

영상에서의 분할정보를 사용한 스테레오 조밀 시차맵 생성

이 범 종[†] · 박 종 승^{**} · 김 정 규^{***}

요 약

스테레오 비전은 시차가 있는 양안 영상으로부터 깊이맵에 해당하는 시차맵을 생성하고 시차맵과 카메라 정보로부터 3차원 구조를 복원하는 기법이다. 시차맵 생성은 정합비용을 계산하고, 전체 정합 비용을 최소화하여 시차를 계산하는 단계로 이루어진다. 본 논문에서는 스테레오 영상으로부터 빠르고 안정된 시차맵을 생성하기 위해서 후처리 과정으로 각 스캔라인에 대해서 분산을 이용하여 분할한 후에 분할 영역 정보를 사용하여 객체간 영역을 구분할 수 있도록 한다. 시차맵의 계산 시에 영역의 균일성 정보를 사용하면 잘못된 매치가 발생되지 않도록 억제할 수 있다. 조밀 시차맵을 생성하기 위해서는 시차 계산에 실패한 픽셀들에 대해서도 인접 픽셀의 값을 사용한 보간 기법을 통한 홀 메우기로 시차값을 계산하여 조밀한 시차맵이 형성되도록 한다. 실제 환경에서의 다양한 스테레오 영상에 대한 실험결과를 제안된 시차맵 생성과 홀 메우는 방법이 기존의 방법보다 안정적이고 다양한 응용에 적용될 수 있음을 보여준다.

키워드 : 조밀 시차맵, 스테레오 비전, 홀 메우기

Computation of Stereo Dense Disparity Maps Using Region Segmentation

Bum-Jong Lee[†] · Jong-Seung Park^{**} · Chung-Kyue Kim^{***}

ABSTRACT

Stereo vision is a fundamental method for measuring 3D structures by observing them from two cameras placed on different positions. In order to reconstruct 3D structures, it is necessary to create a disparity map from a pair of stereo images. To create a disparity map we compute the matching cost for each point correspondence and compute the disparity that minimizes the sum of the whole matching costs. In this paper, we propose a method to estimate a dense disparity map using region segmentation. We segment each scanline using region homogeneity properties. Using the segmented regions, we prohibit false matches in the stereo matching process. Disparities for pixels that failed in matching are filled by interpolating neighborhood disparities. We applied the proposed method to various stereo images of real environments. Experimental results showed that the proposed method is stable and potentially viable in practical applications.

Keywords : Dense Disparity Map, Stereo Vision, Hole-Filling

1. 서 론

최근 실감나는 영상의 구현으로 가상환경에서 실제처럼 느낄 수 있도록 하는 3차원 그래픽 기술, 가상현실에서의 사람이 공간적으로 상호작용을 할 수 있는 체감 상호작용 기술, 실 환경과 가상 환경의 혼합 환경을 구현하는 증강현실 기술, 양안 영상으로부터의 입체 데이터 획득을 위한 스테레오 비전 기술 등에 대한 연구가 활발히 진행되고 있다. 특히 3차원 영상기술이 발전함에 따라 3차원 응용분야에 대한 관심이 고조되고 있다. 3차원 영상 응용분야를 구현하기

위해서는 고전적인 스테레오 비전 기술에 대한 이해가 필요하다. 스테레오 비전은 컴퓨터 비전에서 가장 광범위하게 연구되고 있는 분야 중에 하나이며, 같은 장면의 시차가 다른 두 영상에서 정합되는 점들을 찾아 이 점들의 차이를 표현하는 시차맵을 생성하는 것을 목표로 하고 있다. 스테레오 비전에서 한 장면의 두 영상에서 대응되는 점을 찾는 대응관계 문제(correspondence problem)는 스테레오 계산에서 가장 중요한 단계에 속한다. 한 장면에 대해서 두 개 이상의 스테레오 영상이 주어진다면, 대응관계는 그 점이 각 영상 평면의 점들 중에서 어디에 투영되었는지를 찾아내는 것이다. 다시 말해서, 참조 영상의 점이 다른 영상들에서 어느 부분에 나타나는지를 찾아내는 것을 말한다.

대응관계 문제를 풀기 위한 방법으로는 영역 기반(region-based) 방법과 특징 기반(feature-based) 방법으로 나눌 수 있다. 영역 기반의 스테레오 기법은 한 영상에서 한 픽셀의 이웃하는 픽셀들의 밝기 패턴과 다른 영상에서 대응되는 픽

※ 본 연구는 지식경제부 지방기술혁신사업(RTI05-03-01)의 지원으로 수행되었음.

† 준 회 원 : 인천대학교 컴퓨터공학과 석사

** 중 심 회 원 : 인천대학교 컴퓨터공학과 조교수(교신저자)

*** 정 회 원 : 인천대학교 컴퓨터공학과 교수

논문접수 : 2007년 9월 4일

수정일 : 1차 2008년 8월 19일

심사완료 : 2008년 9월 5일

셀의 이웃하는 픽셀들의 밝기 패턴 사이의 연관성을 사용한다. 즉 한 영상에서의 각 픽셀에 대해 다른 영상에서 정합되는 위치의 이웃 픽셀들의 밝기의 유사도로 일치점을 찾아내는 것이다. 영역 기반의 기법은 각 픽셀의 명암값을 바로 사용하기 때문에 한 장면에 대한 여러 카메라의 위치에 따라 명암값이 변하는 것에서 나타나는 왜곡 등에 매우 취약하다는 단점이 있다. 다른 물체와의 가려짐(occlusion)으로 인한 깊이 정보의 오류가 나타난다는 단점도 있다. 매칭에 있어서의 탐색 범위의 설정과 윈도우의 크기 설정은 계산 시간뿐만 아니라 정확도에도 큰 영향을 미친다. Koo 등은 보다 안정된 대응관계 문제의 해결을 위해 적응적인 탐색 범위와 윈도우의 크기를 사용함으로써 정확도를 높이는 방법을 제안하였다[1]. Stefano 등은 좀 더 신뢰할 수 있는 정합을 찾으려면 이전의 정합을 버리는 방식을 제안하였다[2]. 유일성 제약으로부터 신뢰할 수 없는 매치들을 탐지하며 이러한 매치들은 보다 더 나은 매치들로 대체된다. 속도의 향상을 위해서 일반적인 양방향 매칭 기법 대신에 단방향 매칭 기법을 선택하였다.

