

논문-99-4-2-09

가상스튜디오에서 실사 TV 카메라의 3-D 기준 좌표와 추적 영상을 이용한 카메라 파라메타 추출 방법

한기태*, 김희율*

Camera Parameter Extraction Method for Virtual Studio Applications by Tracking the Location of TV Camera

Gi-Tae Han* and Whoi-Yul Kim*

요 약

가상스튜디오에서 보다 현실감 있는 영상을 제작하기 위해서는 전경화면의 객체와 컴퓨터그래픽으로 지원하는 가상 배경 화면간의 동기화를 정확히 실현하는 것이 중요하다. 본 논문은 TV 카메라(실사 카메라)에 대한 추적 영상좌표를 이용하여 동기화에 필요한 카메라 파라메타의 추출방법을 제안한다. 추적 영상에 대한 역투시 방정식과 실사 카메라에 대한 기준좌표와 추정좌표간의 3-D 변환 식으로부터 카메라 파라메타 추출 식을 유도하고, 유도된 식을 이용하여 추출한 파라메타와 실사 카메라에 적용한 파라메타 간의 일치정도를 비교하여 제안한 방법에 대한 타당성을 보인다.

Abstract

In order to produce an image that lends realism to audience in the virtual studio system, it is important to synchronize precisely between foreground objects and background image provided by computer graphics. In this paper, we propose a method of camera parameter extraction for the synchronization by tracking the pose of TV camera. We derive an equation for extracting camera parameters from inverse perspective equations for tracking the pose of the camera and 3-D transformation between base coordinates and estimated coordinates. We show the validity of the proposed method in terms of the accuracy ratio between the parameters computed from the equation and the real parameters that applied to a TV camera.

I. 서 론

컴퓨터 비전 분야는 최근 들어 정보통신 및 컴퓨터 기술의 발달과 함께 다양한 산업 분야에 응용되고 있다. 방송기술 역시 중요한 응용분야가 되었으며 가상 스튜디오는 이러한 컴퓨터 비전 분야의 응용이라고 볼 수 있다. 가상 스튜디오란 실제의 스튜디오에서 연기를 하는 배우가

그래픽으로 만들어진 가상의 무대에서 연기하는 것처럼 보이게 하는 것이 목적이다^{[1][2]}.

가상 스튜디오를 현실감 있게 구현하기 위해서는 실제 객체를 바라보고 있는 TV 카메라(실사 카메라)의 움직임에 따라 그래픽 배경 영상에 맞추어져 있는 가상 카메라 역시 같은 정도의 움직임을 가지고 움직여 주어야 한다^{[3][4][5][6][13][14]}. 그래야만 전경화면의 배우 객체가 배경화면과 동기화가 이루어져 자연스러운 영상을 연출할 수가 있다. 이를 위해서는 무엇보다도 전경화면을 촬영하는 카메라의 움직임 변수인 카메라 파라메타를 찾아 가상 카메라

* 한양대학교 전자공학과 영상공학 연구실
Image Engineering Lab., Dept. of Electronic Eng., Hanyang Univ.

라에 전달하는 것이 중요하며, 이것을 실시간으로 정확하게 추출하는 것이 본 논문의 목적이다.

현재의 가상 스튜디오 기술은 대부분 메모리 헤드라는 카메라에 부착된 기계적 장치를 통하여 측정된 회전과 줌의 정보를 이용하고 있으며^{[4][7]}, 최근에는 영상처리 기법을 이용하여 카메라 파라메타를 추출하고자하는 연구도 진행되고 있다^[13]. 그러나 전자는 카메라위치가 고정되기 때문에 Dolly In/Out이 불가능하여 다양한 화면 연출이 어렵고 카메라에 부착된 상태에서 작업이 이루어져야 하므로 카메라마다 추적 장비를 부착할 경우 많은 경제적 부담이 따른다. 후자는 다양한 패턴 인식 기법이 동원됨으로 인한 처리 시간의 과다 소요 및 부정확한 파라메타 추출 가능성이 있고 카메라 위치에 따라 패턴이 너무 크게 나타날 경우 기준점이나 패턴의 추적을 잃어 버려 기하학적 변화량을 찾지 못할 수도 있다는 제약을 가지고 있다^[14].

본 논문에서 제안한 방법은 기존 방법의 경제적 부담과 제약점을 해결하면서 실시간으로 정확한 카메라 파라메타를 추출하는데 초점을 맞추었다. 이를 구현하기 위해서는 TV 카메라 외에 카메라를 추적하는 추적용 CCD 카메라가 부수적으로 소요된다. 추적 카메라가 캡처한 영상에서 실사 카메라에 대한 위치를 쉽게 얻기 위한 수단으로 실사 카메라의 중심으로부터 좌우 대칭되는 일정한 거리의 위치와 중심 전방의 특정 위치에 LED를 부착하여 추적 목표점으로 이용한다. 파라메타 추출은 실사 카메라에 부착된 LED에 대한 미리 설정된 스튜디오상의 실세계 기준 좌표와 추적 영상에 나타난 LED 좌표간의 기하학적 변환 관계로부터 추출한다. 줌 값은 일반적으로 전경화면을 촬영하는 실사 카메라의 전경(FG: Fore Ground) 영상과 함께 얻을 수 있으므로 연구대상에서 제외하였다.

본 논문에서는 영상의 합성과정은 배제하고 합성에 필요한 카메라 파라메타 추출 방법만을 제안한다. 추적 카메라(tracking camera)는 스튜디오 천장의 중앙에 고정되어 있고 줌은 없으며 렌즈에 대한 왜곡(distortion)은 레디알 왜곡^{[15][16]}만을 고려한다. 실험은 실제 스튜디오 상에서 실사 카메라에 적용한 파라메타 값과 제안한 방법으로 추출한 파라메타 값의 정확도 비교로 방법의 타당성을 보인다.

II. 스튜디오 세트의 환경 구성

본 논문에서는 스튜디오 촬영세트내의 실사 카메라에 부착된 특정 객체의 공간좌표 변화와 추적 영상에 나타난 특정 객체 좌표간의 기하학적 관계로부터 카메라 파라메

타 추출 식을 유도하여 기존 방식의 문제점을 해결한다.

가상스튜디오 촬영 세트의 구성은 기존의 고가 장비를 배제하고 실시간으로 처리하기 위한 환경으로 구성한다. 그림 1에서 볼 수 있듯이 촬영 카메라 외에 CCD 카메라 1대를 천장의 중앙에 매달고, 촬영 카메라의 헤드에 영상처리 및 카메라 파라메타 추출시 참조점으로 이용하기 위한 그림 2와 같은 3개의 LED를 부착한다. 전경화면의 객체를 촬영하는 TV 스튜디오 카메라는 일반적으로 롤(roll)이 고려되지 않기 때문에^{[7][14]} 3개의 LED 만을 추적 목표점으로 이용한다.

