가 비선형 형태로 나타나기 때문이다. 이와 같이 확률 분포만으로는 대응성을 정확히 판별할 수 없는 환경에서 밝기 변화에 강인한 공간 정보를 식 (3.16)과 같은 형태로 엔트로피에 곱해줌으로써 높은 엔트로피 유사도를 보정하고 엔트로피가 보전하지 못하는 경계선 영역을 명확히 보전한다. 공간 정보를 포함한 엔트로피 유사도 척도는 식(3.19)로 나타낼 수 있다.

$$C_p(p, D_p) = C_{local}(p, D_p)C_{entropy}(p, D_p) \quad (3.19)$$

유사도 척도  $C_p(p, D_p)$ 는 Haar 웨이블릿 특징 벡터 유사도 척도인  $C_{local}(p, D_p)$ 와

엔트로피 유사도 척도인  $C_{entropy}(p, D_p)$ 를 곱하여 얻는다.

제안한 유사도 척도에서 Haar 웨이블릿 특징 벡터와 엔트로피의 역할은 그림 3.10과 같다. 조명이 다른 Moebius 영상에 대해 기존의 엔트로피 유사도 척도로 정합한 결과는 밑의 변위도와 같다. 위의 변위도는 Haar 웨이블릿 특징 벡터를 이용하여 정합한 결과이다. Haar 웨이블릿 특징 벡터는 앞서 언급한 특성으로 인해 조명이 다른 조건에서 엔트로피보다 좋은 성능을 보인다. 이러한 특성을 엔트로피 유사도에 결합함으로써 오른쪽 변위도와 같은 양호한 결과를 얻을 수 있다. 지역적인 밝기 변화로 인해 엔트로피 유사도가 커졌을 때 밝기 변화에도 유사도를 잘 반영하는 지역 특징 벡터의 유사도를 결합하여 대응점이라 판단되면 엔트로피 유사도를 감소시켜 정확한 정합을 유도할 수 있다. 따라서 Haar 웨이블릿 특징 벡터를 이용한 유사도는 오정합을 많이 발생한 엔트로피의 유사도를 보정해주는 역할을 한다.

Haar 웨이블릿 특징 벡터를 결합한 엔트로피 유사도 척도는 기존의 MI에 비해 최종 변위도를 얻는데 걸리는 수렴 횟수를 감소시킨다. 이런 효과를 보이기 위한 수렴 곡선의 예는 그림 3.11과 같다. 그래프에서 사각형을 포함한 곡선은 기존의 MI의 오정합 화소 비율이고, 마름모를 포함한 곡선은 제안한 Haar 웨이블릿 특징 벡터 결합 엔트로피 유사도 척도의 오정합 화소 비율이다. 변위는 명시한 유사도 척도만을 사용하여 얻었고 반복횟수별로 오정합 화소 비율을 계산하여 표시하였다. 사용한 영상은 Aloe와 Cloth4 영상이고 각각 그림 3.11의 (a)와 (b)이다. 두 영상의 조명이 다른 경우 Aloe의 경우 그림 3.11의 (a)와 같이 제안한 유사도 척도는 공간 정보를 이용하여 기존의 MI보다 오정합 화소 비율을 낮췄고, Aloe 영상에서 제안한 유사도 척도의 수렴 횟수는 약 5회였고 기존의 MI는 8회 반복해도 수렴하지 못했다. 이는 Cloth4 영상에서도 유사하게 나타났다. 제안한 유사도 척도의 수렴 횟수는 4회였다.

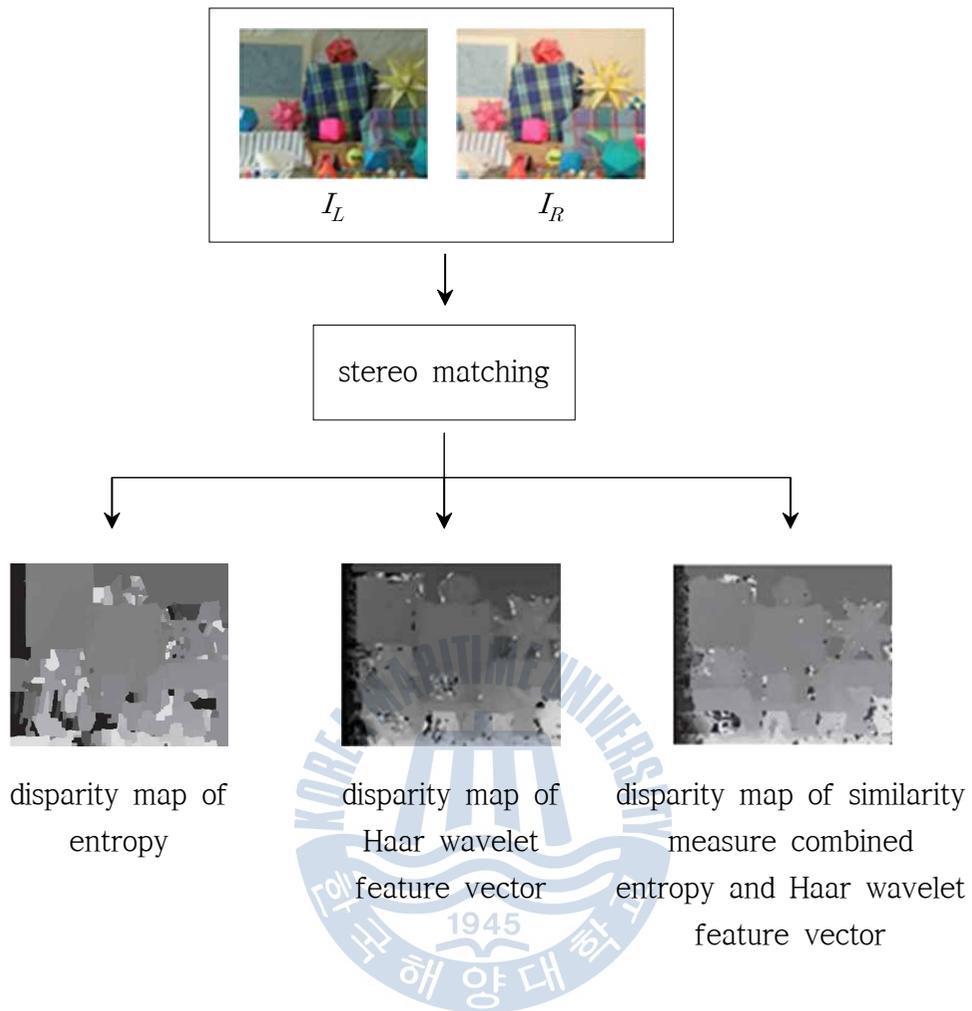
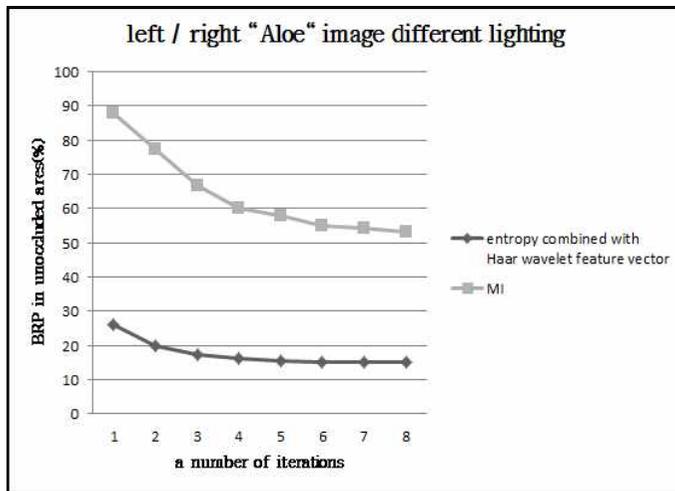


그림 3.10. Moebius 영상에 대한 엔트로피와 Haar 웨이블릿 특징 벡터의 변위도

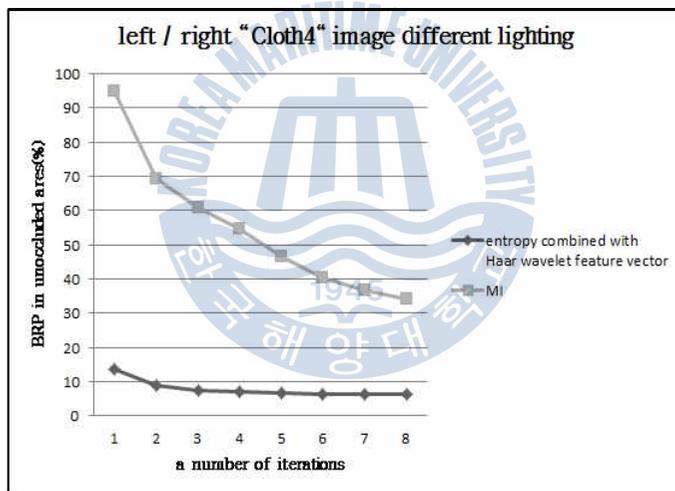
( $I_L$ 과  $I_R$ 은 Middlebury Stereo Vision Lab에서 공개)

Fig. 3.10. Disparity maps of entropy and Haar wavelet feature vector about Moebius image pair

(Image courtesy of  $I_L$  and  $I_R$  Middlebury Stereo Vision Lab)



(a)



(b)

그림 3.11. 조명이 다른 환경에서 MI와 Haar 웨이블릿 특징 벡터를 결합한 엔트로피 유사도 척도의 오차율 수렴 속도 비교를 위한 수렴 곡선 그래프

(a) Aloe (b) Cloth4

Fig. 3.11. Convergence curve graphs for comparison of error rate convergence velocity of similarity measure MI and entropy combined with Haar wavelet feature vector

(a) Aloe (b) Cloth4

### 3.3 제안한 Census 변환 기반의 유사도 척도

#### 3.3.1 Census 변환

Census 변환은 화소의 명암도에 대한 대소 관계를 이용하여 이진 비트로 변환하고, 유사도는 이진 비트간의 비트 연산으로 얻는다. Census 변환의 특징은 화소값을 그대로 사용하지 않고 특정 기준에 대한 해당 화소의 대소 관계를 이용하여 표현하기 때문에 명암도를 그대로 사용했을 때 발생하는 모호함을 해결해 줄 수 있고, 대응점 주변의 구조 정보를 표현할 수 있기 때문에 경계선 영역 등에서 좋은 성능을 보인다.

스테레오 정합에서 Census 변환은 좌영상과 우영상에 대해 각각 적용된다. 각 영상에서 대응점과 그 주변 화소의 대소 관계를 이진 비트로 변환하여 유사도를 계산할 수 있다. 스테레오 정합에서 Census 변환을 이용하여 유사도를 계산하는 과정의 예는 그림 3.12와 같다. 대응점을 중심으로 그림과 같이  $3 \times 3$  크기의 윈도우를 형성하여 기준 화소 자신을 제외한 대소 관계를 나타내는 8비트 길이의 이진 비트를 좌영상과 우영상에서 각각 생성한다. 이 때 기준이 되는 중심화소의 값보다 크면 1을 할당하고, 그렇지 않으면 0을 할당한다. 이렇게 얻은 비트들을 순서대로 나열한 비트열이 중심 화소의 Census 변환이다. 그림 3.12와 같이 좌우영상의 대소관계가 동일하다면 각각의 이진 비트는 11100000이 된다. 이렇게 얻은 이진 비트에 이진 비트에 배타적 논리합(XOR) 연산을 통해 해밍거리를 계산하여 두 대응점 간의 유사도를 얻을 수 있다.

현재 화소  $p$ 를 포함하는 주변 화소의 집합을  $N(p)$ 라 하고  $I(p)$ 가 점  $p$ 에서의 명암도일 때, 주변 화소를 이용한 화소  $p$ 의 Census 변환 결과인 이진 비트는 식 (3.20)와 같이 표현할 수 있다.

$$S(p) = \otimes_{q \in N(p)} \zeta(I(p), I(q)) \quad (3.20)$$

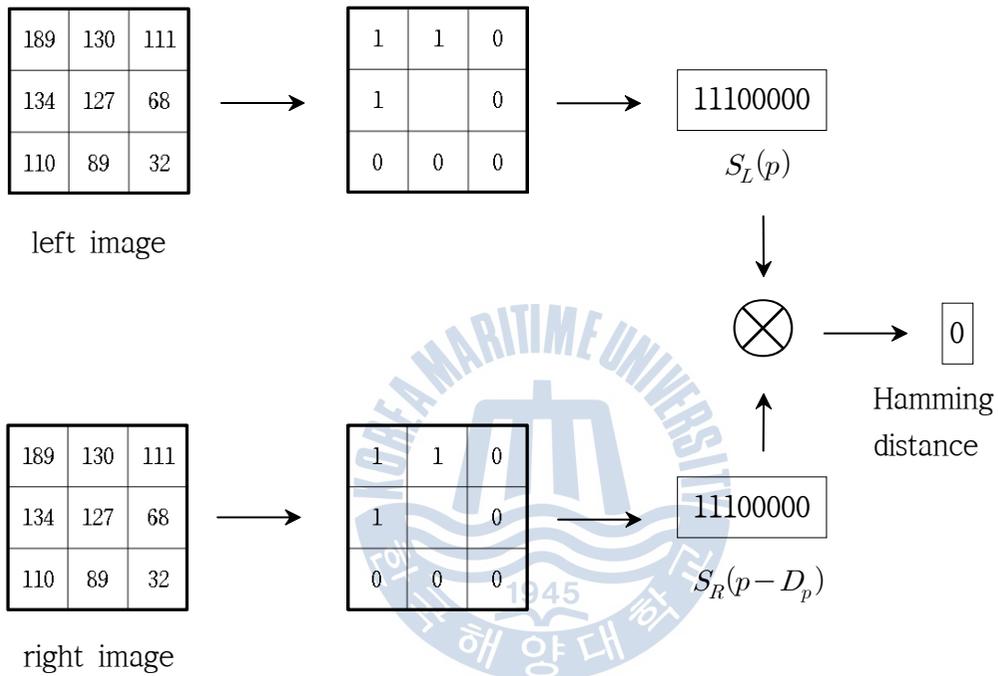


그림 3.12. Census 변환을 이용한 유사도 계산의 예

Fig. 3.12. An example of computing similarity using Census transform

여기서  $I(p)$ 와  $N(p)$ 에 포함된  $I(q)$ 를 비교하여 비트열을 만들어내는 연속연산자 (concatenation operation)인  $\otimes$ 를 사용한다. 그리고 임의의 두 명암도  $\alpha$ 와  $\beta$ 에 대한 비교 함수  $\zeta$ 는 식(3.21)과 같다.

$$\zeta(\alpha, \beta) = \begin{cases} 0, & \alpha \leq \beta \\ 1, & otherwise \end{cases} \quad (3.21)$$

여기서 비교 함수의 값을 결정하는 대소관계는 반대로 설정되어도 무관하다.

위 과정을 거쳐서 얻어진 대응점의 이진 비트열  $S(p)$ 는 지역의 구조적 특징을 나타낸다. 그러나 Census 변환시 중심화소의 역할이 절대적이기 때문에 중심화소가 큰 잡음의 영향을 받으면 오정합이 발생하고, 중심화소와 주변화소의 명암도가 같을 경우 명확히 변환할 수 없다는 문제점을 가지고 있다. 이러한 문제점은 Census 변환시 중심화소와 주변화소들과의 평균값을 계산하여 중심화소값 대신 비교할 때 사용하면 개선시킬 수 있다. 그러나 평균값을 취해도 주변 화소값이 평균과 동일해지는 경우가 발생하므로 적당한 오프셋을 설정하여 평균값에 더한 후 비교하는 방법을 사용한다.  $n \times m$  영역에 대해, 중심화소  $p = (x_p, y_p)$ 에 대한 평균 Census(Mean-Census) 변환  $\Gamma(x_p, y_p)$ 은 식 (3.22)와 같다.

$$\Gamma(x_p, y_p) = \bigotimes_{i=n} \bigotimes_{j=m} [\zeta\{\bar{I}(x_p, y_p) + \Delta, I(x_p + i, y_p + j)\}] \quad (3.22)$$

여기서,  $\bar{I}(x_p, y_p)$ 는 영역의 평균값이고  $\Delta$ 는 오프셋이다.

평균 Census 변환은 중심 화소가 잡음의 영향을 받아도 영역의 평균값을 기준으로 주변 화소와 비교하는 연산을 수행하기 때문에 비교적 강인하다. 하지만 지역적인 밝기 변화가 있는 영상에 적용했을 때 결과는 좋지 않다.

### 3.3.2 제안한 Census 변환을 이용한 유사도 척도

제안한 방법은 지역적인 밝기 변화가 발생한 스테레오 영상에서도 무질감 영역에 대해 오정합을 줄이고 물체의 경계선을 보전할 수 있도록 변형된 Census 변환을 이용한다. 기존 Census 변환과 다른 점은 변환 대상 영역을 중심화소를 공유하는 4개의 영역으로 나누고 각각의 영역의 변량과 평균을 구하여 변량이 가장 작은 영역의 평균을 변환 대상 영역의 평균으로 사용하고 이진 비트 생성 시 중심화소도 포함하는 점이다. 영역별 변량과 평균을 이용함으로써 지역적으로 명암도 변화량이 일정하지 않을 때나 잡음 등으로 인해 중심 화소값의 왜곡이 발생했을 때 이진 비트열이 변하는 단점을 보완하고, 이진 비트열 생성 시 중심화소도 포함시킴으로써 대응점의 직접적인 정보가 배제된 기존의 방법보다 구조 정보를 더 상세히 표현할 수 있고 평탄한 영역에서 좋은 성능을 보인다.

현재 화소  $p$ 와 그 주변 영역에 대한 명암도의 분산  $V$ 은 식 (3.24)와 같다.

$$V = \sum [I(p)]^2 - \frac{[\sum I(p)]^2}{N} \quad (3.24)$$

여기서  $I(p)$ 는 화소  $p$ 의 명암도이고,  $N$ 은 분산을 구하는 영역의 총화소수이다.

제안한 Census 변환은 명암도 분산과 평균을 각각의 영역에 대해 구하고 가장 적은 분산을 가지는 영역의 평균을 이진 비트열 생성의 기준으로 사용한다. 변형된 Census 변환의 예는 그림 3.13와 같다. 좌영상에서 분산이 1206으로 가장 적은 세번째 영역의 평균인 115를 기준값으로 정하고 주어진  $3 \times 3$ 영역에 대해 Census 변환을 수행한다. 이 때 중심 화소의 값이 어떠한 요인으로 150까지 변했을 경우, 평균 Census 변환에 의한 결과는 001001111에서 011101111로 변하지만 제안한 변형된 Census 변환의 이진 비트는 변하지 않는다. 우영상의 영역은 본래 좌영상과 명암도 분포가 같은 영역이지만 표시된 부분의 화소값이 지역적으로 변하였다. 이런 경우 명암도의 대소 관계가 변하기 때문에 기존의

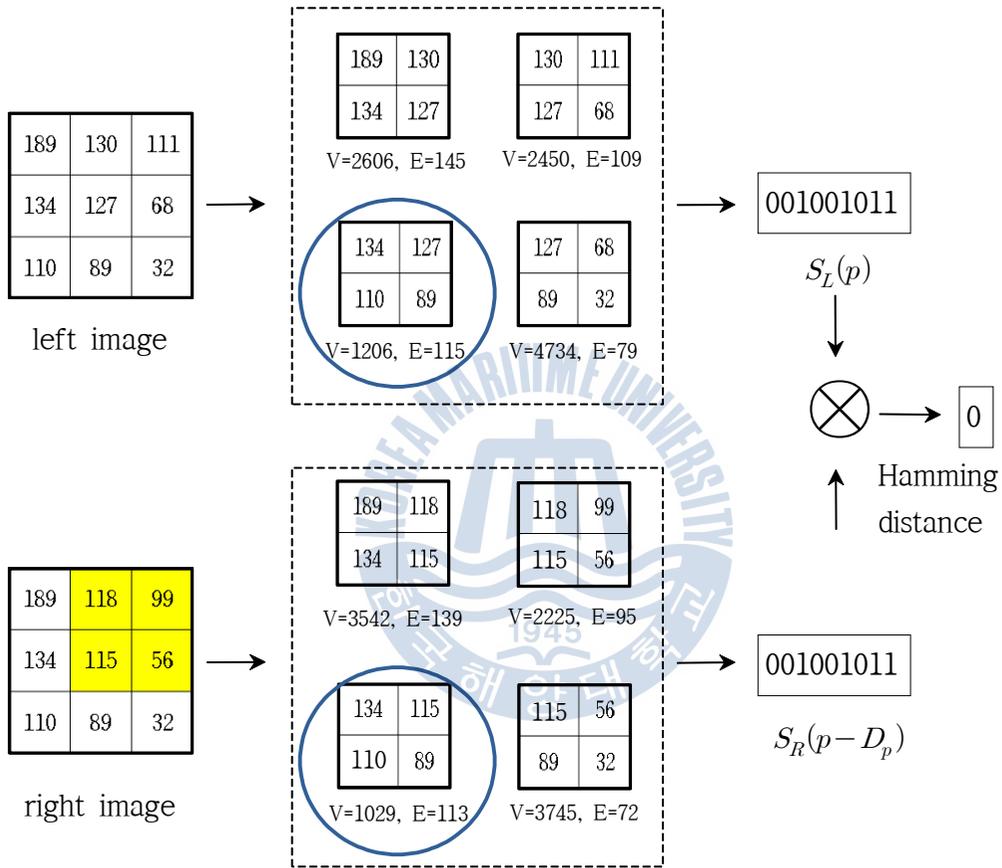


그림 3.13. 변형된 Census 변환을 이용한 유사도 계산의 예

Fig. 3.13. An example of computing similarity using modified Census transform

Census 변환은 좌영상과 우영상에 대해 서로 다른 이진 비트를 생성한다. 반면 제안한 Census 변환은 이러한 변화에도 좌영상과 동일한 이진 비트를 생성할 수 있다.

좌영상과 우영상에서 각각 변환된 비트열들의 유사도를 계산하기 위한 해밍 거리는 식(3.25)와 같다. 해밍거리는 두 비트열  $S_L(p)$ 과  $S_R(p-D_p)$ 의 배타적 논리합을 계산하여 얻는다.

$$C_X(p, D_p) = \text{Hamming}\{(S_L(p), S_R(p-D_p))\} \quad (3.25)$$

여기서  $S_L(p)$ 는 좌영상 화소  $p$ 의 비트열이고,  $S_R(p-D_p)$ 는 변위  $D_p$ 만큼 떨어진 우영상 화소의 비트열이다. 배타적 논리합을 이용한 Census 변환 유사도 척도는 식 (3.26)과 같이 강건 함수(robust function)를 이용하여 계산한다.

$$C_{census}(p, D_p) = \rho(C_X(p, D_p), \lambda) \quad (3.26)$$

여기서 강건 함수  $\rho(c, \lambda)$ 는 식 (3.27)과 같다.

$$\rho(c, \lambda) = 1 - e^{-\frac{c}{\lambda}} \quad (3.27)$$

여기서  $\lambda$ 는  $\rho(c, \lambda)$ 를  $[0,1]$  범위에서 변하게 하는 상수이다.

Census 변환을 통한 정합 성능 향상 효과의 예는 그림 3.14와 같다. Haar 웨이블릿 특징이 결합된 엔트로피 유사도 척도는 기존의 엔트로피가 가지지 못하는 구조 정보를 포함하고 있으나 지역적인 밝기 변화가 발생한 경우에는 결합된 구조 정보는 엔트로피의 오정합을 보정하는 용도로 사용된다. 그래서 공간 정보를 포함하는 Census 변환을 유사도로 결합함으로써 그림 3.14와 같이 무질

감 영역과 변위 불연속 영역에서 정확한 정합이 가능하다.

제안한 화소단위 유사도 척도는 Haar 웨이블릿 특징 벡터, 엔트로피, 그리고 변형된 Census 변환 유사도 척도를 화소단위로 결합한 식 (3.28)과 같다.

$$E_p(p, D_p) = C_{local}(p, D_p)C_{entropy}(p, D_p) + C_{census}(p, D_p) \quad (3.28)$$

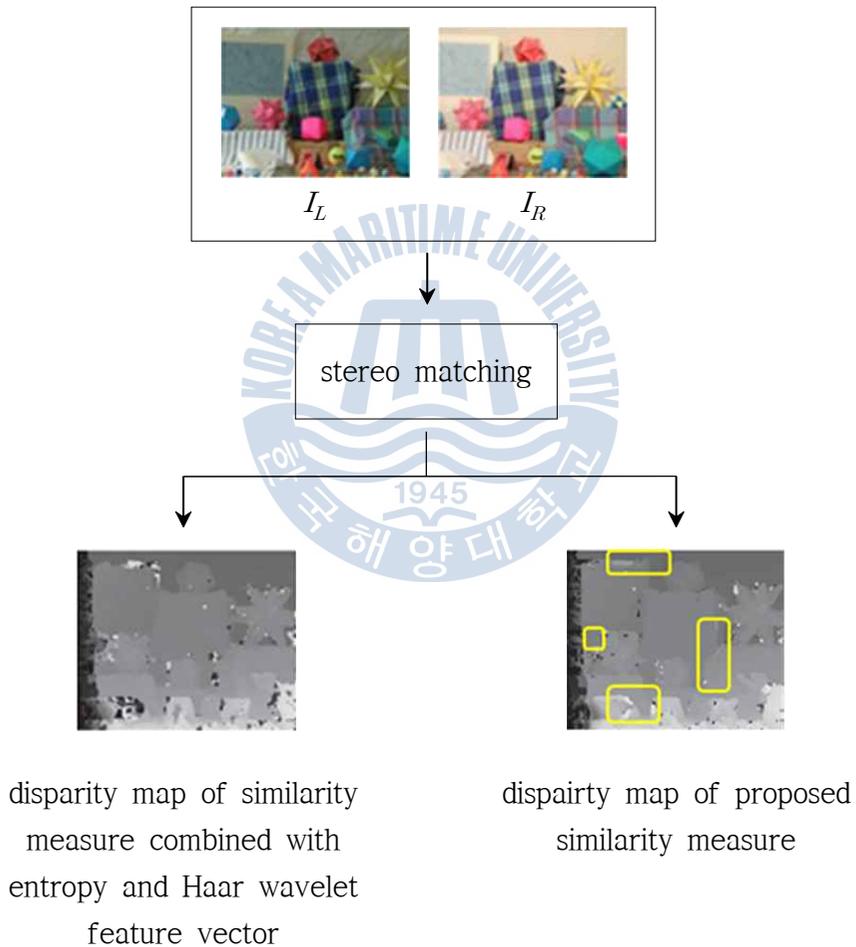


그림 3.14. 변형된 Census 변환의 효과

Fig. 3.14. Effect of modified Census transform

## 제 4 장 8방향 동적 계획법을 이용한 변위 추정

본 장에서는 동적 계획법을 8방향으로 적용하여 변위를 추정하는 최적화 방법을 제안한다. 제안한 최적화 방법은 8방향에 대해 최소 누적 에너지를 구하고 이들을 통합한 에너지에 대해 최적화를 수행하여 각 경로에 분포된 변위들을 추정한다. 이는 한 화소의 변위를 구하기 위해 여러 방향으로 최적화를 수행하므로 스트리킹 현상의 완화 및 경계선 주변에서 일정하게 분포하는 변위를 추정하는 성능이 우수하고, 이웃 화소와의 변위 차이 정도에 따라 에너지 함수의 증가량을 결정하므로 점진적인 변위 변화를 반영할 수 있도록 한다.

