

워터셰드 영역병합을 이용한 스테레오 정합의 베이지언 접근방법

길우성, 김 신형, 장 종환
배재대학교 정보통신공학과

e-mail : wskil@hotmail.com , zeros@mail.pcu.ac.kr , jangjw@mail.pcu.ac.kr

A Bayesian Approach to Stereo Matching via Merging Watershed Regions

WooSung Kil, ShinHyung Kim, JongWhan Jang
Dept of Information and Communication Engineering
Paichai University

요 약

본 논문은 세그멘테이션 기반의 스테레오 정합에서 복잡한 장면 정합 시 발생하는 오 정합을 최소화 하는 방법을 제안한다. 이를 위하여, 스테레오 영상의 좌측 영상에 대해 워터셰드 영상 분할을 이용하여 정합을 위한 feature 를 생성한 다음, 베이지언 프레임워크를 적용하여, 각각의 영역을 비슷한 변이 정보를 가진 것들로 병합한다. 생성되는 정합 패치들은 정합의 모호성이 작게 되어 오 정합이 현저히 줄어 들 뿐만 아니라, 영역간의 콘트라스트가 적은 영상에서도 신뢰할 만한 변이 영상을 생성하게 된다.

1. 서론

가상현실, 3DTV 등의 3 차원 영상정보 시스템의 구현에 있어 2 차원 영상으로부터 3 차원 정보를 추출하기 위하여 스테레오 정합 방법이 활발히 연구되고 있다. 일반적으로, 스테레오 정합의 결과로 생성되는 변이 영상은 변이의 에지를 잘 유지하고, low 텍스처 영역에서 변이의 부드러움(smoothness)를 유지하는 것이 관건이다. 이를 위하여, 영상을 의미 있는 영역의 집합으로 분할 한 후, 각 영역을 서로 정합하는 패치 정합 방법이 연구되고 있다[1][2]. 하지만 비슷한 깊이를 가졌어도 강도의 변화가 큰 영상은 과 분할된 작은 크기의 패치를 정합 해야 하기 때문에 오 정합이 발생되기 쉽고, 더욱이 패치 사이의 contrast 가 작은 경우 정합의 모호함이 커져 정확한 변이 예측이 어렵게 된다.

이러한 문제를 해결하기 위하여 본 논문은 다음 세 단계의 방법을 제안한다. 첫째로 좌측의 카메라로부터 입력되는 영상을 워터셰드 영역으로 분할한다[3][4][5]. 다음으로 초기 윈도우 기반 정합 방법으로 얻어진 변

이 정보를 이용하여 각 영역을 병합한다. 이 과정에서 기준 패치와 서로 비슷한 변이 특성을 가진 영역은 베이지언 확률을 적용하여 사후 확률이 최대가 되는 병합 영역이 정합 패치로 선택된다. 마지막으로 병합된 영역을 서로 정합하여 변이 영상을 구한다.

본 연구의 주된 이점은 패치를 이용한 스테레오 정합에 있어 변이 정보를 함께 고려하여, 영상의 복잡도에 관계없이 신뢰 할 만한 결과를 얻을 수 있다는데 있고, 결과는 실험을 통하여 증명 하였다.

2. 워터셰드 영상 분할

워터셰드 영상 분할 알고리즘은 영상의 그래디언트 (gradient)정보를 이용하여 분할된 영역을 구분하는 방법이다. 일반적인 패치 정합 알고리즘은 분할된 각 영역에 대한 대응(correspondence)를 구함으로써, 각 패치의 변이, 즉 feature disparity 를 찾는다. 본 논문에서는 분할된 각각의 워터셰드 영역에 대한 feature disparity 를 찾음으로써, 영상에 대한 변이 정보를 구한다. 그러나 기본적인 워터셰드 분할 방법은 영상의 잡음과, 그래디언트 영상에서의 불규칙한 영상 강도의 변화에

민감하게 동작하여, 과 분할(over-segmentation)이 발생하게 된다. 이렇게 과 분할된 영역들은 오른쪽 영상에서 그에 대응되는 패치를 찾기에 충분치 않기 때문에, 이웃 패치와 함께 병합하여 정합 하여야 안정된 결과를 얻을 수 있다.

