

# 영상기반 실시간 가상광고 시스템

박혁래, 박종일  
한양대학교 전자전기컴퓨터공학부  
서울특별시 성동구 행당동 17번지.

## Real-Time Vision-Based Virtual Advertisement System

Hyuk-Rae Park and Jong-Il Park  
Hanyang University  
E-mail: kaien99@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

### 요약

최근 스포츠 중계를 중심으로 가상광고가 큰 주목을 받고 있다. 가상광고는 현실에 존재하지 않는 광고를 실영상 내에 삽입하여 마치 실제로 존재하는 물체처럼 보여주는 기술이다. 본 논문에서는 한 대의 카메라로 촬영한 영상으로부터 실시간으로 특징점 추출 및 추적을 행하여, 카메라 Projection 행렬을 복원하고, 이를 이용하여 범용적인 크로마키(chroma-key) 기법으로 영상합성 함으로써 센서와 같은 부가장비가 필요 없는 영상기반의 가상광고 시스템을 구현하였다. 일반 PC상에서 알고리즘의 유효성과 실시간 동작성을 확인하였다.

### 1. 서론

최근 영상처리기술과 컴퓨터의 성능의 급격한 발달로 영상처리의 실시간 처리가 가능하게 됨으로써 TV의 스포츠 중계 중에 실제 영상에는 없으나 화면상에만 보이는 광고가 등장하고 있다. 이를 가상광고(Virtual advertisement)라 하며 스포츠 중계를 중심으로 많이 이용되고 있다. 실 경기장에는 없는 로고 및 그래픽정보 등을 마치 경기장에 있는 것처럼 자연스럽게 합성하여 경기관전에 도움을 준다.

가상광고 시스템은 카메라(camera)의 움직임 추적(tracking) 할 수 있는 센서(sensor)를 이용 카메라의 움직임을 감지하여 가상광고의 위치를 고정하고 광고용 그래픽 정보를 크로마키(chroma-key)기법으로 영상합성하며, 2차원적인 광고뿐만 아니라 3차원적인 그래픽 정보도 실시간으로 합성이 가능하다[9].

카메라의 센서를 이용하면 카메라의 움직임 데이터인 회전과 이동정보를 쉽고 빠르게 전송할 수 있는 장점이 있지만 카메라에 부가적인 정밀한 센서를 부착해야 하는 단점을 가지고 있다. 또한 카메라의 움직임을 정밀하게 감지하기 위해서는 고가의 센서를 사용해야 한다. 이에 본 논문에서는 영상기반(vision-base)에서 3차원적인 위치정보를 알고 있는 고정되어 있는 4개 이상의 특징점을 추출 및 추적하여 점대점 매칭방식을 이용, 위치를 고정할 수 있는 카메라 Projection 행렬을 복원하여 고가의 카메라 센서를 사

용하지 않는 가상광고 시스템을 실현하고자 한다.

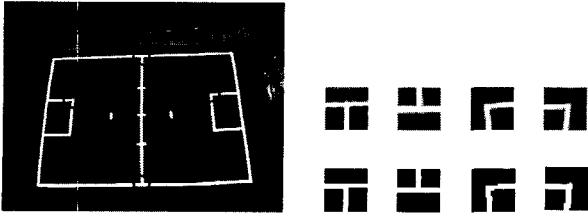
또한 특징점의 위치가 운동선수와 같은 물체에 가려지거나 사라져 특징점 위치정보가 최소 개수 이하로 떨어져 P 행렬 복원이 불가능 할 경우 임의의 점의 움직임 추정(global motion estimation)을 이용하여 P 행렬을 복원함으로써 점대점매칭방식의 단점을 보완하였으며, 영상합성시 범용적인 크로마키기법을 이용하여 합성해야 할 부분의 명도와 채도가 불균일한 부분에 대한 효과적인 영상합성을 구현하였다.

### 2. 특징점 추출 및 추적

#### 2.1 특징점 추출

특수한 마커를 이용하면 특징점의 위치를 쉽게 추출할 수 있지만 가상광고 시스템이 가정하고있는 운동경기장 같은 곳에는 마커를 위치시킬 수 없다. 이에 입력받은 실제영상에서 특징점을 추출하기 위해 경기장의 코너를 특징점으로 사용하였다. 코너는 국제규격에 의해 경기장의 위치가 일정하므로 실좌표계(world coordinate)상의 위치를 계산할 수 있다. 점대점매칭에 의해 카메라 P 행렬을 복원하기 위해서는 최소한 4개 이상의 정확한 특징점 위치정보가 필요하다[4].