제안한 최적화 방법의 순서도는 그림 4.1과 같다. 모든 과정은 좌영상과 우영상에 대해 각각 수행한다. 3장에서 설명한 제안한 유사도 척도를 계산하여 주어진 변위 탐색 범위에 대한 3차원 DSI(Disparity Space Image)를 생성하여 정의된 화소 단위 에너지 함수를 계산한다. 3차원 DSI를 이용해 8방향에 대해 각각 화소 단위 누적 에너지를 구하고 이들을 통합하여 통합 에너지를 최소화하는 변위를 얻는다. 이 과정에서 좌영상을 기준으로 하는 변위도  $D_L$ 과 우영상을 기준으로 얻은 변위도  $D_R$ 을 얻는다. 이 두 변위도를 이용하여 폐색 영역을 검출한 후 보간하는 과정을 거쳐 최종 변위도  $D_L$ 을 얻는다.

### 4.1 동적 계획법

동적계획법은 수학자 리처드 벨만에 의해 1953년에 제시된 최적화 방법의 하나이다. 동적 계획법은 주어진 문제를 시간적 혹은 공간적으로 여러 개의 작은 문제 또는 하위 단계들로 나누고 각각의 최적해를 구한 후, 이들 간의 점진적인 관계식을 통하여 해들을 통합함으로써 주어진 문제의 해를 구한다. 동적 계획법을 이용한 스테레오 정합은 영상에서 경계선 등의 특징을 추출하여 그 특

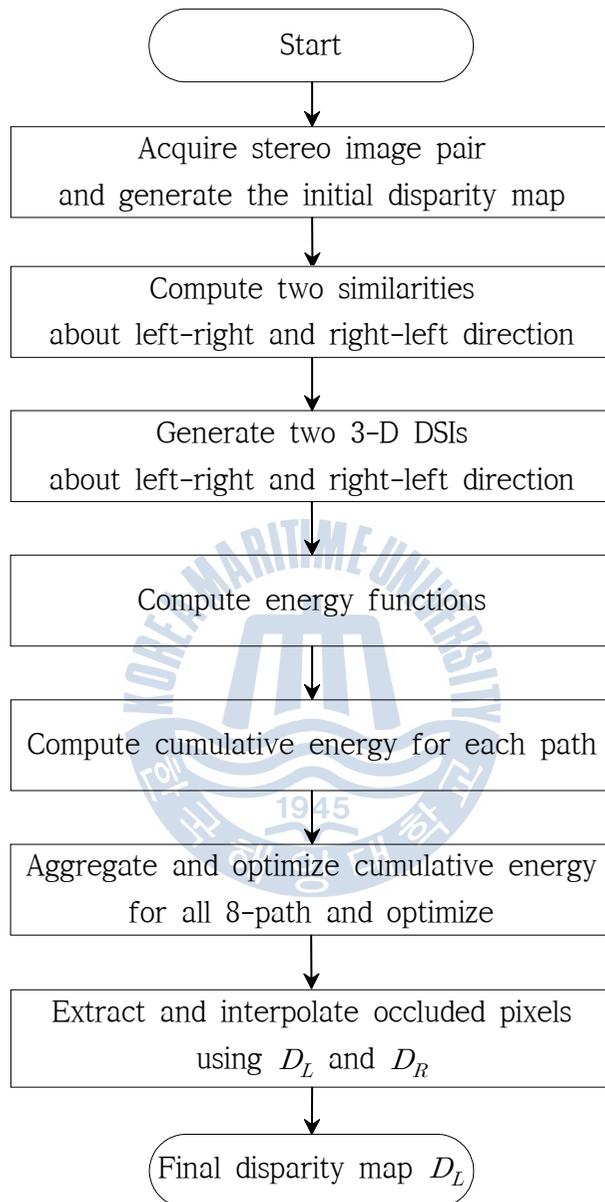


그림 4.1. 제안한 최적화 기법의 순서도

Fig. 4.1. A flowgraph of proposed optimization method

징에 대해 대응점을 찾는데 사용했다. 이런 방법은 정합 과정에서 보간과정을 거쳐야하기 때문에 변위도가 전체적으로 조밀하지 못하다. 그래서 최근에는 전체영상의 명암도나 컬러값을 이용하는 동적 계획법이 많이 연구되고 있다.

각 라인별로 변위의 범위 내에서 좌영상과 우영상의 화소값의 차이를 이용한 2차원 DSI를 만든다. 이때 화소값의 차이는 정합 비용 함수를 이용하여 계산하며, 2차원 배열의 한쪽 끝에서 반대쪽 끝까지 비용 함수를 계산하고, 최소의 값을 찾아 이동하면서 경로(path)를 만든다. 이후 만들어진 경로를 이용해서 각 화소별 변위를 추정한다. 최대 변위가  $D_{\max}$  이고  $m \times n$  크기의 영상인 경우 DSI는  $m \times n \times (D_{\max} + 1)$  크기의 3차원으로 만들어진다.

동적 계획법을 적용하기 위해 2차원 DSI를 만드는 예를 그림 4.2로 보였다.  $5 \times 10$  크기의 영상에 탐색 방향이 수평 방향일 때, 변위 탐색 범위가 0~3이고 AD를 유사도 척도로 사용한다면 적용되는 변위만큼 우영상의 첫번째 라인의 화소들을 이동시켜 유사도를 계산한다. 변위가 0일 경우, 이동없이 좌영상과 우영상의 각 화소의 위치에서 유사도를 계산하여 DSI의  $D=0$ 인 라인에 나열한다. 변위가 1일 때 우영상 라인을 왼쪽으로 한칸씩 이동하여 얻은 유사도를  $D=1$ 라인에 나열한다. 이와같은 방법으로 탐색 범위의 모든 변위에 대해 유사도를 계산하여 나열하면 라인당  $4 \times 10$  크기, 즉  $4 \times 10 \times 5$  크기의 3차원 DSI를 생성하게 된다. 3차원 DSI는 그림 4.3과 같다. 스테레오 영상의 크기에 깊이 방향으로 변위가 한 차원을 구성하는 형태이다.

DSI 생성 후, 정의된 에너지 함수를 계산하여 에너지 함수를 최소화시키는 경로를 찾아 그 경로상에 있는 DSI의 변위를 선택한다. 동적 계획법에서 사용하는 에너지 함수의 기본 형태는 식 (4.1)과 같다.

$$E(p, D_p) = C(p, D_p) + \tau[D_p - D_{p-1}] \quad (4.1)$$

여기서  $C(p, D_p)$ 는 DSI내에서 위치가  $(p, D_p)$ 에 해당하는 값으로 유사도 척도를

40	40	50	50	90	90	90	20	20	20
30	20	50	70	60	80	40	50	80	10
30	20	50	40	70	50	30	60	80	10
20	50	40	70	50	60	30	50	80	40
20	50	50	30	80	90	40	20	60	10

50	60	40	40	40	50	50	80	90	90
50	60	30	20	50	70	60	80	40	50
10	50	30	10	20	30	20	50	70	60
20	20	70	50	60	20	80	30	80	40
20	50	50	80	60	30	40	80	30	60

left image

right image



d=0	10	20	10	10	50	40	40	60	70	70
d=1	x	20	10	20	50	40	40	60	70	70
d=2	x	x	0	10	50	50	50	30	30	60
d=3	x	x	x	0	30	50	50	20	30	30

DSI (Disparity Space Image)

그림 4.2. 동적 계획법을 위한 2차원 DSI 생성

Fig. 4.2. Generating 2-D DSI for dynamic programming

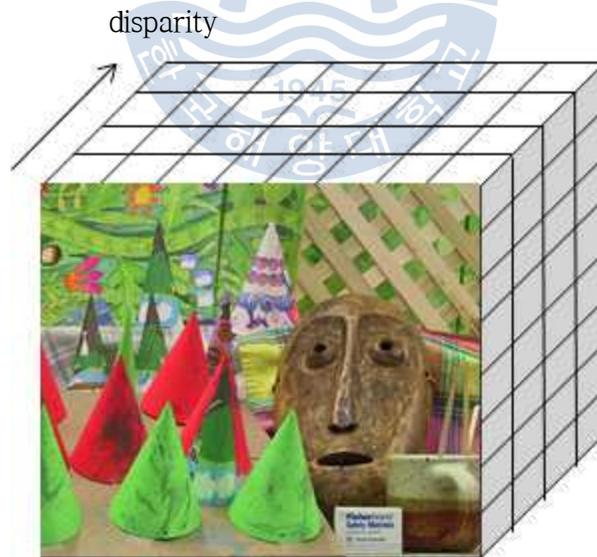


그림 4.3. 3차원 DSI

Fig. 4.3. 3-dimensional DSI

이용하여 계산한다.  $\tau$ 는 변위 불연속에 대한 벌점(penalty)으로 변위 연속성이 위배되거나 폐색 영역이라 판단되는 화소에 대해 정해진 상수만큼 에너지를 증가시키도록 설정된다.

동적 계획법에는 일대일 정합(one-to-one)과 일대다(one-to-many) 정합 방식이 있다. 일대일 정합은 폐색 영역과 정합 대상 화소를 구분하면서 이들 중 정합 비용을 최소화하는 변위를 계속적으로 탐색하는 방법이다. 한 화소의 대응점은 반드시 하나만 존재해야 한다는 유일성 제약 조건을 만족해야 한다.

일대다 정합은 한 영상에 투영된 영역은 기울기를 갖는 표면에서 다른 영상에 투영된 영역보다 좁거나 넓다는 전제로 이뤄진다. 기울어진 표면은 한 영상의 극상선에 위치한  $n$ 개의 화소와 다른 영상의 극상선에 위치한  $n \neq k$ 인  $k$ 개의 화소로 투영될 수 있다. 이 때 한 영상의  $n$ 개의 화소들은 다른 영상의  $k$ 개 화소들의 부분으로 정합되어야 한다. 일대다 정합은 다중 정합이 가능하도록 유일성 제약 조건을 무시하는 방법이다. 일대다 정합에서는 좁은 영역이 넓은 영역에 정합되거나 반대로 넓은 영역이 좁은 영역에 정합되기 위해 하나의 화소가 다른 영상에서 하나 이상의 화소들과 정합될 수 있다.

일대일 정합은 일대다 정합에 비해 계산량이 적으므로 빠른 결과를 얻을 수 있지만 거친 변위를 구할 수 있다. 반면 일대다 정합은 일대일 정합보다 상세한 결과를 얻을 수 있지만, 속도가 느리고 잡음에 매우 민감하여 불연속 지점을 판단하기 어렵다. 그래서 제안한 방법에서는 일대일 정합을 8방향으로 적용하였다.

## 4.2 제안한 8방향 동적 계획법

일반적으로 영상에서 명암도는 물체의 경계선 방향으로 일정하다. 그러므로 경계선 방향에 따라 그 방향에 맞는 동적 계획법을 적용하면 변위 추정에 대한 오차를 줄일 수 있다. 탐색 방향으로 발생하는 오류는 주로 탐색 방향과 일치

하지 않는 경계선 영역과 폐색 영역에서 발생한다. 탐색 방향과 일치하지 않는 경계선 영역의 변위는 다른 탐색 방향에 대한 변위도를 이용하여 해결할 수 있고, 대응점이 존재하지 않는 폐색 영역은 양방향 변위도를 이용하여 검출할 수 있다.

기존의 동적 계획법의 대표적인 단점은 탐색 방향으로 발생하는 줄무늬 형태의 스트리킹 현상이다. Cone 영상에 기존의 동적 계획법을 적용하여 얻은 스트리킹 현상이 발생한 변위도는 그림 4.4와 같다. 수평 방향으로 정합을 수행한 결과 스트리킹 현상이 발생하였다. 특히 이런 문제는 기존 동적 계획법의 탐색 방향에 대해 물체의 경계선이 많이 기울거나 수직인 영역에서 두드러진다.

동적 계획법을 적용하는 방향과 경계선 방향이 일치했을 때 그 영역에 대해서는 스트리킹 현상이 완화되고 경계선 주변의 변위가 정확히 구해진다. 이는 일반적으로 경계선 영역에 존재하는 변위는 그 방향을 따라 일정한 값을 갖기 때문이다. 이러한 특성은 동적 계획법을 다방향으로 적용할 수 있는 근거가 된다.

4방향에 대해 동적계획법을 적용하여 얻은 Cone 영상의 변위도는 그림 4.5와 같다. 이 중 좌측상단에서 우측하단 방향으로 동적 계획법을 적용하여 얻은 변위도는 그림 4.5의 (a)와 같다. 변위도의 좌측하단 영역에서 대응점을 찾지 못하거나 변위를 잘못 추정하였다. 이는 이 영역에 있는 물체의 경계선이 탐색 방향과 거의 직각을 이루고 있기 때문이다. 이와 같이 탐색 방향에 대해 변위가 급격히 변하게 되면 변위 연속성 에너지항에 의한 벌점이 높아져서 탐색 방향의 에너지가 커지기 때문에 변위를 잘못 추정하거나 폐색 영역으로 판단할 여지가 크다. 이 영역의 변위를 잘 추정한 변위도는 그림 4.5의 (b)이다. 이 변위도는 동적 계획법을 좌측하단에서 우측상단 방향으로 적용하여 얻었다. 변위도의 좌측하단에 위치한 꼬갈 모양의 물체들의 경계선이 탐색 방향과 비슷하기 때문에 이 영역에 대한 변위가 그림 4.5의 (a)보다 정확하다.

Cone 영상에서 우측하단의 작은 상자와 컵의 모양은 직사각형 형태이다. 이 영역에 대해 수직방향으로 동적 계획법을 적용하여 얻은 변위도는 그림 4.5의

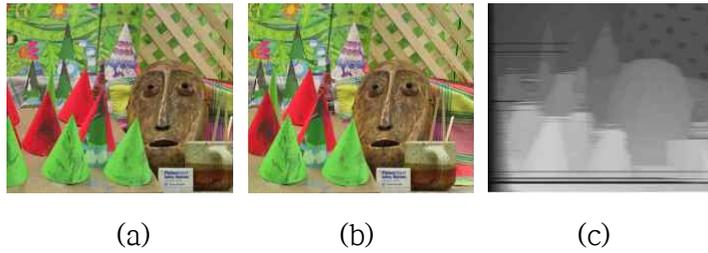


그림 4.4. Cone 영상과 1방향 동적 계획법에 의한 변위도

(a) 좌영상 (b) 우영상 (c) 변위도

((a)와 (b)는 Middlebury Stereo Vision Lab에서 공개)

Fig. 4.4. Cone image pair and disparity map using 1-path DP

(a) left image (b) right image (c) disparity map

(Image courtesy of (a) and (b) Middlebury Stereo Vision Lab)

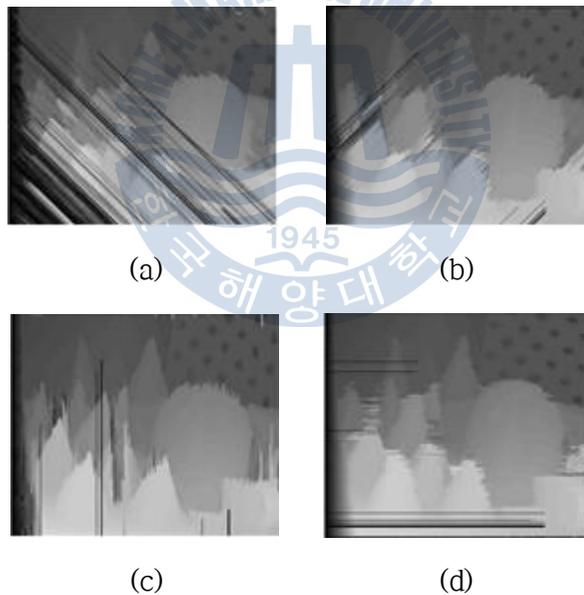


그림 4.5. 방향별 동적 계획법을 수행하여 얻은 Cone 영상의 변위도

(a) 좌상-우하 (b) 우상-좌하 (c) 상-하 (d) 좌-우

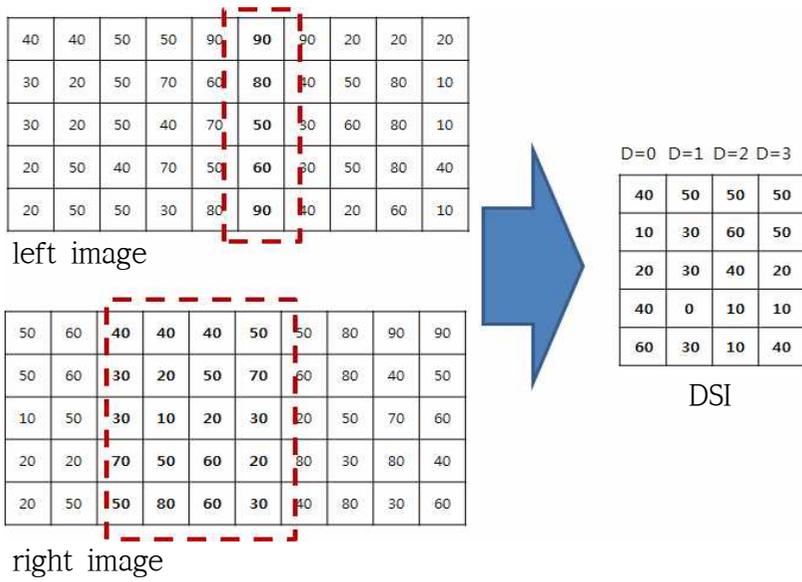
Fig. 4.5. Disparity maps of DP in each direction about Cone image pair

(a) left top-right bottom (b) right top-left bottom (c) top-bottom (d) left-right

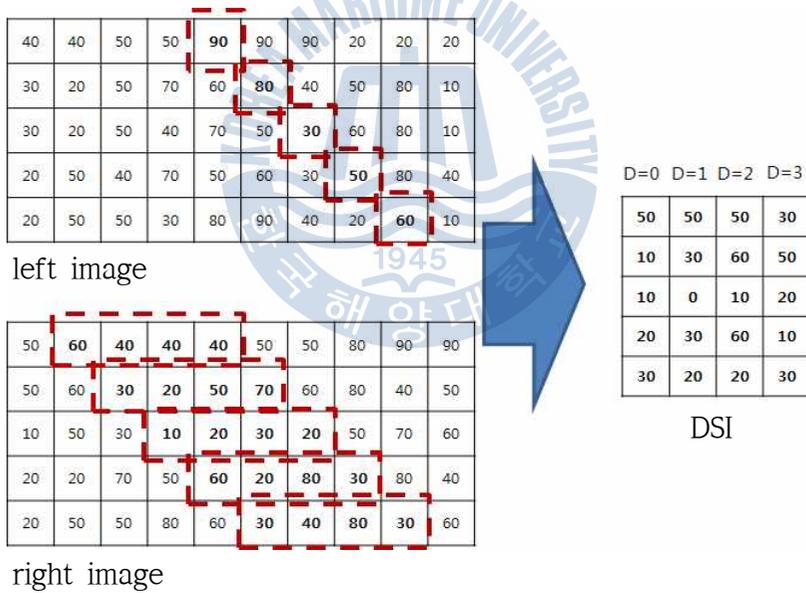
(c)와 같으며 작은 상자와 컵의 수직방향 변위가 잘 구해졌다. 그러나 탐색 방향과 직각을 이루는 수평방향의 경계선 영역에서 스트리킹 현상이 보인다. 이 방향에 대한 변위도는 그림 4.5의 (d)이다. 수직 방향으로 스트리킹 현상이 발생했지만 수평방향의 변위를 잘 추정하였다. 이와 같이 경계선 영역의 변위는 탐색 방향과 경계선 방향이 일치할수록 정확하고 그렇지 않은 영역의 변위는 다른 탐색 방향에 대해 계산된 변위를 참고하여 오정합을 줄인다.

8방향으로 동적 계획법을 적용하기 위해 변위 탐색 방향에 대한 DSI을 생각해 볼 수 있다. 상단에서 하단 방향의 DSI와 좌측상단에서 우측하단 방향의 DSI는 각각 그림 4.6의 (a)와 (b)이다. 변위의 범위가 0~3이고 유사도 척도가 AD일 때 좌영상의 6번째 열은 우영상의 6번째 열에서 변위만큼 좌측으로 이동한 열과 유사도를 계산하여 DSI에 나열되고 이는 그림 4.6의 (a)와 같다. 수직 방향의 경우 DSI의 크기는  $5 \times 4 \times 10$ 이다. 대각선 방향에 대한 DSI도 수직방향과 동일하게 우영상에서 변위만큼 이동시킨 대각선 방향의 화소라인에 대해 유사도를 계산하여 DSI를 만든다. 이는 그림 4.6의 (b)와 같다. 이 때 DSI의 크기 또한 수직 방향과 동일하다.

한편 DSI는 탐색 방향에 대해 각각 생성할 필요가 없다. 수평방향, 수직방향 및 대각선 방향에 대해 형성된 DSI는 각 차원을 구성하는 순서만 다르다. 따라서 하나의 3차원 DSI를 생성하여 해당 방향에 대해 동적 계획법을 적용하면 된다. 3차원 DSI에서 제안한 동적 계획법을 각 방향에 대해 탐색하는 예는 그림 4.7과 같다. 현재 화소  $p$ 를 기준으로 8방향에 대해 누적 에너지를 구한다. 여기서 방향 변수를  $a$ 라 두고 0~7까지 방향에 대한 번호를 부여한다. 방향 변수의 값은 기존의 동적 계획법의 적용 방향인 극상선 왼쪽에서 오른쪽으로 진행하는 방향을 0으로 하고 시계방향으로 회전하는 순서로 증가한다. 화소  $p$ 를 기준으로 0번 방향과 4번 방향에 대한 DSI 탐색 영역은 그림 4.8의 (a)이다. 1번 방향과 5번 방향에 대한 DSI 탐색 영역은 그림 4.8의 (b), 2번과 6번은 그림 4.8의 (c), 그리고 3번과 7번 방향은 그림 4.8의 (d)이다. 이와 같이 제안한 방법은 3차원 DSI를 생성하여 각 방향에 대한 누적 에너지를 고려하며 해당 화소의 정확



(a)



(b)

그림 4.6. 수직 방향과 대각선 방향에 대한 DSI 생성

(a) 수직 방향 (b) 대각선 방향

Fig. 4.6. Generating DSI about vertical and diagonal direction

(a) vertical direction (b) diagonal direction

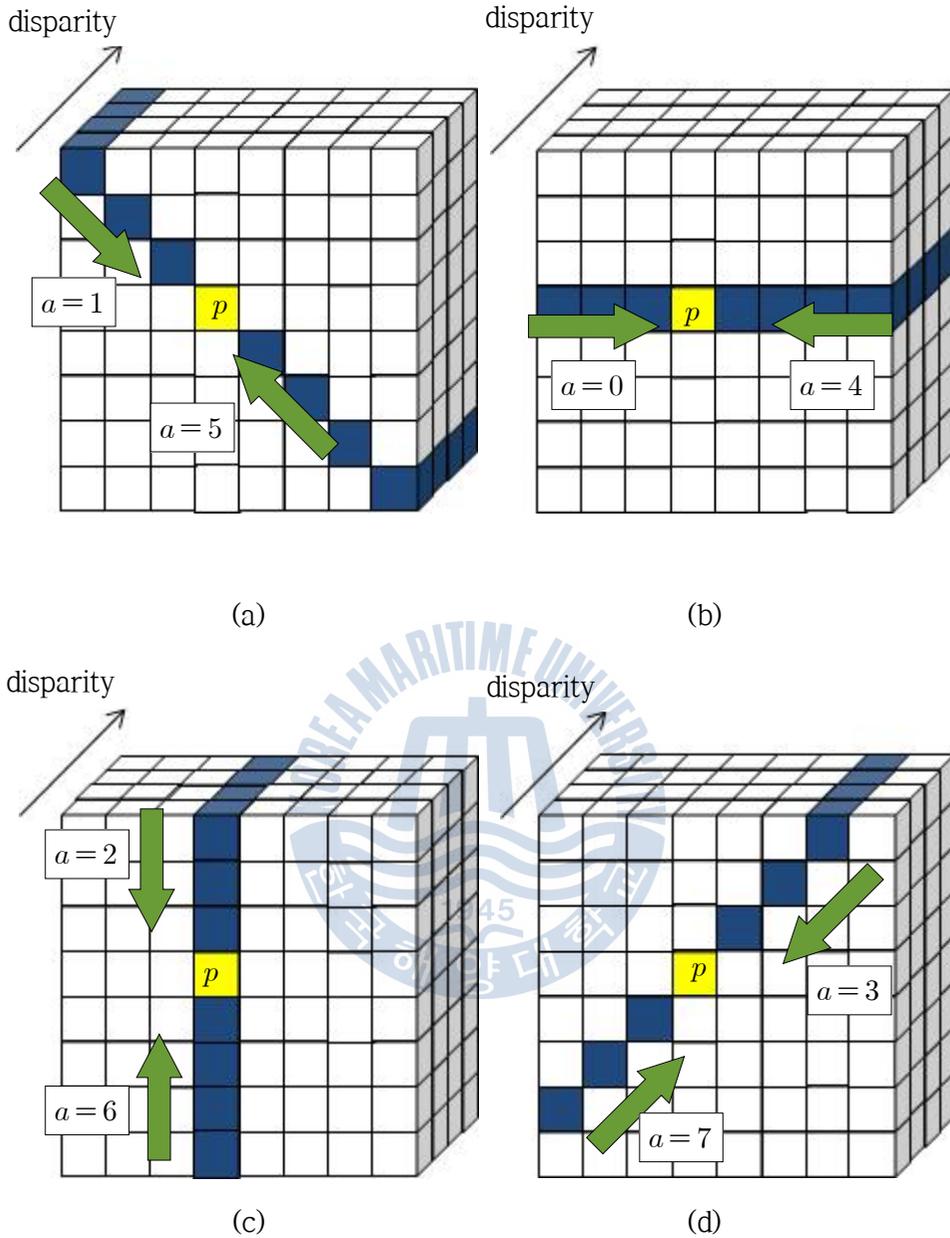


그림 4.7. 8경로에 대한 누적 에너지 통합

(a) a=0와 4 (b) a=1와 5 (c) a=2와 6 (d) a=3과 7

Fig. 4.7. Cumulative energy aggregation in 8-path

(a) a=0 and 4 (b) a=1 and 5 (c) a=2 and 6 (d) a=3 and 7

한 변위를 추정한다.

