

SAD 정합 알고리즘 수정을 통한 지역기반 스테레오정합의 복잡도 감소 기법

장용준, 호요성
 광주과학기술원
 {yjchang, hoyo}@gist.ac.kr

Complexity Reduction for Local Stereo Matching Method Using Modified SAD algorithm

Yong-Jun Chang, Yo-Sung Ho
 Gwangju Institute of Science and Technology (GIST)

요 약

기존의 지역기반 스테레오 정합 방법은 정합에 사용하는 정합창 크기에 따라 다양한 결과를 갖게 된다. 특히 사용하는 정합창의 크기가 커질수록 영상의 잡음이 강인하지만 객체의 경계부분이 모호해지는 단점을 갖고 있다. 본 논문은 고정된 크기의 정합창을 사용하는 지역기반 스테레오 정합 방법과 다른 방법을 제안한다. 제안한 방법은 영상의 경계를 검출하는 알고리즘을 이용하여 경계부분에는 작은 크기의 정합창을 사용한 변이값을 적용하고 경계가 없는 부분은 큰 크기의 정합창을 사용하여 얻은 변이값을 적용하도록 하였다. 경계를 검출하는 과정에서 본 논문은 AND 연산을 사용하여 최대한 객체의 테두리만을 나타내는 경계값을 구하도록 하였다. 또한 두 가지 크기의 정합창을 이용함으로써 발생하는 복잡도 증가를 감소시키기 위해 기존의 SAD 연산 알고리즘을 수정하여 복잡도를 감소시켰다. 본 논문에서 사용한 정합창의 크기는 5×5 와 15×15 이며 실험결과 제안한 방법은 15×15 정합창을 사용한 결과와 비교하여 변이지도에서 객체의 경계부분은 더 잘 살리면서 수행시간을 줄여 효율적인 정합결과를 얻어냈다.

1. 서론

최근 몇 년간 3차원 엔터테인먼트 분야는 많은 관심을 받아왔고 현재까지도 지속적인 연구가 이루어지고 있다. 그 중 대중적으로 많은 관심을 받고 있는 분야 중 하나가 바로 3차원 영상 분야라고 할 수 있다. 최근 3차원 영화 흥행에 힘입어 3차원 영화 산업이 발달하게 되고 이에 따라 3차원 영화를 제작하는 일들이 크게 늘어나게 되었다. 3차원 영상을 만들기 위해서는 한 장면에 대해 여러 시점에서 촬영한 영상이 필요하다. 여러 시점의 영상을 얻기 위한 방법으로 여러대의 카메라를 이용한 촬영 방법도 있지만 기존의 촬영된 영상들을 이용하여 중간시점 영상을 합성하는 방법도 있다. 중간시점 영상 합성을 위해서는 촬영한 영상의 깊이 정보가 필요한데 깊이정보 획득 방법은 크게 능동적 깊이측정 방법과 수동적 깊이측정 방법 두 가지로 나뉜다. 능동적 깊이측정 방법은 TOF (Time of Flight) 카메라를 이용하여 객체의 깊이를 직접 측정하는 방법이며 수동적 깊이 측정방법은 한 객체에 대해 다양한 시점에서 촬영한 영상을 이용하여 각 시점에 따른 변이차를 계산하고 깊이를 측정하는 방법이다.

각각의 깊이정보 획득 방법은 몇 가지의 장점과 단점들을 지니고 있다. 먼저 능동적 깊이측정 방법의 경우 적외선 및 레이저를 이용하여 객체의 깊이를 직접적으로 측정하므로 비교적 정확한 깊이값을 측정할 수 있지만 레이저의 측정 범위를 넘어선 객체가 존재하는 실외공간에서는 능동적 깊이측정 방법을 사용하는 것은 한계가 있다. 반면에 수동적 깊이측정 방법은 이

미 촬영된 다시점 영상에 대해 시점에 따른 객체간의 변이차를 이용하여 깊이를 측정하기 때문에 공간의 제약을 받지 않는다는 장점이 있다. 하지만 객체의 패턴이나 텍스처가 없는 부분에 대해서는 정확한 변이값을 측정하지 못한다. 이러한 수동적 깊이측정 방법 중 하나가 스테레오 정합을 이용한 깊이측정 방법이다. 일반적으로 스테레오 정합 방법은 지역기반 스테레오 정합 방법과 전역기반 스테레오 정합 방법 두 가지로 나뉜다.

