

論文2002-39SP-5-6

# 스테레오 영상을 이용한 물체 추적 방법

## (An Object Tracking Method using Stereo Images)

李學贊\*, 朴章漢\*\*, 南宮璉\*\*, 南宮在贊\*\*

(Hak-Chan Lee, Chang-Han Park, Yun Namkung, and Jae-Chan Namkung)

### 요 약

본 논문에서는 자동물체 추적 시스템의 효율성을 높이기 위해 스테레오 영상을 이용한 물체 추적 시스템을 제안하였다. 기존의 물체 추적 시스템에서 사용한 방법들은 최적의 특성을 지니고 있지만, 많은 계산량을 요구하는 단점이 있다. 또한 단안에 의한 영상의 경우 물체의 다양한 변화에 대한 예측과 추적이 어렵다. 따라서 본 논문에서는 양안에 의해 얻어진 스테레오 영상에 변이(translation)와 회전(rotation)에서의 예측이 어려운 단점을 보완한 블록 정합 알고리즘을 적용함으로써 실시간 물체의 변화의 추적 능력을 지니면서도 그 계산량을 줄일 수 있는 추적 방법을 제시하였다. 실험을 통하여 회전(25개)과 전이(28개), 그리고 회전과 전이의 혼합(50개)된 영상에서 각각 88%, 89%, 88%의 인식률을 얻었으며, 평균 88.3%의 인식률을 얻음으로써 본 연구의 유용성을 입증하였다.

### Abstract

In this paper, we propose a new object tracking system using stereo images to improve the performance of the automatic object tracking system. The existing object tracking system has optimum characteristics, but it requires a lot of computation. In the case of the image with a single eye, the system is difficult to estimate and track for the various transformation of the object. Because the stereo image by both eyes is difficult to estimate the translation and the rotation, this paper deals with the tracking method, which has the ability to track the image for translation for real time, with block matching algorithm in order to decrease the calculation. The experimental results demonstrate the usefulness of proposed system with the recognition rate of 88% in the rotation, 89% in the translation, 88% in various image, and with the mean rate of 88.3%.

**Key Words**: auto-object tracking system, stereo image, translation, rotation, block matching algorithm

### I. 서 론

동영상에서의 움직임 추적에 대한 연구는 오랫동안

컴퓨터 시각 분야에서 연구되어 왔다. 컴퓨터 비전의 궁극적 목표는 인간 시각 시스템의 기능을 컴퓨터로 구현하는데 있으며, 이를 구현하기 위해서는 입력된 영상정보로부터 3차원 물체를 찾아내거나 인식하는 것이 가장 기본적인 과제라 할 수 있다.

\* 正會員, 崇實大學校 컴퓨터科學科  
(Dept. of Computer Science, SoongSil Univ.)

\*\* 正會員, 光云大學校 컴퓨터工學科  
(Dept. of Computer Engineering, Kwangwoon Univ.)

※ 이 논문은 2001년도 광운대학교 교내학술연구비에 의하여 연구되었음.

接受日字:2002年2月21日, 수정완료일:2002年7月11日

2차원 영상으로부터 3차원 정보를 추출하는 과정은 매우 중요한 단계로서 하나의 카메라를 이용하는 단안 시법과 두 개의 카메라를 이용하는 양안 시법이 있는 후자를 일반적으로 스테레오 비전이라고 한다.

자동 물체추적 시스템(automatic target tracking system)<sup>[14]</sup>에 인간 시각계(human visual system)<sup>[9~11,14]</sup>

를 모방한 스테레오 시각장치를 적용함으로써 작업자가 현장의 상황이나 작업 전개를 보다 명확하게 알 수 있어 작업의 효율을 극대화할 수 있다.

일반적으로 원격 작업에 사용되는 영상 정보는 마치 사람이 한눈을 감고 사물을 보는 것과 같은 2차원 평면 영상이다. 그러나 입체 영상은 2차원 평면 영상과는 달리 사물에 대한 원근감, 즉 심도 정보를 지니고 있기 때문에 이를 원격 작업에 이용하면 작업자는 마치 현장에 직접 투입되어 작업을 수행하는 것과 같은 현장감을 얻게 되어 원격 작업을 쉽게 수행할 수 있다. 영국 AEA<sup>[1]</sup> 및 미국 ORNL<sup>[2]</sup> 등에서 수행한 고 방사선 지역이나 우주에서 사용될 로봇의 시각 장치의 원격 조작성 실험에 관한 연구 보고서에서는 입체 영상은 평면 영상에 비해 30% 이상 원격 작업 능력을 향상시키는 것으로 보고하고 있다.

물체 추적 시스템은 주로 레이더나 적외선 감지기 등으로 측정된 정보를 이용하여 표적의 탐지나 식별, 그리고 추적을 수행하게 되는데, 이 과정은 추적 시스템의 대상 물체들이 지니고 있는 현재의 위치, 속도, 그리고 가속도 등에 관한 정확한 정보들을 추출하는 작업을 의미한다.

또한 동영상에서의 움직임 추정 방법은 크게 세 가지로 분류할 수 있는데, 블록 기반(block-based) 추정 방법, 화소기반(pixel-based) 추정 방법, 광류기반(optical flow-based) 추정 방법 등이 있다. 블록 기반 방법은 H.261, MPEG-1, MPEG-2 등과 같은 영상 부호화 표준안에 적용되어 널리 사용되고 있지만, 회전(rotation)과 확대 및 축소(translation)를 보상하지 못하기 때문에 눈에 매우 거슬리는 블록화 잡음을 초래한다는 단점이 있다. 그리고 화소기반 움직임 추정 방법은 객체의 움직임을 정확하게 표현할 수 있다는 장점이 있지만, 모든 화소의 움직임 벡터를 전송해야 한다는 어려움과 각각의 많은 화소에 대하여 실제 움직임을 찾아야 한다는 단점이 있다. 한편, 광류기반(optical flow-based) 추정 방법은 시간에 따라 변화하는 영상의 강도 패턴으로부터 2차원적 벡터 분포를 계산하여 움직임을 추정하는 방법으로 지금까지 제안되어 온 많은 동영상 해석 방법들의 근간이 되고 있다. 이러한 광류기반 방법은 계산 속도가 오래 걸린다는 단점이 있다. 스테레오 영상에서 깊이 정보를 추출하는 방법 중 블록 정합의 경우 깊이 정보를 가지는 물체의 경계영역에서 발생하는 폐색(occlusion) 영역에서 오류 변위를

가진다.<sup>[6-9]</sup> 스테레오 영상에서의 폐색 영역은 실제 물체와의 구별을 어렵게 한다.

