

# 인간시각체계에 기초한 교차시각 스테레오 영상의 깊이 추출

정회원 이 적 식\*

## Depth Extraction of Convergent-Looking Stereo Images Based on the Human Visual System

Juck Sik Lee\* *Regular Member*

요 약

현재까지 스테레오 시각 분야에서 사용된 카메라 구조는 대부분 두 광축이 평행한 시각 모델이었으나, 본 논문에서는 인간시각체계에 기반한 교차시각 스테레오 카메라 모델을 사용한다. 교차시각에서 획득한 스테레오 영상을 고정점에 대해서 4개의 영역으로 구분한다. 이러한 사분면 형태의 영역 분할 근거를 인간시각체계에서 살펴보고, 그 타당성을 기하학적으로 입증할 것이다. 그리고 분할된 한 쌍의 스테레오 영상으로부터 합성 영상 조각들을 형성한 후, 수정된 캡스트럼 작용과 피크 검출 알고리즘을 수행하여, 수평 및 수직 변위를 추출한다. 측정된 변위 벡터와 교차시각 스테레오 카메라 모델의 변수로부터 합성영상의 3차원 정보를 계산한다. 여러 스테레오 영상에 대해서 제안 방법의 결과가 인간시각체계처럼 고정점 부근에서 정확함을 보여주었다.

ABSTRACT

A camera model with optical axes parallel has been widely used for stereo vision applications. A pair of input images are obtained from a convergent-looking stereo camera model based on the human visual system in this paper, and each image is divided into quadrant regions with respect to the fixation point. The reasoning of quadrant partitions is based on the human visual system and is proven by a geometrical method. Image patches are constructed from the right and left stereo images. A modified cepstrum filter is applied to the patches and disparity vectors are determined by peak detection algorithm. The three-dimensional information for synthetic images is obtained from the measured disparity and the convergent stereo camera model. It is shown that the experimental results of the proposed method for various stereo images are accurate around the fixation point like the human visual system.

### I. 서론

3차원 물체의 구조를 추론하기 위해서 인간이 사용하는 많은 단서 중에서, 양안 입체 시각은 가장 큰 비중을 차지한다. 이 방법은 왼쪽과 오른쪽 망막에 투영되는 2차원 스테레오 영상들로부터 3차원 위치 정보를 결정하는 것이다. 한 쌍의 스테레오 영상이 주어질 때 3차원 정보를 계산하기 위해서는

실세계 물체점에 해당하는 대응점을 좌우 영상에서 찾아야 한다. 이 과정이 정합이며 컴퓨터 시각에서는 크게 두 가지 방법을 이용한다. 어떤 크기의 창을 이용하여 두 영상내 밝기 신호의 상관관계를 측정하여 가장 높은 상관계수를 갖는 영역을 서로 정합된 것으로 결정하는 영역 기반 방법과 영상에서 추출된 특징으로 정합을 시도하는 특징 기반 방법이 있다.

지금까지 제안된 대부분 양안 입체 알고리즘들이

\* 경기대학교 전자기계공학부 전자공학전공(jslee@kuic.kyonggi.ac.kr)  
논문번호 : K01222-1127, 접수일자 : 2001년 11월 27일

사용한 카메라 모델은 두 광축이 평행인 구조를 갖고 있다. 이러한 평행시각 스테레오 카메라 모델에서 획득한 좌우 영상으로부터 대응점을 찾기 위해서는 epipolar 직선상에서 단일 탐색 방향이 요구된다. 평행 스테레오 영상에서 깊이를 인지하는 여러 가지 방법을 Dhond와 Aggarwal이 설명하였고<sup>[1]</sup>, Jones는 다양한 특징으로 영상을 정합하는 것은 일반적인 최적화 문제의 부분으로 고려하여 기존의 방법들을 비교하였다<sup>[2]</sup>. 이러한 접근 방법은 교차시각구조를 갖는 인간의 시각 체계와 같은 생물학적 모델로 받아들이기에는 어려움이 있기 때문에 서서히 교차시각 스테레오 카메라 모델에 대한 관심이 고조되고 있다<sup>[3-5]</sup>. 교차시각에서는 epipolar 직선들이 영상평면의 수평 방향과 평행이 아니다. 그러므로 기존의 교차 영상 정합은 epipolar 직선들을 평행이 되도록(즉 영상평면이 기준선과 평행) 하는 정류(rectification) 과정을 거친 후, 평행시각과 동일한 방법으로 정합을 시도하였다<sup>[6]</sup>. 이 방법의 단점은 정류된 영상의 픽셀좌표가 원 영상에서 정수가 아닌 픽셀 좌표에 대응하기 때문에 픽셀 위치 계산시 보간이 필요하다는 것이다.

인간 시각 체계는 평행시각이 아니라 교차시각과 유사하며, 머리를 좌우상하로 움직임으로써 주된 관심 물체를 두 눈의 광축이 만나는 고정점에 위치시켜, 망막의 중심와(fovea)에 투영되도록 하여 정확한 3차원 정보를 추출할 수 있다. 인간의 입체시각 경로와 원리의 이해는 원숭이와 고양이 등을 실험 대상으로 한 연구 결과에서 찾아볼 수 있다. 많은 포유동물의 시각피질(visual cortex)에서 양안 변위(binocular disparity)에 반응하는 신경세포의 존재가 발견되었고, 이 세포들이 입체시각(stereopsis)에 중요한 역할을 담당하는 것으로 여겨져 왔다<sup>[7]</sup>. 일차 시각피질에는 자극 위치와 방향에 민감한 단순세포(simple cell)들이 있으며, 이 세포들의 수용영역(receptive field)은 왼쪽과 오른쪽 망막 둘 다 존재하며 수용영역의 함수 형태에 따라서 크게 두 가지 변위 추출 모델을 신경생리학적인 측면에서 제안하였다. 두 수용영역의 모양은 같지만 위치가 다른 위치이동(position shift) 모델과 위상이 다른 위상이동(phase shift) 모델이며, 어느 모델로도 양안 변위를 산출할 수 있음을 보여질 수 있다<sup>[8]</sup>. 그리고 Yeshurun과 Schwartz는 시각피질에서 모든 방향 성분을 포함하는 안우위기동(ocular dominance column)이 교대로 나타나는 생물학적인 구조를 모방하여 좌우 스테레오 영상을 배열시킨 후, 캡스트

럼 필터(cepstrum filter)를 적용시켜서 나타난 출력에서 최대값을 찾아서 양안 변위를 산출하는 방법을 제안하였지만<sup>[9]</sup>, 좌우 영상을 분할하여 사용하지는 않았다.

본 논문에서는 평행시각에서 얻어진 영상을 입력으로 사용하지 않고, 인간 시각 체계에 기반한 교차시각 스테레오 영상을 사용하였다. 평행 시각모델의 일반화된 교차 시각모델은 카메라의 초점 거리 또는 기준선 거리가 길어질수록 거리 오차는 감소하고, 깊이 분해도에서 더 좋은 성능을 나타낸다<sup>[10]</sup>. 이러한 교차 시각 스테레오 카메라 모델에 대해서 II장에서 간단히 고찰하고, III장에서는 스테레오 영상을 사분면 형태로 분할하기 위한 근거를 인간시각체계의 영상신호 전달 과정에서 파악할 것이며, 아울러 IV장에서는 영상 분할의 타당성을 기하학적으로 살펴본다. 재배열된 영상에 대해서 캡스트럼 필터를 사용하여 좌우 스테레오 영상에 존재하는 변위 벡터를 산출하는 알고리즘이 V장에서 언급될 것이고, 컴퓨터로 발생된 스테레오 입력 영상들과 자연영상에 대한 모의실험 결과와 해석이 VI장에서 다루어지고, 마지막으로 결론을 맺는다. 본 논문의 주된 기여는 시각피질에서 안우위기동의 생물학적 구조를 고려했을 뿐만 아니라 망막에서 대뇌피질까지 영상신호가 분리되어 전달되는 인간시각체계의 과정을 알고리즘으로 구현한 것이다. 이것은 대부분 기존 스테레오 영상 정합방법이 평행시각 모델을 사용하여 영상처리의 관점에서 접근한 것과 확연히 구별된다.

