

특집논문-02-07-1-01

색상패턴 추적을 이용한 실시간 증강영상 시스템

박성춘*, 남승진*, 오주현*, 박창섭*

A Real-time Augmented Video System using Chroma-Pattern Tracking

Sung-Choon Park*, Seung-Jin Nam*, Ju-Hyun Oh*, and Chang-Seob Park*

요 약

최근에 TV 방송에서 가상스튜디오나 가상캐릭터와 같은 가상현실(VR: Virtual Reality) 기술이 자주 사용되고 있으며, 증강현실(AR: Augmented Reality) 기술에 대한 관심도 높아지고 있다. 본 논문에서는 증강현실 기술을 방송에 응용한 가상스크린 시스템에 대해 소개한다. 가상스크린 시스템은 움직이는 색상패턴 패널을 추적하여 실시간으로 그 위에 동영상상을 합성하는 증강영상 시스템이다. KBS 기술연구소에서는 가상스크린 시스템을 개발하고 'K-비전'이라 이름지었다. 이 시스템은 사용자가 들고 움직이는 패널에 동영상이나 그래픽 영상 등을 보여줄 수 있는데, 보여지는 모든 영상은 카메라의 움직임과 패널의 움직임에 따라 정확하게 입혀진다. 패널 추적을 위하여 블럽 분석(blob analysis)이나 특징 추적(feature tracking)과 같은 영상처리 기술을 이용한다. K-비전은 모든 타입의 카메라와 사용 가능하며, 특별한 부가장치가 필요하지 않다. 센서를 부착하지 않아도 되고, 캘리브레이션(calibration) 과정 또한 필요하지 않다. K-비전은 선거개표 방송, 다큐멘터리, 오락 프로그램 등 생방송 프로그램에서 활용한다.

Abstract

Recently, VR(Virtual Reality) applications such as virtual studio and virtual character are widely used in TV programs, and AR(Augmented Reality) applications are also being taken an interest increasingly. This paper introduces a virtual screen system, which is a new AR application for broadcasting. The virtual screen system is a real-time video augmentation system by tracking a chroma-patterned moving panel. We have recently developed a virtual screen system, 'K-vision'. Our system enables the user to hold and move a simple panel on which live video, pictures or 3D graphics images can appear. All the images seen on the panel change in the correct perspective, according to movements of the camera and the user holding the panel, in real-time. For the purpose of tracking panel, we use some computer vision techniques such as blob analysis and feature tracking. K-vision can work well with any type of camera, requiring no special add-ons. And no need for sensor attachments to the panel, no calibration procedures required. We are using K-vision in some TV programs such as election, documentary and entertainment.

I. 서 론

지난 수년간 디지털 영상 분야는 놀라운 기술적 발전을 거듭하여 왔다. 특히, 최근의 영화에서 볼 수 있듯이 컴퓨터 그래픽과 영상합성 기술의 발전은 실제 영상과 가상의 그래픽 영상을 구분할 수 없을 정도로 정교한 합성 영상을 제작할 수 있게 하였다.

TV 방송에서도 1990년대 중반부터 3차원 컴퓨터 그래픽을 이용한 가상의 세계를 선보이기 시작하였다. 영화제작에서와 같은 포스트 프로덕션(post-production) 방식으로 사전 제작하여 사용하기도 하였으나, 선거개표 방송이나 스포츠 중계에서 실시간으로 그래픽을 생성하여 실제 영상과 합성하여 보여주는 기술들이 등장하였다. 실시간 합성 기술의 대표적인 예로 가상스튜디오(virtual studio), 가상캐릭터(virtual character), 가상이미징(virtual imaging) 등을 들 수 있다.

이 기술들은 컴퓨터 그래픽스, 특히 가상현실(VR: Virtual

* 한국방송 기술연구소
Technical Research Institute Korea Broadcasting System

Reality) 기술을 기반으로 하고 있다. 가상현실은 3차원 그래픽을 활용하여 사용자에게 현실감 있는 가상의 세계를 제공하는 기술로 각종 센서를 이용하기도 한다. 가상현실의 변형으로 증강현실(AR: Augmented Reality)이 있는데, 가상의 그래픽을 실제 환경에 합성하여 보여줌으로써, 실제 환경에 대한 이해를 돕는 것이 목적이다^[1]. 가상현실은 가상의 그래픽이 주인데 비해서 증강현실은 실제 환경이 주가 되고 가상의 그래픽이 일부 첨가된다는 차이가 있다.