따라서 본 논문에서는 폐색 영역으로부터의 실제 물체 추출 방법과 블록기반 추정 방법을 보완함으로써 향상된 스테레오 물체 추적 시스템을 제시하고자 한다. 본 논문의 구성은 2장에서 스테레오 영상에 대한 것을 설명하고, 3장에서 제안한 스테레오 영상을 이용한 물체 추적에 대한 것을 설명하였다. 4장에서는 제안한 방법의 모의 실험 결과를 보여주고, 마지막으로 5장에서 결론을 맺는다.

## II. 스테레오 영상

사람의 두 눈은 사물을 가까이 보았을 때 하나의 사물을 다른 시점에서 받아들여 뇌에 보낸다. 뇌는 이 한 쌍의 이미지들 사이의 근소한 차이점을 인식해서 공간(거리와 부피)감을 만들며, 입체사진은 이 최적의 시스템을 모방하여 공간감을 실제처럼 재생산할 수 있다. 두 눈에서 각각 받아들여진 한 쌍의 그림을 뇌는 2개의 이미지를 3차원의 관점으로 융합시킬 수 있다. 이것을 스테레오 영상이라고 한다.<sup>[9]</sup> 스테레오 영상에는 최소한 좌우로 구성된 두 장 이상의 평면 영상이 필요하며, 이들 영상은 한 대의 카메라를 사용하여 시분할적으로 획득할 수도 있으나 일반적으로 두 대 이상의 영상 센서를 이용하여 얻는다.

### 1. 깊이 인식

깊이의 거리감을 얻는 것은 일반적으로 깊이 인식(depth perception)이라 불리우며, 입체시(stereoscopic vision)라는 말은 두 눈을 사용한 깊이의 지각이라는 한정된 의미로 사용되는 경우도 있다. 우리들이 주위의

표 1. 깊이감을 초래하는 시각의 요인  
Table 1. Main cause of the eye sight to bring the depth.

양안에 의한 결과	단안에 의한 결과
폭주 양안시차	초점 조절
	운동시차
	시야의 크기
	공기투시
	선원근법
	텍스처 구매
	그림자
	중첩 진출색-후퇴색

공간을 깊이를 가지고 입체적으로 파악할 때, 여러 가지 실마리가 사용된다. 이것에는 두 눈(양안)을 이용하는 것에 따른 것과, 외 눈(단안)에 의해서도 생길 수 있는 것이 있으며 표 1에 제시한 바와 같다.

#### (1) 양안에 의한 결과

어떤 대상을 바라볼 때, 양안은 내향으로 회전하여, 그 대상 위에서 만난다(그림 2.1). 이 같은 양안의 작용을 폭주(vergence)라 하며, 이때 대상 A와 시선이 이루는 각을 폭주각이라 부른다. 폭주에 의한 깊이와 실마리는, 물론 근거리에서 큰 효과를 나타내는데 20m까지 유효하다. 그러나 거리가 멀어지면 폭주각이 작아지므로 이 효과는 급격히 작아진다. 또, 2개의 눈이 떨어진 위치에 있으므로 해서, 어떤 물체를 바라보았을 때 양안의 망막상은 같아지지 않고, 주시점으로부터 떨어진 위치에서는 대체로 간격이 생긴다.

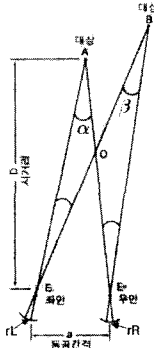


그림 1. 폭주와 양안시차

Fig. 1. Vorgence and Parallax.

양안에서 이 같은 간격량의 차이를 양안시차라 한다. 그림 1에서 A를 주시하고 있을 때, 이것과는 다른 방향에서 다른 깊이를 갖는 대상 B와의 사이의 양안시차는 좌, 우안에서의 간격량의 차, 즉 각도로 표현하면  $rL - rR$  로 주어진다. 또, 그림에서 B가 양안에 대해 떠는 각을  $\beta$ 로 하였을 때, 3각형 ELOA 및 3각형 EROB에 주목하면(내각의 합이 같으므로) 상기한 양안시차는  $\beta - \alpha$ 로 표현할 수 있다

#### (2) 단안에 의한 결과

단안에 따른 깊이의 결과에는 눈의 렌즈의 두께를 바꾸어 핀트 조절을 하는데 따른 효과가 있다. 이것은 관찰거리가 2~3m이내의 근거리일 때만 유효하다. 이 같이 관찰하는 사람의 위치와의 상대적인 변화에 따라 생기는 대상물의 움직임과 차이를 운동시차(motion

parallax)라 부른다. 이같이 움직임과의 차이에서 초래되는 깊이 판단의 효과는 조건에 따라 양안시차와 같은 정도로 유효하며, 현재의 TV나 영화 등 2차원의 화면 속에서의 깊이감을 부여하는데 유효한 실마리로 되어 있다.

### 3. 스테레오 카메라

입체 영상에서는 좌,우로 구성된 두 장 이상의 평면 영상이 필요하며, 이들 영상은 한 대의 카메라를 사용하여 시분할적으로 획득할 수도 있으나 일반적으로 두 대 이상의 영상센서를 이용하여 동시에 촬영하여 얻는다. 양안 입체카메라는 사람의 눈과 같이 두 개의 영상센서를 이용하여 좌,우 영상으로 구성되는 입체영상을 동시에 획득하는 장치를 말하며, 일반적으로 입체카메라라고 하면 양안 입체카메라를 지칭한다. 입체카메라는 좌,우 영상센서의 배열방법에 따라 평행축, 교차축, 그리고 수평 이동축 방법으로 구분된다.

평행축 입체카메라는 양안식 입체영상 카메라 중 가장 단순한 형태로 두 대의 영상센서를 사람의 눈 간격과 비슷한 거리로 평행하게 고정시켜 두고 영상을 획득할 수 있게 설계되어 있다. 평행축 입체카메라는 좌,우 영상센서 사이의 거리가 일정하게 고정되어 있어 사람의 눈동자 움직임, 즉 주시각 제어 기능이라고 하는 물체의 거리변화에 따른 입체 영상 시차 조절 기능이 없다.