## II. 교차시각 스테레오 카메라 모델

사용된 카메라 모델은 인간 시각과 유사한 교차시각이며, 그림 1에 나타나 있다. 물체의 좌우, 상하 반전을 피하기 위해서 초점 앞에 영상 평면을 위치시킨다. 렌즈의 초점이 투영의 중심이 되며, 두 렌즈 중심사이의 거리를 기준선 거리  $b$ 라고 하고, 이 기준선 중앙에 기준 좌표계  $(X, Y, Z)$ 의 원점이 위치한다. 두 광축이 평행 모델처럼 평행이 아니고 왼쪽 카메라의 광축은 오른쪽으로 교차각(convergent angle)  $\theta_l$  만큼, 오른쪽 카메라는 왼쪽으로  $\theta_r$  만큼 기울어진 구조를 갖고 있어서 어떤 고정점(fixation point)에서 만난다. 두 영상평면의 좌표는 우측이  $+x$ , 좌측이  $+y$  방향이다.

스테레오 영상이란 동일한 물체에 대해 인간의 두 눈과 같이 조금 다른 위치에서 얻어진 한 쌍의

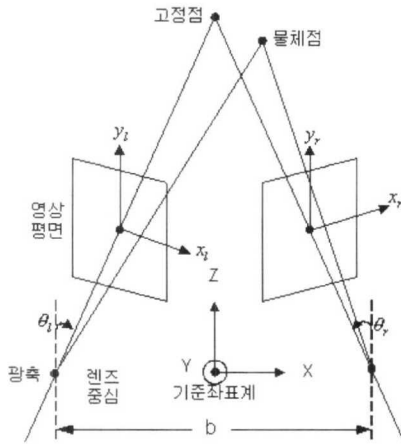


그림 1. 교차시각 스테레오 카메라 모델

영상을 말한다. 즉 3차원 공간상의 점들이 좌우 영상 평면에 각각 투영될 경우, 왼쪽 영상의 한 점이 두 카메라의 상대적인 위치와 회전등에 의해서 오른쪽 영상에는 왼쪽 영상의 위치와 다르게 나타난다. 이때 영상좌표의 상대적인 위치 차이를 양안 변위(binocular disparity)라고 한다. 공간상의 한 물체 점에 대응하는 좌우 영상 평면 좌표를 각각  $(x_l, y_l)$ ,  $(x_r, y_r)$ 라고 하면, 교차와 평행 카메라 모델이 확연히 다른 점은 평행시각에서는 수평 변위(disparity)인  $x_l - x_r$ 의 함수로만 깊이  $Z$ 가 결정되지만, 교차시각에서는 정확한 대응 영상 좌표를 알아야만 깊이를 산출할 수 있다는 것이다. 다시 말하면 좌우 영상평면에서 대응점 좌표,  $(x_l, y_l)$ 와  $(x_r, y_r)$ 를 좌우 영상에서 각각 찾고 카메라의 기하학적 변수들이 주어진다면 교차시각에서 기준좌표계에 대한 물체의 3차원 위치 정보를 계산할 수 있다 [10].

### III. 인간 시각 체계에 기초한 스테레오 영상 재배열

인간의 시각 신호 전달 과정이 그림 2에 나타나 있다[11]. 외부에 빛이 존재하면 자연 환경의 광학적인 영상은 인간의 시각 구조를 구성하는 각막, 수양액, 수정체를 통하여 좌우 망막(retina)에 투영된다. 투영된 망막 영상들은 망막의 중심와(fovea)을 기준으로 두 부분으로 구분된다. 즉, 코에 가까운 비측망막(nasal retina)과 관자놀이에 가까운 이측망막(temporal retina) 부분이다. 망막에 투영된 영상은

광수용체(photoreceptor), 양극세포(bipolar cell), 신경절세포(ganglion cell)를 거쳐 시신경(optic nerve)으로 전달되고, 시신경교차(optic chiasm)에서 신호가 나누어진다. 이측부분의 신호(그림 2에서 왼쪽 망막의 5,6,7과 오른쪽 망막의 2,3,4)는 교차없이 시삭(optic tract)으로 전달되고, 비측부분의 신호(왼쪽 망막의 1,2,3,4와 오른쪽 망막의 5,6,7,8)는 시신경교차에서 서로 교차가 일어나 시삭으로 전달된다. 다시 말하면, 좌측망막의 비측신호는 우측 시삭으로, 우측망막의 비측신호는 좌측 시삭으로 경로를 변경한다. 그러므로 좌측망막의 비측부분과 우측망막의 이측부분에 존재하는 신호들은 동일한 왼쪽시각계의 영상 신호(숫자 1,2,3,4)를 나타낸 것으로 우측시각 신호를 구성한다. 유사하게 좌측시각의 신호들은 우측망막의 비측부분과 좌측망막의 이측부분에서 나온 신호로 구성되어 있으므로 동일한 오른쪽시각계의 영상(숫자 5,6,7,8)을 표현한 것이다. 이렇게 동일한 시각계 영상을 포함하는 시삭신호는 외측슬상체(lateral geniculate body: LGB)를 경유하여 시방선(optic radiation)으로 전달된다.

LGB는 6개의 층으로 구성되며 왼쪽과 오른쪽 망막의 신호가 교대로 3층씩 차지한다. 즉, 시신경교차에서 교차된 신호(비측신호)와 교차되지 않은 신호(이측신호)는 교대로 다른 층에 할당된다. 외측슬상체는 해당 시각계 신호를 또 다시 위쪽과 아래쪽 두 부분의 시각계로 구분한다. 시방선에는 두 경

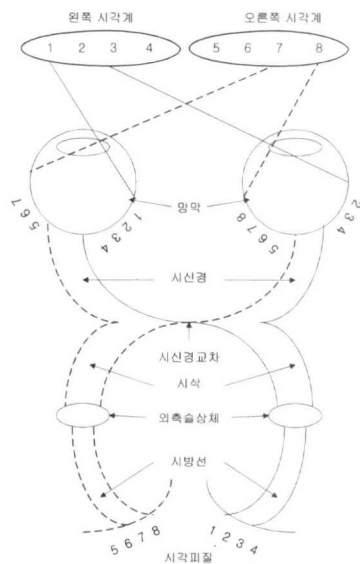


그림 2. 시각신호 전달 과정

로가 존재하며, 그 중 하나인 하측시방선은 Meyer 경로로 불리며 위쪽 시각영역에 해당하는 영상신호를 후두부에 존재하는 하측 뇌구(calcarine fissure)에 전달하고, 다른 경로인 상측시방선은 아래쪽 시각영역의 영상신호를 상측 뇌구에 전달한다. 그러므로 전체 시각영역은 시신경교차에서 좌우 시각계로, 시방선에서 상하 시각계로 구분되어, 결과적인 시각계는 사분면으로 나누어진 형태를 갖는다. 사분면으로 구분된 신호들은 대뇌의 좌우측 후두엽에 도달하여 일차시각피질(primary visual cortex)에서 특징정합 과정을 거쳐서 영상에 존재하는 여러 가지 정보를 추출 가능하게 한다<sup>[12]</sup>. 시각 경로 장애에 의해서 발생하는 시각 손상의 결과로부터 시각 영역이 사분면으로 분할됨을 확인할 수 있다<sup>[11]</sup>.