현재 화소  $p$ 에서 탐색 방향을 따라 다음 점의 변위를 결정할 때, 누적 에너지를 최소화하는 지점을 경로로 선택한다. 그리고 8방향에 대한 현재 화소  $p$ 까지의 누적 에너지들을 더함으로써 통합 에너지를 얻고 그 통합 에너지를 최소화시키는 변위들을 화소  $p$ 까지의 경로로 선택한다.

3장에서 제안한 엔트로피와 공간 정보를 결합한 새로운 에너지 함수는 기존의 에너지 함수의 형태인 식(4.2)를 기본으로 한다.

$$E(D) = E_{data}(D) + E_{smooth}(D) \quad (4.2)$$

여기서 유사도 척도항  $E_{data}(D)$ 는 화소단위 유사도 척도  $E_p(p, D_p)$ 의 총합으로 얻을 수 있다. 따라서 Haar 웨이블릿 특징 벡터를 결합한 엔트로피 항과 변형된 Census 변환 유사도 항으로 구성되며 식 (4.3)과 같다.

$$\begin{aligned} E_{data}(D) &= \sum_p E_p(p, D_p) \\ &= \sum_p \{C_{local}(p, D_p)C_{entropy}(p, D_p) + C_{census}(p, D_p)\} \end{aligned} \quad (4.3)$$

변위의 연속성을 보장하기 위한 현재 화소  $p$ 의 변위 연속성 에너지항  $E_{sm}(p, D_p)$ 은 화소  $p$ 와 그 이웃화소영역  $N(p)$ 의 화소  $q$ 의 변위차를 이용하여 계산한다. 이는 식 (4.4)와 같다.

$$E_{sm}(p, D_p) = \sum_{q \in N(p)} \alpha(D_p, D_q) \quad (4.4)$$

여기서  $D_q$ 는 화소  $q$ 의 변위이고 벌점함수(penalty function)  $\alpha(D_p, D_q)$ 는 두 화소의 변위차의 클수록 큰 벌점을 부여하는 역할을 한다. 우변의 첫번째 항은

작은 변위차에 대해 작은 벌점을 부여함으로써 점진적으로 변하는 변위를 반영하도록 하고, 두번째 항은 큰 변위차에 대해 큰 벌점을 줌으로써 변위 불연속 영역을 보전하는 역할을 한다. 이 때 추정된 변위  $D_p$ 와  $D_q$ 를 검증하는 의미에서 두 화소의 유사도를 이용하여 벌점을 결정하며 식 (4.5)와 같다.

$$\alpha(D_p, D_q) = \begin{cases} P_1 & , \quad |D_p - D_q| = 1 \\ P_2 / [ |E_p(p, D_p) - E_q(q, D_q)| + 1 ] & , \quad |D_p - D_q| > 1 \\ 0 & , \quad o.w. \end{cases} \quad (4.5)$$

여기서  $P_1$ 과  $P_2$ 는 상수이고  $P_1 < P_2$ 의 관계를 갖는다. 벌점함수의 값이  $P_1$ 인 경우 변위 차이가 1, 작은 변위 차이에 대해 상대적으로 벌점을 적게 해줌으로써 작은 변위의 차이를 허용하는 동시에 점진적인 변위의 변화를 반영할 수 있도록 한다. 두 화소의 변위 차이가 클 때, 두 화소의 유사도의 차이가 크면 큰 변위 차이에 대해 벌점을 완화하고, 유사도의 차이가 적을 경우 정확한 변위가 아니라고 판단하고 벌점을 강화한다. 이는 변위 차이가 클 때 에너지 함수가 무조건 증가하는 현상을 방지할 수 있기 때문에 전체 에너지 함수의 최적화 과정에서 발생하는 오차를 줄이고 변위 불연속 영역에서의 정합 성능을 높일 수 있다.

전체 에너지 함수인 식 (4.2)의 변위 연속성 항  $E_{smooth}(D)$ 는 식 (4.6)과 같이 화소단위 변위 연속성 항의 총합으로 얻을 수 있다.

$$E_{smooth}(D) = \sum_p E_{sm}(p, D_p) \quad (4.6)$$

전체 에너지 함수는 식 (4.4)와 식 (4.5)의 합, 즉 화소단위 유사도 척도 에너지항과 화소단위 변위 연속성 에너지항의 총합이며 식 (4.7)과 같다.

$$E(D) = \sum_p [E_p(p, D_p) + E_{sm}(p, D_p)] \quad (4.7)$$

현재 화소  $p$ 의 변위를 추정하기 위해 8방향에 대해 누적 에너지를 구하여 통합한다. 8방향은 화소  $p$ 를 중심으로 하며 영상의 각 모서리에서부터 시작한다. 이 때  $a$ 를 방향 변수라 칭하고 각 방향에 대해 0부터 7까지의 정수를 할당한다.

8방향의 누적 에너지는 3차원 DSI의 값을 이용하여 계산한다. 이 때 종점 화소  $p$ 까지의 누적 에너지를 최소화시키기 위해 현재 탐색 방향에서 이전 화소  $p'$ 과의 변위 차이를 고려한다. 즉 화소 단위 에너지를 계산하는데 할당된 현재 변위  $D_p$ 와 이전 화소에서의 변위와의 차이를 클수록 큰 벌점을 부여한다.

방향 변수  $a$ 에 대한 누적 에너지  $R_a(p, D_p)$ 는 현재 방향에 대한 이전 화소  $p'$ 에서의 누적 에너지들 중 최소 에너지를 합하여 계산한다. 이는 식 (4.8)과 같다.

$$R_a(p, D_p) = [E_p(p, D_p) + E_{sm}(p, D_p)] + \min[R_a(p', D_p), P_1 R_a(p', D_p - 1), P_1 R_a(p', D_p + 1), P_2 \min_i R_a(p', i)] \quad (4.8)$$

여기서  $R_a(p, D_p)$ 는 방향 변수  $a$ 에 대해 현재 화소  $p$ 까지 변위  $D_p$ 를 적용하여 구한 화소 단위 에너지 함수의 총합을 뜻한다.  $E_p(p, D_p)$ 와  $E_{sm}(p, D_p)$ 는 각각 변위  $D_p$ 에 대한 현재 화소  $p$ 의 화소 단위 유사도 척도 에너지 항과 변위 연속성 에너지항이다. 방향  $a$ 에 대한 이전 화소  $p'$ 까지의 누적 에너지는 다음 3개의 누적 에너지 중에서 가장 작은 것을 선택한다.

$R_a(p', D_p)$ 는 변위  $D_p$ 에 대한 이전 화소  $p'$ 까지의 누적 비용이고,  $R_a(p', D_p - 1)$ 와  $R_a(p', D_p + 1)$ 는 각각 이전 화소  $p'$ 에서 근소한 차이를 보이는 변위에 대한 누적 비용이다.  $\min_i \{R_a(p', i)\}$ 는 이전 화소  $p'$ 에서 최소 누적 비용

을 갖게 하는 변위에 대한 누적 비용이다. 누적 에너지를 최소화하는 이웃 화소의 변위와 현재 변위  $D_p$ 와의 차이가 1이면 작은 벌점을 부여하고 1이상이면 큰 벌점을 부여하여 변위 불연속 영역이라 판단한다.

모든 방향에 대해 누적 에너지가 구해지면 에너지를 통합하는 과정을 수행한다. 변위  $D_p$ 에 대한 현재 화소  $p$ 의 통합 에너지  $U(p, D_p)$ 는 8방향의 누적 에너지를 모두 합하여 얻으며 식 (4.9)와 같다.

$$U(p, D_p) = \sum_a R_a(p, D_p) \quad (4.9)$$

여기서 모든 방향  $a$ 에 대한 누적 비용  $R_a(p, D_p)$ 의 총합이 최소가 되는  $D_p$ 를 현재 화소  $p$ 의 변위로 선택한다.

제안한 방법은 폐색 영역을 검출하기 위해 좌영상을 기준으로 얻은 에너지 함수와 우영상을 기준으로 얻은 에너지 함수를 따로 통합한다. 두 방향으로 통합하여 얻은 좌영상 기준의 변위도와 우영상 기준의 변위도를 각각  $D_L$ 와  $D_R$ 이라고 한다면 좌영상의 현재 화소  $p$ 에 대응되는 우영상의 화소  $q$ 의 변위를 이용하여 최종 변위를 얻는다. 현재 화소  $p$ 의 변위  $D_p$ 는 식 (4.10)과 같다.

$$D_p = \begin{cases} D_{Lp}, & |D_{Lp} - D_{Rq}| \leq 1 \\ D_{oc}, & o.w. \end{cases} \quad (4.10)$$

여기서  $D_{Lp}$ 는 좌영상 기준의 변위도에서 화소  $p$ 의 변위이고,  $D_{Rq}$ 는 우영상 기준의 변위도에서 화소  $q$ 의 변위이다.  $D_{oc}$ 는 양방향 변위들의 차이가 클 경우, 폐색 영역 후보가 됨을 의미한다. 폐색 영역 후보 화소는 모든 처리 과정 수행 후 주변의 가장 작은 변위를 할당받는다.

## 제 5 장 다해상도 기반의 스테레오 정합

정확한 변위도를 효율적으로 얻기 위해 다해상도 기반의 스테레오 정합 기법을 제안한다. 스테레오 영상의 영상 피라미드를 구축하여 저해상도 레벨부터 정합을 수행하여 계산복잡도를 줄일 수 있고 기존의 다해상도 기법보다 정확한 변위도를 얻을 수 있다. 기존의 다해상도 기법은 최저해상도 영상에서 얻은 변위도 자체를 원영상 크기까지 확대하기 때문에 속도는 빠르지만 조밀한 변위도를 얻기 어렵다. 제안한 방법은 모든 영상 피라미드 계층에 대해 정합을 수행함으로써 정확한 변위도를 얻을 수 있다.

### 5.1 가우시안 영상 피라미드

영상 피라미드 기법은 여러 해상도의 영상을 표현할 수 있으며 개념적으로도 간략한 구조를 지닌다[56,57]. 영상 피라미드는 피라미드 형상으로 감소하는 해상도의 영상들을 정렬시킨 구조를 갖는다. 가장 아래의 레벨 0는 기저 영상(basis image)으로 원본 영상이다. 영상 피라미드는 계층이 올라갈수록 기저 영상의 해상도를 감소시켜 얻은 근사 영상(approximation image)이 위치한다. 이때 레벨 0부터 피라미드 위쪽으로 한 단계씩 올라갈수록 영상의 크기는  $1/4$ 로 감소한다.

일반적인 피라미드 기법은 현재 레벨이 1단계일 경우 레벨 0단계의 근사 영상과 1단계의 예측차 영상(prediction image)을 이용하여 0단계의 영상을 복원할 수 있다. 레벨 0의 입력 영상을 다운표본화(downsampling)를 하여 레벨 1단계의 근사 영상을 획득한다. 획득된 레벨 1단계의 근사 영상을 다시 업표본화(upsampling)를 하고 보간을 거친 후 레벨 0의 입력영상과 차를 계산하여 레벨 0의 예측차 영상을 구한다. 레벨 0단계의 근사 영상과 레벨 1단계의 예측차 영상으로 레벨 1단계의 원 영상을 복원할 수 있는데 예측차 영상을 구할 때 사용

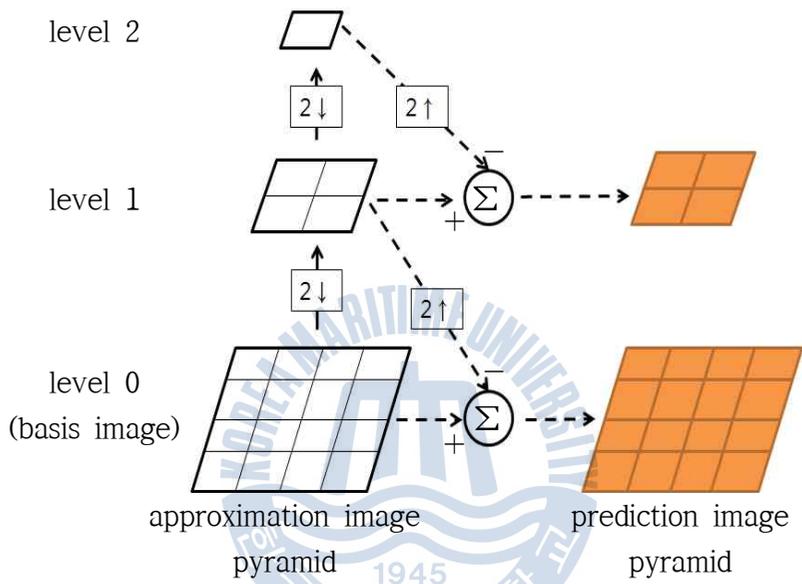


그림 5.1. 영상 피라미드의 예 (3단계)

Fig. 5.1. An example of image pyramid (3 levels)

한 보간 필터를 근사 영상에 적용시킨 후 예측차 영상과 근사 영상을 합하여 원영상으로 복원한다. 이러한 영상 피라미드의 예는 그림 5.1과 같다. 좌측은 기저 영상을 포함한 근사 영상 피라미드이고, 우측은 예측차 영상 피라미드이다. 레벨 0의 기저 영상을 다운표본화하여 3단계 영상 피라미드를 구성하고 있고, 기저 영상과 레벨 1단계의 근사 영상을 업표본화한 영상의 차를 이용하여 레벨 0의 예측차 영상을 얻을 수 있다.

일반적인 영상 피라미드는 계산의 편의를 위해 단순한 다운표본화 과정을 거쳐 근사 영상을 얻는다. 이는 피라미드 상위 계층으로 갈수록 근사 영상의 정보 손실을 야기하기 때문에 스테레오 정합에 적용했을 때 조밀한 변위도를 얻기 어렵다. 이러한 문제는 가우시안 피라미드를 이용하여 해결할 수 있다.

가우시안 피라미드는 영상 피라미드 구성할 때 가우시안 필터를 이용한다 [14]. 다음 단계 근사 영상을 얻을 때, 즉 다운표본화를 수행할 때 소실되는 화소들의 정보를 가우시안 필터를 이용하여 남아있는 화소에 가중합하여 포함시킨다. 가우시안 피라미드에서 화소값을 계산하는 예는 그림 5.2와 같다.  $g_0$ 에서  $g_2$ 까지 3단계 피라미드를 구성하는 과정에서 두번의 다운표본화 과정을 거치게 된다.  $g_1$ 은  $g_0$ 을 다운표본화하여 얻을 수 있는데  $g_0$ 의 화소 5개의 정보를  $g_1$ 의 한 화소에 포함시킨다. 따라서 다운표본화 과정에서 화소들이 삭제되어도 남아있는 화소에 삭제된 화소들의 정보가 남아있게 된다. 이러한 과정은  $g_2$ 를 생성할 때도 그대로 적용된다. 이와 같이 구성되는 가우시안 피라미드와 일반 피라미드의 차이는 그림 5.3과 같다. 일반 피라미드인 오른쪽 영상은 근사 영상에서 세밀한 정보를 손실하지만, 왼쪽의 가우시안 피라미드의 경우 근사 영상으로 축소됨에도 영상의 세밀한 정보를 포함하고 있는 것을 확인할 수 있다.

제안한 방법은 가우시안 피라미드를 이용하여 영상을 계층화한다.  $g_0$ 를 원영상으로 두고 한단계씩 축소하는 과정에서 가우시안 필터를 사용하여 삭제되는 주변 화소의 정보를 포함할 수 있다.  $k-1$ 단계의 가우시안 피라미드 영상을  $g_{k-1}$ 이라 하면 4배 축소된 근사 영상  $g_k$ 의 화소값은 식 (5.1)과 같이 계산된다.

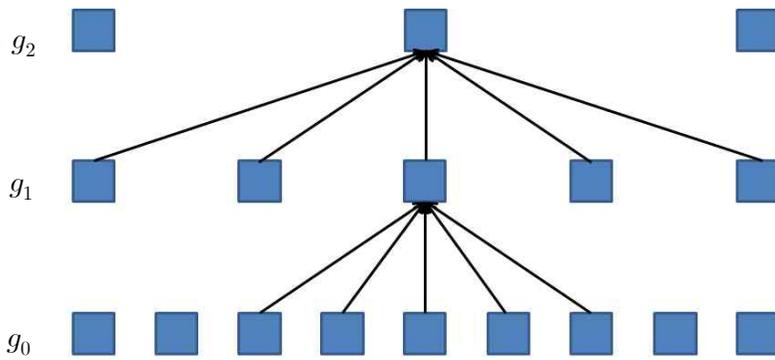


그림 5.2. 가우시안 피라미드의 화소값 계산

Fig. 5.2. Pixel value evaluation of gaussian image pyramid

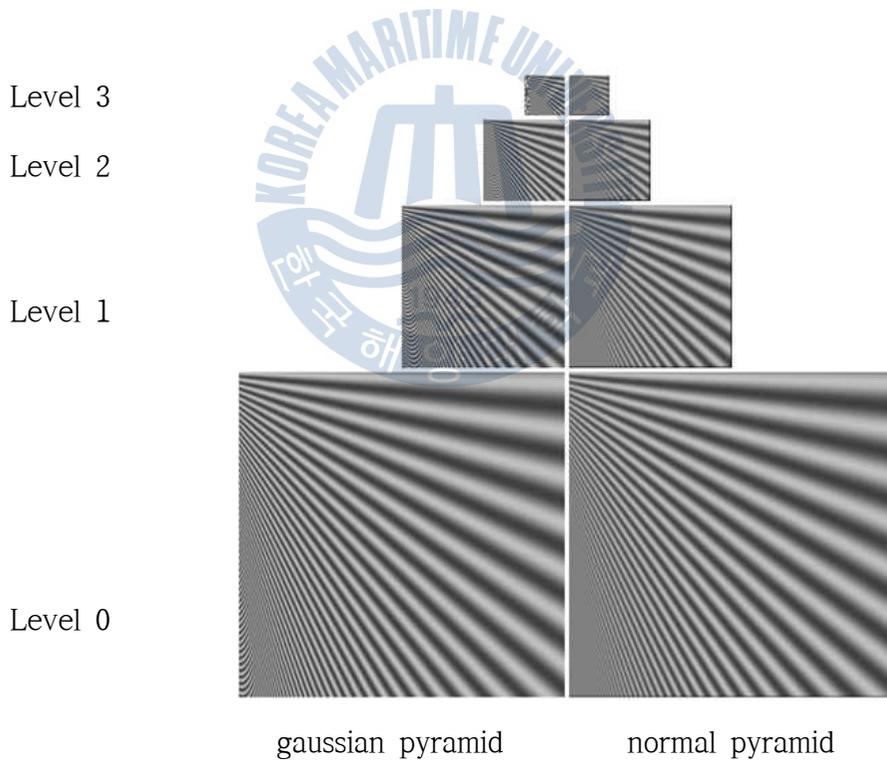


그림 5.3. 가우시안 피라미드와 일반 피라미드

Fig 5.3. Gaussian pyramid and conventional pyramid

$$\begin{aligned}
 g_k(i,j) &= REDUCE(g_{k-1}) \\
 &= \sum_{m=-2}^2 \sum_{n=-2}^2 w(m,n)g_{k-1}(2i+m,2j+n)
 \end{aligned}
 \tag{5.1}$$

여기서  $w(m,n)$ 은  $5 \times 5$ 크기의 가우시안 필터에서  $(m,n)$ 위치에 대한 근사화된 가우시안 가중치이다.  $g_k$ 에서  $g_{k-1}$ 를 복원하려면 예측차 영상이 필요하다. 가우시안 피라미드의 예측차 영상들은 라플라시안 피라미드를 형성한다.  $g_{k-1}$ 를 복원하기 위한 예측차 영상  $L_{k-1}$ 은  $g_k$ 을 4배 확대한 후  $g_{k-1}$ 와의 차를 이용해 얻을 수 있다. 3단계 가우시안 영상 피라미드의 예는 그림 5.4와 같다. 좌측은 근사 영상으로 구성된 3단계 가우시안 피라미드이고, 우측은 예측차 영상인 라플라시안 영상 피라미드이다. 여기서  $L_0$ , 즉 기저 영상 크기의 라플라시안 영상은 기저 영상인  $g_0$ 과 1단계 근사 영상  $g_1$ 을 4배 확대한 영상의 차를 통해 얻을 수 있다.

가우시안 피라미드는 일반적인 피라미드와 달리 가우시안 필터를 이용하므로 상위 레벨에서 표본화에 의한 에일리어싱을 방지할 수 있고, 가우시안 필터가 저역통과필터의 역할을 하기 때문에 아웃라이어에 의한 영향을 줄일 수 있다. 이러한 장점은 제안한 방법의 Haar 웨이블릿 특징을 얻는데 도움이 된다. 일반적인 피라미드의 근사영상에 대해 Haar 웨이블릿 특징을 얻으면 고주파 성분을 많이 포함하게 되는데, 가우시안 피라미드를 적용한 근사 영상에서는 샘플링에 의한 고주파 성분을 줄일 수 있기 때문에 Haar 웨이블릿 응답의 변별력이 좋아진다. 그리고 이러한 장점은 지역적인 밝기 변화가 발생한 영상에 대해 영상 피라미드 상에서 주변의 밝기 정보를 분산시키는 효과가 있다.

## 5.2 제안한 다해상도 기반 스테레오 정합

기존의 다해상도 기법은 최상위 단계, 즉 최저해상도 영상에서 정합을 시도

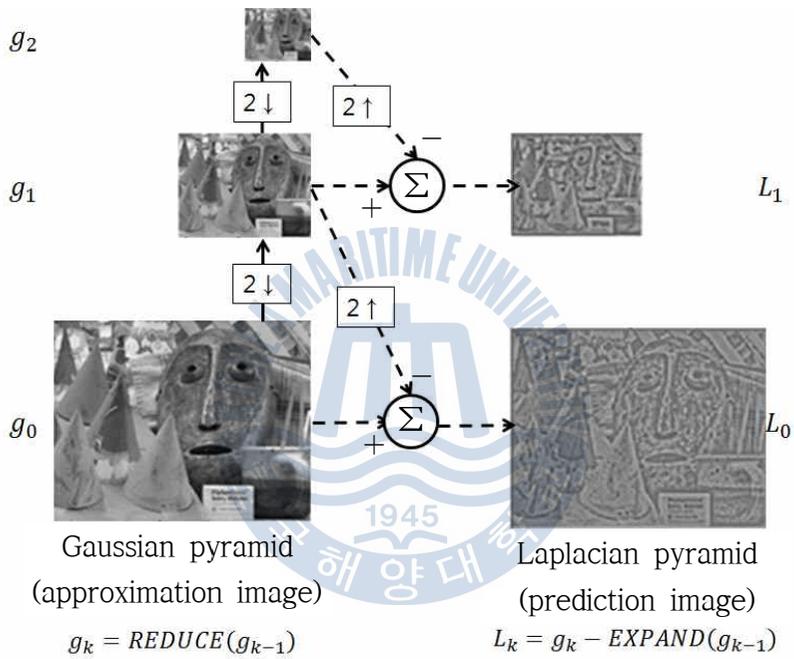


그림 5.4. 가우시안 영상 피라미드

Fig. 5.4. A gaussian image pyramid

하여 얻은 변위도를 원영상 크기까지 확대한다. 이 과정에서 정합은 최저해상도 스테레오 영상에 대해서만 수행하고, 원영상 크기의 변위도는 상위 해상도에서 얻은 변위도의 변위를 전파하고 보간하는 과정으로 얻는다. 기존 다해상도 기법을 그림 5.5의 (a)에 나타내었다. 3계단 다해상도 구조를 가질 때 최상위 계층에서 정합을 시도하여 최상위 단계 해상도의 변위도를 얻고 영상의 크기를 점차 복원하며 변위도를 원영상 크기로 증가시킨다. 이 때 중간 단계 근사 영상을 이용하여 경계선 정보 등 구조 정보를 이용하여 변위를 정제하기도 한다. 그리고 이러한 특징 정보들을 적극적으로 이용하기 위해 특징 기반 기법의 정합 요소를 주로 사용한다.

기존의 다해상도 기법은 해상도가 감소한 영상에 대해 정합을 수행하기 때문에 전역적인 최적화 기법을 이용해도 변위 추정까지 걸리는 시간이 짧다는 장점이 있어서 실시간 스테레오 정합에 주로 응용된다. 그러나 최상위 단계에서 잘못 추정된 변위가 전파되면서 최종 변위도의 정확도에 영향을 끼치기 때문에 조밀한 변위도를 얻기 힘들다는 단점이 있다.

제안한 다해상도 기법은 영상 피라미드를 구성하는 각 단계별로 정합을 수행한다. 최상위 단계에서 바로 아래 단계로 전파된 변위를 토대로 정합을 수행하여 변위도를 개선한다. 제안한 방법의 다해상도 기법은 그림 5.5의 (b)와 같다. 각 계층별로 앞 장에서 소개한 유사도 척도와 최적화 기법을 수행한다. 단계간 이동시 3차 회선 보간법을 이용해 변위도를 보간하고 다음 단계 정합을 수행한다. 이러한 방법은 기존의 다해상도 기법보다 계산량이 늘어나지만 변위도의 정확도를 보장할 수 있는 장점이 있다.

기존의 다해상도 기법[56]과 제안한 다해상도 기법으로 얻은 변위도는 그림 5.6과 같다. Tsukuba 영상에 대해 각각의 방법으로 변위를 얻었을 때 기존의 다해상도 기법에 비해 제안한 방법의 변위도가 더 조밀함을 알 수 있다. 특히 기존의 방법은 얇은 물체의 정보가 영상 피라미드를 구축하여 축소하는 과정에서 거의 소실되어 스탠드 거치대에 대한 정합이 거의 이뤄지지 않은 반면, 제안한 방법은 비교적 잘 표현되었다.

