

論文2000-37CI-4-5

## 움직이는 관찰자용 3차원 디스플레이 방법 (3D Display Method for Moving Viewers)

許慶茂\*, 金明新\*

(Kyung-Moo Huh and Myeong-Sin Kim)

### 요 약

본 논문에서는 관찰자의 위치가 변하더라도 이를 실시간 추적하여 관찰자의 시점에 정확히 대응하는 입체 영상을 구현할 수 있는 방법을 제안하였다. 즉 고성능의 하드웨어 장비가 아닌, 일반적으로 사용되는 개인용 컴퓨터에서도 인간의 두 눈을 찾아내고 관찰자의 움직임을 추적할 수 있도록 하는 알고리즘을 제안하였으며, 또한 관찰자의 위치에 따라 달라지는 물체의 모습을 표현해 주기 위해 유한한 다수의 입력 영상 정보를 이용하여 입체 영상을 제작하고 관찰자의 위치 이동에 정확히 대응하는 영상을 디스플레이하는 방법을 제안하였다. 본 논문에서 제안한 방법을 통해 평균 0.39초의 짧은 시간내에 약 97.5%까지 정확히 두 눈의 위치를 찾을 수 있었으며, F16 모델을 사용하여 여러 관찰자의 시점에 대응하는 3차원 디스플레이 실험 결과를 보임으로써 본 방법의 우수함을 보였다. 그리고 실제 로봇을 이용하여 좌우 카메라로 얻은 좌우 영상과 인벤터를 통해 버퍼에서 렌더링되는 스테레오 영상과의 유사도를 측정하여, 관찰자의 시점에 대응하여 렌더링되는 3차원 영상이 최적의 시점 대응 영상임을 확인하였다.

### Abstract

In this paper we suggest a method of detecting the two eyes position of moving viewer by using images obtained through a color CCD camera, and also a method of rendering view-dependent 3D image which consists of depth estimation, image-based 3D object modeling and stereoscopic display process. Through the experiment of applying the suggested methods, we could find the accurate two-eyes position with the success rate of 97.5% within the processing time of 0.39 second using personal computer, and display the view-dependent 3D image using F16 flight model. And through the similarity measurement of stereo image rendered at z-buffer by Open Inventor and captured by stereo camera using robot, we could find that view-dependent 3D picture obtained by our proposed method is optimal to viewer.

### I. 서 론

앞으로의 멀티미디어 기술의 주요한 분야로 예상되는 3차원 영상 구현에 있어서, 3차원 영상 기술의 실용화를 위한 핵심적인 기반 기술 구축과 구현 방법에

관한 연구가 현재 활발하게 이루어지고 있다. 이렇게 디스플레이에 있어서 입체적인 3차원 영상을 요구하는 시대가 가까이 다가 올 것으로 예상하는 가운데, 이러한 3차원 디스플레이를 위한 여러 방법들이 현재 연구 및 검토되고 있지만 아직은 3차원 TV 등으로의 실용화가 이루어지지 않은 실정이다. 3차원 디스플레이 방법은 크게 나누어, 양쪽 눈에 서로 다른 화상을 보임으로써 입체 영상을 구현하는 방법과 공간상에 3차원 영상을 구현해 내는 방법의 2가지로 분류할 수 있는데, 그 중 가장 많이 쓰이는 방법이 양쪽 눈에 서로 다른 화상을 보이게 하되 시역(viewing-zone)에 위치한 인간의 눈에서는 완벽한 입체 영상이 되도록 하

\* 正會員, 檀國大學校 電子工學專攻

(Dept. of Electronic Eng Dankook University)

※ 본 연구는 한국과학재단 핵심전문연구(과제번호 981-0910-042-2) 지원으로 수행되었음.

接受日字:2000年 2月 4日, 수정완료일:2000年 6月9日

는 전자의 방법으로서 이의 대표적인 방법이 오토스테레오스코픽(autostereoscopic) 방법이다. 이러한 오토스테레오스코픽 방법은 고성능의 하드웨어나 어떠한 추가적인 장치 없이도 입체 영상을 즐길 수 있다는 장점을 가지고 있으나, 입체 영상을 즐길 수 있는 가시영역이 매우 협소하고, 좌우가 바뀌는 영역이 존재하는 것 등이 가장 큰 문제점으로 알려져 있다. 즉 현재로서 오토스테레오스코픽의 가장 큰 문제점은 일정하게 정해진 위치에서만 입체 영상을 감상할 수 있다는 것으로서, 움직이고 있는 사람은 입체 영상을 볼 수 없다는 점이다.

따라서 움직이고 있는 관찰자가 계속 입체 영상을 볼 수 있게 하기 위해서는 관찰자의 움직임을 추적하여 움직임에 따른 입체 영상의 보상을 실시간적으로 수행시켜야 한다. 또한 이러한 방법은 홀로그램이나 가상현실 등의 응용에서 관찰자의 위치 이동에 따른 표시영상의 변화도 구현할 수 있어 그 활용도가 매우 높다.

