

# 주시각이 제어된 스테레오 물체추적 시스템의 광-디지털적 구현

정회원 고정환\*, 이재수\*\*, 김은수\*

## Opto-Digital Implementation of Convergence-Controlled Stereo Target Tracking System

Jun-Hwan Ko\*, Jae-soo Lee\*\*, Eun-soo Kim\* *Regular member*

### 요 약

본 논문에서는 계층적 디지털 알고리즘과 광 BPEJTC를 이용하여 주시각을 제어함으로써 이동표적을 적응적으로 추적할 수 있는 새로운 광-디지털 스테레오 물체추적 시스템을 제안하였다.

즉, 제안된 시스템에서는 먼저, 순차적인 입력영상으로 부터 배경정합을 통해 표적을 탐지하고 이어서, 영상 차분 필터, 논리곱 연산 및 모폴로지 필터를 이용하여 구성된 표적 투영마스크를 이용하여 표적물체를 영역화한 다음 최종적으로, 광 BPEJTC(binary phase extraction joint transform correlator)를 이용하여 표적물체의 위치정보를 추출해냄으로써 이를 이용한 실시간 카메라 주시각 제어 및 물체추적이 가능한 새로운 광-디지털 스테레오 물체추적 알고리즘을 제시하였다. 또한, 본 논문에서 새로이 제안된 표적물체 추출 및 카메라 주시각 제어 알고리즘의 광학적 구현을 통해 적응적 스테레오 물체 추적시스템의 실시간적 구현 가능성을 제시하였다.

### ABSTRACT

In this paper, a new opto-digital stereo object-tracking system using hierarchical digital algorithms and optical BPEJTC is proposed. This proposed system can adaptively track a moving target by controlling the convergence of stereo camera. Firstly, the target is detected through the background matching of the sequential input images by using optical BPEJTC and then the target area is segmented by using the target projection mask which is composed by hierarchical digital processing of image subtraction, logical operation and morphological filtering. Secondly, the location's coordinate of the moving target object for each of the sequential input frames can be extracted through carrying out optical BPEJTC between the reference image of the target region mask and the stereo input image. Finally, the convergence and pan/tilt of stereo camera can be sequentially controlled by using these target coordinate values and the target can be kept in tracking. Also, a possibility of real-time implementation of the adaptive stereo object tracking system is suggested through optically implementing the proposed target extraction and convergence control algorithms.

### 1. 서론

이동표적의 실시간적 추적기술은 지난 수십년 동안

안 여러 응용분야에서 많은 연구가 이루어져 왔으며, 최근에는 산업 및 군사기술의 발달과 더불어 무인 자동화 시스템의 요구가 높아짐에 따라 이에 대한 관심이 고조되고 있다<sup>[1-3]</sup>.

\* 광운대학교 전자공학과 국가지정 3차원 영상미디어 연구실 (eskim@daisy.gwu.ac.kr)

\*\* 김포대학 전자정보계열 (jslee@kimpo.ac.kr)

논문번호 : 010366-1203, 접수일자 : 2001년 12월 3일

※ 본 연구는 과학기술부 국가지정연구실 사업의 지원으로 수행되었습니다.

특히, 스테레오 로봇 비전(stereo robot vision) 기술을 이용한 물체추적 및 원격작업은 로봇에게 사람과 같은 입체시각을 부여함으로써 마치 현장에 있는 것과 같이 자연스러운 작업 처리가 가능함으로 기존의 2차원 시각에 비해 30% 이상의 원격작업 효율을 높여주는 것으로 알려지고 있다.<sup>[1]</sup> 이와 같은 스테레오 로봇 비전 시스템에는 입체시각을 부여하는 스테레오 카메라 기술, 스테레오 입력 영상을 처리하여 표적물체를 인지하는 물체인식 기술 그리고 표적물체를 연속적으로 추적하여 관측자가 표적물체를 계속 추적, 감시할 수 있는 스테레오 물체추적 기술등이 포함된다.

일반적으로, 양안시차에 기반한 스테레오 물체추적 시스템(stereo object tracking system)에서 좌,우 두 눈의 초점면에 표적물체의 시점이 일치하지 않게 되면 표적물체가 두 개로 겹쳐 보이게 됨으로써 눈의 피로감 가중과 작업효율의 저하를 가져오게 된다. 따라서, 추적물체에 스테레오 시차(stereo disparity)<sup>[5]</sup>가 존재할 경우에는 이 시차를 제거하게 되는데 이것을 주시각 제어(convergence angle control)<sup>[5-7]</sup>라고 한다. 즉, 일반적인 스테레오 물체추적 시스템에서는 양안시차<sup>[7-8]</sup>가 존재하면서 추적물체에 대한 스테레오 시차를 없애주기 위한 주시각 제어기술과 추적물체가 항상 카메라 시야(FOV: field of view) 중앙에 위치하도록 하는 카메라 팬/틸트 제어(FOV 제어) 기술등 두 가지 기능이 요구된다. 기존의 스테레오 물체 추적 시스템에서 주시각 제어 및 추적물체의 FOV 제어를 위해서는 스테레오 입력영상의 위치좌표를 찾아내기 위한 기준영상의 추출이 필수적이다. 이러한 기준영상의 영역추출 알고리즘은 대부분 2차원 순차 입력영상을 이용하여 이전 영상과 현재 영상사이의 관계 등으로부터 배경 및 잡음을 제거함으로써 추적물체를 식별하는 방법등이 사용되고 있다. 즉, 이러한 2차원적 영상처리에서 기준영상을 추출하기 위한 방법으로는 전, 후 프레임의 비교에 의한 차영상(image difference)방법, 사전에 정의된 특정 모델의 특징을 이용하는 모델기반(model-based) 방법, 광류(optical flow)에 의한 방법 및 블록매칭(block matching) 방법 등이 연구되고 있다<sup>[1-2]</sup>. 하지만 차영상에 의한 방법은 영상 검출기가 움직이고 있는 경우에는 이동 표적을 검출하는데 문제가 있다. 모델 기반에 의한 방법은 입력되는 영상이 미리 저장된 모델의 물체와 달리 상대적 명암이나 약간 변형되는 경우에는 매칭에 대한 어려움의 제약이 있다. 또한 광류에

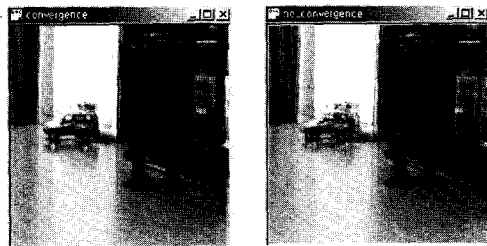
의한 방법은 화소 단위의 정합에 의한 방법이기 때문에 영상의 이동 변위가 클 경우 계산상의 오류가 발생하여 응용할 수 없고 일반적인 원근감을 갖는 정합은 어려운 것으로 보고되고 있으며, 블록에 기반한 방법<sup>[4]</sup>은 각 블록간을 비교하여 움직임을 추정하는 방법으로 가장 부합되는 블록을 찾기 위한 계산량과 전,후 프레임의 배경에 변화가 있는 경우에는 이동 물체를 추출하기가 어려운 단점이 있다.