대뇌의 모든 피질은 기본적으로 6개 층으로 구성되며, 특히 4C 층은 LGB로부터 대부분 입력을 받으며 여기서 두 망막신호가 서로 이웃하게 배열되어 양안시각을 일으킨다. 더욱이 Hubel과 Wiesel은 후두엽의 피질에 극소 전극을 삽입하여 동일한 방향에 반응하는 수직으로 배열된 세포 기둥(column)을 발견하였다<sup>[13]</sup>. 18 내지 20개의 이웃한 기둥들이 모여서 초기둥(hypercolumn)을 형성하여 모든 방향성분의 특징을 인식한다. 즉 하나의 기둥은 약 10°의 분해능을 가지고 있으며 어떤 입력 영상에 대해서 약 18개의 방향성 필터에 의해서 해석된다고 볼 수 있다.

시각피질에는 동일한 물체점 영상에 해당하는 좌우 두 망막 신호가 하나의 안우위기둥에 전송되어 양안 변위가 추출되는 것으로 알려져 있다<sup>[14]</sup>. 좌우 망막의 모든 대응점 상의 물체는 당연히 융합(fusion)되어 인지될 수 있으며, 더욱이 인간의 양안 시각은 대응점 주변의 작은 영역(Panum의 융합영역)내에 존재하는 물체도 융합할 수 있으며 중심에서 멀어질수록 이 영역은 넓어진다<sup>[15]</sup>.

요약하면, 자연환경의 광학적 영상은 고정점을 중심으로 좌우 상하 반전되어 망막에 투영되고, 일차적으로 시신경교차에서 좌우 반으로 구분되고, 부가적으로 외측슬상체에서 상하로 나누어진 두 시방선을 따라서 전달되므로, 각각의 일차시각피질에 전달되는 영상신호는 자연 영상이 좌우상하 사분면으로 나누어진 신호 형태를 갖추고 있다. 사분면의 원점은 망막의 중심위에 그리고 동시에 자연영상의 고정점에 해당된다. 제안된 연구는 사분면 형태로 구분된 1차 시각 피질 내의 구조적 특징과 연관되어 있을 것으로 추정되어지는 양안 변위·측정 알고리

즘을 탐구하는 것이다.

#### IV. 인간시각 영역의 사분면 분할과 근거

앞에서 언급한 것처럼 전체 시각계 영역에 대한 영상 신호는 시신경교차와 외측슬상체에서 각각 좌우, 상하 영역으로 분할되므로, 사분면의 형태로 일차시각피질에 도달하여 여러 영상처리 과정을 겪게 된다. 교차시각에서 좌우측 망막영상에서 각 영상을 사분면으로 분할하면, 서로 동일한 사분면에서만 정합점을 찾으면 되기 때문에 탐색영역이 훨씬 감소하여 처리속도 향상과 오차 감소 효과를 얻을 수 있다. 이러한 사분면 영역 분할의 타당성을 기하학적으로 살펴보고자 한다.

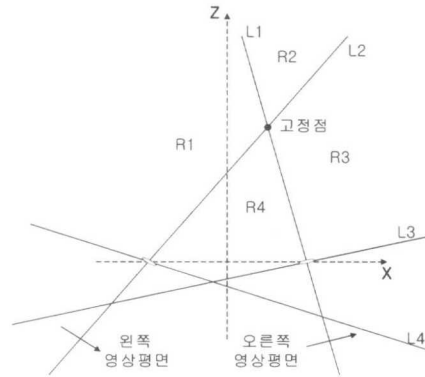


그림 3. 영상평면에 투영되는 시각영역 구분

물체가 3차원에 놓여있는 위치에 따라서 실제적으로 영상평면에 투영되는 사분면을 고려하면 그림 3과 같고, 표시된 직선들의 식들은 다음과 같다.

$$\begin{aligned}
 L1: Z &= \left(\frac{b}{2} - X\right) \cot \theta, \\
 L2: Z &= \left(X + \frac{b}{2}\right) \cot \theta, \\
 L3: Z &= \left(X - \frac{b}{2}\right) \tan \theta, \\
 L4: Z &= -\left(X + \frac{b}{2}\right) \tan \theta,
 \end{aligned} \tag{1}$$

L1과 L2는 각각 오른쪽과 왼쪽 렌즈의 광축을 나타내고, L3은 오른쪽 렌즈 중심을 통과하면서 오른쪽 영상평면과 평행인 직선, L4는 왼쪽 렌즈 중심을 통과하면서 왼쪽 영상평면과 평행인 직선을 나타낸다. 표시된 R1에서 R4는 위 4 직선에 의해서 구분되는 렌즈 앞에 위치한 영역들이다. 직선 L1과 L2는 각각 오른쪽 영상평면과 왼쪽 영상평면의 x

좌표의 부호를 결정한다. 즉  $L1$ 의 오른쪽에 물체가 놓이면  $x_r < 0$ 이고, 왼쪽에 놓이면  $x_r > 0$ 이 된다. 각 영역이 영상평면에 투영되는 위치는 다음과 같다.

$$\begin{aligned} R1: x_l > 0, x_r > 0, \quad R2: x_l > 0, x_r < 0 \\ R3: x_l < 0, x_r < 0, \quad R4: x_l < 0, x_r > 0 \end{aligned} \quad (2)$$

영상평면의  $y$ 좌표는 물체의 기준좌표계  $Y$ 에 의해서 결정되며,  $Y > 0$ 이면  $y_l < 0, y_r < 0$ 이고,  $Y < 0$ 이면  $y_l > 0, y_r > 0$ 이다. 그리고 영상평면의 사분면에 해당하는 시각영역은 다음과 같다.

- 1 사분면 ( $x > 0, y > 0$ ):  $Y < 0$ 인  $R1$  영역
- 2 사분면 ( $x < 0, y > 0$ ):  $Y < 0$ 인  $R3$  영역
- 3 사분면 ( $x < 0, y < 0$ ):  $Y > 0$ 인  $R3$  영역
- 4 사분면 ( $x > 0, y < 0$ ):  $Y > 0$ 인  $R1$  영역

$R2$ 와  $R4$  영역은 두 영상평면의  $x$ 좌표의 부호가 다르며 주어진 고정점에서 인간시각의 변위기울기(disparity gradient) 제한인 2를 초과하는 영역이고,  $L1$  또는  $L2$  직선 위에 존재하는 점들은 고정점에 대해서 수평 변위기울기가 2이다. 불투명한 물체의 표면에서 수평 변위기울기가 2를 초과하는 점들이 존재하면, 그 점들은 어느 한 쪽 영상에는 투영되지 않는다. 내부 영역  $R2$ 와  $R4$ 에서는 기울기가 2를 초과하고, 이 영역들은 horopter를 고려한 경우에 Panum 융합영역에 포함되지 않기 때문에 융합될 수 없으며, 동일 물체가 다른 공간 위치에 나타나 보이는 복시(diplopia)를 유발시키고, 또한 정합제약으로 많이 사용하는 순서조건이 성립하지 않은 곳이다<sup>[15]</sup>.