TV 방송에서 그래픽과 실제 영상을 실시간으로 합성하기 위하여 일반적으로 크로마키 장비나 가상스튜디오를 사용하고 있는데, 일기예보와 같이 크로마키를 사용하는 방법은 배경에만 그래픽을 사용할 수 있고, 카메라를 움직이는 경우에는 진행자 화면과 배경 그래픽 화면이 분리되어 현실감이 없는 단점이 있다. 가상스튜디오의 경우에는 특수한 설비를 갖춘 스튜디오에서만 제작이 가능하다는 불편함이 따르며, 카메라의 움직임에는 자연스러운 연동이 가능하나, 움직이는 실제 물체에는 그래픽을 연동시켜 표시할 수 없다.

이러한 단점을 보완하기 위하여 움직이는 물체에 그래픽을 연동시켜 합성할 수 있는 시스템을 연구 개발하고, 'K-비전'이라 이름지었다. 이 시스템은 움직이는 색상패턴 패널을 추적하여 실시간으로 그 위에 동영상상을 합성하는 증강영상 시스템으로 가상스크린(virtual screen), 가상윈도우(virtual window), 가상패널(virtual panel), 가상빌보드(virtual billboard) 등으로 불린다. 외국의 상용 시스템으로는 프랑스 Symah Vision사의 EPSIS^[2], 이스라엘 Orad사의 MobileSet^[3] 등이 있다.

가상스크린 시스템은 사용자가 들고 움직이는 패널에 동영상이나 그래픽 영상 등을 합성하여 보여줄 수 있는데, 패

널에 보여지는 모든 영상은 카메라의 움직임과 패널의 움직임에 따라 정확하게 입혀진다. 그림 1에 가상스크린 시스템 K-비전의 개념을 보이고 있다.

본 논문의 구성은 다음과 같다. II장에서 시스템 구성에 대해 살펴보고 III장에서 색상패턴 패널의 추적 알고리즘에 대해 설명하며, IV장에서 운용 소프트웨어를 소개하고 V장에서 방송활용 사례를 살펴본다. 마지막으로 VI장에서 결론을 맺는다.

II. 시스템 구성

K-비전 시스템은 외관상 PC를 기반으로 하는 본체와 프레임 딜레이(frame delay), 크로마키어(chromakeyer)로 구성된다. 본 장비는 입출력이 디지털이기 때문에 아날로그 카메라, 아날로그 VCR, 아날로그 스위처 등을 사용하는 스튜디오에서는 A/D 및 D/A 변환 장치가 별도로 필요하다. 그림 2에 시스템 구성을 보이고 있다.

본체는 윈도우NT를 탑재한 PC를 기반으로 구성되는데, 2개의 펜티엄 III CPU와 512MB의 메인메모리를 장착하고 있으며, 내부에는 그림 2에서와 같이 보드 타입의 모듈들이 내장된다.

첫 번째는 카메라 영상을 실시간으로 획득하여 이진영상화(binazation)와 블럽분석(blob analysis, 색상과 밝기가 유사한 영역의 특성분석) 등의 영상처리를 수행하는 영상 획득 및 처리 모듈이다. 두 번째 모듈인 디지털 영상효과 모듈과 세 번째 모듈인 영상 입출력 및 합성 모듈은 서로 연결되어 있으며, 삽입될 영상의 모양을 영상처리 결과에 따라 변형시킨다. 스크린 패널을 추적하는 알고리즘은 다음장에서 설명하기로 한다.

삽입될 영상은 VCR, 문자발생기, 카메라 등과 같이 외부로부터 입력되기도 하고, 시스템 하드디스크에 저장된 내부의 동영상도 가능하다. 동영상은 M-JPEG 포맷으로 압축되며, 압축 오디오도 포함한다. 압축률은 아날로그 방송 화질은 4:1(5MB/sec), 디지털 방송 화질은 2.84:1(7MB/sec) 정도이다. 따라서, 18GB 용량의 하드디스크에는 40분내지 1시간 분량의 동영상이 저장될 수 있다. 동영상의 압축 및 재생은 영상 입출력 및 합성 모듈에서 수행된다.