3. 베이지언 확률을 이용한 워터셰드 영역 병합

3.1 초기 변이 영상

베이지언 병합을 수행하기 전에, 각 영역에서의 초기 변이 정보를 구한다. 초기의 변이 정보를 이용하여 분할된 두 영역이 같은 변이를 갖고 있는지를 판단하여, 같은 변이 특성을 갖는 패치들은 서로 병합된다. 초기의 변이 영상은 (식 1)과 같이 일반적인 윈도우 정합 방법을 이용하였다.

$$D_{init}(x,y) = \arg \min_d \left\{ \sum_{w=0}^{n-1} \{I_l(x+w,y) - I_r(x+w+d,y)\}^2 \right\} \quad (\text{식 1})$$

I_l 과 I_r 은 각각 스테레오 좌, 우 영상을 나타낸다. 좌측 영상에서 고정된 크기의 윈도우 w 는 탐색 구간을 따라가면서 윈도우 내의 화소들과 가장 유사한 화소의 위치 d 를 변이 영상에 할당한다. 다음으로, 각 패치에 대하여 (식 2)와 같이 초기 변이 영상의 평균을 구한다.

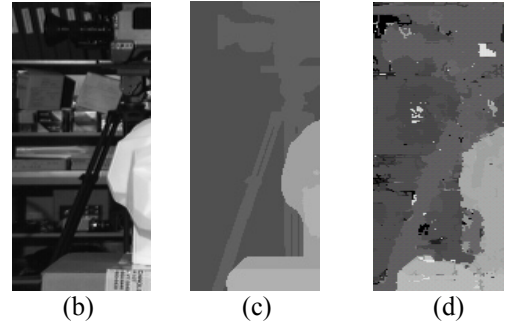
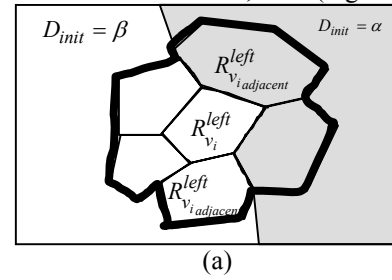
$$\hat{\theta}_v = \frac{1}{N} \sum_{(x,y) \in R_v^{left}} D_{init}(x,y) \quad (\text{식 2})$$

여기서 N 은 패치 내에 포함된 화소의 수를 나타내고, R_v^{left} 는 좌측 영상에서 워터셰드로 분할된 각 영역으로, 왼쪽 영상이 K 개의 영역으로 나뉘어 졌을 때 $\{R_v^{left} | v = \{v_j | j=1,2,\dots,K\}\}$ 와 같고, 각 패치에서의 변이의 평균 벡터는 $\theta_v = [\theta_{v_1}, \theta_{v_2}, \dots, \theta_{v_k}]$ 으로 표현할 수 있다.

만일 서로 다른 변이를 가진 두 패치가 병합될 경우, 변이 영상은 투영 왜곡(projective distortion)이나, 영역 확장 현상(fattening effect)등이 일어나기 쉽다. 패치에 대한 변이 벡터는 스테레오 정합을 위한 패치 병합에 있어 매우 중요한 정보를 제공한다. <그림 1>은 패치 내의 변이 변화를 고려하지 않고 병합한 결과를 보여준다. <그림 1>(a)에서 병합된 패치는 그 내부에 (식 1)을 통하여 구한 초기 변이 $D_{init}=\alpha$, $D_{init}=\beta$ 를 포함하고 있다. 따라서 변이 영상은 (d)와 같이 실제의 변이 경계를 유지하지 못하고 밖으로 확장되어 나타난다. 이러한 현상은 Kanade의 적응 윈도우 방법[6]에서 윈도우 내에서의 변이 변화에 따른 투영 왜곡 현상과 유사하며, 패치 정합에 있어 병합된 영역의 내부는 변이 변화를 되도록 포함하지 않아야 한다.