특징 기반의 스테레오 기법은 명암 영상으로부터 특징들을 획득하고 각 특징들을 매칭하는 기법이다. 이 기법은 주변의 빛의 변화에 더 안정적이다. 사용되는 특징으로는 모서리 점들을 연결하여 만든 모서리 세그먼트가 흔히 사용된다. 특징 기반의 스테레오 기법은 정합되는 특징의 속성만 비교하기 때문에 영역 기반에 비해서 속도가 빠르다는 장점이 있다. Taylor는 특징 기반의 스테레오 매칭 결과로부터 물체의 표면을 복원하는 방법을 제시하였다[3]. 아주 작은 영역에 대한 매칭 결과가 주어지더라도 특징들의 매칭 결과가 정확하다면 전체 표면을 재구성할 수 있음을 보여준다. 특징 기반 스테레오 매칭에 있어서 특징점의 선정은 매칭의 정확성을 결정하는 매우 중요한 과정이다. Lew 등은 학습을 통한 최적의 특징점을 선택하는 방법을 제안하였다[4]. 대응되는 영상에서 특징점의 모션과 명암의 불연속에 민감한 문제를 해결하는 방법도 있다[5]. 이 방법은 영역 기반 분석 모듈과 에지 기반 스테레오 분석 모듈 등의 여러 분석 모듈들을 혼합하여서 계층적인 장면의 추론이 가능하도록 하였다. 두 에피폴 라인의 1차원적인 스캔라인 매칭에 있어서도 특징 기반 스테레오 매칭 접근이 연구되었다. McKeown 등은 1차원 파형의 정합을 위한 계층적인 매칭 알고리즘을 제안하였다[6].

기존의 스테레오 정합 방법들은 각각 특정 응용에 따른 장단점을 가지고 있다. 본 논문에서의 연구의 목표는 가상 현실과 같은 시각적인 표현이 중요한 응용에서 안정적으로 사용될 수 있는 시차맵 생성 기법의 개발이다. 본 논문에서 제안하는 방법은 두 장의 스테레오 영상을 입력으로 하여 조밀 시차맵을 생성한다. 두 입력 스테레오 영상에 대해서 SSD 방법을 사용하여 정합비용을 계산하고 비용을 집성하여 시차맵을 계산하고, 분산 기법을 이용하여 스캔라인 별로 분할한 후에 분할 구간별로 평균을 적용하여 객체간의 좀더 명확한 구분이 되도록 하여 영역 내에서의 일관성 있는 매칭이 시도되도록 한다. 매칭이 실패한 영역들에 대해

서는 후처리 작업으로 홀 메우기를 실시한다. 홀 메우기에서는 인접 영역들의 색상정보를 보간하여 홀의 각 픽셀의 색상값을 결정한다.

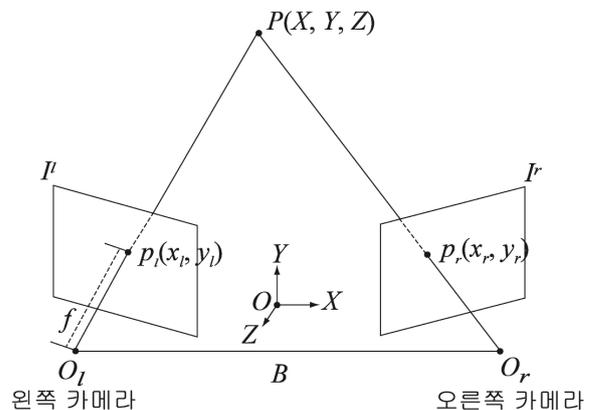
2. 관련 연구

본 절에서는 스테레오 매칭을 통한 시차맵 생성의 관련 연구들을 설명하고 장단점 및 문제점을 기술한다. (그림 1)은 스테레오 비전에 대한 카메라 설정을 보여주는 그림이다. 두 장의 스테레오 입력 영상에 대해서 왼쪽의 영상을 I^l 이라 하고, 오른쪽의 영상을 I^r 이라 하고, 두 입력 영상에 대한 카메라의 중심(optical center)은 각각 O_l 과 O_r 이라 하자. P 는 장면의 3D점을 의미하며, 그 점이 두 스테레오 영상에 각각 p_l 과 p_r 로 투영된다. O 는 평면직각좌표계(Cartesian coordinate system)에서의 원점이고, f 는 초점 거리(focal length), B 는 두 카메라 중심 사이의 거리인 베이스라인 길이(baseline distance)를 의미한다.

스테레오 비전 기술은 스테레오 영상에 대해서 3차원 정보인 시차맵을 계산하는데 매우 유용하며 다음과 같은 세 단계 절차에 의해 수행된다[7].

- (1) 정합 비용 계산
- (2) 비용 집성
- (3) 시차 계산 및 최적화

왼쪽 영상에서의 픽셀 (x,y) 이 오른쪽 영상에서는 (x',y') 의 위치에서 보여진다면 그 변이의 차이인 $(x'-x,y'-y)$ 를 시차(disparity)라고 하고 (d_x, d_y) 로 표시한다. 정확한 시차를 계산하는 것이 스테레오 매칭의 목표이다. 첫 번째 단계인 정합 비용(matching cost) 계산 단계는 좌우의 영상에서 대응되는 부분을 찾기 위해 픽셀 간의 유사도를 계산한다. 왼쪽 영상에서의 한 픽셀 (x,y) 에서의 시차가 (d_x, d_y) 라면 그 픽셀은 오른쪽 영상에서 $(x+d_x, y+d_y)$ 의 위치에 나타나야 한다. 시차가 얼마나 정확한지에 대한 측도는 왼쪽 영상의 픽셀 (x,y) 의 밝기와 오른쪽 영상의 픽셀 $(x+d_x,$



(그림 1) 스테레오의 기하 구조 및 카메라 설정 상태

$y+d_y$)의 밝기의 차이로 정의할 수 있다. 픽셀 기반의 정합 비용을 계산하는 가장 대표적인 방법으로는 식 (1)과 같이 좌우 영상의 명암값의 차이의 제곱으로 비용을 계산하는 명암 차의 제곱(SD; Squared Intensity Differences)[8]과 식 (2)와 같이 좌우 영상의 명암값의 차이의 절대값으로 비용을 계산하는 명암 차의 절대값(AD; Absolute Intensity Differences)[9]이 있다.