영상좌표축들에 대해서도 고려해보면, 우선 좌우측 영상 평면의 원점은 고정점에 대응되는 것으로 항상 일치한다. 고정점의 좌표는 아래와 같다.

$$X = \frac{b}{2} \frac{\sin(\theta_l - \theta_r)}{\sin(\theta_l + \theta_r)}, \quad Z = \frac{b}{\tan \theta_l + \tan \theta_r} \quad (3)$$

기준좌표계와 영상좌표계 사이의 관계식으로부터  $y_l = 0$ 이면  $Y = 0$ 이고 역시  $y_r = 0$ 가 된다<sup>[10]</sup>. 즉 어느 영상의 수평축에 투영되는 물체는 다른 영상의 수평축에 역시 투영된다. 그러나  $x_l = 0$  좌표를 갖는 물체부분이 일반적으로  $x_r = 0$ 에 투영이 되지 않는다. 다시 말하면 수직축들은 서로 대응관계가 성립되지 않는다. 이것을 그림으로 보면 좌측 영상에서는 좌측 렌즈와 고정점을 지나는 수직평면에 존재하는 물체가 영상 평면의 수직축에 투영되지만,

우측영상에서는 우측 렌즈와 고정점을 지나는 수직 평면에 존재하는 물체가 수직축에 투영된다. 그러므로 각 수직축에 투영되는 물체가 서로 다를 수 있다. 이러한 물체의 위치들은 변위 기울기의 제한에 놓여있다. 만약 고정점 부근을 폐색(occlusion)시키는 물체가 두 카메라 앞에 없고, 고정점에서 물체의 기울기가 수직방향으로 거의 영에 가까우면, 고정점을 지나고  $X-Z$  평면에 수직인 직선상의 물체 위치들이  $x_l = 0$  그리고  $x_r = 0$ 를 동시에 만족시켜서 서로 대응관계를 성립하게 한다. 이 수직선은 3차원에서  $L1$ 을 포함하고  $X-Z$  평면에 수직인 평면과  $L2$ 를 포함하고  $X-Z$  평면에 수직인 평면이 만나는 직선이다.

변위기울기 조건을 만족하는 물체의 경우에 고정점 근처의 국부적인 영역에서  $x_l = 0$ 이면 일차적인 근사치로  $x_r = 0$ 가 된다고 가정하여 대응점을 찾을 수 있다. 인간 시각은 교차각들을 정확히 산출할 수가 없기 때문에 정확한 3차원 위치보다는 상대적인 위치를 파악하며, 대부분 양안세포(binocular cell)들은 영변위(zero disparity)에 반응한다<sup>[15]</sup>. 그러므로 인간 시각에서 좌우 교차각이 다르면 불편함을 느끼며, 무의식적으로 고개를 돌려서 서로 일치하도록 하여 고정점을  $X=0$ 에 위치시킨다. 전반적인 물체의 특성을 파악하기 위해서는 고정점을 물체 주변으로 이동한다.

고정점의 지속적인 이동에 대한 근거로 인간시각 시스템은 한 곳에 2-3초 이상 초점을 두면 영상이 희미해진다. 그래서 무의식적으로 눈을 3/10 - 5/10 초 간격으로 계속 움직여서 망막에 투영되는 영상을 변하게 한다. 새와 곤충들은 눈을 움직일 수가 없기 때문에, 새는 지속적으로 머리를 움직이고 곤충은 심장과 호흡기관의 움직임에 기인하여 몸이 움직인다. 더욱이 물체가 움직이면 곤충은 사물을 보다 분명히 볼 수 있다<sup>[16]</sup>.

## V. 캡스트럼 필터를 이용한 변위 벡터 추출 알고리즘

Yeshurun과 Schwartz가 시각피질의 안우위기동에 좌우 영상이 교대로 나타나는 구조를 스테레오 영상 정합에 처음으로 적용하였다<sup>[9]</sup>. 좌우 영상에서 작은 영상조각을 추출하여 나란히 재배열한 후, 캡스트럼 필터를 사용하여 양안 변위를 추출하였다. 그들은 안우위기동에 투영된 영상  $f(x, y)$ 를 단순화

하기 위해서 왼쪽 영상  $s(x, y)$ 와 오른쪽 영상  $s(x - W, y)$ 를 나란히 결합시킨 것으로 가정하였다. 즉, 오른쪽 영상은 왼쪽 영상을 안우위기동 세포의 폭  $W$  만큼 이동되어진 영상이다. 그러므로 합성 영상  $f(x, y)$ 는 다음과 같이 표현할 수 있다.

$$f(x, y) = s(x, y) * \{ \delta(x, y) + \delta(x - W, y) \} \quad (4)$$

여기서  $\delta(x, y)$ 는 2차원 delta 함수이다. Fourier 변환 후, 로그를 취하면

$$\log F(u, v) = \log S(u, v) + \log(1 + e^{-j2\pi Wu}) \quad (5)$$

식(5)의 스펙트럼은  $(W, 0)$ 만큼 이동된 위치에서 최대값을 가질 것이다. 위 경우는 나란히 놓인 두 개의 동일한 영상들에 대한 간단한 경우이며, 모의 실험에서 이와 같은 경우에는 거의 완벽하게 깊이 정보를 추출함을 알 수 있었다.

좌우 영상에 수평 및 수직 변위가 존재하는 한 쌍의 영상 조각의 경우에는 다음과 같이 표현할 수 있다. 우선  $s_l(m, n)$ 을 왼쪽 기준 디지털 영상에서 추출한 크기가  $M \times N$ 인 부분 영상으로 가정한다.  $s_l(m, n)$ 을 아래쪽으로  $p$ 만큼, 오른쪽으로  $q$ 만큼 순환 이동(circular shift)한 영상을 오른쪽 디지털 영상  $s_r(m, n)$ 으로 정의하고, 왼쪽 부분 영상에 추가하여 다음과 같이 크기가  $M \times 2N$ 인 합성영상  $f(m, n)$ 을 형성한다.

$$f(m, n) = [s_l(m, n), s_r(m, n)] \quad (6)$$

합성영상 조각  $f(m, n)$ 에 대해 이산 주파수 변환(DFT)을 취하면,

$$F(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{2N-1} f(m, n) e^{-j2\pi(\frac{mk}{M} + \frac{nl}{2N})} = S_l(k, \frac{l}{2}) \left[ 1 + e^{-j2\pi(\frac{pk}{M} + \frac{(N+q)l}{2N})} \right] \quad (7)$$

여기서  $S_l(k, l)$ 는  $s_l(m, n)$ 의 이산 주파수 변환이다.  $\log F(k, l)$ 를 역 Fourier 변환하면 아래와 같은 캡스트럼을 얻을 수 있다.

$$\text{IDFT} [\log F(k, l)] = \frac{1}{2MN} \sum_{k=0}^{M-1} \sum_{l=0}^{2N-1} \left[ \log S_l(k, \frac{l}{2}) + \log(1 + e^{-j2\pi(\frac{pk}{M} + \frac{l(N+q)}{2N})}) \right] \cdot e^{j2\pi(\frac{km}{M} + \frac{ln}{2N})} \quad (8)$$

식 (8)의 둘째 항을 정리하면 다음과 같다.

$$\sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{2MNi} \sum_{k=0}^{M-1} \sum_{l=0}^{2N-1} e^{j2\pi(\frac{m-pi}{M}k + \frac{n-(N+q)i}{2N}l)} = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{i} \delta[m - pi, n - (N + q)i] \quad (9)$$