따라서, 본 논문에서는 광 BPEJTC(binary phase extraction joint transform correlator)와 변형된 영상 차분필터 및 모폴로지 필터의 일종인 Convex hull 필터를 통해 이동표적을 영역화하고, 광학적 실험을 통해 배경잡음에 강하고 실시간적 주시각 제어가 가능한 새로운 광-디지털 스테레오 물체추적 시스템을 제안하고 제안된 시스템의 알고리즘적 분석과 광-디지털적 실험결과에 대해서 논하고자 한다. 즉, 제안된 시스템에서는 먼저, 광 BPEJTC를 이용하여 순차적인 두 입력 프레임간의 상관을 통해 배경을 정합한 다음, 전후 프레임 영상에 변형된 영상 차분 필터를 이용하여 표적영역을 탐지하고 변형된 Convex hull 필터를 적용함으로써 표적을 영역화하게 된다. 그리고, 영역화된 기준영상과 스테레오 입력 영상간에 광 BPEJTC<sup>[9-11]</sup>를 수행함으로써 추적물체의 위치정보를 구하고 최종적으로, 이를 이용하여 스테레오 카메라의 주시각 및 팬/틸트를 제어함으로써 이동물체의 실시간적 추적이 이루어지게 된다. 실험결과 convex hull 처리를 도입함으로써 기존의 차분필터와 BMA 등에서 발생하는 영역화 오류를 감소시켰으며, 광학적 구성이 가능한 BPEJTC를 통해 실시간으로 위상 상관관이 가능한 상관 침투치를 이용하여 정확히 추적물체를 추출하고 이를 이용하여 스테레오 물체 추적 시스템의 주시각 제어를 할 수 있었다.

## II. 스테레오 물체추적 시스템

일반적으로, 사람이 이동물체를 추적할 때 추적물체에 좌,우 두 눈의 시점을 일치시키는 것과 마찬가지로 스테레오 카메라를 이용한 물체추적 시스템에서 추적물체의 좌, 우 시점을 일치시키기 위해서는 카메라의 스테레오 시차가 "0"이 되어야 하며 이를 통해 눈의 피로감이 없는 자연스러운 관측과 작업 효율의 개선이 이루어 지게 된다. 반면에 좌,우 두 눈의 초점면에 추적물체의 시점이 일치하지 않아 스테레오 시차가 존재하는 경우는 추적물체가

이중으로 보이게 되어 눈에 피로감이 가중되게 되어 작업효율이 크게 떨어지게 된다. 그림 1(a), (b)는 추적물체에 대해 좌우 카메라의 스테레오 입력 영상이 주시점에 정확히 일치된 경우와 주시점이 다르기 때문에 스테레오시차가 생겨 화면에 2개의 영상이 겹쳐 나타난 예를 각각 보인 것이다.



(a) (b)

그림 1. 스테레오 카메라 영상 (a)주시각이 제어된 경우, (b)주시각이 제어되지 않은 경우

스테레오 물체추적 시스템은 추적물체를 인식, 추출하여 추적하고 추적물체를 화면의 중앙에 위치시키는 면에서는 기존 2차원 물체추적 시스템의 기능과 같지만 좌우 두 카메라의 스테레오 영상에서 추적물체에 대하여 주시점을 일치시키고 동시에 추적해야 하는 점이 다르다. 즉, 주시각 제어 및 스테레오 카메라의 팬/틸트 제어 기능이 추가된다.

여기서, 주시각 제어는 스테레오 카메라의 시점을 3차원 공간의 동원점(표적물체)상에 일치시키는 것으로 이 점은 스테레오 시차가 "0"이 되어 관측자의 좌우 두 눈의 초점이 일치하는 점이라고 할 수 있다. 또한, 스테레오 카메라의 팬/틸트 제어는 물체가 이동함에 따라 스테레오 카메라의 팬과 틸트를 제어하여 사람의 눈동자 움직임과 같이 추적 물체가 항상 좌우 카메라 시야(FOV)의 중앙에 놓이도록 하는 기능을 말한다.

일반적으로, 스테레오 물체추적 시스템에서 두 카메라의 입력영상으로 부터 추적물체의 현재 위치정보의 추출이 가능하면 카메라의 주시각 제어 및 팬/틸트 제어가 동시에 이루어 질 수 있는 반면에, 스테레오 영상처리에서 주시각 제어를 위한 스테레오 시차만 추출하는 경우에는 추적을 위한 물체의 현재위치 정보를 따로 구해야만 한다.

기존의 상관기법을 이용한 스테레오 영상처리의 경우는 추적물체의 현재 위치좌표의 추출이 가능함으로 이를 이용한 주시각 및 팬/틸트의 동시제어가 가능하다.

### III. 제안된 광-디지털 스테레오 물체추적

그림 2는 본 논문에서 새로이 제안한 광-디지털 스테레오 물체추적 시스템의 흐름도를 나타낸 것으로 크게 3 단계로 이루어진다.

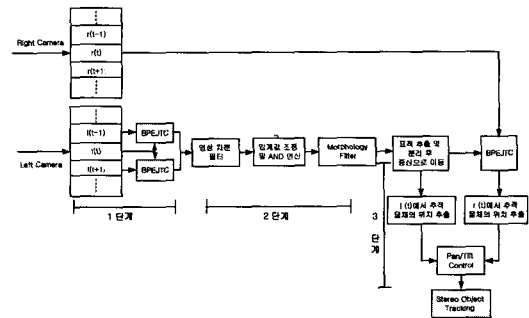


그림 2. 제안된 스테레오 물체추적 시스템의 흐름도

즉, 1 단계에서는 광 BPEJTC를 이용하여 연속적인 세 프레임 입력영상 사이의 상관을 통해 카메라의 이동 속도와 방향에 대한 상대적 움직임을 찾아내고 이를 이용하여 두 영상 사이의 배경을 정합하게 되며 이는 제 2 단계의 배경분리에 사용되게 된다. 2 단계에서는 영상 차분필터와 논리곱 연산 등을 적용하여 두 영상 사이의 순수한 표적 운동 성분만을 얻어 내고, 모폴로지 필터를 이용하여 마스크를 구성하여 기준영상에 투영시킴으로써 이동표적만을 영역화하게 된다.

마지막 3 단계에서는 영역화된 기준영상과 스테레오 입력 영상간에 광 BPEJTC를 실행하여 추적물체의 위치값을 구하고 여기서 구한 표적물체의 위치값을 가지고 최종적으로 스테레오 카메라의 주시각 제어뿐만 아니라 스테레오 카메라의 팬/틸트를 제어 하게된다.

#### 1. 제1 단계 : 표적물체의 배경정합

일반적으로, 순차적인 입력영상간의 배경이 일치되는 경우에는 영상 차분필터 만으로 이동표적의 움직임을 쉽게 검출할 수 있지만 배경과 표적이 모두 움직이는 실질적인 경우에는 배경제거에 의한 표적검출을 어렵게 한다.

따라서, 본 논문에서는 새로이 제안된 알고리즘의 효율성을 분석하기 위하여 카메라 변위가 표적 변위보다 큰 경우와 카메라의 변위가 표적 변위보다 작은 경우로 구분하여 추적 시나리오를 구성하였다.

그림 3은 카메라가 움직일 경우 t+1 프레임에서

의 좌,우 입력영상인  $l(t+1)$ ,  $r(t+1)$  을 나타낸 것으로 왼쪽은 카메라 변위가 표적 변위보다 큰 경우이며, 오른쪽은 표적의 변위가 카메라 변위보다 큰 경우를 각각 나타낸 것이다.