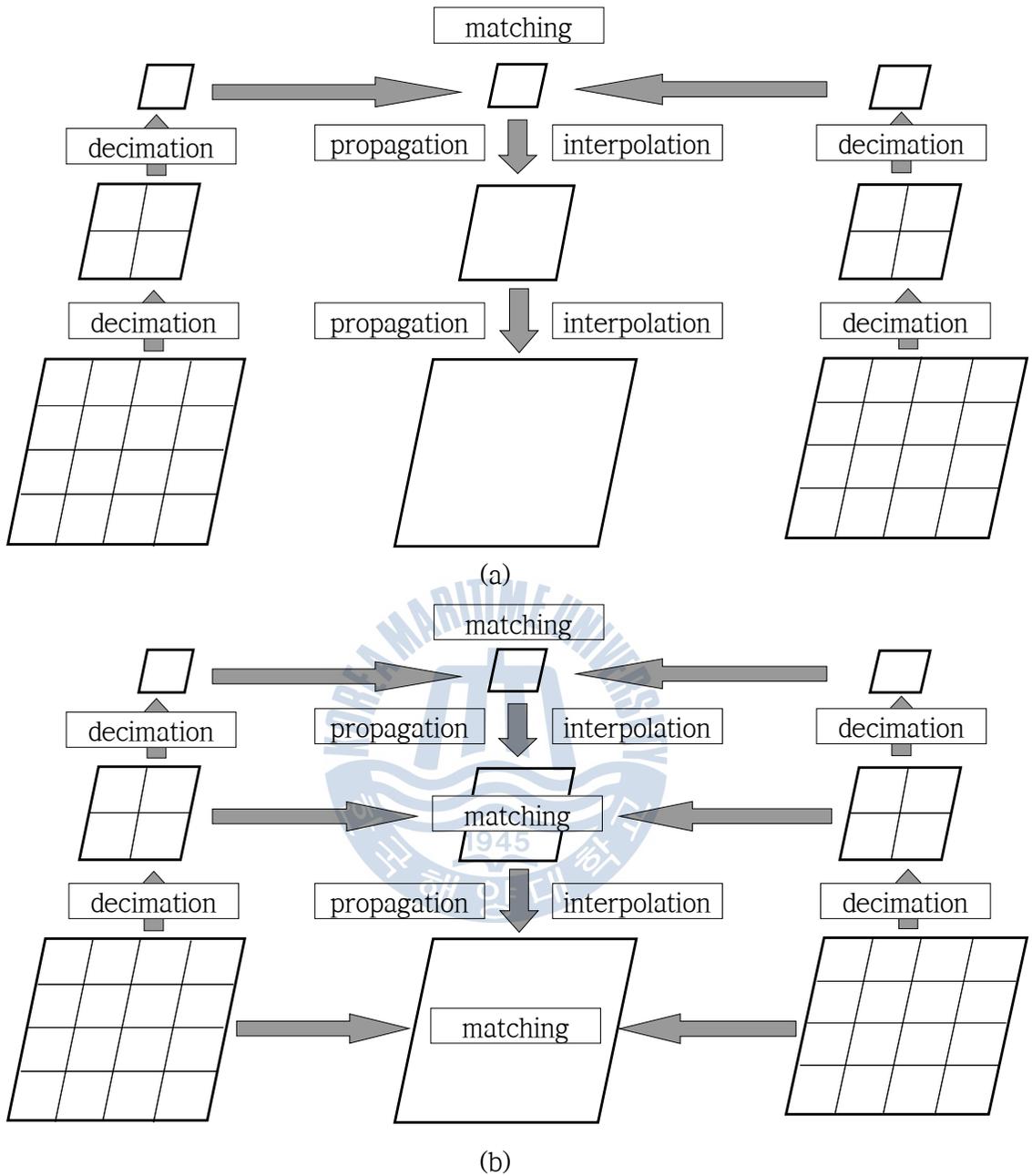


그림 5.5. 기존의 다해상도 기법과 제안한 다해상도 기법

(a) 기존의 다해상도 기법 (b) 제안한 다해상도 기법

Fig. 5.5. Conventional multiresolution and proposed multiresolution method  
 (a) conventional multiresolution method (b) proposed multiresolution method

제안한 다해상도 기법을 기반으로 유사도 척도와 최적화 방법을 적용한 제안한 방법의 순서도는 그림 5.7과 같다. 가우시안 영상 피라미드를 스테레오 영상과 변위도에 대해 구축하고, 현재 레벨  $j$ 의 해상도와 변위 탐색 범위에 대해 유사도 척도를 계산하여 양방향 3차원 DSI를 생성한다. DSI를 토대로 에너지 함수를 계산하고 8방향에 대한 누적 에너지를 최적화하여 변위를 추정하기 위해 8방향 동적 계획법을 적용한다. 각각의 변위 공간에 대해 최적화 과정으로 얻은 변위도  $D_L$ 과  $D_R$ 를 이용하여 폐색 영역을 검출한다. 마지막으로 영상 피라미드의 레벨  $j-1$ 로 이동하면서  $D_L$ 을 다음번 최적화의 변위도로 사용한다.

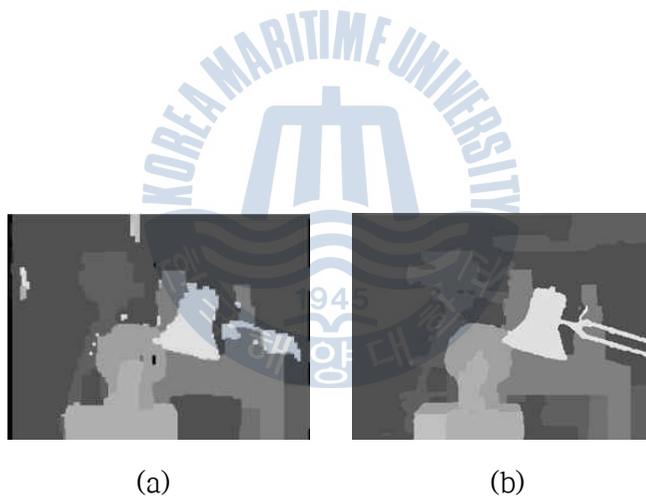


그림 5.6. 기존 다해상도 기법과 제안한 다해상도 기법의 변위도

(a) 기존 다해상도 기법의 변위도

(b) 제안한 다해상도 기법의 변위도

Fig. 5.6. Disparity maps of conventional and proposed multiresolution method

(a) disparity map of conventional method

(b) disparity map of proposed method

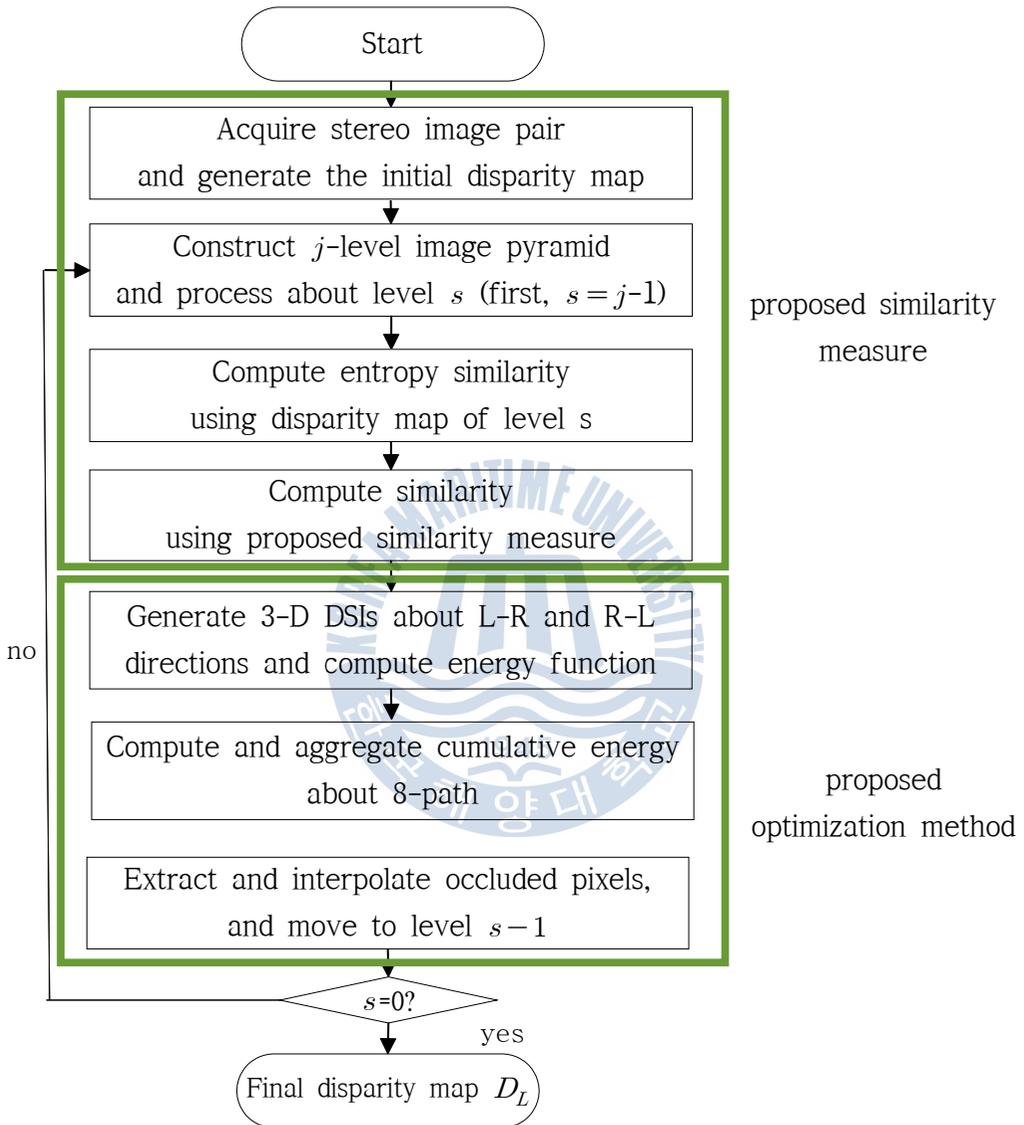


그림 5.7. 제안한 방법의 처리 단계

Fig. 5.7. Summary of processing step of the proposed method

## 제 6 장 실험 및 고찰

실험을 위해 랜덤 닷 스테레오그램(RDS : Random Dot Stereogram), 환경의 변화 없이 획득한 표준 영상, 환경의 변화가 발생한 표준 영상 및 실험실에서 획득한 실제 영상으로 제안한 방법과 기존의 정합 방법에 대해 정합 성능과 계산시간을 비교하였다.

RDS 영상 실험은 제안한 방법의 오차율과 변위도를 기존의 방법들과 비교하기 위해 잡음이 있는 경우와 없는 경우에 대해 수행하였다. 환경 변화가 없는 표준 영상 실험은 조명과 카메라 노출 등 모든 환경 요인들이 동일한 스테레오 영상을 사용하였다. 환경의 변화가 발생한 표준 영상 실험은 카메라 노출 시간과 조명의 변화에 대한 제안한 방법의 성능을 보기 위해 수행하였다. 그리고 실험실에서 획득한 실제 영상은 좌우 카메라의 조명 조건이 동일하고 조리개 개방 정도를 변화시켜 여러가지 조명 환경을 조성하였다. 마지막으로 제안한 다해상도 기법의 효율성을 보이기 위해 기존의 방법들과 계산 시간을 비교하였다. 실험에 사용한 RDS 영상과 실제 촬영 영상은 한국해양대학교 영상처리연구실에서 제작하고 획득했고, 그 외 영상들은 Middlebury college의 Stereo Vision Lab에서 공개한 영상들이다[68].

### 6.1 정합 성능 평가 방법

정합 성능은 정량적인 방법과 정성적인 방법으로 나눠서 평가하였다. 정량적인 방법은 객관적인 비교 분석을 위해 오정합 화소 비율(BPR : bad pixel rate)와 평균제곱오차(MSE : mean square error)를 계산하여 제시하였다[5,68]. 오정합 화소 비율은 비폐색 영역과 변위 불연속 영역에 대해 계산하였고, 비폐색 영역과 변위 불연속 영역에 대한 검증된 기준이 없는 영상들은 변위도 전체에 대한 오정합 화소 비율과 평균제곱오차를 계산하였다.

오정합 화소는 실험으로 얻은 변위도에서 사실 변위도(ground truth)의 변위와 정해진 문턱치 이상 차이를 발생시키는 화소이다. 오정합 화소 비율(BPR)은 이를 전체 화소에 대한 비율로 계산하여 얻을 수 있고 식 (6.1)과 같다.

$$BPR = \frac{1}{n_c} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f(x,y) \quad (6.1)$$

여기서 오정합 화소 여부를 판별하는 함수  $f(x,y)$ 는 다음과 같이 정의한다.

$$f(x,y) = \begin{cases} 1 & , |D(x,y) - D_{gt}(x,y)| > \theta \\ 0 & , o.w. \end{cases} \quad (6.2)$$

여기서  $D(x,y)$ 는 실험으로 얻은 변위도의 위치  $(x,y)$ 의 변위이고  $D_{gt}(x,y)$ 는 사실 변위도의 위치  $(x,y)$ 의 변위이다. 문턱치  $\theta$ 는 모든 실험에서 1로 설정하여 실험으로 얻은 변위가 사실 변위와 1화소 넘는 차이가 발생하면 오정합 화소로 간주하였다.

$M \times N$  크기의 영상에 대한 변위도의 평균제곱오차는 식(6.3)과 같다.

$$MSE = \frac{1}{MN} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} [D(x,y) - D_{gt}(x,y)]^2 \quad (6.3)$$

정성적인 방법은 사실 변위도에 입각하여 제안한 방법으로 얻은 변위도와 다른 정합 방법에 의한 변위도를 비교하였다. 그리고 실험 영상의 질감 특성이나 특징 분포를 고려하여 적용한 방법들의 성능과 장단점을 설명하였다.

본 장의 모든 실험에서 제안한 방법이 사용한 Haar 웨이블릿 특징 벡터는 그림 3.8의 (d)와 같은 36차원의 형태로 생성하였다. 그 외 모든 실험에 대해서 제안한 방법에서 사용하는 매개 변수는 표 6.1과 같다.

**표 6.1** 제안한 방법의 매개변수 설정

**Table 6.1** Parameter setting about proposed method

parameter entry	value
the number of image pyramid levels	3
standard deviaton $\sigma$ for entropy	6
window size for Census transform	$5 \times 5$
$R_1, R_2, \Sigma_1, \Sigma_2$	7,3,3,1
window size $m \times m$	$13 \times 13$
Haar wavelet filter size $h \times h$	$4 \times 4$
penalty constant $P_1, P_2$	1.5, 3.8

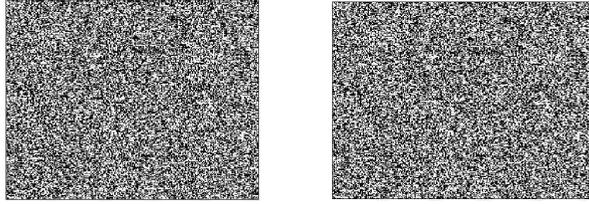
## 6.2 스테레오 정합 실험

### 6.2.1 RDS 영상 실험

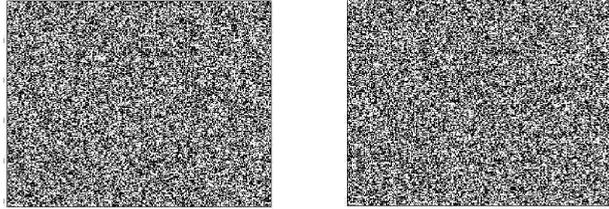
#### 1) 실험 영상과 실험 방법

실험에 사용한 영상은 0 또는 255의 명암도를 갖는 화소들의 비율을 각각 50%로 설정하고 화소들의 위치를 무작위로 분포시켜 만든 50% RDS 합성 영상이다. RDS 영상은 시차를 제외한 다른 매개변수들이 동일하므로 제안한 방법의 성능과 특성을 명확히 확인할 수 있다. 실험은 잡음 없는 경우와 잡음이 20% 첨가된 경우에 대해 수행하였다.

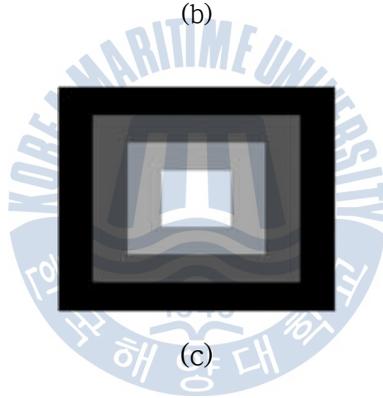
정성적인 성능 평가는 NCC[16]와 MI[37]를 이용한 방법으로 얻은 변위도와 제안한 방법으로 얻은 변위도와 비교하였고, 정량적인 성능 평가를 위해 사실 변위도에 대한 오정합 화소 비율과 평균 제곱 오차를 계산하였다. 실험에 사용한 RDS 영상과 사실 변위도는 그림 6.1과 같다. 그림 6.1의 (a)는 잡음이 없는 영상이고, 그림 6.1의 (b)는 우영상에 20%의 salt-and-pepper 잡음을 첨가한 스테레오 영상이다. 사실 변위도는 그림 6.1의 (c)와 같다. 변위도는 크기가 다른



(a)



(b)



(c)

그림 6.1. 실험에 사용한 50% RDS 영상들과 사실 변위도

(a) 잡음이 없는 좌영상과 우영상

(b) 잡음이 없는 좌영상과 20% salt-and-pepper 잡음이 섞인 우영상

(c) 사실 변위도

Fig. 6.1. 50% Random Dot Stereogram sets and ground truth

(a) left image and right image without noise

(b) left image without noise and right image with 20% salt-and-pepper noise

(c) ground truth

평면들이 피라미드 형태로 쌓여있고 밝을수록 변위가 큰 것을 의미한다. 제안한 방법과 실험에 사용한 다른 방법들은 잡음이 없는 경우와 잡음이 있는 경우에 대해 동일하게 적용되었다. NCC를 이용한 방법은 WTA로 변위를 추정하였고 MI는 그래프 컷(GC)[58]을 이용하여 변위를 추정하였다. 그 밖에 RDS 실험에 적용한 영상의 속성과 다른 방법들의 실험 조건은 표 6.2와 같다.

표 6.2 50% RDS에 대한 실험 조건

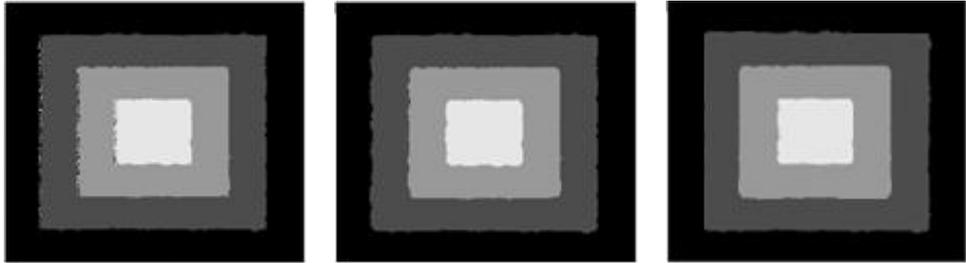
Table 6.2 Experimental condition about 50% RDS

entry	value
image size	256 × 256
real range of disparity	[0, 9]
search range of disparity	[0, 12]
window size of NCC	7 × 7
standard deviaton $\sigma$ for MI	10

## 2) 50% RDS 실험에 대한 결과

잡음이 없는 50% RDS 실험에서 NCC를 이용하여 얻은 변위도는 그림 6.2의 (a)와 같다. NCC는 변위도의 각 평면 좌측에서 일정한 두께의 오정합을 발생시켰다. 이는 윈도우 기반 정합에서 흔히 나타나는 현상이고 오정합 영역의 두께는 적용한 윈도우의 크기와 밀접한 관계를 가진다. 기존의 MI를 이용하여 얻은 변위도는 그림 6.2의 (b)이고 제안한 방법으로 얻은 변위도는 그림 6.3의 (c)와 같다. 제안한 방법의 변위도는 MI를 이용하여 얻은 변위도와 비교했을 때 시각적으로 큰 차이를 보이지 않았으나 평면들의 경계선을 비교적 뚜렷하게 표현하였다. 제안한 유사도 척도는 기존의 MI에서 고려하지 않는 공간 정보를 사용하기 때문에 RDS와 같이 복잡한 질감을 띄는 영상에서 물체의 경계선과 같은 구조적인 특성을 잘 반영한다.

우영상에 20%의 salt-and-pepper 잡음을 첨가하여 수행한 실험에서 NCC로



(a) (b) (c)

그림 6.2. 50% RDS 변위도

(a) NCC (b) MI (c) 제안한 방법

Fig. 6.2. Disparity maps of 50% Random Dot Stereogram

(a) NCC (b) MI (c) proposed method



(a) (b) (c)

그림 6.3. 20% 잡음이 첨가된 50% RDS의 변위도

(a) NCC (b) MI (c) 제안한 방법

Fig. 6.3. Disparity maps of 50% Random Dot Stereogram with 20% noise

(a) NCC (b) MI (c) proposed method

얻은 변위도는 그림 6.3의 (a)과 같다. 우영상에 분포된 잡음으로 인해 NCC는 오정합을 많이 발생시켰다. MI를 이용하여 얻은 변위도는 그림 6.3의 (b)와 같다. MI는 NCC에 비해 잡음에 영향을 비교적 덜 받았지만 평면의 경계선 모양을 정확히 표현하지 못했다. MI는 명암도 분포 특성을 이용하는 유사도 척도이기 때문에 정합점 주변 화소의 변형에 영향을 덜 받지만 공간 정보를 반영하지 않기 때문에 물체의 경계선과 같은 변위가 변하는 영역의 구조적인 특성을 정확히 반영하기 어렵다. 제안한 방법으로 얻은 변위도는 그림 6.3의 (c)이다. 잡음이 없는 경우와 큰 차이가 없었으며 기존의 MI보다 경계선을 잘 보전했다. 이는 경계선 영역에서 정합 성능이 좋은 변형된 Census 변환의 장점이 잘 반영되었기 때문이라 판단된다. 본 실험에서 얻은 변위도들의 오정합 화소 비율과 평균제곱오차는 표 6.3과 같다.

**표 6.3** 50% RDS 실험에서 얻은 MI, NCC와 제안한 방법의 변위도의 오차율

**Table 6.3** Error rates about disparity maps of MI, NCC, and proposed method in test using 50% RDS

	50% RDS		50% RDS with 20% noise	
	BPR(%)	MSE	BPR(%)	MSE
MI[37]+GC[58]	1.40	1.23	6.33	3.72
NCC[16]+WTA	6.21	3.41	11.89	10.42
<b>proposed method</b>	<b>1.13</b>	<b>0.75</b>	<b>2.87</b>	<b>2.56</b>

오정합 화소 비율과 평균제곱오차를 통해 MI와 제안한 방법의 성능이 NCC보다 우수하였다. MI는 화소단위 유사도 척도이고 제안한 방법 또한 화소단위 엔트로피를 기반으로 하는 유사도 척도를 사용하기 때문에 대응점 주변 화소의 변형이 발생해도 오정합의 빈도가 적다. 잡음이 발생한 경우에서도 MI와 제안한 방법의 오정합 화소 비율은 각각 6.33%와 2.87%로 11.89%으로 계산된 NCC보다 작았다. 특히 제안한 방법은 기존의 MI의 오정합 화소 비율을 2배 이상

감소시킨 사실을 통해 엔트로피 기반의 유사도 척도에 공간 정보를 포함시켰을 때 잡음에 강인하고 정합 성능이 향상됨을 알 수 있다. 또한 제안한 방법의 평균제곱오차는 두 실험에서 각각 0.75와 2.56으로 가장 작았다. 이는 에너지 함수의 변위 연속성을 적용할 때 잘못 추정된 변위의 영향이 적음을 의미한다.

## 6.2.2 환경 변화가 없는 표준 영상 실험

### 1) 실험 영상과 실험 방법

실험에 사용된 영상은 조명과 카메라 노출 시간이 일정한 환경의 변화가 없는 표준 영상 20개이고, 합성 영상과 실제 촬영 영상이 포함되어 있다. 이 중 Tsukuba, Venus, Teddy, Cone 등 4개의 영상은 비폐색 영역과 변위 불연속 영역에 대한 정보가 공개되어 있어 이 영역에 대해 오정합 화소 비율을 계산하였고 다른 실험과 구별하기 위해 4개의 영상을 set A라 칭하였다. 그 외 16개의 영상에 대해서는 검증된 비폐색 영역과 변위 불연속 영역의 기준이 없고 전체 변위도만 공개되어 있는 까닭에 전체 변위도에 대한 오정합 화소 비율과 평균 제곱오차를 계산하였다. 이 16개 영상은 set B로 분류하였다.

실험에 사용한 표준 영상들은 그림 6.4와 그림 6.5와 같다. set A의 4개 영상과 사실 변위도들은 그림 6.4이고 set B의 16개의 영상과 사실 변위도들은 그림 6.5와 같다. set A를 구성하는 영상들은 물체들의 변위 분포가 고르면서 최대 변위가 40화소 미만이라 시차에 따른 기하학적 변환이 거의 없고 합성 영상 Venus가 포함되어 있어서 카메라 표본화 등의 영향없이 알고리즘 자체의 정합 성능을 평가할 수 있다. set B의 영상들은 최대 변위가 70화소 이상이고 대부분의 물체들이 비스듬히 기울어져 있기 때문에 물체 내에서 점진적으로 변하는 변위를 추정하는 성능을 평가할 수 있다.