제안방법의 성능을 확인하기 위하여 모형축구장을 만들어 실험을 수행하였다. 그림 1에 특징점을 추출하는 예를 나타내었다. 다이아몬드형태로 4개의 특징점의 위치를 추출하였다. 코너는 경기장의 선이 수직선과 수평선이 만나는 점의 형태로 되어 있는 특징이 있기 때문에, 특징점이 되는 코너가 있는 일정영역을 지정하고 기하학적인 정보를 추출할 수 있는 Hough 변환을 이용, 대표되는 두 개의 직선을 구하고 두 직선의 교점을 특징점 위치로 지정하여 정확한 특징점 위치를 추출해 낸다.



(a)

(b)

그림 1. 특징점 모양 (a)모형축구장의 코너위치,

(b)코너를 확대한 영상(상)과 2진 영상(하)

## 2.2 특징점 추적

특징점이 추출된 다음에는 이전 프레임과의 유사성에 의거하여 특징점을 추적하는 Lucas & Kanade 방식 [8]을 이용하여 특징점을 추적하게 된다. 그러나, 특징점들이 경기장의 선수들에 의해 가려지거나 화면밖으로 사라지는 경우 특징점의 추적이 불가능하거나 잘못된 특징점위치를 추적하게 될 가능성이 생긴다.

이에 초기에 특징점 위치의 일정영역의 색상테이블을 만들어 선수들이 특징점에 근접할 경우 추적을 중단하고 다른 점들에 의해 복원된 카메라 P 행렬을 이용하여 위치정보를 저장하고 있다가 가려졌던 특징점이 다시 나타날 경우 저장된 위치정보를 이용하여 다시 추적함으로써 특징점의 추적시에 나타나는 문제점을 해결하고 있다.

## 3. Projection 행렬 복원

### 3.1 점대점매칭방식

카메라 P(Projection)행렬은 현실세계를 가상공간상에 일치시킬 수 있는 상관관계이다. 이 P 행렬은  $(x, y, z)'$ 가 실좌표계(world coordinate)이고  $(x', y', z')$ 가 영상좌표계(pixel coordinate)이면  $3 \times 4$ 의 U 와 같은 행렬형태로 나타낼 수 있다[2].

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = U \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ u_{21} & u_{22} & u_{23} & u_{24} \\ u_{31} & u_{32} & u_{33} & u_{34} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (1)$$

$$X = \frac{x'}{z'}, \quad Y = \frac{y'}{z'}$$

여기서, F는 초점거리,  $(X, Y)'$ 는 이미지배열상의 좌표계이다. P 행렬을  $u_1, u_2, u_3, u_4$  단위벡터형태로 나타낼 수 있는데, 실좌표계의 좌표가 z축이 0인 위치가  $P_i(x_i, y_i, 0)'$  ( $i=0, \dots, 3$ )이면  $u_3$  성분은 제거되고 (2)와 같이 9개의 성분으로 줄일 수 있다.

$$\begin{pmatrix} x'_i \\ y'_i \\ z'_i \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & u_{14} \\ u_{22} & u_{22} & u_{24} \\ u_{31} & u_{32} & u_{34} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \quad (2)$$

$$X_i = \frac{x'_i}{z'_i} F, \quad Y_i = \frac{y'_i}{z'_i} F$$

카메라의 F를 고려하지 않으면 (3)과 같은 V(model view matrix)는 실좌표계의 4개의 좌표와 영상좌표계의 점대점매칭하여 V 행렬을 복원할 수 있다[4-6].

$$\begin{pmatrix} x'_i \\ y'_i \\ z'_i \end{pmatrix} = V \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \quad (3)$$

$$X_i = \frac{x'_i}{z'_i}, \quad Y_i = \frac{y'_i}{z'_i}$$

이는 (3)을  $v_{33}$ 을 1로 가정하고 8개의 변수 v에 대하여 정리하면  $8 \times 8$ 형태의 행렬을 구할 수 있는데 이것의 역행렬과 영상좌표계의  $8 \times 1$  행렬을 곱하여 V 행렬을 복원할 수 있다[4]. 그러나, F 성분을 무시하였기 때문에 3차원적인 영상복원을 할 수 없으므로, (2)와 (3)에 의해 F와  $u_1$ 과  $u_2$ 의 수직관계에 의해  $u_3$  성분을 복원하여 (1)의 P 행렬을 복원한다. 이러한 방식을 이용하면 같은 평면상의 위치정보만으로도 P 행렬을 복원하여 3차원적인 표현이 가능하게 된다[6].