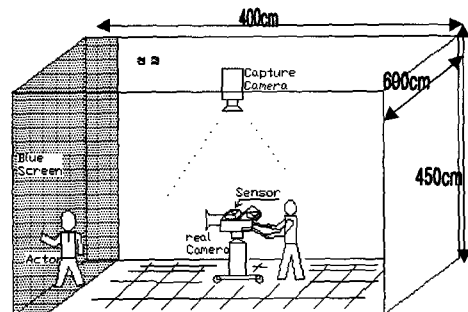


그림 1. 스튜디오 세트 구성
Fig. 1. Studio set configuration

본 논문에서 추출하고자하는 카메라 파라메타는 카메라의 움직임(pan, tilt)과, 틸트(tilt)각, 붐(boom) 값 및 실사카메라의 위치(position)로 이것들은 3개의 LED에 대한 실제 스튜디오 상에서의 기준 위치 좌표와 추적 영상으로부터 계산한 좌표로부터 추정된다. 추출대상 파라메타인 펜각은 배경(BG: Back Ground) 영상의 좌우 이동에 영향을 주며, 틸트각과 붐값은 상하이동에 영향. 그리고 실사카메라의 위치좌표는 배경 영상의 확대 및 축소에 영향을 준다.

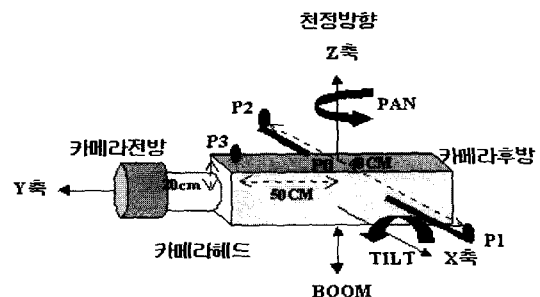


그림 2. TV 카메라 헤드상의 LED(P1,P2,P3) 위치
Fig. 2. LED position(P1,P2,P3) on real camera head

실사 카메라의 중심 좌우에 부착된 2개의 LED는 카메라의 X축에 해당되고, 중심 전방에 부착된 LED는 Y축에 해당되며, 카메라의 중심으로부터 천장 방향은 Z축에 해당된다.

카메라의 움직임 종류와 그림 2에서의 좌표축 변환 관계는 표 1과 같다.

표 1. 카메라 움직임종류와 좌표축간의 관계
Table 1. Relationship between camera parameter and coordinate axis

카메라 움직임	좌 표 축	관계LED	효 과
PAN	Z축을 좌우로 회전	P ₁ ,P ₂ ,P ₃	X,Y값에 변화
TILT	X축을 앞뒤로 회전	P ₃	X,Y,Z값에 변화
BOOM	Z축 위아래로 이동	P ₁ ,P ₂ ,P ₃	Z값에 변화
POSITION	X,Y축으로 이동	P ₁ ,P ₂ ,P ₃	X,Y값에 변화

또한 스튜디오 세트에서 사용되는 TV 카메라는 롤이 없어, 그림 2에서 Y축에 대한 회전이 발생하지 않으므로 점P₁과 P₂에 대한 Z값은 항상 동일함을 전제로 한다. 실사 카메라의 기준 좌표는 스튜디오 그라운드상의 어떤 위치를 잡아도 문제는 없으며 한번 설정된 기준 좌표는 스튜디오 세트에 변화가 있기 전까지 프로그램 내에 설정해 놓고 사용할 수 있다. 실험의 편의를 위하여 본 논문에서는 스튜디오 그라운드상의 정 중앙(추적 카메라의 정 중앙 아래)에서 실사카메라의 Y축이 전방(블루스크린)을 향하고, X축 상의 P₀가 추적 카메라와 수직일 때의 3차원 좌표를 실사 카메라의 기준 좌표로 사용한다.

III. 파라메타 추출 식 유도

폐쇄 공간 좌표와 영상 좌표간의 변환식을 기반으로 카메라 파라메타 추출식을 유도한다. 3차원 실세계 좌표상의 한 점 $p(x,y,z)$ 을 투시사영(perspective projection)하여 영상평면(image plane) 상에 나타난 좌표를 $p'(x',y')$, 사영 참조 점(projection references point)을 z_{vp} , z_{pp} 와 같은 z 축선 상에 있는 영상평면상의 점을 z_p , 그리고 초점거리를 f 라 할 때 다음과 같은 그림에서 투시변환(perspective transformation) 방정식을 얻을 수 있다.

그림 3은 3차원상의 점 p 가 영상평면상의 점 p' 로의 사영을 나타내고 있으며, 실세계 좌표로부터 영상좌표로의

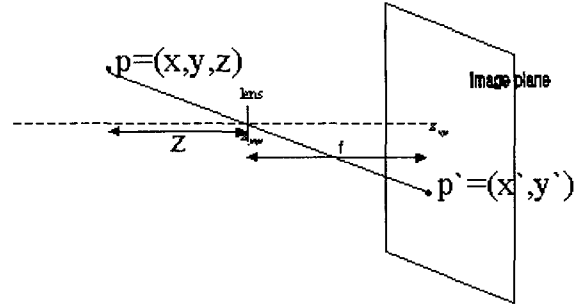


그림 3. 실세계 좌표와 영상 플랜과의 관계
Fig. 3. Relationship between real world and image plane

투시변환 식^{[8][9][10][11]}은 다음과 같다.

$$\begin{aligned} x' &= x \left(\frac{z_{pp} - z_{vp}}{z_{pp} - z} \right) = x \left(\frac{d_p}{z_{pp} - z} \right), \\ y' &= y \left(\frac{z_{pp} - z_{vp}}{z_{pp} - z} \right) = y \left(\frac{d_p}{z_{pp} - z} \right) \end{aligned} \quad (1)$$

여기서 $d_p = z_{pp} - z_{vp}$ 는 사영 참조 점과 영상평면간의 거리로 초점거리 (f)를 말한다.

식 3을 3차원 동차좌표(homogeneous coordinates)를 사용하여 행렬 형태로 표현하면

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ h \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -z_{vp}/d_p & z_{pp}/d_p \\ 0 & 0 & -1/d_p & z_{pp}/d_p \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (2)$$

가 되고, 동차 factor $h = (z_{pp} - z)/d_p$ 이 된다.