제안한 유사도 척도의 성능을 확인하기 위해 기존의 유사도 척도들과 동일한 최적화 기법을 적용하여 결과를 비교하였다. 이를 위해 사용한 기존의 유사도 척도는 NCC와 MI이고 제안한 유사도 척도는 HEC라 칭하였다. 최적화 방법은



그림 6.4. 비폐색 영역과 변위 불연속 영역의 척도가 있는 set A의 영상들과 사실 변위도

Fig. 6.4. Stereo image pairs and ground truths of set A which has measurements about unoccluding regions and disparity discontinuity regions

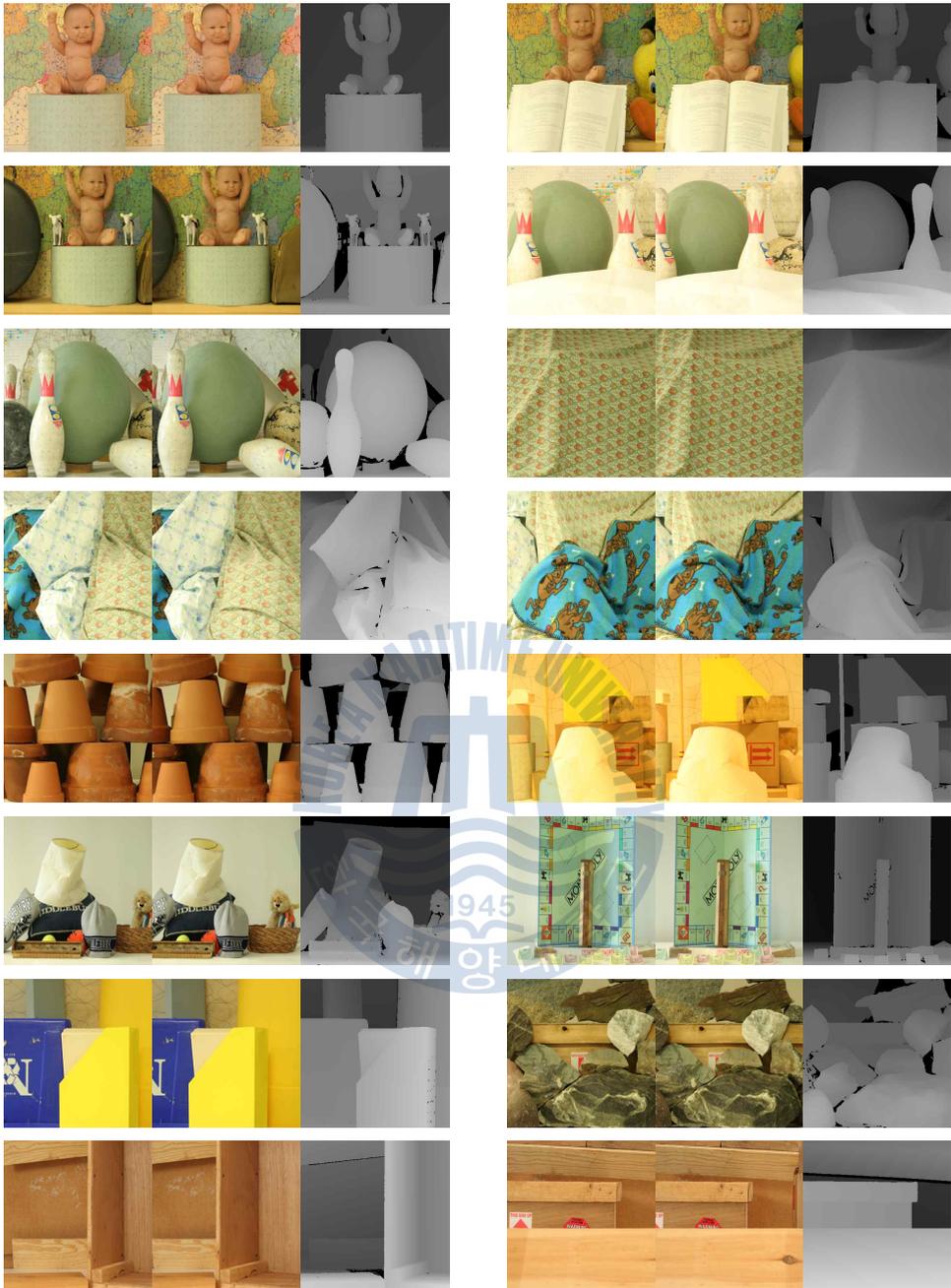


그림 6.5. 전체 변위도의 오정합 화소 비율과 평균제곱오차를 계산한  
set B의 영상들과 사실 변위도

Fig. 6.5. Stereo image pairs and ground truths of set B which has evaluated  
BPR and MSE about all disparity maps

그래프 컷을 적용하였다. 제안한 최적화 방법의 성능은 제안한 유사도 척도를 그래프 컷에 적용한 결과와 비교하여 측정하였다. 제안한 최적화 방법의 명칭은 구분을 위해 8path DP라 표기하였다. 그리고 제안한 방법의 성능을 보이기 위해 신뢰 동적 계획법(RDP)[25], 양방향 신뢰 확산법(DoubleBP)[18] 및 적응 신뢰 확산법(AdaptingBP)[17] 등 기존 정합 방법들과 비교하였다.

## 2) 환경 변화가 없는 표준 영상에 대한 실험 결과

합성 영상 Venus의 좌영상과 사실 변위도는 그림 6.6의 (a)와 (b)와 같다. 이 영상은 프로그램으로 제작하였기 때문에 촬영시 발생하는 아웃라이어나 카메라 표본화 등의 영향을 받지 않아 이상적인 스테레오 영상에 대한 알고리즘 성능을 확인할 수 있다. Venus 영상은 좌상단의 무질감 영역을 포함하고 있고 우측의 신문으로 인한 폐색 영역이 크다. 그리고 물체들이 깊이 방향으로 비스듬히 놓여져 있기 때문에 변위가 같은 물체 내에서도 다르게 나타난다.

오정합 화소 비율을 계산하기 위해 사용된 사실 변위도, 기존의 MI, RDP, DoubleBP, AdaptingBP를 이용하여 얻은 변위도는 그림 6.6의 (b)~(f)와 같다. MI와 AdaptingBP는 우측 신문의 변하는 변위를 반영하지 못했다. RDP는 곳곳에서 약한 스트리킹 현상을 보이고, RDP 이외의 기법들은 우측상단 무질감 영역의 변위를 정확히 구하지 못했다. 각 물체의 경계선을 보전하는 성능은 실험에 사용한 모든 기법들이 크게 차이가 없으나 제안한 방법이 물체 내에서 변위가 변하는 특성을 잘 반영했음을 제안한 방법으로 얻은 변위도인 그림 6.6의 (i)로 확인할 수 있다.

NCC와 제안한 유사도 척도인 HEC의 성능을 비교하기 위해 이들을 그래프 컷에 적용하여 변위도를 얻었고 이는 각각 그림 6.6의 (g)와 그림 6.6의 (h)와 같다. NCC는 폐색 영역이나 물체 경계 근처에서 변위가 급격히 변하면 유사도를 정확히 반영하지 못하므로 배경과 변위의 차이가 큰 우측 신문의 경계에서 성능이 좋지 않았다. 그림 6.6의 (g)에서 발생한 신문의 경계에서 발생한 오정합이 그림 6.6의 (h)에서 개선되었다. 제안한 유사도 척도는 경계 영역에서 주

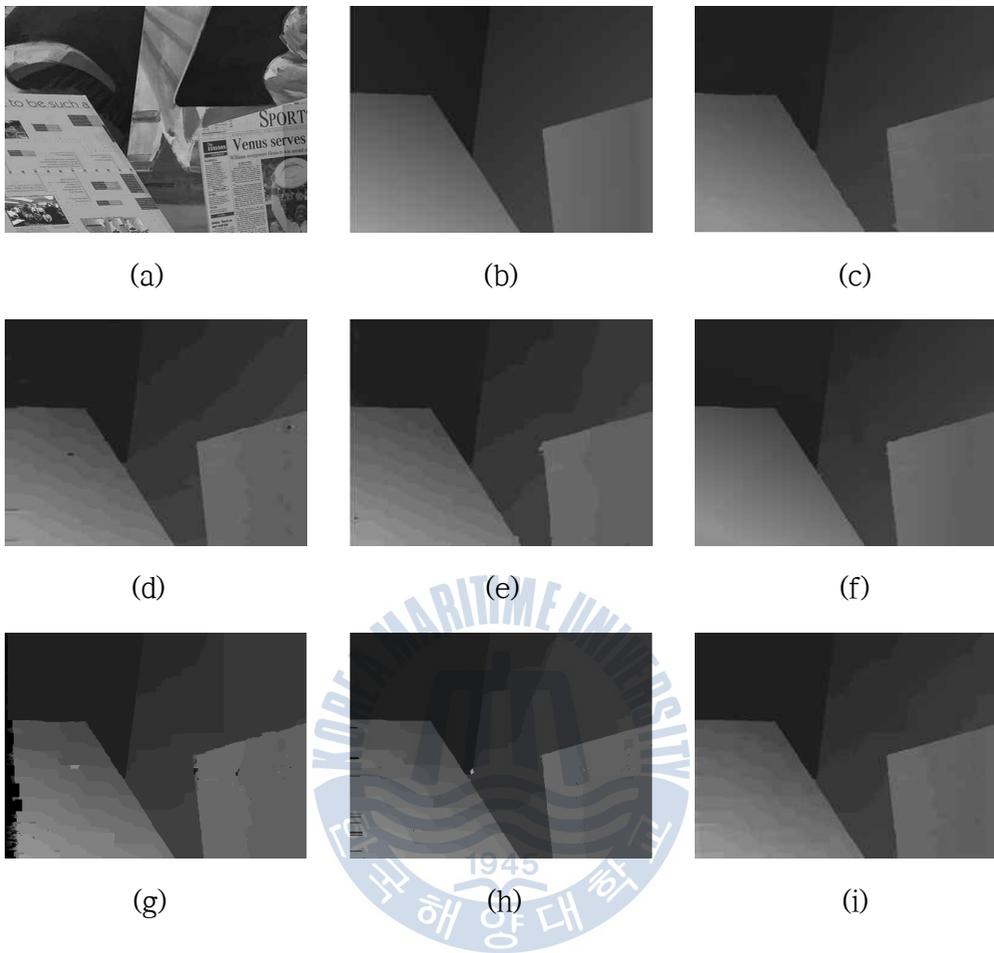


그림 6.6. Venus 영상의 좌영상, 사실 변위도와 변위도들

(a) 좌영상 (b) 사실 변위도 (c) MI (d) RDP (e) DoubleBP

(f) AdaptingBP (g) NCC+GC (h) HEC+GC (i) 제안한 방법(HEC+8path DP)

((a)와 (b)는 Middlebury Stereo Vision Lab에서 공개)

Fig. 6.6. Left image of Venus, ground truth and disparity maps

(a) left image (b) ground truth (c) MI (d) RDP

(e) DoubleBP (f) AdaptingBP (g) NCC+GC (h) HEC+GC

(i) proposed method(HEC+8path DP)

(Image courtesy of (a) and (b) Middlebury Stereo Vision Lab)

변 물체의 화소값이 변해도 그 영향을 적게 받는 화소단위 유사도 척도이고 Haar 웨이블릿 특징 벡터와 Census 변환을 이용하여 공간 정보를 반영하기 때문에 기존의 NCC와 MI에 비해 경계선 영역에서 정합 성능이 우수하다.

그래프 컷을 최적화 방법으로 적용했을 때 우측상단 뒷배경의 변위를 정확히 추정하지 못함을 그림 6.6의 (g)와 그림 6.6의 (h)를 통해 확인할 수 있다. 이는 그래프 컷이 경계선과 경계선 사이의 변위를 동일하게 추정하는 특성에 기인한 오정합이다. 그래프 컷은 동일한 변위를 가지는 화소 사이에 다른 변위가 존재할 때 해당 화소들을 배제하는 과정을 거치므로 점진적으로 변위가 변하는 영상에서 조밀한 변위를 얻기 힘들다. 제안한 방법의 변위도는 그림 6.6의 (i)와 같다. 8path DP가 대각선 방향으로 변위가 일정한 우측 상단 영역 등 경계선 영역의 변위를 다른 방법들보다 비교적 잘 추정하였다. 또한 그래프 컷이 반영하지 못한 우측 상단의 점진적으로 변하는 변위를 잘 반영하였다. 이는 제안한 최적화 방법이 에너지 통합시 변위 차이에 대한 에너지 증가량을 변별적으로 계산하는데 기인한다.

실제 촬영한 표준 영상 중 Tsukuba 영상의 좌영상은 그림 6.7의 (a)와 같다. 이 영상은 뒷배경에 다양한 패턴의 질감이 분포되어 있고, 얇은 물체들이 비슷한 색의 배경과 겹쳐져 있다. 따라서 다양한 형태의 질감을 유사도가 잘 반영하는지 여부와 변위 불연속 영역에 대해 변위를 추정할 수 있는 성능을 볼 수 있다.

사실 변위도는 그림 6.7의 (b)와 같다. MI, RDP, DoubleBP, AdaptingBP, 그리고 제안한 방법으로 얻은 변위도는 그림 6.7의 (c)부터 그림 6.7의 (f)의 순서로 나열하였다. 기존의 MI 기법으로 얻은 변위도에서 스탠드 주변과 뒷배경에서 오정합이 많이 관찰되었다. 특히 상단 중앙의 책장틀 부분에서 변위가 일정하지 않았고 우측 상단의 질감이 없는 배경 부분에서도 오정합이 발생하였다. RDP로 얻은 변위도는 그림 6.7의 (d)와 같다. 동적 계획법의 단점인 스트리킹 현상이 발생하였다. DoubleBP와 AdaptingBP에 의한 변위도는 그림 6.7의 (e)와 그림 6.7의 (f)와 같다. 영상 중앙의 캠코더 주변과 영상의 좌측 상단의 책장틀

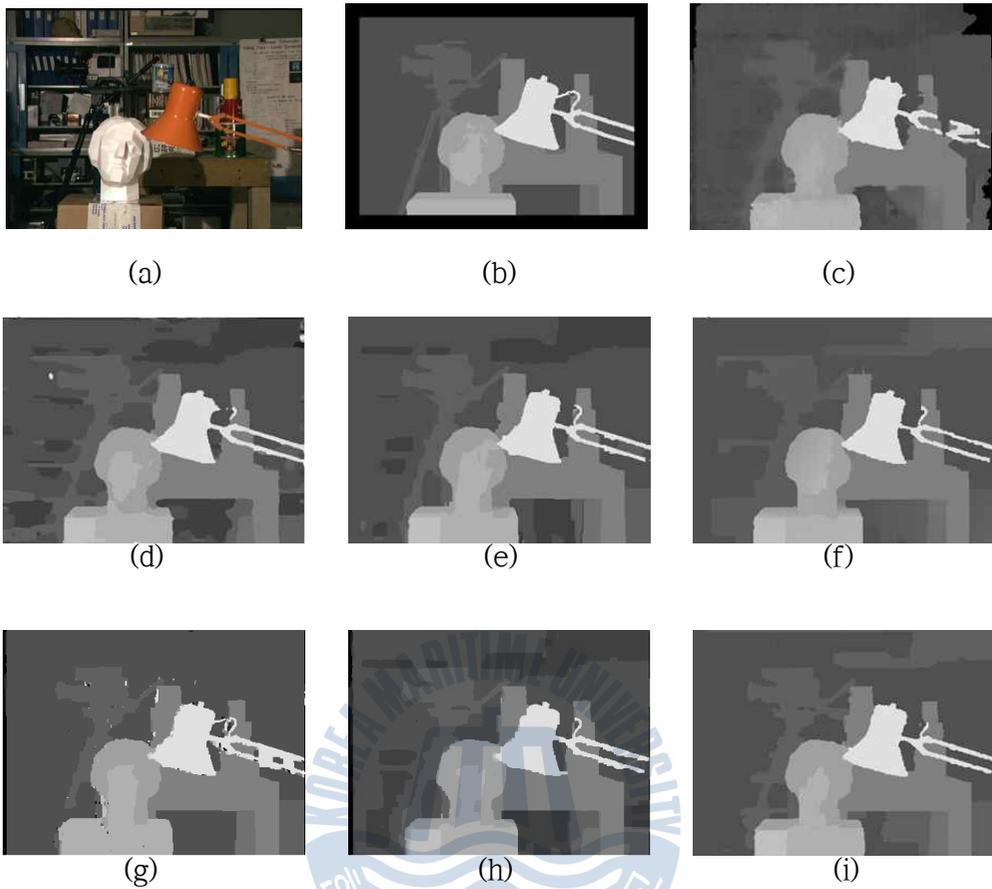


그림 6.7. Tsukuba 영상의 변위도

(a) 좌영상 (b) 사실 변위도 (c) MI (d) RDP (e) DoubleBP (f) AdaptingBP  
 (g) NCC+GC (h) HEC+GC (i) 제안한 방법(HEC+8path DP)

((a)와 (b)는 Middlebury Stereo Vision Lab에서 공개)

**Fig. 6.7.** Disparity maps of Tsukuba

(a) left image (b) ground truth (c) MI (d) RDP

(e) DoubleBP (f) AdaptingBP (g) NCC+GC (h) HEC+GC

(i) proposed method(HEC+8path DP)

(Image courtesy of (a) and (b) Middlebury Stereo Vision Lab)

영역에서 오정합을 발생시켰다. 이는 특정 영역에 대한 정확하지 않은 정보가 최종 변위도에 영향을 주는 BP의 특성으로 인해 발생했다.

HEC를 그래프 컷에 적용하여 얻은 변위도는 그림 6.7의 (h)이다. 그림 6.7의 (g)와 비교했을 때 스탠드와 캠코더의 경계선 영역에서 더 좋은 성능을 보였다. 특히 스탠드 거치대와 뒷배경의 흰 게시판이 겹치는 질감이 약한 영역에서 MI와 NCC는 오정합을 유발했지만 HEC는 배경과 물체를 비교적 잘 구분하였다. 이를 통해 제안한 유사도 척도의 공간 정보가 다양한 질감 영역과 경계선 영역의 정합 성능을 향상시킨다는 사실을 확인하였다.

제안한 방법으로 얻은 변위도는 그림 6.7의 (i)이다. 제안한 최적화 방법인 8path DP는 스트리킹 현상을 개선하고 변위를 정확히 추정하였다. 특히 변위 불연속 영역에서 다른 방법들보다 성능이 우수하였다. 이 영상과 같이 물체의 경계선들이 직선의 형태로 존재하는 영상에서 제안한 최적화 기법의 변위 추정 성능이 좋다.

실험에 적용한 방법들의 성능을 정량적으로 비교하기 위해 총 20개의 영상 중 Venus, Tsukuba, Teddy, Cone 영상에 대해서는 비폐색 영역과 변위 불연속 영역에 대한 오정합 화소 비율을 각각 구하여 그 평균을 계산하였고, 나머지 16개 영상에 대해서는 사실 변위도와와의 오정합 화소 비율과 평균제곱오차의 평균을 계산하였다. 이를 표 6.4에 나타내었다.

**표 6.4** 표준 영상의 오정합 화소 비율(%)과 평균제곱오차(pixels)의 평균

**Table 6.4** Average BPR(%) and MSE(pixels) of testbed images

	set A		set B	
	unocc.(%)	disc.(%)	BPR(%)	MSE
RDP[25]	2.14	6.72	4.75	0.47
DoubleBP[18]	1.86	6.01	4.19	0.23
AdaptingBP[17]	1.98	6.59	4.23	0.26
NCC+GC[58]	6.98	13.84	11.4	0.64
MI+GC[37]	3.04	11.33	7.50	0.62
HEC+GC	2.79	8.83	4.42	0.31
<b>proposed method</b>	<b>2.05</b>	<b>5.96</b>	<b>3.63</b>	<b>0.14</b>

set A로 분류된 영상들에 대한 실험에서 제안한 유사도 척도인 HEC의 성능은 비폐색 영역과 변위 불연속 영역에서 NCC와 MI보다 좋았다. 그리고 변위 불연속 영역에서 제안한 방법의 유사도 척도의 오정합 화소 비율은 NCC에 비해 5.01% 낮았고 기존의 MI와는 2.5%의 차이를 보였다. 이는 제안한 유사도 척도가 Haar 웨이블릿 특징 벡터와 Census 변환을 통해 계산한 공간 정보를 활용함으로써 다양한 질감 영역 및 경계선 영역 등 구조적인 특성을 띠는 영역에서 기존의 유사도 척도보다 우수한 성능을 보임을 알 수 있다.

변위 불연속 영역의 오차율은 제안한 최적화 방법을 적용했을 때 8.83%에서 5.96%으로 감소하였다. 이는 기존의 경계선 영역의 변위를 잘 추정한다고 알려져 있는 그래프 컷에 비해 제안한 방법의 변위 추정 성능이 더 우수함을 의미한다. 그리고 제안한 최적화 방법으로 얻은 변위 불연속 영역의 오차는 실험에 적용한 모든 방법들 중 가장 낮다. 제안한 최적화 방법이 경계선 방향의 일정한 변위를 잘 추정하고 변위 불연속 영역을 잘 보존한다는 사실을 수치를 통해 확인하였다.

set B에 대한 실험에서 제안한 유사도 척도의 오정합 화소 비율은 전체 변위도에 대해 NCC보다 약 7% 낮았고 MI보다 3%정도 낮았다. 그리고 제안한 최적화 방법을 적용했을 때 4.42%에서 3.63%로 더 감소하였다. set B에 대해 제안한 방법의 전체 오정합 화소 비율과 평균제곱오차가 가장 작았다. 실험 결과 제안한 방법은 환경 변화가 없는 영상에 대해 우수한 성능을 보였다.

### 6.2.3 환경 변화가 발생한 표준 영상 실험

#### 1) 실험 영상과 실험 방법

실험에서 가정한 환경 변화는 조명의 변화, 카메라 노출 시간의 변화이다. 조명과 카메라 노출 시간은 각각 세가지 경우를 색인으로 구분하였다. 카메라 노출 시간에 대한 색인과 구체적인 노출 시간은 표 6.5와 같다.

표 6.5 카메라 노출 색인에 대한 시간 조건(msec)

Table 6.5 Exposed time condition about camera exposure index (msec)

	e_1	e_2	e_3
Aloe	200	800	3200
Art	250	1000	4000
Dolls	250	1000	4000
Cloth4	200	800	3200
Moebius	200	800	3200
Laundry	200	800	3200
Rock1	250	1000	4000
Reindeer	250	1000	4000

조명은 L\_1, L\_2, L\_3으로 표시된다. L\_1은 표준광을 이용하여 정면에 비치도록 했고 L\_2는 조명의 색을 달리했다. L\_3은 L\_1과 동일한 조명이지만 방향이 다르기 때문에 그림자의 위치로 인한 지역적인 밝기 변화를 발생시킨다. 실험에서 각각의 환경 변화에 따른 성능을 따로 확인하기 위해 조명 변화에 대한 실험은 카메라 노출 시간을 색인 e\_2에 고정하여 조명 색인에 대한 좌영상과 우영상 조합을 구성하였고, 카메라 노출 시간에 대한 실험에서는 조명 색인을 L\_1에 고정하고 카메라 노출 색인에 대한 실험 영상 조합을 구성하였다. 단 Aloe 영상에 대해서는 카메라 노출 시간을 e\_1에 고정하였다. 조명 색인과 카메라 노출 색인에 대한 영상의 예는 그림 6.8과 같다. 한 장의 영상은 조명과 카메라 노출 색인의 조합을 이용해 9종류를 만들어낼 수 있다.

실험에 사용한 영상은 Aloe, Art, Dolls, Cloth4 등 8세트이고 그림 6.9와 같다. 실험에 사용한 영상들의 최대 변위는 가장 적은 영상이 40이고 가장 큰 영상은 94이다. 성능 평가는 실험으로 얻은 변위도의 비폐색영역에 대한 오정합 화소 비율을 계산하였고, 시각적인 비교를 위해 변위도를 제시하였다. 제안한 방법의 유사도 척도인 HEC의 성능을 다른 유사도 척도들과 비교하기 위해 HEC를 포함한 모든 유사도 척도를 그래프 컷에 적용한 결과를 제시하였고, 최종적으로

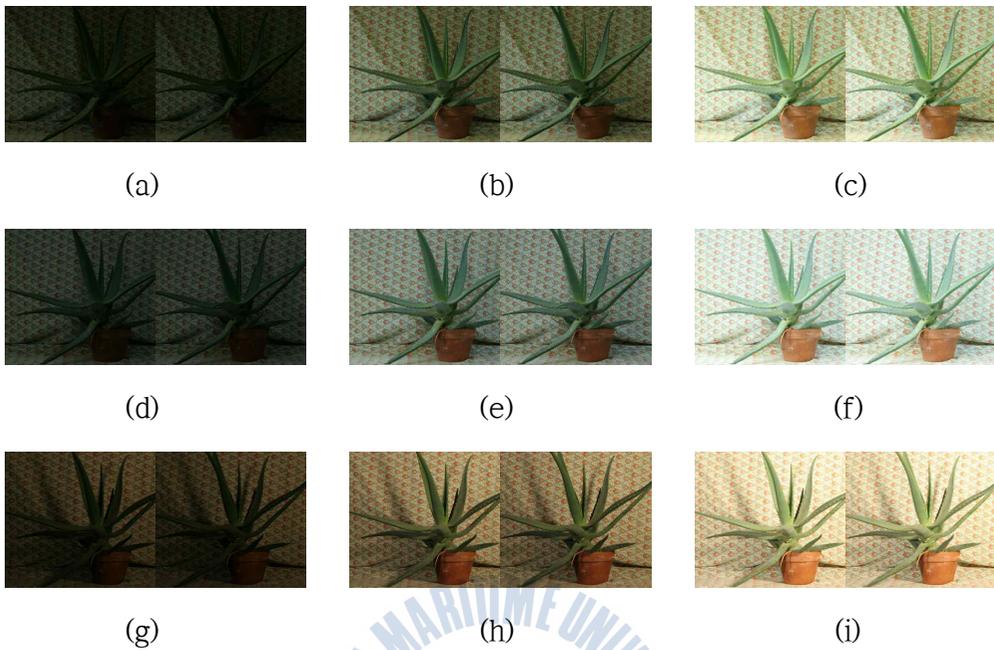


그림 6.8 조명과 카메라 노출 색인에 대한 영상의 예

- (a)  $I_1$  &  $e_0$  (b)  $I_1$  &  $e_1$  (c)  $I_1$  &  $e_2$   
 (d)  $I_2$  &  $e_0$  (e)  $I_2$  &  $e_1$  (f)  $I_2$  &  $e_2$   
 (g)  $I_3$  &  $e_0$  (h)  $I_3$  &  $e_1$  (i)  $I_3$  &  $e_2$

(Middlebury Stereo Vision Lab에서 공개)

Fig. 6.8. An example of test image pair about illumination index and camera exposure index

- (a)  $I_1$  &  $e_0$  (b)  $I_1$  &  $e_1$  (c)  $I_1$  &  $e_2$   
 (d)  $I_2$  &  $e_0$  (e)  $I_2$  &  $e_1$  (f)  $I_2$  &  $e_2$   
 (g)  $I_3$  &  $e_0$  (h)  $I_3$  &  $e_1$  (i)  $I_3$  &  $e_2$

(Image courtesy of Middlebury Stereo Vision Lab)



(a)



(b)



(c)



(d)

그림 6.9. 실험 영상과 사실 변위도

(a) Aloe (b) Art (c) Dolls (d) Cloth4

(e) Moebius (f) Laundry (g)Rock1 (h) Reindeer

(Middlebury Stereo Vision Lab에서 공개)

Fig. 6.9. Test image pairs and ground truths

(a) Aloe (b) Art (c) Dolls (d) Cloth4

(e) Moebius (f) Laundry (g)Rock1 (h) Reindeer

(Image courtesy of Middlebury Stereo Vision Lab)



(e)



(f)



(g)



(h)

그림 6.9. (계 속)  
Fig. 6.9. (continued)

제안한 방법의 결과와 비교하였다. 비교를 위해 사용된 기존의 유사도 척도는 NCC[58], MI[37], LoG/BT[45], Rank/BT[46]과 AWSM[34]이다.

## 2) 카메라 노출 시간의 차이가 발생한 영상 실험 결과

카메라 노출 시간의 차이로 발생한 영상의 변형은 전역적으로 나타난다. 노출 시간을 극단적으로 다르면 밝거나 어두운 저대비 영상을 얻는다. 실험에 사용한 카메라 노출 시간은 표 6.5에 제시한 3단계로 설정하였다. 예를 들어 좌영상의 색인이  $e_1$ 이고 우영상의 색인이  $e_2$ 일 경우  $e_1/e_2$ 로 표기하였다.

실험에 사용한 Aloe 영상은 그림 6.10의 (a)와 (b)이다. 이 영상은 뒷배경의 질감이 반복되는 형태를 띠고 화분 왼쪽의 뒷배경이 앞으로 돌출이 되어있어 다른 영역과 변위의 차이가 미세하게 발생한다. 화분은 무질감 영역이고 알로에 줄기들이 시차에 따라 폐색 영역으로 존재한다. 그리고 왼쪽 영상의 카메라 노출 시간이 짧아서 전체적으로 어둡다.