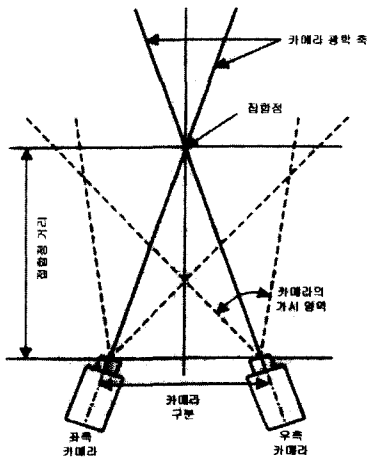
교차축 입체카메라는 토우드인(toed-in) 또는 폭주식 입체카메라라고도 하며, 물체의 거리 변화에 따라 주시각 제어가 가능하도록 설계되어 있다. 이 스테레오 카메라는 물체의 거리 변화에 따라 영상센서의 광 축을 회전시켜 항상 물체의 상이 좌우 영상센서의 중심에 맺히도록 주시각을 제어한다. 이는 가까운 물체를 볼 때 안쪽으로 모이고 멀리 있는 물체를 볼 때 벌어지는 사람의 눈동자 움직임을 모방한 것이다. 교차축 스테레오 카메라는 좌우 영상센서를 회전하여 주시각을 조절한다는 점에서는 사람의 눈과 같다. 그러나 교차축 입체카메라가 주시각을 제어할 때 사람 눈에서와는 달리 상이 맺히는 좌,우 영상 센서들 사이의 간격이 변하게 된다. 이러한 영상 센서간 간격 변화는 교차축 입체카메라로 입체영상을 관측할 때 발생하는 영상 왜곡의 원인이 된다.<sup>[13]</sup>

평행 이동축 입체카메라는 교차축 방식과 같이 관측 물체의 거리 변화에 따라 카메라의 주시각 제어가 가

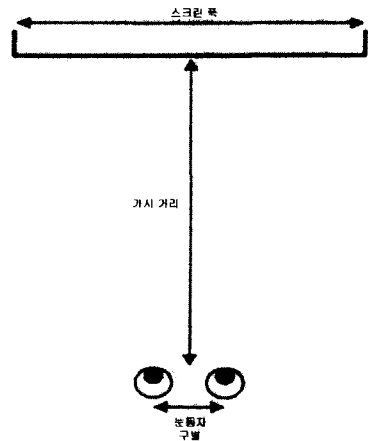
능한 카메라이다. 그러나 평행 이동축 입체카메라는 영상센서의 광축을 회전시켜 주시각을 제어하는 교차축 방식과는 달리 렌즈를 영상센서로부터 분리시킨 후 렌즈에 대해서 영상센서가 평행하게 수평으로 움직이게 하여 주시각을 조절하도록 설계되어 있다. 이러한 영상센서의 평행 이동에 의한 주시각 제어 방식은 좌,우 영상센서 사이의 간격 변화량이 작아 교차축 방식에 비해 상대적인 영상 왜곡이 작다.

이들 입체카메라가 사람의 눈에서와 같이 자연스럽게 선명한 양질의 입체영상을 획득하기 위해서는 인간의 시각 기능중 최소한 세 가지 기능, 즉 양안시차(binocular parallax) 기능, 초점제어 기능, 그리고 주시각 제어 기능을 갖추어야 한다.<sup>[3,4]</sup> 양안시차 기능은 두 대의 영상센서를 이용하는 양안식 입체카메라의 경우 항상 만족하며, 양안시차 기능에 의해 발생하는 관측 영상에 대한 시차량은 물체의 거리 변화와 두 영상 센서간 간격에 따라 달라진다. 사람의 좌,우 눈 사이의 간격은 평균 65mm이다. 초점제어 기능은 물체의 원근에 따라 초점을 맞추어 관측 영상을 선명하게 보도록 하는 기능으로, 사람의 눈의 경우 먼 물체를 볼 때 수정체의 두께가 가늘어지고 가까운 물체를 볼 때는 두꺼워진다. 주시각 제어 기능은 물체를 볼 때, 좌,우 눈의 광축을 회전시켜 두 눈의 망막 중앙부에 상이 맺히도록 하는 것으로 주 관심 물체만을 응시하도록 하며, 관측 물체를 자연스럽게 볼 수 있게 한다.

그림 2는 입체 영상 시스템에서 사용되는 경우로서 (a) 입체 카메라, (b) 입체 영상 디스플레이를 나타낸 그림이다.



(a) 입체 카메라  
(a) Stereo camera



(b) 입체 영상 디스플레이  
(b) Stereo display

그림 2. 입체 영상 시스템  
Fig. 2. Stereo image system.

### III. 스테레오 영상을 이용한 물체 추적

그림 3은 본 연구에서 제안한 스테레오 영상을 이용한 물체 추적 시스템에 대한 전체 구성도이다.

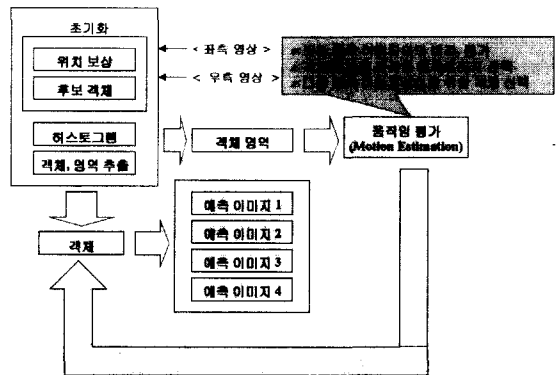


그림 3. 스테레오 영상을 이용한 물체 추적 시스템  
Fig. 3. Image tracking system by using stereo image.

#### 1. 스테레오 영상에서의 시차정보 추출

스테레오 카메라는 인간 시각계에 근거를 두고 피사체에 카메라를 평행하게 하여 스테레오 영상을 얻을 수 있다.

평행축 스테레오 카메라를 구성할 경우 스테레오 영상은 수평시차만 존재하고 수직시차는 존재하지 않는다. 교차식 스테레오 카메라의 경우 영상의 중앙으로부터 멀어질수록 수직 시차 성분이 존재(키스톤 변형)하고 폭주각이 클수록 수직시차의 성분은 늘어난다. 교차

식 스테레오 카메라 구성에서 나타나는 수직시차는 입체화상을 보기 나쁘게 하는 요인으로 나타나고 있다.

평행축 스테레오 영상에서 물체가 카메라로부터의 거리가 멀어질수록 수평시차 성분이 작아지고 거리가 가까울수록 수평시차의 성분은 늘어나는 특징을 가지고 있다. 그림 4는 스테레오 영역과 평면 영역을 나타낸 것이다.

입체 영상에는 입체영역 외에도 입체감이 형성되지 않는 평면영역(mono area)과 입체영상을 획득할 때 입체카메라의 좌,우 영상센서의 에피폴라선(epipolar line)이 제대로 정렬되지 않아 발생하는 불일치 영역(unmatched area)이 존재한다. 이러한 평면 영역과 불일치 영역은 입체영상을 관측할 때 관측자로 하여금 구역질, 어지러움과 같은 입체영상 관측 거부감을 느끼게 하는 원인이 된다.<sup>[5]</sup>

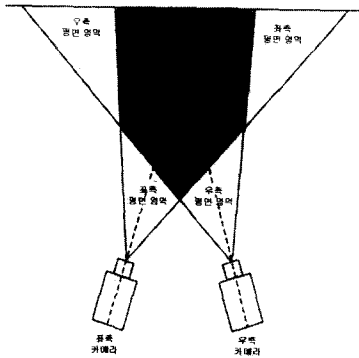


그림 4. 스테레오 영역과 평면 영역  
Fig. 4. Stereo and mono arcs on a stereo camera.

