

# 투영된 패턴과 원영상의 합성을 이용한 스테레오 매칭

장지호, 정재찬, 조재일  
한국전자통신연구원  
e-mail : changjh@etri.re.kr

## Texture projected Stereo matching using DoE pattern and original image

Jiho Chang, Jae-chan Jeong, Jae-il Cho  
Electronics and Telecommunication Research Institute

### 요 약

기존의 스테레오 매칭 시스템은 얻어지는 영상에 따라 disparity의 결과가 크게 차이를 나타내게 된다. 이러한 오류를 줄이고자 일정한 패턴을 주사하여 스테레오 매칭을 계산하는 방식인 액티브 스테레오 기법이 대두되고 있다. 본 논문에서는 이러한 액티브 스테레오 매칭을 사용시에 얻어질 수 있는 패턴 영상과 패턴이 없는 원영상을 서로 조합하여 스테레오 매칭을 수행함으로써 정확한 disparity를 얻고자 한다. 이러한 영상의 합성시에 두 영상의 비율에 따라서, 얻어지는 영상의 밝기 구성에 따라서 disparity결과의 차이와 이를 실제 시스템을 구성시에 필요한 점에 대해서 확인하고자 한다.

### 1. 서론

스테레오 매칭은 기본적으로 두 개의 카메라로부터 얻어진 영상으로부터 서로 대응되는 점들을 찾아 이에 대한 영상에서의 거리를 disparity라 하여, 이를 통해서 물체와의 거리를 알아내는 방식이다.[4] 대부분의 스테레오 매칭에서는 일반 카메라를 이용하여 영상정보를 획득하기 때문에 두 카메라에서 일반적으로 영상의 발생하여지는 문제점에 의해서 거리정보의 정확도와 정합율(매칭 성공률)이 달라지게 된다. 특히 대부분의 매칭 알고리즘이 텍스처 정보가 많을수록 정확하게 나오며, 반대로 텍스처가 반복되거나 없는 영역에서는 매칭에 실패하거나 거리 정도의 정확도가 크게 떨어지게 된다. 기존의 패시브 스테레오 매칭 시스템의 단점을 보정하기 위하여 다양한 방법을 사용하고 있다. 이중에서 일정한 패턴을 주사하고 이를 활용하여 거리 정보를 계산하는 액티브 스테레오 방식이 오류를 줄이는데 가장 확실한 방법으로 대두되고 있다.[1][3] 또는 MS의 Kinect와 같은 DoE 패턴을 조사하고 이와 카메라의 삼각법을 사용하여 Disparity를 구하는 방식도 있다. 일반적으로 액티브 스테레오 방식은 일반적으로 얻어진 영상으로부터 패턴만을 분리하여 거리정보를 얻게 된다. 하지만 이러한 방식은 패턴이 주사되지 않은 폐색영역이나 패턴이 잘 반사되지 않는 부분에서 많은 오류가 발생하게 되며, 야외환경과 같은 프로젝션되는 패턴과 같은 과정의 빛이 들어오는 곳에서도 패턴이 보이지 않아서 거리 정보를 획득할 수가 없게 된다. 본 논문에서는 이러한 점을 해결하고자 패턴과 원본 이미지의 텍스처가 혼재한 상

태에서의 스테레오 매칭을 가정하고자 한다. 이때 원본 영상과 패턴이 합성될 시에 두 영상간의 상대적인 밝기 정보의 비율에 의해서 disparity 정보를 얻을 때 영향을 미치게 된다. 본 논문에서는 실제로 촬영된 패턴 영상을 통해서 원본과 패턴의 세기 정도가 스테레오 매칭 결과에 미치는 영향을 알아보고, 두 정보를 모두 사용하는 시스템을 구성할 경우에 카메라의 셋팅과 두 영상정보를 합성하는 방법에 대해서 논의하고자 한다.

### 2. 패턴과 원영상의 합성

액티브 방식의 스테레오 매칭의 사용하기 위해선 광원을 사용하여 패턴을 조사하여야 한다. disparity를 측정하는 피사체가 물체의 경우에는 광원의 사용에 대해서 크게 문제가 없으나, 피사체가 사람의 경우에는 가시광을 이용한 방법은 크게 제약을 받게 된다. 본 논문에서는 Kinect와 유사하게 IR영역에서 DoE를 사용하여 패턴을 조사하는 것을 사용하도록 한다. 특히 DoE를 사용함으로써 점단위의 패턴을 사용가능하는 것도 장점이다. 위와 같이 IR대역의 패턴의 성능을 높이기 위해선 필터를 사용하게 되고, 원영상도 IR대역을 받아야 하므로 IR광원이 추가로 필요하게 된다. 이를 한 시스템에서 구현하고자 영상의 프레임 단위에 따라서 한번은 DoE 패턴 조사를 위한 레이저를 다음 프레임에서는 IR LED를 점등하는 방식으로 구현하고자 한다. 다음의 그림은 본 논문에서의 시스템을 사용하여 얻어진 패턴영상과 LED광원을 통한 원영상을 나타내고 있다.



