

## 경계영역 정확도 향상을 위한 영상분할 기반 스테레오 매칭

문지훈 호요성

광주과학기술원

{jhm, hoyo}@gist.ac.kr

## Image Segment-Based Stereo Matching for Improving Boundary Accuracy

Ji-Hun Mun Yo-Sung Ho

Gwangju Institute of Science and Technology

## 요약

3차원 영상을 생성하기 위해 스테레오 매칭을 통해 깊이 정보를 획득한다. 이때 발생하는 경계영역과 텍스처가 부족한 부분의 깊이정보 부정확성 문제를 해결하기 위해 영상 분할 기반 스테레오 매칭 방법을 제안한다. 일반적으로 사용하는 윈도우 기반 스테레오 매칭 결과를 기반으로 분할된 영상 내에서 최적의 변위 값을 재 할당함으로써 깊이정보의 정확성을 향상시킬 수 있다. Mean-shift는 참조 영상에서 화소 간 평균값 차이가 최대가 되는 영역들을 반복적으로 찾는다. 유사한 평균값을 갖는 영역들을 기반으로 영상을 분할하는 것을 Mean-shift를 이용한 영상분할 이라고 한다. 분할된 영상은 각 영역을 대표하는 패치 구조를 가지고 있어 참조 영상에 포함되어있는 잡음에 강인한 특성을 지닌다. 스테레오 매칭을 통해 화소별로 변위 값을 할당해주는 대신, 분할된 영상을 이용하여 각 분할 영역에 동일한 변위 값을 할당한다. 분할된 영상에 동일한 변위 정보를 할당할 경우 객체와 배경의 경계영역에서 잘못된 변위 값이 할당되는 경우가 발생한다. 이러한 경계 영역의 변위정보 부정확성을 보완하기 위해 화소의 기울기 향을 비용 값 계산 과정에 추가하여 단점을 보완한다. 최종 비용 값 계산을 통해 획득한 초기 변위 지도에 중간 값 필터를 적용하여 분류된 영역에 동일한 변위 값을 할당한다. 제안한 방법을 적용하여 경계영역의 정확도가 향상된 최종 변위 지도를 획득한다.

## 1. 서론

스테레오 매칭은 3차원 콘텐츠 제작과 다양한 분야에 적용 가능한 덕분에 많은 연구들이 진행되고 있다 [1,4]. 스테레오 매칭은 일반적으로 크게 4단계를 거쳐 진행된다. 주어진 스테레오 영상으로부터 윈도우 기반 혹은 전역기반 매칭과정을 통해 비용 값을 계산한다. 각 변위 탐색 범위별로 계산된 비용 값들을 합친 뒤, 그 중 비용 값을 최소로 갖는 변위정보를 해당 화소의 변위 값으로 할당한다. 마지막으로 획득한 변위지도에 후처리 필터를 이용해 잡음 영역이나 폐색영역을 처리하는 과정을 거친다.

Scharstein과 Szeliski [2]는 스테레오 매칭 알고리즘을 분류하고 지역과 전역 두 가지 방법으로 분류하여 정의했다. 지역 기반 스테레오 매칭 알고리즘에서 변위정보의 비용 값은 정해진 윈도우 크기 내에서의 화소 밝기와 색상 값에 의해 달라진다. 모든 지역 기반 알고리즘은 모든 비용 값을 합치는 과정을 통해 간접적 영상 평활화를 수행한다. 반면에 전역 기반 스테레오 매칭의 경우, 비용 값 계산 시 평활화 향이 명시되어 있으며 데이터향의 합으로 이루어진 비용 값을 최적화 하는 과정을 거친다. 일반적으로 널리 알려진 전역 기반 스테레오 매칭 방법에는 dynamic programming [3, 5], belief propagation [6] 그리고 graph cuts [7] 과 같은 방법이 있다. 지역 기반 알고리즘과 다르게 전역 기반 알고리즘은 하나의 화소 값에서 주변의 모든 화소 값으로부터 비용 값을 추정하므로 최적화 과정이 필요하다.

일반적으로 지역기반 알고리즘은 윈도우 내에서의 모든 비용 값을 더하거나 평균값을 사용함으로써 최종 비용 값을 합산한다. 가장 호

용적인 지역 기반 비용 값 추정 방법으로는 비 정규화된 필터를 이용하여 선행 시간 내에 계산이 가능한 알고리즘이 있다. Yoon과 Kweon은 깊이 영상의 경계영역을 보존할 수 있는 경계 영역 인식 필터를 사용하여 경계 영역의 변위정보 정확도를 높였다 [8].

