

블록 매칭 기반에서 디스페리티의 블록킹 현상을 줄이는 알고리즘

송원석, 유용현, 문정환, 김민기

고려대학교 전자정보공학과

전화 : 02-3290-3977 / 핸드폰 : 018-321-0136

A blocking effect reduction method for disparity computation under the block matching algorithm

Won Seok Song, Yong Hyun Yoo, Jeoung Hwan Moon, Min Gi Kim

Dept. of Elec. & Info. Engineering, Korea University

E-mail : originalstone40@hotmail.com

Abstract

Stereo matching is an important technique in the area of computer vision. There are already many theorems to find disparity map using stereo images. Usually, disparity is searched by using block matching. However block matching result in blocking effects caused by using fixed size window for computing pixel correlations. This paper suggests an efficient method to remove the blocking effect in stereo matching procedures.

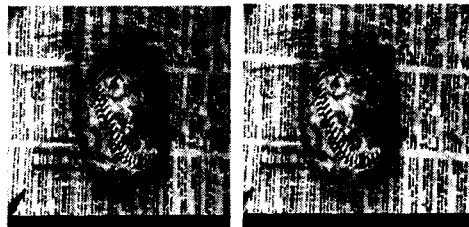
I. 서론

디스페리티(disparity) 정보추출의 목적은 삼차원 물체를 표현하는 오른쪽, 왼쪽 두개의 이차원 영상에서 상호 대응된 화소값의 위치를 찾아 그 차이를 추출함으로써 깊이에 대한 일차적인 정보를 얻어내는데 있다. 스테레오 매칭 알고리즘은 일반적으로 특징(feature based)과 영역(area based)을 기반으로 하는 두가지 방법으로 나누어진다. 특징 기반은 두 이미지의 깊이 정보를 얻기위해 보간법을 사용하여 특징부분을 찾는 방법이며 영역 기반은 찾는 윈도우 안에서 두 이미지 픽셀의 상관도를 이용하여 찾는 방법을 말하며 두 영상에서 대응점의 유사성을 판단하기 위해서 좌측 영상의 고정된 블록을 기준으로 우측 영상을 탐색하면서 비용을 구한다. 하지만

이러한 방법으로 대응점을 찾아가는 방법은 블록을 기준으로 하였고 있었기 때문에 구하고자 하는 디스페리티 맵의 블록 현상이 나타나게 된다. 이러한 블록킹 현상은 블록 단위로 대응되는 MPEG의 motion vector를 찾는 방법에서도 찾아 볼수 있다.

본 논문은 이러한 현상을 줄여주는 방법으로 영상의 특성을 고려하여 블록의 크기를 조절해 비용을 탐색하는 방법을 제안하고자 한다.

II. Matching Method(정합 방법)



(a) Right image plane (b) Left image plane
그림 1. 왼쪽, 오른쪽 입력 영상(256*256)

2.1 Cost Function(비용함수)

두 영상에서 대응점의 유사성을 판단하기 위해 일반적으로 블록 정합을 주로 사용하며 좌측 영상의 고정된 창을 기준으로 우측영상을 탐색하면서 블록내의 화소들의 밝기들을 서로 비교해 유사성을 판단한다. 이때 유사성을 비교하기 위해 두개의 비용함수(최소화·최대화 비용함수)에 의한 디스페리티맵을 비교한다. 최소화 비

용 함수인 SSD(sum of squared differences)와 최대화 비용함수인 Fast Cross Correlation(fast ZNCC)에 대한 디스페리티를 살펴보자.

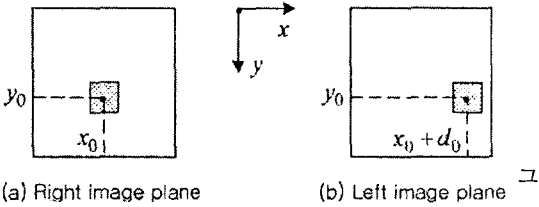


그림 2. 영역기반 정합의 기본 구조

본 연구에서는 스테레오 영상을 위해 비수렴 모델을 사용하였기에 그림 2에서와 같이 epipolar line이 같은 scan line y_0 에 존재한다. 창을 크기를 $K \times L$ 이라고 하고 변이 d 는 수평선을 따라 이동한 거리라고 하면 SSD는 다음과 같이 정의할 수 있다.