$$(I^l(x,y) - I^r(x+d_x, y+d_y))^2 \tag{1}$$

$$|I^l(x,y) - I^r(x+d_x, y+d_y)| \tag{2}$$

한 픽셀에서의 시차는 그 픽셀을 중심으로 하는 영상 블록을 고려하여 계산한다. 영상 블록을 윈도우라고 하며 윈도우 내의 각 픽셀들의 오차의 총 합을 중심 픽셀에서의 정합 비용으로 간주한다. 블록 기반의 정합 비용을 계산하는 대표적인 방법으로는 식 (3)과 같이 계산되는 정규화된 교차상관(NCC; Normalized Cross Correlation)[10]과 식 (4)와 같이 계산되는 차의 제곱 합(SSD; Sum of Squared Differences)과 식 (5)에 의해 계산되는 차의 절대값 합(SAD; Sum of Absolute Differences)이 있다. 정합 비용 단계에서는 위의 방법들을 사용하여 모든 픽셀과 일정 범위의 시차에 걸쳐서 시차공간 영상(DSI; Disparity Space Image)을 만들게 된다.

$$\frac{\sum_{x,y} (I^l(x,y) - \bar{I}^l) \cdot (I^r(x+d_x, y+d_y) - \bar{I}^r)}{\sqrt{\sum_{x,y} (I^l(x,y) - \bar{I}^l)^2 \cdot \sum_{x,y} (I^r(x+d_x, y+d_y) - \bar{I}^r)^2}} \tag{3}$$

$$\sum_{x,y} (I^l(x,y) - I^r(x+d_x, y+d_y))^2 \tag{4}$$

$$\sum_{x,y} |I^l(x,y) - I^r(x+d_x, y+d_y)| \tag{5}$$

두 번째 단계인 비용 집계(cost aggregation) 단계에서 공간과 윈도우 기반의 방법은 시차공간 영상에서 올바른 영역에 걸쳐 합계나 평균의 방법으로 정합 비용을 집계한다. 올바른 영역은 고정된 시차에서 2차원 공간에서의 영역이거나 또는 3차원 공간에서의 영역이 될 수 있다. 2차원 공간의 경우는 정사각 윈도우, 이동할 수 있는 윈도우[11], 적응적인 크기의 윈도우[12] 등과 같은 복합 윈도우를 사용하여 집계한다. 3차원 공간의 경우는 제한된 시차의 차이[13]를 이용하거나, Prazdny의 연관성 원리[14]를 사용하여 집계한다.

고정된 영역으로 집계하는 것은 직사각형의 윈도우의 경우에는 2차원이나 3차원 회선(convolution)을 사용하거나, 효율적인 박스필터를 사용하여 수행될 수 있다. 이동할 수 있는 윈도우의 경우에는 최소필터를 사용하여 효율적으로 구현될 수 있다. 반복 확산(iterative diffusion)과 같은 집계 방법은 각 픽셀의 비용에 이웃하는 픽셀의 비용들의 가중치 값을 반복적으로 더함으로써 구현된다[15].

마지막 단계인 시차(disparity)를 계산하고 최적화하는 단계는 공간 방법, 전역 최적화, 동적 계획법으로 나눌 수 있다. 공간 방법에서는 정합 비용계산 단계와 집성 단계를 강조하며, 시차 계산은 간단하게 각 픽셀에서 최소의 비용을 갖는 값과 관련한 시차를 선택함으로써 이루어진다. 공간 방법은 각 픽셀에서 WTA(Winner-Take-All) 기법을 적용하여 수행한다. 공간 방법의 단점은 실제 참조 영상에 대해 단지 하나의 정합만을 허용하지만, 다른 영상에서 점 은 여러 점과 정합될 수 있다는 것이다.

전역 최적화는 거의 대부분 시차를 계산하는 단계에서 집계하는 단계까지 수행하거나 집계 단계를 생략한다. 많은 전역 최적화는 에너지 기반의 방법[16]으로 공식화되며, 에너지 기반의 방법은 다소 복잡하긴 하지만 전체의 정합 결과가 에너지 식을 최소화하는 시차를 구하기 때문에 가장 좋은 성능을 보인다[17]. 최근, 전역 최적화 문제를 특별한 방법으로 풀어내기 위한 방법들이 제안되었다[18].

전역 최적화의 또 다른 방법은 동적 계획법에 기반하는 방법이다. 동적 계획법은 독립적인 스캔라인에 대해 최소 값을 찾을 수 있다. 동적 계획법은 조밀 스캔라인의 최적화 문제에 초점을 맞추고 있다[19]. 동적 계획법은 두 대응되는 스캔라인들 사이의 정합 비용들로 이루어진 행렬을 사용하여 최소 비용의 경로를 계산한다.

3. 제안된 조밀 시차맵 생성 기법

스테레오 영상으로부터 시차맵을 생성하기 위해서 본 논문에서는 영역 기반의 방법을 사용한다. 영역 기반의 방법으로는 색상의 유사도(similarity)에 기반하여 주어진 윈도우에서 픽셀의 가중치를 조정하는 가변 가중치를 사용하여 대응관계를 찾는 방법[20]이 있고, 영역에 기반하여 두 영상 사이의 공통의 정보를 최대화하여 유사도를 측정하는 방법[21]이 있다. Veksler[22][23]는 유용한 윈도우의 크기와 형태를 발견하였지만, 너무 많은 인자값들을 정의해야 하는 단점이 있다. 제안하는 시차맵 생성 방법에서는 SSD 비용 함수로 정합 비용을 계산하고 스캔라인 매칭 시에 영역 정보를 사용하여 잘못된 매치가 일어나지 않도록 억제시켜서 정확한 매칭 결과를 얻을 수 있도록 한다.

3.1 분할 및 평균을 사용한 시차맵 생성

본 논문에서 제안하는 시차맵 생성 방법에 있어서 유사도 측정을 위해서 SSD 비용 함수를 사용한다. SSD 비용 함수는 식 (4)와 같이 계산되며, 그 원리는 한 영상의 각 위치에 대해서 대응 영상에서 대응 후보 위치들에서의 픽셀값의 차이의 제곱한 값을 사용하여 두 픽셀 간의 차이를 양의 실수로 표현하게 된다. SSD 비용 함수를 사용하여 시차공간 영상을 생성한 후에 WTA 방법으로 시차를 계산할 경우에 객체간의 가려짐 등의 이유로 인해 잘못된 매치가 발생할 수 있다. 이러한 잘못된 매치들은 객체간의 구분을 애매하게 만들며, 동일 객체를 여러 영역으로 분리

시킬 수도 있다. 이러한 객체간의 구분의 애매함을 제거하기 위해서 본 논문에서는 영역의 균일성(homogeneity) 정보를 사용하여 잘못된 매치가 발생하지 않도록 억제하는 방법을 제안한다. 생성된 시차맵에 대해서 분산을 이용하여 시차맵을 평가할 경우 객체간의 구분이 되는 부분 즉, 분산 값이 크게 변하는 부분을 객체의 외곽선으로 판정할 수 있다. 이러한 분산 정보를 이용하여 일정 임계치 이상의 변화에 대해서 객체들을 구분한다. 객체들을 구분한 후에는 같은 객체에 대해서 균일성을 보장하기 위해 임계치 이하의 분산 변화량에 대해서는 같은 객체로 판정하고, 그 영역들에 대해서 평균의 방법으로 객체의 균일성을 보장하는 방법을 제안한다.

