

모노-스테레오 영상을 이용한 스테레오 시차 추출

종신회원 이재수*, 정회원 홍권의**, 종신회원 안정근***

Stereo Disparity Extraction using a Mono-Stereo Image

Jae-soo Lee* *Lifelong Member*, Kwon-eui Hong** *Regular Member*,
Jeong-keun Ahn*** *Lifelong Member*

요약

본 논문에서는 평행식 스테레오 카메라 모델의 중앙에 카메라 렌즈를 하나 더 둔 단일 카메라의 모노-스테레오 시스템을 이용하여 스테레오 시차를 추출하였다. 모노 영상은 좌, 우 스테레오 영상의 중심점 역할을 하여 영상처리 시 기준 영상으로 사용하였고, 모노 영상을 기준 영상으로 좌, 우 스테레오 영상을 정합하여 그 결과를 스테레오 시차로 제시하였다. 실험결과 좌, 우측 영상간 스테레오 시차의 계산값과 실험치가 거의 일치하였고, 따라서 단일 카메라의 모노-스테레오 시스템이 스테레오 물체 추적기, 자동 이동로봇 시스템 및 다시점 3D 디스플레이 등에 응용 가능할 것이다.

Key Words : Mono-Stereo Camera, Stereo Vision System, Stereo Matching, Stereo Disparity, LoG, MAD

ABSTRACT

In this paper, we extracted stereo disparity using mono-stereo system of single camera added a camera lens on the parallel stereo camera model. We used the mono image as the reference image, because it plays a role in image processing of the left and right stereo image, and we show a stereo disparity as a result of the stereo matching using the mono image changes into the reference image. From the experimental results, the calculating value and the experimental value of the stereo disparity for the left and right image are almost the same, and therefore, in case of implementation the mono-stereo system of a single camera, the stereo object tracker, the automatic moving robot system and the multi-views 3D display system will be applied practical application.

I. 서론

사람의 시각 체계는 서로 다른 방향의 두 눈에서 획득한 영상을 적절히 조합하여 뇌에 전달하고, 정보처리 과정을 거쳐 물체의 깊이감, 거리정보 및 형상 등 3차원 정보를 획득하여 사람들에게 입력되는 영상들을 입체적으로 인식하게 해준다.

이와 같이 사람의 시각 체계를 모방하여 만든 것이 스테레오 비전 시스템으로, 인간 두 눈의 기하학

적 특성을 고려하여 3차원 공간상에 설치된 카메라로부터 얻어진 좌, 우 영상으로부터 3차원 정보를 획득하여 영상을 처리하게 된다.

그림 1은 인간의 시각 시스템을 스테레오 비전 시스템으로 모델링하여 비교한 것이다.

스테레오 비전 시스템을 로봇이나 컴퓨터 비전 및 입체 시스템 등에 응용하기 위해서는 3차원 정보를 구하는 것이 가장 중요하다. 즉, 스테레오 영상인 좌, 우 영상에서 상호간의 대응점인 정합점을

* 본 논문은 2008학년도 김포대학의 연구비 지원에 의하여 연구되었습니다.

* 김포대학 정보통신과 부교수(jslee@kimpo.ac.kr), ** 김포대학 정보통신과 부교수 ***김포대학 멀티미디어과 부교수
논문번호 : 08057-0830, 논문투고일자 : 2008년 8월 30일

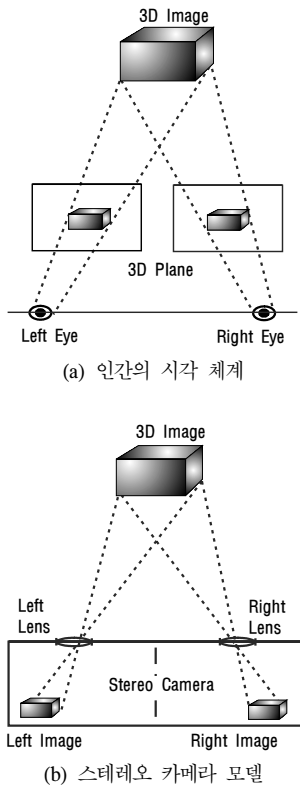


그림 1. 인간의 시각체계와 스테레오 카메라
Fig. 1. Human vision system and stereo camera

찾아내고, 정합점으로부터 3차원 정보를 얻게 된다. 이때 좌, 우 스테레오 영상으로부터 대응점을 추출하는 정합과정이 가장 어렵고 중요하다¹¹.