일반적으로 인간의 얼굴은 매우 복잡한 패턴을 갖고 있기 때문에 얼굴의 특징을 추출하고 규정하기는 상당히 어렵다. 그러나 눈은 얼굴 내에서 가장 뚜렷하고 단순한 특징을 가지고 있으므로 관찰자의 움직임을 추적하는데 가장 중요한 정보가 된다. 더구나 오토스테레오스코픽 방식의 3차원 영상의 감상에 있어서는 관찰자의 두 눈의 위치가 직접적으로 관계되고 절대적인 역할을 하고 있으므로, 얼굴 자체를 검출하는 것보다는 두 눈의 위치를 직접 검출하는 것이 중요하고 필요하다. 3차원 영상을 위해 눈의 위치를 검출하기 위한 방법으로서 Imai 등은 [1]에서 관찰자의 눈 바로 위에 자기 위치 센서를 부착하여 위치를 검출하고 그에 따라 영상 이동 장치를 구동시키는 방법을 시도하였다. 그러나 이 방법은 그 시스템 구성이 복잡하고 한 쪽 눈의 위치만 검출함으로써 인해 정확도가 떨어지는 단점이 있었다. 그리고 Yoon 등은 [2]에서 흑백 카메라를 사용하여 워크스테이션으로 구현하였으며, 눈 검출 성공률은 91.7%를 기록하였지만, 이 방법에서는 안경 착용시와 30° 이상 얼굴이 기울어졌을 때에는 검출이 곤란한 것으로 나타났다. 또한 Yokoyama 등은 [3]에서 컬러 영상으로부터 얼굴과 특징점을 검출하고, 얼굴내 특징점들(눈, 눈썹, 입, 코)의 궤적을 구하는데까지 수행하였는데, Indigo2 워크스테이션을 사용하였으며 그 처리시간은 한 프레임당 20초가 걸렸다. 그리고 Wu 등은 [4]에서 퍼지 패턴 매칭을 사용하여 얼굴과

유사한 영역을 검출하고, 얼굴 후보 영역을 검사한 후 얼굴의 특징점을 추출하였는 바, 좌우 두 눈의 중심 좌표와 입의 중심 좌표를 구하였다. 여기에서도 Indigo2 워크스테이션을 사용하여 구현하였는데, 이 실험 결과에서는 80장의 정지 영상을 실험하여 약 90%의 성능을 보였으며, 얼굴을 검출하는데 까지만한 프레임 당 2초가 걸렸다.

본 논문에서는 고성능의 하드웨어 장비가 아닌, 일반적으로 사용되는 개인용 컴퓨터에서도 인간의 두 눈을 찾아내고 관찰자의 움직임을 추적할 수 있도록 하는 알고리즘을 제안하고, 이를 적용한 실험 결과를 통해 관찰자의 두 눈 위치 검출 속도와 검출 성공률이 매우 향상되었음을 보인다. 또한 생동감과 현실감 있는 영상을 제공하기 위해서는 관찰자의 위치에 따라 달라지는 물체의 모습을 표현해 주는 다시점 입체 영상 표시 기술이 필요한 바, 이를 위해 유한한 다수의 입력 영상 정보를 이용하여 입체 영상을 제작하고 관찰자의 위치 이동에 정확히 대응하는 영상을 합성하여 이를 디스플레이하는 방법을 제안한다.

## II. 두 눈 위치 검출 알고리즘

두 눈 위치 검출 알고리즘은 크게 전처리 과정과 두 눈의 중심의 픽셀 좌표를 얻는 눈 검출과정으로 나누어 수행하도록 하였다.

### 1. 전처리 과정

전처리 과정은 입력 영상으로부터 얼굴 영역의 위치를 결정하는 얼굴 영역 검출 과정과 조명 상태를 판단하여 알맞은 문턱값을 결정하는 과정, 그리고 얼굴 영역의 적색(red) 값의 평균 정규화 색상 계산 과정으로 나누어 수행하였다.

#### (1) 색상과 조명

칼라 입력 영상에서 한 화소는 각각 [0-255] 사이의 값을 가지는 R, G, B의 조합으로 표현되며, 이 모델은 하나의 색상뿐만 아니라 빛의 휘도까지 포함된 값이다. 따라서 이런 빛의 휘도 성분을 제외시킨 정규화 색상 모델을 사용한다. 색상을 이용하여 일련의 작업을 하려고 할 때 큰 문제로 나타나는 것이 조명의 영향이다. 조명의 변화에 색상은 영향을 받지 않으나 휘도는 크게 영향을 받으므로, 이러한 조명에 의한 영향을 제거시킬 수 있는 정규화 과정을 거칠 필요가 있

다. 즉 식 (1)의 정규화 과정을 거치게 되면 순수한 컬러 정보만을 지니게 된다.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \quad (1)$$

(2) 얼굴영역 검출 과정

본 연구에서는 순수 칼라의 영상 이미지를 사용하였으며, 한 픽셀의 R 레벨값을  $Rv$ , G 레벨값을  $Gv$ , B 레벨값을  $Bv$ 로 표기하였다. 얼굴 영역 검출 과정에서는 입력 영상에서 피부색과 가까운 부분(식 (2)의 피부색 조건에 만족하는 부분으로서, 이 조건은 실험을 통해 얻었음)만을 남기고 모두 제거한다.