3.2. 제안된 방법

본 연구에서 정합하고자 하는 각 패치와 인접한 패치와의 관계를 정의하기 위하여, RAG(region adjacency



<그림 1> 변이를 고려하지 않고 병합된 영역에 대한 스테레오 정합의 예. (a): 변이 경계를 포함한 병합 패치 (b): 원 좌측 영상 (c): ground truth 영상 (d): (a) 패치를 이용한 변이 영상

graph)를 이용하였다. RAG의 각 노드는 워터셰드 패치에 대응된다. 각각의 노드에 대하여 만일 인접한 패치가 있다면 두 노드는 arc로 연결되어, 패치 사이의 관계를 쉽게 알 수 있다. 다음으로 RAG 상에서 각 영역에 대한 라벨 필드 $L = \{l_v | l_v \in [1,2,\dots,N_l], v \in V\}$ 를 정의한다. 여기서 $V = \{v_j | j=1,2,\dots,K\}$ 이다. 이전 절에서 구한 각 영역의 변이 벡터 $D = \{\theta_v | v \in V\}$ 와 함께 (식 3)과 같이 베이지언 법칙을 사용하여 사후확률을 구한다.

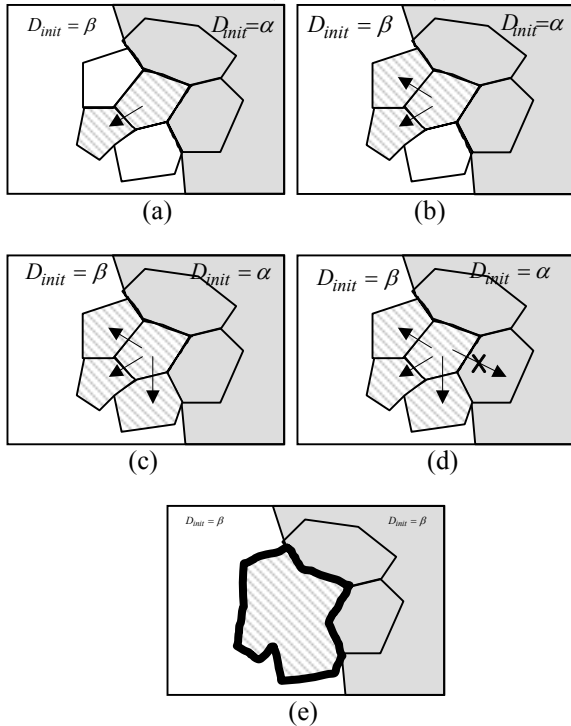
$$P(L|D) \propto P(D|L) \cdot P(L) \quad (\text{식 3})$$

여기서 $P(D|L)$ 은 조건부 확률 분포로서, 주어진 병합 영역, 즉, 각 연결된 노드가 최소의 변이 분산을 갖고 있어야 하고, 다음과 같이 표현될 수 있다.

$$P(D|L) = \exp \left(- \sum_{v \in V} \frac{1}{2\sigma_l} \sum_{p \in R_v^{left}} \|D_{init}(p) - \theta_v\|^2 \right) \quad (\text{식 4})$$

여기서 p 는 $p \in R_v^{left}$ 에 포함된 각 화소를 나타낸다. 이를 위하여 기준 패치와 인접한 패치의 변이 분산을 고려하여 <그림 2>와 같이 단계적으로 병합 패치를 생성한다. 각각의 단계에서 병합된 영역은 초기 변이 정보에 대한 분산을 구하여, 변이의 분산이 일정 값 이상이 되면 병합이 정지된다. 그림에서, 기준 패치는 β 값을 갖는 초기 변이에 위치하고 있고, α 값의 변이