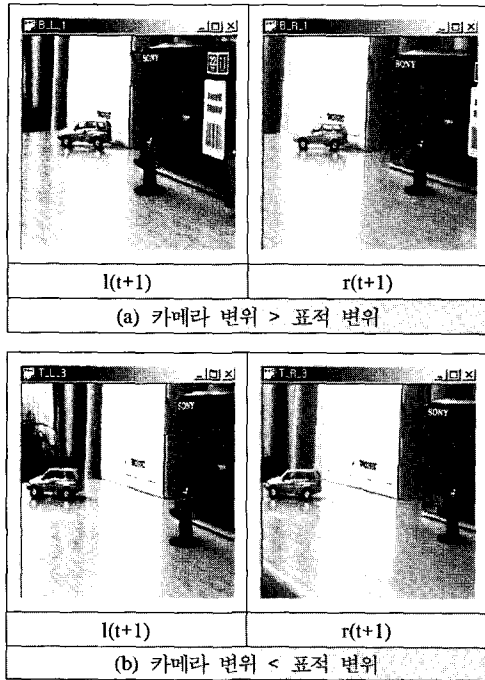


그림 3. 카메라가 움직이는 경우의 스테레오 입력영상

그림 4는  $t$  프레임의 좌측 입력영상인  $l(t)$ 를 기준으로 한 광 BPEJTC의 구성도를 나타낸 것이며 이를 이용하여 순차적인 입력영상인  $l(t-1)$ ,  $l(t)$  혹은  $l(t)$ ,  $l(t+1)$ 사이의 상관을 통해 배경과 표적의 상관도를 구하게 된다.

광 BPEJTC 시스템은 기존의 광 JTC 시스템에서 발생하는 과도한 DC나 상관오류 등을 제거하고 분리 조건에 보다 탄력적으로 대처할 수 있도록 제안된 위상형 광 상관기 시스템이다.[8] 즉, 광세기 분포로 나타나는 JTPS(joint transform power spectrum)에서 자기상관 성분과 동일영상 평면에서 나타나는 상호상관 성분을 효과적으로 제거하여 위상함수만 추출함으로써 상관 판별력을 보다 개선하고 광효율을 극대화한 시스템이다. 광 BPEJTC 시스템에서 이진위상 JTPS를 구성하기 위해서는 광 JTC의 JTPS 이외에 기준 및 입력영상에 대한 각각의 광 세기 분포가 필요하므로 그림 4의 CCD3과 CCD4를 통해 기준 및 입력영상에 대한 광 세기를 얻을 수 있도록 구성하였다.

광 BPEJTC 입력평면의 SLM1 영상은 렌즈에 의해 푸리에 변환되어 JTPS 형태로 CCD1에 검출되게 된다.

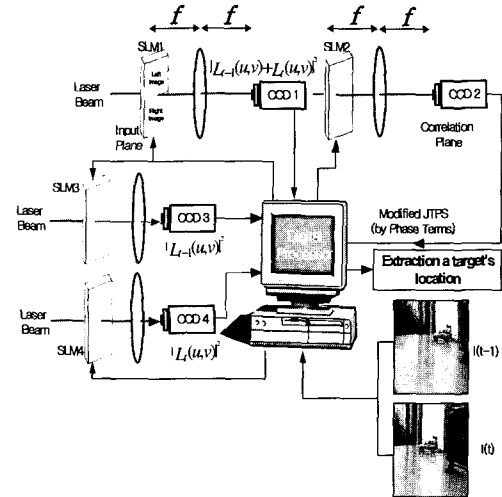


그림 4. 좌측영상에 대한 광 BPEJTC 시스템 구성도

그림 4의 입력평면은 그림 5와 같이 2단으로 분리하여 광 BPEJTC 입력평면으로 나타낼 수 있다.

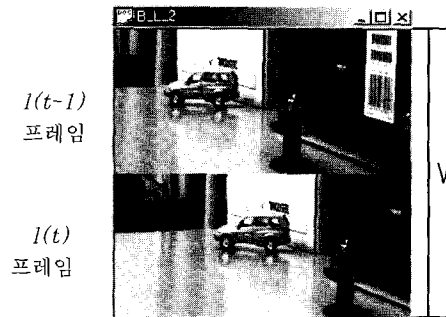


그림 5. 광 BPEJTC의 입력평면

임의의 시간에서 두 영상을 동시에 관찰하면 영상간의 거리는 영상 간격  $L$ 에 실제 이동 거리가 더해진 상태로 나타나게 된다. 식(1)과 식(2)의  $l_{t-1}(x,y)$ ,  $l_t(x,y)$ 는 기준영상인 이전영상과 현재영상을 각각 나타낸다. 그리고 식(1)에서  $t_1(x,y)$ 는 영상내에 포함된 표적을  $b_1(x,y)$ 은 두 프레임간에 동일하게 나타나는 배경을  $e_1(x,y)$ 는 입력장치에 의동으로 대상 프레임에는 존재하나 기준 프레임에서는 사라진 배경영상을 각각 나타낸다.

$$l_{t-1}(x,y) = t_1(x,y) + b_1(x,y) + e_1(x,y) \quad (1)$$

$$l_i(x, y) = t_2(x + x_T, y + y_T) + b_2(x + x_B, y + y_B) + e_2(x, y) \quad (2)$$

그리고, 식(2)에서  $t_2(x + x_T, y + y_T)$ 는 현재영상에 있는 표적이  $x$ 와  $y$ 축으로 ( $x_T, y_T$ )만큼 이동한 것을,  $b_2(x + x_B, y + y_B)$ 는 팬틸트의 흔들림에 따라 배경영상인  $b_1(x, y)$ 가 ( $x_B, y_B$ )만큼 이동한 것을 각각 나타낸 것이다. 또한,  $e_2(x, y)$ 는 표적의 이동과 검출기의 요동으로 기준 프레임에서 새로이 입력된 배경영상을 나타낸다. 여기서,  $x_T, y_T, x_B, y_B$ 는 상하 프레임 간격  $W/2$ 가 포함된 값이다.

이와 같이 구성된 두 프레임을 동시에 푸리에 변환하여 에너지 검출기로 검출할 수 있는 광세기 분포인 JTPS는 식(3)과 같이 나타낼 수 있다.

$$E(u, v) = |L_i(u, v)|^2 + |L_{i-1}(u, v)|^2 + L_i(u, v) L_{i-1}^*(u, v) + L_i^*(u, v) L_{i-1}(u, v) \quad (3)$$

위식에서,  $L(u, v)$ 는  $l(x, y)$ 의 푸리에 변환식음, \*는 복소공액을 각각 나타낸다. 식 (3)에서 세번째와 네번째 항이 두 프레임간의 상관스펙트럼으로 이 항을 추출하여 이치화 시키고 이진값으로 변환된 JTPS를 역푸리에 변환함으로써 최종적으로 BPEJTC 상관출력을 얻을 수 있다. 즉, 식 (4)는 재구성된 JTPS로써 영문 대문자는 소문자로 나타낸 공간영역 영상의 푸리에 변환함수를 나타낸 것이다. 여기서, 표적과 표적, 배경과 배경간의 자기상관은 높은 상관 침투치로 나타나며 나머지항들은 서로 다른 영상간에 발생한 상호상관 성분으로 비교적 작은 값으로 나타나게 된다.

$$\begin{aligned} &L_i(u, v) L_{i-1}^* + L_i^* L_{i-1}(u, v) \\ &= 2 |L_i(u, v)| |L_{i-1}(u, v)| \\ &\quad \cos[\phi_{L_i}(u, v) - \phi_{L_{i-1}}(u, v)] \\ &= T_2(u, v) \exp[-j2\pi(ux_T + vy_T)] \\ &\quad [T_1^*(u, v) + B_1^*(u, v) + E_1^*(u, v)] \\ &\quad + T_2^*(u, v) \exp[j2\pi(ux_T + vy_T)] \\ &\quad [T_1(u, v) + B_1(u, v) + E_1(u, v)] \\ &\quad + B_2(u, v) \exp[-j2\pi(ux_B + vy_B)] \\ &\quad [T_1^*(u, v) + B_1^*(u, v) + E_1^*(u, v)] \end{aligned}$$

$$\begin{aligned} &+ B_2^*(u, v) \exp[j2\pi(ux_B + vy_B)] \\ &\quad [T_1(u, v) + B_1(u, v) + E_1(u, v)] \\ &\quad + E_2(u, v) [T_1^*(u, v) + B_1^*(u, v) \\ &\quad + E_1^*(u, v)] + E_2^*(u, v) [T_1(u, v) \\ &\quad + B_1(u, v) + E_1(u, v)] \end{aligned} \quad (4)$$