<피부색 조건>

$$Rv > Gv, Rv > Bv, Bv > Gv \quad (2)$$

이렇게 처리된 이미지를 보면 얼굴 영역뿐만 아니라 다른 여러 부분들이 남게 되는데, 피부색 조건에 만족하는 픽셀이 가장 많이 그리고 넓게 분포되어 있는 부분을 얼굴 영역으로서 검출하고, 검출된 영역의 상하 좌우의 얼굴영역 좌표를 구한다.

(3) Threshold 레벨 값 결정 과정

입력 이미지는 조명의 상태에 큰 영향을 받으므로, 이 점을 보완하기 위하여 얼굴 영역 검출 과정에서 얻은 이미지로부터 조명 상태를 파악한다. 즉, 얼굴 영역 안의 피부색 픽셀들을 그레이 레벨로 변환하여 색상 성분을 없애고 조명의 휘도 정보를 찾아낸다.

$$Level = \frac{1}{i} \sum_{k=1}^i \frac{(Rvk + Gvk + Bvk)}{3} \quad (3)$$

여기서 Level은 얼굴 전체의 평균 Gray 레벨(0-255)을 나타내고,  $i$ 는 얼굴영역 내부의 피부색 픽셀의 수를 의미하며,  $Rvk, Gvk, Bvk$ 는 얼굴 영역 내부의 피부색 픽셀들의 R, G, B 레벨값을 나타낸다.

(4) 얼굴 영역의 적색(R) 값의 평균 정규화 색상 계산

2.2.2절의 두 눈의 위치 결정 부분에서 입을 검출하는데 사용하기 위하여 우선 얼굴영역의 적색 값의 평균 정규화 색상을 구한다. 다음 식 (4)에 의해 얻어진  $Qred$ 가 얼굴 영역의 적색 값의 평균 정규화 색상이다.

$$Qred = \frac{1}{i} \sum_{k=1}^i \frac{Rvk}{(Rvk + Gvk + Bvk)} \quad (4)$$

2. 눈 검출 과정

눈 검출 과정은 전처리 과정에서 얻은 얼굴 영역 부분 안에서 우선적으로 눈과 비슷한 크기와 색상 정보를 가지고 있는 부분(유사 눈)을 추출해내는 유사 눈 추출 과정과, 이렇게 찾은 유사 눈들로부터 실제 두 눈의 조건과 일치하는 것을 찾음으로써 두 눈의 위치를 결정하는 과정으로 나누어 수행하였다.

(1) 유사 눈 추출 과정

눈동자는 주위의 색과는 구별되게 검은 색의 ( $Rv, Gv, Bv$  값이 작은) 픽셀들로 구성되어 있으므로, 먼저 눈동자의 기준이 될 수 있는  $Rv, Gv, Bv$  값의 범위를 정하였다. 이 부분에서 조명의 영향을 감소시키기 위해 전처리 과정에서 얻은 Level 값에 대한 문턱값(threshold value, TV)를 결정한다. 그리고나서 유사 눈의 조건의 범위에 드는 픽셀 들이 가로와 세로에 얼마나 연속적으로 분포되어 있는지를 분석함으로써 눈동자 부분 영역을 파악해 내도록 하였다. 표 1은 Level 값에 대한 문턱값으로서 실험을 통해 구하였다. 본 연구에서는 카메라와 관찰자의 거리를 통상적인 관찰거리인 150~180(cm) 사이로 설정하여, 기준  $Rv, Gv, Bv$  값의 범위와 유사 눈의 가로, 세로의 기준 픽셀 수의 범위를 설정하였다.

표 1. Level 값에 대한 문턱값  
Table 1. Threshold value about the level value.

Level 값	90 이하	90~100	100~110	110~120	120~140	140~160	160 이상
TV	60	65	70	75	80	85	90

<유사 눈의 조건>

(a) 유사 눈의 기준 픽셀의  $Rv$  값의 범위 :  
 $Rv < TV$  (5)

(b) 유사 눈의 가로(W)와 세로(L) 크기의 기준 범위 :  
 $5 < W < 20 \quad 3 < L < 15$  (픽셀) (6)

(c) 유사 눈에서의 인접 픽셀간의 색상 레벨값의 차이 :  

$$\begin{aligned} |Rv(x_{k+1}, y_k) - Rv(x_k, y_k)| &< 15, \\ |Rv(x_k, y_{k+1}) - Rv(x_k, y_k)| &< 15, \\ |Gv(x_{k+1}, y_k) - Gv(x_k, y_k)| &< 20, \\ |Gv(x_k, y_{k+1}) - Gv(x_k, y_k)| &< 20, \\ |Bv(x_{k+1}, y_k) - Bv(x_k, y_k)| &< 20, \\ |Bv(x_k, y_{k+1}) - Bv(x_k, y_k)| &< 20 \end{aligned} \quad (7)$$

이상 3가지의 유사 눈의 조건은 실험을 통해 얻은 결과로서, 이 조건들을 만족하는 부분을 얼굴영역 안의 이미지를 분석하여 검출해낸다. 검출된 유사 눈들에 대하여 각각의 수평(X축) 및 수직(Y축) 중심좌표와 중심 픽셀 값을 구한다. 유사 눈들은 사각형의 모양으로 검출되므로 중심 좌표는 쉽게 계산되어진다. 그림 1은 입력 영상 전체를 분석하여 유사 눈의 조건에 부합되는 모든 부분을 흰색으로 변화시킨 것이다.