최근에는 이전 스테레오 매칭 알고리즘보다 뛰어난 결과로 인해 영상 분할 기반 스테레오 매칭 방법들이 제안되고 있다 [9]. 영상 분할 방법은 분할된 영역과 초기에 획득된 변위 값들의 공간적 유사성이 높은 특징을 이용한다. 영상 분할을 이용한 스테레오 매칭은 크게 4단계로 나눌 수 있다. 지역기반 스테레오 매칭 알고리즘을 이용하여 초기 변위 지도를 획득한 후 각각의 스테레오 영상의 분할을 수행한다. 분할된 영상에 초기에 획득한 변위정보에 중간 값 필터를 이용하여 분할된 각각의 평면에 변위정보를 새롭게 정의한다.

본 논문에서는 분할된 영상에 동일한 값을 할당하여 획득한 변위 지도에서 발생하는 경계영역 부정확성 문제를 해결하기 위해 영상의 기울기 정보를 사용한다. 초기 변위 지도를 획득하기 위해 계산하는 비용 값에 영상 기울기 정보 향을 추가하여 변위지도를 획득한다. 영상의 기울기 정보를 고려하여 변위 값을 계산하기 때문에 경계영역과 같이 영상의 기울기 차이가 큰 영역에서의 값이 정확히 구별된다. 기울기 정보를 고려해 획득한 초기 변위 값을 분할된 영상의 각 패치에 동일한 값을 할당한 후, 중간 값 필터를 적용하여 최종 변위지도를 획득한다.

## 2. 윈도우 기반 스테레오 매칭

사람은 양안을 통해 받아들이는 영상들의 시차를 통해 깊이감을

느낀다. 스테레오 영상에서 모든 화소들에 대해 적절한 변위정보를 찾기 위해서는 많은 변위 후보 값들이 존재한다. 지역기반 스테레오매칭 알고리즘은 미리 정해둔 윈도우 크기를 기반으로 해당 화소의 주변영역의 정보를 이용하여 비용 값을 최소로 하는 변위를 찾음으로써 효율적으로 변위 후보들을 줄인다. 가장 일반적인 윈도우 기반 비용 값 계산 방법으로는 sum of absolute differences (SAD)와 sum of squared differences (SSD)와 같이 화소의 색상 값을 이용한 방법들이 있다.

지역 기반 매칭을 실행할 때 식 (1)과 같이 윈도우 크기 안에서만 주어진 화소들의 색상 값들로부터 비용을 추정하기 때문에 객체의 경계 부근에서 부정확한 변위 값을 획득하는 문제가 발생했다.

$$Cost_{SAD}(x,y) = \sum_{(i,j) \in N(x,y)} |I_1(i,j) - I_2(i+d,j)| \quad (1)$$

경계 영역의 변위 부정확성 문제를 해결하기 위해 화소들의 기울기 변화를 고려하는 항을 식 (2)와 같이 나타낸다. 기울기 변화를 영상 내에서 고려하기 위해서는 현재 화소와 주변 화소 값의 차이를 알아야 한다. 하나의 화소를 중심으로 총 8방향의 이웃 화소를 가지고 있지만 계산량 증가로 인한 복잡도 문제를 고려하여 x축과 y축의 이웃 화소들과의 차이만 계산한다.

$$Cost_{Grad}(x,y) = \sum_{(i,j) \in N_x(x,y)} |\nabla_x I_1(i,j) - \nabla_x I_2(i+d,j)| + \sum_{(i,j) \in N_y(x,y)} |\nabla_y I_1(i,j) - \nabla_y I_2(i+d,j)| \quad (2)$$

$N(x,y)$ 는 윈도우 크기 내에 포함된 화소들을 나타내며 본 논문에서는  $7 \times 7$  크기를 갖는 윈도우를 사용했다.  $N_x(x,y)$ 는 영상의 맨 우측 영역을 제외한 영역을 나타내고  $N_y(x,y)$ 는 영상의 맨 아래 영역을 제외한 영역을 의미한다. 또한  $\nabla_x$ 는 x축 정방향 화소 기울기 값을,  $\nabla_y$ 는 y축 정방향 화소 기울기 값을 나타낸다. 색상 영상을 이용하여 비용 값을 계산하기 때문에 각각 모든 채널에서의 색상 불일치 정도를 반영할 수 있다.

