에지 적응형 방법을 이용한 스테레오 정합에 관한 연구

한기서* · 염진수* · 허창우* · 류광렬*

*목원대학교

A Study on the Stereo Matching which using an Edge-Adaptive Method

Ki-seo Han* · Jin-su Yeom** · Chnag-wo Hur*** · Kwang-ryol Ryu****

* Mokwon University

E-mail : hanki0624@nate.com

요 약

스테레오 정합을 하기 위해서 먼저 두 이미지의 밝기를 정규화 하였고, 밝기가 정규화 된 영상을 이용하여 Canny방식으로 Edge를 검출한 후 에지 적응형 스테레오 정합 방식으로 두 영상을 정합하여 변이도 영상들을 얻었다. 그 중 Canny Edge 방식이 다른 Edge 검출 방식보다 선명한 영상을 추출했다. Canny Edge 방식으로 획득한 스테레오영상을 중위수 필터를 사용하여 잡음을 제거한 후, 필터링 전·후의 영상을 PSNR(Peak Signal to Noise Ratio)로 비교하여 본 결과 개선된 영상을 획득하였다.

키워드

Edge, Stereo Matching, Disparity, Canny

Ⅰ. 서 론

컴퓨터 비전은 영상장비로 획득한 이미지를 컴퓨터를 사용하여 분석하고 해석하는 학문이다. 컴퓨터 비전은 이미지처리, 패턴인식, 사진측량으로 크게 세 가지 분야로 나눌 수 있다. 첫째, 이미지처리는 원본 이미지로부터 보다 좋은 이미지를 얻어내는 이미지 강화, 이미지 데이터 전송 및 저장을 위한 이미지 압축, 원본 이미지로부터 현재 알고 있는 필요하지 않은 정보를 제거하는 복원, 이미지 특징들을 추출하는 과정이 이에 해당된다. 둘째, 패턴인식은 기하학적인 특징을 이용하여 이미지를 여러 유형별로 분류하거나 정해진 특징을 가진 이미지를 찾는 기술을 말한다. 셋째, 사진측량은 비접촉 이미지로부터 정확하고 신뢰할 만한 데이터를 얻는 것을 말한다. 본 논문은 스테레오 카메라를 이용하여 이미지를 획득하고 그 이미지를 이용한 스테레오 비전에 대하여 다루려고 한다.

우리가 실물을 입체적으로 볼 수 있는 것은 두 눈의 양안시차(Binocular Disparity)에 의한 것이다. 이는 두 눈이 다른 시점의 영상을 동시에 인식하고 인간의 뇌가 이 두 영상의 차이를 이용하여 입체 영상을 효과적으로 합성, 인식하기 때문이다. 눈으로 보는 실제의 이미지는 3차원 실세계를 두 개의 2차원 평면에 투영시킨 모습인데, 사실은 우리의 두뇌에서 이러한 두 개의 평면 영상을 이용해 입체감을 느끼게 되는 것이다. 이렇게 평면 영상으로부터 3차원 영상을 얻어내는 변환을 원근적 변환(perspective transformation), 또는 이미징 변환(imaging transformation)이라고 한다. 변환은 다-대-일(many-to-one) 변환이기 때문에 일반적으로 하나의 이미지만 가지고는 원래의 3차원 세계를 복원하는 것이 거의 불가능하다.

그러나, 인간이 두 개의 눈으로 3차원을 인식하듯이 공간적으로 떨어진 위치에 있는 두 개의 카메라를 가지고 동일한 장면으로부터 이미지 쌍을 만들어낸다면, 두 이미지에서 볼 수 있는 3차원 세계의 부분들은 다시 복원할 수 있다. 이러한 과정을 스테레오 비전(stereo vision)이라고 한다. 스테레오 비전을 컴퓨터로 구현하는 일은 지난 30년간 많은 연구가 진행됐음에도 불구하고 쉽게 해결되지 않고는 있지만 실현만 된다면 인간의 시각 시스템과 같이 보행, 항해 등에 사용할 수 있음은 물론이고, 다양한 분야에 적용할 수 있을 것이다. 최근 연구에서는 인간이 3차원을 인식하는 원리인 자연적인 제약 조건들을 이용하는 스

테레오 비전 구현을 위한 고속 병렬 알고리즘들이 연구되고 있다.