기존의 MI를 이용하여 얻은 변위도가 그림 6.10의 (d)이다. 전체적으로 정합이 잘 이뤄졌지만 곳곳에 아주 밝거나 어두운 부분에서 오정합이 발생하였다. Rank/BT와 LoG/BT는 밝기 차이가 극단적인 영상에서 오정합이 많이 발생했다. LoG/BT는 아주 밝거나 어두운 영역에서 명암도의 2차 미분치를 정확히 계산하지 못하고, Rank/BT는 Rank 변환을 이용하기 때문에 카메라 노출 시간의 변화로 대비의 폭이 좁아진 영상에서는 기존의 명확한 차이를 보이던 중간 순위 명암도의 순위가 같아지는 현상 등이 발생하므로 유사도를 정확히 반영하지 못한다. 특히 Rank/BT의 경우 알로에 잎의 경계선 부근에서 오정합이 많다.

제안한 방법으로 얻은 변위도는 그림 6.10의 (i)이다. 기존의 MI를 이용한 방법과 큰 차이가 없지만 경계선 부근에서 오정합이 감소하였다. 특히 화분의 윗부분 알로에 줄기가 시작되는 영역에서 다른 방법들에 비해 비교적 정확한 변위를 얻었다. 이 영역에서는 밝기 차이뿐만 아니라 대비의 차이도 발생하였다. 공간 정보와 컬러 정보를 이용한 적응적 가중치를 이용하는 AWSM은 이러한 문제를 개선했으나 시차로 인해 변한 배경 영역의 질감을 정확히 반영하지 못

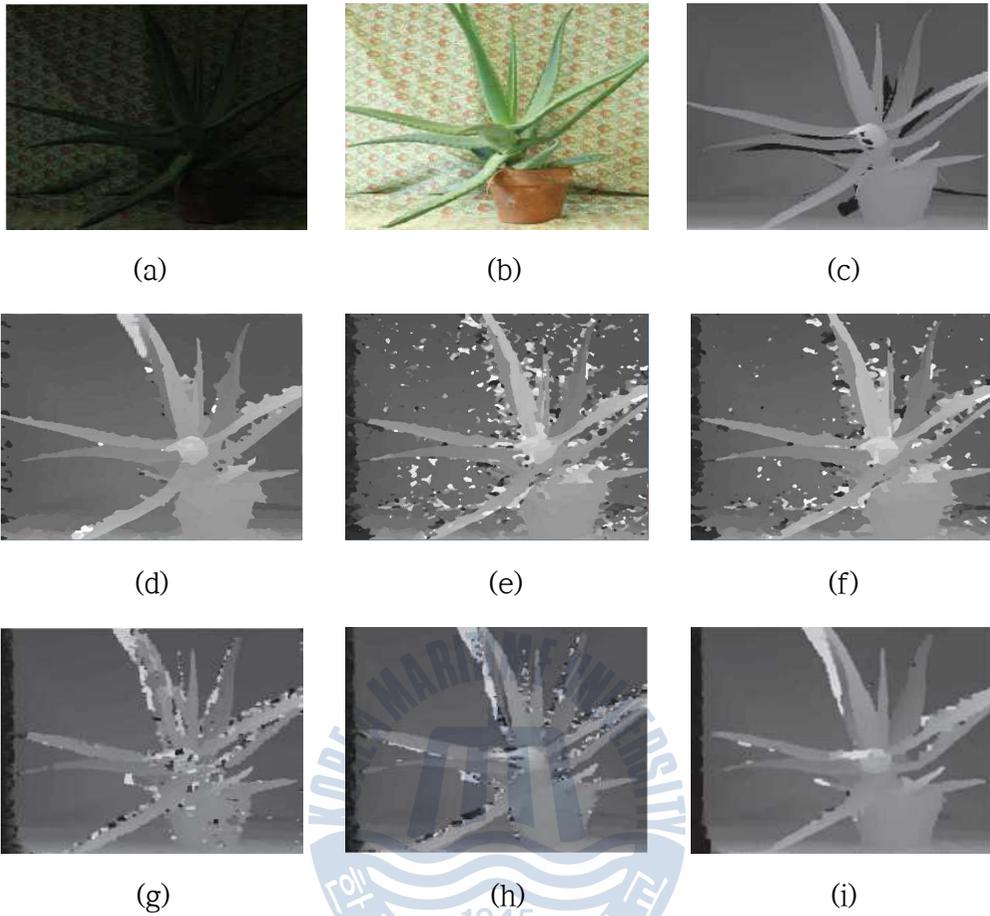


그림 6.10. Aloe 영상 및 변위도

(a) 좌영상 (b) 우영상 (c) 사실 변위도 (d) MI (e) LoG/BT  
 (f) Rank/BT (g) NCC (h) AWSM (i) 제안한 방법(HEC+8path DP)  
 ((a), (b), (c)는 Middlebury Stereo Vision Lab에서 공개)

**Fig. 6.10.** Image pair of Aloe and Disparity maps

(a) left image (b) right image (c) ground truth (d) MI (e) LoG/BT  
 (f) Rank/BT (g) NCC (h) AWSM (i) proposed method(HEC+8path DP)  
 (Image courtesy of (a),(b) and (c) Middlebury Stereo Vision Lab)

해 배경과 겹쳐지는 부분에서 오정합을 발생시켰다. 8개의 영상에 대한 비폐색 영역의 오정합 화소 비율을 구하여 제안한 방법의 성능이 모든 조합에 대해 우수한 영상들에 대한 오차율은 표 6.7과 같다.

**표 6.7** 제안한 방법이 우수한 성능을 보인 영상들의 카메라 노출 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

**Table 6.7** BPR(%) of unoccluded region about each camera exposure index combinations of outstanding images

		e_0/e_0	e_1/e_0	e_2/e_0	e_2/e_1	e_2/e_2
Aloe	NCC[58]	11.63	6.73	7.42	6.55	6.83
	LoG/BT[45]	27.31	38.14	41.92	14.37	11.55
	Rank/BT[46]	7.71	13.92	17.36	13.85	4.52
	AWSM[34]	4.17	4.03	6.01	3.94	4.21
	MI[37]	5.95	4.37	5.97	4.42	3.35
	HEC+GC	4.13	4.03	5.51	5.43	5.27
	<b>proposed method</b>	<b>4.12</b>	<b>4.02</b>	<b>3.72</b>	<b>3.56</b>	<b>3.14</b>
Art	NCC	20.43	17.67	21.25	15.13	15.37
	LoG/BT	38.42	61.27	72.38	33.18	12.05
	Rank/BT	15.24	14.21	16.87	15.24	11.27
	AWSM	9.73	9.00	13.82	11.69	10.93
	MI	17.05	18.07	20.83	18.21	17.74
	HEC+GC	9.13	9.27	12.17	12.27	10.62
	<b>proposed method</b>	<b>7.27</b>	<b>6.37</b>	<b>10.62</b>	<b>9.41</b>	<b>9.81</b>
Dolls	NCC	13.27	12.73	15.71	12.45	11.24
	LoG/BT	23.54	36.18	57.18	19.33	14.58
	Rank/BT	9.52	7.58	12.00	7.14	5.52
	AWSM	6.14	5.84	8.88	6.85	4.30
	MI	15.14	13.71	20.24	11.58	17.73
	HEC+GC	9.48	7.15	10.42	8.34	6.63
	<b>proposed method</b>	<b>4.24</b>	<b>4.21</b>	<b>6.24</b>	<b>5.47</b>	<b>4.02</b>
Moebius	NCC	14.62	12.72	15.92	14.17	13.24
	LoG/BT	23.17	24.37	32.51	16.72	16.96
	Rank/BT	12.76	13.00	15.27	10.24	10.13
	AWSM	7.29	10.31	10.72	10.25	8.15
	MI	10.81	11.96	17.27	14.47	11.74
	HEC+GC	7.91	8.51	10.43	10.67	9.24
	<b>proposed method</b>	<b>7.14</b>	<b>7.27</b>	<b>7.91</b>	<b>7.84</b>	<b>7.38</b>

전반적으로 모든 방법들의 오차율이  $e_2/e_0$  색인 조합, 가장 밝고 가장 어두운 영상을 정합하는 경우 가장 크다. 이 색인 조합에 대한 오차율과 노출 시간이 동일한 조합의 오차율의 편차를 살펴보면 제안한 유사도 척도인 HEC가 동일한 최적화 기법을 적용했을 때 다른 유사도 척도들에 비해 가장 적었다. 오차율의 편차는 Dolls 영상에서 2.7%로 가장 컸다. 같은 영상에서 MI는 9% 이상, AWSM은 3% 정도의 편차를 보였다. 제안한 방법은 공간 정보를 활용하여 유사도에 계산하기 때문에 밝기 변화에 대해 기존의 유사도 척도들보다 영향을 적게 받았다.

제안한 유사도 척도는 기존의 MI의 오차율을 감소시켰고, 그 차이는 Art 영상과 Dolls 영상에서 확실히 나타났다.  $e_2/e_0$  조합에서 MI와 제안한 유사도 척도의 오차율 차이는 두 영상에서 각각 8.16%와 9.82%였다. 이 영상들은 작은 물체들이 영상에 조밀하게 위치되어 있다. 이들 영상은 다른 영상에 비해 변위 불연속 영역이 많기 때문에 공간 정보를 포함한 제안한 유사도 척도의 성능이 더 좋았다.

다른 방법들과의 성능차이는 Art 영상과 Moebius 영상에서 두드러진다. Art 영상은 얇은 물체가 많고, Moebius 영상은 변위가 같은 물체 내에서 경계선이 많이 검출되는 특징이 있다. 실제로 Dolls 영상에서 GC를 적용한 기법들의 오정합율은 AWSM과 제안한 방법에 비해 높게 나타남을 알 수 있다. 제안한 방법의 결과를 종합해보면 GC보다 제안한 최적화 기법의 성능이 우수함을 알 수 있다.

각 영상의 색인조합  $e_0/e_0$ 와  $e_2/e_0$ 에 대한 제안한 방법의 오차율 편차는 Aloe에서는 0.4%, Art는 2.35%, Dolls에서는 2%, 그리고 Moebius에서는 0.77%이다. 제안한 방법의 오차율 편차가 가장 작기 때문에 다른 방법들에 비해 전역적인 밝기 변화에 대해 강인하다고 할 수 있다.

기존 방법들과 제안한 방법의 성능이 대등하게 나타난 영상에 대한 실험 결과를 표 6.8에 보였다.

**표 6.8** 제안한 방법의 성능이 대등하게 나타난 영상들의 카메라 노출 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

**Table 6.8** BPR(%) of unoccluded region about each camera exposure index combinations of even images

		e_0/e_0	e_1/e_0	e_2/e_0	e_2/e_1	e_2/e_2
Cloth4	NCC[58]	2.42	1.69	1.75	1.64	1.72
	LoG/BT[45]	7.31	8.92	16.82	3.16	1.87
	Rank/BT[46]	1.6	1.16	2.37	1.17	1.10
	AWSM[34]	1.62	1.17	2.36	1.79	1.82
	MI[37]	2.61	2.57	2.83	2.64	2.70
	HEC+GC	1.62	1.42	2.02	1.34	2.26
	<b>proposed method</b>	<b>1.75</b>	<b>1.06</b>	<b>1.47</b>	<b>1.32</b>	<b>1.97</b>
Laundry	NCC	27.15	23.37	25.15	25.91	25.47
	LoG/BT	23.51	25.97	30.28	20.42	27.00
	Rank/BT	20.14	21.76	22.41	21.24	20.46
	AWSM	13.92	11.97	19.62	16.84	12.17
	MI	21.51	20.71	22.15	22.71	23.64
	HEC+GC	14.71	11.24	20.52	15.10	14.18
	<b>proposed method</b>	<b>14.41</b>	<b>12.15</b>	<b>18.32</b>	<b>14.24</b>	<b>11.15</b>
Rock1	NCC	17.12	12.71	18.72	13.91	13.57
	LoG/BT	14.37	16.75	27.51	24.72	26.41
	Rank/BT	17.34	13.12	18.16	17.24	14.27
	AWSM	14.52	15.48	16.82	14.15	13.64
	MI	17.16	14.82	19.77	15.72	14.93
	HEC+GC	17.24	16.72	17.21	14.10	14.20
	<b>proposed method</b>	<b>15.52</b>	<b>15.21</b>	<b>16.49</b>	<b>14.21</b>	<b>13.10</b>

제안한 유사도 척도는 e\_0/e\_0과 e\_2/e\_2 색인에서 기존 방법인 AWSM과 오차율이 비슷하거나 다소 높았다. 이는 대비의 축소로 인해 Haar 웨이블릿 특징 벡터의 엔트로피 보정 성능이 저하됐기 때문이라 판단된다. 한편 각 영상의 색인 조합에 대한 오차율 변화 추이를 보면 제안한 방법의 편차가 가장 적음을 알 수 있다. 또한 모든 영상에 대한 e\_2/e\_0 색인에서 제안한 방법의 오차가 기존의 방법들보다 낮았다. 이 색인은 밝기 차이가 가장 극단적으로 발생하는 경우인데 제안한 방법의 오차율이 Cloth4에서 1.47%와 Laundry 영상에서

18.32%, Rock1 영상에서 16.49%로 가장 낮았다. 몇몇 색인에 대해 제안한 방법보다 AWSM의 오차율이 낮았지만 오차율의 차이는 0.5% 미만이었다.

제안한 방법의 성능이 좋지 않은 영상에 대한 오차율은 표 6.9와 같다.

**표 6.9** 제안한 방법의 성능이 미진한 영상의 카메라 노출 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

**Table 6.9** BPR(%) of unoccluded region about each camera exposure index combinations of inferior image

		e_0/e_0	e_1/e_0	e_2/e_0	e_2/e_1	e_2/e_2
Reindeer	NCC[58]	15.24	14.16	12.27	11.15	9.27
	LoG/BT[45]	17.24	22.37	33.92	12.57	7.61
	Rank/BT[46]	12.71	12.48	14.37	10.42	9.41
	AWSM[34]	11.72	8.14	6.82	4.16	3.62
	MI[37]	15.27	13.72	20.15	20.07	10.57
	HEC+GC	14.01	10.14	12.72	10.24	13.95
	<b>proposed method</b>	<b>13.12</b>	<b>9.34</b>	<b>11.73</b>	<b>8.41</b>	<b>9.37</b>

Reindeer 영상에서 제안한 방법의 오정합율이 큰 차이를 보였다. 이 영상은 다른 영상에 비해 최대 변위가 커서 변위가 큰 물체에 대해 약한 기하학적 왜곡이 발생했고, 변위가 큰 물체가 두 영상에서 컬러 차이가 큰 배경에 각각 위치해있다. 특히 가장 앞에 있는 물체의 변위가 94화소로 실험에 사용한 영상들 중 가장 크다. 표에서 계산한 오정합 화소 비율은 구해진 변위와 사실 변위도 와서 변위 차이가 1화소 이상일 경우 오정합 화소로 구분한다. 이 경우 해당 영상과 같은 큰 변위를 갖는 영역에 대해 약간의 소수점 차이가 발생해도 오정합 화소로 분류되는 결과를 초래한다. 실제로 이러한 특성을 알아보기 위해 오정합 화소의 임계치를 2로 설정하고 오차율을 계산했을 때, 모든 방법들의 오차율이 개선되는 효과를 보였고, 제안한 방법의 오차율은 큰 폭으로 감소하였다. 특히 e\_2/e\_0 색인에 대한 AWSM의 오차는 5.71%로 나타났고, 제안한 방법

의 오차는 6.43%였다. 그리고  $e_1/e_0$  색인에 대해서는 AWSM이 6.22%, 제안한 방법이 6.08%로 오히려 더 좋은 결과를 보였다. 그리고 최대 오차율의 편차는 제안한 방법이 3.72%로 8.1%를 나타낸 AWSM보다 낮았다. 제안한 방법의 오정합 화소 비율은 AWSM보다 높게 나타났으나 환경 변화에 대한 성능 차이는 제안한 방법이 더 적었다.

8개 영상에 대한 실험을 통해, 제안한 유사도 척도와 최적화 기법은 두 영상의 밝기 차이가 극단적일 때 다른 방법들에 비해 오차율의 변화가 적었다. 그리고 실험 영상 조합에 따른 오차율의 편차가 다른 방법들에 비해 적었다. 반면 제안한 방법의 오차율이 아주 밝거나 어두운 영상에서 오차율이 소폭 증가하는 것을 볼 수 있었는데, 이는 극단적으로 밝거나 어두운 영상에서 Census 변환이 Rank 변환과 비슷한 단점을 가지고 있기 때문이다. 이런 점을 제외하면 제안한 방법의 성능이 전반적으로 우수하게 나타났다.

### 3) 조명의 차이가 발생한 영상 실험 결과

실험에 사용한 영상과 실험으로 얻은 변위도들을 그림 6.11에 나타내었다. 좌영상과 우영상의 조명이 다른데, 우영상의 조명 방향이 영상의 오른쪽에서 왼쪽으로 변경되어 좌영상에 없는 그림자가 관찰된다. 실험에 사용한 조명 색인은  $I_1$ ,  $I_2$ ,  $I_3$ 로 구성하였고 좌우영상의 조명 조합은  $I_1/I_2$ 와 같은 형식으로 표기하였다. 실험에 사용한 영상은 카메라 노출 시간 변화 실험에 사용한 영상과 동일하다.

기존의 MI를 이용해 얻은 변위도는 그림 6.11의 (d)와 같다. 밝기 변화에 강인하다고 알려진 MI의 성능이 좋지 않음을 확인할 수 있다. Rank/BT와 NCC를 이용하여 얻은 변위도는 (e)와 (f)이다. 두 방법 모두 조명이 변한 환경에서 좋은 성능을 보였지만 화분 상단과 오른쪽 알로에 잎 부근에서 오정합이 많이 발생한다. 그림 6.11의 (a)에는 없는 그림자가 (b)의 중앙에 보인다. 제안한 방법으로 얻은 변위도가 그림 6.11의 (i)이다. 물체의 경계선이 다른 방법들에 비해 잘 보전되었음을 알 수 있다. 해당 영상 화분 윗부분에서는 앞에서 수행한 카

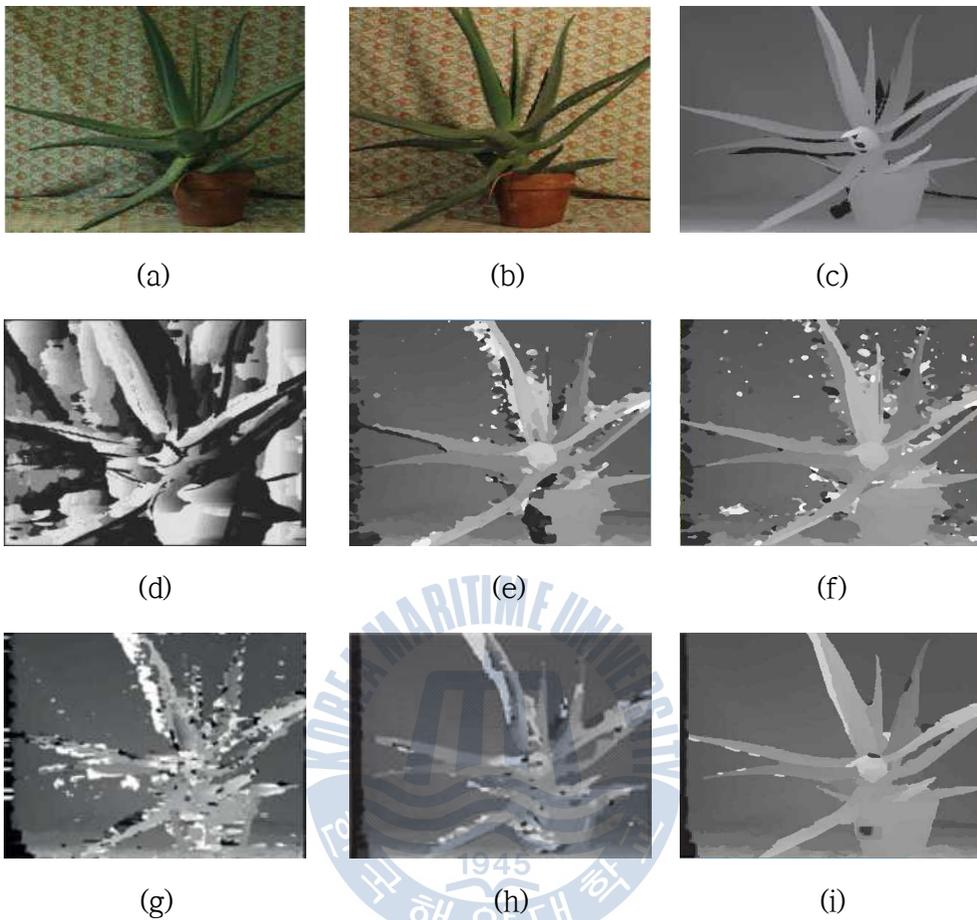


그림 6.11. Aloe 영상 및 변위도

(a) 좌영상 (b) 우영상 (c) 사실 변위도 (d) MI (e) LoG/BT (f) Rank/BT  
 (g) NCC (h) AWSM (i) 제안한 방법(HEC+8path DP)  
 ((a),(b)와 (c)는 Middlebury Stereo Vision Lab에서 공개)

Fig. 6.11. Image pair of Aloe and Disparity maps

(a) left image (b) right image (c) ground truth (d) MI (e) LoG/BT  
 (f) Rank/BT (g) NCC (h) AWSM (i) proposed method(HEC+8path DP)  
 (Image courtesy of (a), (b) and (c) Middlebury Stereo Vision Lab)

메라 노출 시간의 차이로 발생한 특징이 유사하게 나타난다. 기존의 방법들은 이 영역에서 오정합이 많이 발생하였다. 그러나 전반적으로 제안한 방법의 변위가 다른 방법에 비해 정확히 구해졌다.

Dolls 영상에 대해 실험한 결과를 그림 6.12에 보였다. 이 영상은 모든 물체들의 경계가 곡선 형태이고 군데군데 폐색영역이 분포된 특징이 있다. 그리고 이 영상은 물체들의 변위 차이가 크지 않고 물체의 질감이 다양하기 때문에 경계선이 많이 검출되는 물체에 대해 변위를 일정하게 추정해야 한다.

제안한 유사도 척도와 GC를 적용하여 얻은 변위도는 그림 6.12의 (h)이다. 기존의 MI를 비롯한 다른 유사도 척도들보다 오정합이 적음을 알 수 있다. 제안한 유사도 척도는 조명의 변화에도 영상 전반적으로 유사도를 잘 반영했다고 볼 수 있다. 기존 방법들의 오정합이 발생한 영역은 주로 조명의 변화로 인해 물체의 명암도 차이가 발생한 곳이다. 특히 왼쪽 영상에서 가장 앞에 있는 아기 인형의 머리나 오른쪽 뒤에 있는 인형의 머리에서 명암 차이가 발생한 것을 볼 수 있다. GC는 이와 같은 영역을 변위 불연속 영역으로 간주하기 쉽기 때문에 그림 6.12의 (i)를 제외한 나머지 변위도에서 오정합이 많이 발생하였다.

제안한 최적화 기법을 적용하여 얻은 변위도는 그림 6.12의 (i)이다. 이 영상은 같은 물체의 경계선이 많이 분포되어 있기 때문에 GC에 비해 제안한 방법의 경계선 보전 성능이 부각되었다. 실제로 제안한 방법은 변위가 변하는 영역에서의 오정합을 많이 감소시켰다. 단 영상 중앙에 보이는 아기 머리 부분에서 오정합이 발생하였다. 아기 머리 부분의 컬러 변화 형태는 좌영상에서 하이라이트 부분까지 밝기가 점진적으로 증가하지만 우영상에서는 평탄하다. 이로 인해 각각의 영역에서 Haar 웨이블릿 특징 벡터가 다르게 계산되어 오정합이 발생했다고 판단된다.

8개의 영상과 조명 조합에 대한 제안한 방법과 다른 방법들의 오정합율을 표로 나타내었다. 모든 조명 조합에 대해 제안한 방법의 성능이 우수한 영상들의 오차율은 표 6.10과 같다.