그러므로 영상 평면상의 사영좌표는 동차좌표로부터 다음과 같이 얻을 수 있다.

$$x' = x_h/h, \quad y' = y_h/h \quad (3)$$

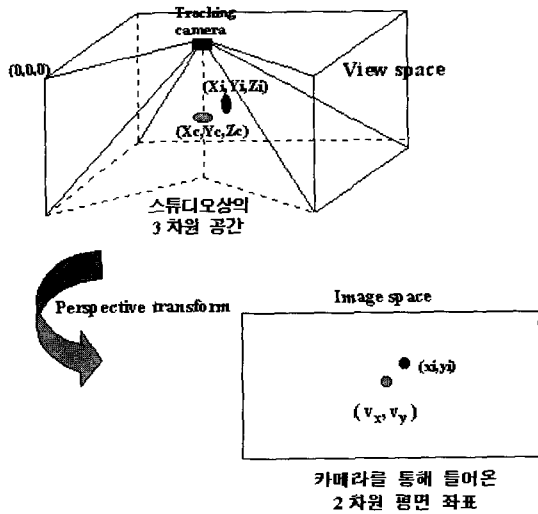
위의 식에서 z_{pp} 를 z축 상의 0의 위치로 놓으면 d_p 는 초점거리와 동일하게 되므로 이를 본 논문에서 제안하는 카메라의 실세계 좌표를 구하는 방법에 적용한다.

그림 1처럼 폐쇄 공간인 스튜디오 상에서 추적 카메라의 수직 정 중앙 아래에 위치한 기준 실사 카메라 영상의 중심 좌표를 x, y 라하고, 기준 실사 카메라의 3차원 중심 (P_0) 좌표를 X, Y, Z 라하며 초점거리를 f 라 할 때 기준 영

상의 중심 좌표(x, y)는 다음의 식으로 나타낼 수 있다.

$$[x \ y]^T = [x \times f \times 1/Z \quad y \times f \times 1/Z] \quad (4)$$

실제적으로 추적 카메라로 들어오는 스튜디오의 3차원 폐쇄 공간(시야공간:View Space)상 특정 객체의 한 점



(v_x, v_y) : 영상의 중심 좌표, (x_i, y_i) : 특정 위치의 영상 좌표
 (X_c, Y_c, Z_c) : 실세계 view space내의 중심 좌표
 (X_i, Y_i, Z_i) : 특정위치의 실세계 좌표, f 초점거리

그림 4. 폐쇄 3-D 공간과 2-D 영상 평면
 Fig. 4. Closed 3-D space and image plane.

좌표와 이에 해당하는 영상 좌표간의 관계를 그림 4에서 나타내고 있으며, 이들간의 관계식을 아래에서 유도한다.

그림 4에서 시야공간상 특정 위치의 영상좌표는 투시사영 식으로부터

$$x_i = v_x \left(1 + \frac{(X_i - X_c) \times f}{Z_i}\right), y_i = v_y \left(1 + \frac{(Y_i - Y_c) \times f}{Z_i}\right) \quad (5)$$

가되며, 영상에서의 특정 위치에 대한 실세계 좌표는 역투시 변환(inverse perspective transform)에 의해

$$X_i = X_c + \frac{Z_i \times (x_i - v_x)}{f \times v_x}, Y_i = Y_c + \frac{Z_i \times (y_i - v_y)}{f \times v_y} \quad (6)$$

이 된다.

IV. 카메라 파라메타 추출 방법

추적 영상에서 3개의 LED에 대한 좌표 값과 스튜디오 상 실사 기준 카메라의 LED 좌표 값을 이용하여 FG와 BG를 동기화하는데 필요한 카메라 파라메타를 추출하고자 한다. 여기서는 영상처리기법^[12]을 사용하여 추적 영상으로부터 카메라에 부착된 LED의 영상 위치를 찾고, 영상에서의 LED 좌표(P_1, P_2, P_3)와 앞에서 유도한 식(6)을 사용하여 가상 스튜디오 구현에 필요한 각 파라메타 추출식의 유도 및 추출 방법을 제안한다. 파라메타 추출과정은 그림 5와 같은 절차를 거친다.

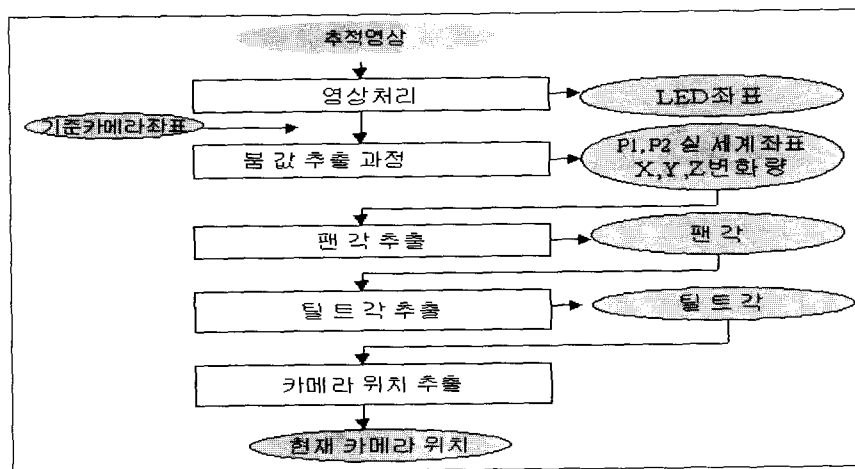


그림 5. 카메라 파라메타 추출 과정
 Fig. 5. Camera parameter extracting flow chart.

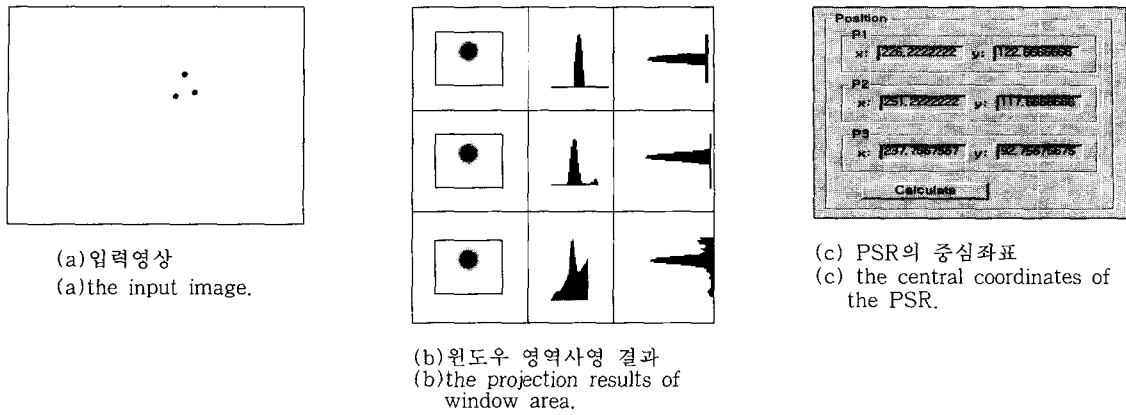


그림 6. 영상처리 결과
Fig. 6. The result of image processing.