본 논문에서 제안하는 스테레오 영상으로부터 시차맵을 생성하는 전체 과정이 (그림 2)에 나타나 있다. 가장 먼저 두 영상의 스캔라인을 맞춰주기 위해 보정(rectification)을 한다. 보정은 두 영상에서의 대응되는 에피폴 라인들이 동일한 행으로 배치되도록 변환하는 절차이다. 두 영상의 보정은 두 영상의 에피폴 기하학을 사용하여 변환 관계를 구한 후에 정합 영상을 참조 영상에 맞추게 된다. 두 스테레오 카메라가 직각으로 정확히 배치되어 있다면 촬영되는 두 영상의 각 행이 에피폴 라인에 일치하므로 별도의 보정 과정이 필요없다. 그러나 촬영시의 두 카메라의 배치가 정확한 직각이 아닌 경우에는 보정 과정을 수행하도록 한다. 보정된 두 스테레오 영상에서의 매칭은 단순히 각 행의 매칭 문제로 간주될 수 있으므로 매칭을 빠르게 수행할 수 있다. 빠른 매칭의 수행을 위해서는 두 카메라를 직각 스테레오 장치에 부착하여 촬영하도록 하여 입력 영상들의 보정 과정이 생략될 수 있도록 하자.

다음으로, 식 (4)의 SSD 비용 함수를 사용하여 두 영상에 대한 유사도를 측정한다. 이 단계에서는 두 영상이 보정되어 있다고 가정하므로 두 영상에 대해 동일 행의 라인들을 대응되는 스캔라인으로 하여 정합비용을 계산하고, 이들은 집성한다. 이 단계에서는 시차의 단계별로 식 (4)의 결과인 시차공간 영상을 생성하게 된다. 보정된 후에는 매칭을 스캔라인별로 수행하므로 시차는 하나의 스칼라 값으로 표현된다. 시차공간 영상은 3차원 배열의 형태이고 첫 번째와 두 번째 인덱스는 기준 영상에서의 위치 (r,c) 이고, 세 번째 인덱스는 위치 (r,c) 에서의 시차 d 이다. 시차공간 영상에서의 각 픽셀은 시차 d 에 대한 유사성 측도값을 가진다.

다음으로, 시차를 계산하게 되는데, 시차별로 생성된 시

차공간 영상에 대해서 각 픽셀에서 비용이 가장 적은 픽셀을 선택하는 WTA(winner-take-all) 방법을 사용하여 시차를 생성한다. WTA는 각 지점 (r,c) 에서의 시차공간 영상에서 최소 정합 비용을 가지는 시차 d 를 선택한다. 선택된 시차 d 를 (r,c) 에서의 시차로 지정한다.

SSD 비용 함수로 계산된 정합 비용만을 사용하여 시차맵을 생성하게 되면, 다른 물체와의 가려짐이나, 매치에 실패한 픽셀들 때문에 시차맵이 거칠게 나타난다. 본 논문에서는 이러한 문제를 해결하기 위해 영역 분할을 실시하여 세그먼트들을 생성한다. 영역 분할에서의 영역 균일성 측도는 균일한 밝기값의 분포 영역에 기반한다. 각 지점에서의 밝기값의 평균과 분산의 변화율을 구하고 이로부터 평균 밝기가 변하는 영역을 세그먼트 경계로 지정한다. 예지의 강도를 동시에 고려하면 보다 정확한 영역들을 구할 수 있다[24].

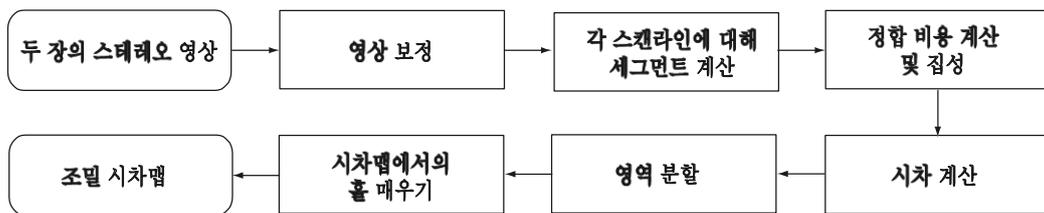
영역 균일성 측도로 밝기의 균일성을 고려한다. 영상의 위치 (r,c) 에서의 분산 $s(r,c)$ 는 식(6)과 같이 정의한다.

$$s(r,c) = \frac{1}{n} \sum_{p \in W(r,c)} (p - m_{rc})^2 \tag{6}$$

여기서 $W(r,c)$ 는 (r,c) 를 중심으로 하는 윈도우 내의 모든 픽셀들을 포함하는 픽셀 집합이고 N 은 $W(r,c)$ 내의 픽셀들의 개수이고 m_{rc} 는 $W(r,c)$ 의 모든 픽셀들의 밝기의 평균이다.