$$SSD_{i,j,d}(f,g) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} (f_{m,n} - g_{m+d,n})^2 \quad (2.1)$$

좌우 이미지 f, g 에 대해 u_f, u_g 는 창에 존재하는 밝기의 평균일때 영 평균(zero mean)으로 정규화된 상관성(cross correlation)기반의 fast ZNCC는 다음과 같다.

$$C(i,j,d) = \frac{cov_{ij,d}(f,g)}{\sqrt{var_{ij}(f)} \times \sqrt{var_{ij,d}(g)}} \quad (2.2)$$

여기에서

$$cov_{ij,d}(f,g) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} (f_{m,n} - u_{f,j})(g_{m+d,n} - u_{g,d,j}) \quad (2.3)$$

$$var_{ij}(f) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} (f_{m,n} - u_{f,j})^2 \quad (2.4)$$

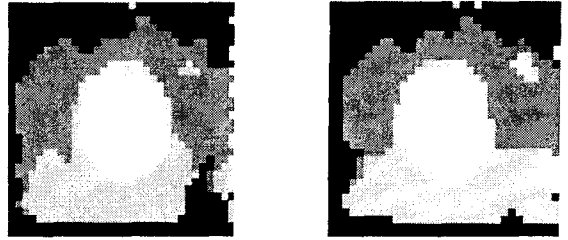
$$var_{ij,d}(g) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} (g_{m+d,n} - u_{g,d,j})^2 \quad (2.5)$$

i, j 는 이미지의 row, column을 의미하고 d 는 epipolar line을 따라 오른쪽 영상에서 창의 움직임을 말한다. 평균과 분산의 빠른 계산을 위해 box-filtering 기술을 이용하면 식 (2.3)은 (2.6)과 같이 고칠 수 있다.

$$cov_{ij,d}(f,g) = \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} f_{m,n} g_{m+d,n} - W u_{f,j} u_{g,d,j} \quad (2.6)$$

W 는 correlation window의 크기 즉, $(2K+1)(2L+1)$ 을 의미한다.

원본 영상에 대한 각각의 disparity map을 살펴보면 다음과 같다. window size는 8×8 설정하였고 탐색 범위는 좌·우 30 픽셀씩 움직였다.



(a) SSD (b) fast ZNCC

그림 3. 디스페리티 맵

각각의 비용함수에서 디스페리티를 살펴보면 SSD보다는 두 영상의 correlation에 의한 fast ZNCC가 신뢰도가 높음을 살펴볼 수 있다.

2.2. Multi-Resolution Image Pyramid

블록 매칭시 블록의 크기나 탐색 범위는 영상의 디스페리티는 결정하는 중요한 요소이다. 일정 영역의 영상 조각을 가지고 탐색 범위안에서 다른 영상과 상관관계를 찾는 방법이기 때문에 영상의 탐색 범위를 줄여 나갈 수 있다면 탐색 시간을 줄임과 동시에 더욱 정교한 디스페리티를 얻을 수 있을 것이다.

이러한 생각을 기반으로 다해상도 분해를 이용하여 탐색 범위를 줄여 보았다.

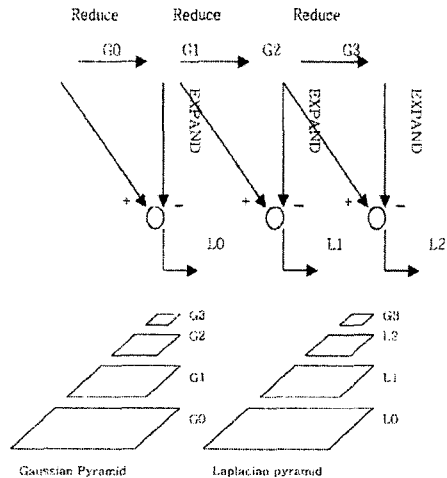


그림 4. 피라미드 다이어그램

라플라시안(Laplacian) 피라미드와 가우시안(Gaussian) 피라미드를 이용하며 sub image를 만들고 그 각각 이미지에서 대한 디스페리티 맵을 구하였다.