따라서, 최종 출력 상관값은 식(5)와 같이 나타낼 수 있으며 식에서  $Edge[\cdot]$ 는 영상의 위상신호를 발생시킨 이상적인 경계를 나타낸 것이다. 결과적으로, 복잡한 환경을 가진 경우는 표적의 크기를 고려해 볼 때 배경성분이 표적성분보다 많이 존재하므로 최대 상관값의 위치는 배경 즉, 카메라의 이동성분으로 정의할 수 있다. 여기서,  $\otimes$ 는 상관을 그리고 \*는 컨볼루션을 각각 나타내며 실제 검출된 배경의 움직임 정도는  $\Delta x_B = x_B, \Delta y_B = y_B - W/2$ 로 구할 수 있다.

$$\begin{aligned} c(x, y) &= \{ Edge[t_1(x, y)] \otimes edge[t_2(x, y)] \} \\ &\quad * \delta(x - x_T, y - y_T) \\ &\quad + \{ Edge[t_1(x, y)] \otimes edge[t_2(x, y)] \} \\ &\quad * \delta(x + x_T, y + y_T) \\ &\quad + \{ Edge[b_1(x, y)] \otimes edge[b_2(x, y)] \} \\ &\quad * \delta(x - x_B, y - y_B) \\ &\quad + \{ Edge[b_1(x, y)] \otimes edge[b_2(x, y)] \} \\ &\quad * \delta(x + x_B, y + y_B) \end{aligned} \quad (5)$$

## 2. 제 2 단계 : 표적의 영역화

순차적인 3개 프레임의 입력영상이  $l_{i-1}(x, y), l_i(x, y), l_{i+1}(x, y)$ 로 주어지고, 광 BPEJTC를 통해 구한  $l_{i-1}(x, y)$ 와  $l_i(x, y)$ 간의 배경이동 정도가  $\Delta x_B, \Delta y_B$ 로 주어지는 경우, 영상간의 배경을 없애는 용이한 방법으로 영상 차분필터를 사용하게 되는데 그 결과는 식 (6), (7)와 같이 영상 사이에 움직임 성분이 나타나게 된다.

$$d_1(x, y) = |l_i(x, y) - l_{i-1}(x + \Delta x_B, y + \Delta y_B)| \quad (6)$$

$$d_2(x, y) = |l_i(x, y) - l_{i+1}(x - \Delta x_B, y - \Delta y_B)| \quad (7)$$

다음으로, 영상 차분필터의 결과로부터 공통된 움직임을 구별해 내기 위하여  $d_1$ 과  $d_2$ 의 값을 서로

논리곱 연산(AND operation)을 취함으로써 표적을 검출할 수 있다. 이런 과정을 통하여 연속으로 입력된 세 영상으로부터 운동성분을 찾아내게 된다. 즉, 표적 운동 성분값을  $T_m(x,y)$ 라고 하면 식(8)과 같이 나타낼 수 있으며 식에서  $B_i$ 는 이진화 과정을 의미한다.

$$T_m(u,v) = B_i \{ d_1(x,y) \text{ AND } B_i \{ d_2(x,y) \} \quad (8)$$

효율적인 이진화를 위하여 확률적인 방법으로 영역 전체의 분산을 최소화하는 방법을 사용하였다<sup>[12]</sup>. 즉, 제안된 방법에서는 영상차분 결과를 이진화 시키기 위하여 각 그룹내 분산(within-group variance)을 최소화하거나, 그룹간 분산(between-group variance)을 최대로 하는 임계값  $\tau$ 를 전 범위에 걸쳐 확률적으로 선택함으로써 전체 영상의 분산은 식(9)와 같이 주어지게 된다. 여기서,  $q_1(\tau)$ 는 임계값  $\tau$ 보다 작은 부분에서의 확률분포를,  $q_2(\tau)$ 는 임계값  $t$ 보다 큰 경우의 확률값을 각각 나타낸다. 그리고,  $v_1(\tau)$ 는 첫 번째 영역의 평균값을 그리고,  $v_2(\tau)$ 는 두 번째 영역의 평균값을 의미한다. 따라서, 각 영역내 분산을 최소화하며 각 영역들간의 분산차  $\sigma_B$ 가 최대가 되는  $t$ 를 찾음으로써 임계값을 결정하게 되며 그 결과는 식 (10)과 같이 주어진다.

$$\sigma_{total}^2 = [ q_1(\tau) \sigma_1^2(\tau) + q_2(\tau) \sigma_2^2(\tau) + \{ q_1(\tau) [ v_1(\tau) - v ]^2 + q_2(\tau) [ v_2(\tau) - v ]^2 \} \quad (9)$$

$$\sigma_B = q_1(\tau) q_2(\tau) [ v_1(\tau) - v_2(\tau) ]^2 \quad (10)$$

일반적으로, 모폴로지 필터는 영상처리 과정에서 없어지지 않고 남아있는 배경 등에 대한 잡음을 없애거나 윤곽선 검출, 필터링, 세션화 등의 처리에 광범위하게 사용되고 있다.<sup>[6]</sup>

본 논문에서는 움직임 검출 결과로부터 기준영상에서의 표적 영역화를 위하여 Convex hull 필터와 클로징(closing) 기법의 순차적인 반복과정을 적용하였다. Convex hull 필터와 클로징 기법을 영상 전반에 적용함으로써 경계선을 평활시킬 뿐만 아니라 각 경계면 안에 존재하는 빈 공간을 채움으로써 표적 마스크를 만들게 된다. 한편, 기존의 Convex

hull 필터는 그림 6(a)와 같이 4개의 기본 구성요소로 이루어져 있으나 이러한 기존의 구성요소들을 각 영상에 적용하는 경우 실제 표적보다 큰 잉여적인 외곽 패턴이 발생하여 정확한 영역화를 기대할 수 없다.

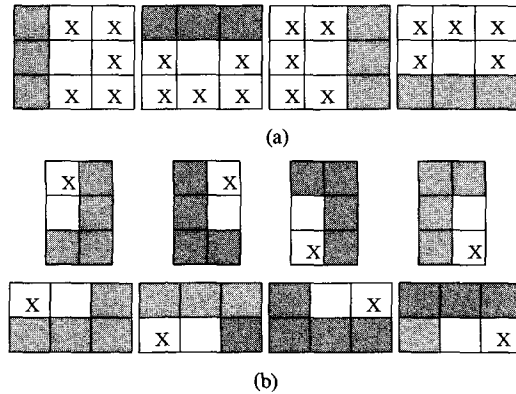


그림 6. Convex hull 필터 (a) 기존의 구성요소, (b) 변형된 구성요소

따라서, 가능한 원래 표적에 가까운 영상을 추출하기 위하여 본 논문에서는 그림 6(b)와 같이 8개의 변형된 마스크를 설정하였다. 입력영상을  $A$ , 구성요소를  $B^i$  라고 하고 Convex hull 필터를 적용하기 전의 초기조건으로써  $X_0^i = A$  라고 정의하면 전체 수행과정은 식 (11), (12)과 같이 나타낼 수 있다.