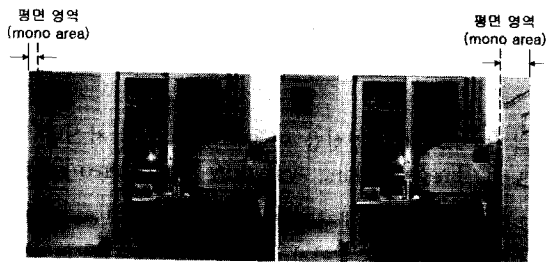


그림 5. 스테레오 영상에서의 평면 영역  
Fig. 5. Mono area of the stereo image.

그림 4와 그림 5에서 볼 수 있듯이, 평면 영역은 주로 입체영상을 구성하는 좌우 평면영상의 가장자리에 나타나며, 불일치 영역은 좌우 평면영상의 상하 영역에 나타난다. 그림 5는 편이상 교차축에 대한 평면영역을 표현하였으나 수평 이동축에서도 동일하게 평면영역이

존재한다. 평면영역과 불일치영역과 같은 비입체영역을 줄이기 위해서는 영상획득 이전에 입체카메라의 캘리브레이션 작업이 선행되어야 한다.

본 논문에서 시차 정보의 추출 방법은 고정된 크기의(fixed size block) 블록정합 알고리즘을 사용하였다. 블록정합에서 사용되는 기저함수는 MAD(Mean Absolute Difference)와 MSD(Mean Square Difference)를 주로 사용하는데, 본 논문에서는 계산량을 줄이면서 시차 추출을 위해 SAD(Sum of Absolute Difference)를 기저함수로 사용하였다. 공간상의 한 점이 좌우 두 영상에 투영되었을 때 두 점을 연결하는 가상의 선을 에피폴라선이라 한다.<sup>[6,7]</sup> 우측영상의 대응 블록은 좌측 영상과 동일한 y축의 좌표를 가지는 에피폴라 선상에서 정합된다. 식(1)은 스테레오 영상의 시차추출을 위해 사용된 SAD를 나타낸다.

$$SAD = \sum_i^{N_x} \sum_j^{N_y} [I_L(i, j) - I_R(i+k, j)] \quad (1)$$

여기서  $N_x$ ,  $N_y$ 는 블록의 크기를 나타낸다.  $I_L$ 은 좌측 영상의 블록을 나타내고  $I_R$ 은  $I_L$ 에 대응되는 우측 영상의 블록을 나타낸다.

2. 스테레오 영상의 배경 시차 보상

스테레오 영상의 배경 시차 보상은 스테레오 영상에는 포함되어 있는 평면 영역과 입체카메라의 에피폴라선이 제대로 정렬되지 않아 발생하는 불일치 영역을 최소화하는 과정이다. 평면 영역은 입체 영상을 구성하는 좌우 평면 영상의 좌우 가장자리에 분포하며 상하 불일치 영역은 영상의 상하 가장자리에 분포한다.

따라서 시차에 의해 발생하는 좌,우 영상의 불일치를 보상하는 과정이 선행되어야 한다. 이를 보상하기 위해, 본 논문에서는 블록 크기(N)를 4\*4으로 하는 블록정합 알고리즘을 이용하였다.

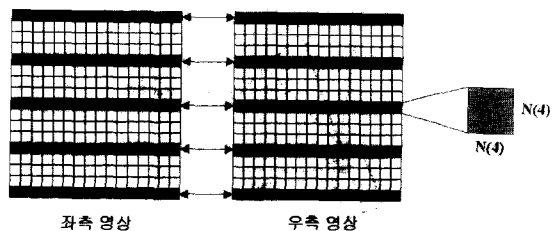


그림 6. 스테레오 영상에서의 배경 시차 보상  
Fig. 6. Background parallax computation of stereo image.

그림 6은 스테레오 영상에서의 배경 시차 보상을 나타내고 있다.

먼저, 해당열의 SAD를 식(2)을 이용하여 각각 구하고,

$$SAD_k = \sum_i \sum_j^{N_x} [I_L(i, j) - I_R(i+k, j)] \quad (2)$$

각 동일 열에서의 최소 SAD는 식(3)과 같이 구한다.

$$SAD_{MIN}^p = MIN(\sum SAD_k) \quad (3)$$

마지막으로 전체 영상의 배경 시차인 최소 SAD를 식(4)을 이용하여 얻을 수 있다.

$$SAD_{MIN} = MIN(\sum SAD_{MIN}^p) \quad (4)$$

따라서, 이 때의 k가 좌우 영상의 배경 시차가 된다. 그림 3.5는 배경 시차를 이용하여 얻은 보상된 스테레오 영상을 나타내고 보상된 이미지를 식(5)와 식(6)과 같이 쓸 수 있다.

$$right^* = right_{i-k} \quad (5)$$

$$left^* = left_{i+k} \quad (6)$$

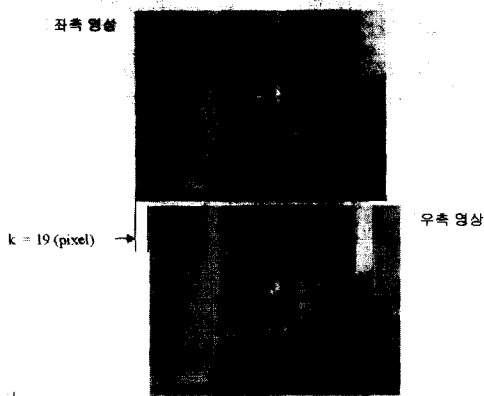


그림 7. 보상된 스테레오 영상  
Fig. 7. Compensated stereo image.

### 3. 스테레오 영상의 후보 객체 영역 추출

후보 객체 리스트란 스테레오 영상에서의 시차 정보와, 폐색 영역에 의해 보정된 왼쪽 영상과 오른쪽 원 영상, 또는 왼쪽 원 영상과 보정된 오른쪽 영상과의 차가 발생하는데, 이것은 실제 객체를 나타내는 영역이

될 수 있지만 폐색 영역과 같이 객체가 될 수 있는 모든 영역을 말한다.

스테레오 영상은 인간의 시각계를 모방한 시스템으로 제작되기 때문에 두 영상은 매우 흡사한 영상정보를 가진다. 두 카메라가 바라보는 점을 기준(가상 표시면)으로 한다면 두 카메라는 동일한 영역을 바라보기 때문에 두 영상은 대부분이 일치된 영상정보로 구성된다. 가상 표시면을 기준으로 양안시차가 큰 물체의 경우 좌우로 일정간격 떨어져 있는 카메라의 위치 차이로 인해 물체를 보는 면이 약간 달라진다.