지역기반 스테레오 정합 방법은 좌측시점에서 촬영한 좌영상과 우측시점에서 촬영한 우영상, 이 두 가지 영상의 화소간 정합비용을 계산하여 깊이값을 측정하는 방법을 뜻한다. 정합비용은 SAD (Sum of Absolute Differences), SSD (Sum of Squared Differences), NCC (Normalized Cross Correlation) 등과 같은 유사도 측정 방법을 이용하여 계산한다. 이들 중 SAD와 SSD를 이용한 정합방법은 최소 정합비용을 갖도록 하는 정합위치를 찾는 것이고 NCC는 두 화소간의 유사도를 측정하여 정합위치를 찾는 방법으로 최대 정합비용을 갖도록 하는 정합위치를 찾는다. 지역기반 스테레오 정합은 전역기반 스테레오 정합에 비해 정합속도가 빠르다는 장점이 있지만 영상 내부에서 수직방향의 패턴을 갖는 영역 또는 텍스처가 없는 영역에 대해서는 정합률이 떨어지는 단점이 있다.

전역기반 스테레오 정합 방법은 위에서 설명한 지역기반 스테레오 정합의 단점을 어느 정도 보완한다. 이 방법은 MRF (Markov Random Field) 기반의 에너지 함수를 이용하여 최적의 변이값을 추정한다. 이때 변이값은 에너지 함수를 최소화 시키는 값으로 결정이 된다. 에너지 최적화를 위한 방법으로는

BP (Belief Propagation), GC (Graph Cuts) 등의 방법이 있다 [1, 2]. 최적화 과정에서 전역기반 스테레오 정합 방법은 비용을 발생시키는데 이로 인해 정합과정에서의 전체적인 복잡도가 증가하게 되는 단점이 있다.

이처럼 두 가지의 스테레오 정합 방법은 각각 장점과 단점을 지니고 있으며 단점들을 보완하기 위해 많은 연구가 진행 중이다. 본 논문 역시 기존의 지역기반 스테레오 정합의 단점인 정합창 크기 증가에 따른 경계 모호성 문제 해결을 위해 객체의 경계유무에 따라 두 가지의 정합창을 이용하여 얻은 변이값을 선택적으로 적용하도록 하였다. 또한 변이값을 계산하는 과정에서 발생하는 복잡도 증가를 줄이기 위해 SAD 연산 알고리즘을 수정하였다.

2. 정합창을 이용한 스테레오정합

지역기반 스테레오 정합 방법을 사용할 때 화소값 하나 하나에 대해서 정합비용을 찾는 것은 정합과정에서의 모호성을 야기한다. 현재 화소값과 매우 유사한 값을 지닌 화소가 하나가 아닌 여러 위치에서 존재할 수 있기 때문이다. 이러한 이유로 정합과정의 모호성을 줄이기 위해 블록 단위의 정합창을 사용하여 지역기반 스테레오 정합을 한다. 블록 단위의 정합창은 정합비용을 계산할 때 현재 위치에서의 화소값만 이용하는 것이 아닌 주변의 다른 화소들 또한 정합비용 계산에 사용하여 지역기반 스테레오 정합의 정확도를 높여준다. 정합창의 크기는 $3 \times 3, 5 \times 5$ 등 사용자의 필요에 따라 결정이 된다.

그림 1을 보면 좌영상에 있는 정합창은 기준이 되는 참조 정합창이 되고 우영상에 있는 정합창은 비교대상이 되는 비교 정합창이 된다. 동일한 탐색선 상에서 참조 정합창은 (x, y) 위치에 고정되어있으며 비교 정합창은 탐색범위 내에서 정합비용을 계산하며 이동하게 된다. 만약 비교 정합창이 $(x-i, y)$ 위치에서 최적의 정합비용을 갖는다면 좌영상의 (x, y) 에 위치한 화소는 i 라는 변이값을 갖게 된다.