일반적으로 스테레오 정합 결과에서 얻어진 정합점의 상대 거리를 변위(disparity)값으로 나타낼 수 있고, 이 값으로부터 물체까지의 거리(distance)나 깊이(depth)와 같은 3차원 정보를 구하게 된다.

스테레오 정합 방법에는 여러 가지 알고리즘이 사용되고 있지만, 기존 방법들은 크게 특징기반 방식(features based matching)과 영역기반 방식(areas based matching)으로 분류할 수 있다.

특징기반 방식은 정합과정에서 영상의 밝기 대신 스테레오 영상에서 추출한 특징 점들을 사용하여 정합 가능점들을 찾은 다음, 물리적 현상에 바탕을 둔 몇 가지의 제한 조건들을 적용하여 정합 가능점 중에서 이 조건들을 가장 잘 만족 시키는 점을 찾음으로써 두 영상을 정합하는 방식이다²⁻⁴.

영역기반 방식은 정합하려는 곳을 포함하는 일정한 크기의 기준 영역을 설정하여 이 영역을 상대 영상의 탐색범위 내에서 이동시키면서 상호 연관

정도를 구해 이 연관성이 최대가 되는 점을 찾음으로써 변위 값을 구하는 방식이다⁵.

정합 알고리즘으로 사용되고 있는 것들은 MAD (mean absolute difference), SAD(sum of absolute difference)^{6,7}, MSE(mean squared difference), SSD(sum of squared difference)^{8,9} 등이 있다.

본 논문에서는 정합의 정밀도를 높이기 위해 단일 카메라의 모노-스테레오 영상을 이용하여 스테레오 영상과 중앙 영상인 모노영상, 즉, 스테레오 영상뿐만 아니라 스테레오 카메라의 중앙 영상을 획득하였다. 이 모노 영상을 기준 영상으로 좌, 우 스테레오 영상을 정합하여 그 결과 값으로 스테레오 시차를 추출하였다.

II. 모노-스테레오 비전 시스템

일반 스테레오 비전 시스템의 카메라 모델은 그림 2와 같이 교차식과 평행식의 2가지 방법이 있다.

이 두 가지 방식 중 가장 간단한 방식은 평행 카메라 방식으로 평행한 렌즈의 광축과 평행한 카메라의 광축을 가진 두 카메라를 이용하는 방식이다.