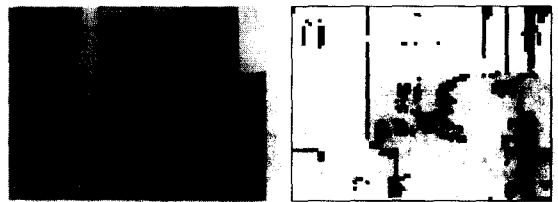


그림 8. 후보 객체 리스트  
Fig. 8. Candidate-object list.

또한, 시차의 크기에 따라 물체의 경계영역에서 볼 때, 한쪽 영상에서는 보이지만 다른 쪽 영상에서는 시차를 가지는 물체로 인해 폐색(occlusion) 영역이 생긴다. 스테레오 영상에서 폐색 영역은 주로 물체의 경계영역에서 발생한다. 따라서 폐색 영역은 객체가 될 수 있는 영역이 될 수 있다.

그림 8은 스테레오 영상에서의 폐색 영역과 시차를 이용한 후보 객체 리스트를 나타내고 있다. 이러한 폐색 영역은 시차가 큰 물체의 경계일수록 시차에 비례하는 넓은 영역의 폐색 영역을 차지한다. 폐색 영역이 생기는 위치는 기준영상을 어느 쪽으로 하느냐에 따라 물체의 오른쪽 또는 왼쪽으로 달라지며, 물체가 가상 표시면의 앞에 있느냐 뒤에 있느냐에 따라서도 폐색 영역의 위치는 바뀐다.

카메라와 물체사이의 거리 변화는 카메라에 비치는 물체크기에 영향을 미쳐 시차 변화를 일으키게 하는 전이(translation)가 발생한다. 그림 9는 카메라와 물체간의 거리 변화에 따른 시차 변화를 나타내고 있다. 카메라로부터 물체의 거리가 가까울수록 좌,우측 영상간의 시차는 커지게 되고 거리가 멀어질수록 시차가 작아지는 것을 쉽게 볼 수 있다. 식(7)~식(9)는 후보 객체 리스트를 구하는 수식을 나타내고 있다.

$$image_{diff} = left - right^* \quad (7)$$

$$Image_{co} = \begin{cases} 0 & image_{diff} < threshold \\ 1 & otherwise \end{cases} \quad (8)$$

$$threshold = MBSIZE \times 150 \quad (9)$$

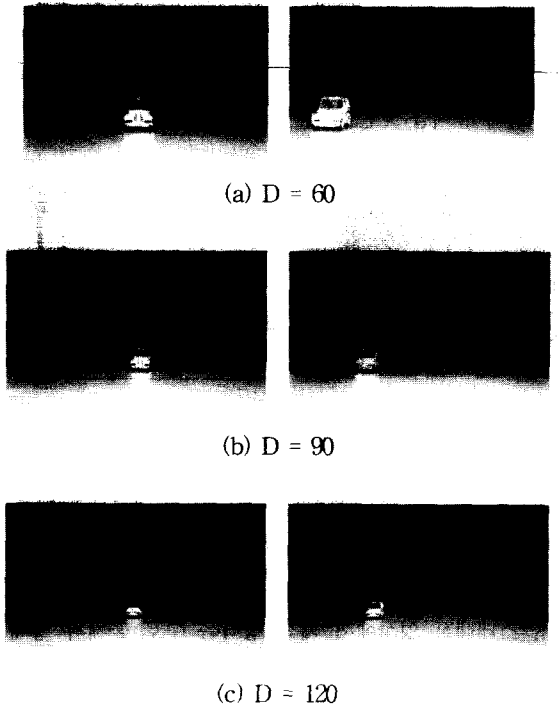


그림 9. 카메라로부터의 거리(D)에 따른 물체의 시차 변화

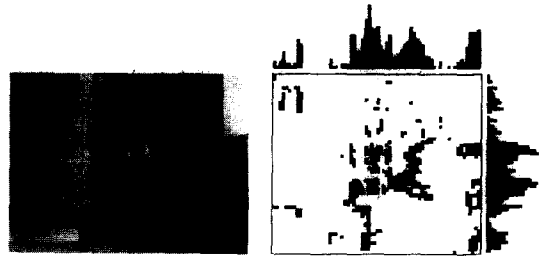
Fig. 9. Parallax alteration by D the distance from camera to image.

4. 스테레오 영상에 의한 객체 추출

스테레오 영상에서 배경과 객체를 분리하기 위해서는 먼저, 실제 움직이는 객체의 확장된 영역을 추출하는 과정이 필요하다. 후보 객체 리스트의 히스토그램 계수 변화는 객체의 움직임 및 다양한 변화를 예측할 수 있게 해 준다.

(1) 후보 객체 추출 방법

히스토그램은 디지털 영상 처리에서 가장 간단하면서 유용한 툴 중에 하나로, 영상의 명도 내용을 요약해 빈도수를 표현하고 상대적인 빈도를 길이로서 표현하기 위해서 막대그래프 모양으로 나타낸다. 그림 10은 후보 객체 리스트의 히스토그램을 나타내고 있다.



(a) 원 영상 (b) 원 영상에 대한 히스토그램

그림 10. 후보 객체 리스트에 대한 히스토그램  
Fig. 10. Histogram of candidate object list.

(2) 객체 추출 방법

입력 영상으로부터 얻은 전, 후 영상의 히스토그램은 객체가 존재할 영역을 구하는데 중요한 요소가 된다. 블록 정합 알고리즘을 이용한 전 영역 탐색은 실시간 물체 추적 시스템을 구현하는데 비효율적인 방법이다. 히스토그램은 계수가 변하는 시작점(first)과 변화가 더 이상 존재하지 않는 끝점(last)을 찾아 각각 x, y축에 대한 후보 영역을 구함으로써 국부적 탐색을 용이하게 함과 동시에 연산량을 감소 시키는 역할을 한다. 그림 11은 후보 영역과 객체 추출을 나타내고 있다.

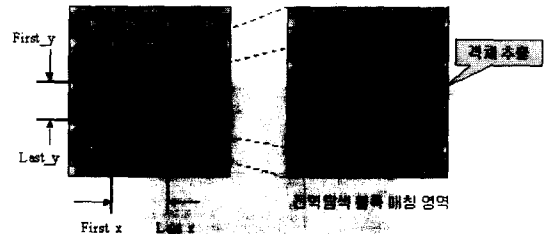


그림 11. 후보 영역과 객체 추출

Fig. 11. Candidate area and Object extraction.