Ⅱ. 에지 검출에 의한 스테레오 정합 제안

1. 시스템 구성

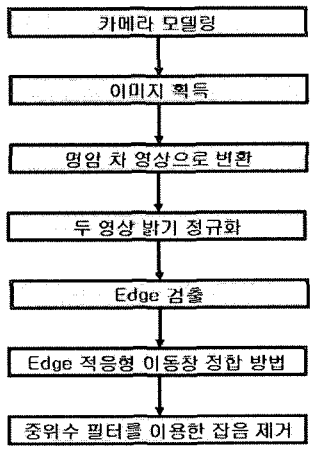

그림 1. 시스템 흐름도

2. 두 영상간의 밝기 정규화

영상은 날씨, 촬영 시각과 같은 획득 당시의 환경에 의하여 전체 픽셀의 밝기에 차이가 생길 수 있으며, 밝기의 차이에 의하여 히스토그램의 패턴도 영향을 받게 된다. 광원에 의한 픽셀의 밝기가 달라질 경우 히스토그램의 분포위치만 이동하게 된다. 히스토그램의 모양이 유사하더라도 분포 위치가 변하면 상관 계수나 변이 계수 측정 시 많은 오차가 발생된다. 그러므로 영상 밝기 정규화 과정을 통하여 히스토그램 분포의 위치를 동일하도록 만들어줄 필요가 있다. 매칭의 대상이 되는 좌·우 영상 간에 밝기 정규화를 시킬 때 어떤 쪽의 영상을 기준으로 정규화를 하더라도 상관이 없다. 두 영상의 밝기 조절은 픽셀의 밝기에 일정한 값을 더하거나 빼는 과정을 뜻하는 것이다.

Left 영상의 표준화는 식(2.1)과 같이 표현한다.

$$f' = (x,y)\frac{(f_L(x,y)-m_L)}{\sigma_L} \qquad \text{식(2.1)}$$

Left 영상과 마찬가지로 Right 영상의 표준화도 식(2.1)과 같이 표현 한다.

$$f' = (x,y)\frac{(f_R(x,y)-m_R)}{\sigma_R} \qquad \text{식(2.2)}$$

Left 영상과 Right 영상 표준화가 같다고 놓는다.

$$\frac{(f_L(x,y)-m_L)}{\sigma_L} = \frac{(f_R(x,y)-m_R)}{\sigma_R} \qquad \text{식(2.3)}$$

Left 영상을 기준으로 하면 두 영상의 밝기 분포에 대한 정규화 과정은 식(2.4)과 같다.

$$f_L(x,y) = m_L + (f_R(x,y)-m_R)*\frac{\sigma_L}{\sigma_R} \qquad \text{식(2.4)}$$

히스토그램 정규화는 전체적인 밝기 분포만을 고려하기 때문에 조명의 밝기가 변화할 때에는 그림자 영역과 밝은 영역의 대비를 완화시켜 좋은 결과를 나타낼 수 있다