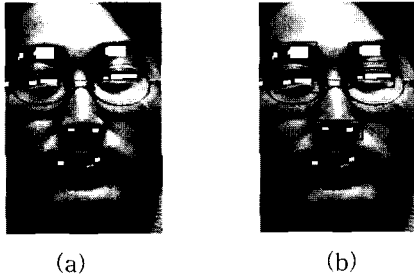


그림 1. 유사 눈 추출 결과 영상

(a) 결과 영상, (b) 유사 눈들의 그룹화

Fig. 1. Extracted images of the candidate eyes  
(a) extracted images, (b) grouping of the candidate eyes.

이렇게 검출된 유사 눈들은 그림 1의 (a)에서 보이는 것과 같이 약 10~30개 정도이고, 그 분포를 보면 눈, 눈썹, 입, 그리고 어두운 부분 등 일정 영역에 집중되어 있음을 알 수 있다. 유사 눈들이 일정영역에 모여 있음을 감안하여, 두 눈 사이의 거리를 감안해 유사 눈들 간의 거리가 40 픽셀보다 작은 경우의 유사 눈들을 하나의 그룹으로 만들고, 한 그룹 당 하나의 유사 눈만을 추출하도록 하는 간소화 단계를 거친다. 그림 1의 (b)는 유사 눈들의 그룹을 일정 영역으로 묶은 것이고, 이를 살펴보면 한 영역 안에 복수개의 유사 눈이 있음을 알 수 있다. 간소화 단계에서는 이 여러 개의 유사 눈들의 중심 픽셀의  $R_v, G_v, B_v$  값을 비교하여 그 중에서 가장 낮은 레벨 값(가장 검은 색상)을 갖는 유사 눈만을 남기고 나머지를 제거한다. 이 간소화 단계를 거친 후 남게 되는 유사 눈의 수는 10개 보다 작게 된다.

#### (2) 두 눈 위치 검출 과정

두 눈 위치 검출 과정은 간소화 단계를 거쳐 추출된 유사 눈들 중에서 두 눈의 조건과 일치하는 부분을 찾아내는 과정이다. 두 눈의 조건은 다음과 같이 설정하였다.

#### <두 눈의 조건>

두 눈으로 판단하기 위한 조건은 다음의 (a)~(d)와 같다.

(a) 관찰자와 카메라 사이의 거리를 앞서 언급한 바와 같이 150~180(cm) 정도로 가정했을때의 두 눈 사이의 픽셀 거리( $d$ )는  $40 \leq d \leq 80$ 의 범위에 있어야 한다. 이 거리 조건에 만족하는 두 유사 눈의 쌍들을 모두 구한다.

(b) 얼굴 안에서 두 눈은 대칭이다. 그러므로 각 유사 눈을 중심으로 일정 영역을 나누어 히스토그램 분석을 통하여 두 유사 눈의 대칭 여부를 판단한다. (a)의 거리 조건을 만족시키는 두 유사 눈의 쌍들 중에서 가장 대칭이라 판단되는 5개의 유사 눈의 쌍들을 선정한다. 여기서 대칭 정도의 판단은 R, G, B 각각의 색상 정보로써 판단한다. 즉 영역 안의 각각의 R, G, B 분포를 만들고 두 유사 눈 쌍의 각각의 R, G, B 분포의 차이 값을 계산하여, 모든 레벨(0~255)에 대하여 전부 합한 값이 가장 작은 것부터 5개의 유사 눈 쌍을 선정한다.

(c) 눈동자는 얼굴 내에서 가장 낮은 색상을 갖는다. 즉 b)의 과정을 거쳐 나온 유사 눈의 쌍들로부터, 각 유사 눈의 중심 픽셀의 레벨값을 비교하여 그 값이 가장 작은 쌍부터 순위를 정한다.

(d) 위의 (a)~(c)의 조건에 의해 순위가 결정된 유사 눈의 쌍으로부터, 정확한 두 눈의 쌍인가를 확인하기 위하여 입과의 위치관계를 확인한다. 일반적으로 사람의 얼굴은 그림 2와 같은 구조로 이루어져 있다. 이 그림에서 보이는 것처럼 사람의 두 눈 사이의 거리를  $d$ 라 할 때 일반적으로 입은 두 눈의 중심( $m$ )으로부터 약  $(4/3)d$ 만큼 아래에 위치하게 된다. 위의 (a)~(c)의 조건에 의해 판단된 두 유사 눈의 중심 좌표들로부터 두 유사 눈 사이의 중심점  $m$ 과 입의 중심점  $S$ 를 간단한 계산에 의하여 찾아 낼 수 있다. 이  $S$ 를 중심으로 입 영역을 설정하여 입 영역 안에서의 적색(R) 값 평균 정규화 색상(Qredlip)을 구한다. 입 영역의 크기는 가로가  $d$ 이고 세로가 15 픽셀이 되도록 설정하였다. 앞에서 정한 순위에 따라 Qredlip을 구하여 전처리 단계에서 구한 Qred와 비교한다.  $Qred < Qredlip$ 의 조건이 성립되면 입이  $S$ 점을 중심으로 위치해 있다고 판단하고, 이 때의 유사 눈 쌍을 두 눈이라 결정한다. 그렇지 않은 경우는 다음 순위의 유사 눈 쌍을 대상으로 위의 과정을 계속 수행한다.