후보 영역의 크기를 결정하는 길이(length)를 다음과 같이 정의하고

$$Length = |first - last| \quad (10)$$

길이 정보를 이용하여 움직임에 대한 예측을 할 수 있다. 이전 영상 히스토그램 계수(length\_pre)와 현재 영상의 히스토그램 계수(length\_cur)를 비교하여 히스토그램 계수의 변화가 있을 경우는 객체의 움직임을, 히스토그램 계수의 변화가 없을 경우 움직임이 없다는 것을 의미한다.

5. 예측 이미지

블록 정합 알고리즘의 특징은 카메라로부터 물체가 점점 가까워지거나 멀어질 경우 상관성이 떨어지는 변이(translation)와 물체의 회전(rotation)의 변화를 알기 어려운 단점을 가지고 있다. 시차 정보는 이를 보완할 수 있는 중요한 정보가 될 수 있는데, 움직이는 물체를 추적하기 위해서는 물체의 변화되는 모습을 예측할 수 있어야 한다. 변이의 경우 카메라로부터 객체의 거리가 가까워지거나 멀어지게 되면 카메라에는 각각 확대나 축소의 모양으로 나타나기 때문에 초기 객체의 크기를 가지고 블록 정합 알고리즘을 이용한다면 예상치 못한 결과를 초래하게 된다. 때문에 시차정보를 이용하면 이러한 결과를 보완할 수 있다. 물체의 양안 시차가 커진다는 것은 물체가 카메라 방향으로 이동하는 것을 의미하고 반대로, 물체가 카메라로부터 점차 멀어지게 되면 양안 시차는 작아지게 된다. 따라서 실제 객체의 가로, 세로의 비가 선형 배율로 확대, 축소된 영상으로 보여지게 되는 것이다. 회전의 경우, x축 회전과 y축 회전이 있는데 객체가 회전할 경우 x, y축 중 어느 한쪽만의 히스토그램 계수가 변하는 성격을 갖고 있다.

표 2. 객체 변화  
Table 2. Object alteration.

	Length_x 변화	Length_y 변화
변이(translation) (선형 배율)	O	O
회전(rotation)	O(X)	X(O)
변이(translation)와 회전(rotation) (비선형 배율)	O	O

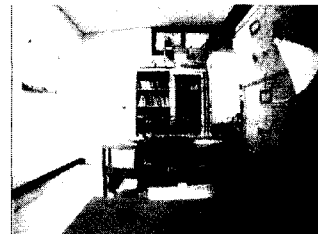
예측 이미지는 표 3에 있는 정보를 가지고 생성되며, 예측 이미지의 수는 연산량에 영향을 주기 때문에 그 수를 8개로 제한하였다.

움직이고 있는 물체는 방향성을 가지고 운동하게 된다. 즉, 현재 움직이고 있는 방향에서 갑자기 반대방향이나 혹은 다른 방향으로의 전환이 어렵다. 만약, 객체가 영상의 앞쪽으로 이동하고 있다면 확대 영상에 예측이미지 수를 늘려 적중률을 높임으로써 보다 정확한 움직임 추출을 할 수 있다.

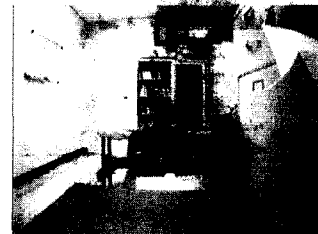
IV. 모의 실험

1. 실험

본 논문에서 실험을 위해 320X240 크기를 갖는 256 그레이 레벨의 좌, 우측 이미지 103개씩을 가지고, Visual C++6.0을 이용하여 작성하였다. 스테레오 카메라의 간격은 약 6cm이고, 카메라와 객체와의 거리는 약 2m이며, 펜티엄-III와 주 메모리 256MB인 컴퓨터 사용하여 실험하였다.



(a) 좌측 영상



(b) 우측 영상

그림 12. 스테레오 영상  
Fig. 12. Stereo image.

그림 12는 실험을 위해 좌, 우측에서 각각 촬영한 실제 사진이다.



그림 13. 생성된 후보객체 리스트  
Fig. 13. Created candidate object list.

그림 13은 스테레오 영상에서 생기는 배경시차를 보정한 후 좌, 우측 영상의 차를 통하여 객체가 될 수 있



는 영역의 객체들을 표현한 것이다. 배경 이미지가 제거된 후보 객체 리스트는 객체와 객체 사이에 가려지는 영역인 폐색 영역과 실제 객체들로 구성되는데 스테레오 영상에서의 시차 정보는 이것을 가능하게 해주는 결정적인 요소가 된다. 실제로 후보객체 리스트에서는 폐색 영역과 객체의 구별이 불가능하다. 때문에 이전 영상과 현재 영상의 히스토그램 계수차를 이용하여 폐색 영역과 객체를 분리하고 후보객체에 대한 후보영역을 얻을 수 있다.

그림 14와 그림 15는 스테레오 영상에 대한 히스토그램을 나타내고 있다. 그림 14는 그림 12에 대한 이전 영상의 히스토그램이고, 그림 16은 그림 12의 현재 영상에 대한 히스토그램을 이다. 그림 16은 이전, 현재 영상에 대한 두 히스토그램의 계수를 비교하여 계수의 변화가 있는 가로축과 세로축 영역의 길이를 후보 영역으로 선택하였다.



그림 14. 이전 영상 후보객체 리스트에 대한 히스토그램  
Fig. 14. Histogram of the previous image candidate object.

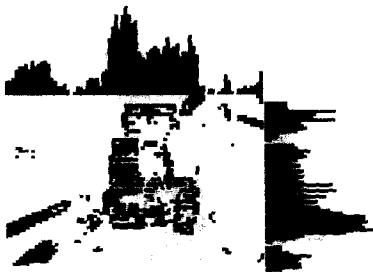
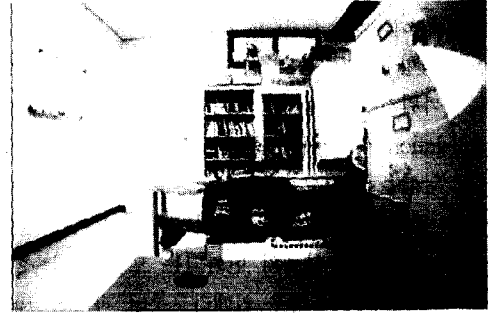


그림 15. 현재 영상 후보객체 리스트에 대한 히스토그램  
Fig. 15. Histogram of the current image candidate object.

후보영역에서 객체를 추출하기 위해 BMA연산을 수행하였으며 정밀한 위치 탐색을 위해 전역탐색 블록 매칭을 이용하였다. 그림 16은 추출된 객체를 나타내고 있다.



..... 후보영역  
———— 분리된 객체

그림 16. 분리된 객체  
Fig. 16. Separated object.