초기에만 이러한 방식을 이용 P 행렬을 복원하고, 나머지 6개점의 위치를 추정하여 10개점의 위치정보가 획득되면 매 프레임마다 특징점 추적을 통해 나온 N개( $N=10-n$ , n:특징점 추적중에 사라지거나 가려지는 특징점의 개수)의 위치정보로 최소제곱방식을 이용하여 P 행렬을 복원함으로써 오차율을 최소화 할 수 있으며 또한 특징점이 선수에 의해 가려지거나 화면밖으로 사라지는 경우의 문제점을 해결할 수 있다.

### 3.2 GME(global motion estimation) 방식

점대점매칭방식을 사용하여 P 행렬을 복원할 경우 N이 4 이하로 떨어질 경우 P 행렬 복원을 할 수 없는 문제점이 발생한다. 이에 경기장 주위의 임의의 점들의 움직임 정보를 이용하여 P 행렬을 복원하였다.

GME는 카메라가 이동이 없고 줌과 회전만 있다는 가정하에서 N이 4 이하로 떨어지기 직전의 P 행렬에서 회전과 이동성분을 분리하고, 움직임 정보를 이용하여 카메라의 회전변화량 정보를 추출[9]하고 분리된 P 행렬의 회전성분에 곱하여 카메라의 움직임을 추정하여 P 행렬을 복원하는 방식이다.

먼저 카메라 P 행렬은 (4)와 같은 형태로 나타낼 수 있다[2].

APnD

$$= \begin{pmatrix} F & 0 & Cx \\ 0 & F & Cy \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & x_t \\ r_{21} & r_{22} & r_{23} & y_t \\ r_{31} & r_{32} & r_{33} & z_t \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

$$= \begin{pmatrix} r_{11}F + r_{31}Cx & r_{12}F + r_{32}Cx & r_{13}F + r_{33}Cx & x_tF + z_tCx \\ r_{21}F + r_{31}Cy & r_{22}F + r_{32}Cy & r_{23}F + r_{33}Cy & y_tF + z_tCy \\ r_{31} & r_{32} & r_{33} & z_t \end{pmatrix}$$

여기서 A는 카메라 intrinsic 행렬, D는 3D Euclidean 변환행렬, Pn는 정규화된 카메라의 P 행렬이다.

F는 점대점매칭을 통하여 구하고, Cx, Cy를 영상의 중심이라 했을 때, 회전과 이동성분을 가지고 있는 D 행렬은 (4)와 이전 프레임에서 복원된 P 행렬과 비교하여 분리한다. 카메라의 움직임을 f(zoom),  $\alpha$ (tilt),  $\beta$ (pan),  $\gamma$ (rotation)라 하면 3D 회전행렬은  $R=R_\alpha R_\beta R_\gamma$ 와 같이 나타낼 수 있다. 카메라움직임의  $\gamma$ 가 작고, F가 상당히 큰 경우 실좌표와 영상좌표는 (5)와 같은 상관관계를 가지게 된다[3].

$$\begin{aligned} x'' &= f(x' + \gamma y' - \beta F) = a_1 x' + a_2 y' + a_3, \\ y'' &= f(-\gamma x' + y' + \alpha F) = -a_2 x' + a_1 y' + a_4 \end{aligned} \quad (5)$$

(5)의 관계를 이용 임의의 점들에 대한 현재 프레임과 이전 프레임간의 ( $x''$ ,  $y''$ )의 위치변화에 대한 최소제곱법으로 (6)에 의해  $a_1, a_2, a_3, a_4$ 를 구하고 (5)에 의해 변화량  $\alpha, \beta, \gamma$ 를 구하여 분리된 3D 회전행렬에 곱하여 P 행렬을 복원한다.

$$\min_a \sum_{i=1}^m \left[ (x_i'' - (a_1 x_i' + a_2 y_i' + a_3))^2 + (y_i'' - (-a_2 x_i' + a_1 y_i' + a_4))^2 \right] \quad (6)$$

#### 4. 범용적인 크로마키기법

전통적인 크로마키 기법은 배경앞부분(foreground)의 영상을 파란배경이나 그린배경에서 촬영한다. 그리고, 단순히 배경이 되는 푸른색과 그린색의 색상정보를 이용하여 색을 제거하고 배경이 되는 영상을 대신 삽입하여 영상합성을 하는 기법이다.