$$X_0^i = (X \cap B^i) \cup A \quad i=1,2,\dots,7,8, \quad (11)$$

$$C(A) = \bigcup_{i=1}^8 D^i \quad (12)$$

식 (11)에서  $\cap$ 는 hit-or-miss 변환[6]을 나타낸다. 구성요소를 하나씩 적용시킨 결과를  $X_k^i$  라고 할 때, 하나의 구성요소에 대해서 완전하게 수렴되는 결과를 찾기 위한 일시적 과정을  $D^i = X_{conv}^i$  라고 표현하게 된다. 따라서, 이는 새로 구해진 결과  $X$ 가 이전값에 수렴되는 조건인  $X_k^i = X_{k-1}^i$  이 만족될 때까지 계속해서 반복적으로 계산을 한다는 의미로써 CONV 란 첨자를 삽입하게 된다. 결국, 최종적인 Convex hull 필터 결과인  $C(A)$ 는 구성요소를 전부 적용시킨 영상 결과  $D^i$ 를 전부 합함으로써 얻을 수 있다. 따라서, 표적검출 결과로부터 이동영역 마스크를 설정하기 위하여 Convex

hull 과 클로징 과정이 결합된 새로운 모폴로지 필터를 식 (8)에 추가 적용함으로써 식 (13)와 같은 표적 마스크  $T_{mask}(x, y)$ 를 설정하게 된다.

$$T_{mask}(x, y) = MORPHOLOGY[T_m(x, y)] \quad (13)$$

후단의 인식과정을 고려하여 표적위치를 좌표값으로 영역화하기 위하여 식 (13)에 의해 구해진 표적 마스크를 기준영상  $l_i(x, y)$ 로 투영시키면 표적만을 부각시킬 수 있는데 검출된 마스크를  $x, y$  방향으로 각각 투영하는 방법을 새로이 도입하였다.

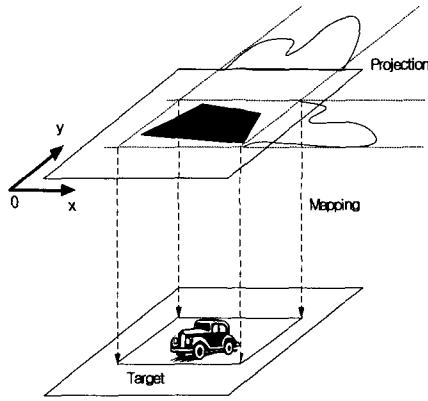


그림 7. 표적 영역화를 위한 투영

즉, 그림 7은 제안된 투영과정을 나타낸 것으로 움직임 검출값으로부터 먼저  $x$ 축으로 투영시켜  $x$ 축상에 존재하는 표적의 위치를 알아낸 다음 다시  $y$ 축으로 투영시킴으로써  $y$ 축상에 존재하는 표적의 정확한 위치를 찾아내어 배경으로부터 표적만을 분리해 낼 수 있게 된다.<sup>[9]</sup> 또한, 배경이 분리된 좌측 영상의 윈도우 마스크에서 추적 물체의 위치좌표  $(\pm \Delta x_l, \pm \Delta y_l)$ 를 추출하여 중앙으로 위치하도록  $(\pm \Delta x_l, \pm \Delta y_l)$ 만큼 이동시킨 후에, 광 BPEJTC 입력 평면의 기준영상으로 사용하게된다. 그리고 이동값인  $(\pm \Delta x_l, \pm \Delta y_l)$ 은 좌측 카메라의 팬/틸트 제어값이 된다.

### 3. 제 3단계 : 표적물체의 위치정보 추출

표적 물체의 위치정보를 구하기 위하여 표적 마스크의 기준영상  $T_{mask}(t)$ 과 우측 영상인  $r(t)$ 간에 광BPEJTC를 통해 두 물체간에 떨어진 상대거리 값을 구하여 우측 카메라의 주시각 제어값으로 사용하게 된다.

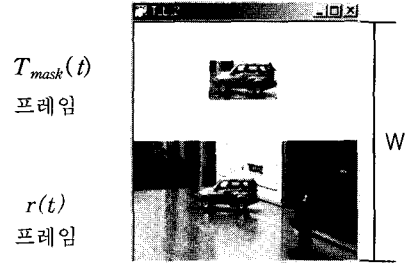


그림 8. 광 BPEJTC의 입력평면

여기서, 표적의 위치정보를 추출하기 위해 사용되는 광 BPEJTC의 수행절차는 제 1단계에서 설명한 배경정합 과정과 같으므로 생략하고, 그 결과값에 대해서 논의하기로 한다.

즉, 광 BPEJTC 입력평면  $[T_{mask}(t), r(t)]$ 에서 SLM1의 영상은 렌즈에 의해 푸리에 변환되어 JTPS 형태로 CCD1에 검출되게 된다. 그림 4의 입력평면은 그림 8과 같이 2단으로 분리하여 광 BPEJTC 입력평면으로 나타낼 수 있다. 그림 4에서 광 BPEJTC 입력으로 사용되는 기준영상  $T_{mask}(t)$  과 입력영상  $r(t)$ 의 추적물체의 위치좌표는  $(x, y + \frac{w}{2}), (x - \Delta x_r, y - [\frac{w}{2} - \Delta y_r])$ 로 각각 주어진다. 그리고 이진 위상 형태로 변형된 JTPS를 SLM2의 입력으로 사용하고 역 푸리에 변환하여 상관 평면에서 얻은 결과는 식 (14)와 같이 얻을 수 있다.

$$\begin{aligned} C_{BPEJTC}(x, y) &= \mathcal{F}^{-1} E_{BPEJTC}(u, v) \\ &= Edge[T_{mask}(x, y)] \\ &\quad \otimes Edge[T_{mask}^*(x, y)] \\ &\quad * \delta[x + \Delta x_r, y + (w - \Delta y_r)] \\ &\quad + Edge[T_{mask}(x, y)] \\ &\quad \otimes Edge[T_{mask}^*(x, y)] \\ &\quad * \delta[x - \Delta x_r, y - (w - \Delta y_r)] \quad (14) \end{aligned}$$

식 (14)는 그림 4의 상관평면인 CCD2 에서 상관 침두치로 검출되며 이 침두치는 중심좌표에 있는 기준영상과 우측영상 간에 떨어진 상대거리로 침두치의 위치좌표를 얻을 수 있다. 또한, 이 위치좌표로부터 우측영상의 추적물체 위치좌표  $(\pm \Delta x_r, \pm \Delta y_r)$ 값과 우측 카메라의 팬/틸트 제어값  $(\pm \Delta x_r, \pm \Delta y_r)$ 을 각각 구할 수 있다.

식 (14)에서,  $Edge[\cdot]$ 는 영상의 이상적인 경계선 추출함수를 나타낸 것으로, 식(15)에서 보면 푸리에

변환하여 진폭성분을 제거하고 다시 역 푸리에 변환하여 영상의 경계추출 결과를 얻을 수 있음을 알 수 있다. 결과적으로 광 BPEJTC 는 DC와 배경으로 인하여 발생하는 상관오류를 제거할 수 있을 뿐만 아니라, 공간 정합필터에서 우수한 peak-to-sidelobe를 나타내는 특성을 가지고 있다.

$$Edge [r(x, y)] = \mathcal{F}^{-1} \{Ph[R(u, v)]\}$$

$$Where, Ph[R(u, v)] = \left( \frac{R^*(u, v)}{R(u, v)} \right)^{0.5}$$

$$= 1 \cdot \exp[-j\Phi_r(u, v)] \quad (15)$$

따라서,  $l(t)$  영상의 전처리에서 구한 좌측 카메라의 제어값인  $(\pm \Delta x_l, \pm \Delta y_l)$ 과 광 BPEJTC를 통해 구한 우측 영상의 제어값인  $(\pm \Delta x_r, \pm \Delta y_r)$ 으로 부터 스테레오 카메라의 팬/틸트를 제어함으로써 스테레오 물체추적이 가능하게 된다.