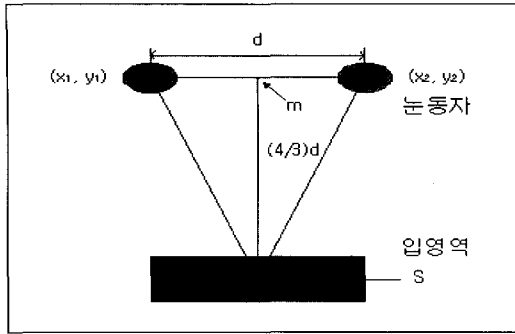


그림 2. 얼굴 구조  
Fig. 2. Face structure

### III. 두 눈 위치 검출 실험 결과

관찰자의 두 눈 위치 검출 실험에서는 다양한 관찰자를 대상으로 관찰자가 움직이는 환경 하에서 수행하였다. 입력되는 영상은 320×240의 순수 컬러이고, 관찰자는 검은 눈동자를 가진 동양인으로서 카메라를 주시하고 있으며, 관찰자와 카메라 사이의 거리  $L$ 은  $1.50 < L < 1.85(m)$ 의 범위 내에 있다는 조건하에서 실험을 수행하였다.

본 연구에서 사용한 방법은 관찰자의 변화와 거리의 변화에도 상당한 정도의 융통성을 보이며, 적응해 나갈 수 있었다. 모두 8명의 대상으로부터 80장의 입력 이미지를 얻어 실험을 실행한 결과 그중 78장의 이미지에서 두 눈의 위치를 정확하게 찾아 낼 수 있었다. 즉 97.5%까지 정확히 두 눈의 위치를 찾을 수 있었으며, 330MHz급 Pentium II PC를 사용하여 구현하였는데 두 눈 위치 검출에 걸린 수행 시간은 평균적으로 약 0.39초가 소요되었다. 앞서 언급한 기존의 방법들과 비교해보면 알 수 있듯이, 제안한 방법이 PC를 사용하면서도 수행 시간과 검출 성공률 면에서 훨씬 우수한 성능을 보이고 있음을 알 수 있다.

### IV. 3차원 영상 디스플레이

실제 영상을 이용하여 3차원 물체를 모델링하고 이를 이용해 관찰자의 두 눈의 위치에 대응하는 3차원 영상을 디스플레이하기 위해서, 본 연구에서는 실제 영상에서 깊이 정보를 추출하고, 3차원 물체 모델링을 하며, 스테레오 영상 표시를 하는 단계로 나누어 수행하였다.

#### 1. 깊이 정보 추출

실제 영상에서 깊이 정보를 얻는 방법으로는 스테레오 정합(stereo-matching) 방법 중 MR-MBS (Multi-Resolution Mutiple-Baseline Stereo)방식<sup>[9]</sup>을 사용하였다. 스테레오 정합 방법은 공간상의 같은 점이 2차원 입력 영상에서는 서로 다른 2차원 영상 좌표에 맷히는, 이른바 좌우 영상간의 좌표 차이를 이용하여 깊이 정보를 얻어내는 방법이다. 그 중에서 MR-MBS 방식은 영상의 해상도를 변화시키면서 정합을 해나가는 방식으로 고정해상도에서 처리하는 방식보다 연산 시간을 단축시킨다는 장점을 가지고 있으며, 단순한 스테레오 정합 방식에 비하여 폐색 영역 처리에 대해 우수하며, 경계선 연장 문제 해결에도 장점이 있는 방식이다.

#### 2. 3차원 물체 모델링

##### (1) 물체 추출

모델링을 하기 위한 첫 번째 단계는 디스페리티 맵에서 배경과 물체를 분리하는 작업으로 배경과 물체가 충분한 거리를 두고 떨어져 있다는 가정 하에서 디스페리티 히스토그램을 이용하여 물체를 배경과 분리하였다. 히스토그램을 분석하면 배경 부분이 되는 작은 디스페리티를 가지는 부분에서 많은 픽셀이 나타나고, 디스페리티가 약간 떨어진 이후에 다시 많은 픽셀 수를 가진 디스페리티 영역이 나타난다. 일반적으로 배경과 물체 사이에 충분한 거리가 있으며 물체는 배경보다 가까운 영역에 있게 되므로, 디스페리티 히스토그램에서 문턱값 이하의 화소를 가진 디스페리티는 노이즈 혹은 배경과 물체 사이의 공간이라고 생각하고 물체를 추출하게 된다. 이러한 방법으로 디스페리티 히스토그램을 이용하여 배경과 물체를 분리해 내고, 배경 부분에는 0 값을 넣어 물체의 정보만을 추출한 영상을 만들었다. 이 때 구한 디스페리티 맵을 그대로 모델화 하게 되면 피크(peak) 노이즈들로 인해 모델의 형태가 깨끗이 나오지 않게 되므로, 이러한 현상을 피하고 물체의 형태를 잘 보존하기 위하여 메디안 필터를 사용하여 주변보다 크거나 작은 피크 노이즈를 제거하였다.