를 갖는 이웃 패치와 인접해 있다. 패치는 많은 신호 성분을 포함하기 위하여 인접 패치와 병합을 시도하



<그림 2> Bayesian 병합 과정 (a), (b), (c): 기준 패치에 대하여 각 라벨 필드에 대해 변이의 분산을 검사 (d) 변이의 변화가 커지면 병합이 중지 (e) 선택된 정합 패치

지만, 변이가 서로 다른 패치를 만나면 병합된 영역 내에 변이의 분산이 커지므로, 병합이 중지된다. 따라서 선택되는 병합 영역은 내부에 변이의 변화가 없는 영역이 된다. 다음으로 (식 3)의 사전 확률 분포 $P(L)$ 로 다음과 같이 라벨 필드에서 두 패치간의 영상 강도의 변화를 고려한다.

$$P(L) = \exp \left(- \sum_{v \in L'} \frac{1}{2\sigma_l} \sum_{p \in R_v^{left}} (I_l(p) - \bar{I}_v)^2 \right) \quad (\text{식 5})$$

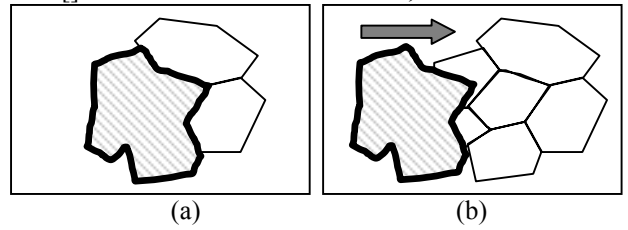
따라서 사후 확률을 최대로 하는 라벨 필드는 다음의 식으로 나타낼 수 있다.

$$\hat{L} = \arg \min_L \left(\frac{1}{2\sigma_l} \sum_{v \in L'} \sum_{p \in R_v^{left}} \{ D_{mit}(p) - \theta_v \}^2 + \{ I_l(p) - \bar{I}_v \}^2 \right) \quad (\text{식 6})$$

즉, 두 영역이 강도의 차이가 적고, 변이의 차이가 적을 때, 사후 확률이 최대가 된다. 즉, 두 영역 사이에 영상 강도가 작아도 워터셰드 알고리즘으로 인해 분할이 되었어도, 사후 확률이 커지기 때문에 병합될 가능성이 커지게 된다.

4. 병합된 영역의 스테레오 정합

이전 절에서 다룬 정합 패치는 적응적으로 선택된 윈도우[]와 비슷한 기능을 한다. 즉, 패치 내에 변이의



<그림 3> 병합된 영역의 스테레오 정합 과정 (a) 왼쪽 영상 (b) 오른쪽 영상

변화를 최소화 함으로서 변이 예측의 불확실성을 최소화 하게 된다. 그러나 세그멘테이션에 의하여 생성된 패치를 이용하게 되면 영역의 형태 정보가 부가되기 때문에, 즉, 각 패치는 영상의 경계 정보와 넓이 등을 갖고 있으므로, 더욱 의미 있는 정합 프리미티브가 된다. 이와 같이 선택된 패치를 갖고 다음과 같이 정합을 시도한다.