$Cost_{SAD}$ 와  $Cost_{Grad}$ 의 합으로 이루어진 최종 비용 값에서 서로 다른 가중치를 부여하기 위해 가중치 인자  $\omega$ 를 사용한다.  $\omega$ 의 최적 값을 찾기 위해 좌영상 기준 스테레오 매칭 결과와 우영상 기준 스테레오 매칭 결과를 이용하여 교차검사를 수행한다. 최대 신뢰 화소 수를 갖게 하는  $\omega$ 값을 사용한다. 식 (3)에서 최대 신뢰 화소수를 가질 때 최소의 비용 값을 갖게 되므로 기존의 지역 기반 스테레오 매칭 방법과 부합한다.

$$C(x,y) = (1-\omega) * C_{SAD}(x,y) + \omega * C_{Grad}(x,y) \quad (3)$$

그림 1은 식 (3)을 Cambridge data set의 스테레오 영상에 적용하여 획득한 초기 변위 지도를 나타낸다. 일반적으로 매칭 비용을 계산하기 위해 식 (1)만 사용한 경우와 다르게 식 (3)을 사용하여 영상의 기울기 항을 추가 한 결과, 스테레오 매칭 결과 영상에서 객체 또는 배경

의 경계 영역이 두드러지게 나타난 것을 확인할 수 있다.

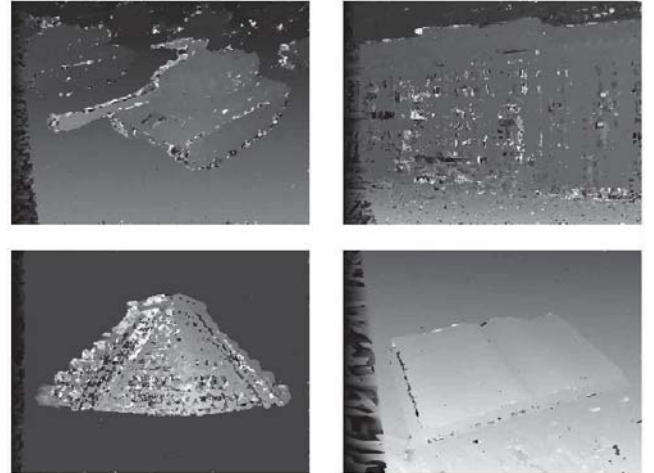


그림 1. 영상의 기울기를 고려한 스테레오 매칭 결과

### 3. 입력 영상의 색상 분할

스테레오 입력 영상으로부터 화소들의 기울기를 이용하여 초기 변위 정보를 획득한 뒤, 이 변위 영상들의 경계 영역에 나타나는 부정확한 변위 값을 보정해야 할 필요가 있다. 이를 위해 본 논문에서는 우선 mean-shift[10]를 이용한 영상의 색상 분할 방법을 입력 영상에 대해 적용시켜 동일한 색상을 갖는 영역(homogeneous) 혹은 회색 영역으로 나눈다.

이 알고리즘은 깊이 불연속 영역이나 객체의 경계 영역에서의 변위 값이 일정하게 변하는 것을 가정하여 분할을 수행한다. 색상 분할 알고리즘을 이용하여 최대한 많은 부분으로 분류되도록 하는 것이 추후 변위 정보를 보정하는데 좋은 역할을 한다. 다양한 부분으로 분류된 영역을 통해 기존의 가정이 성립할 확률을 높일 수 있기 때문이다.

Mean-shift를 이용한 영상의 색상 분할 방법은 우선 영상 내 각각의 화소들에 대해 mean-shift를 수행한다. 이 과정은 식 (4)가 최대값을 가질 때 까지 반복적으로 수행하게 된다. 식 (4)가 최대값을 가짐은 영상에서 유사 색상 영역의 경계까지 정확하게 찾을 수 있게 됨을 의미한다.

$$y_{j+1} = \sum_{i=1}^n x_i \cdot g(\| \frac{y_j - x_i}{h} \|^2) / \sum_{i=1}^n g(\| \frac{y_j - x_i}{h} \|^2) \quad (4)$$

식 (4)에서  $y_{j+1}$ 은 최대 분포를 갖는 확률의 연속되는 중앙을 의미하고  $y_j$ 는 현재 중심 위치 그리고  $x_i$ 는 주변에 분포된 확률들을 의미한다. 현재 화소에서의 중심이 되는 색상 영상의 중심과 주변 화소들의 색상 차이를 이용하여 미리 지정해둔 문턱 값 이상 차이가 발생하는 경우 서로 다른 영역이라고 판단하게 된다.