영상분석과 변환 등의 과정에서 시간적인 지연이 생기기 때문에, 원래의 카메라 영상은 프레임 딜레이를 사용하여 지연시켜야 하며, 이에 따라 카메라 영상과 관련된 오디오도 함께 지연시켜야 한다.

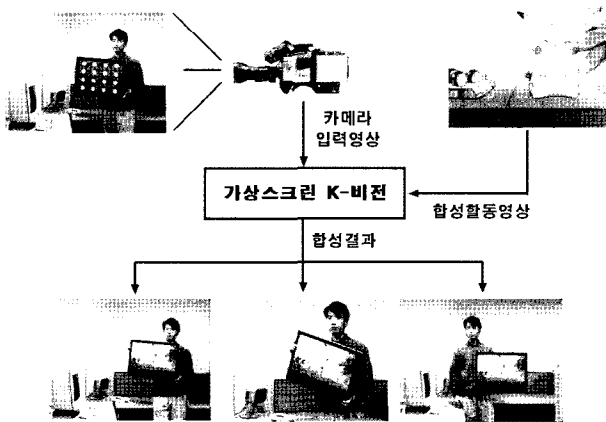


그림 1. 가상스크린 시스템 K-비전의 개념
Fig. 1. The concept of a virtual screen system, K-vision

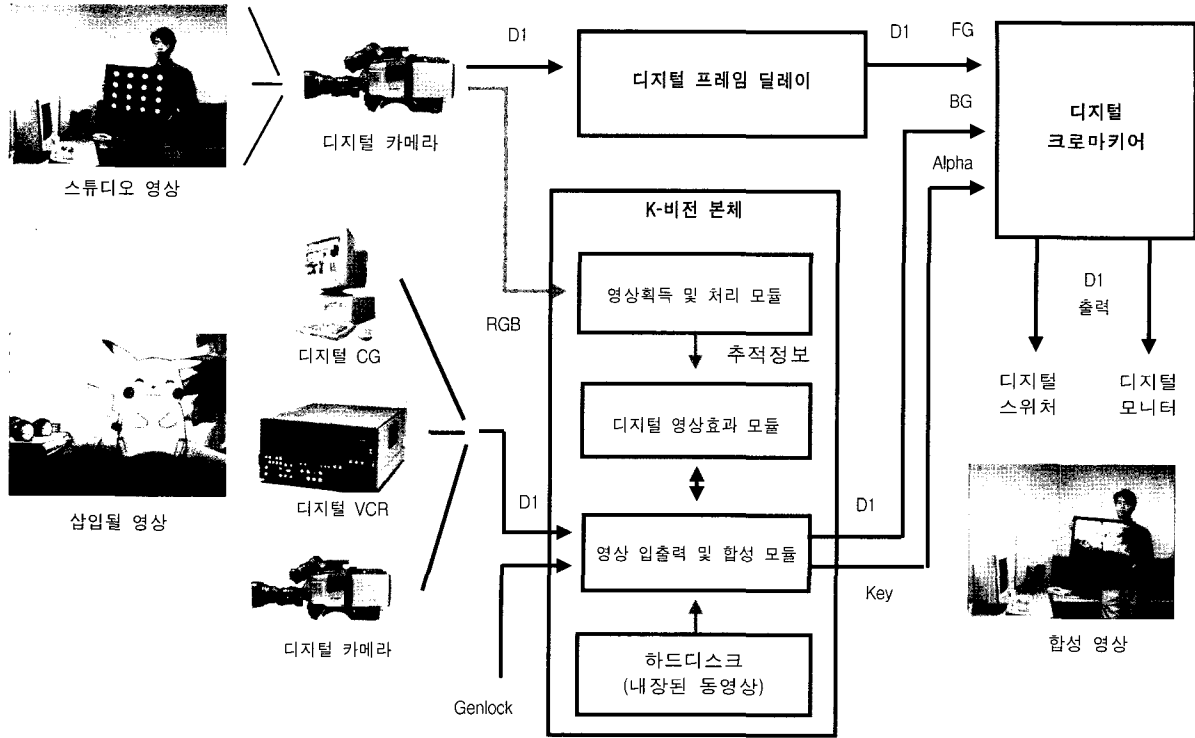


그림 2. K-비전 시스템 구성
Fig. 2. The configuration of K-vision system

이렇게 지연된 카메라 영상은 변형된 삽입 영상과 크로마키어를 통해서 합성된다. 이 때, 스크린과 비슷한 색상을 갖는 바깥 부분의 영상을 보전하기 위해서 변형된 삽입 영상에 포함된 알파키(alpha key) 정보를 사용한다.