3. 에지 적응형 스테레오 정합 방법

물체의 경계선 부근에서는 가려진 영역이 발생하기 때문에 한 화소의 Local support를 설정할 때 깊이 값의 불연속선을 포함하지 않도록 잡아줄 필요가 있다. 이를 위한 방법으로 이동 창 정합방법에서는 다수의 윈도우를 이용한 것이다. 어떤 화소의 왼쪽에 불연속선이 있을 때에는 정합창을 오른쪽으로 이동시켜 설정한다. 다른 경우에도 마찬가지 원리를 적용하면 된다. 그런데 불연속선이 존재하는지 여부를 미리 알 수가 없으므로 모든 이동창에 대해 전부 정합을 시도하고, 그 중에서 가장 작은 오차를 보이는 것을 선택함으로써 최적의 이동창과 깊이 값을 구할 수 있다. 단 계산량이 이동창의 개수에 따라 증가하는 것이 단점이다. 따라서 이동창을 사용하는 것과 같은 효과를 얻으면서 계산량을 획기적으로 줄이는 노력이 필요하다. 본 논문에서는 에지의 위치를 파악하여 에지에 따라 이동창의 종류를 알맞게 선택하는 방법을 제안한다.

그림 2를 살펴보면 각 에지의 특성에 따라 짝지여진 정합창이 특정한 규칙에 의해서 짝지여졌다는 것을 알 수가 있다. 즉 정합창 내에서 에지 성분의 화소의 값을 포함하지 않도록 하였으며 가려진 영역이 생기게 되는 부분에 대해서만 중점적으로 정합을 실행하도록 되어 있다. 이러한 방식에 의해 정합 시 다양한 이동창을 적응적으로 사용하면서도 알고리즘의 수행시간을 단축시키고, 잘못된 정합 값이 추출되는 확률을 줄이게 된다.

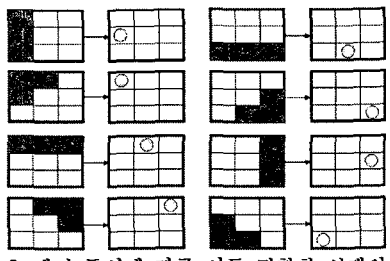

그림 2. 에지 특성에 따른 이동 정합창 선택의 예시

4. 중위수 필터를 이용한 잡음 제거

획득하고자 하는 영상은 영상 자체에 잡음이 있거나 처리과정에서 잡음이 섞이는 등으로 인해 좋지 못한 결과를 보이게 된다. 잡음 제거는 손상 부분의 영향을 최소화하기 위한 작업이다. 즉, 영상의 손상에도 불구하고 가장 이상적인 이미지를 대상으로 얻을 수 있는 결과에 근접하기 위한 과정이다. 잡음은 그 특성상 주변 픽셀의 픽셀 밝기와 급격한 픽셀의 밝기 차를 보이게 된다. 이러한 잡음의 성질을 이용하여 영상내의 잡음을 제거하는 방법으로 중위수 필터를 이용하여 잡음을 제거하였다.

2	2	2
4	10	3
5	4	4

2	2	2
4	P	3
5	4	4

2 2 2 3 4 (4) 4 5 10

그림 3. 중위수 필터에 의하여 픽셀의 밝기를 구하는 방법

그림 3은 영상에서 3x3 영역내의 9개의 픽셀을 오름차순으로 정렬한 결과를 나타내고 있다.

9개의 픽셀 중, 왼쪽에서 5번째 픽셀의 밝기 4가 중위수이다. 즉, P는 4의 값을 갖는다. 그림 3에서 픽셀의 밝기 10을 갖는 픽셀은 잡음성분이다. 이 잡음 성분을 4로 변경되면서 잡음이 제거되었다. 주변과 비교하여 극단적으로 농도의 차이가 잇는 것은 크기의 순서로 나열했을 때, 좌나 우로 치우치게 되어 중위수로 선택되지 않기 때문이다. 이와 같이 중위수 필터는 어떤 픽셀 주변의 영역 내에 픽셀 농도의 중간 값을 구하여 원하는 픽셀의 밝기로 처리하는 것이다.

Ⅲ. 실험 및 고찰

본 논문에서 스테레오 정합 프로그램의 성능을 실험하기 위한 영상으로 Tsukuba 대학에서 제공하는 영상을 이용하였다. 이 영상은 스테레오 카메라를 이용한 영상과 깊이 맵을 제공하고 있기 때문이다.