즉 mean-shift를 이용한 색상 영상 분할은 기본적으로 특징 공간에서 화소 간 기울기 차이가 최대가 되는 영역을 찾아가는 과정을 필요로 한다. 특징 공간은 좌표 공간과 영역 분할 분석에 관련된 것들과의 조합으로 이루어져 있다. Mean-shift를 이용한 영상분할의 장점은 각 화소 색상 차이의 최대값을 반복적으로 계산해나가는 알고리즘을

사용하므로 영상에 포함된 객체 영역 경계 영역을 효율적으로 찾아낼 수 있다는 점이다. 그림 2는 mean-shift를 이용해 색상 영상을 효율적으로 분류한 결과를 나타낸다.

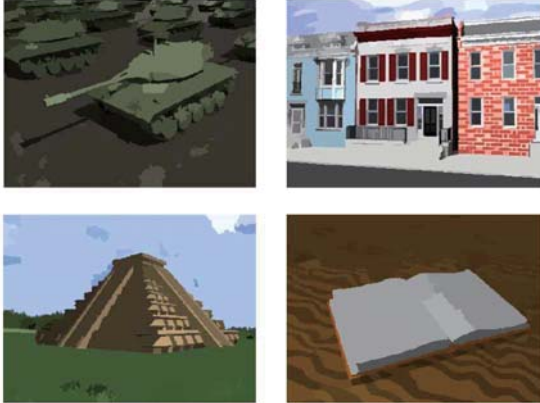


그림 2. Mean-shift를 이용한 영상 분할 결과

그림 2에서 알 수 있듯이 영상에 있는 객체 내부의 유사한 영역은 동일한 하나의 영역으로 나타나 있으며, 객체의 경계 부분은 명확하게 분류되어 있음을 확인할 수 있다.

#### 4. 초기 변위 지도 보완

경계 영역을 고려한 윈도우 기반 스테레오 매칭결과 영상에서 나타나는 경계영역의 변위 값 부정확성과 객체 내에서의 변위 값 오류를 수정하기 위해서 분할된 색상 영상정보를 사용한다. 초기 변위 정보와 분할된 색상 영상을 이용한 변위지도 신뢰도검사를 통해 각 영상을 대표하는 적절한 변위정보를 도출 할 수 있다. 이 과정은 크게 두 가지 단계로 나뉘어 처리된다. 첫째로 색상 분할 영역과 초기 변위지도와의 매칭 관계를 형성하고, 매칭 되는 영역의 값을 중간 값 필터를 이용하여 채워나가는 과정을 수행한다.

##### 1. 분할 영역과 초기 변위 매칭

윈도우 기반 스테레오 매칭 결과와 색상 분할 영상과의 매칭 관계를 형성하기 전, 좌영상 기준 스테레오 매칭 결과와 우영상 기준 스테레오 매칭 결과를 이용하여 각 화소별 최소의 비용 값을 갖게 하는 Winner-Takes-All(WTA)알고리즘을 적용한다. 이를 통해 얻어진 새로운 변위 값을 이용하여 개선된 변위 지도를 생성한다.

그림 3은 좌영상과 우영상을 기준으로 스테레오 매칭을 수행한 결과를 이용하여 최소 비용 값을 갖는 새로운 변위지도를 획득한 결과를 보인다. 그림 3에서 확인할 수 있듯이 그림 1의 결과에 비해 객체 내부의 세밀한 영역이나 객체의 경계부분의 변위 값이 다른 영역의 변위 값에 비해 정확하게 다름을 알 수 있다.

이 방법을 이용하여 얻게 되는 정보는 개선된 변위 지도가 아닌 좌, 우영상을 기준으로 스테레오 매칭을 수행한 결과 변위 지도에서 나타나는 공통된 잡음 영역이다. 잡음 영역이 검출된 변위 정보를 개선하기 위해 mean-shift를 이용한 분할 영상을 사용한다. 이 과정은 변위 지도의 정확성을 높이기 위해 동일 변위정보를 갖는 평면을 반복적으

로 잡음이 검출된 변위 지도에 맞춰나간다.