식 (9)은 공간위치  $m$ 이  $p$ 의 배수,  $n$ 이  $N + q$ 의 배수에서 피크값들을 갖는다. 인간시각에서는 두 눈이 수평적으로 위치하므로 일반적으로 수직방향의 변위는 작고 수평방향의 변위는 크다. 실제 영상에서는 데이터 값이 순환적으로 이동하지 않기 때문에 큰  $q$ 에 대해서는  $N$ 을 크게 설정해야 하지만, 큰 영상조각의 크기는 공간적인 정확성을 저하시키는 원인이 된다. 영상에 존재하는 수평방향 변위  $q$ 의  $+$ 와  $-$  방향의 변화량이 동일하다고 가정하면  $-N/3 < q < N/3$ 인 경우에 유일하게 변위를 결정할 수 있으며 탐색 범위는  $N$ 을 기준으로  $\pm N/3$ 이 된다. 이것은  $N$ 과  $2N$  사이에서 양의 변위 ( $q_+$ )와 음의 변위 ( $q_-$ )에 의해서 최대값이 동일한 위치에서 일어나는 경우를 고려한 식  $N + q_+ = 2(N + q_-)$ 와  $q_- = -q_+$ 에 의해서 유도된다. 최대값이  $N$ 보다 작은 경우에서 발견되면 음의 수평 변위,  $N$ 보다 큰 경우에서 발견되면 양의 수평변위가 됨을 알 수 있다. 이론적으로 수직과 수평방향 변위를 갖는 변위 벡터  $(pi, (N + q)i)$ 는 모든 자연수  $i$ 에 대해서 피크값을 갖고,  $i=1$ 일 때 최대값을 갖는다.  $-N/3 < q < N/3$ 인 조건이 없는 경우에는  $N$ 과  $2N$  사이에서 최대값이 발견되면 변위 벡터  $(pi, (N + q)i)$ 에서  $i$ 값을 알 수가 없기 때문에 수평, 수직 변위를 유일하게 결정할 수가 없다. 수직 및 수평 방향의 변위 탐색 범위는 인간의 양안 시각에서 Panum의 융합영역에 대응되는 것으로 고려할 수 있기 때문에, 고정점 근처에서는 작은 범위, 먼 곳에서는 큰 범위가 필요하다.

식 (9)은 합성영상에서 좌우 영상이 서로 순환 이동된 영상인 경우에는 적합하지만, 실제 영상은 이 조건을 만족하지 않는다. 그러므로, 보다 실제적이고 잡음에 강한 알고리즘을 구현하기 위해서 epipolar 직선에 적용한 Smith와 Nandhakumar의 방법<sup>[17]</sup>을 수평 및 수직 변위가 존재하는 영상에서도 적용하기 위해서 아래와 같이 합성 영상조각을 수정하여 본 논문에서 사용할 것이다.

$$f_{p,q}(m, n) = [s_l(m, n), s_r(m + p, n + q)] \quad (10)$$

여기서  $s_x(m, n)$ 은 왼쪽 영상에서 변위를 계산하고자 하는 영상점을 중심으로 크기가  $M \times N$ 인 부분 영상이고,  $s_x(m+p, n+q)$ 는 왼쪽 영상 중심점을 수직 방향으로  $p$ 만큼, 수평 방향으로  $q$ 만큼 이동된 중심점을 기준으로 크기가  $M \times N$ 인 오른쪽 부분 영상이며,  $f_{p,q}(m, n)$ 은 이 두 부분 영상을 수평으로 나란히 합성한 영상 조각이며 크기가  $M \times 2N$ 이다.  $p$ 와  $q$ 의 범위는  $-d_{m1} \leq p \leq d_{m2}$ ,  $-d_{n1} \leq q \leq d_{n2}$ 이며, 각각 상하, 좌우 방향으로 영상 내 존재하는 최대 변위 범위를 나타낸다.  $(p, q)$ 의 모든 조합 쌍에 대해서  $f_{p,q}(m, n)$ 의 캡스트럼을 구한 후, 캡스트럼이 최대가 되는  $p, q$ 를 찾으면, 이 값들이 각각 두 영상 사이의 수직 및 수평 변위가 된다.

$$(p, q) = \max_{p,q} \{ \text{cepstrum of } f_{p,q}(0, N) \} \quad (11)$$

여기서  $(0, N)$ 을 사용하는 이유는 우측에 첨가되는 영상이 좌측 영상과 동일할 경우, 수직 방향으로 0, 수평방향으로  $N$ 만큼 이동하면 최대 캡스트럼을 얻기 때문이다. 본 논문에서는 직사각형의 영상 조각을 언급하였지만, 이러한 조각들은 임의의 형태가 될 수 있다.

## VI. 실험 결과

앞에서 언급된 인간시각의 변위 추출 알고리즘을 검증하기 위해서 무작위 점 스테레오 영상(random dot stereogram), 컴퓨터로 합성한 원통 영상, 자연 영상인 Pentagon 영상을 사용하였다.

### 1. 수정 캡스트럼의 유용성

우선 캡스트럼 자체만 사용한 식 (9)와 캡스트럼에 수직, 수평 변위 범위를 포함하는 수정된 식

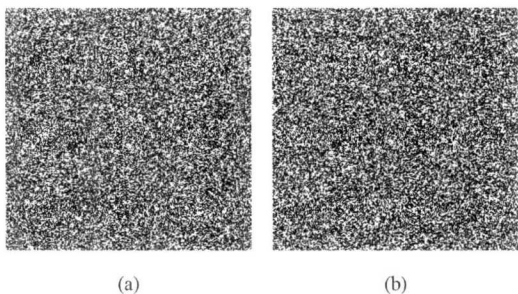


그림 4. 무작위 점 스테레오 영상  
(a) 왼쪽 영상, (b) 오른쪽 영상

(11)의 성능을 비교 평가하기 위하여 무작위 점 영상을 사용하였다. 그림 4는 크기가  $256 \times 256$ 이며, 두 층으로 구성된 무작위 점 스테레오 영상이다. 내부 층은 크기가  $52 \times 52$ 이며, 왼쪽 영상을 기준으로 오른쪽 영상의 수평 변위가 왼쪽으로 +5 픽셀, 수직 변위가 위쪽으로 +4 픽셀을 갖도록 하였고, 외부 층은 크기가  $152 \times 152$ 이며 수평 변위가 왼쪽으로 +2 픽셀, 수평 변위가 위쪽으로 +2 픽셀을 갖도록 발생하였다.

사용된 합성영상 조각의 크기는 캡스트럼만을 고려한 그림 5의 (a)와 (b)에서는  $15 \times 22$ , 수직 및 수평 변위 범위를 고려한 그림 5의 (c)와 (d)는  $7 \times 22$ 이다. 왼쪽과 오른쪽 영상에서 추출한 각각의 부분 영상 조각들을 단순히 이웃하게 위치시켜서 한 쌍의 영상으로 합성하여 변위를 추출한 것으로, 앞에서 설명한 교차 시각 스테레오 영상에 대한 정합 방법으로 변위를 구한 것은 아니다. 그림 5의 (a)와 (b)는 캡스트럼만을 이용하여 측정된 수평 변위, 수직 변위를 각각 나타낸다. 정확히 측정되면 회색으로 표시된 (a)의 내부층 값은 5, 검게 표시된 외부층 값은 3이고, (b)의 내부층 값은 4, 외부층 값은 2이다. 각 층의 내부에서도 잡음이 존재함을 볼 수 있다. 수정 캡스트럼 경우인 그림 5의 (c)와 (d)에서는 두 층의 경계선을 제외하면 정확한 변위들을 계산하였음을 알 수 있다. 물론 (a)와 (b) 경우, 보다