영역 균일성 측도로 밝기의 균일성과 더불어 밝기의 변화량을 고려한다. 밝기의 변화량은 (r,c) 에서의 그래디언트(gradient)의 크기(magnitude)로 측정된다. 위치 (r,c) 에서의 수평과 수직 방향으로의 그래디언트를 G_x 와 G_y 라고 하면 위치 (r,c) 에서의 그래디언트의 크기는 $e(r,c) = \sqrt{G_x^2 + G_y^2}$ 이다. 보다 빠른 계산을 위해서 그래디언트의 크기를 $e(r,c) = |G_x| + |G_y|$ 로 근사할 수 있다. 그래디언트의 크기가 클수록 예지의 강도가 클을 의미하므로 그래디언트의 크기도 영역의 균일성 측정에 사용될 수 있다. 위치 (r,c) 에서의 균일성 측도 $H(r,c)$ 는 식 (7)과 같이 정의한다. $H(r,c)$ 는 1 이하의 양수값을 가지며 1에 가까울수록 영역이 균일함을 의미하며 값이 작아질수록 다른 영역과 인접한 위치임을 의미한다.

$$H(r,c) = \frac{1}{1 + s(r,c) * e(r,c)} \tag{7}$$



(그림 2) 제안된 방법의 전체 과정


```

입력 : 홀이 메워지지 않은 시차맵 ( $I^d$ )
출력 : 홀을 메운 조밀 시차맵 ( $I^f$ )
인자 :  $width$ ,  $height$ (시차맵의 폭과 높이),  $winSizeMax$ (윈도우 최대 크기)
begin
 $winSize := 1$ 
for (시차맵  $I^d$ 에서의 각 위치  $(r,c)$ 에 대해) do
    if ( $I^d(r,c)$ 가 홀일 경우) then
        repeat {
            위치  $(r,c)$ 에서 크기가  $\Phi winSize$ 인 윈도우를 설정.
            설정 윈도우 내에서  $(r,c)$ 를 포함하는 동일영역의 최소사각형 찾기.
            if (최소사각형을 찾을 수 없으면) then
                 $winSize$ 를 1 증가시켜 다시 시도함.
            else {
                 $I^f(r,c) :=$  최소사각형 내에서의 쌍선형 보간으로 시차값 결정.
                루프를 빠져나감.
            }
        } until ( $winSize < winSizeMax$ );
    else
         $I^f(r,c) := I^d(r,c)$ 
    end if
end for
end
    
```

(그림 5) 인접 픽셀 정보를 사용한 홀 메우기 알고리즘

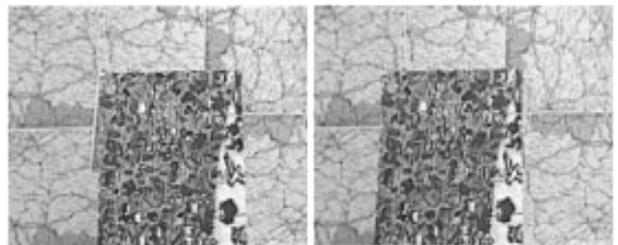
픽셀을 찾는다. 이 픽셀과 동일한 영역 내의 인접 픽셀들로 최소사각형을 구성하도록 시도한다. 최소사각형이란 그 홀 픽셀을 포함하는 가장 작은 사각형이다. 최소사각형이 구성이 가능하면 그 홀 픽셀의 시차를 인접 픽셀들의 시차로부터 보간하여 구한다. 인접 픽셀들이 동일 영역의 픽셀들이므로 매우 정확한 시차값들을 얻을 수 있다.

인접한 최소사각형을 구성할 수 없는 경우라면 영상의 변두리에 위치한 홀이거나 영역들 사이에 위치한 홀에 있는 픽셀인 경우이다. 이 경우는 두 스테레오 영상에서 한 쪽 영상에서만 보이는 위치이거나 가려진 위치에 해당한다. 이러한 위치들에서는 강제로 보간을 하게 된다면 틀린 결과를 얻을 수 있는 위치들이므로 홀 메우기를 수행하지 않도록 한다. 홀 메우기의 전체적인 절차가 (그림 5)에 있다.

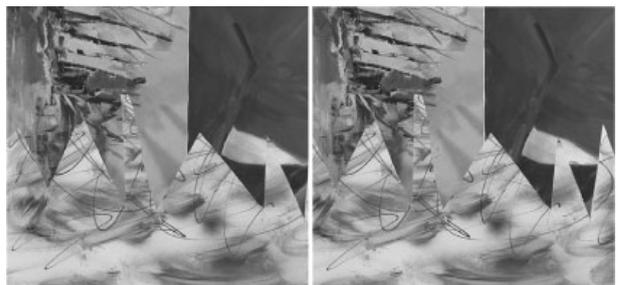
4. 실험 결과

제안된 방법의 유효성을 증명하기 위해 본 논문에서 제안된 조밀 시차맵 생성 기법을 윈도우 플랫폼에서 구현하였다. PC는 2.4GHz 코어2 프로세서로 1GB DDR RAM의 메모리를 장착하고 512MB DRAM의 GeForce 6800 GPU의 그래픽 카드를 사용하였다. 제안된 방법의 유효성을 보이기 위한 스테레오 영상으로는 (그림 6)~(그림 9)의 영상을 사용하였다. 모든 입력 영상들은 직각인 스테레오 고정 장치로 촬영되어서 보정 과정은 별도로 수행할 필요가 없다.

조밀 시차맵 생성을 위한 제안된 기법의 기존의 방법들



(그림 6) 284×216 크기의 Map의 좌우 스테레오 영상



(그림 7) 434×380 크기의 Sawtooth의 좌우 스테레오 영상



(그림 8) 384×288 크기의 Tsukuba의 좌우 스테레오 영상

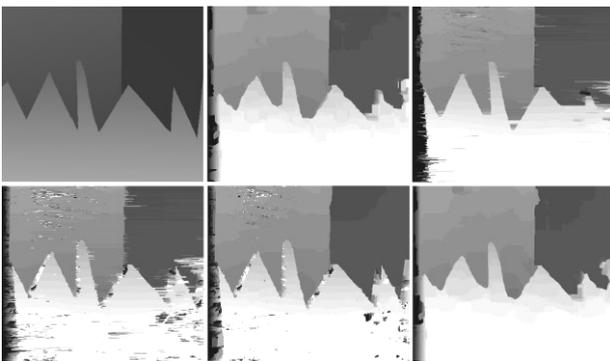


(그림 9) 434×383 크기의 Venus의 좌우 스테레오 영상

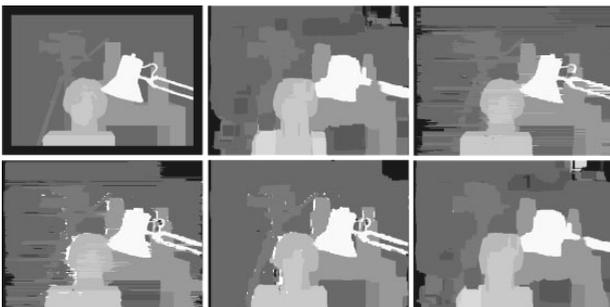
과의 비교를 위해서 Scharstein와 Szeliski[7]의 실험 데이터 및 코드를 사용하였다. (그림 10)~(그림 13)은 각각 (그림 6)~(그림 9)의 영상을 입력으로 시차맵을 생성한 결과로서 윗줄의 왼쪽은 실제 시차맵(ground truth), 가운데는 21×21 크기의 이동 윈도우(shiftable window) 방식의 SSD, 오른쪽은 DP(Dynamic Programming), 아랫줄의 왼쪽은 SO(Scanline Optimization), 가운데는 GC(Graph Cuts), 오