1. 영상처리 과정

그레이 영상인 추적 영상을 좌에서 우 및 상에서 하 방향으로 스캔하면서 각 화소의 값이 특정 값(threshold value) 이상인 것만을 배열에 저장한 후 배열로부터 뭉쳐진 점 영역(PSR: Point Support Region)의 크기보다 작은 영역을 버리게 되면 빛이나 조명에 의해 부가되었던 노이즈가 제거된 점 영역만이 표시된 영상이 된다^[12]. 추출된 점 영역들의 PSR로부터 각각 평균좌표를 계산한 후 평균좌표와 가장 가까운 좌표에 $m \times n$ (예: 30×40) 화소의 윈도우를 씌어 그 영역의 X와 Y축 방향에 대한 사영 결과로 PSR의 중심점을 구하고 이 점을 각 LED에 대한 영상 중심 좌표로 이용한다. 영상처리 알고리즘은 다음과 같으며 그 결과에 대한 예를 그림 6으로 나타내고 있다.

[알고리즘 1]

- ① 각 화소에 대하여 특정 기준값(threshold value)으로 thresholding를 수행한다.
- ② PSR의 크기가 일정 기준값(area threshold value) 이상인 것만 검출한다.
- ③ 선택된 PSR에 대한 평균 좌표를 구한다.
- ④ 평균 좌표와 가장 가까운 화소를 중심으로 $m \times n$ 화소 크기의 윈도우를 씌운다.
- ⑤ 가로축과 세로축에 대하여 각각 사영을 수행한다.
- ⑥ 각 축 방향으로 사영한 결과의 최고점(peak point)을 중심으로 $r\%$ 의 영역을 선택한다.
- ⑦ intensity 에 따라 가중치가 반영된 선택 영역의 평균 좌표를 구한다.

2. 붐값 및 팬각 추출

'Rigid Body Transformation'^[11]에 의하여 식(6)의 역투시사영 관계식으로부터 Z값을 찾을 수 있다. 기준 카메라 좌표로부터 현재 실사 카메라 좌표로의 X, Y, Z축에 대한 이동량을 d_x, d_y, d_z 라 할 때 실사 카메라에 대한 기준카메라 위치와 영상으로부터 추정된 카메라 좌표로부터 Z값 변동 여부와 변화량을 찾는다. 그림 2에서 P_1 과 P_2 의 실제계 기준 좌표를 (X_1^B, Y_1^B, Z_1^B) 와 (X_2^B, Y_2^B, Z_2^B) 라 할 때 P_1 과 P_2 의 거리는 X, Y, Z축의 회전이나 이동에 불변함으로 아래와 같은 알고리즘 2를 적용하여 Z값의 변화량 d_z 를 추정한다.

[알고리즘 2]

기준 객체상의 P_1 과 P_2 의 거리가 $D_0 = \sqrt{(X_2^0 - X_1^0)^2 + (Y_2^0 - Y_1^0)^2}$ 이고, 기준 객체상의 P_1 과 P_2 의 Z좌표 값이 Z^B 일 때

- ① 식(6)에 실사 카메라에 대한 영상의 p'_1 과 p'_2 좌표와 Z^B 을 적용하여 거리 D_a 를 계산한다.

$$D_a = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$$

- ② 거리차($dif = |D_a - D_0|$)를 계산한다.
- ③ dif 가 threshold value 이하이면 Z^B 를 그대로 적용한다 (boom 값 변동 없음).
- ④ ③의 조건을 만족하지 않으면 dif 가 threshold value 이하로 될 때의 Z값 변화량 d_z 를 아래와 같은 탐색 알고리즘을 적용하여 찾는다(단, t_z 는 일정 범위 안의 z값 변화량이며, X_i 와 Y_i 산출 식은 식(6)의 확장임).

```

loop:   tz 초기값, 증분값, 최대값
        Xi = Xc +  $\frac{(Z^B \pm tz) \times (x_i - v_x)}{f \times v_x}$ ,
        Yi = Yc +  $\frac{(Z^B \pm tz) \times (y_i - v_y)}{f \times v_y}$ 
        tz = z 변화값(범위내)
        Db =  $\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2}$ 
        거리차(dif) = |Db - Do|
        if dif < threshold value
            dz = tz
            break
        end
end loop
    
```

실세계상의 P₁과 P₂의 기준 좌표에 대한 중심좌표를 (X_b, Y_b, Z^B)라 할 때, (X^B₁, Y^B₁, Z^B)와 (X^B₂, Y^B₂, Z^B)로부터 X_b=(X^B₁+X^B₂)/2, Y_b=(Y^B₁+Y^B₂)/2로 나타낼 수 있으며, x_c=(x₁+x₂)/2, y_c=(y₁+y₂)/2 와 d_z값을 이용하여 X축과 Y축에 대한 변화량 d_x, d_y을 아래의 식(7)으로부터 구할 수 있다.

$$\begin{aligned}
 d_x &= X_c + \frac{(Z^B \pm d_z) \times (x_c - v_x)}{v_x \times f} - X_b \\
 d_y &= Y_c + \frac{(Z^B \pm d_z) \times (y_c - v_y)}{v_y \times f} - Y_b
 \end{aligned}
 \tag{7}$$

Z축에 대한 회전량인 팬각은 알고리즘 2를 통해 얻은 3차원 공간상의 점P₁과 점P₂에서 XY 평면상의 좌표인 P₁(X₁, Y₁), P₂(X₂, Y₂)로 구한 기울기에 arctan을 취하여 추출한다.

$$\begin{aligned}
 \text{기울기} &= (Y_2 - Y_1) / (X_2 - X_1) \\
 \text{팬각}(\alpha) &= \arctan(\text{기울기}) \times 180 / \pi
 \end{aligned}
 \tag{8}$$

3. 틸트각 및 카메라위치 좌표 추출

틸트각은 객체의 X축에 대한 회전 량으로 점 P₃의 이동 및 회전 변환 식으로 유도되는 X, Y, Z 계산식과 영상 의 p'₃에 대한 역 투시변환으로 유도한 X, Y 계산식을 이

용하여 구할 수 있다. 실사 카메라 기준 위치에서 P₃에 대한 좌표를 P₃=[X^B₃ Y^B₃ Z^B₃]^T라하고, 위치 및 회전 변동 후 P₃에 대한 좌표를 P'₃=[X₃ Y₃ Z₃]^T라 할 때 각 회전 및 이동 량을 행렬로 표현하면 다음과 같다.