그러나, 가상광고에 쓰이는 크로마키방식은 야외에서 촬영되고 제거할 색상이 한정되어 있지 않고 또한 채도와 명도가 일정하지 않기 때문에 적용하기 어렵다.[9] 이에 범용적인 크로마키기법을 제안하였다. 범용적인 크로마키기법의 알고리즘은 제거해야 할 부분 일정영역의 색상정보(RGB)를 입력받고 색상크기로 나열한 후 색상정보를 미디언필터(median filter)처리하여 실제영상에서 들어오는 임펄스 잡음성분을 제거한다. 제거 후 YCbCr 형태로 변환(conversion)한다[1].

YCbCr 변환 후 다시 기울기(Cb/Cr) 순서대로 나열하

여 최소값(Amin)과 최대값(Amax)을 구하고, 최소값과 최대값의 이용 평균기울기(Am)와 XZ 변환에 필요한  $\theta$ 를 구한다. 최대기울기에서 평균기울기의 차로 A를 구하고 YCbCr 성분을 다시 원점에서의 거리 순서대로 나열하여 D를 구한다.  $\theta$ 를 이용한 XZ 축으로의 변환 후 A가 0보다 클 때,  $D < X$ ,  $-AX < Z < AX$  두 조건을 만족하는 XZ 영역으로 키색상(key colour)을 인식하게 된다. 이렇게 인식된 키색상영역으로 실제영상을 YCbCr 형태로 변환하여 픽셀(pixel)단위로 비교하면서 그림 2와 같은 XZ 영역의 실선부분에 있으면 배경으로 인식하여 광고용영상의 색상정보를 입력하여 영상합성을 수행한다. 이와 같은 크로마키기법 이용하여 영상을 합성하면 특정색 부분에 대해서만 광고영상이 삽입되므로 배경에 있는 선수와 같은 물체가 광고영상 영역에 들어오면 물체에 대해 영상합성을 수행하지 않으므로써 광고가 경기장의 바닥에 존재하는 느낌을 줄 수 있다.

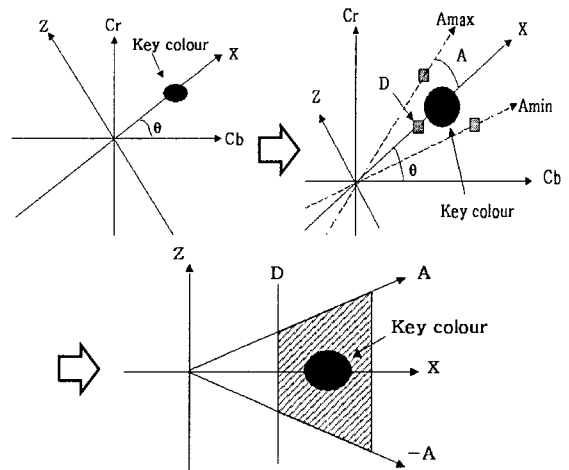


그림 2. 키색상 인식하는 과정

표 1. 실험환경

하드웨어	CPU: Intel Pentium-III 550 RAM: 256MB Grab board: Matrox Meteor II Camera: Canon MV1 디지털카메라
소프트웨어	Visual C++ 6.0 IPL(Image Processing Lib.), OpenCV
영상	320×240(pixel)

#### 5. 실험 및 결과

실시간 가상광고 시스템의 실험은 표1과 같은 실험환경을 사용하였다. 합성할 실제영상을 카메라로부터 640×480의 동영상을 PC상에서 Grab board를 이용 320×240으로 입력받고[7] 광고용 영상으로 정지영상(still image)을 입력받아 가상광고를 수행하였다.

초기화 과정으로 동영상의 정지영상을 만들고 정지영상에 마우스를 이용 특징점의 대략위치를 순서대로

4개 지정하여 특징점을 추출하고 나머지 6개의 특징점들은 P행렬에 의해 자동으로 위치가 생성되고 추적된다. 추적된 특징점의 위치정보를 이용 P행렬을 복원하여 광고용 영상의 위치고정과 기하학적인 변형을 위해 워핑(warping)[8]을 수행하여 주변과 자연스러운 영상합성을 구현하였다.