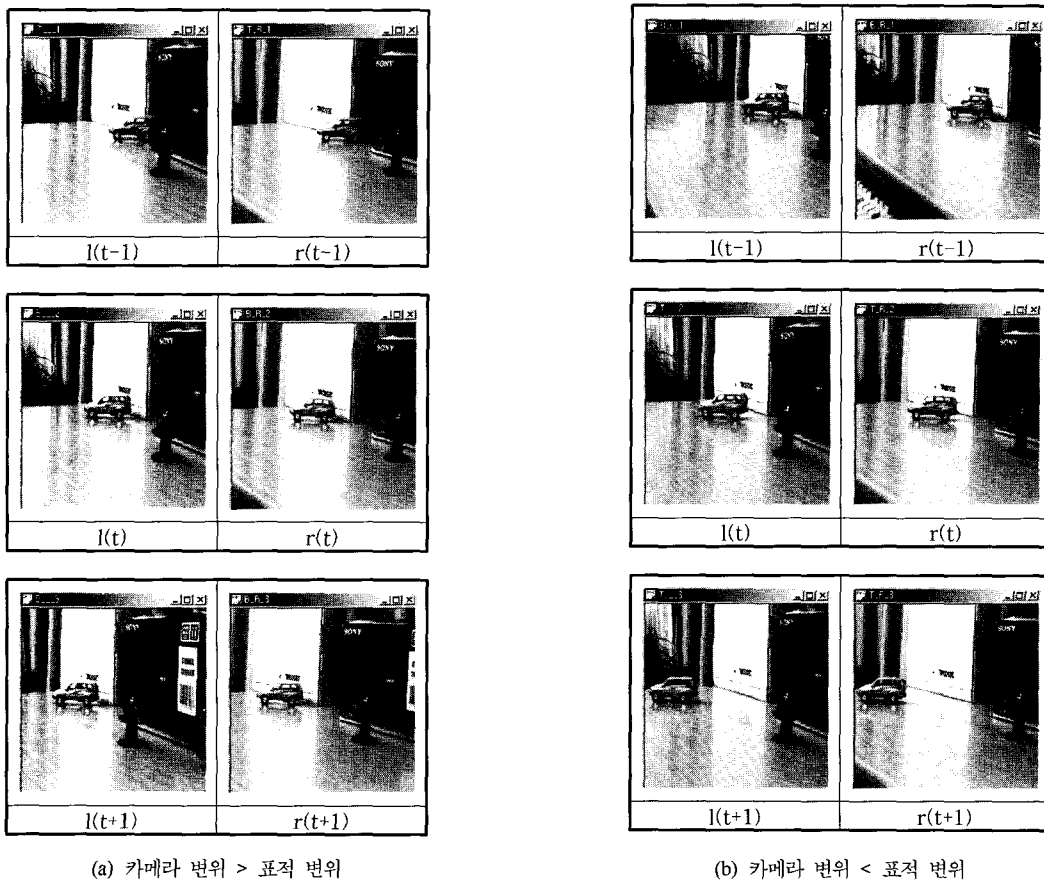
#### IV. 실험 및 결과 고찰

본 실험에서 좌, 우측 카메라의 영상 입력용 카메라는 동경전자(주)의 CS-8239B 카메라를 사용하였으며, 이동표적(자동차)을 포함하는 입력영상 데이터의 크기는 256x256 픽셀로 사용하였다. 영상 저장용 프레임 그래픽은 Matrox사의 Metro II/4와 Metro II MC/2의 2개를 사용하여 저장하였으며 디지털 시스템은 펜티엄 III-800 (256MB)를 사용하였다.

##### 1. 제 1단계 : 표적물체의 배경정합 실험

제한된 알고리즘의 효율성을 보이기 위한 시나리오인 카메라 변위가 표적 변위보다 클 경우와 카메라 변위가 표적 변위보다 작은 경우로 구분하여 실험하였다.

그림 9는 카메라가 움직일 경우의 입력영상을 나



(a) 카메라 변위 > 표적 변위

(b) 카메라 변위 < 표적 변위

그림 9. 스테레오 입력 영상



타낸 것으로 왼쪽은 카메라 변위가 표적 변위보다 큰 경우이며 오른쪽은 표적의 변위가 카메라 변위보다 큰 경우를 나타낸 것이다.

배경정합을 위한 기준영상은 2 번째 프레임인 I(t)로 하였으며 기준영상과 비교하여 순차적으로 입력되는 영상들을 연속적으로 상관시켜서 얻은 최대 상관값 위치와 카메라의 이동변위 성분을 표1에 나타내었다.

표 1. 이동변위 성분

카메라 변위 > 표적 변위		
1, 2 번째 영상	상관값 위치	(117, 116)
	변위 성분	$\Delta x_f = -16, \Delta y_f = 5$
2, 3 번째 영상	상관값 위치	(164, 96)
	변위 성분	$\Delta x_f = 17, \Delta y_f = 16$

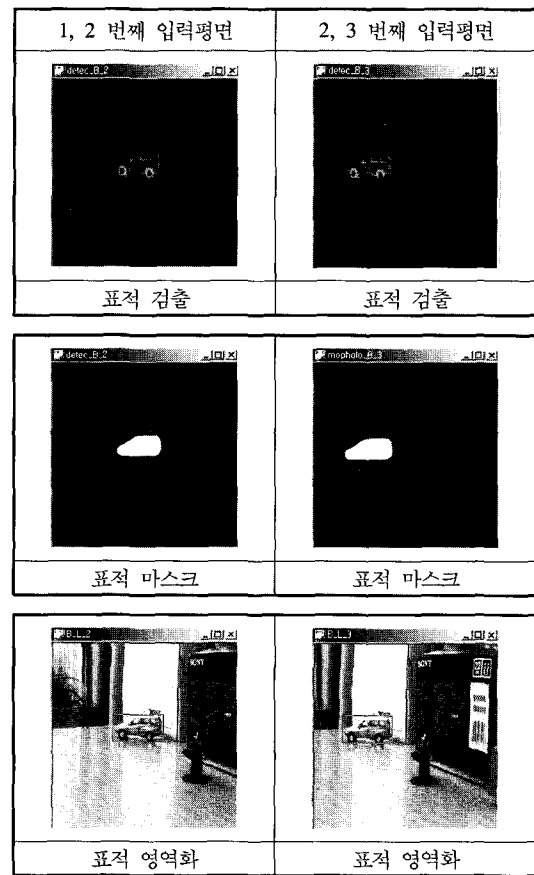
  

카메라 변위 < 표적 변위		
1, 2 번째 영상	상관값 위치	(104, 123)
	변위 성분	$\Delta x_f = -5, \Delta y_f = 0$
2, 3 번째 영상	상관값 위치	(96, 117)
	변위 성분	$\Delta x_f = 0, \Delta y_f = 1$

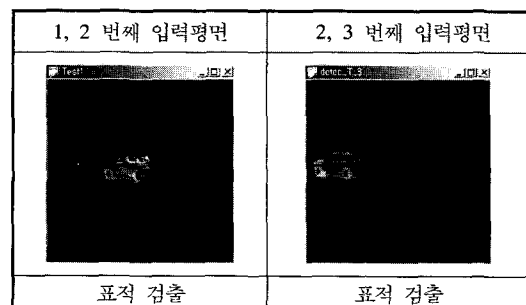
그리고, 변위성분의 부호는 대상영상과 기준영상 사이에 배경정합을 위하여 상대적으로 움직여야 하는 방향으로 주어지게 된다. 즉, 표 1에서 카메라 변위가 표적 변위보다 클 때 1, 2 번째 영상을 상관시킨 결과를 보면 x방향으로 음의 값을, y 방향으로는 양의 값의 결과가 나왔다. 이는 기준 영상을 중심으로 할 때 x 방향으로는 기준영상보다 오른쪽에 위치하고, y 방향으로는 왼쪽에 위치한다는 것을 의미한다. 따라서, 첫 번째 영상 [I(t-1)]을 두 번째 영상 [I(t)]의 배경과 정합시키기 위하여 왼쪽으로 16 화소, 아래쪽으로 5화소를 이동시켜 주면 영상간 배경이 일치하게 된다. 그리고, 이치화 과정, 논리곱 연산등을 이용하여 두 결과값을 결합시켜 공통된 움직임 부분만을 계산해 내게 된다.