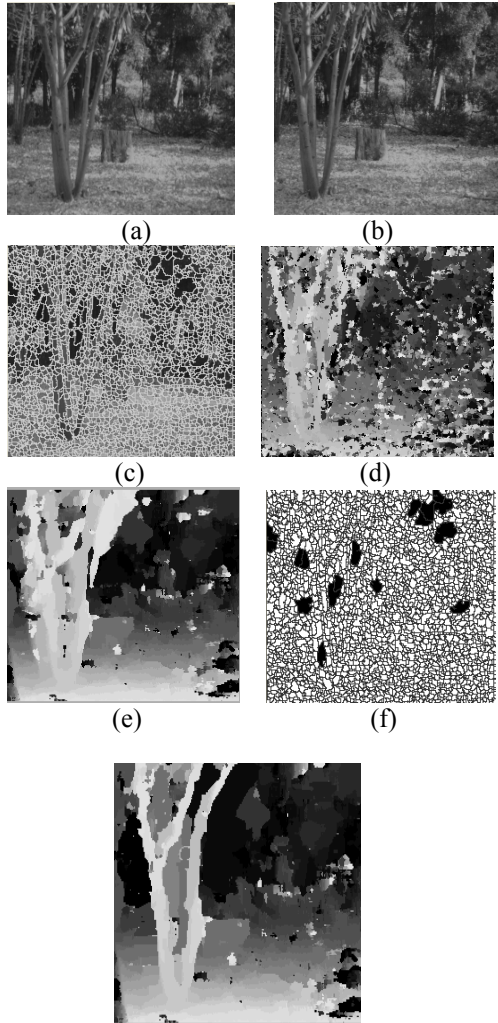
$$D(R_{v_L}^{left}) = \arg \max_{d \in D} \text{Sim}(R_{v_L}^{right}, d) \quad (\text{식 7})$$

병합된 패치는 우측 영상에서 유사도를 측정하고, 변이 값은 좌측 영상의 대응되는 패치에 할당이 된다.

5. 실험 결과

본 논문에서 제안한 적응 패치 정합 방법은 펜티엄 IV 2.0GHz, 메모리 512Mbyte, windows XP 환경에서 Visual C++을 이용하여 구현 하였다.

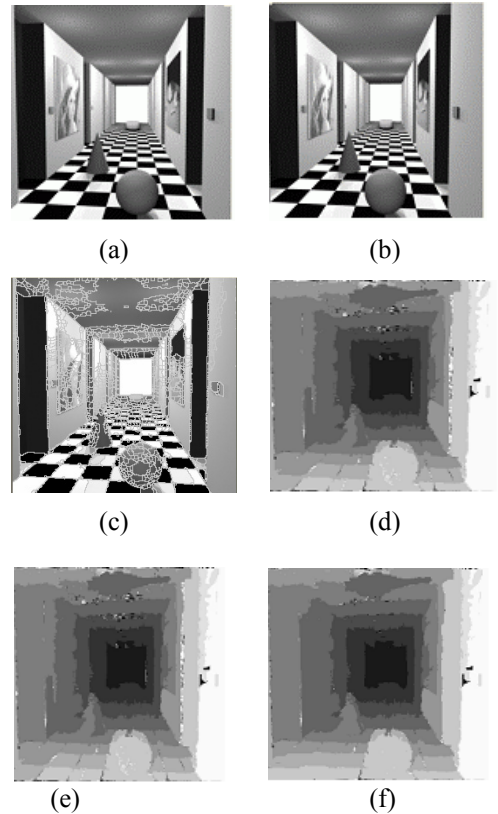
[그림 6]은 tree 영상에 대한 실험 결과를 나타낸다. (a)는 좌측 영상, (b)는 우측 영상을 나타내고 (c)는 왼쪽 영상에 대하여 워터셰드 알고리즘으로 분할된 영상을 나타낸다. Tree 영상은 영상 전 영역에 걸쳐 intensity의 변화가 크게 나타나기 때문에 (c)와 같이 영상의 대부분이 작은 크기를 가진 패치로 분할된다. 일반적인 패치 정합 알고리즘에서 작은 크기를 갖는 패치는 (d)와 같이 많은 오 정합을 발생시키지만 (e)와 같이 윈도우 기반 정합 방법은 윈도우 내에 정합을 위한 많은 신호를 포함하여 좋은 결과를 낸다. 하지만 윈도우 기반 정합에서 변이의 경계는 윈도우의 크기에 비례하여 확장되어 나타나기 때문에 변이의 경계를 유지하지 못하게 된다. 본 논문에서 제안하는 방법의 결과는 (g)와 같다. High 텍스처로 판별된 패치는 정합의 모호함으로 많은 오정합이 발생하지만 본 알고리즘으로 병합된 패치를 사용한 결과(g)는 정합 에러가 현저히 적어짐을 알 수 있고, 병합된 패치는 변이의 변화를 포함하지 않기 때문에 변이의 경계를 비교적 정확히 유지시키는 것을 알 수 있다. corridor 영상에 대한 정합 결과는 (e), (f)와 같다. 인접 영역과의 병합은 반복과정으로 수행되며 반복 과정이 많아질수록 병합된 패치는 더 많은 신호를 포함하게 되어 정합 모호함이 현저히 줄어든다.