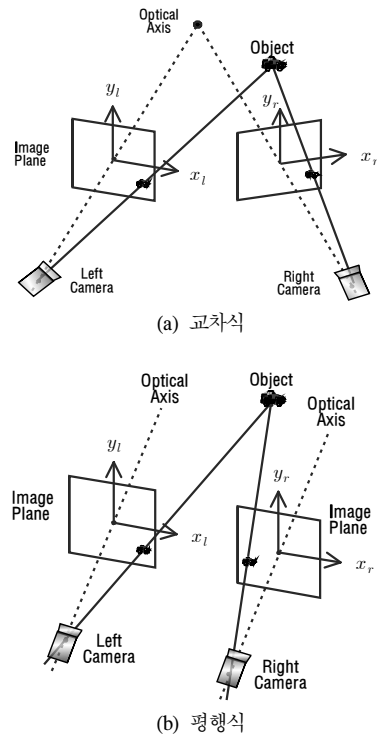


그림 2. 스테레오 카메라 모델
Fig. 2. Stereo camera model

이 방법에서 얻어지는 좌, 우 영상 사이의 공통 시역은 광축 사이의 거리에 따라서 매우 작아지며, 수직 시차가 없기 때문에 시차 방정식이 간단하여 영상처리에 편리하다. 교차식 방식은 두 카메라 축과 렌즈 축이 교차하도록 서로를 향해 안쪽으로 회전하는 카메라를 가지고 있어 소위 주시점이라 불리는 곳에서 모이게 된다. 이 방식은 카메라들 사이의 공통 시역을 최대화 할 수 있기 때문에 실제로 많이 사용되고 있지만, 수평과 수직 시차 방정식이 복잡하고 수학적인 분석이 어렵다.

그림 3, 4의 단일 카메라의 모노-스테레오 모델은 일반 평행식 스테레오 카메라 모델의 중앙에 카메라를 하나 더 두어 중앙에서 모노로 촬영한 영상을 이용하여 영상처리를 할 수 있도록 한 시스템이다.

단일 카메라에서 스테레오 영상과 중앙의 모노 영상을 얻을 수 있기 때문에 카메라를 2개 사용하는 스테레오 시스템에서 발생하는 좌, 우 영상의 캘리브레이션 문제 및 명암도 문제 등에서 유리하다.



그림 3. 단일 카메라의 모노-스테레오 렌즈
Fig. 3. Mono-stereo camera lens of single camera

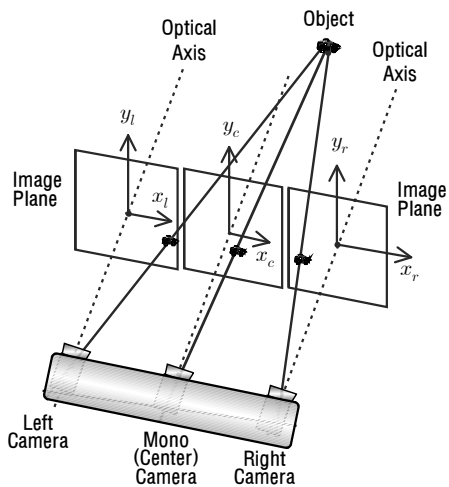


그림 4. 모노-스테레오 카메라 모델
Fig. 4. Mono-stereo camera model

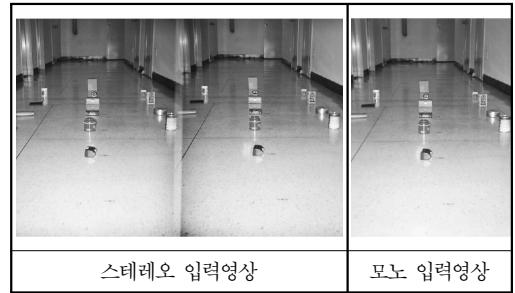


그림 5. 단일 카메라의 모노-스테레오 입력 영상
Fig. 5. Mono-stereo input images of single camera

그림 5는 단일 카메라의 모노-스테레오 시스템으로부터 획득한 스테레오 영상과 모노 영상을 보여 주고 있다. 이 모노-스테레오 입력 영상은 그림 8과 같이 중앙의 모노 영상과 좌, 우 영상으로 분리하여 스테레오 영상으로 사용하게 된다. 그림 5는 단일 카메라의 모노-스테레오 시스템으로부터 획득한 스테레오 영상과 모노 영상을 보여 주고 있다. 이 모노-스테레오 입력 영상은 그림 8과 같이 중앙의 모노 영상과 좌, 우 영상으로 분리하여 스테레오 영상으로 사용하게 된다. 이와 같이 평행식 단일 카메라로부터 획득한 중앙 영상인 모노영상과 좌측영상 및 우측영상을 스테레오 영상처리에 사용함으로써 좌, 우 영상의 명암도 문제, 캘리브레이션 문제 및 수직시차 문제 등에 영향 없이 영상처리시 간편하게 처리할 수 있다. 이렇게 함으로써 스테레오 영상으로부터 스테레오 정합의 정밀도를 개선할 수 있으며, 모노영상과 스테레오 영상을 이용하면 쉽게 다수의 중간영상을 생성할 수 있으므로 다시점 디스플레이 시스템에도 적용할 수 있을 것이다. 또한, 스테레오 물체 추적시 모노영상을 기준영상으로 사용하여 좌, 우 스테레오 영상의 추적물체 추출시 정확하게 추출할 수 있으므로 스테레오 물체 추적 등에도 적용할 수 있을 것이다.