표 1. 실험 환경

Hardware	CPU:Intel dual 2.8 Ghz RAM : 2GB
Software	Visual C++ 6.0 Matlab 7.0
Image	Pentagon image -Resolution:512 x 512 -Max disparity:16 Pixel

스테레오 정합의 성능을 테스트하기 위해 Tsukuba 대학에서 제공하는 펜타곤 영상을 이용하였다.

에지 적응형 스테레오 정합 방법을 이용하여 다음과 같이 하였다.

라플라시안(Laplacian), 라플라시안의 단점을 보안하기 위하여 가우시안의 라플라시안 여파기(LoG) 그리고 캐니(Canny) Edge 검출 방법으로 서로 비교하였다.

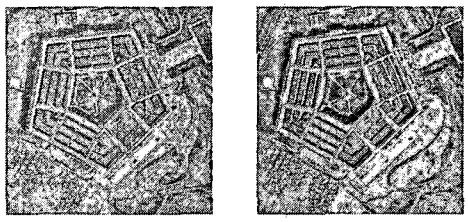

그림 4. 입력 좌·우 영상

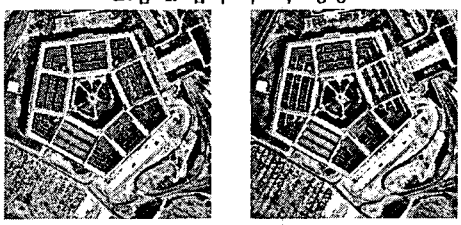

그림 5. 밝기 정규화 후의 좌·우 영상

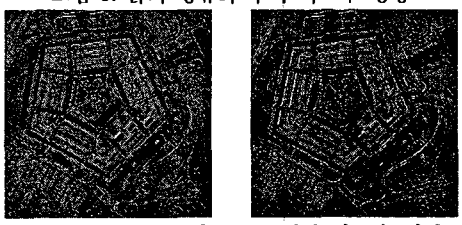

그림 6. Laplacian 마스크를 취한 좌·우 영상

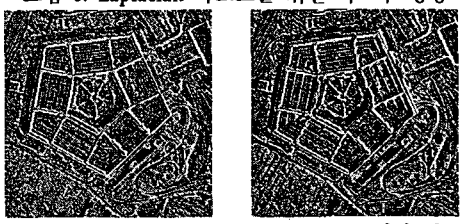

그림 7. Laplacian of Gaussian 마스크를 취한 좌·우 영상

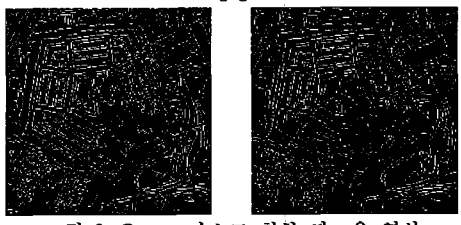

그림 8. Canny 마스크 취한 좌·우 영상

그림 8은 캐니 마스크 방법에 의한 결과이다.

에지 검출 방법은 이미지픽셀의 캐니 마스크 방법이 세밀하지 못한 라플라시안, LoG와 DoG 방법 보다 검출 면을 확인하면 더 선명하게 에지가 검출된 것을 육안으로 확인 할 수 있다. 그러나 캐니 마스크 방법이 다른 방법에 비해 객관적인 영상의 비교 기준으로는 적합하지 않다. 따라서 성능 평가 지수로 입력 스테레오 좌·우영상과 에지 적응형 스테레오 정합으로 얻은 스테레오 정합 영상들의 PSNR(Peak Signal to Noise Ratio)를 비교하였고 사용된 식은 아래 식으로 표현할 수 있다.

$$PSNR = 20\log_{10}\left(\frac{NM(fmax - fmin)}{\sum_{y=0}^{m-1}\sum_{x=0}^{N-1}(f(x,y) - f^*(x,y)2)}\right)$$

여기서 $fmax$와 $fmin$는 입력 영상의 에너지 평균값이고 $f(x,y)$와 $f^*(x,y)$은 입력영상 평균값과 스테레오 정합 영상을 나타내며 x,y는 영상의 열(Row)과 행(Column)을 나타낸다.