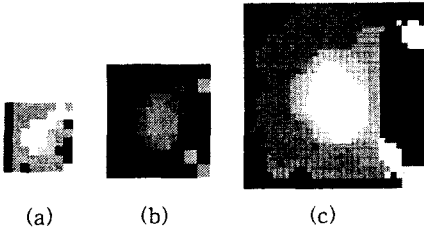


그림 5. 피라미드를 이용한 SSD

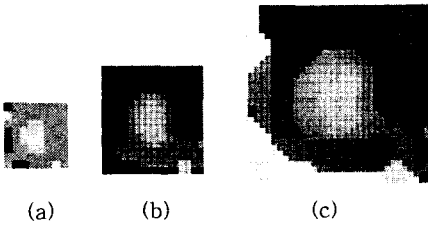


그림 6. 피라미드를 이용한 ZNCC

그림 5와 6에서 보는 바와 같이 3단계 피라미드 영상을 사용하여 디스페리티 맵을 구했다. 각각의 (a)는 그림 4의 가우시안 피라미드로 얻어진 좌우의 G2 이미지에서 찾은 디스페리티이며 (b)는 좌우의 G1 이미지에서 전 레벨 G2에서 찾은 디스페리티 (a)의 값을 이용하여 탐색범위를 줄여 상태에서 찾은 디스페리티이고 (c) 또한 역시 전 레벨 G1의 디스페리티 (b)의 값을 이용하여 탐색 범위를 줄인 상태에서 좌우의 가우시안 피라미드 G0의 영상에서 추출한 디스페리티 결과이다.

그림 5의 (a) 64*64 이미지는 좌우 양방향 30픽셀을 탐색 범위를 설정한 경우이다. (b) 128*128 이미지의 경우 탐색 범위는 (a)의 각각 블록에서 추출한 디스페리티를 탐색 기준으로 하여 좌우 양방향 17픽셀을 적용한 경우이다. (c) 256*256 경우 (b)와 마찬가지로 전 레벨의 디스페리티를 기준으로 하여 좌우 양방향 1픽셀을 탐색 범위로 적용한 경우이다. 그림 3 (a)는 원본 이미지를 양방향 30의 탐색 범위를 가졌지만 피라미드 영상은 양방향 1로써 디스페리티를 찾는 것이 가능하다는 것을 볼수 있었다. 하지만 제일 작은 영상에서 잘못된 디스페리티가 나오면 영상이 커질수록 그 에러는 증폭되어 디스페리티를 찾는데 영향을 미침을 볼수 있다.

III. 에지를 이용한 조밀한 디스페리티 구현

지금까지 보아온 모든 디스페리티 맵을 살펴보면 8*8의 일정한 블록의 상관 관계에 의한 정합으로 인하여 디스페리티 맵이 부드럽지 못하고 각이지는 blocking 현상을 볼 수가 있다. 이러한 현상을 없애주기 위해 원본 영상의 에지와 구해진 디스페리티 맵의 에지(edge)를 비교하여 blocking이 일어나는 부분에서 보다 작은 블록을 설정하여 디스페리티를 다시 구한다.



그림 7. blocking 현상이 일어나는 곳의 에지정보

위에서 구한 에지를 찾기 위해 Robert filter를 사용하였다. 그림 7의 (a)는 원본 왼쪽 이미지의 에지 (b)는 디스페리티 맵의 에지 (c)는 (a)와 (b)를 이용하여 blocking이 일어나는 곳을 찾은 부분이다.

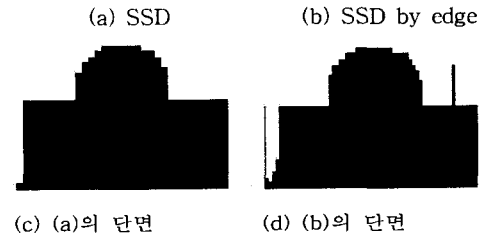
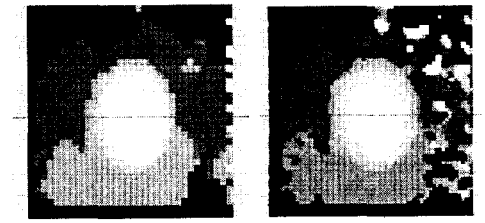
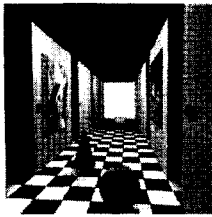