(그림 1) DoE 패턴과 원영상

서로 상이한 영상을 합성하는 방법으로는 Truncated summation, Alpha blending, Guassian 등의 방법이 사용되어 지고 있다. [1][2] 이 중 Alpha blending은 간단한 방법으로 구현이 가능하기 때문에 다양한 영상처리 분야에서 사용되고 있다. 하지만 얻고자 하는 결과에 대해서는 입력되어지는 파라미터에 따라서 많이 다르기 때문에 적합한 파라미터를 찾기 위한 노력이 필요하다. Alpha blending은 일반적으로 다음의 식으로 표현이 가능하며, 식에서 alpha는 실험적으로 구해야 한다. 그림 2는 위 시스템에서 획득된 영상을 이용하여 각각의 alpha 비율에 따라서 합성된 결과를 나타낸다.

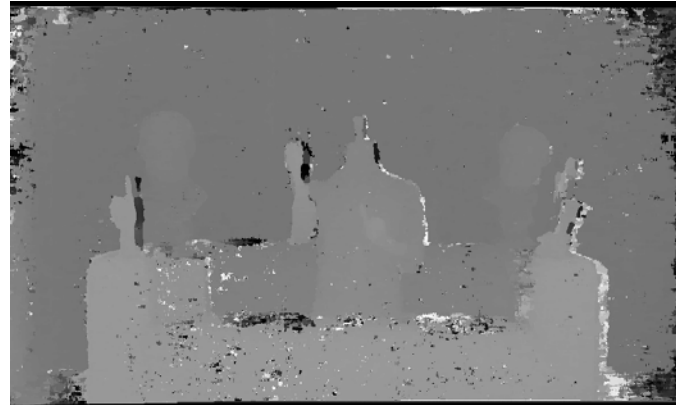


(그림 2) Alpha 비율에 따라서 합성된 영상

### 3. 스테레오 매칭 알고리즘

스테레오 매칭 알고리즘은 D. Scharstein and R. Szeliski의 논문[4]에서 matching cost computation, cost aggregation, disparity computation, refinement의 4 단계로 알고리즘을 구분될 수 있다. 본 논문에서는 matching cost 로서 AD-Census를 사용하고 cost aggregation으로

서 adaptive weight 방식인 Information permeability filtering(PF)를 사용하였다. PF는 파라미터가 간단하고 연산량이  $O(1)$  수준이다. disparity computation 방법으로는 WTA(Winner Takes All), Post-processing 기법은 사용하지 않았다. 상기의 나열된 알고리즘은 구조가 간단하며 구현이 용이하고 시뮬레이션의 시간이 오래 걸리지 않는 장점이 있다.



(그림 3) 스테레오 매칭으로 얻어진 Disparity영상

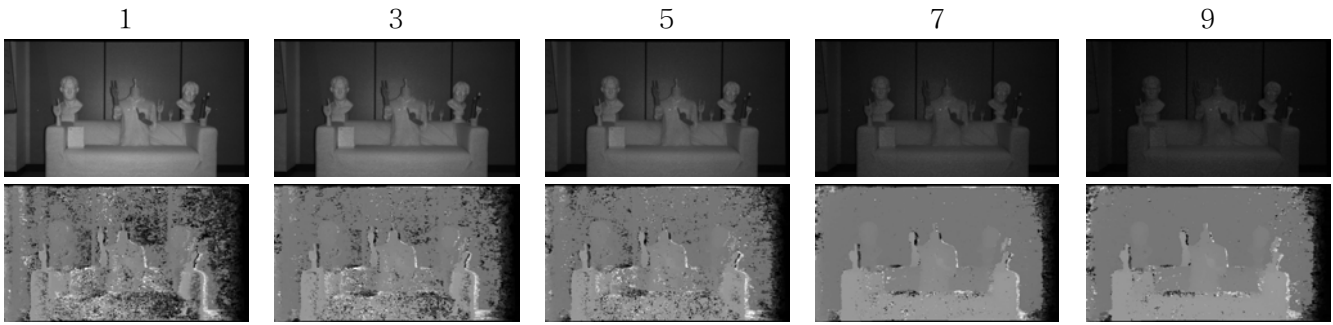
그림 3은 2절에서 소개된 합성된 영상을 상기의 알고리즘을 사용하여 얻어진 disparity 결과이다. 또한 다양한 알고리즘의 경우 입력되어지는 영상에 대한 강인도가 높으므로 입력 영상에 따른 disparity의 결과 차이가 크게 나타나지 않게되기 때문에 간단한 알고리즘으로 수행하기로 했다.