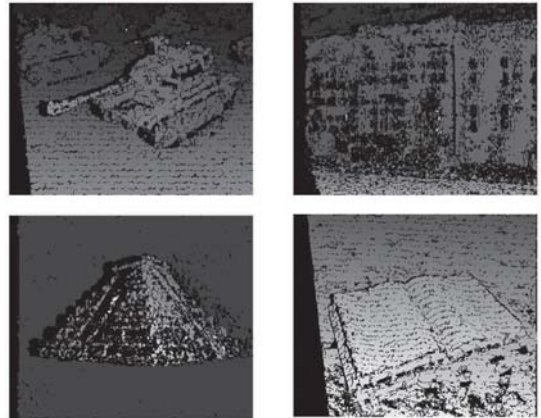


그림 3. 좌, 우영상 매칭 결과를 이용한 잡음 영역 검출

그림 2의 분할된 결과 영상에서 확인할 수 있듯이 객체 내부의 중요하지 않은 영역은 동일한 색상 값을 갖도록 되어있고 객체의 경계 영역은 확연히 다른 색상 값을 가짐을 알 수 있다. 즉 객체의 내부와 외부의 변위정보를 정확히 구별하기 위해 잡음이 검출된 변위 정보와 분할 영상을 매칭 시켜가며 새로운 변위 지도를 생성하게 된다.

분할된 영상에 사용될 매칭 비용 값은 각각 분할된 부분 별로 다르게 계산한다. 비용 값은 각 분할 부분에 속해 있는 화소들이 가지고 있는 비용 값을 총합함으로써 얻을 수 있다. 식 (5)는 이 과정을 나타낸다.:

$$Cost_{segment}(C_p, S) = \sum_{(x,y) \in S} Cost(x,y) \quad (5)$$

식 (5)에서  $C_p$ 는 각 화소가 가지고 있는 비용 값,  $S$ 는 분할된 영상의 각 분할 부분을 나타내며  $Cost_{segments}$ 는 분할영상의 각 부분별 총 비용 값을 의미한다.

##### 2. 중간 값 필터를 이용한 변위지도 개선

잡음이 검출된 변위 지도와 분할된 영상이 서로 매칭 되는 관계를 형성하고, 각 분할된 영역별로 비용 값을 최소가 되게 하는 변위를 선택함으로써 일정한 변위 값을 갖는다. 각 영역별로 일정한 변위 값을 할당받았지만 동일한 영역이라 할지라도 서로 다른 변위 값이 할당되는 경우가 발생 할 수 있다. 이 경우에는 중간 값 필터를 사용하여 해당 영역에 적절한 변위 값을 재 할당해준다.

Mean-shift를 이용해 분할된 영역의 수를 파악한 뒤, 동일 영역이라고 판단되는 영역과 일치하는 변위 값을 찾는다. 일치하는 영역을 찾아 변위 값을 할당하기 위해서는 그림 2의 결과와 그림 3의 결과를 서로 비교하여 정확한 매칭 지점을 판별해야 한다.

이 과정에서 특정 영역으로 분류된 부분들은 단일 변위 값을 갖는 경우도 있지만, 서로 다른 변위 값을 갖는 경우도 발생할 수 있다. 이러한 영역에서는 중간 값 필터를 적용시켜 기준에 할당된 변위 값과 큰 차이가 나지 않는 값을 재 할당해준다. 일반적으로 중간 값 필터는 영상에서 윤곽선 영역을 검출하기 전 존재하는 미세한 잡음을 제거하기 위해 사용된다. 미세 잡음을 제거하기 잡음 영역을 중심으로 커널을 사

용하여 주변의 이상적인 값들과 유사한 값을 갖도록 한다.

중간 값 필터를 분할된 영상에 반복적으로 사용함으로써 분할 영상을 필터를 통해 획득한 값을 이용하여 변위지도를 생성해 나간다. 이러한 과정을 거쳐 획득한 변위지도는 그림 4에 나타나 있다.

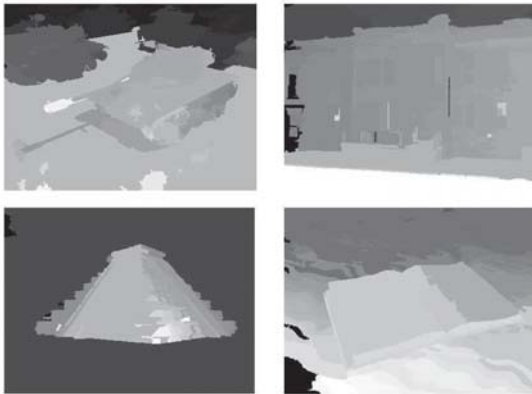


그림 4. 중간 값 필터를 이용한 새로운 변위 지도

중간 값 필터를 이용한 결과 객체 내부에서 유사한 위치에 있는 영역은 동일한 변위 값을 사용하고 객체와 배경 즉 객체의 경계 부분에서는 명확히 구분되는 변위 값을 갖게 됨을 확인 할 수 있다. 그림 3에서 검출된 변위 지도의 오류 값들 또한 효율적으로 제거할 수 있었으며 특히 경계 영역은 변위 값의 번짐이나 뭉개짐 현상 없이 정확한 값을 획득했음을 알 수 있다.