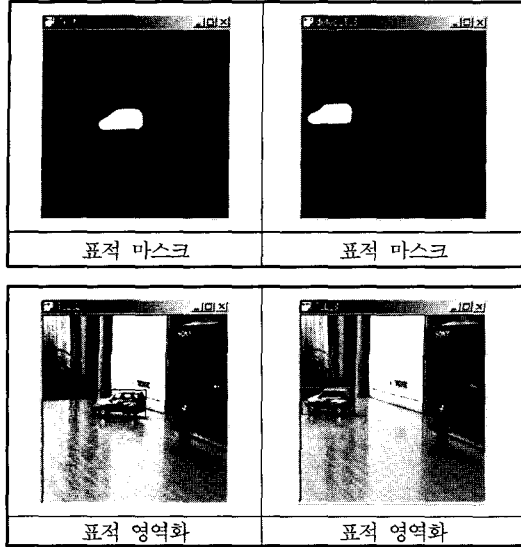
2. 제 2 단계 : 표적물체의 이동성분 추출 실험  
입력영상 사이의 상관을 통해 배경을 정합시킨 후 영상 차분필터를 이용함으로써 배경을 효과적으로 제거하게 되는데 그림 10은 검출된 표적의 움직임 값을 이용하여 Convex hull 필터와 클로징 기법을 영상 전반에 적용함으로써 경계선을 평활시킬 뿐만 아니라 각 경계면 안에 존재하는 빈 공간을

채움으로써 표적 마스크와 이 결과값을 기준영상에 투영시켜 매핑함으로써 영역화된 결과를 보여준 것이다. 그림 10을 보면 배경잡음에 관계없이 추적하고자 하는 표적물체의 효과적인 검출과 그에 따른 표적 마스크가 적응적으로 이루어지고 있으며, Convex hull 처리후에 일정 크기의 영역화 마스크만을 선택하기 때문에 전경과 배경에 존재하는 다른 물체와 관계없이 적응적인 표적탐지가 가능함을 알 수 있다.



(a) 카메라 변위 > 표적 변위





(b) 카메라 변위 < 표적 변위

그림 10. 좌측영상의 표적검출 및 영역화

3. 제 3 단계 : 표적물체의 위치추출 실험

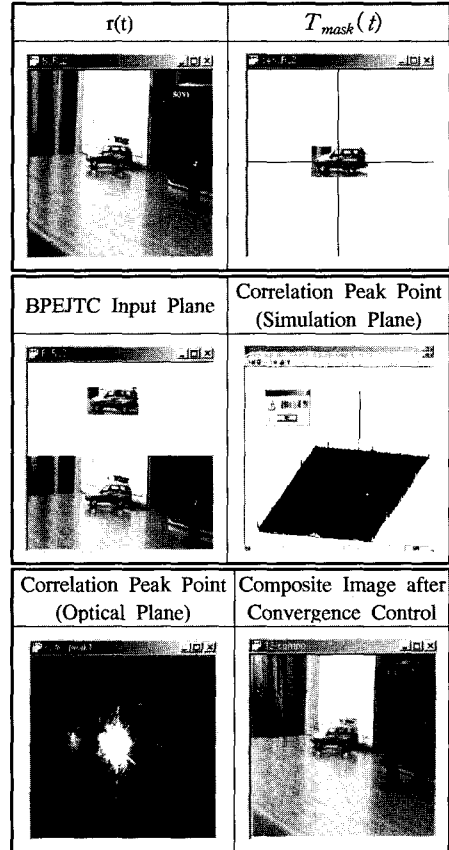
그림 11은 카메라 변위가 표적 변위보다 클 때의 표적검출 및 영역화된 좌측영상 [  $T_{mask}(t)$  ]과 현재 입력되는 우측영상 [  $r(t)$  ]간에 광 BPEJTC를 실행하기 위한 입력 평면과 광 BPEJTC를 실행한 후 상관 평면에서 나타나는 좌, 우측 영상의 상관 침두치를 결과를 각각 나타낸 것이다. 이러한 상관 침두치의 위치 좌표값을 이용함으로써 좌, 우 영상의 추적물체가 화면의 중심으로 부터 떨어진 거리인 ( $\pm \Delta x_l, \pm \Delta y_l$ )와 ( $\pm \Delta x_r, \pm \Delta y_r$ )를 구할 수 있다. 표 2는 순차적인 4개 프레임의 좌, 우 입력영상에 대해서 본 논문에서 제안한 추적 알고리즘을 사용하여 최종적으로 구한 추적물체의 이동된 위치 좌표 ( $x, y$ )를 나타낸 것이다. 여기서, 위치좌표는 입력영상(256x256)의 중심좌표(128,128)를 기준으로 하여 x, y축 방향으로 이동된 추적물체의 위치를 각각 나타낸 것이다.

표 2. 추적물체의 위치좌표 추출

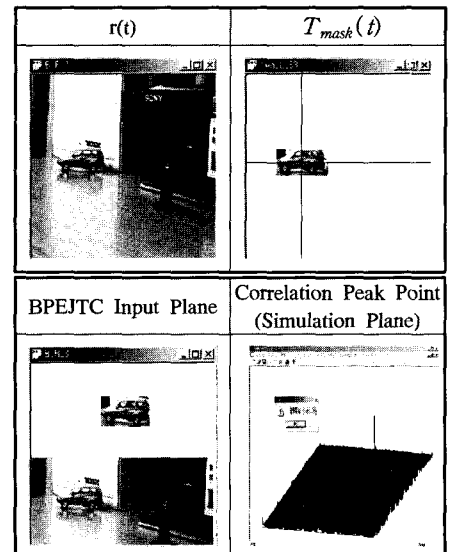
Frame Coordinate	카메라 변위 > 표적 변위	
	Left	Right
1,2 번째(x, y)	(-10, -14)	(-9, 0)
2,3 번째(x, y)	(-51.5, -7)	(-49, -20)

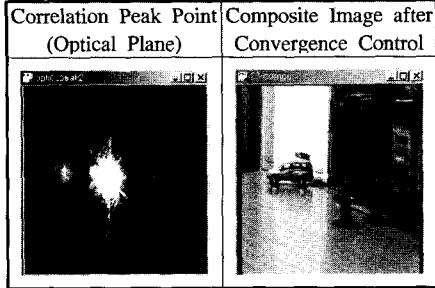
Frame Coordinate	카메라 변위 < 표적 변위	
	Left	Right
1,2 번째(x, y)	(-21, -4.5)	(-30, -9)
2,3 번째(x, y)	(-87, -1)	(-91, -13)

그리고 그림 11의 우측 하단에 있는 각각의 영상은 그림 9의 입력 영상에서 제안된 방법을 사용하여 구한 표적물체의 이동성분인 표 2를 통해 추적한 후에 합성된 스테레오 영상을 나타낸 것이다.