실험결과, 그림 3처럼 카메라의 pan(좌우움직임)에 대해서도 광고영상이 위치할 곳(모형축구장의 센터)에 벗어남이 없이 고정되어 있으며 기하학적인 변형과 중심선이 보임으로써 워핑과 크로마키에 의한 영상합성이 구현됐음을 확인 할 수 있다. 3차원적인 광고영상에 대해 축구장의 센터에 육면체를 삽입하여 3차원영상 복원을 구현하였다. 일반 PC상에서 초당 16프레임 이상의 가상광고가 가능하였다.

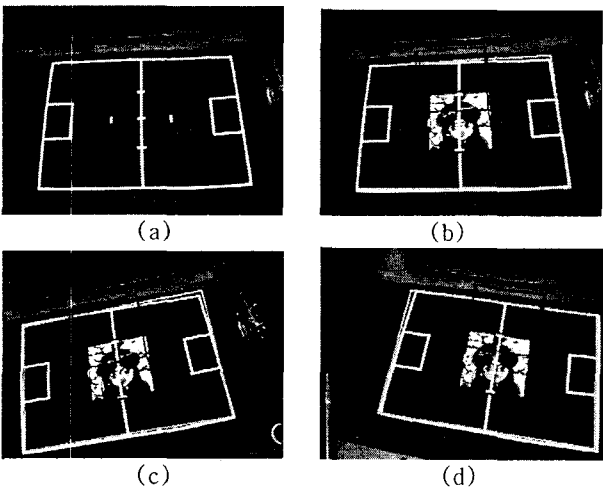


그림 3. 실험결과 (a)합성용 실제영상 (b)육각형의 3차원 광고결과와 2차원영상의 워핑된 결과 (c)(d) 카메라를 pan 했을 때의 가상광고결과

## 6. 결론

본 논문에서는 영상기반 가상광고 시스템을 개발하였다. 먼저 가상광고용 요소기술인 특징점 추적과 추출, 카메라 Projection 행렬 복원, 범용적인 크로마키기법에 의한 영상합성을 이용 전체시스템에 대한 실험을 통하여 영상기반 가상광고 시스템용 알고리즘을 개발 및 입증하였다. 기존의 센서기반의 가상광고시스템과 달리 부가적인 카메라 센서를 제거했으며 영상기반으로 인한 특징점이 물체에 가려지는 문제점을 최소제곱법에 의한 점대점매칭방식을 이용하여 해결했으며 특징점이 4개 이하로 떨어질 경우에 대해서는 GME방식을 이용하여 해결하였다. 또한 범용적인 크로마키기법을 이용하여 다양한 배경화면에 대해 간단하면서도 고속의 우수한 영상합성결과를 얻을 수 있었다.

## 참고 문헌

- [1] C. P. Sandbank Eds. , *Digital Television*, John Wiley & Sons, pp.539-557, 1990.
- [2] Gang Xu, Zhengyou Zhang, *Epipolar Geometry in Stereo, Motion and Object Recognition*, Kluwer Academic Publishers, 1996.
- [3] J. Park, N. Yagi, "Estimation of camera parameters from image sequence for model-based video coding." IEEE Trans. Circuit and systems for video technology, vol.4, No.3, pp288-296, 1994.
- [4] T. Okuma, K. Kiyokawa, H. Takamura and N. Yokoya, "An augmented reality system using a real-time vision based registration." Proc. ICPR 1998, pp.1226-1229, 1998.
- [5] T. Okuma, K. Sakaue, H. Takemura, N. Yokoya, "Real-time camera parameter estimation from images for a mixed reality system." Proc. ICPR 2000, Vol.4, pp482-486, 2000.
- [6] Y. Nakazawa, S. Nakano, T. Komatsu and T. Saitou, "A system for composition fo real moving images and CG images based on image feature points." The Journal of the institute of image information and television engineers, Vol.51, No.7, pp1086-1095, 1997. (in Japanese.)
- [7] "Matrox Imaging Library Version 6.0, User Guide & Command Reference" , Matrox Electronic Systems Ltd., 1999.
- [8] "Intel Image Processing Library Reference Manual", Intel Ltd..
- [9] <http://www.orad.tv/index.htm>