(a)좌측 영상 (b)우측 영상 (c) watershed segmentation 영상 (d) 패치 정합 알고리즘 (e) SSD 영상 (f) high textured 추출 영상 (g) 적응 패치 정합 알고리즘
[그림 3] Tree 영상에 대한 실험 결과

6. 결론

본 논문에서는 스테레오 영상을 정합하는 방법으로 정합에 사용되는 프리미티브를 베이지언 확률로서 선택된 세그먼트 패치를 사용하였다. 일반적으로 하나의 객체는 동일한 intensity 를 갖는 영역이 많고, 각 영역은 비슷한 변이를 갖을 가능성도 크다. 따라서 우리는 영상 분할을 통해 영상을 비슷한 intensity 를 갖는 영역으로 나누고 그들을 서로 정합 시키는데, 각 영역의 내부는 서로 같은 변이를 갖는다는 가정을 하였으므로 한번의 패치 정합만으로 그 내부의 모든 화소에 대한 변이를 찾을 수 있다. 따라서 기존의 영역기반 정합 방법에서, 초기의 작은 크기의 윈도우를 사용하여, 좌, 우로 한 단계씩 한 화소 만큼 확장하여 모든 픽셀에 대해 정합을 시도한 것과 달리 정합에 있어 더욱 빠른 계산 속도를 보여 주었다. 또한 분할된 영역은 객체의 윤곽을 유지시켜주기 때문에 일반적인 영역기반 스테레오 정합에서의 문제로 지적되었던 영역 확장 현상이 현저히 줄어들 수 있었다.



(a)좌측 영상 (b)우측 영상 (c) watershed segmentation 영상 (d) 패치 정합 알고리즘 (e) 적응 패치 정합 알고리즘(반복 1 회) (f) 적응 패치 정합 알고리즘(반복 5 회)
[그림 4] Corridor 영상에 대한 실험 결과

6. 참고 문헌

- [1] M. Abbasi-Dezfouli and T. Freeman, "Patch matching in stereo-images based on shape," In ISPRS International Archives of Photogrammetry and Remote Sensing, vol. 30, pp.1-8, 1994.
- [2] M. Bleyer and M. Gelautz, "A layered stereo algorithm using image segmentation and global visibility constraints," ICIP 2004.
- [3] Vincent L. and Soille P., "Watershed in digital spaces: an efficient algorithm based on immersion simulations," IEEE Trans. Patt. Anal. Mach. Intell. vol. 13, no. 6, pp. 538-598, 1991
- [4] K. Haris, S. N. Efstratiadis, N. Maglaveras, and A. K. Katsaggelos, "Hybrid image segmentation using watersheds and fast region merging," IEEE Trans. Image Process., vol. 7, no. 12, pp. 1684-1699, Dec. 1998.
- [5] J. B. T. M. Roerdink and A. Meijster, "The watershed transform: definitions, algorithms and parallelization strategies," Fund. Inform., vol. 41, pp. 187-228, 2000.
- [6] T. Kanade and Okutomi M., "A stereo matching algorithm with an adaptive window: theory and experiment," IEEE Trans. Patt. Anal. Mach. Intell., vol. 16, no. 9 pp.920-932, 1994.