그림 9은 Laplacian Edge 검출을 이용하여 스테레오 정합을 한 결과 영상이다. 잡음이 많고 영상이 뚜렷하지 않은 것을 볼 수 있다. 그림 10에서도 Laplacian 보다는 좋은 영상을 얻었지만 Canny보다 선명하지 못한 것을 알 수 있다.

그림 9. Laplacian 정합 영상

그림 10. Laplacian of Gaussian 정합 영상

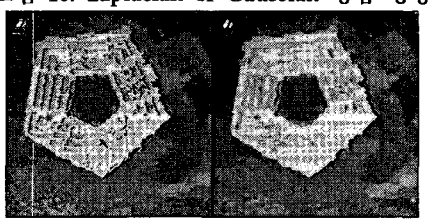

그림 11. Canny 정합 영상

그림 12. 중위수 필터를 사용한 결과

표 2. 입력 영상과 스테레오 정합 영상들의 PSNR

	영상 정합 결과 영상		
	라플라시안	LoG	Canny
PSNR (dB)	20.07	20.81	21.88

각 영상을 비교한 결과 Canny Edge 검출을 통한 스테레오 정합이 선명하게 나왔음을 알 수 있다. 또한 표 2은 영상 정합으로 얻은 영상들의 PSNR 값을 나타낸다. 비교 결과 캐니 마스크 방법이 라플라시안 방법에 비해 약 1.81dB이 향상되었고, LoG 방법에 비해 약 1.07dB이 향상되었다.

마지막으로 잡음 제거를 위하여 제안된 방식으로 정합하여 얻어진 영상에 중위수 필터를 적용한 결과 그림 11과 같이 잡음이 많이 없어지고 보다 선명한 영상을 얻게 되었다. 중위수 필터를 사용하여 잡음 제거 후 PSNR값이 22.71dB을 얻었다. 캐니로 정합 후 검출된 영상보다 중위수 필터로 제거한 영상이 약 0.83dB 정도 향상 되었다.

Ⅳ. 결 론

본 논문에서는 스테레오 영상의 정합을 위한 기존의 에지 검출을 이용하여 특징 점을 검출하고 검출된 특징 점을 이용하여 경계선 및 영역 정보를 이용한 에지 적응형 스테레오 정합을 제안하였다.

기존의 다수의 이동창을 이용한 방법은 한 화소의 깊이 값을 산출하기 위해 모든 이동창을 이용하여 수행시간이 매우 긴 단점이 있는 반면, 에지 적응형 정합 방법을 사용한 경우 적절한 하나의 정합창만을 이용하기 때문에 수행시간이 단축된다. 여러 가지 에지 검출 방법을 이용하여 스테레오 정합 결과 캐니 검출에 의한 정합 방법이 다른 에지 검출 방식보다 0.47dB~1.81dB정도 향상된 깊이 값을 산출함을 볼 수 있었다. 그리고 캐니 검출에 의한 정합을 한 후, 중위수 필터로 잡음을 제거한 결과 기존 캐니 검출에 의한 정합 방법보다 0.83dB정도 향상된 영상을 얻을 수 있었다.

스테레오 정합은 컴퓨터 비전에서 중요한 요소인 3차원 정보를 추출하는 알고리즘이다. 스테레오 매칭 방법으로 얻어진 영상을 이용하면 영상합성이나, 화상 전송 그리고 게임 시스템에서의 사용자 인식, 물체 인식, 보안 시스템 로보틱스와 같은 컴퓨터 비전 분야에 적용될 수 있으며, 앞으로의 발전 가능성은 무한하다 할 수 있다.