##### (2) 삼각형 메쉬 생성 및 텍스처 매핑

물체만을 추출해 낸 후에 삼각형의 메쉬(mesh)를 이용하여 물체를 모델링하였는데, 삼각형 메쉬를 만들어내기 위해서는 물체의 깊이 정보를 잘 표현할 수 있

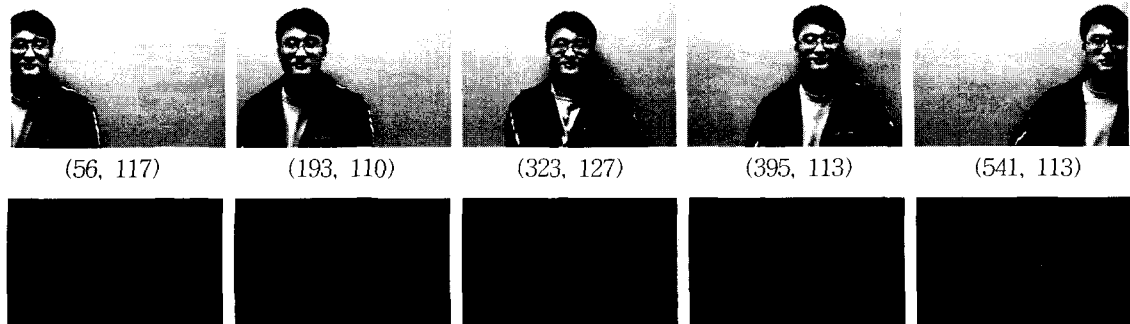


그림 3. 관찰자의 이동에 대한 시점 대응 3차원 영상 렌더링

Fig. 3. View-dependent 3D picture rendering for moving viewer.

는 위치에 삼각형 메쉬의 꼭지점이 되는 정점(vertex)의 위치를 찾아내어야 한다. 이렇게 모델을 생성해 내는 방법으로서 규칙적으로 정점을 추출한 후에 각 정점을 꼭지점으로 갖는 삼각형 메쉬를 만들고, 삼각형 메쉬 내부의 깊이 정보의 변위를 실제 깊이 맵과 비교하여 문턱값 이하가 될 때까지 삼각형 메쉬를 더 작은 삼각형 메쉬로 나누어주거나, 주변 메쉬가 가지는 깊이 변위가 문턱값 이하일때는 메쉬를 병합하는 방법을 취하였다. 그리고 모델에 현실감을 부여하기 위하여 텍스처 매핑 작업을 하였다. 텍스처 매핑은 모델을 만들기 위하여 사용한 중간 영상의 컬러 영상을 모델에 입혀 주어 실제 물체와 같은 느낌을 주도록 하였다. 또한 영상의 효율적인 합성을 위해 Open Inventor 환경에서 영상을 렌더링하였다.

### 3. 스테레오 영상 표시

스테레오 영상을 표시하기 위해 LCD shuttering 방법을 사용하였다. 3차원 모델을 수학적 연산을 통하여 좌, 우 영상을 만들고, 만들어진 두 개의 영상을 초당 30프레임으로 보여주기 위하여 1초에 120필드를 화면에 표시하며, 각 필드가 표시 될 때 나오는 동기 신호에 맞추어 좌, 우 영상을 표시한다.

## V. 최적의 시점 대응 3차원 디스플레이 실험 결과

### 1. 임의 시점에서의 3차원 디스플레이 실험 결과

관찰자의 위치 추적을 위해 CCD 카메라로 입력되는 영상은 640×480의 해상도를 가지며, 실험을 통해 관찰자의 두 눈의 중심 좌표를 정확히 검출할 수 있는 영역은 72cm×41cm의 영역이다. 관찰자와 카메라 사이

의 거리를 1.5m로 고정시키고, 관찰자를 상하, 좌우로 이동시켜 위치 추적 알고리즘을 적용시킨 결과로 나온 두 눈의 사이의 중심 좌표를 3차원 영상을 표시해주는 컴퓨터에 전송한다. 좌표값을 받은 컴퓨터에서는 관찰자의 이동에 해당하는 값을 회전값으로 바꾸어 3차원 모델을 합성하고 화면에 디스플레이 한다. 그림 3은 관찰자의 이동에 대한 시점 대응 영상을 렌더링한 결과이다.

렌더링되는 영상은 인벤터에 의해 Z-버퍼의 back-left 버퍼에서 좌측 영상이, back-right 버퍼에서 우측영상이 렌더링되어 화면상에 교대로 보여지는 것이므로 입체안경을 착용한 경우에만 3차원으로 볼 수 있다. 시점 대응 렌더링 영상이 관찰자에게 보이는 최적의 영상인지를 확인하기 위하여 좌우 Z-버퍼에서 렌더링되는 좌우 스테레오 영상과 로봇을 이용하여 좌우 카메라로 얻은 영상과의 유사도를 측정하였다.