III. 모노-스테레오 영상의 정합

모노-스테레오 영상 정합에서 우선, 정합의 정밀도를 개선하고, 스테레오 영상 간 동일한 배경을 제거하기 위해 LoG(Laplacian of Gaussian) 필터를 이용하여 스테레오 영상과 모노 영상의 에지 영상을 구하였다. 그리고 이 에지 영상들을 이용하여 모노 영상과 스테레오 영상간 정합점을 구하고, 이 정합 값인 변위 값으로 스테레오 시차를 구하였다.

3.1 모노-스테레오 영상의 에지 추출

영상에서 물체를 구분하기 위해서는 우선적으로 배경 부분은 무시되고, 각 물체의 경계선에 의해서 물체를 쉽게 인식하게 된다. 즉, 획득한 영상에 미분을 행하면 경계선 부분에서 고 주파 잡음과 함께 에지 성분이 생기게 되고, 이 에지 영상에 가우시안 필터를 적용하면 고 주파성분의 잡음을 제거할 수 있다.

이와 같이 가우시안과 미분특성을 이용하여 에지 정보를 추출하게 되는데 이 방법이 LoG 방법이다.

우리가 획득한 영상은 2차원 영상 $I(u, v)$ 이므로 표준편차를 이용하여 식 (1)과 같이 2차원 가우시안 분포로 나타낼 수 있다.

$$I(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{(u^2 + v^2)}{2\sigma^2}} \quad (1)$$

여기에서 σ 는 표준편차로 가우시안 분포의 크기를 변화 시킬 수 있으며, 라플라시안 후에 에지의 크기를 결정하는 성분이 될 수 있다.

식 (1)에 2차 미분인 라플라시안을 적용(LoG)하면, 식 (2)와같이 되고, 이 수식으로부터 영상의 변화가 심한부분의 에지성분을 추출할 수 있게 된다.

$$\nabla^2 I(u, v) = \frac{\partial^2 I(u, v)}{\partial u^2} + \frac{\partial^2 I(u, v)}{\partial v^2} = \frac{1}{2\pi\sigma^2} \left(\frac{u^2 + v^2}{\sigma^2} - 2 \right) e^{-\frac{(u^2 + v^2)}{2\sigma^2}} \quad (2)$$

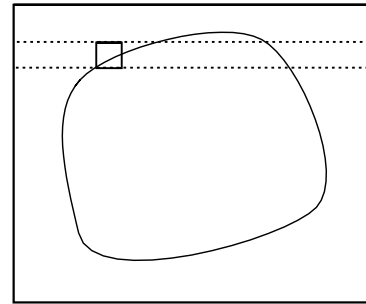
식 (2)에서 표준편차 σ 의 값이 커지면 얇은 에지를 추출할 수 있고, 작아지면 가는 에지를 추출할 수 있다. 그리고 에지성분 이외에 고립된 잡영들이 존재하는데 이는 화소귀납 알고리즘의 다중회귀 분석법^[10]을 이용하여 제거하였다.

3.2 모노-스테레오 영상의 정합

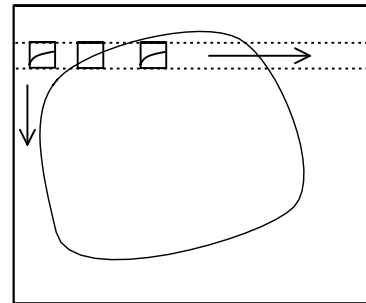
영역기반의 정합에서 영상의 명암이 뚜렷하거나 형상의 변위가 연속적으로 완만하게 변하는 경우가 가장 우수하며, 계산 과정이 국수적으로 이루어 질 수 있도록 정합의 밀도를 임의로 조절 할 수 있다.