기준 좌표상의 객체 중심 P₀를 실세계 좌표의 원점으로 이동시킬 변화량(T₀)은

$$T_0 = [1 \ 0 \ 0 \ -x_0; 0 \ 1 \ 0 \ -y_0; 0 \ 0 \ 1 \ -z_0; 0 \ 0 \ 0 \ 1]^T$$

이 되고, 이동 후 객체의 X축을 실세계 좌표의 X축과 일치시키기 위한 회전 변화량(R₀)은

$$R_0 = [\cos \mu \ -\sin \mu \ 0 \ 0; \sin \mu \ \cos \mu \ 0 \ 0; 0 \ 0 \ 1 \ 0; 0 \ 0 \ 0 \ 1]^T$$

이 된다.

여기서 μ는 기준 객체에서 점 P₁과 P₂의 기울기에 arctan을 취하여 구할 수 있으며, w₁=cos μ, w₂=sin μ, v₁=cos(-μ), v₂=sin(-μ) 로 놓으면,

$$R_o = \begin{bmatrix} w_1 & -w_2 & 0 & 0 \\ w_2 & w_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad R_o^{-1} = \begin{bmatrix} v_1 & -v_2 & 0 & 0 \\ v_2 & v_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

이 된다.

Z축에 대한 회전 변화량인 펜각의 변화량(R_p)은

$$R_p = [\cos \alpha \ -\sin \alpha \ 0 \ 0; \sin \alpha \ \cos \alpha \ 0 \ 0; 0 \ 0 \ 1 \ 0; 0 \ 0 \ 0 \ 1]^T$$

으로, a₁=cos(α), a₂=sin(α) 로 놓으면

$$R_p = \begin{bmatrix} a_1 & -a_2 & 0 & 0 \\ a_2 & a_1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

으로 표현할 수 있다. 또한 X축에 대한 회전 변화량인 구 하고자하는 틸트각의 변화량(R_t)은

$$R_t = [1 \ 0 \ 0 \ 0; 0 \ \cos \theta \ -\sin \theta \ 0; 0 \ \sin \theta \ \cos \theta \ 0; 0 \ 0 \ 0 \ 1]^T$$

로 표현할 수 있는데, 여기서 θ₁=cos θ, θ₂=sin θ 로

놓으면,

$$R_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \theta_1 & -\theta_2 & 0 \\ 0 & \theta_2 & \theta_1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

이 된다. 기준 좌표 상에서 현재 단계로의 이동 변화량을

$$T = [1 \ 0 \ 0 \ d_x; 0 \ 1 \ 0 \ d_y; 0 \ 0 \ 1 \ d_z; 0 \ 0 \ 0 \ 1]^T$$

로 나타낼 때, P_3 에 대한 3-D 이동 및 회전 변환 식은 아래의 식(9)로 나타낼 수 있다.

$$P'_3 = T \cdot T_0^{-1} \cdot R_0^{-1} \cdot R_p \cdot R_t \cdot R_0 \cdot T_0 \cdot P_3 \quad (9)$$

이동 및 회전 변환 식(P'_3)에서 미지수는 θ_1 과 θ_2 이고, 다른 변수들은 앞의 과정을 통하여 구해지는 값들이므로 P'_3 식을 전개하여 θ_1 과 θ_2 에 대한 계수 및 상수 형태로 놓으면

$$X_3' = p_1 \times \theta_1 + p_2 \times \theta_2 + p_c$$

$$Y_3' = q_1 \times \theta_1 + q_2 \times \theta_2 + q_c$$

$$Z_3' = g_1 \times \theta_1 + g_2 \times \theta_2 + g_c$$

가 되고, X_3' , Y_3' , Z_3' 을 식(6)에 대입하여 풀면

$$A_x = v_x \times f \times p_1 + g_1 \times v_x - g_1 \times x_3,$$

$$B_x = v_x \times f \times p_2 + g_2 \times v_x - g_2 \times x_3,$$

$$C_x = v_x \times f \times X_c - g_c \times (v_x - x_3) - v_x \times f \times p_c,$$

$$A_y = v_y \times f \times q_1 + g_1 \times v_y - g_1 \times y_3,$$

$$B_y = v_y \times f \times q_2 + g_2 \times v_y - g_2 \times y_3,$$

$$C_y = v_y \times f \times Y_c - g_c \times (v_y - y_3) - v_y \times f \times q_c$$

라 할 때

$$A_x \times \theta_1 + B_x \times \theta_2 = C_x, \quad A_y \times \theta_1 + B_y \times \theta_2 = C_y$$

형태가 되어, 위식에서 θ_1 과 θ_2 에 대한 방정식의 해를 구하면

$$\theta_2 = \frac{(C_x \times A_y - A_x \times C_y)}{(B_x \times A_y - A_x \times B_y)}$$

가 되므로, 구하고자하는 틸트 각 (θ)은 $\theta = \arcsin(\theta_2) \times 180/\pi$ 혹은 $\theta = \arccos(\theta_1) \times 180/\pi$ 로 얻을 수 있다.

실사 TV 카메라의 중심 위치는 P_0 가 이동한 위치 좌표 (P'_0)가 되므로 알고리즘 2를 통하여 얻은 X_1, Y_1, X_2, Y_2 및 Z값으로부터 구한다. P'_0 을 (X'_0, Y'_0, Z'_0)라 할 때, $X'_0 = (X_1 + X_2)/2$, $Y'_0 = (Y_1 + Y_2)/2$, $Z'_0 = Z^B + d_z$ 이 된다.