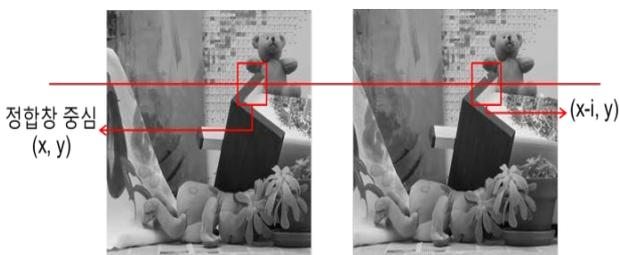


그림 1. 정합창을 이용한 스테레오정합

최적의 정합비용을 얻어내기 위한 방법으로 유사도 측정방법을 이용하는데 유사도 측정방법들 중 하나로 SAD정합 알고리즘을 예로 들 수 있다. SAD의 수식은 식 (1)과 같다.

$$SAD(d) = \sum_i \sum_j |I_L(x + d + i, y + j) - I_R(x + i, y + j)| \quad (1)$$

식 (1)에서 d 는 좌영상과 우영상 사이의 변이값을 나타내고 x 와 y 는 영상 내 임의의 좌표값이다. i 와 j 는 탐색범위 그리고 I_L 과 I_R 각각은 좌영상과 우영상을 표현한 것이다.

3. 정합 창 크기에 따른 변이값 추정

지역기반 스테레오 정합은 사용하는 정합창의 크기에 따라 다양한 결과값을 갖는다. 작은 크기의 정합창 사용은 다양한 패턴과 복잡한 텍스처 등 영상 내 비교적 세밀한 부분의 변이값을 구하는데 적합하나 영상이 지니고 있는 잡음에는 약하다는 단점이 있다. 반대로 큰 크기의 정합창을 사용할 경우 영상내부의 잡음에는 강한 반면 객체의 경계와 복잡한 패턴을 지닌 영상에 대해서는 작은 크기의 정합창을 사용할 때 만큼 세밀하게 표현해내지 못한다는 단점이 있다. 그림 2는 $5 \times 5, 7 \times 7, 11 \times 11, 15 \times 15$ 크기의 정합창을 이용하여 변이지도를 얻은 결과이다.

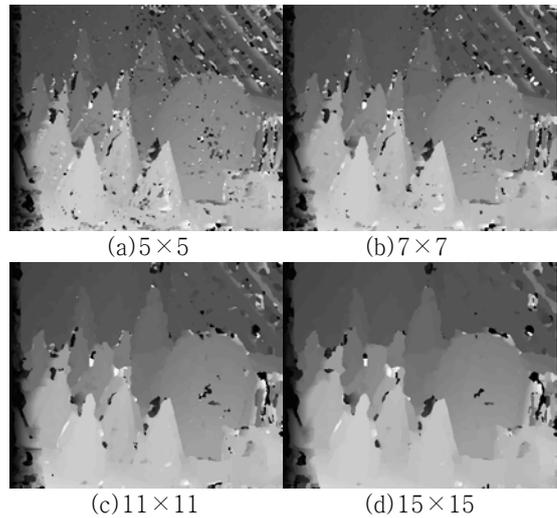


그림 2. 정합창 크기에 따른 변이지도

그림 2(a)와 그림 2(b)를 보면 변이지도의 경계부분은 잘 나오는 것을 확인할 수 있으나 잡음이 많이 생기는 것을 알 수 있다. 반면에 그림 2(c)와 그림 2(d)의 결과를 보면 잡음은 거의 사라졌지만 객체의 경계부분이 모호해진 것을 확인할 수 있다. 아래의 표 1과 표 2는 각각 정합창 크기에 따른 정합 오차율 BPR (Bad Pixel Rate)과 스테레오정합에 걸리는 시간을 표로 정리한 것이다.

표 1. 정합창 크기에 따른 BPR

Test Images	Resolution	5×5	7×7	11×11	15×15
		BPR_all (%)			
Cones	450×365	31.53	28.37	26.00	25.93
Teddy	450×375	33.48	31.13	30.37	30.20
Tsukuba	384×288	12.88	11.18	10.04	10.18
Venus	434×383	22.88	17.32	11.16	9.09

표 2. 정합창 크기에 따른 수행시간

Test Images	Resolution	5×5	7×7	11×11	15×15
		Time (sec.)			
Cones	450×375	0.50	0.96	2.40	4.34
Teddy	450×375	0.50	1.00	2.44	4.46
Tsukuba	384×288	0.09	0.18	0.46	0.78
Venus	434×383	0.19	0.37	0.78	1.45

표 1과 표 2에서 정합크기가 커질수록 BPR은 감소하지만 정합을 수행하는데 걸리는 시간이 늘어나는 것을 확인할 수 있다. 기존 SAD 정합방식은 정합창의 크기를 N 이라 할 때 N^2 의

복잡도를 갖는다. 따라서 SAD 정합과정의 복잡도를 줄이기 위해 정합창 내부의 모든 화소를 연산에 사용하는 것이 아닌 일부 후보군을 선택하여 SAD 정합연산에 사용하는 방법도 제시되고 있다 [3].