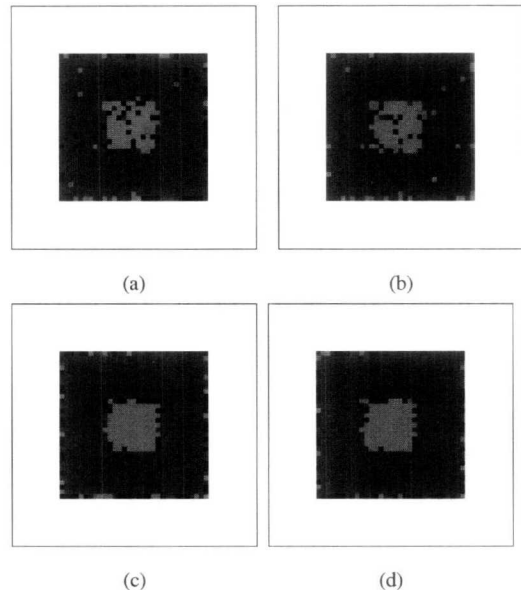


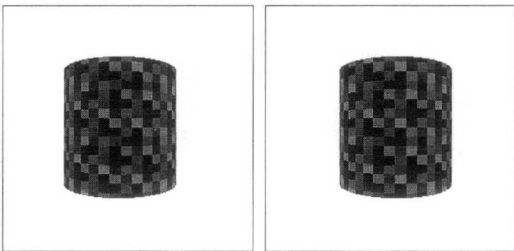
그림 5. 무작위 점 영상의 변위 (a) 캡스트럼의 수평 변위, (b) 캡스트럼의 수직 변위, (c) 수정 캡스트럼의 수평 변위, (d) 수정 캡스트럼의 수직 변위

큰 영상 조각을 사용하면 더 좋은 결과를 얻을 수 있으나 수직 방향으로 이미 2배 정도의 영상 조각을 사용하였지만 (c)와 (d)보다 성능이 저하됨을 볼 수 있다. 작은 정사각형의 모양들은 계산 시간을 감소시키기 위해서 어느 중심 픽셀에서 계산된 변위를 그 픽셀을 중심으로 5x5인 크기로 주변을 동일 변위값으로 대체한 것이다.

수정 캡스트럼에 나타나는 경계선 부근에 대한 변위 오차는 무작위 점 영상 자체의 특성에 기인한 것이다. 즉, 무작위 점 영상을 발생할 때, 층 내부는 기준 영상에서 복사하여 사용하고, 층 외부는 다시 무작위 수를 발생하여 채우기 때문에 두 층의 경계선 부근은 왼쪽과 오른쪽 영상이 상당히 다르다. 물체는 동일하고 물체와 배경의 경계선 부근이 좌우 영상에서 다를 때, 경계선 부근에서 희미하게 물체가 인식되는 현상은 인간시각에서도 느낄 수 있다. 물론 경계선 부근의 변위 오차를 감소시키기 위해서 일반적인 영상처리에서 사용되는 보간법을 적용할 수 있지만, 알려진 인간시각체계의 정보 전달 과정을 충실히 실험하기 위해서 어떠한 보편적인 영상처리 필터도 사용하지 않았다. 결과적으로 캡스트럼만을 사용한 알고리즘은 잡음에 민감하고, 수정된 캡스트럼은 좌우 영상이 크게 다른 경계선 부분만 제외하면 정확히 변위를 추출함을 알 수 있다. 다른 영상에 대해서도 수정된 캡스트럼의 성능이 우수함을 볼 수 있었다.

2. 원통 합성영상

여러 개의 작은 직사각형 영역으로 구성된 원통이 3차원 공간에 존재한다고 가정하고, 왼쪽 망막과 오른쪽 망막에 맺히는 영상을 획득하였다. 이것은 교차 시각 스테레오 카메라 모델에서 기준 좌표계와 영상 좌표계의 관계를 나타내는 식들을 이용하여 컴퓨터로 발생시킬 수 있으며, 그 결과가 그림 6에 크기 256x256인 좌우 스테레오 원통 영상으로

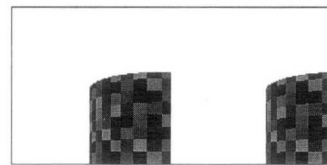


(a) (b)

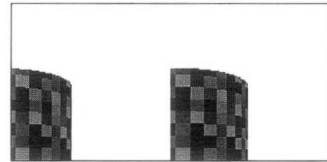
그림 6. 원통 스테레오 영상  
(a) 왼쪽 영상, (b) 오른쪽 영상

나타내었다. 사용된 카메라 변수는 두 카메라 사이의 기준선거리 6.5 cm, 초점거리 2.5 cm, 카메라의 시야각도 60°, 교차각 6.183°이다. 원통의 반지름은 10cm이고, 그 중심 좌표는 (X, Y, Z) = (0, 0, 40cm)이다.

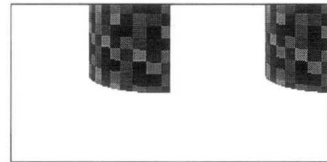
3차원 공간에 위치한 원통 물체에 교차 시각 스테레오 카메라 모델을 적용하면, 고정점은 각 영상 평면의 중심 픽셀에 투영되어진다. 이 영상 평면의 중심을 기준으로 좌측 상부(2사분면)와 하부(3사분면) 그리고 우측 상부(1사분면)와 하부(4사분면)로 4개 영역으로 분할된 영상을 각 사분면 별로 그림 7에 보여준다. 각 사분면에서 왼쪽 부분은 왼쪽 영상에서, 오른쪽 부분은 오른쪽 영상에서 추출한 것이다. 각 분할된 영상에 대해서 수정 캡스트럼 알고리즘으로 수직, 수평 변위를 찾으면 동일 물체점에 대한 해당 대응점 (x<sub>l</sub>, y<sub>l</sub>)와 (x<sub>r</sub>, y<sub>r</sub>)을 얻을 수 있다. 수직방향 픽셀 변위는 -3에서 3까지, 수평방향 변위는 -9에서 9까지 변화시켰고, 영상조각의 크기



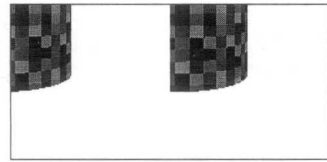
(a)



(b)



(c)



(d)

그림 7. 분할된 원통 영상 (a) 2사분면, (b) 1사분면, (c) 3사분면, (d) 4사분면



는 7x11을 사용하였다. 대응 영상 좌표와 카메라 모델식으로부터 3차원 좌표를 계산하였고, 깊이 정보인 Z좌표가 사분면별로 그림 8에 나타나 있다. 그림에서 밝게 나타난 부분이 가까운 거리, 어두운 부분이 먼 거리를 의미한다.