그림 8. 에지 정보 활용 유무에 따른 디스페리티 맵과 히스토그램

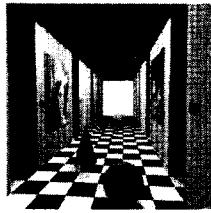
그림 8의 (a)는 블록을 8*8로 고정된 상태에서의 디스페리티 맵이며 (b)는 원본 디스페리티 맵에서 에지 정보를 이용하여 blocking 현상이 일어나는 곳을 찾고 그곳에서 블록을 4*4로 재설정하여 찾은 디스페리티를 적용시킨 결과이다. (c)에 나타난 단면을 살펴보면 (b)의 영상이 훨씬 조밀한 공의 표현을 하고 있다는 것을 확인할수 있다.

IV. 실험 및 결과 분석

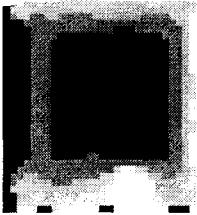
다른 이미지에서 블록의 크기를 조절하면서 디스페리티가 어떻게 변하는지를 살펴보고자. 실험한 이미지는 corridor image와 pentagon image이며 블록의 크기는 8*8과 4*4이다.



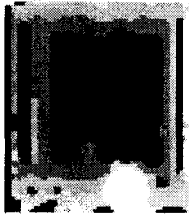
(a) corridor left



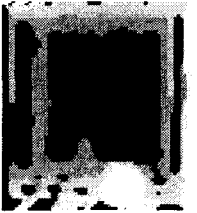
(b) corridor right



(c) 8*8 SSD



(d) 8*8 SSD by edge



(e) 4*4 SSD



(f) 4*4 SSD by edge

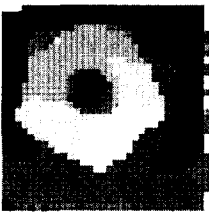
그림 9. corridor door image의 디스페리티



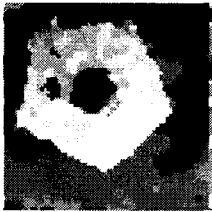
(a) pentagon left



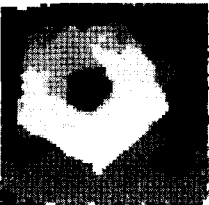
(b) pentagon right



(c) 8*8 SSD



(d) 8*8 SSD by edge



(e) 4*4 SSD



(f) 4*4 SSD by edge

그림 10. pentagon image의 디스페리티

V. 결론

본 논문에서 디스페리티의 blocking현상을 없애기 위해 조절된 블록을 가지고 비용을 구하는 방법을 제시하였다. 이러한 조절된 블록을 위해 원본 영상과 블록 정합으로 얻어진 디스페리티 맵 사이의 에지 정보를 이용하는 방법을 제안하였고 그 방법으로 실험한 결과 blocking이 생기는 부분에 좀더 부드러운 영상을 얻을 수가 있었다.

추후의 연구 과제는 매칭하는 속도와 신뢰도를 높이기 위해서 이미지 피라미드에서 구한 각각의 디스페리티에 blocking이 일어난 부분을 찾아 좀더 세밀한 블록을 설정하여 재과정을 하는 부분을 생각할 수 있다. 또한 전체적인 영상의 분포를 감안하여 분산을 통해 좀더 세밀화된 블록을 설정하여 디스페리티를 구하는 방법을 적용한다면 더욱 부드러운 디스페리티를 구할 수 있을 것이다.

참고문헌

- [1]Peter J. Burt and Edward H. Adelson, "The Laplacian Pyramid as a Compact Image Code", *IEEE Transactions on Communications*, April, 1983
- [2]Takeo Kanade and Masatoshi Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment", *IEEE Trans. on Pattern analysis and Machine Intelligence*, September, 1994
- [3]김병우 "다해상도 영상피라미드를 이용한 조밀한 디스페리티 맵 추출", *고려대학교 대학원 공학석사 학위논문*