본 논문에서는 에지 성분을 추출한 영상을 통하여 정합하기 때문에 명암의 변화가 뚜렷하므로 정확하게 매칭 되어 쉽게 정합점을 찾을 수 있다. 정합을 위한 윈도우 대응 블록은 임의의 크기로 정하여 정합시킬 수 있으며, 이 대응 블록은 비용함수를 최소로 하는 블록을 선택하면 된다.

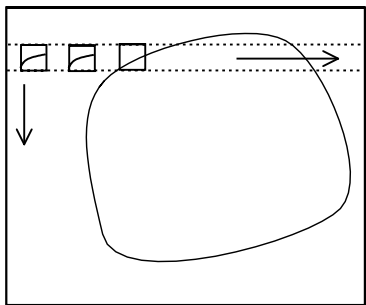
그리고 그림 6과 같이 두 영상의 특정 영역간 유사성으로 대응점을 찾는 정합은 에피폴라 선상에서



(a) 모노 영상



(b) 우측 영상



(c) 좌측 영상

그림 6. 에피폴라 선상의 탐색 영역
Fig. 6. Search region on the epipolar line

이루어지며, 모노 영상의 대응 블록인 $(N_x \times N_y)$ 의 작은 윈도우와 좌, 우 영상의 동일한 에피폴라 선상의 탐색 영역에서 각 화소에 대한 특성을 조사하여 대응점을 찾아 정합점을 구하게 된다.

본 논문에서 비용 함수는 식 (3), (4), (5)의 MAD를 사용하였으며, 대응 블록은 3×3 의 윈도우를 사용하였다.

$$MAD_{LC} = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I_C(m, n) - I_L(m + x, n + y))^2 \quad (3)$$

$$MAD_{CR} = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I_C(m, n) - I_R(m + x, n + y))^2 \quad (4)$$

$$MAD_{LR} = \frac{1}{M \cdot N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (I_L(m, n) - I_R(m + x, n + y))^2 \quad (5)$$

여기서 MAD_{LC} , MAD_{CR} 및 MAD_{LR} 은 각각 좌측 영상과 중앙(Center)의 모노 영상간 MAD값, 모노 영상과 우측 영상간 MAD값 및 좌측영상과 우측 영상간의 MAD값이고, $I_C(m,n)$, $I_L(m,n)$ 및 $I_R(m,n)$ 은 각각 중앙의 모노영상, 좌측영상 및 우측영상의 빛의 세기인 명암 값이다. MAD는 변위의 변화에 따른 두 영상의 밝기 값에 대한 차를 구한 후, 절대값에 대한 합을 구하는 방식으로 에피폴라 선상에서 수행되며, 하나의 화소에 대한 정합점의 값은 MAD가 최소로 되는 값을 택하여 정하게 된다.

3.3 모노-스테레오 영상의 시차 추출

스테레오 비전 시스템에서 정확한 스테레오 정합을 통해 스테레오 시차를 추출하는 것은 중요하다, 왜냐하면 이 시차 정보는 3차원 정보으로써 거리정보, 깊이정보 및 스테레오 디스플레이를 위한 영상합성 정보 등에 사용하기 때문에 매우 중요하다.

그리고 그림 4의 세 카메라 렌즈 모델을 그림 7과 같이 세 영상을 하나의 영상 평면에 나타내어 시차 정보를 나타낼 수 있다.