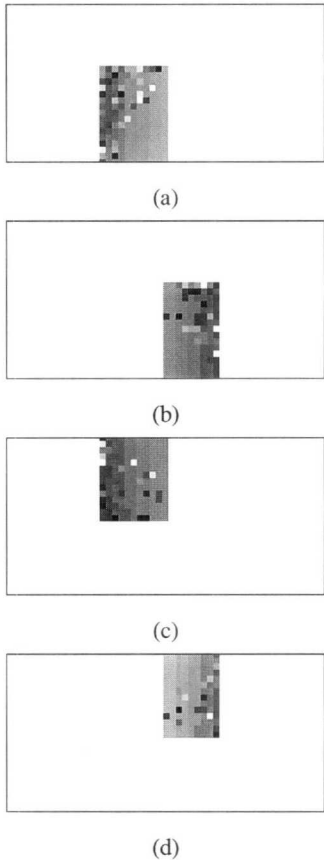


그림 8. 분할된 원통의 깊이 (a) 2사분면, (b) 1사분면, (c) 3사분면, (d) 4사분면

각 영역별로 추출된 깊이를 통합하여 나타낸 결과가 그림 9에 보여준다. 고정점 부근에서는 정확히 깊이를 추출하였고, 고정점에서 멀어질수록 원통의 곡률 때문에 좌우 영상 차이가 심하여, 깊이 오차가 발생함을 알 수 있다. 원통 물체의 깊이는 거의 연속적으로 변화되지만 디지털 영상의 양자화에 기인하여 깊이가 불연속적으로 표시된다. 이것은 변위 자체가 디지털 픽셀 차이로 계산되어 정수값만을 허용하기 때문에 발생하는 깊이분해능의 한계에 기인한 것이다. 보간법을 이용하여 subpixel 단위로 변위를 계산하면 더 정확한 깊이를 산출할 수 있을 것이다. 대응점으로부터 계산된 원통의 X 좌표와 Y

좌표가 그림 10의 (a), (b)에 각각 나타내었다. X좌표는 우측이 +방향이고, Y 좌표는 위쪽이 +방향이며, 두 좌표 모두 +방향을 어둡게 표시하였다. 고정점 부근에서 영상은 어느 시각 영역에서도 모두 나타나기 때문에 고정점 부근에서는 다른 이웃한 사분면의 값을 동시에 사용하였다. 고정점에 대한 주변 물체점들의 변위 기울기가 2보다 작으면 위와 같은 영역 사분할이 타당함을 인간 시각에서도 볼 수 있다<sup>[15]</sup>.

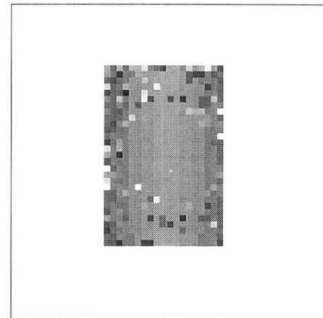


그림 9. 원통의 전체 깊이(Z 좌표)

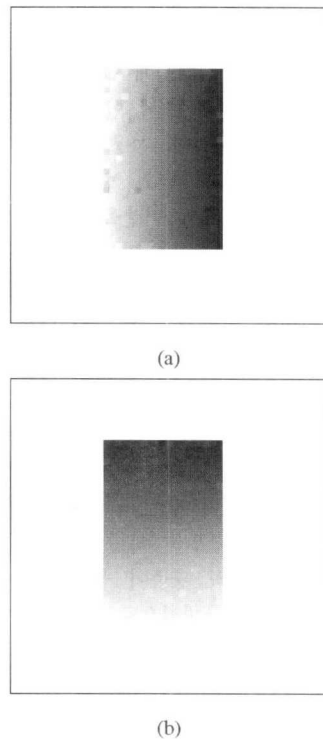


그림 10. (a) 원통의 X 좌표, (b) 원통의 Y 좌표

표 1은 각 사분면의 한 영상좌표에 대하여, 실제

기준좌표와 제한한 알고리즘으로 계산한 기준좌표의 정량적인 3차원 정보를 보여준다. 영상 픽셀 좌표  $(i, j)$ 는 원점이 영상의 좌측 위쪽에 위치하며  $i$ 는 행,  $j$ 는 열을 나타낸다. 기준좌표계 원점에서 가장 가까운 원통 표면의 좌표는  $(0, 0, 30\text{cm})$ 이다. 실제 좌표와 계산한 좌표 사이의 차이는 아주 작으며, 변위 계산시 디지털 픽셀에 기인한 양자화오차 범위 내에 있음을 알 수 있다.

표 1. 실제 좌표와 제한한 알고리즘으로 계산한 좌표

왼쪽 영상 $(i, j)$ :pixel	실제 좌표 $(X, Y, Z)$ : cm	계산한 좌표 $(X, Y, Z)$ : cm
(120,140)	(1.82,0.96,30.17)	(1.88,0.98,30.53)
(109,109)	(-2.43,2.45,30.30)	(-2.43,2.47,30.45)
(140,109)	(-2.43,-1.77,30.30)	(-2.43,-1.78,30.45)
(140,140)	(1.82,-1.79,30.17)	(1.88,-1.81,30.53)

### 3. 교차시각 자연 영상

널리 알려진 그림 11의 교차시각 Pentagon 스테레오 영상을 사용하였다. 이 영상의 특징은 원거리에서 촬영된 것으로, 영상에 존재하는 물체 깊이의 변화가 측정거리에 비해서 상당히 작다. 카메라 모델의 변수들을 모르기 때문에 정확한 3차원 정보를 계산할 수 없고, 수평과 수직 변위만을 그림 12 (a)와 (b)에 각각 나타내었다. 사용된 합성 영상조각의 크기는  $7 \times 22$ 이며, 최대 변위 범위는 수직방향으로  $\pm 3$ , 수평방향으로  $\pm 7$  픽셀을 사용하였다. 중심 픽셀에서 계산된 변위를 그 픽셀을 중심으로  $3 \times 3$ 인 크기로 주변을 동일 변위값으로 사용하였다. 깊이 (Z)는 수평변위에 의해서 주로 결정되므로, 수평변위 영상으로부터 대체적인 물체 윤곽과 깊이 차이를 알 수 있다. 밝은 부분은 정면에서 보았을 때 관측자로부터 가까운 거리를 나타내고, 어두운 부분은 더 먼 거리를 의미한다. 특히 우측 상단부  $135^\circ$  방향의 도로는 주변 보다 낮게 위치하므로 검게 표시되어 있다. 오각형의 우측 하단 부분이 조금 희게 나타난 것은 카메라의 광축이 물체와 수직이 아니고 조금 기울어져 상대적으로 좌측 상단부분보다 가까운 거리임을 알 수 있다.

물론 인간시각작용과 다르게 영상처리가 목적이면 입력 영상 또는 변위 영상에 어떤 필터를 적용하여 보다 시각적으로 우수한 출력을 얻을 수 있을 것이다. 그러나 본 논문의 목적은 영상처리 그 자체가 아니라 인간시각작용에 기반한 영상정합이므로

일반적으로 영상처리에 사용되는 어떠한 필터 작용도 포함하지 않았다.

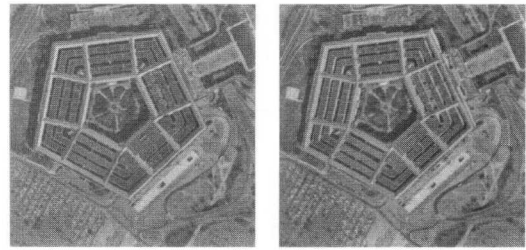
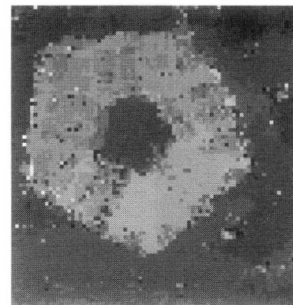
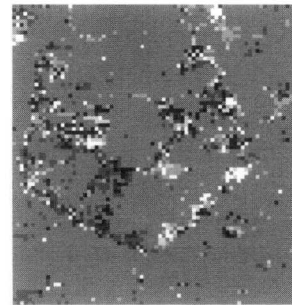


그림 11. Pentagon 스테레오 영상 (a) 왼쪽 영상, (b) 오른쪽 영상



(a)



(b)

그림 12. Pentagon의 변위 (a) X 변위, (b) Y 변위

입체 영상정합에 대한 여러 방법들이 제안되었지만, 어떠한 방법도 제한한 알고리즘처럼 입체 영상 데이터의 형식을 인간 시각 피질의 안우위기동 형태의 모방과 고정점을 기준으로 하여 영상들을 사분면으로 분할하여 정합을 시도하지는 않았다. 논문에서 수행한 알고리즘은 이러한 인간 시각 모델에 근거하여 주어진 스테레오 영상에서 거리 정보를 얻기 위한 방법으로 사용되었다. 그러나 눈을 통하

여 뇌로 전달된 영상을 인지하는 것과 관련된 스테레오 정합 반응은 일차시각피질 뿐만 아니라 이차피질에서도 발생할 수 있으므로 보다 정확한 방법은 이러한 것을 고려해야 할 것이다<sup>[12]</sup>.