4. 제안한 지역기반 스테레오정합 방법

그림 2와 표 1, 표 2를 보면 5×5 크기의 정합창을 사용하였을 때 BPR 값은 매우 높지만 변이지도 상에서 객체의 경계부분은 잘 살리는 것을 볼 수 있다. 반면에 15×15 크기의 정합창은 BPR 값이 낮아지지만 객체의 경계부분은 잘 살리지 못한다. 이를 개선하기 위해 하나의 정합창만 사용하는 것이 아닌 두 개의 정합창을 사용하여 최종 변이지도의 질을 높이는 방법도 제안되었다 [4]. [4]의 경우 두 가지의 정합창을 이용하여 각각에 대한 변이지도를 만들어 놓은 후 두 변이지도의 변이값 연속성 측정으로 최종 변이지도를 얻도록 하였다. 이때 변이지도의 연속성 측정방법은 현재 화소와 주변 화소간의 유사도 비용을 계산하여 비용이 더 낮은 화소의 변이값을 선택하도록 하였다. [4]의 방법은 최종 변이지도의 질은 향상시키지만 주변과의 연속성만을 고려하여 변이지도를 구하기 때문에 최종 변이지도에서 객체의 경계부분이 명확하지 않다는 단점이 있다. 따라서 이러한 단점을 개선하기 위해 영상의 경계에 따른 적응적 변이값 선택 방법을 제안한다.

제안한 방법은 영상의 경계부분에서는 5×5 크기의 정합창을 사용하고 경계 이외의 부분에서는 15×15 크기의 정합창을 사용하도록 하여 최종 변이지도의 질을 높이도록 하였다. 이때 경계 검출법은 Canny 경계 검출법을 사용하였다. Canny 경계 검출법을 사용한 검출 결과는 그림 3(b)에 나와있다. 그림 3(b)는 원본영상에 대한 경계 검출 결과를 나타낸다. 그림 3(b)를 보면 물체의 경계 내부에도 패턴과 텍스처에 의한 경계선들이 그려져있는 것을 볼 수 있다. 이는 큰 크기의 정합창 보다 상대적으로 잡음을 많이 발생시키는 5×5 정합창의 사용 횟수를 증가시킨다. 따라서 최종 결과영상에 잡음이 많이 생기게 되는 결과를 낳을 수 있으므로 최대한 물체의 경계부분만 작은 크기의 정합창이 적용되도록 물체 내부의 경계들을 없애주는 작업이 필요하다.

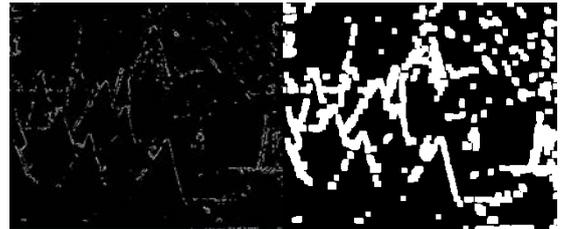


(a) 원본영상 (b) Canny 경계 검출 결과
그림 3. Canny 경계 검출 결과

물체의 경계부분만 추출하기 위해 본 논문에서는 기존의 5×5 정합창으로 만든 변이지도를 이용하였다. 5×5 정합창을 이용하여 얻은 변이지도에 Canny 경계 검출 알고리즘을 적용하여 경계를 검출한다. 그 후 원본영상에서 검출한 경계영상과 변이지도에서 얻은 경계영상을 이용하여 식 (2)와 같은 AND 연산을 적용한다.