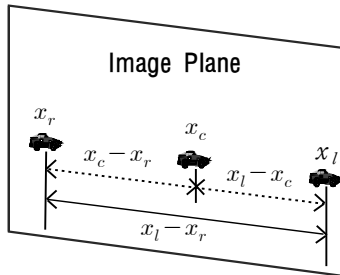


그림 7. 한 영상 평면의 스테레오 시차
Fig. 7. Stereo disparity of single image plane

이 스테레오 시차를 수식으로 나타내면 식 (6), (7), (8)과 같이 나타낼 수 있고 이 수식을 이용하여 각각의 시차를 구할 수 있다.

$$D_{lc} = |(x_l - x_c)| \tag{6}$$

$$D_{cr} = |(x_c - x_r)| \tag{7}$$

$$D_{lr} = |(x_l - x_r)| \tag{8}$$

여기에서 D_{lc} 은 왼쪽 영상과 중앙의 모노 영상간의 시차이고, D_{cr} 은 모노 영상과 오른쪽 영상간의 시차이다. 그리고 D_{lr} 은 왼쪽 영상과 오른쪽 영상간의 시차이다.

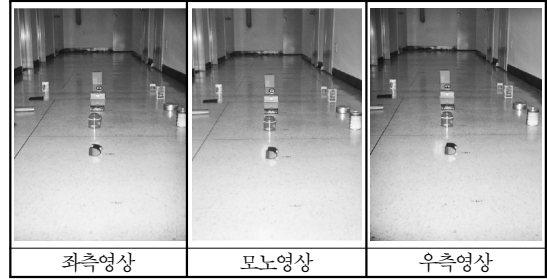


그림 8. 분리된 모노-스테레오 입력 영상
Fig. 8. Divided mono-stereo input images

IV. 실험 및 결과 고찰

그림 8은 본 실험에 사용된 영상으로 단일카메라의 모노-스테레오 시스템으로부터 획득한 영상(그림 5.)을 모노 영상과 좌, 우 영상으로 분리한 것이다.

그림 9의 (a)는 그림 8의 모노-스테레오 입력 영상에 식 (2)를 적용하여 LoG 필터에 의해 에지를 추출한 영상이고, 이 추출한 에지 영상을 정합하여 스테레오 시차를 추출하게 된다. 그림 9의 (b)는 그림 9의 (a)를 식 (3), (4), (5)로 정합을 하여 식(6), (7), (8)에 의해 스테레오 시차를 구하고, 이 시차를 이용해 두 영상을 합성한 영상이다. 즉, 중앙 영상인 모노 영상과 좌측 영상, 모노 영상과 우측 영상 및 좌측 영상과 우측 영상 간 스테레오 시차를 구하여 시차가 클수록 그레이 레벨이 진하게 나오도록 합성한 영상이다.

그림 10은 4개의 물체에 대하여 좌측-모노 영상간(Left-Mono) 및 모노-우측 영상간(Mono-Right) 스테레오 시차의 실험치와 이들 실험치를 더하여 좌측-우측 영상간 계산 값(Left-Right 계산값)을 구하고, 좌측-우측 영상간 실험치(Left-Right 실험치)를 구하여 나타낸 것이다. 그림 10에서 좌, 우측 영상간 스테레오 시차의 계산값과 실험치를 비교해 보면 거의 일치하게 나타난다. 즉, 모노 영상을 통한 좌, 우 영상의 스테레오 정합과 좌, 우측간의 직접적인 정합이 거의 일치함을 알 수 있다.

따라서 스테레오 영상처리시 모노 영상을 기준 영상으로 사용하면, 스테레오 물체 추적시 쉽게 추적물체를 인식하여 정확하게 추적할 수 있고, 또한 이 모노 영상을 이용하여 다수의 중간 영상을 임의적으로 쉽게 생성할 수 있으므로 다시점 디스플레이 시스템에도 적용할 수 있을 것이다.