주어진 고정점에서 획득한 한 쌍의 스테레오 영상에 대해서 지금까지 영상 전역에 정합을 시도하여 영상 중심에서는 깊이 정확도를 유지할 수 있었으나, 중심에서 멀어질수록 부정확해지는 경향을 볼 수 있었다. 실제적인 인간 시각과 다른 점을 지적하면, 첫째로 인간 시각은 픽셀 자체의 크기가 일정하지 않고, 중심에서는 크기가 아주 작아서 매우 정확히 위치 변화를 측정할 수 있으나, 중심에서 멀어질수록 픽셀의 크기가 지수적으로 커져서 변위 측정에 오차를 유발한다. 둘째로 인간은 넓은 영역의 깊이를 파악할 때, 양안의 교차각을 변화시킴으로써 고정점을 여러 영역 주변에 위치하여 여러 개의 스테레오 영상을 획득한 후, 각 스테레오 영상마다 정합 과정을 거쳐서 관심 있는 영상 중심의 물체에 대한 깊이를 정확히 산출하고 최종적으로 여러 고정점에 대한 깊이 정보를 통합하여 완전한 3차원 물체의 정보를 획득한다. 인간시각모델을 이용한 영상정합은 위 사실들을 고려하면 보다 더 정확한 결과를 얻을 수 있을 것이다.

## Ⅶ. 결 론

눈을 통하여 망막에 투영되는 영상 신호는 시신경을 통해서 뇌로 전달되어지고, 이 정보는 뇌의 일차시각피질에서 좌안과 우안의 신호가 교대로 띠모양의 형태로 구성되어지는데, 이러한 형태의 안우 위기동 세포는 스테레오 인지에 대한 민감한 반응을 나타낸다. 교차 시각에서 두 눈을 통해 전달되는 동일한 물체 정보가 영상평면에 투영될 때, 수직과 수평 변위 모두가 존재하기 때문에 평행시각 모델처럼 수평 변위만을 고려하여 거리를 산출할 수가 없다. 본 논문은 교차시각 스테레오 영상에 캡스트럼 필터를 적용하기 위해서 주어진 영상을 중심와를 기준으로 4개의 영역으로 분할하였다. 각각 분할된 영역에 대해서 대응하는 좌우 영상 조각들을 추출하여 한 쌍의 영상 조각으로 합성하였고, 이와 같은 영역들에 수정된 캡스트럼 필터를 이용한 알고리즘을 적용하여 스테레오 영상에 대한 3차원 정보를 추출하였다. 여러 종류의 스테레오 영상에 대한 실험결과로부터 인간시각체계와 유사한 제안 방법은 고정점 부근에서 정확한 정합을 보여주었다.

논문에서는 하나의 고정점에 대해서 얻어진 한 쌍의 스테레오 영상만을 사용하여 일련의 알고리즘을 적용하였지만, 광범위한 영역에 대해서 정확한 깊이 정보를 산출하기 위해서는 고정점의 위치를 다양하게 이동시켜서, 각 고정점 부근의 국부적인 영역에 대한 여러 스테레오 영상들을 얻고, 제안한 방법으로 각 스테레오 영상을 처리하여 부분적인 깊이 정보를 추출한 후, 여러 고정점에 대한 깊이 정보를 하나로 통합하는 일이 필요하다.

## 참 고 문 헌

- [1] U. R. Dhond and J. K. Aggarwal, "Structure from stereo-A review," *IEEE Trans. Systems, Man Cybernetics* 19(6), pp. 1489-1510, 1989.
- [2] G. A. Jones, "Constraint, Optimization and Hierarchy: Reviewing Stereoscopic Correspondence of Complex Features," *Computer Vision and Image Understanding*, Vol. 65, No. 1, pp. 57-78, 1997.
- [3] C. Capurro, F. Panerai and G. Sandini, "Dynamic vergence using log-polar images," *International Journal of Computer Vision* 24(1), pp. 79-94, 1997.
- [4] 이용범, 최영수, 윤종진, 진성일, "양안식 능동시각장치의 자세 및 주시각 제어 연구," 제11회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 387-392, 1999.
- [5] 방태식, 이적식, "캡스트럼 필터를 이용한 교차시각 스테레오 영상의 3차원 정보추출," 제11회 영상처리 및 이해에 관한 워크샵 발표 논문집, pp. 299-304, 1999.
- [6] D. Papadimitriou and T. Dennis, "Epipolar line estimation and rectification for stereo images," *IEEE Trans. Image Processing*, 3(4): 672-676, April 1996.
- [7] G. C. DeAngelis, I. Ohzawa and R. D. Freeman, "Depth is encoded in the visual cortex by a specialized receptive field structure," *Nature* 352, pp. 156-159, 1991.
- [8] D. J. Fleet, H. Wagner and D. J. Heeger, "Neural Encoding of Binocular Disparity: Energy Models, Position Shifts and Phase Shifts," *Vision Research* 36, pp. 1839-1858, 1996.

- [9] Y. Yeshurun and E. L. Schwartz, "Cepstral filtering on a columnar image architecture: A fast algorithm for binocular stereo segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 759-767, July 1989.
- [10] 이적식, "교차 시각 스테레오 카메라 모델 해석," *대한전자공학회 논문지*, 제33권 10호, pp. 50-62, 1996.
- [11] B. Pansky, D. Allen, G. Budd, Review of Neuroscience, 2nd Edition, McGraw-Hill, 1988.
- [12] E. L. Schwartz, "Spatial mapping in the primate sensory projection: analytic structure and relevance to perception," *Biological Cybernetics*, Vol. 25, pp. 181-194, 1977.
- [13] M. W. Matlin and H. J. Foley, Sensation and Perception, Allyn and Bacon, 1992.
- [14] R. F. Thompson, The Brain: A Neuroscience Primer, W. H. Freeman and Company, New York, 1993.
- [15] I. P. Howard and B. J. Rogers, Binocular Vision and stereopsis, Oxford University Press, New York, 1995.
- [16] S. Sinclair, How Animals See: Other Visions of Our World, Croom Helm, England, 1985.
- [17] P. Smith and N. Nandhakumar, "An Improved Power Cepstrum Based Stereo Correspondence Method for Textured Scenes," *IEEE Trans. PAMI*, Vol. 18, No. 3, March 1996.

이 적 식(Juck Sik Lee)

정회원



1979년 2월 : 경북대학교

전자공학과 학사

1987년 1월 : Texas A&M U.

Elec. Eng. M.S.

1991년 5월 : Texas A&M U.

Elec. Eng. Ph. D.

1979년~1984년 :국방과학연구소

연구원

1992년 3월~현재 : 경기대학교 전자공학전공 부교수  
<주관심 분야> 영상처리, 컴퓨터시각, 생체신호해석