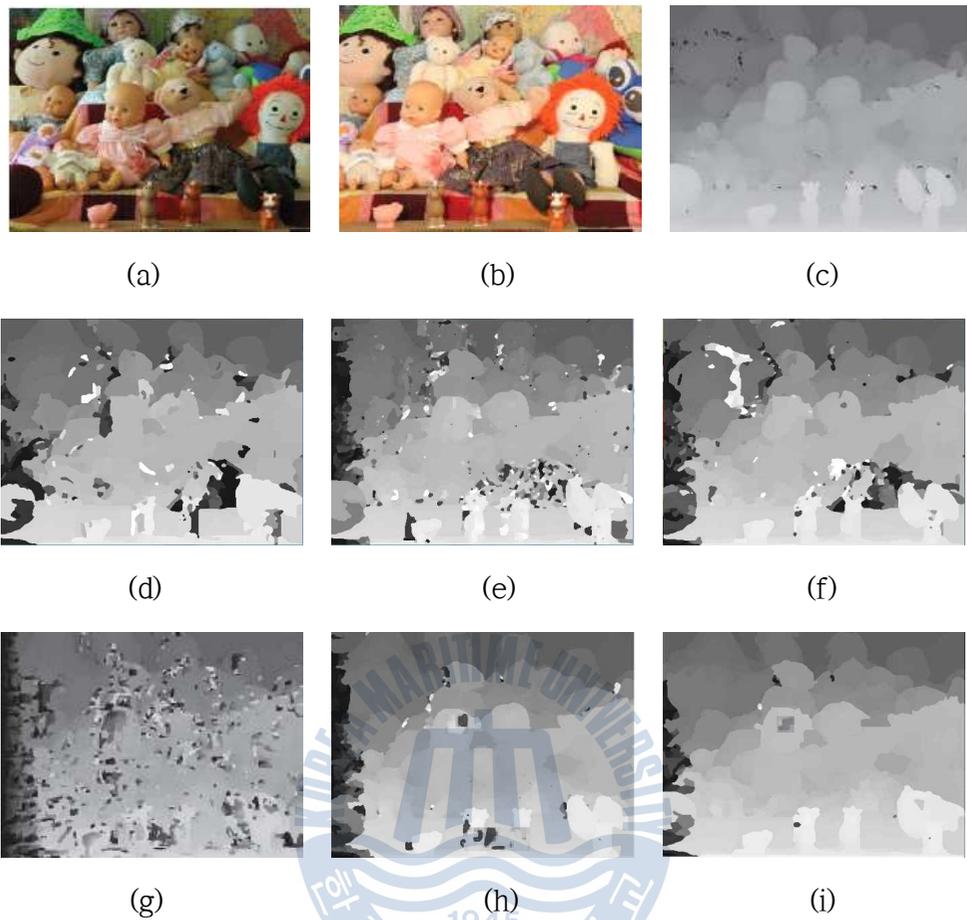


그림 6.12. Dolls 영상 및 변위도

(a) 좌영상 (b) 우영상 (c) 사실 변위도 (d) MI (e) LoG/BT (f) Rank/BT  
(g) NCC (h) AWSM (i) HEC+8path DP

((a),(b)와 (c)는 Middlebury Stereo Vision Lab에서 공개)

**Fig. 6.12.** Image pair of Dolls and Disparity maps

(a) left image (b) right image (c) ground truth (d) MI (e) LoG/BT  
(f) Rank/BT (g) NCC (h) AWSM (i) HEC+8path DP

(Image courtesy of (a), (b) and (c) Middlebury Stereo Vision Lab)

표 6.10 제안한 방법의 성능이 우수한 영상들의 조명 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

Table 6.10 BPR(%) of unoccluded region about each illumination index combinations of outstanding images

		I_1/I_1	I_2/I_1	I_3/I_1	I_3/I_2	I_3/I_3
Aloe	NCC[58]	9.33	11.41	22.73	17.83	9.82
	LoG/BT[45]	14.27	15.92	26.41	19.73	13.82
	Rank/BT[46]	6.84	8.12	19.24	16.48	6.62
	AWSM[34]	3.46	7.41	13.24	10.24	4.34
	MI[37]	5.31	18.14	72.54	63.75	6.21
	HEC+GC	4.37	6.62	11.25	8.46	5.69
	<b>proposed method</b>	<b>3.36</b>	<b>5.92</b>	<b>9.91</b>	<b>9.11</b>	<b>3.79</b>
Dolls	NCC	10.12	11.42	35.27	28.51	9.40
	LoG/BT	17.71	17.31	26.15	19.91	15.54
	Rank/BT	6.14	9.24	33.15	24.15	6.41
	AWSM	4.31	8.85	20.21	22.14	5.21
	MI	17.13	32.48	63.29	42.75	14.53
	HEC+GC	5.37	7.41	12.48	13.54	5.62
	<b>proposed method</b>	<b>4.12</b>	<b>7.24</b>	<b>12.42</b>	<b>13.45</b>	<b>4.28</b>
Cloth4	NCC	3.74	6.71	10.92	3.89	3.69
	LoG/BT	3.37	4.12	7.27	3.97	3.68
	Rank/BT	3.05	4.88	7.39	3.84	3.75
	AWSM	2.82	6.94	7.87	3.21	2.09
	MI	3.48	37.41	43.17	8.41	2.71
	HEC+GC	2.63	3.35	4.34	4.08	2.31
	<b>proposed method</b>	<b>2.33</b>	<b>3.12</b>	<b>3.67</b>	<b>2.32</b>	<b>1.83</b>
Rock1	NCC	3.74	6.71	10.93	3.89	4.48
	LoG/BT	3.37	4.84	7.95	4.34	3.86
	Rank/BT	2.90	4.64	8.01	4.14	3.55
	AWSM	1.82	6.97	11.87	3.21	2.09
	MI	3.14	11.84	39.12	15.27	3.18
	HEC+GC	2.15	3.24	7.68	4.27	2.38
	<b>proposed method</b>	<b>1.63</b>	<b>2.35</b>	<b>2.47</b>	<b>2.32</b>	<b>1.33</b>
Laundry	NCC	22.14	34.21	42.91	30.45	24.51
	LoG/BT	11.24	21.45	32.51	19.42	18.24
	Rank/BT	20.24	26.91	39.72	27.15	19.56
	AWSM	11.42	22.14	28.35	19.24	11.56
	MI	13.57	37.85	61.47	41.75	15.61
	HEC+GC	11.38	18.75	27.75	15.26	12.62
	<b>proposed method</b>	<b>10.72</b>	<b>15.32</b>	<b>24.93</b>	<b>14.21</b>	<b>10.63</b>

조명 색인이 다르면 MI의 오정합율이 증가했다. 기존의 MI는 두 영상의 조명 변화로 인해 불균일하게 밝기가 변했을 때 유사도를 잘 반영하지 못하기 때문이다. 이러한 특성은 Aloe 영상에 대한 실험에서 조명 색인이  $I_1/I_1$ 로 같을 때 5.31%였던 오차율이 조명 색인이  $I_3/I_1$ 로 달라졌을 때 72.54%로 급격히 증가했다는 점에서 확실히 드러난다. 동일한 조건에서 제안한 유사도 척도인 HEC의 오정합율은 4.37%에서 11.25%로 증가폭이 감소하였다. 다른 영상에 대해서도 HEC의 오정합율은 조명 색인이 같을 경우 MI와 근소한 차이를 보였지만 다른 경우 오정합율의 증가폭이 확연히 감소하였다. 또한 HEC는 모든 영상의 조명 색인  $I_3/I_1$ 에서 실험에 사용한 유사도 척도들 중 가장 우수한 성능을 보였다. 이러한 결과를 통해 지역적으로 불균일한 밝기 변화가 발생한 경우 Haar 웨이블릿 특징 벡터로 얻은 주파수 특성의 공간 정보는 기존의 엔트로피 기반 유사도 척도의 오정합을 보완한다는 사실을 알 수 있다.

Aloe 영상의 색인이  $I_3/I_1$ 일 때 제안한 유사도 척도를 GC와 8path DP에 적용하여 계산한 오차율은 각각 11.25%과 9.91%였다. 그 외 다른 영상들의 동일한 색인에 대한 실험에서 제안한 최적화 방법의 성능이 더 좋았고, 이 중 Rock1 영상의 동일한 색인에서 제안한 최적화 방법과 GC의 오차율 차이가 5.21%로 가장 컸다. 이는 대부분 질감이 풍부한 암석들로 구성되어 있는 Rock1 영상은 동일한 물체 내에서 경계선이 많이 검출되기 때문에 기존의 GC는 에지 화소를 변위 불연속 영역으로 분류하여 오정합이 증가하기 때문이다.

색인별 오차율의 변화는 모든 영상에서  $I_3/I_1$  색인에 대한 오차가 가장 크고  $I_1/I_1$  색인의 오차가 가장 작았다. Dolls 영상을 제외한 4개의 영상에서 이 색인들에 대한 오차율의 차이는 제안한 방법에서 가장 작았기 때문에 조명 변화에 가장 강인한 방법이라고 할 수 있다. Dolls 영상에서는 제안한 방법의 편차가 1.88% 높지만  $I_3/I_1$ 에서 제안한 방법의 오차가 가장 작았기 때문에 제안한 방법이 환경 변화에 민감하다고 할 수는 없다.

기존의 방법과 제안한 방법의 성능이 대등하게 나타난 영상에 대한 오차율은 표 6.11과 같다.

표 6.11 제안한 방법의 성능이 대등한 영상들의 조명 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

Table 6.11 BPR(%) of unoccluded region about each illumination index combinations of even images

		I_1/I_1	I_2/I_1	I_3/I_1	I_3/I_2	I_3/I_3
Art	NCC[58]	19.96	21.93	42.17	38.15	19.11
	LoG/BT[45]	28.12	28.38	41.32	39.67	26.86
	Rank/BT[46]	16.67	17.85	38.46	31.92	13.00
	AWSM[34]	11.26	13.28	23.99	20.24	11.12
	MI[37]	17.14	32.42	63.29	42.72	14.53
	HEC+GC	12.42	13.23	21.25	18.34	12.24
	<b>proposed method</b>	<b>11.33</b>	<b>11.75</b>	<b>20.92</b>	<b>16.41</b>	<b>11.36</b>
Moebius	NCC	10.52	14.21	22.18	19.63	11.48
	LoG/BT	15.52	16.25	21.38	19.85	12.42
	Rank/BT	9.21	10.35	20.92	18.21	9.75
	AWSM	6.27	8.75	14.62	12.51	5.15
	MI	15.76	24.72	41.48	31.57	13.76
	HEC+GC	8.05	9.42	16.72	13.72	9.05
	<b>proposed method</b>	<b>6.71</b>	<b>8.54</b>	<b>15.23</b>	<b>11.42</b>	<b>5.17</b>

조명 색인이 같을 때 Art 영상에 대한 실험에서 AWSM가 제안한 방법보다 오정합 화소 비율이 근소하게 작았다. 조명 색인 I\_1/I\_1과 I\_3/I\_3에 대한 차이가 각각 0.07% 0.24%이므로 월등히 성능이 좋다고 볼 수는 없다. 오히려 조명이 다른 세가지 조합에서 제안한 방법의 성능이 더 좋았다.

Moebius 영상의 조명 색인 I\_3/I\_1에 대한 제안한 방법의 오정합 화소 비율은 AWSM보다 약간 높았다. 조명 색인 I\_2/I\_1과 I\_3/I\_2에 대한 성능은 제안한 방법이 더 우수했지만, 오정합이 가장 많이 발생하는 색인 I\_3/I\_1에 대해서는 AWSM이 더 우수하였다. Moebius 영상은 조명 변화로 인해 발생한 밝기 변화의 불균일한 정도가 크지 않기 때문에 공간 정보를 이용하는 제안한 방법의 성능이 AWSM의 적응형 가중치보다 부각되지 못했다고 판단된다.

제안한 방법의 성능이 미진한 영상에 대한 오차율은 표 6.12와 같다.

**표 6.12** 제안한 방법의 성능이 미진한 영상에 대한 조명 색인 조합별 비폐색 영역의 오정합 화소 비율(%)

**Table 6.12** BPR(%) of unoccluded region about each illumination index combinations of inferior image

		I_1/I_1	I_2/I_1	I_3/I_1	I_3/I_2	I_3/I_3
Reindeer	NCC[58]	24.31	30.14	37.28	22.27	10.48
	LoG/BT[45]	11.28	21.44	32.53	19.47	18.20
	Rank/BT[46]	20.14	25.38	37.59	19.61	17.34
	AWSM[34]	10.42	22.16	26.8	17.07	11.52
	MI[37]	11.47	34.81	62.57	77.24	13.72
	HEC+GC	11.42	24.28	25.18	19.26	12.68
	<b>proposed method</b>	<b>11.42</b>	<b>24.24</b>	<b>24.27</b>	<b>19.28</b>	<b>12.05</b>

Reindeer 영상에서는 실험에 사용한 모든 방법의 오정합율이 다른 영상에 비해 컸다. 이는 카메라 노출 관련 실험과 마찬가지로 이 영상은 두 영상의 시차가 커서 변위가 큰 물체에 대한 기하학적 왜곡이 발생했기 때문이다. 표 (6.12)의 오정합 화소 비율들은 사실 변위도와의 변위 차이가 1화소 이상일 때 오정합 화소로 구분하기 때문에 해당 영상과 같은 큰 변위를 갖는 영역에 대해 약간의 소수점 차이가 발생해도 오정합 화소로 분류된다. 그리고 MI를 제외한 다른 방법들과 제안한 방법의 오차율의 차이가 크지 않다. 따라서 기하학적 변화를 동반한 영상의 특성으로 인해 조명의 변화가 각 방법의 성능에 큰 영향을 주지 않았다고 판단된다.

#### 6.2.4 실제 영상 실험

##### 1) 실험 영상과 실험 방법

제안한 방법의 성능을 검증하기 위해 실험용 표준 영상이 아닌 실제 촬영한 영상에 대해 적용하였다. 스테레오 영상은 BASLER사의 SCA1000-30FC 2개를

이용해 촬영하였고 이득과 감마계수는 각각 507과 1.0으로 설정하였다. 렌즈의 초점은 0.4m로 동일하게 맞췄고 조리개를 조절하여 환경 변화 조건을 조성했다. 실험에 사용한 영상은 CCD, BOOK, DIP, PIG이며 그림 6.14에 나타내었다. 모든 실험은 환경 조건이 동일한 상태에서 촬영한 영상과 한쪽의 조리개 수치에 변화를 주고 촬영한 영상에 대해 각각 수행하였다. 이 중 CCD와 BOOK 영상 실험은 조리개의 변화를 준 경우에 대해 기존 방법의 변위도와 비교하여 제안한 방법의 성능을 정성적으로 평가하였다. 조리개 조절은 전역적인 밝기 변화를 초래하므로 앞 절에서 적용한 LoG/BT, Rank/BT, NCC 그리고 AWSM을 적용하여 비교하였다. 각 영상 획득시 카메라 간격은 PIG 영상은 7cm, CCD 영상은 4cm, BOOK 영상과 DIP 영상은 3cm였다.

CCD 영상은 그림 6.14의 (a)이다. 왼쪽부터 동일 조건에서 촬영한 좌영상, 우영상, 그리고 조리개를 닫은 우영상이다. 이 영상은 뒷배경의 깊이가 작고 중앙의 뒷배경이 앞으로 돌출되어 있어서 약간의 변위차가 발생한다. 배경 바로 앞에 있는 물체가 뒤로 누어 있기 때문에 물체내에서 점진적인 변위차가 발생한다. 동일 조건에서 촬영한 영상에 대해 제안한 방법의 변위도를 보였고, 조리개를 닫은 우영상과의 정합 성능은 기존 방법들과의 변위도를 비교하였다. 실험에 사용한 BOOK 영상은 그림 6.14의 (b)이다. 이 영상은 좌영상에 대해 조리개를 닫고 촬영하여 변위도를 얻었다. 이 영상은 가장 앞에 있는 의자를 제외하고 뒷배경으로 존재하는 책장과 책들의 작은 변위차이를 잘 반영하는지 확인하기 위한 목적으로 사용하였다. 그리고 책장의 공백영역은 무질감의 형태를 띠기 때문에 무질감을 표현하는 유사도 척도의 성능을 확인할 수 있는 영상이다. DIP영상은 실험 영상 중 기준선이 가장 짧고 배경의 질감이 다양하기 때문에 표준 영상과 구성이 비슷하다. PIG 영상은 변위가 가장 크기 때문에 폐색영역이 크고 기하학적인 왜곡이 발생한 경우에 대한 성능을 확인하기 위해 사용하였다.



(a)



(b)



(c)



(d)

그림 6.14. 실제 영상

(a) CCD (b) BOOK (c) DIP (d) PIG

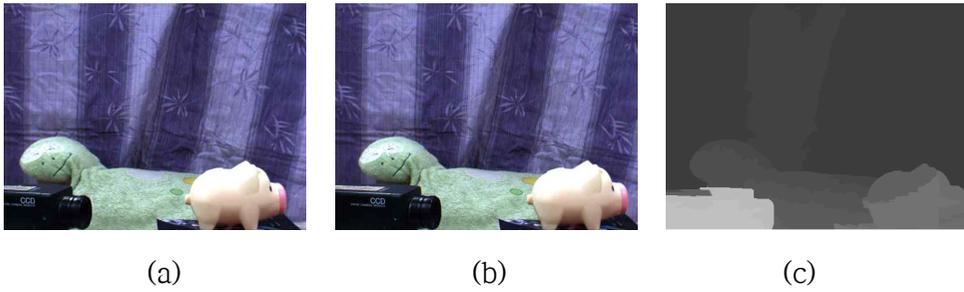
Fig. 6.14. Real image sets

(a) CCD (b) BOOK (c) DIP (d) PIG

## 2) 실제 영상 실험에 대한 결과

실험에 사용한 CCD 영상과 제안한 방법의 변위도는 그림 6.15와 같다. 두 카메라의 조리개 개방이 동일한 경우 제안한 방법은 깊이 방향으로 눕혀진 공룡 인형의 점진적으로 변하는 변위를 잘 추정하였고 왼쪽의 저금통의 모양을 뚜렷하게 표현했다. 그리고 뒷배경에서 변위가 다른 부분도 잘 찾았다. 우영상의 조리개를 조절하여 어둡게 하고 수행한 실험의 결과는 그림 6.16과 같다. 앞 절에서 수행한 카메라 노출 시간에 대한 실험과 유사한 결과들을 보였다. LoG/BT의 변위도는 오정합 영역이 길게 펼쳐져 있고, Rank/BT는 경계선 영역과 비폐색 영역에서 잡음 형태의 오정합을 발생시킨 것을 그림 6.16의 (b)와 (c)를 통해 알 수 있다. NCC는 CCD 카메라와 저금통 등 질감이 없는 평탄한 영역에서 오정합을 발생시켰고, AWSM은 공간 정보를 이용하여 NCC의 문제점을 크게 개선하였으나 인형과 저금통이 겹쳐지는 부분에서 오정합이 발생하였다. 제안한 방법의 변위도는 그림 6.16의 (f)와 같다. 제안한 방법은 AWSM과 같이 공간 정보를 포함하는 유사도 척도를 사용하기 때문에 제안한 방법의 변위도와 AWSM의 변위도와 큰 차이가 없어 보이지만 AWSM에서 오정합이 발생한 영역에서 경계선을 잘 보전하였다. 그리고 그림 6.15의 (c)와 그림 6.16의 (f)를 비교했을 때 제안한 방법은 환경 변화에 따른 변위도의 시각적인 차이가 거의 없음을 확인하였다.

실험에 사용한 BOOK 영상과 제안한 방법의 변위도는 그림 6.17과 같다. 환경의 변화가 없을 때 제안한 방법의 변위도는 뒷배경의 작은 변위차를 잘 표현하였고 책장과 책들의 경계선을 잘 표현하였다. 꽂혀있는 책들의 깊이가 다른 특성도 잘 반영하였다. BOOK의 좌영상을 어둡게하여 앞 절과 동일한 방법을 실험하였다. 실험 영상과 각 방법으로 얻은 변위도는 그림 6.18로 나타냈다. 변위가 같은 작은 단위의 물체들의 경계선 부근에서 LoG/BT와 Rank/BT의 오정합이 두드러진다. 그리고 앞에 있는 담요의 체크 무늬에서도 경계선이 검출되기 때문에 이 영역에서도 오정합이 발생했다. NCC는 경계선이 뭉개지는 문제를 보였다. 공간 정보를 적용한 AWSM과 제안한 방법의 변위도가 다른 기법들에



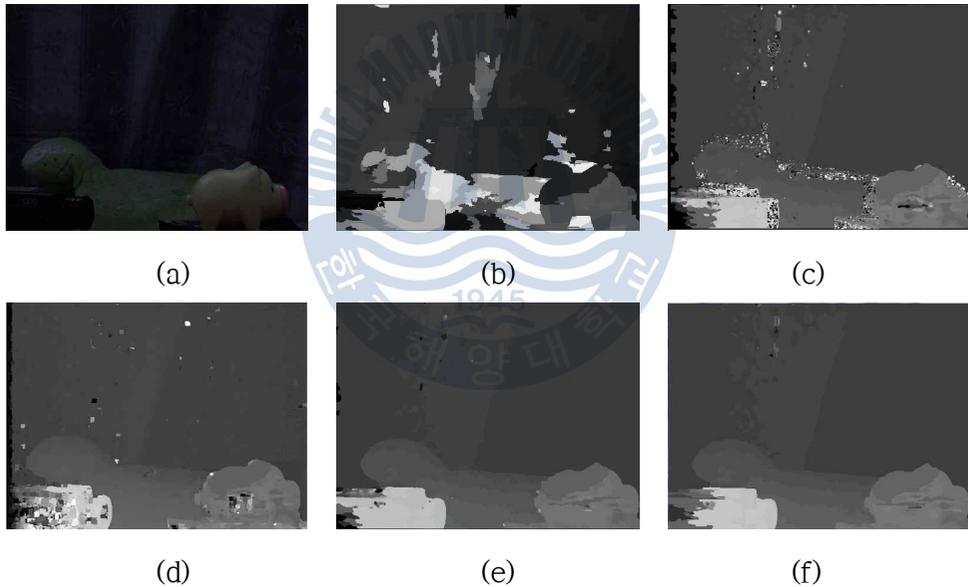
(a) (b) (c)

그림 6.15. CCD 영상과 제안한 방법의 변위도

(a) 좌영상 (b) 우영상 (c) 제안한 방법의 변위도

Fig. 6.14. Real stereo images and disparity maps

(a) left image (b) right image (c) disparity map of proposed method



(a) (b) (c)  
(d) (e) (f)

그림 6.16. CCD 영상의 우영상과 변위도

(a) 우영상 (b) LoG/BT (c) Rank/BT

(d) NCC (e) AWSM (f) 제안한 방법

Fig. 6.16. Right image of real stereo images and disparity maps

(a) right image (b) LoG/BT (c) Rank/BT

(d) NCC (e) AWSM (f) proposed method

비해 우수한 성능을 보인다. 제안한 방법은 책꽂이의 빈 부분에서 오정합이 발생하였다. 반면 AWSM은 의자의 경계영역과 왼쪽의 빈 책장 영역에서 오정합이 크게 발생했다. DIP 영상과 PIG 영상을 이용한 실험은 환경 변화에 대한 제안한 방법의 강인함을 보이기 위해 수행하였다. DIP 영상과 제안한 방법의 변위도는 그림 6.19와 같다. DIP 영상은 환경이 같을 경우와 우영상의 밝기가 변한 경우의 변위도가 큰 차이를 보이지 않았고 제안한 방법의 장점인 경계선 보전 성능이 잘 드러났다. PIG 영상과 변위도는 그림 6.20과 같다. PIG 영상의 중앙에 있는 영역에서 정합이 정확히 이뤄지지 않았다. 이 영상에서 변위가 가장 큰 물체는 100에 가깝다. 그리고 다른 영상과는 달리 배경의 변위도 크다. 이러한 특성으로 인해 좌영상과 우영상에서 약한 기하학적 왜곡이 관찰된다. 좌영상에 없는 책의 옆면이 우영상에서 관찰되고, 책과 저금통 사이 영역의 배경 질감이 두 영상에서 다르다. 이로 인해 Haar 웨이블릿 특징 벡터와 Census 변환의 오차가 커지기 때문에 최적화 기법에서는 이 영역을 일종의 폐색 영역으로 둘 가능성이 크다.

실제 촬영 영상의 실험을 통해 표준 영상 실험에서 증명된 제안한 방법의 전역적인 밝기 변화에 강인한 성능과 경계선 영역에서의 우수한 정합 성능을 확인하였다.



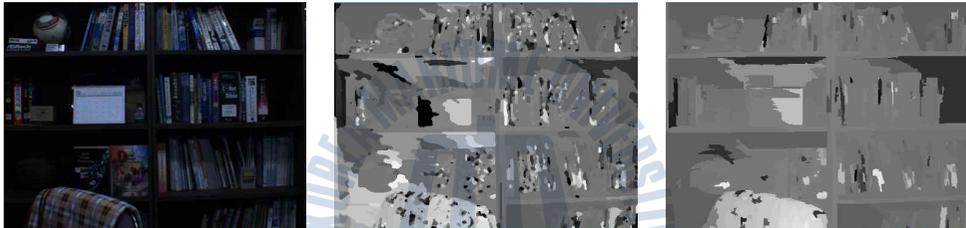
(a) (b) (c)

그림 6.17. BOOK 영상과 제안한 방법의 변위도

(a) 좌영상 (b) 우영상 (c) 제안한 방법의 변위도

Fig. 6.17. BOOK image pair and disparity maps

(a) left image (b) right image (c) disparity map of proposed method



(a)

(b)

(c)



(d)

(e)

(f)

그림 6.18. BOOK 영상의 조리개 닫힌 좌영상과 변위도들

(a) 조리개 닫힌 좌영상 (b) LoG/BT (c) Rank/BT

(d) NCC (e) AWSM (f) 제안한 방법

Fig. 6.18. A left image with closed aperture of BOOK pair and disparity maps

(a) left image with closed aperture (b) LoG/BT (c) Rank/BT

(d) NCC (e) AWSM (f) proposed method



(a) (b) (c)

그림 6.19. DIP 영상과 제안한 방법의 변위도

(a) 좌영상 (b) 우영상 (c) 제안한 방법의 변위도

Fig. 6.19. DIP stereo images and disparity maps

(a) left images (b) right images (c) disparity map of proposed method



(a) (b) (c)

그림 6.20. PIG 영상과 제안한 방법의 변위도

(a) 좌영상 (b) 우영상 (c) 제안한 방법의 변위도

Fig. 6.20. PIG stereo images and disparity maps

(a) left images (b) right images (c) disparity map of proposed method

### 6.3 계산 속도

제안한 방법의 계산상의 효율성을 보이려고 계산 시간을 측정하여 제시하였다. 동일한 영상을 처리하여 변위도를 얻는데 걸리는 시간을 기존의 방법들과 비교하였고 제안한 방법의 유사도 척도를 얻는데 걸리는 시간과 정합을 1회 수행하는 시간을 측정하여 제시하였다. 그리고 제안한 기법에 다해상도 기법을 적용했을 때 소요되는 시간을 적용하지 않았을 때 걸리는 시간과 비교하였다. 실험에 사용된 영상은 Aloe 영상이고 계산시간을 산출한 PC 환경은 Intel i5 2.1Ghz의 CPU에 4GB 메모리였다.

제안한 방법의 계산시간을 다른 기법들과 비교하기 위해 각각의 계산시간을 표 6.13에 나타내었다.

**표 6.13** Aloe 영상에 대한 다른 방법들과 제안한 방법의 계산 시간 비교

**Table 6.13** Computational time comparison with various methods about Aloe image pair

matching method entry	computational time(sec)
SAD(7×7)+GC	0.93
ZSAD(7×7)+GC	2.01
NCC(9×9)+GC	1.22
ZNCC(9×9)+GC	2.37
AWSM(31×31)+GC	8.05
LoG/BT(9×9)+GC	0.73
Rank/BT(5×5)+GC	0.76
HEC+GC	1.73
<b>proposed method</b>	<b>1.26</b>

최적화 방법으로 GC를 이용한 방법들 중 제안한 방법의 순위가 5번째였다.

이는 가장 빠른 BT기반의 기법에 비해 2배 이상 차이가 난다. 제안한 방법은 엔트로피와 Haar 웨이블릿 특징 벡터, Census 변환을 결합한 유사도 척도를 사용한다. 그리고 그 중 계산 복잡도가 특히 높은 Census 변환을 이용하므로 그 래프 컷과 같은 전역적 최적화 방법에 적용하면 계산 복잡도가 더 증가한다.

다른 상관도 기반 기법들 중에서 ZSAD와 ZNCC는 창 크기에 비해 계산 시간이 크게 증가한 것을 볼 수 있다. 이는 상관도 기반의 기법들은 유사도 척도를 계산하면서 명암도나 컬러값을 그대로 사용하는 경향이 있고 결국 큰 정수값을 처리하면서 CPU의 캐시메모리에 과부하를 많이 유발한다. 제안한 방법의 유사도 척도들은 각각 최대치가 1이하로 계산되고 5×5 크기의 비교적 작은 윈도우를 사용하기 때문에 과부하 유발이 적다. 제안한 방법을 다해상도 기법에 적용했을 때 GC를 적용했을 때보다 25%정도 계산시간이 감소되었다. 8방향으로 동적 계획법을 적용하고 폐색 영역을 추정하기 위해 양방향으로 정합을 수행하지만 다해상도 기법을 통해 계산의 효율성을 증가시킬 수 있었다.