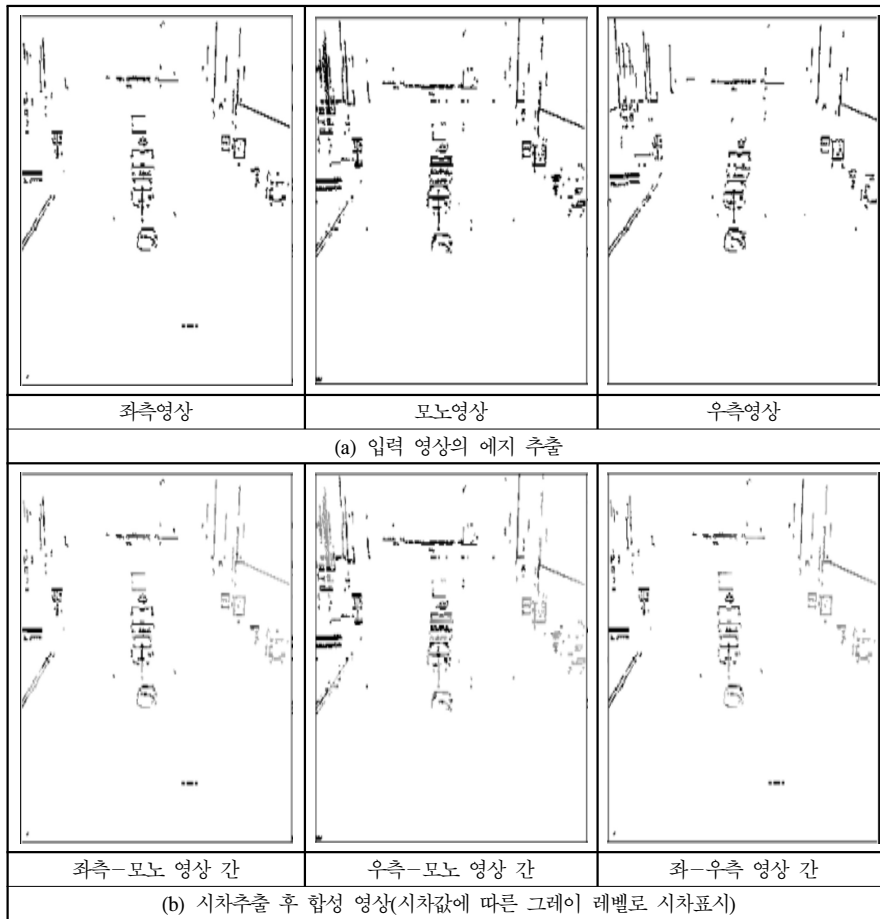


그림 9. 에지추출 및 시차합성 영상
Fig. 9. Edge detection and disparity composited images

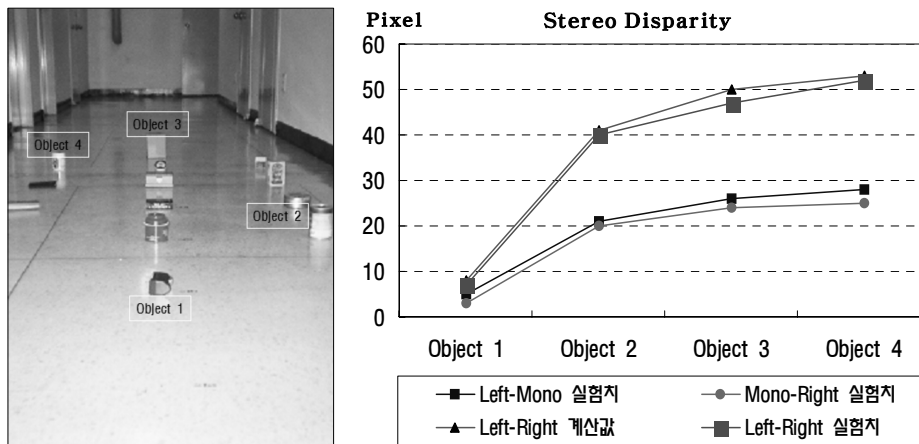


그림 10. 물체 4개의 스테레오 시차 비교
Fig. 10. Stereo disparity comparison for the four objects

V. 결론

스테레오 비전 시스템에서 정확한 스테레오 정합을 통해서 스테레오 시차를 추출하는 것이 중요하다. 스테레오 시차는 3차원 정보처리를 위한 값으로써 거리정보, 깊이정보 및 스테레오 디스플레이를 위한 영상합성 정보 등에 사용하기 때문에 오차가 없어야 정확한 3차원 정보처리가 가능하다.