BMA연산에서 발생하는 회전과 전이 문제는 움직임 추적에 문제가 된다. 예측 이미지는 다음 움직임의 예측 이미지로써, 분리된 객체를 회전과 전이에 대해 단계적으로 이미지를 4-6개씩, 방향성에 비례하여 예측 이미지를 생성, 다음 이미지에서의 움직임 추적을 하며 갱신시키면서 움직임 추적을 수행하였다.

예측 이미지를 위한 객체의 변화는 다음과 같은 조건을 따른다.

**\* 변이(Translation)**

단일 이동 물체의 Length<sub>x</sub>, Length<sub>y</sub> 의 변화가 동시에 선형적인 비율로 변화했을 때,

**\* 회전(Rotation)**

단일 이동 물체의 Length<sub>x</sub> 나 Length<sub>y</sub> 중 어느 한쪽만 변화했을 때,

**\* 변이과 회전(Translation & Rotation)**

단일 이동 물체의 Length<sub>x</sub>, Length<sub>y</sub>가 동시에 비선형적인 비율로 변화했을 때,

그림 17은 스테레오 영상에서의 변이와 함께 움직이는 물체의 추출을 보여주고 있다.

그림 187(a)의 그림은 움직이기 이전의 원 영상을 그림 187(b)는 객체의 회전이 발생한 원 영상을 나타내고 있으며 이 영상을 가지고 실험결과를 그림 19에 보여주고 있다.

표 3를 보면, 실제로 물체의 움직임을 추적하는데 소요되는 대상 적용은 배경 시차 보상, 후보 객체 리스트 생성, 후보 영역 검출, 예측 이미지 생성, 움직임 벡터 계산 등이다. 본 시스템에서는 초기화 과정인 배경 시

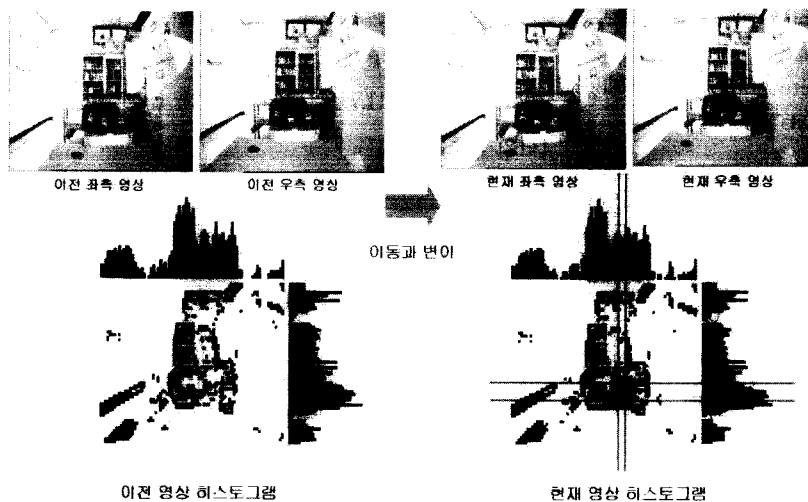


그림 17. 이동과 변이에서의 물체 추출  
Fig. 17. Tracking of moving and translation.



그림 18. 회전된 객체  
Fig. 18. rotated object.

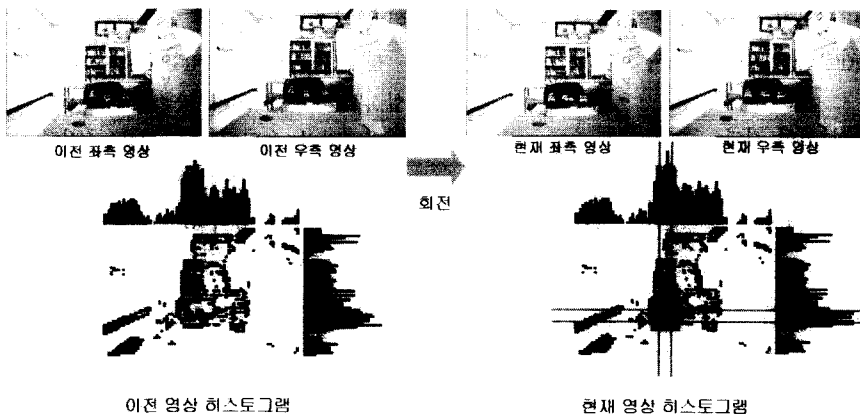


그림 19. 회전된 객체의 물체 추적  
Fig. 19. Tracking of rotated object.

표 3. 물체의 움직임을 추적하는데 소요되는 시간

Table 3. Required time for tracking to moving of object.

[단위 : 초]

처리 내용	처리 시간	프레임당 처리 시간
배경 시차 보상	0.11	0.11/50 = 0.0022
후보 객체 리스트 생성	2.1	2.1/50 = 0.42
후보 영역 검출	0.19	0.19/50 = 0.0038
예측 이미지 생성	3.3	3.3/50 = 0.066
움직임 벡터 계산	0.42	0.42/50 = 0.0084

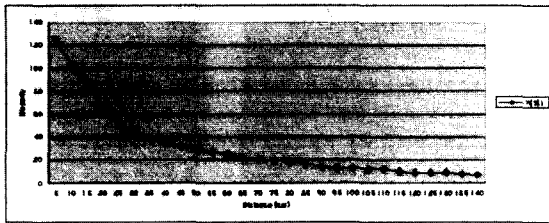


그림 20. 객체의 이동에 따른 변하는 시차 그래프  
Fig. 20. Changing disparity graph for moving of object.

차 보상을 제외하고, 초당 8.17 프레임을 처리할 수 있다.

그림 20은 객체의 이동에 따라 변하는 시차를 보여 주고 있으며, 예측 이미지 생성시 크기를 결정하는 중요한 요소가 된다. 시차의 변화를 보면, 지수 형태를 띠고 있어 선형식을 적용해서 정확한 거리를 측정하기에는 어려움이 있지만, 물체가 카메라와 가까워질수록 시차의 변화가 점점 커져 거리의 변화를 쉽게 찾을 수 있었다. 회전하는 물체는 물체의 모양과는 상관 없이 시차의 변화가 발생하지 않고, 단지 좌, 우측 영상의 예측 이미지의 상관상을 비교하여 변화를 찾았다.

표 4. 모의 실험 결과

Table 4. Experimental results.

예측 이미지 조건	전체 영상 수	제안한 알고리즘		블록 매칭 알고리즘	
		검출 성공	검출 실패	검출 성공	검출 실패
회전	25	22	3	19	6
변이	28	25	3	20	8
회전과 변이	50	44	6	35	15
백분율(%)		88.3	11.7	76.7	23.3

표 4는 모의 실험한 결과이며, 제안한 알고리즘이 회전, 변이, 회전과 변이에서 각각 88.3%의 검출 성공과 11.7%의 검출 실패를 볼 수 있다.