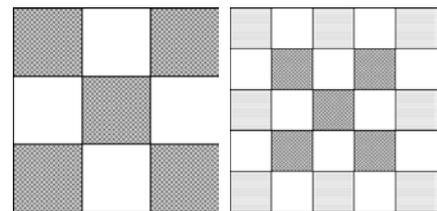
$$I_0(x,y) \cap I_D(x,y) = \begin{cases} 255, (I_0(x,y) = I_D(x,y)) \\ 0, (I_0(x,y) \neq I_D(x,y)) \end{cases} \quad (2)$$

식 (2)에서 $I_0(x,y)$ 는 원본영상의 Canny 경계검출 결과영상이고 $I_D(x,y)$ 는 변이지도의 Canny 경계검출 결과영상이다. 따라서 같은 좌표에서 두 영상의 화소값이 같으면 새로운 영상의 같은 위치에 255의 화소값을 넣어 주고 두 영상의 화소값이 서로 다르면 새로운 영상의 같은 위치에 0의 화소값을 넣어준다. AND 연산 결과는 그림 4(a)와 같다. AND 연산 후에 얻은 결과영상은 그림 4(b)처럼 경계부분을 팽창시켜 물체의 경계가 더 잘 보이도록 하였다.



(a) AND 연산 결과 (b) 팽창연산 적용
그림 4. AND 연산 후 결과영상

본 논문에서는 경계의 유무에 따라 선택적으로 변이값을 적용하도록 제안하였다. 따라서 사전에 5×5와 15×15의 정합창을 이용하여 변이지도를 먼저 구하게 되는데 이 과정에서 연산의 복잡도가 증가한다. 복잡도 감소를 위해 그림 5와 같이 SAD 연산 알고리즘을 수정하였다.



(a) 3×3 (b) 5×5
그림 5. SAD 연산 후보군

3×3 정합창을 기반으로 하여 그림 5(a)와 같이 총 5개의 연산 후보군을 정하고 3×3 이상의 정합창에 대해서는 5개의 3×3 연산 후보군을 포함한 최외곽의 화소들을 그림 5(b)처럼 하나씩 건너뛰어 선택하도록 한다 [4]. 따라서 정합창의 크기가 N 일 경우 연산의 복잡도는 $\frac{N^2+1}{2}$ 으로 기존의 N^2 복잡도 보다 줄어들게 된다.

마지막으로 결과영상에서 남아있을 수 있는 불연속 지점들을 보완하기 위해 수평방향과 수직방향으로 변이값의 연속성을 확인하여 잡음을 제거하도록 하였다. 변이값의 연속성을 확인하는 식은 아래 식 (3)과 같다.

$$\begin{aligned} Cost_{H1}(x,y) &= abs(D(x-1,y) - D(x,y)) \\ Cost_{H2}(x,y) &= abs(D(x+1,y) - D(x,y)) \\ Cost_{V1}(x,y) &= abs(D(x,y-1) - D(x,y)) \\ Cost_{V2}(x,y) &= abs(D(x,y+1) - D(x,y)) \end{aligned} \quad (3)$$

표 3. 결과값 비교

Test Images	Resolution	5×5			15×15			Proposed Algorithm		
		Time (sec.)	BPR_all (%)	BPR_disc. (%)	Time (sec.)	BPR_all (%)	BPR_disc. (%)	Time (sec)	BPR_all (%)	BPR_disc. (%)
Cones	450×375	0.50	31.53	33.52	4.34	25.93	39.90	2.70	26.19	35.19
Teddy	450×375	0.50	33.48	35.23	4.46	30.20	40.72	2.63	29.61	37.24
Tsukuba	384×288	0.09	12.88	20.86	0.78	10.18	34.61	0.48	10.19	29.42
Venus	434×383	0.19	22.88	28.61	1.45	9.09	40.63	0.89	11.04	36.00
Average		0.32	25.19	29.56	2.76	18.85	38.97	1.67	19.26	34.46