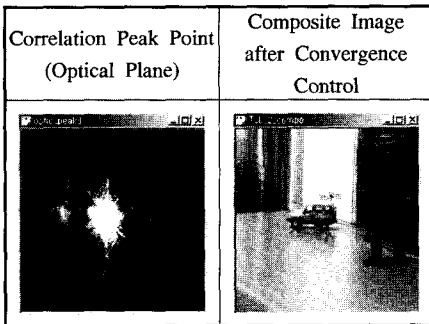
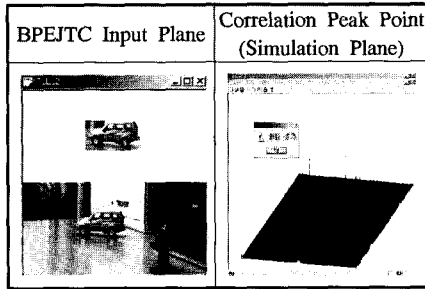
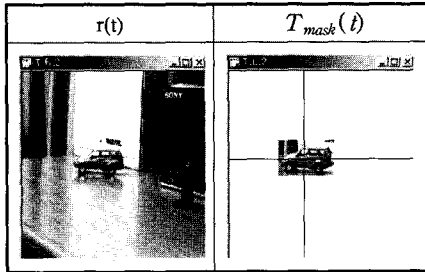


(a) 카메라 변위 > 표적 변위  
(1,2 번째 입력 평면)

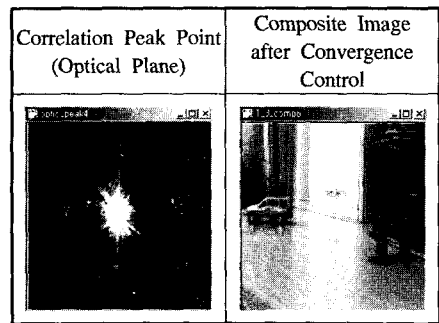
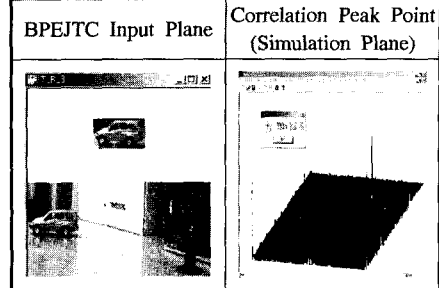
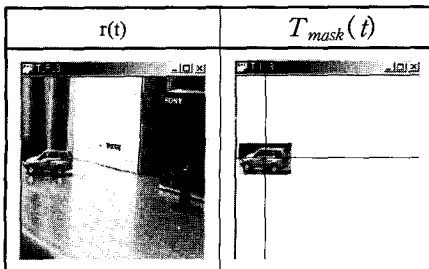




(b) 카메라 변위 > 표적 변위  
(2,3 번째 입력 평면)



(c) 카메라 변위 < 표적 변위  
(1,2 번째 입력 평면)



(d) 카메라 변위 < 표적 변위  
(2,3 번째 입력 평면)

그림 11. 광 BPEJTC 시뮬레이션 결과

그림 11에서 보면 추적물체(자동차)는 본 논문에서 제안한 추적 알고리즘을 통해 스테레오 카메라의 주시각 및 팬/틸트가 정확히 제어되어 좌, 우 시점이 표적물체에 일치되는 하나의 합성영상으로 나타나지만 그 이외의 전경 및 배경 물체들은 스테레오 시차로 인해 2개로 겹쳐 보임을 알 수 있다.

결과적으로, 본 논문에서는 계층적인 디지털 영상 처리 알고리즘과 광 BPEJTC 시스템을 상호보완적으로 사용함으로써 스테레오 카메라의 주시각 및 팬/틸트를 제어할 수 있는 새로운 광-디지털 스테레오 물체추적 시스템을 제시하였으며 또한, 본 논문에서 새로이 제안된 표적물체 추출 및 스테레오 카메라 주시각 제어 알고리즘의 광학적 구현을 통해 적응적 스테레오 물체 추적시스템의 실시간적 구현 가능성도 제시하였다.

## V. 결 론

본 논문에서는 스테레오 물체추적의 새로운 접근 방법으로 계층적 디지털 알고리즘과 광 BPEJTC를 상호 보완적으로 이용한 새로운 하이브리드 광-디지털 스테레오 물체추적 시스템을 제안하였다.

즉, 제안된 물체추적 시스템에서는 먼저, 순차적으로 입력되는 표적영상에 대한 배경정합을 수행하

고 이어서, 영상 차분 필터, 논리곱 연산 및 모폴로지 필터를 이용하여 구성된 표적 투영마스크를 통해 표적물체를 영역화한 다음 최종적으로, 광 BPEJTC를 이용하여 표적물체의 위치정보를 추출하였다. 그리고, 추출된 표적의 위치정보를 이용하여 카메라의 주시각 및 팬틸트를 제어함으로써 이동물체의 실시간적 추적이 가능한 새로운 광-디지털 스테레오 물체추적 시스템을 제시하였다. 또한, 본 논문에서 새로이 제안된 표적물체 추출 및 스테레오 카메라 주시각 제어 알고리즘의 광학적 구현을 통해 적응적 스테레오 물체 추적시스템의 실시간적 구현 가능성을 제시하였다.

참 고 문 헌

[1] K. T. Kim, J. S. Lee, S. H. Kim, E. S. Kim, "Human-like stereo vision system based-on optical JTC", *SPIE Optical Pattern Recognition*, vol.3466, pp.259-266, 1998

[2] Stephen T. Barnard, and William B. Thompson, "Disparity analysis of images", *IEEE Trans. PAMI*, vol. PAMI-2, no.4, pp.333-340, July, 1980

[3] R.C Jane, "Segmentation of frame sequences obtained by a moving observer," *IEEE Trans. PAMI*, vol.6, no.5, pp.624-629, 1984

[4] J. S. Lee, J. H. Ko, E. S. Kim, "Real-time stereo object tracking system by using block matching algorithm and optical binary phase extraction joint transform correlator", *Optics Communications*, vol. 191, pp.191-202, 2001

[5] Y. Mae, S. Yamamoto, Y Shirai and J. Miura, "Optical flow based real-time object tracking by active vision system," *Proc. of 2nd Japan-France Congress on Mechatronics*, vol.2, pp.545-548, 1994

[6] Rafael C. Gonzalez and Richard E. Woods, "Efficient block motion estimation using integral projections", *Digital Image Processing, Addison-Wesley*, 1992

[7] C. Tam, T.S. Yu, A, Gregory, D. Juday, "Autonomous real time object tracking with an adaptive JTC," *Opt. Eng.*, vol.29, no.4, pp.314-320, 1990

[8] T. J. Olson and D. J. Coombs, "Real-time

vergence control for binocular robots," *Intl. J. of Computer Vision*, vol.7, no.1, pp.67-89, 1991

[9] Gosta H. Granlund and Hans Knutsson, "Signal Processing for Computer Vision", *Kluwer Academic Publishers*, 1995

[10] J. H. Ko, J. S. Lee, C. W. Seo, E. S. Kim, "Target extraction using stereo disparity information in stereo vision", *Proc. of the 32nd ISR(International Symposium on Robotics)*, vol.2, pp. 724-727, 2001

고 정 환(Jung-hwan Ko)

정회원



1999년 2월 : 광운대학교 제어계측공학과 졸업  
 2001년 8월 : 광운대학교 전자공학과 석사  
 2001년 9월~현재 : 광운대학교 전자공학과 박사과정

<주관심 분야> 3D 영상처리, 로봇 비전, 컴퓨터 비

이 재 수(Jae-soo Lee)

정회원

현재 : 김포대학 전자정보계열 조교수  
 한국통신학회지 논문지 제26권 9호 참조

김 은 수(Eun-soo Kim)

정회원

현재 : 광운대학교 국가지정 3차원 영상미디어 연구실 및 전자공학부 교수  
 (<http://3dmedia.kwangwoon.ac.kr>)  
 한국통신학회지 논문지 제26권 11호 참조