V. 실험 및 결과 분석

실험은 스튜디오 공간상에서 TV 카메라의 이동 및 파라메타의 변동시 추적 카메라에 캡처된 영상을 가지고 실행한다. 스튜디오는 그림 1과 같이 가로(600cm)×세로(400cm)×높이(450cm)인 직육면체로 구성하고 천장의 중앙에 실사 카메라를 추적하기 위한 CCD 카메라를 달았으며, 추적 영상으로부터 기하학적 정보를 추출하는데 이용되는 실사 카메라 부착용 LED는 그림 2에서처럼 1번과 2번의 간격을 40cm, 1, 2번의 중심과 3번의 거리는 50cm 및 높이는 카메라 중심 축으로부터 Z축 방향으로의 카메라 두께를 고려하여 20cm 정도 차이가 나도록 제작하였다. 높이를 모두 같게 할 수도 있는데 이때는 카메라 축과 LED위치까지의 높이를 보정하여 주면 같은 결과를 얻게된다. 기준 좌표가 될 카메라 위치는 천장과 1, 2번 중심(P_0)까지의 거리를 300cm로 하여 실사 기준 카메라 좌표는 $P_1(280, 200, 300)$, $P_2(320, 200, 300)$, $P_3(300, 150, 280)$ 일 때 $P_0(300, 200, 300)$ 가 되며 이때 팬 및 틸트 각은 0° 에 맞추어져 있다. 실험은 팬각, 틸트각, 붐 및 위치 이동이 동시에 일어나는 스튜디오 상황을 고려한 파라메타의 복합 변경 시를 대상으로 한다. 추적 카메라가 바라보는 3차원 스튜디오 공간상의 중심 좌표는 X_c, Y_c, Z_c 로 $P_c = (300, 200, 225)$ 가 되며, 추적 영상은 400×300 픽셀로 캡처 한다. 영상처리에 있어서 PSR의 중심좌표 값은 화소에 대한 서브 픽셀의 크기를 고려하여 소수점 4째 자리까지 이용하였다. 제안한 방법에서의 추정 값은 기준 카메라 파라메타로부터의 변화된 값이 되며, 가상카메라 보정 값은 전 단계의 추정 값과 현 단계 추정 값간의 차이 값이 된다. 정확도는 실사 카메라의 기준 값으로부터 각 단계까지의 실사 카메라에 적용한 파라메타 변화의 누적 량(R)과 제안한 방법으로 추정된 가상 카메라 파라메타 값(E)을 가지고 아래의 방법으로 계산한다.

P_i : 현 단계의 실사 카메라 파라메타 변화량
 C_i : 각 단계까지 실사 카메라에 적용된 카메라 파라메타 변화량의 누적 값
 R_i : 각 단계까지 실사 카메라에 적용된 카메라 파라메타 절대 변화량의 누적 값

n : 실험 회수

$$\text{정확도} = \left(1 - \frac{1}{n} \times \sum_{i=1}^n \frac{|C_i - E_i|}{R_i}\right) \times 100 \quad (10)$$

$$C_i = \sum_1^i P_i \quad R_i = \sum_1^i |P_i|$$

E_i : 각 단계에서 제안한 방법으로 추출한 파라메타 값(기준 값과의 차이)

실험 자료는 실사 카메라를 X축과 Y축으로 임의의 거리만큼 이동한 후, 팬과 틸트각 및 붐 값을 변경한 상태에서 추적 카메라를 통하여 캡처된 영상으로 단계별 파라메타 변경에 따른 6개의 추적 영상을 이용하였다.

그림 7은 각 단계에서의 카메라에 대한 추적 영상이고, 표 2는 카메라에 적용한 실제 파라메타 값과 제안한 방법

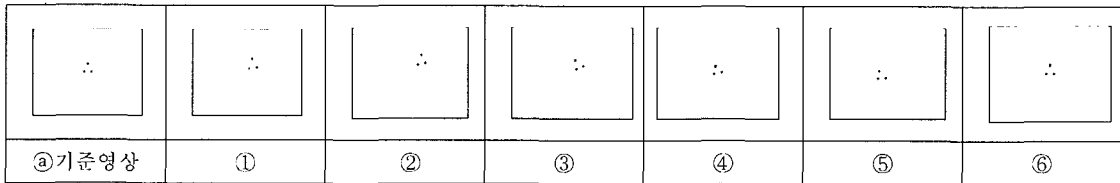


그림 7. 각 단계별 추적 영상
 Fig. 7. Sequence of the tracked image.

표 2. 복합 변화에 대한 파라메타 추정 (L: left R: right U: up D: down)
 Table 2. The parameter extraction for complex variation(L:left R:right U:up D:down)

번호	실사카메라의 파라메타 변동 값				제안한 방법으로 추정한 값 [가상 카메라 보정 값]				정확도
	pan	tilt	boom	이동	pan(°)	tilt(°)	boom(cm)	이동(cm)	
①	L10	U10	D15	(+30,-30)	-10.0327 [L 10.0327]	8.2410 [U 8.2410]	316.0 [D 16.0]	(+30.03, -30.10) [30.03, -30.10]	pan: (97.5%)
②	L5	U10	U20	(+30,-30)	-14.7425 [L 4.7098]	19.9648 [U 11.7238]	296.70 [U 19.30]	(+60.65, -60.35) [30.62, -30.25]	
③	L15	D20	U10	(-20,+20)	-29.9197 [L 15.1772]	-0.4284 [D 20.3832]	284.30 [U 12.4]	(+40.05, -39.62) [-20.60, 20.73]	tilt: (92.8%)
④	R15	U20	D10	(-40,+40)	-14.0171 [R 15.9026]	28.3153 [U 28.8437]	295.00 [D 10.7]	(0.00, 0.026) [-40.05, 39.65]	boom: (93.3%)
⑤	R5	D10	D10	(30,+30)	-9.8729 [R 4.1442]	8.5821 [D 19.7332]	307.0 [D 12.0]	(-30.26, +30.40) [30.26, 30.43]	이동: (94.2%)
⑥	R10	D10	U5	(+20,-20)	0.0389 [R 9.9118]	0.7780 [D 7.8041]	299.90 [U 7.1]	(-9.97, +10.04) [20.29, -10.36]	

표 3. 실사카메라 위치 추정(실제위치, [추정위치])
 Table 3. The position estimation of real TV camera(real position, [estimated position])