### 4. 실험 결과

실험은 영상을 얻을 시에 카메라의 파라미터를 조정하여 서로 다른 노출 및 Gain을 가지는 영상들의 조합을 사용하여 영상을 합성하고 3절에서 설명되어진 스테레오 매칭 알고리즘을 통해 얻어진 Disparity의 변화를 살펴보고자 하였다. 본 실험에서는 groundtruth를 획득하기가 용이하지 않기 때문에 각 disparity간의 상대적인 차이를 사용하여 비교를 하고자 한다.

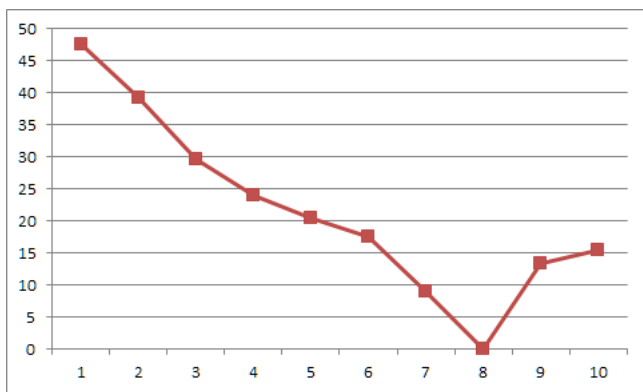
그림 4는 패턴과 원본영상의 비율에 따른 Disparity 결과를 나타낸 것이다. 촬영된 영상에서 벽의 경우 Textureless 구간이기 때문에 패턴보다 원본영상의 비율이 클수록 에러가 나타나게 되며, 패턴의 비율이 높아질수록 에러가 줄어드는 것을 확인할 수 있다.

정성적으로 평가하기에 7번 영상이 다른 영상에 비하여 좋은 결과를 나타내기 때문에 이를 groundtruth처럼 가정하고 각 영상에 대해서 에러율을 측정하였다. 본 논문에서는 전체 영역에 대하여 disparity 값이 1이상 차이난 픽셀의 개수를 측정하여 전체 픽셀에 대한 비율로 에러율을 나타내었다. 패턴에 비해 원본 영상이 다음의 그림 5는 7번 영상을 기준으로 에러율이 증가함으로서 패턴의 비율



(그림 4) 패턴과 원영상의 비율에 따른 Disparity

이 낮거나 너무 높은 경우에는 에러율이 증가함을 알 수 있다.



(그림 5) 패턴 비율에 따른 에러율

### 5. 결론 및 향후 연구 방향

본 논문에서는 기존의 패시브 방식의 스테레오 매칭을 극복하고자 제안된 액티브 스테레오 매칭 방식에서 DoE 이용한 패턴을 사용한 환경을 사용하였다. 또한 이렇게 구성된 환경에서 원본 영상과의 패턴 영상과의 비율을 서로 달리 하여 스테레오 매칭을 수행한 결과 어느 한쪽이 강한 것 보다는 두 영상의 특징이 잘 나타내는 영상이 가장 좋은 결과를 얻는다는 것을 확인하였다. 차후에 다양한 스테레오 매칭의 수행을 통한 검증으로 액티브 방식의 스테레오 매칭에서의 결과값을 예측하고, 이를 시스템 구성시에 활용하고자 한다.

### Acknowledgement

"본 연구는 미래창조과학부방송통신위원회의 ETRI 연구개발지원사업의 일환으로 수행되었음 [11921-03001, "Beyond 스마트TV 기술개발"]"

### 참고문헌

[1] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo", IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI),

vol. 25, no. 8, pp. 993 - 1008, AUGUST 2003.

[2] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light", in Proc. Computer Vision and Pattern Recognition (CVPR), vol. 1, 2003, pp. 195 - 202.

[3] Jongwoo Lim, "Optimized Projection Pattern Supplementing Stereo Systems", IEEE International Conference on Robotics and Automation, Kobe, Japan, May 12-17, 2009

[4] D. Scharstein and R. Szeliski. "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms." International Journal of Computer Vision, 47(1/2/3):7-42, April-June 2002.

[5] Cigla, C., Alatan, A.A. "Efficient Edge-Preserving Stereo Matching", ICCV LDRMC, pages 696-699, 2011