본 논문에서는 단일 카메라의 모노-스테레오 시스템을 이용하여 스테레오 시차를 추출하였다. 모노-스테레오 카메라 모델은 일반 평행식 스테레오 카메라 모델의 중앙에 카메라 하나 더 두어 중앙의 모노 영상을 이용하여 스테레오 영상처리를 할 수 있도록 한 시스템이다. 이 모노 영상은 좌, 우 스테레오 영상의 중심점 역할을 하여 영상처리시 기준 영상으로 사용하고, 이 기준 영상으로 좌, 우 스테레오 영상을 정합하여 그 결과를 스테레오 시차로 제시하였다. 실험을 통해 좌, 우측 영상간 스테레오 시차의 계산값과 실험치를 비교해 보면 거의 일치하게 나타나고 있다.

앞으로 스테레오 영상처리시 모노 영상을 기준 영상으로 사용한 스테레오 물체추적 시스템이나 다시점 디스플레이를 위한 다수의 중간영상 생성 등에 대한 연구가 계속되어야 할 것이다.

참 고 문 헌

- [1] U. R Dhond and J. K Aggarwal, "Structure from Stereo - a review," IEEE Trans. on system, Man, and Cybernetics, Vol.19, No.6, pp.1189-1510, Nov./Dec. 1989
- [2] H. H. Baker and T. O. Binford. "Depth from edge and intensity-based stereo." Proc. 7th int Joint Conf. Artif. Intell., pp.631-636. 1981.
- [3] W.E.L. Grimson, "Computational experiments with a feature based stereo algorithm." IEEE Trans. Pattern Anal. Mach. Intell., Vol. PAMI-7, pp.17-34. 1985.
- [4] Y. Ohta and T. Kanade. "Stereo by intra- and inter-scanline search using dynamic programming." IEEE Trans. Pattern Anal. Mach. Intell., Vol. PAMI-7, pp.139-154, 1985.

- [5] R. D. Arnold and T. O. Binford, "Geometric constraints in stereo vision." Proc. SPIE, Vol.238, pp.281-292, 1980.
- [6] M. Okutomi and T. Kanade, "A locally adaptive window for signal matching," Int. J. of Computer Vision, Vol.7, No.2, pp.113-162, Apr. 1992.
- [7] T. Kanade and M. Okutomi, "A stereo matching algorithm with an adaptive window: theory and experiment," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.16, No.9, pp.920-932,
- [8] L. H. Matthies, R. Szeliski and T. Kanade, "Kalman filter-based algorithm for estimating depth from image sequences," Int. J. of Computer Vision, Vol.3, pp.209-236, 1989.
- [9] D. Scharstein and R. Szeliski, "Stereo matching with non-linear diffusion," Proc. Computer Vision and Pattern Recognition, pp.343-350, California. Jun. 1996.
- [10] Soo-In Kim, Jae-Soo Lee, Young-Woo Shon, "Distance Measurement of the Multi Moving Object using Parallel Stereo Camera in the Video Monitoring System". Journal of KIEE Vol.18, No.1, pp.137-145, 2004.

이 재 수 (Jae-soo Lee)

중신회원

한국통신학회 논문지 제32권 8호 참조

홍 권 의 (Kwon-eui Hong)

정회원



1983년 2월 건국대학교 전자공학과 졸업

1984년 8월 대우중공업 중앙연구소 전자기술부

1985년 2월 건국대학교 전자공학과 석사

1997년 2월 건국대학교 전자공학과 박사

1997년 3월~현재 김포대학 정보통신과 부교수

<관심분야> 광통신, 광정보처리, 영상처리

안 정 근 (Jeong-keun Ahn)

중신회원

한국통신학회 논문지 제33권 8호 참조