번호	P1	P2	P3	정확도
①	(3.1030, 1.7347, 3.1500) [3.1034, 1.7339, 3.1600]	(3.4970, 1.6653, 3.1500) [3.4973, 1.6642, 3.1600]	(3.2205, 1.2493, 2.8662) [3.2207, 1.2488, 2.8681]	X축: 94.15 %
②	(3.4068, 1.4518, 2.9500) [3.4456, 1.4300, 2.9180]	(3.7932, 1.3482, 2.9500) [3.7533, 1.3191, 2.9180]	(3.4961, 1.0122, 2.5911) [3.5034, 1.0978, 2.6290]	
③	(3.2268, 1.7000, 2.8500) [3.2523, 1.6708, 2.8430]	(3.5732, 1.5000, 2.8500) [3.6373, 1.4492, 2.8430]	(3.1500, 1.1670, 2.6500) [3.1500, 1.1684, 2.6464]	Y축: 92.65 %
④	(2.8003, 2.0535, 2.9500) [2.8003, 2.0535, 2.9500]	(3.1932, 1.9482, 2.9500) [3.1998, 1.9465, 2.9500]	(2.8961, 1.6122, 2.5911) [2.8990, 1.6231, 2.5188]	Z축: 85.69 %
⑤	(2.5030, 2.3347, 3.0500) [2.5004, 2.3383, 3.0700]	(2.8970, 2.2653, 3.0500) [2.8944, 2.2697, 3.0700]	(2.6205, 1.8493, 2.7662) [2.6180, 1.8478, 2.7982]	
⑥	(2.7000, 2.1000, 3.0000) [2.7001, 2.1003, 2.9990]	(3.1000, 2.1000, 3.0000) [3.0042, 2.1006, 2.9990]	(2.9000, 1.6000, 2.8000) [2.9007, 1.6022, 2.7918]	

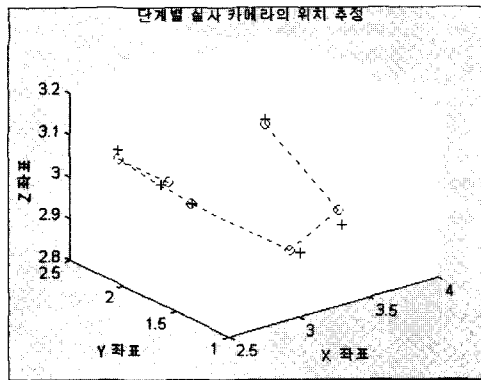


그림 8. 단계별 실사카메라의 위치와 제안된 방법으로 추정된 카메라 위치 비교

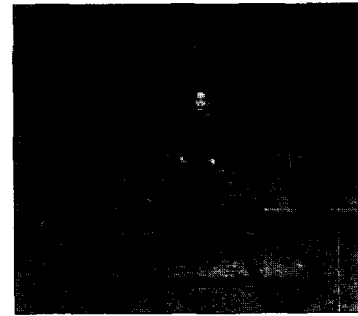
Fig. 8. Compare between the real camera position and the estimated camera position in each step.

을 통하여 추정된 파라메타 값의 비교를 보이고 있고, 제안한 방법에 대한 타당성 분석을 위해 앞에서 유도한 식 (10)을 이용하였다.

표 3과 그림 8은 실사 카메라의 좌표와 추정된 카메라 위치 좌표의 일치 정도를 나타내고 있으며 각 단계별 좌표 이동 량의 누계와 각 단계에서 실사카메라 좌표와 추정 카메라 좌표간의 오차를 누적한 값간의 비율을 가지고 카메라 위치 추정에 대한 정확도를 계산하여 보이고 있다.

그림 9는 가상 카메라 파라메타를 보정하지 않은 상태에서의 합성 영상과 추적 영상으로부터 카메라 파라메타를 추출하여 가상 카메라 파라메타를 보정한 상태에서 합성한 영상 예를 보이고 있다. (a)는 표 2와 3 및 그림 8의 4번째 단계인 번호④에 해당되는 것으로 기준 카메라 파라메타가 적용된 가정된 최초 합성영상과 같은 결과이고, (b)는 실사 TV 카메라가 전방으로 이동하여 FG는 줌인(zoom-in) 효과로 확대가 되었으나 BG 영상은 전 단계에 적용된 카메라 파라메타를 그대로 사용했기 때문에 동기화가 되지 않은 영상이며, (c)와 (d)는 실사 카메라에 대한 추적 영상으로부터 카메라 파라메타를 추출하여 BG 영상에 적용되었던 전 단계의 가상 카메라 파라메타를 보정하기 때문에 동기화가 이루어져 자연스러운 영상으로 연출된 합성영상이다. 그림 9의 (d)는 6번째 단계인 번호⑥에 해당된다.

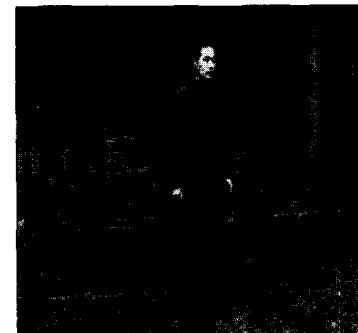
실험 결과를 분석해보면 영상처리를 할 때 서브 픽셀의 계산을 고려하여 정확한 PSR의 중심 좌표를 산출할 시는 90% 이상의 정확도를 갖는 파라메타 값을 추출할 수 있지만, 서브 픽셀의 계산을 고려하지 않거나 영상처리의



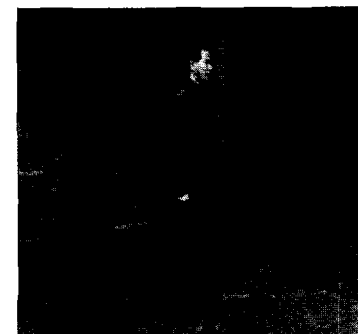
(a)기준카메라 위치에서 본 가정된 최초합성영상



(b)실사카메라가 전방 이동했으나 배경영상과 동기화 되지 않은 경우



(c)우측으로 팬이 발생했을 때 동기화된 경우



(d)우측 팬과 전방 이동의 동시 발생시 동기화된 경우

그림 9. 가상스튜디오 시스템에서의 합성 영상 예
Fig. 9 Examples of composed images in the virtual studio system

부정확성때문에 PSR의 중심에서 벗어난 좌표를 가지고 파라메타를 계산할 시는 특히 틸트 값의 추출에 문제를 일으킬 수 있었다. 제안한 시스템에서 실험한 결과는 평균적으로 팬각은 4.5/180도 이내, 틸트 각은 12.96/180도 이내, 붐 값은 6.7/100cm이내, 카메라 위치에 대한 좌표는 X축에 대하여는 5.85/100cm, Y축에 대하여는 7.25/100cm, Z축에 대하여는 14.31/100cm 이내의 오차 안에서의 정확도를 가지고 파라메타를 추출할 수 있었으며, 펜티엄 III 450Mhz 환경에서 영상처리와 변환 및 계산과정을 통하여 파라메타를 추출하는데 0.8초 정도의 시간이 소요됨으로 실시간 처리가 가능함을 볼 수 있었다.