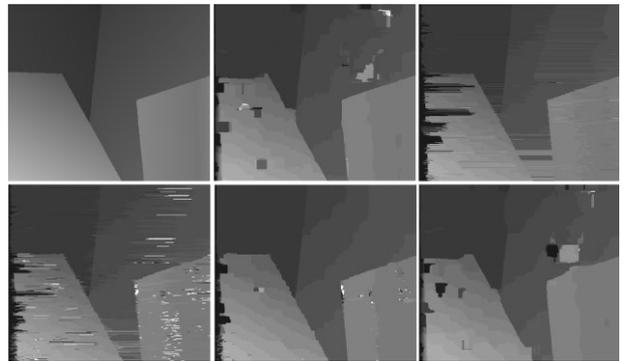
(그림 10) Map의 시차맵의 생성 결과



(그림 11) Sawtooth의 시차맵의 생성 결과



(그림 12) Tsukuba의 시차맵의 생성 결과



(그림 13) Venus의 시차맵의 생성 결과

른쪽은 제안한 방법(PM: Proposed Method)으로 시차맵을 생성한 것이다. 실제 시차맵은 Scharstein과 Szeliski[7]의 연구에서 사용된 데이터를 사용하였다.

<표 1>~<표 4>는 각각 (그림 10)~(그림 13)의 시차맵 생성 결과에 따른 성능을 분석한 결과를 표로 나타낸 것이다. RMS 에러(root mean squared error)는 식 (7)과 같이 계산되며 실제 시차맵과 얼마나 많은 차이를 가지는지를 의미한다. 여기서 N 은 픽셀의 수를 의미하며, I^d 와 I^t 는 각각 시차맵과 실제 픽셀값을 의미한다. 정확하지 않은 정합(bad match)은 식 (8)과 같이 계산되며 부정확하게 정합된 픽셀들의 비율을 나타낸다. 여기서 δ 는 부정확하게 정합된 픽셀을 판별하는 임계치를 의미하며, 본 논문에서는 1.0으로 설정하여 실험하였다.

$$R_{err} = \left(\frac{1}{N} \sum_{(x,y)} |I^d(x,y) - I^t(x,y)|^2 \right)^{\frac{1}{2}} \quad (7)$$

$$B_{err} = \frac{1}{N} \sum_{(x,y)} (|I^d(x,y) - I^t(x,y)| > \delta) \quad (8)$$

각 방법들의 수행 시간이 <표 1>~<표 4>에 있다. 수행 시간은 입력 스테레오 영상으로부터 조밀 시차맵을 생성하는데 걸리는 모든 시간들을 합한 시간이다. 측정된 시간을 보면 제안한 방법(PM)이 기존의 방법들 중에서 가장 빠른 방법인 이동 윈도우 방식의 SSD 방법과 수행시간이 거의 비슷하게 나타나면서, RMS 에러 및 부정확한 정합이 다른 방법에 비해 더 좋은 성능을 나타내는 것을 확인할 수

<표 1> Map의 시차맵 생성 결과에 따른 성능 분석

	SSD	DP	SO	GC	PM
RMS 에러	4.93591	4.88630	6.68418	6.30781	4.46202
Bad match	0.05448	0.05575	0.07069	0.05635	0.04651
시간(초)	4.62500	4.32800	5.46900	829.593	4.51500

<표 2> Sawtooth의 시차맵 생성 결과에 따른 성능 분석

	SSD	DP	SO	GC	PM
RMS 에러	21.9158	21.5816	21.9316	21.8404	21.4713
Bad match	0.09961	0.09721	0.09907	0.09927	0.09954
시간(초)	11.5780	12.2030	15.4690	2866.75	11.9220

〈표 3〉 Tsukuba의 시차맵 생성 결과에 따른 성능 분석

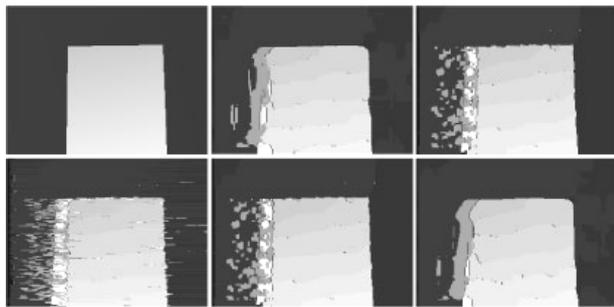
	SSD	DP	SO	GC	PM
RMS 에러	1.92046	1.97323	3.11331	2.92175	1.89033
Bad match	0.11219	0.13135	0.13773	0.11733	0.11055
시간(초)	8.03200	8.23500	10.0780	1253.40	7.87500

〈표 4〉 Venus의 시차맵 생성 결과에 따른 성능 분석

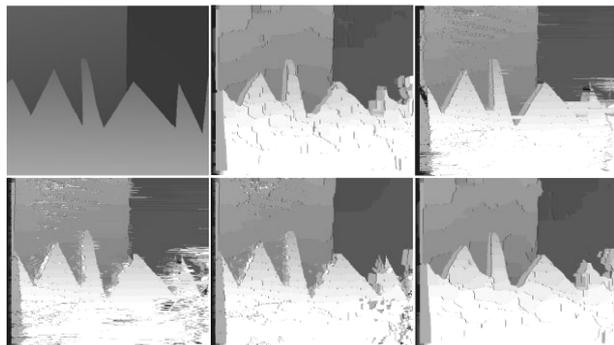
	SSD	DP	SO	GC	PM
RMS 에러	1.67291	2.55647	3.46238	1.84527	1.72354
Bad match	0.05237	0.11038	0.10605	0.03836	0.06312
시간(초)	11.7180	12.2970	15.2660	2565.53	11.9530

있다. SSD 방법에 있어서 수행 시간이 제안 방법보다 더 길어지는 경우도 발생하였는데 이는 비교 대상인 SSD 방법이 정확한 시차맵을 생성할 수 있도록 윈도우의 크기를 충분히 크게 하였고 또한 움직이는 윈도우 방법으로 계산 시간이 더 늘어났기 때문이다.