제안한 기법이 Aloe 영상에 대해 정합을 1회 수행하면서 각 단계별로 수행시간의 분포를 확인하기 위해 단계별 소요시간은 표 6.14와 같다.

**표 6.14** Aloe 영상에 대한 제안한 방법의 단계별 계산 시간

**Table 6.14** Computational time of of each similarity measures and matching for 1 iteration about Aloe image pair

process entry	computational time(msec)
entropy	35
Haar wavelet feature vector	74
Census transform	116
proposed similarity measure	225
matching for 1 iteration	720

HEC 유사도 척도는 엔트로피, Haar 웨이블릿 특징 벡터, Census 변환으로 구

성된다. 각각의 유사도 척도를 계산하고 결합하는데 걸리는 시간이 225msec 정도였다. 이 중 Census 변환의 계산량이 가장 많다. 그 이유는 중심값을 이용해 주변 화소의 대소를 비교하여 이진 비트를 생성하고 해밍거리를 계산하는 과정을 거치기 때문이다. 유사도 척도를 계산하여 정합을 1회 반복하여 변위도를 개선하는데까지 약 720msec의 시간이 소요된다.

제안한 기법을 이용해 최종 변위도를 얻기까지 소요되는 시간을 다해상도 기법의 유무에 따라 각각 산출한 결과와 기존의 다해상도 기법을 적용하여 얻은 계산시간은 표 6.15와 같다.

**표 6.15** 다해상도 미적용시 3회 반복 시간, 기존의 다해상도 기법과 제안한 다해상도 기법의 3계층 처리 시간

**Table 6.15** Computational time of non-hirachical method for 3 iterations, conventional method and proposed hirachical method about 3 levels

entry	computational time(sec)
matching without multi-resolution method (3 iterations)	2.16
matchig with conventional multi-resolution method (3 levels)	0.96
<b>proposed method (3 levels)</b>	<b>1.23</b>

제안한 다해상도 기법의 영상 피라미드가 3단계이기 때문에 원영상 크기에서 3회 반복한 시간과 비교하였다. 그 결과 약 40%의 계산시간 단축 효과를 보였다. 기존의 다해상도 기법을 동일한 조건에서 적용했을 때 0.96초였고 이는 1.24초 걸린 제안한 방법에 비해 빨랐다. 그러나 최종 변위도의 정확도를 고려했을 때 제안한 방법의 효율성이 더 좋다고 할 수 있다.

## 6.4 제안한 방법의 정합 성능에 대한 고찰

### 1) RDS 영상 실험

제안한 방법의 오정합율은 기존의 MI보다 0.27% 적었다. 특히 잡음이 20% 첨가된 영상에 대한 변위도를 통해 제안한 방법이 변위 불연속 영역에서 성능이 좋음을 확인하였다. 이는 제안한 방법의 Census 변환이 경계선 영역의 정합 성능을 향상시킴을 의미한다. 그리고 전체 변위도에 대한 오차율이 2.87%, 6.39%를 기록한 MI보다 우수하였다. RDS 영상의 실험을 통해 제안한 방법의 전체적으로 변위를 추정하는데 좋은 성능을 보임을 알 수 있었다.

### 2) 환경 변화가 없는 표준 영상 실험

그래프 컷을 최적화 방법으로 적용했을 때, 제안한 유사도 척도는 기존의 NCC와 MI보다 경계선 영역과 무질감 영역에서 성능이 우수하였다. 이는 제안한 유사도 척도가 Haar 웨이블릿 특징 벡터와 Census 변환으로 계산한 공간 정보를 이용해 대응점들의 유사도를 잘 표현했기 때문이라 판단된다. 제안한 유사도 척도를 그래프 컷과 제안한 최적화 방법에 적용했을 때, Venus 영상에 제안한 방법을 적용했을 때 점진적인 변위의 변화를 가장 잘 반영하였다. 제안한 방법의 오정합 화소 비율은 변위 불연속 영역에서 5.96%로 가장 작았고, 16개 영상에 대해 계산한 사실 변위도의 오정합 화소 비율 평균도 3.63%로 가장 작았다.

제안한 방법은 기울어진 물체에서 나타나는 점진적으로 변하는 변위를 잘 추정하였고 변위 불연속 영역에서 오정합 화소 비율을 감소시켰다. 경계선 주변의 변위가 경계선 방향으로 일정하다는 특성을 이용하여 동적 계획법을 8방향으로 적용했기 때문에 변위 불연속 영역에서 좋은 성능을 보였고, 주변 화소와의 변위 차이에 따라 에너지 증가량을 변별적으로 적용하여 에너지 함수를 통합하였기 때문에 점진적인 변하는 변위를 잘 추정할 수 있었다고 판단된다.

### 3) 환경 변화가 발생한 표준 영상 실험

카메라 노출 시간이 다른 영상에 대해 제안한 유사도 척도는 MI의 성능을 20%에서 40%까지 개선시켰다. 이는 Haar 웨이블릿 특징 벡터는 대비의 변화에 강인하고 Census 변환 또한 비모수 특성으로 인해 대비가 변한 영상에서도 경계선과 무질감 영역을 잘 표현하기 때문이다. 조명이 다른 영상의 실험에서 제안한 유사도 척도는 지역적으로 밝기 변화가 발생한 경우에 적합한 유사도 척도임을 확인하였다. 이는 Haar 웨이블릿 특징 벡터는 주파수 특성을 이용하여 추출한 공간 정보를 이용하고 변형된 Census 변환은 지역적인 밝기 변화가 발생한 영역에서 이진 비트를 안정적으로 발생시키기 때문이라 판단된다. 한편 조명의 변화로 인해 동일한 영역의 명암도 변화 형태가 크게 달라졌을 때 유사도를 정확히 반영하지 못하는 점은 향후 보완되어야 할 것이다.

### 4) 실제 영상 실험

두 카메라의 조건이 같을 때, 제안한 방법과 기존 방법들의 특징이 환경 변화가 없는 영상 실험과 동일하게 나타났다. 조리개 개방 정도가 다른 영상에 대한 실험에서 제안한 방법은 다른 방법들 보다 강인한 성능을 보였다. 한편 제안한 방법은 다른 방법들과 마찬가지로 시차가 큰 영상에서 발생한 기하학적인 왜곡을 정확히 반영하지 못하였기 때문에 이에 대한 연구가 필요하다.

### 5) 계산 시간

제안한 HEC를 GC에 적용하여 최종 변위도를 얻을 때까지 걸리는 시간을 다른 유사도 척도들과 비교했을 때 공간정보를 포함하지 않는 상관도 기반 유사도 척도들보다 계산량이 많았다. 제안한 최적화 방법은 기존의 GC보다 25%정도 빠르게 처리할 수 있었다. 제안한 유사도 척도 계산에 소요되는 시간은 225ms였고 이 중 Census 변환 유사도를 계산하는데 116ms로 가장 오래 걸렸다. Census 변환에 소요되는 계산시간은 향후 H/W를 이용한 병렬처리 연구를 통해 해결할 수 있을 것이다.

## 제 7 장 결 론

본 논문에서는 환경 변화에 강인하고 효율적인 스테레오 정합을 위해 엔트로피 및 Census 기반의 유사도 척도와 다해상도 기반의 변위 추정을 위한 방법을 제안하였다. 제안한 방법의 유사도 척도는 Haar 웨이블릿 특징을 결합한 엔트로피와 변형된 Census 변환으로 구성되었고 변위 추정 최적화 방법은 8방향으로 동적 계획법을 적용하여 에너지를 통합하는 방법이다. 그리고 이 일련의 과정을 효율적으로 수행하기 위해 가우시안 피라미드를 이용한 다해상도 스테레오 정합 기법을 사용하였다.

유사도 척도는 기존의 엔트로피 기반 유사도 척도가 반영하지 못하는 공간 정보를 Haar 웨이블릿 특징 벡터와 변형된 Census 변환을 이용하여 반영한다. Haar 웨이블릿 특징 벡터는 분포된 명암도의 변화 패턴에 대한 주파수 성분을 계산하여 밝기 변화에 상관없이 영상의 공간 정보를 표현한다. 제안한 유사도 척도는 이러한 공간 정보를 이용하여 지역적인 밝기 변화가 발생한 영상에서 발생하는 엔트로피의 오차를 보정하였고, 최종 변위도를 얻는데 걸리는 수렴 횟수를 줄였다.

Haar 웨이블릿 특징 벡터는 엔트로피 유사도를 보정하는 공간 정보이고 변형된 Census 변환은 영상의 구조 정보를 부여하는 역할을 한다. 기존의 Census 변환은 중심 화소의 변형이 발생했을 때 성능이 좋지 않다. 제안한 Census 변환은 이러한 단점을 개선하기 위해 Census 변환 영역을 4개의 부영상 단위로 나누고 분산이 가장 적은 영역의 평균을 이진 비트 생성의 기준치로 사용하였다. 이는 중심 화소 자체가 변하거나 변환 영역의 화소들이 지역적으로 다르게 변해도 일정한 이진 비트를 생성할 수 있기 때문에 밝기 변화가 발생한 영상에서 경계선 영역과 무질감 영역의 유사도를 잘 반영하였다.

제안한 유사도 척도는 RDS 실험에서 기존의 MI보다 경계선 영역을 잘 보전하였고, 잡음이 첨가된 영상에서 기존의 방법들보다 좋은 성능을 보였다. 환경

변화가 없는 영상 실험을 통해 제안한 유사도 척도는 기존의 유사도 척도에 비해 변위 불연속 영역의 오정합이 작았다. 실제로 변위 불연속 영역 오정합을 평균이 18.84%에서 8.83%로 감소하였다. 전역적인 밝기 변화가 발생한 영상에서 제안한 유사도 척도는 기존의 MI의 오정합율을 최고 40% 감소시켰고, 지역적인 밝기 변화가 발생한 영상에서 최고 54% 감소시켰다. 또한 제안한 방법은 공간 정보를 이용하여 전체적으로 조밀한 변위도를 생성하였고, 특히 경계선 영역의 변위를 잘 반영하였다.

변위 추정 최적화 방법은 8방향으로 동적 계획법을 적용하는 방법을 이용하였다. 제안한 최적화 방법은 신뢰 확산법과 그래프 컷 등 성능이 검증된 전역적 최적화 방법의 높은 계산복잡도 문제를 해결하고 기존 동적 계획법의 단점인 스트리킹 현상을 개선하기 위해 다른 탐색 방향의 누적 에너지를 이용하여 변위를 추정하였다.

실험을 통해 제안한 최적화 방법은 스트리킹 현상을 해결함을 보였고, 동일한 유사도 척도를 이용한 실험에서 그래프 컷보다 비폐색 영역뿐만 아니라 변위 불연속 영역에서 좋은 성능을 보였다. 제안한 최적화 방법은 누적 에너지 통합 과정에서 변위의 차이를 변별적으로 반영하여 점진적인 변위의 변화를 잘 반영하였고 8방향으로 동적 계획법을 적용하여 경계선 영역에서 좋은 성능을 보였다. 환경의 변화가 없는 영상의 실험에서 변위 불연속 영역의 정합 성능이 기존의 그래프 컷보다 우수했고, 환경 변화가 발생한 실험에서 제안한 방법의 경계선 보전 성능이 우수함을 확인하였다.

제안한 방법은 효율적으로 정합을 수행하기 위해 가우시안 피라미드를 이용한 다해상도 기법을 적용하였다. 기존의 다해상도 기법은 최상위 계층에서 얻은 저해상도 변위도를 원영상 크기까지 전파하고 보간하는 과정을 거치기 때문에 계산 복잡도 측면에서 큰 이점이 있으나 최초 변위도가 정확하지 않으면 그 영향이 하위 계층까지 미쳐 정확한 변위도를 얻기 힘들다. 제안한 방법은 각 계층마다 정합을 수행하여 계산 시간을 줄였고 기존의 다해상도 기법보다 정확한 변위도를 보장하였다. 제안한 방법은 3개의 유사도 척도를 이용하므로 유사

도를 계산하는 과정에서 다른 방법들에 비해 시간이 많이 소요되지만 다해상도 기법을 적용하여 이러한 단점을 보완할 수 있었다.

제안한 방법은 환경 변화의 여부에 관계없이 기존의 방법들보다 좋은 성능을 보였다. 한편 지역적으로 불균일하게 밝기 변화가 발생하여 대응점 주변의 명암도 변화 패턴과 방향성이 크게 달라졌을 때 제안한 Haar 웨이블릿 특징 벡터의 유사도가 다르게 반영될 수 있고, 대비가 크게 축소된 영상에서 변형된 Census 변환이 오차를 발생시켰다. 그리고 정합 과정에서 유사도 계산에 걸리는 시간의 비중이 다른 방법에 비해 컸다. 따라서 이와 같은 스테레오 영상의 변형에 대해 강인한 성능을 보일 수 있는 유사도 척도에 대한 연구와 실시간 처리를 위해 GPU나 FPGA 등 하드웨어를 이용한 병렬 처리 기법에 대한 연구가 필요하다.



## 참고 문헌

- [1] E. R. Davies, *Machine Vision, Third Edition : Theory, Algorithms, Practicallities(Signal Processing and its Applications)*, Morgan Kaufmann, 2005.
- [2] U. R. Dhond and J. K. Aggarwal, "Structure from Stereo - A Review," *IEEE Transactions on Systems Man and Cybernets*, Vol. 19, No. 6, pp. 1489-1510, 1989.
- [3] G. Medioni and R. Nevatia, "Segment-based Stereo Matching," *Computer Vision Graphics Image Process*, Vol. 31, pp. 2-18, 1988.
- [4] J. Siebert and C. Urquhart, "Active Stereo: Texture Enhanced Reconstruction," *IET JOURNALS & MAGAZINES Electronics Letters*, Vol. 26, No. 7, pp. 156-166, 1990.
- [5] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *International Journal of Computer Vision*, pp. 7-42, 2002.
- [6] R. Horadu and Th. Skordas, "Stereo Correspondence Through Feature Grouping and Maximum Cliques," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 11, pp. 1168-1180, 1989.
- [7] M. Patel and F. Cohen, "Local Surface Shape Estimation of 3-D Textured Surfaces Using Gaussian Markov Random Fields and Stereo Windows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1091-1098, 1993.
- [8] S. K. Hong and S. J. Cho, "A Study on Genetic Algorithm and Stereo Matching for Object Depth Recognition," *Korean Institute of Navigation and Port Research Journal in Korean*, Vol. 32, No. 5, pp. 355-361, 2008.

- [9] F. Barrera Campo, F. Lumbreras Ruiz, and A. D. Sappa, "Multimodal Stereo Vision System: 3D Data Extraction and Algorithm Evaluation," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6, No. 5, pp. 437-446, 2012.
- [10] J. Sun, N. Zheng, and H. Shum "Stereo Matching Using Belief Propagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 7, pp. 787-800, 2003.
- [11] N. M Nasrabadi, "A Stereo Vision Technique Using Curve Segments and Relaxation Matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 5, pp. 566-572, 1992.
- [12] V. Kolmogorov, A. Criminisi, A. Blake, G. Cross, and C. Rother, "Probabilistic Fusion of Stereo with Color and Contrast for Bilayer Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 9, pp.1480-1492, 2006.
- [13] R. Feris, R. Raskar, C. Longbin, K. Tan, and M. Turk, "Multiflash Stereopsis: Depth-Edge-Preserving Stereo with Small Baseline Illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 1, pp. 147-159, 2008.
- [14] S. K. Hong and S. J. Cho, "A Multiresolution Stereo Matching with Modified Census Transform and Adaptive Window," *2012 Conference on Korea Information Processing Society in Korean*, Vol. 19, No. 2, pp. 572-575, 2012.
- [15] T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 9, pp. 920-932, 1994.
- [16] N. Chang, T. Tsai, B. Hsu, T. Chen, and T. Chang, "Algorithm and

- Architecture of Disparity Estimation with Mini-Census Adaptive Support Weight,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 6, pp. 792-805, 2010.
- [17] K. Pauwels, M. Tomasi, J. Diaz Alonso, E. Ros, and M. M. Van Hulle, “Segment-Based Stereo Matching Using Belief Propagation a Self-Adapting Dissimilarity Measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 61, No. 7, pp. 999-1012, 2012.
- [18] R. Kalarot, J. Morris, and G. Gimel’farb, “Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation, and Occlusion Handling,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 3, pp. 492-504, 2008.
- [19] E. Gudis, G. van der Wal, S. Kuthirummal, and S. Chai, “Establishing Visual Correspondence from Multi-Resolution Graph Cuts for Stereo-Motion,” *IEEE 20th Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pp. 29-32, 2012.
- [20] S. K. Hong and S. J. Cho, “A Stereo Matching Based on A Genetic Algorithm Using Multi-resolution Method and AD-Census,” *The Korea Institute of Signal Processing and Systems Journal in Korean*, Vol. 13, No. 1, pp. 12-18, 2012.
- [21] S. Gurumani, H. Cholakkal, Y. Liang, K. Rupnow, and D. Chen, “Stereo Matching with Reliable Disparity Propagation,” *2011 18th Asia and South Pacific Design Automation Conference*, pp. 305-312, 2011.
- [22] K. J. Yoon and I. S. Kwon, “Adaptive Support-Weight Approach for Correspondence Search,” *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol. 28, No. 4, pp. 650-656, 2006.
- [23] Y. Tseng, J. Hwang and F. H. Sheehan, “Three-dimensional object

- representation and invariant recognition using continuous distance transform neural networks,” *IEEE Transactions on Neural Networks*, Vol. 8, No. 1, pp. 141-147, 1997.
- [24] J. J. Lee, J. C. Shim, and Y. H. Ha, “Stereo correspondence using hopfield neural of new energy function,” *The Journal of the Pattern Recognition Society*, Vol. 27, pp. 1513-1522, 1994.
- [25] M. Gong and Y. Yang “Real-Time Stereo Matching Using Orthogonal Reliability-Based Dynamic Programming,” *IEEE Transactions on Image Processing*, Vol. 16, No. 3, pp. 879-884, 2007.
- [26] Y. Ohta and T. Kanade, “Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7, No. 2, pp. 139-154, 1985.
- [27] S. K. Hong and S. J. Cho, “A Multiresolution Stereo Matching Based on Genetic Algorithm using Edge Information,” *Korea Information Processing Society Journal in Korean*, Vol. 17-B, No. 1, pp. 63-68, 2010.
- [28] K. P. Han, K. W. Song, E. Y. Chung, S. J. Cho, and Y. H. Ha, “Stereo Matching Using Genetic Algorithm with Adaptive Chromosomes,” *The Journal of the Pattern Recognition*, Vol. 34, No. 9, pp. 1729-1740, 2001.
- [29] S. K. Hong, “Stereo Matching Using Edge Information and Genetic Algorithm,” Korea Maritime University Master Degree Thesis, 2008.
- [30] L. Wang, R. Yang, and J. E. Davis, “BRDF Invariant Stereo Using Light Transport Constancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 9, pp. 1616-1626, 2007.
- [31] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs, “Appearance Characterization of Linear Lambertian Objects, Generalized Photometric

- Stereo, and Illumination-Invariant Face Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 2, pp. 230-245, 2007.
- [32] Y. Heo, K. Lee, and S. Lee, “Illumination and Camera Invariant Stereo Matching,” *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2008.
- [33] L. Liu, P. Fuyuan, T. Yan, and W. Yaping, “Scale Invariant Features Extraction for Stereo Vision,” *BIAI Journal of Systems Engineering and Electronics*, Vol. 20, No. 1, pp. 50-54, 2009.
- [34] X. Chang, Z. Zhong, and Q. Zhong, “Apply Two Pass Aggregation to Real-Time Stereo Matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 2, pp. 387-391, 2010.
- [35] P. Viola and W. Wells, “Alignment by Maximization of Mutual Information,” *IEEE Proceeding in Sixth IEEE Intelligent Conference of Computer Vision*, Vol. 24, No. 2, pp. 137-154, 1997.
- [36] G. Egnal, “Mutual Information as a Stereo Correspondence Measure,” *Technical Report MS-CIS-00-20, Computer and Information Science*, University of Pennsylvania, 2000.
- [37] J. Kim, V. Kolmogorov, and R. Zabih, “Visual Correspondence Using Energy Minimization and Mutual Information,” *IEEE Proceeding in Intelligent Conference of Computer Vision*, Vol. 2, pp. 1033-1040, 2003.
- [38] I. Sarkar and M. Bansal, “A Wavelet-Based Multiresolution Approach to Solve the Stereo Correspondence Problem Using Mutual Information,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, Vol. 37, No. 4, pp. 1009-1014, 2007.

- [39] A. Ogale and A. Yiannis, "Robust Contrast Invariant Stereo Correspondence," *In Proceedings of Conference on Robotics and Automation*, pp. 819-824, 2004.
- [40] U. Raghavendra, "Entropy based Log Chromaticity Projection for Real-time Stereo Matching," *2nd International Conference on Communication, Computing and amp; Security*, Vol. 6, pp. 223-230, 2012.
- [41] K. Whisnant, R. Iyer, Z. Kalbarczyk, and P. Jones, "The Effects of an ARMOR-based SIFT environment on the performance and dependability of user applications," *IEEE Transactions on Software Engineering*, Vol. 30, No. 4, pp. 257-277, 2004.
- [42] W. Hongxia, Y. Kejian, G. Feng, and L. Jun, "Normalization Methods of SIFT Vector for Object Recognition," *IEEE Transactions on Image Processing*, Vol. 21, No. 11, pp. 2656-2661, 2002.
- [43] W. Eric and L. Grimson, "Pose and Expression Independent Facial Landmark Localization Using Dense-SURF and the Hausdorff Distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 3, pp. 624-638, 2012.
- [44] C. Brodid and C. Hayes, "DAISY: a decision support design methodology for complex, experience-centered domains," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, Vol. 32, No. 1, pp. 50-71, 2011.
- [45] S. Birchfield and C. Tomasi, "Depth Discontinuities by Pixel-to-Pixel Stereo," *IEEE Proceeding in Sixth IEEE Intelligent Conference of Computer Vision*, pp. 1073-1080, 1998.
- [46] S. Birchfield and C. Tomasi, "A Pixel Dissimilarity Measure That Is Insensitive to Image Sampling," *IEEE Transactions on Pattern Analysis and*

*Machine Intelligence*, Vol. 20, No. 4, pp. 401-406, 1998.

- [47] D. N. Bhat and S. K. Nayar, "Ordinal Measures for Image Correspondence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 4, pp. 415-423, 1998.
- [48] H. Hirschmuller, P. Innocent, and J. Garibaldi, "Real-time Correlation-based Stereo Vision with Reduced Border Errors," *International Journal of Computer Vision*, Vol. 47, No. 1, pp. 229-246, 2002.
- [49] M. Weber, M. Humenberger, and W. Kubinger, "A Very Fast Census-based Stereo Matching Implementation on A Graphics Processing Unit," *2009 IEEE 12th International Conference on Computer Vision Workshops*, pp. 786-793, 2009.
- [50] M. Manzano, D. Ojeda, M. Devy, J. Boizard, and J. Fourniols, "Stereo Vision Algorithm Implementation in FPGA Using Census Transform for Effective Resource Optimization," *DSD '09. 12th Euromicro Conference on Digital System Design, Architectures, Methods and Tools*, pp. 799-805, 2009.
- [51] S. K. Hong and S. J. Cho, "A Multiresolution Stereo Matching Using AD-Census Cost Measure and Genetic Algorithm," *2011 Conference on Korea Information Processing Society in Korean*, Vol. 20, No. 1, pp. 89-94, 2012.
- [52] J. Weng, "Image Matching Using The Windowed Fourier Phase," *International Journal of Computer Vision*, Vol. 11, No. 3, pp. 211-236, 1993.
- [53] S. Berlemont and J. Olivo-Marin, "Combining Local Filtering and Multiscale Analysis for Edge, Ridge, and Curvilinear Objects Detection," *IEEE Transactions on Image Processing*, Vol. 19, No. 1, pp. 74-84, 2010.
- [54] S. K. Hong and S. J. Cho, "A Multiresolution Stereo Matching based on

- Genetic Algorithm Using Local Feature Information,” *2010 Conference on Korea Information Processing Society in Korean*, Vol. 17, No. 2, pp. 758-761, 2012.
- [55] M. Thomas, A. Thomas, and T.M. Cover, *Elements of Information Theory*, John Wiley & Sons, 2005.
- [56] R. Horadu and Th. Skordas, “Depth Map Construction from Range-Guided Multiresolution Stereo Matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 11, pp. 1168-1180, 1994.
- [57] N. M Nasrabadi, “Multiresolution Energy Minimization Framework for Stereo Matching,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 5, pp. 566-572, 2012.
- [58] Y. Wang, C. Tung, and P. Chung, “Efficient Disparity Estimation Using Hierarchical Bilateral Disparity Structure Based Graph Cut Algorithm With a Foreground Boundary Refinement Mechanism,” *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 23, No. 5, pp. 784-801, 2013.
- [59] F. Barrera Campo, F. Lumbreras Ruiz, and A.D. Sappa, “Multimodal Stereo Vision System: 3D Data Extraction and Algorithm Evaluation,” *IEEE Journal of Selected Topics in Signal Processing*, Vol. 6, No. 5, pp. 437-446, 2012.
- [60] G. Medioni and R. Nevatia, “Segment-based Stereo Matching,” *Computer Vision Graphics Image Process*, Vol. 31, pp. 2-18, 1988.
- [61] M. Okutomi and T. Kanade, “A Multiple-baseline Stereo,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, pp. 353-363, 1993.

- [62] G. Stockman, S. Chen, G. Hu, and N. Shrikhande, "Near-Duplicate Keyframe Identification With Interest Point Matching and Pattern Learning," *IEEE Transactions on Multimedia*, Vol. 9, No. 5, pp. 1037-1048, 2007.
- [63] Z. Chen and S. Sun, "A Zernike Moment Phase-Based Descriptor for Local Image Representation and Matching," *IEEE Transactions on Image Processing*, Vol. 19, No. 1, pp. 205-219, 2010.
- [64] N. Shiono and C. Hashimoto, "Threshold-Voltage Instability of n-Channel MOSFET's Under Bias-Temperature Aging," *IEEE Transactions on Electron Devices*, Vol. 29, No. 3, pp. 361-368, 2008.
- [65] S. Thompsom and S. Kagami, "Humanoid Robot Localisation Using Stereo Vision," *In Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots*, pp. 19-25, 2005.
- [66] W. Burgard, D. Fox, D. Henning, and T. Schmidt, "Estimating The Absolute Position of A Mobile Robot Using Position Probability Grids," *In Proceedings of National Conference on Artificial Intalligence*, 1996.
- [67] X. Mei, X. Sun, M. Zhou, S. Jiao, H. Wang, and X. Zhang. "On building an accurate stereo matching system on graphics hardware," GPUCV 2011.
- [68] Middlebury Stereo Vision Lab, Middlebury College, Middlebury, VT 05753, USA, "<http://vision.middlebury.edu/stereo/>" accessed on 20 April 2013.