VI. 결론 및 향후 연구 방향

본 논문에서는 가상 스튜디오 구현시 FG 영상과 BG 영상의 동기화에 필요한 카메라 파라메타를 컴퓨터 비전을 이용하여 추출하는 방법을 제안하였다. 스튜디오 천장에 부착된 추적 CCD 카메라를 통하여 입력되는 2-D 영상이 스튜디오상의 3-D 영역에 사상(mapping)되는 역투시 방정식과 기준 카메라 객체와 현 단계 실사 카메라 객체간의 3-D 회전 변환 식으로부터 각 파라메타(붐값, 팬각, 틸트각, 카메라위치)별로 추정 식을 유도하고, 복합 파라메타 변경시 제안한 방법을 통하여 얻은 파라메타 값과 실제로 적용된 카메라 파라메타 값을 비교하여 정확도를 제시하였다. 이 결과 영상처리시 서브 픽셀을 고려하여 PSR의 중심 좌표를 정확히 추출할 때에는 90% 이상의 정확도로 파라메타 값을 추출할 수 있었지만, 영상처리의 부정확성 때문에 PSR의 중심에서 벗어난 좌표를 가지고 파라메타를 계산할 시는 특히 틸트값의 추출에 문제를 일으킬 수 있었다. 또한 FG 영상과 BG 영상의 합성시 제안한 방법을 통하여 추출한 가상 파라메타의 보정량을 BG 영상에 적용한 경우와 적용하지 않은 경우의 비교를 통하여 제안한 방법에 타당성이 있음을 보였다. 그러므로 제안한 방법은 복잡한 부대시설이나 고가의 장비 없이 실시간으로 가상 카메라에 대한 파라메타를 추출할 수 있으므로 방송국의 가상스튜디오 구현이나 가상대학 구축 시 적용 가능한 방법이라 생각되며, 향후 줌 파라메타에 대한 정확한 추출 방법과 영상처리 시 서브픽셀에 대한 정확도를 향상시킬 수 있는 방법이 연구되어야 한다고 생각한다.

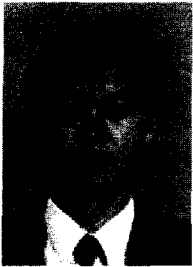
[참 고 문 헌]

- [1] Sim Gibbs, "Virtual Studios: An Overview," *IEEE Multimedia*, vol. 5, no. 1, Jan. 1998.
- [2] Peter Rander and Takeo Kanade, "Virtualized Reality: Constructing Virtual Worlds from Real Scenes," *IEEE Multimedia*, vol. 4, no. 1, Jan-Mar. 1997.
- [3] Andrew Wojdala, "Challenges of Virtual Set Technology," *IEEE Multimedia*, vol. 5, no. 1, Jan-Mar. 1998.
- [4] Masaki Hayashi, "Image Compositing Based on Virtual Cameras," *IEEE Multimedia*, vol. 5, no. 1, Jan-Mar. 1998.
- [5] M. V. SRINIVASAN, S. VENKATESH, and R. HOSIE, "Qualitative estimation of camera motion parameters from video sequences," *Pattern Recognition*, vol. 30, no 4, pp. 593 - 606, 1997.
- [6] P. J. Narayanan and Takeo Kanade, "Virtual World using Computer Vision," *Proceedings of the 1998 Workshop on Computer Vision for Virtual Reality Based Human Communications(CVVRHC '98)*.
- [7] G. A. Thomas, J. Jin, T. Niblett, and C. Urquhart, "A Versatile Camera Position Measurement System for Virtual Reality TV Production," *Conference Publication no. 447, IEEE, 1997 Internal Broadcasting Convention*, pp. 284-289, Sep. 1997.
- [8] Donard Hearn and M. Pauline Baker, *Computer Graphics C Version*, pp. 431-446, Prentice Hall, 1997.
- [9] Ramesh Jand, Rangachar Kasturi, and Brian G. Schunck, *Machine Vision*, pp. 313-316, McGRAW-HILL Internal Inc, 1995.
- [10] A. Murat Tekalp, *Digital Video Image Processing* pp. 19-35, Prentice Hall PTR, 1995.
- [11] Robert M. Haralick and Linda G. Shapiro, *Computer and Robot Vision VOLUME II*, pp. 47-61, Addison-Wesely Publishing Company, 1993.
- [12] Gerhard X. Ritter and Joseph N. Wilson, *Computer Vision Algorithms in Image Algebra*, pp. 125-141, CRC Press, 1996.
- [13] 박성우(포항공대), "가상스튜디오 구현을 위한 실시간 카메라 파라메타 변수 추출", *제10회 영상 처리 및 이해에 관한 워크샵 발표논문집*, pp. 357-362, 1998.
- [14] 고희동(한국과학 기술원 영상미디어 연구 센터), Virtual Studio 기술, *방송공학회지 제2권* pp. 224-234.

[15] 한기태, 장재영, 김이경, 김희율, "중성자 레디오 그래픽 영상 처리 시스템을 위한 왜곡 영상 보정 방법", *한국통신학회 하계종합학술발표회 논문집*, 1998.

[16] 한기태, 이광진, 김희율, "광각 CCD 카메라 렌즈를 위한 간단한 자동 보정 방법", *한국통신학회 추계종합학술발표회 논문집*, 1999.

저 자 소 개



한 기 태

1982. 2 : 충남대학교 (현)컴퓨터과학과 졸업(이학사)
 1990. 2 : 한양대학교 산업대학원 전자계산과 졸업(공학석사)
 1995. 8 ~ : 현재, 한양대학교 전자공학과 박사과정(수료)
 1986. 1 ~ : 1992.8. 한국주택은행 전산부 근무(전산역)
 1992. 9 ~ : 현재, 경원 전문대학 전자계산과(부교수)
 주관심분야 : 컴퓨터비전, 영상데이터베이스, 전자상거래 등



김 희 율

1976. ~ 1980. : 한양대학교 전자공학(공학사)
 1981. ~ 1983. : Pennsylvania State Univ. 전자공학(공학석사)
 1984. ~ 1989. : Purdue University 전자공학(공학박사)
 1989. 9 ~ 1994.2 : University of Texas 조교수
 1994. 3 ~ 1998.2 : 한양대학교 전자공학과 조교수
 1998. 3 ~ 현재 : 한양대학교 전자공학과 부교수
 주관심분야 : 컴퓨터비전, 영상검색, 영상 압축, 워터마크 등