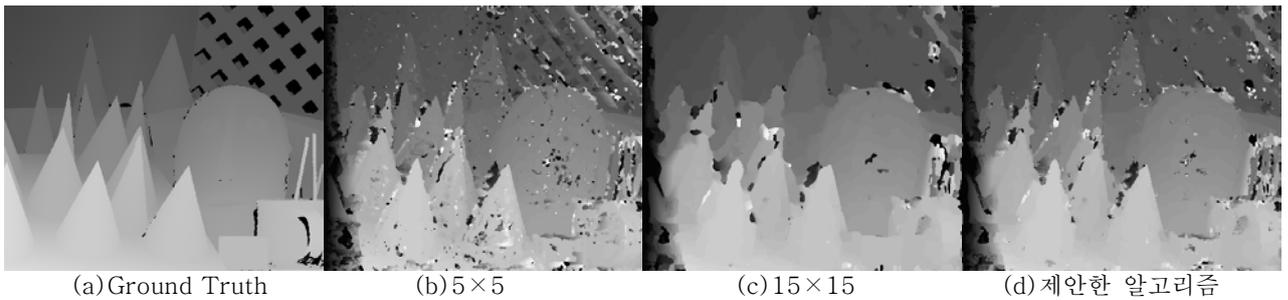


그림 6. 최종 실험결과 영상

식 (3)에서 $Cost_{H1}$ 과 $Cost_{H2}$ 는 수평방향에서 현재 화소의 위치 $D(x, y)$ 를 기준으로 좌우 방향의 연속성을 확인한 비용이다. 그리고 $Cost_{V1}$ 과 $Cost_{V2}$ 는 수직방향에서 현재 화소의 위치를 기준으로 위, 아래의 연속성을 확인한 비용을 나타낸다. 변이지도에서 같은 객체 내부의 화소들은 비슷한 값을 지닐 확률이 높다. 따라서 비용이 크다는 것은 현재 화소와 이웃하는 화소 사이의 변이값 차이가 크다는 것을 의미한다. 이러한 특성을 이용하여 연속성 비용이 임계값을 넘으면 영상 잡음이라고 판단하고 이러한 경우 현재의 화소값을 주변의 화소값과 같은 값으로 대체하도록 한다.

5. 실험결과

제안한 알고리즘을 실험하기 위해 우리는 *Cones*, *Teddy*, *Tsukuba*, *Venus* 총 4개의 영상을 사용하였다. 그림 6(b)와 그림 6(d)를 비교해보면 제안한 알고리즘의 영상에서 잡음이 많이 없어진 것을 확인할 수 있다. 또한 그림 6(c)와 그림 6(d)를 비교해보면 제안한 알고리즘의 영상에서 물체의 경계 주변과 내부에 잡음이 조금 존재하지만 15×15 정합창을 사용한 결과영상보다 물체의 경계를 더 잘 표현했다는 것을 확인할 수 있다. 표 3은 최종 실험결과이며 전체 BPR과 불연속 지점의 BPR값 그리고 전체 수행시간을 각 영상과 사용한 정합창의 크기에 따라 정리하여 나타내었다.

6. 결론

기존의 지역기반 스테레오 정합은 사용하는 정합창의 크기 또는 정합창을 적용하는 방법에 따라서 연산의 복잡도가 증가할 수 있다. 따라서 본 논문은 개선된 변이지도를 얻기 위해 영상의 경계를 기반으로 한 적응적 변이값을 사용하였고 또한 이때 발생하는 연산의 복잡도 증가를 감소 시키기 위해 SAD 정

합 알고리즘을 수정하였다. 실험결과 기존의 5×5 SAD 정합결과보다 제안한 알고리즘의 전체 BPR이 5.93% 감소하였으며 15×15 SAD 정합결과와 비교하여 제안한 알고리즘은 0.41%의 전체 BPR 증가와 약 39.49%의 복잡도 감소를 보였다. 또한 불연속성 관점에서 본 BPR의 경우 오히려 제안한 알고리즘이 15×15 정합창을 사용한 결과값 보다 4.51% 감소하여 물체의 경계부분이 기존의 15×15 정합보다 더 잘 나왔다는 것을 확인할 수 있다.

감사의 글

이 논문은 2014년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2011-0030079)

참고 문헌

- [1] J. Sun, N. N. Zheng, and H. Y. Shum, "Stereo Matching Using Belief Propagation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787-800, July 2003.
- [2] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, Nov. 2001.
- [3] J. Yi and J. S. Kim, "Heuristic Designs of SAD Correlation Algorithm for Vision System," *The Institute of Electronics and Information Engineers Trans. On Computer and Information* vol. 47, no. 5, pp. 61-66, Sep. 2010.
- [4] Y.J. Chang and Y.S. Ho, "Modified SAD Using Adaptive Window Sizes for Efficient Stereo Matching," *International Conference on Embedded Systems and Intelligent Technology*, pp. 9-11, Sep. 2014.