### 2. 임의 시점 영상 렌더링

관찰자의 눈을 정확히 추적할 수 있는 영역을 좌우로 7등분 7등분(4.27° 간격), 상하로 5등분(3.82° 간격)하여 각 지점에서의 시점 대응 영상을 렌더링하였다.



그림 4. 인벤터를 통해 z-버퍼에서 렌더링되는 좌우 영상 (a) 모델의 좌측 버퍼 영상 (b) 모델의 우측 버퍼 영상

Fig. 4. The stereo image rendered at z-buffer by Open Inventor. (a) left buffer image. (b) right buffer image.

2. 로봇 시뮬레이션 및 영상 획득

로봇 팔에 CCD 카메라를 장착하여 실제 F16 모델을 촬영하였다. 로봇은 6축 다관절형으로 삼성 FARAMAN AS1-i를 사용하였고, CIMetrix를 이용하여 로봇을 제어하였다. CIMetrix는 개방형 제어 프로그램으로서, 실제 로봇을 모델링하는 CIMTools, 모델링된 로봇의 작업을 시뮬레이션하는 CIMulation, 로봇을 제어하는 CIMControl, 그리고 CODE API로 구성되어 있다.

임의 시점에서 렌더링시킨 위치와 동일한 11곳의 티칭포인트(TP: teaching piont) 및 각 TP를 기준으로 0.5° 간격으로 TP를 생성하고 시뮬레이션하였으며, 로봇을 구동하여 각 TP에서의 스테레오 영상을 얻었다. 카메라의 간격은 인간의 두 눈 사이의 거리인 65mm로 정하였다. 로봇을 이용하여 얻은 영상은 640×480의 해상도를 가지며 에피폴라 구속조건을 충족시킨다.



그림 5. 로봇을 이용하여 좌우 카메라로 얻은 좌우 영상 (a) 실제 모델의 좌측 영상 (b) 실제 모델의 우측 영상  
Fig. 5. The stereo image captured by stereo camera using robot. (a) left image. (b) right image.

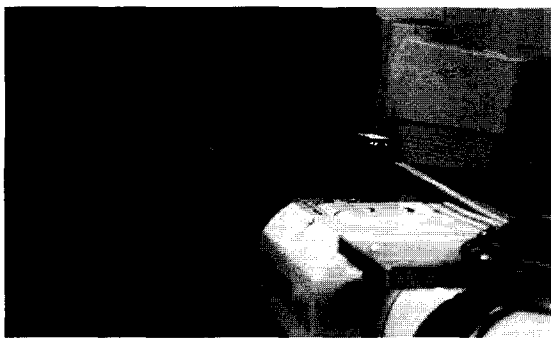


그림 6. 실험 환경  
Fig. 6. Experiment environment

3. 실험 환경

관찰자를 추적하는 거리와 동일한 거리인 1.5m의 위치에 로봇에 CCD 카메라를 설치하고 CIMetrix에 의해 정확한 3차원 좌표를 이동하여 11곳의 TP와 각

TP를 기준으로 0.5° 간격으로 로봇을 이동시켜 F16 모델의 스테레오 영상을 획득하였다.

4. 영상의 유사도 측정 실험 결과

기준이 되는 TP에서의 버퍼의 영상과 로봇으로 얻은 영상과의 유사도를 측정하여 최적의 시점 대응 영상을 구한다. 영상의 비교는 먼저 영상에서 물체를 배경과 분리시킨 후, 물체의 가로 길이를 구하고 물체에서 가장 윗 부분에 해당하는 특징점의 위치를 결정하고 식(8)에 따라 두 영상의 유사도를 구하였다.

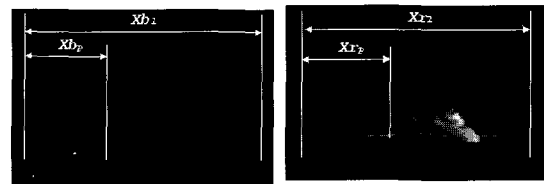


그림 7. 유사도 측정 방법  
Fig. 7. Similarity estimation method

$$Error = \left| \frac{Xb_p}{Xb_i} - \frac{Xr_p}{Xr_i} \right| \times 100 \quad [\%] \quad (8)$$

여기서  $Xb_i, Xb_p$ 은 버퍼 영상의 가로 거리(픽셀)와 특징점의 위치이고,  $Xr_i, Xr_p$ 은 로봇을 이용하여 얻은 영상의 가로 거리(픽셀)와 특징점의 위치이다.

표 2. 버퍼에서 렌더링되는 영상과 로봇을 이용하여 얻은 영상과의 에러 측정 실험 결과

Table 2. Error measurement result between stereo image rendered at z-buffer and stereo image captured by stereo camera using robot.