### V. 결 론

본 연구의 실험 결과 스테레오 영상의 객체 움직임에 따른 물체 추적 시스템의 개발 가능성을 보였으며, 블록 매칭 알고리즘에서의 문제점인 전이와 회전의 문제점을 개선시킬 수 있음을 보였다. 기존의 2D 영상에서의 물체 추적시스템은 거리를 인식할 수 없어 전이를 인식할 수 없었으나 본 연구에서 스테레오 영상의 시차를 이용함으로써 예측이미지를 생성하는데 scaling 정보를 제공함으로써 이를 개선하였다. 또한 예측 이미지는 회전된 영상에서도 성능을 개선시키는 요소라는 것을 볼 수 있었다.

본 논문에서 쓰인 실험 영상은 회전각이 적은 영상을 가지고 실험하였다. 회전각의 크기를 점차 키워 실험한 결과, 작은 각(15~30도)의 회전 영상에서는 추적 성능이 좋은 반면, 회전각이 커질수록 성능이 낮아지는 현상을 보였다. 또한 스테레오 영상을 얻기 위한 두 대의 카메라의 성능이 각기 조금씩 달라 정확한 스테레오 영상의 취득이 어려워 히스토그램에서 쓰레기 값이 포함되어 후보영역의 크기가 커짐에 따라 연산량이 증가하여 성능이 저하되었다.

제안한 방법은 물체 추적 뿐만 아니라, 거리 정보를 이용한 물체의 이동속도를 얻을 수 있기 때문에 차량의 속도 측정이나 도로교통 정보를 얻을 수 있어 기존의 많은 비용이 소요되는 교통 분야의 기술 발전에 기여할 것으로 사료된다.

향후 과제로는 제안한 방법이 연산량을 최소화하여 실시간 처리를 고려하여 설계되었기 때문에 고정된 카메라에서의 움직임은 물체를 추적하는 시스템으로 제한을 두어 실험하여, 물체가 제한된 영역을 벗어날 경우 추적이 불가능하다. 이러한 문제점을 해결하기 위해서는 기존의 알고리즘에 대한 추가의 연구가 필요할 것으로 사료된다.

### 참 고 문 헌

[1] A. Dumbreck, C. Smith, and S. Murphy, "The development and evaluation of a stereoscopic

- television system for remote handling”, *Image Technology Journal of BKSTS*, pp. 8~12, Jan. 1990.
- [2] J. Draper, Y. Fujita, and J. Merdon, “Evaluation of High Definition TV for Remote Task Performance”, Peropt no. ORNL/TM-10301, Oak Ridge National Laboratory, Apr. 1987.
- [3] T. Olson and D. Coombs, “Real-time vergence control for binocular robots”, *International Journal of Computer Vision*, Vol. 7, No. 1, pp. 67~89, 1991.
- [4] P. Smith and N. Nandhakhmar, “An Improved power cepstrum based stereo correspondence method for textured senses”, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 18, No. 3, pp. 338~348, Mar. 1996.
- [5] K. Boff, L. Kaufmann, and J. Thomas, “Handbook of perception and Human Performance”, John Wiley & Sons, 1986.
- [6] L.H. Chen, and W.C. Lin, “Visual surface segmentation from stereo”, *Image and Vision Computing*, Vol. 15, pp. 95~106, 1997.
- [7] V.S. Grinberg, G.W. Podnar, and M.W. Siegel, “Geometry of Bionocular Imaging”, *Proc. SPIE*, Vol. 2177, pp. 56~65, Feb. 1994.
- [8] E. Izquierdo, “Disparity/Segmentation Analysis: Matching with an Adaptive Window and Depth-Driven Segmentation”, *Proc. IEEE trans, Circuits Syst. Video Technol.*, Vol. 7, No. 4, pp. 589~607, Jun 1999.
- [9] U.R. Dhond and J.K. Aggrawal, “Structure from stereo-A review”, *IEEE trans, System. Man and Cybern.*, Vol. 19, pp. 1489~1510, 1989.
- [10] W. Y. Yau, and H. Wang, “Fast Relative Depth Computation for an Active Stereo Vision System”, *proc. Academic Press. Real-time Imaging 5*, pp. 189~202, 1999.
- [11] T. Notoki, H. Isono, and I. Yuyama, “Present Status of Three Dimensional Television Research”, *Proc. IEEE*, Vol. 83, pp. 1009~1201, No. 7, July 1995.
- [12] 김은수, 이승현, “3차원 영상의 기초”, *技多利*, 1998
- [13] 이용범, 진성일, 박순용, “촬영법에 따른 입체영상 재현성 비교”, *대한전자공학회 하계학술대회 논문집*, 제19권 1호, pp. 503~510, 1996
- [14] 이재수, 김규태, 김은수, “광 JTC를 이용한 스테레오 물체추적 시뮬레이터의 구현”, *대한전자공학회*, 제 36권, D편, 8호, pp. 68~78, 1999

## 저 자 소 개

## 李 學 贊(學生會員)

1999년 2월 : 광운대학교 컴퓨터공학과 졸업. 2001년 2월 : 광운대학교 대학원 컴퓨터공학과 석사. 2001년 3월~현재 숭실대학교 대학원 컴퓨터공학과 박사과정 재학 중. <주관심분야 : 스테레오 영상, 신경망,

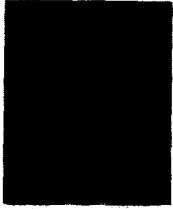
MPEG, 이산신호처리, 실시간 시스템, 내장형 시스템>



## 朴 章 漢(學生會員)

1997년 2월 : 방송통신대학교 전산계산학과 졸업. 1999년 8월 : 광운대학교 전산대학원 소프트웨어공학과 석사. 2001년 3월~현재 광운대학교 대학원 컴퓨터공학과 박사과정 재학 중. <주관심분야 : 스테레오 비전,

영상처리, 영상인식, 신경망, MPEG, 콤포넌트, UML>



南宮 璉(正會員)

1971년 2월 : 광운대학교 응용전자  
과 졸업. 1988년 2월 : 광운대학교  
대학원 컴퓨터공학과 석사. 1998년 9  
월~현재 광운대학교 대학원 컴퓨터  
공학과 박사과정 수료. 현재 인천기  
능대학 전자계산기과 교수 재직 중. <주관심분야 : 신  
경망, 문자인식, 영상처리>



南宮在贊(正會員)

1970년 2월 : 인하대학교 전기공학  
과 졸업. 1976년 8월 : 인하대학교  
대학원 전자공학과 석사. 1982년 2  
월 : 인하대학교 대학원 전자공학과  
박사. 1979년 3월~현재 광운대학교  
컴퓨터공학과 교수. <주관심분야 :  
신경망, 영상처리, 문자인식, 패턴인식, 영상통신,  
MPEG, 3D영상>