### 5. 실험 결과 및 분석

제안하는 방법의 효율성을 입증하기 위해 Cambridge Computer Laboratory에서 제공하는 400×300 크기의 *Tank, Street, Temple* 그리고 *Book* 실험 영상을 사용했다. 4가지 실험영상을 이용하여 제안하는 방법을 이용해 스테레오 매칭을 수행한 결과 객체 내부 영역에서 비슷한 위치에 있는 영역의 경우 동일한 변위 값을 갖게 되었고 경계 영역에서는 배경과 확실하게 분리될 수 있도록 서로 다른 변위 값을 가짐을 확인할 수 있었다.

표 1. BPR 결과 비교

실험 영상	BPR(%)		
	SAD	SSD	Proposed
Tank	11.3	11.1	8.7
Street	10.4	9.8	7.3
Temple	10.9	10.3	8.1
Book	13.2	13.0	10.1

표 1은 실험에 사용된 4가지 영상을 기존 스테레오 매칭 방법인 sum of absolute differences (SAD)와 sum of squared differences (SSD)를 이용한 결과와 제안한 방법을 이용하여 획득한 변위 영상의 Bad pixel rate (BPR)을 검사한 결과를 나타낸다. BPR은 스테레오 매칭을 사용하여 획득한 변위 지도와 원본 깊이 영상의 각 화소를 비교 하였을 때, 각 화소 값의 차이가 1보다 큰 값을 갖는 화소를 BPR영역

으로 판단하는 방법이다.

### 6. 결론

본 논문에서는 기존 윈도우 기반 스테레오 매칭을 수행할 때 발생하는 경계영역의 깊이정보 부정확성과 텍스처가 부족한 영역에서의 오류를 보완하기 위해 mean-shift방법을 이용한 분할 영상 기반 스테레오 매칭 방법을 제안한다. 분할된 영상과 좌, 우영상 기반 스테레오 매칭 결과를 이용해 획득한 변위 지도와 매칭을 통해하여 각 영역에 사용할 변위 값을 정한다. 각 영역에 미리 정해진 변위 값에 중간 값 필터를 사용하여 정확한 변위 값을 찾아 분할된 영역에 재 할당한다. 제안한 방법을 사용하여 실험한 결과, 기존의 스테레오 매칭 결과보다 약 3% BPR 성능이 향상된 것을 확인했다.

### 감사의 글

이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NO. 2011-0030079)

### 참고문헌

- [1] M. Bleyer and M. Gelautz, "A layered stereo matching algorithm using image segmentation and global visibility constraint," ICIP, Vol. 5, pp. 2997-3000, Oct. 2004.
- [2] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithm," IJCV, pp. 131-140, Dec. 2001.
- [3] O. Veksler, "Stereo correspondence by dynamic programming on a tree," CVPR, Vol.2, pp.384-390, June 2005.
- [4] Y. Deng, Q. Yang, X. Lin, and X. Tang, "A symmetric patch-based correspondence model for occlusion handling," ICCV, Vol. 2, pp. 1316-1322, Oct. 2005.
- [5] A. Bobick and S. Intille, "Large occlusion stereo," Journal of Computer Vision, Vol. 33, no. 3, pp. 181-200, May 1999.
- [6] Q. Yang, L. Wang, and N. Ahuja, "A constant-space beliefpropagation algorithm for stereo matching," CVPR, 2010, pp. 1458-1465, June 2010.
- [7] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," PAMI, Vol. 23, no. 11, pp. 1222-1239, Sep. 1999.
- [8] K.J Yoon and I.S. Kweon, "Adaptive support weight approach for correspondence search," PAMI, Vol. 28, no. 4, pp. 650-656, April, 2006.
- [9] L. Hong and G. Chen, "Segment-Based stereo matching using graph cuts," CVPR, Vol. 1, pp. I: 74-81, July 2004.
- [10] D. Comaniciu and P. Meer. "Mean shift: A robust approach toward feature space analysis," IEEE:PAMI, Vol. 24, no. 5, pp. 603-619, May 2002.