각 도 [°]	-1.5	-1.0	-0.5	0	0.5	1.0	1.5	
Left	TP0	0.819	0.331	0.279	0.041	0.392	0.539	0.926
	TP1	0.795	0.624	0.165	0.097	0.187	0.195	0.691
	TP2	0.668	0.478	0.339	0.239	0.301	0.593	0.880
	TP3	1.120	0.983	0.235	0.050	0.451	0.523	0.684
	TP4	0.854	0.312	0.208	0.187	0.337	0.466	0.753
	TP5	0.916	0.456	0.194	0.006	0.434	0.798	0.984
Right	TP6	0.965	0.717	0.398	0.233	0.368	0.798	0.984
	TP0	0.879	0.232	0.108	0.068	0.498	1.001	1.185
	TP1	0.721	0.416	0.354	0.259	0.516	1.157	1.269
	TP2	0.426	0.384	0.258	0.231	0.331	0.888	0.915
	TP3	1.049	0.793	0.227	0.020	0.132	0.294	0.611
	TP4	0.769	0.449	0.273	0.234	0.423	0.699	1.077
	TP5	0.811	0.211	0.199	0.083	0.398	0.936	1.221
TP6	0.662	0.460	0.366	0.322	0.483	0.899	1.315	

표 2는 버퍼에서 렌더링되는 영상과 로봇을 이용하여 얻은 영상과의 비교 에러를 구한 실험 결과이다. 표를 보면 임의의 위치에서 렌더링 시킨 영상과 동일한 위치에서 로봇을 이용하여 얻은 영상과의 에러가 가장 작은 것을 알 수 있다. 따라서, 오차가 가장 작은 TP에서 렌더링되는 영상이 관찰자에게 가장 최적으로 보이는 영상이라 할 수 있다.

## VI. 결 론

본 논문에서는 관찰자의 두 눈의 위치를 검출할 수 있도록 하는 알고리즘을 제안하고, 이의 적용 실험 결과를 제시하였다. 제안한 방법을 통해 0.39초의 짧은 시간 내에 약 97.5%까지 정확히 두 눈의 위치를 찾을 수 있었고, 관찰자의 위치에 따라 달라지는 물체의 모습을 표현해 주는 다시점 입체 영상 표시 기술이 필요한 바, 이를 위해 다수의 입력 영상 정보를 이용하여 관찰자의 위치 이동에 정확히 대응하는 영상을 구현하여 이를 디스플레이하는 방법을 제안하였다. 그리고 실제 로봇을 이용하여 좌우 카메라로 얻은 좌우 영상과 인벤텐터를 통해 버퍼에서 렌더링되는 스테레오 영상과의 유사도를 측정하여, 관찰자의 시점에 대응하여 렌더링되는 3차원 영상이 최적의 시점 대응 영상임을 확인하였다.

한편 헤드 트래킹시 관찰자의 기울임, 거리 변화, 조명의 변화에서도 정확히 두 눈의 위치를 찾고, 다수의 관찰자의 위치를 파악하기 위해서는 더욱 많은 연구가 필요하다고 판단된다.

## 참 고 문 헌

- [1] Hiroshi Imai, Masao Imai, Yukio Ogura, "Eye-Position Tracking Stereoscopic Projector Using Image Shifting Optics," Proceedings of the Third International Display Workshops (IDW), pp. 449-452, Kobe, Japan, Nov., 1996.
- [2] Ho-Sub Yoon, Jung Soh, Min Wang, and Byung-Woo Min, "Human Eye Detection in a Complex Background," The Fourth International Conference on Control, Automation, Robotics and Vision(ICARCV'96), Vol. 1, pp. 380-384, Singapore, Dec., 1996.
- [3] Taro Yokoyama, Haiyuan Wu and Masahiko Yachida, "Automatic Detection of Facial Feature Points and Contours," Proceedings of 5th IEEE International Workshop on Robot and Human Communication, pp.335-340, Tsukuba, Nov., 1996.
- [4] Haiyuan Wu, Takeshi Fukumoto, Qian Chen, and Mashhiko Yachida, "Face Detection and Rotations Estimation using Color Information," Proceedings of 5th IEEE International Workshop on Robot and Human Communication, pp. 341-346, Tsukuba, Nov., 1996.
- [5] J. Weng, T. S. Huang, and N. Ahuja. Motion and Structure from Image Sequences, Springer-Verlag, Berlin, 1993.
- [6] Tomohiko Hattori, "Streographic Image Taking and Output Devices Compatible with Conventional Infrastructure," Proc. of International Workshop on 3D Imaging Media Technology, pp. 43-150, March 1996.
- [7] C. Fermuller and Y. Aloimonos, "The Role of Fixations in Visual Motion Analysis," Int'l J. Computer Vision, 11:2, pp. 165-186, 1993.
- [8] M. J. Swain and M. A. Stricker, "Promising Directions in Active Vision," Int'l J. Computer Vision, 11:2, pp. 109-126, 1993.
- [9] 박남준, "임의 시점 영상 합성을 위한 다해상도 다시점 스테레오 정합 기법," 연세대학교 석사학위 논문, 1997



저 자 소 개



許 慶 茂(正會員)

1979년 서울대학교 전자공학과 졸업 한국과학기술원 전기및전자공학과 석사(1981), 동대학 박사(1989). 현재 단국대학교 공학대학 전자공학과 부교수. 주요 연구분야는 로봇트 제어, 3차원 디스플레이, 시스템 제어, 학습 제어, 디지털 제어기 설계 및 Simulation, Servo Motor 제어



金 明 新(正會員)

1998년 단국대학교 전자공학과 졸업. 현재 동대학원 석사과정. 관심 분야는 3차원 디스플레이