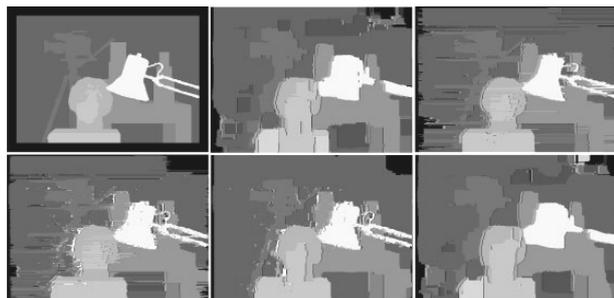
(그림 14)~(그림 17)은 생성된 시차맵에 가려짐(occlusion)을 표시한 결과로서, 왼쪽 영상을 기준으로 오른쪽 영상에



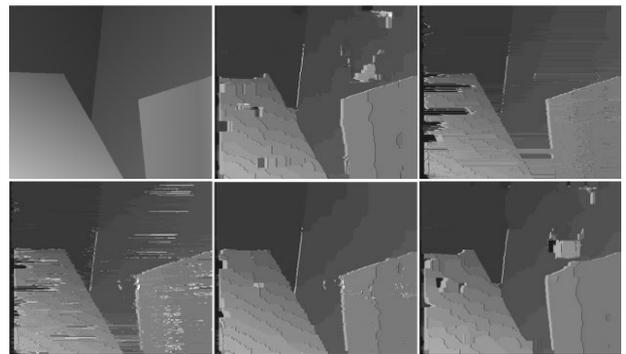
(그림 14) Map의 시차맵에 가려짐을 표시한 결과



(그림 15) Sawtooth의 시차맵에 가려짐을 표시한 결과



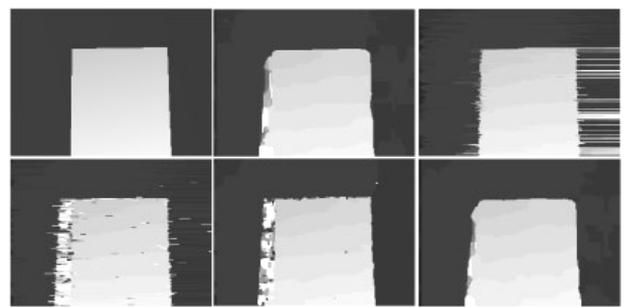
(그림 16) Tsukuba의 시차맵에 가려짐을 표시한 결과



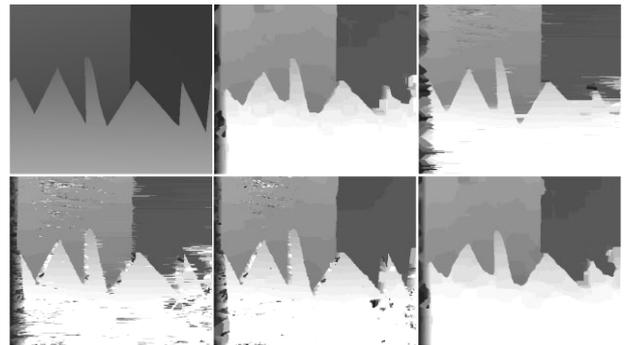
(그림 17) Venus의 시차맵에 가려짐을 표시한 결과

서 정합되는 픽셀을 찾을 수 없는 경우에는 녹색, 반대의 경우에는 파란색으로 표시하였으며, 본 논문에서는 왼쪽 영상을 참조 영상으로 설정하였기 때문에 파란색의 경우는 영역으로 나타나지 않고 선과 같은 경계로 나타나는 것을 확인할 수 있다.

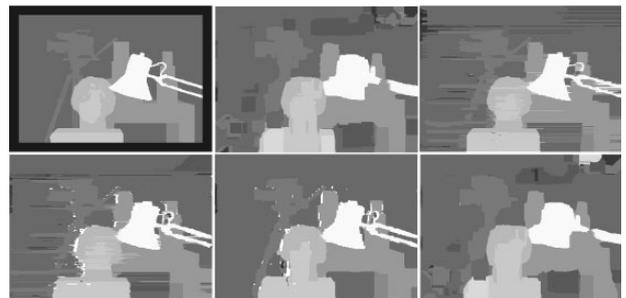
(그림 18)~(그림 21)은 생성된 시차맵에서 홀에 해당



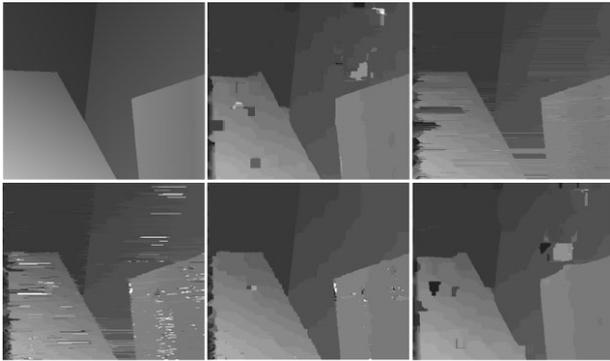
(그림 18) Map의 시차맵에서 홀을 메운 결과



(그림 19) Sawtooth의 시차맵에서 홀을 메운 결과



(그림 20) Tsukuba의 시차맵에서 홀을 메운 결과



(그림 21) Venus의 시차맵에서 홀을 메운 결과

하는 부분들을 인접한 픽셀들로부터 보간하여 홀을 메운 조밀 시차맵을 보여준다.

5. 결 론

본 논문은 두 스테레오 영상으로부터 영역 기반 매칭을 통해 조밀 시차맵을 생성하는 기법을 제안하였다. 스테레오 영상으로부터 빠르고 안정된 조밀 시차맵을 생성하기 위해서 각 스캔라인에 대해서 분산을 이용하여 분할한 후에 각 분할구간 별로 평균을 내어 객체간의 구분을 명확하게 해 주는 과정을 도입하였다. 영역의 균일성 특성에 따라 상이한 영역으로의 잘못된 매칭을 억제하여 올바른 시차맵이 생성되도록 하였다. 3차원 구조를 복원하기 위해서는 시차맵의 모든 픽셀에 대해서 시차를 계산해야 하는데 매칭에 실패한 픽셀들로부터 발생하는 시차맵에서의 홀에 대해서 소속되어야 할 영역의 인접한 사각형을 얻고 쌍선형 보간으로 시차값을 구하여 메우게 된다.

제안된 방법을 다양한 스테레오 영상에 대해서 적용하고, 다른 조밀 시차맵 생성 방법들과 비교하였다. 실험 결과들은 제안된 방법이 다른 방법들에 비해 시간 비용을 많이 들이지 않고도 좋은 결과를 얻을 수 있다는 것을 보여주었다. 제안된 조밀 시차맵 기법은 형상의 실사기반 렌더링과 물체 표면의 조밀한 메시 생성 등의 응용에 사용될 수 있다. 향후 연구 과제로 가려짐에 따른 영역 분할 문제와 물체나 카메라의 움직임에 따른 모션 분할 문제가 해결되어야 할 것이다. 영역의 분할 문제에 있어서 가려짐과 움직임을 탐색하고 그 정보를 분할에 적용할 수 있다면 실용적인 대화형 응용 시스템에도 조밀 시차맵 생성 기법이 적용될 수 있을 것이다.

참 고 문 헌

- [1] Han-Suh Koo and Chang-Sung Jeong, "An Area-Based Stereo Matching Using Adaptive Search Range and Window Size," Proceedings of the International Conference on Computational Science-Part II, pp.44-56, 2001.
- [2] L. Di Stefano, M. Marchionni and S. Mattocchia, "A Fast Area-Based Stereo Matching Algorithm," Image and Vision Computing, Vol.22, No.12, pp.983-1005, 2004.
- [3] Camillo J. Taylor, "Surface Reconstruction from Feature Based Stereo," IEEE International Conference on Computer Vision, Vol.1, pp.184-190, 2003.
- [4] M. S. Lew, T. S. Suang and K. Wong, "Learning and Feature Selection in Stereo Matching," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.16, No.9, pp.869-881, 1994.
- [5] S. B. Marapane and M. M. Trivedi, "Multi-Primitive Hierarchical (MPH) Stereo Analysis," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.16, No.3, pp.227-240, 1994.
- [6] D. M. McKeown and Y. C. Hsieh, "Hierarchical Waveform Matching: A New Feature-Based Stereo Technique," In Computer Vision and Pattern Recognition," pp.513-519, 1992.
- [7] Daniel Scharstein and Richard Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," International Journal of Computer Vision, Vol. 47, No.1-3, pp.7-42, 2002.
- [8] P. Anandan, "A Computational Framework and An Algorithm for The Measurement of Visual Motion," International Journal of Computer Vision, Vol.2, No.3, pp.283-310, 1989.
- [9] Takeo Kanade, "Development of a Video-Rate Stereo Machine," Proceedings of the 1994 ARPA Image Understanding Workshop, pp.549-558, 1994.
- [10] R. C. Bolles, H. H. Baker and M. J. Hannah, "The JISCT Stereo Evaluation," DARPA93, pp.263-274, 1993.
- [11] Aaron F. Bobick and Stephen S. Intille, "Large Occlusion Stereo," International Journal of Computer Vision, Vol.33, No.3, pp.181-200, 1999.
- [12] Sing Bing Kang, R. Szeliski and Jinxiang Chai, "Handling Occlusion in Dense Multi-View Stereo," In Computer Vision and Pattern Recognition, Vol.1, pp.103-110, 2001.
- [13] W. E. L. Grimson, "Computational Experiments with a Feature Based Stereo Algorithm," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.7, No.1, pp.17-34, 1985.
- [14] K. Prazdny, "Detection of Binocular Disparities," Readings in Computer Vision: issues, problems, principles, and paradigms, pp.73-79, 1987.
- [15] D. Scharstein and R. Szeliski, "Stereo Matching with Non-linear Diffusion," International Journal of Computer Vision, Vol.28, No.2, pp.155-174, 1998.
- [16] D. Terzopoulos, "Regularization of Inverse Visual Problems Involving Discontinuities," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.8, No.4, pp.413-424, 1986.
- [17] C. L. Zitnick and T. Kanade, "A Cooperative Algorithm for

Stereo Matching and Occlusion Detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.7, pp.675-684, 2000.

[18] Y. Boykov, O. Veksler and R. Zabih, “Fast Approximate Energy Minimization via Graph Cuts,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.11, pp.1222-1239, 2001.

[19] 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.

[20] Kuk-Jin Yoon and In So Kweon, “Adaptive Support-Weight Approach for Correspondence Search,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, No.4, pp.650-656, 2006.

[21] Z. Zivkovic and B. Krose, “On Matching Interest Regions Using Local Descriptors: Can an Information Theoretic Approach Help?,” British Machine Vision Conference, 2005.

[22] O. Veksler, “Stereo Correspondence with Compact Windows via Minimum Ratio Cycle,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.12, pp.1654-1660, 2002.

[23] O. Veksler, “Fast Variable Window for Stereo Correspondence using Integral Images,” In IEEE Computer Vision and Pattern Recognition, Vol.1, pp.556-561, 2003.

[24] S. Hernandez, K. E. Barner and Y. Yuan, “Region Merging using Homogeneity and Edge Integrity for Watershed-based Image Segmentation,” Optical Engineering, Vol.44, pp.1-14, 2005.



이 범 중

e-mail : leeyanga@incheon.ac.kr
 2006년 인천대학교 컴퓨터공학과(공학사)
 2008년 인천대학교 컴퓨터공학과(공학석사)
 관심분야 : 동작기반 인터페이스, 증강현실, 가상현실



박 종 승

e-mail : jong@incheon.ac.kr
 1992년 경북대학교 전자계산학과(이학사)
 1994년 포항공과대학교 컴퓨터공학과(공학석사)
 1999년 포항공과대학교 컴퓨터공학과(공학박사)
 1999년~2001년 한국전자통신연구원 가상현실연구부 선임연구원
 2001년~2003년 디지털아리아 시각정보처리연구소 책임연구원
 2004년~현 재 인천대학교 컴퓨터공학과 조교수
 관심분야 : 영상처리, 3D 비전, 증강현실, 게임공학



김 정 규

e-mail : ckkim@incheon.ac.kr
 1974년 연세대학교(이학사)
 1981년 일본 치바(千葉)대학(공학석사)
 1985년 일본 토카이(東海)대학(공학박사)
 1989년~현 재 인천대학교 교수
 1999년~2000년 UNSW(Sydney, Australia) Visiting Scholar
 1985년~1989년 KAIST 시스템공학센터 선임연구원
 1984년~1985년 정보과학전문학교 이와사키(岩崎) 학원 강사
 1981년~1985년 일본 토카이(東海)대학 정보기술센터 연구원
 1979년~1981년 일본치바(千葉)대학 영상격측연구센터 연구원