III. 색상패턴 패널 추적 알고리즘

그림 3은 카메라로부터 획득한 영상으로부터 색상패턴 패널의 모양을 인식하는 과정을 나타낸 것이다. 실시간으로 패널의 움직임을 인식하고, 크로마키(chromakey)로 합성하기 위하여 파란색이나 녹색 기반의 원형 패턴을 갖는 패널을 사용한다.

크로마키를 사용하는 이유는 패널이 사람이나 다른 물체에 의해 가려지더라도 합성되는 동영상에 사람이나 물체 뒤의 패널에만 보이게 하기 위하여 사용한다. 또한 패널의 일부가 가려지더라도 추적을 계속할 수 있도록 16개의 원형 특징점을 일정한 간격으로 배치하였다.

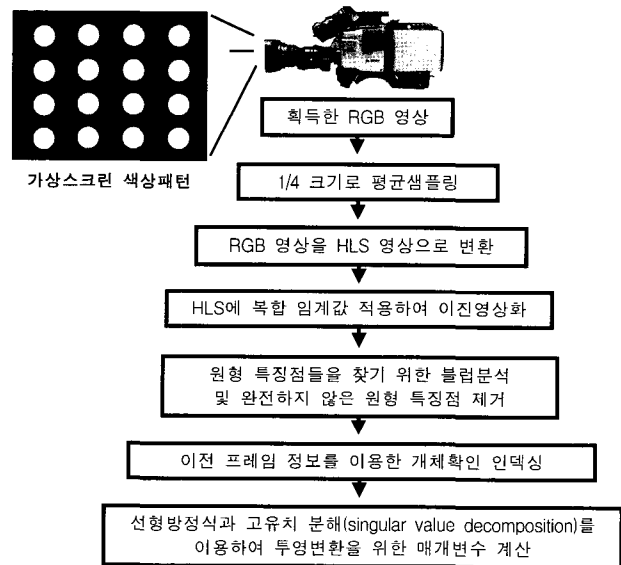


그림 3. 색상패턴 패널 추적 알고리즘
Fig. 3. The algorithm for tracking panel

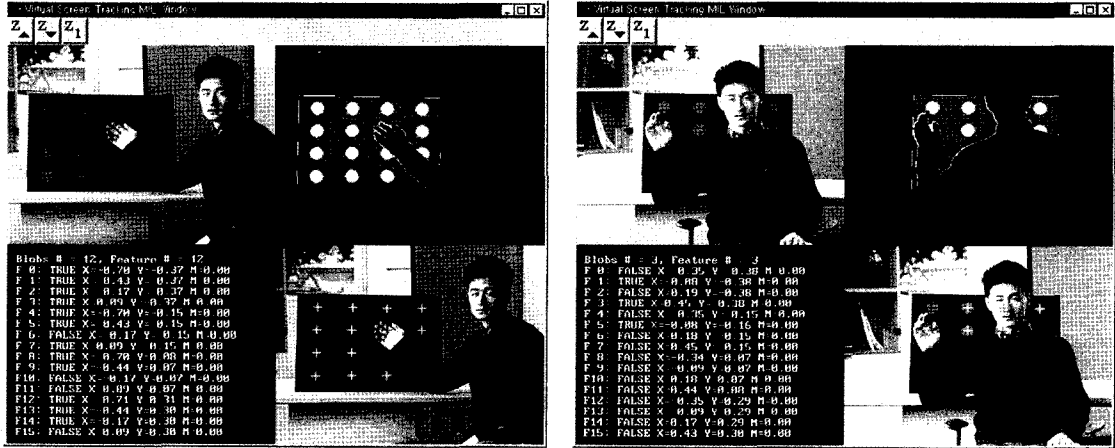


그림 4. 블럽분석을 이용한 특징점 위치 결정
Fig. 4. Determination of the feature positions using blob analysis

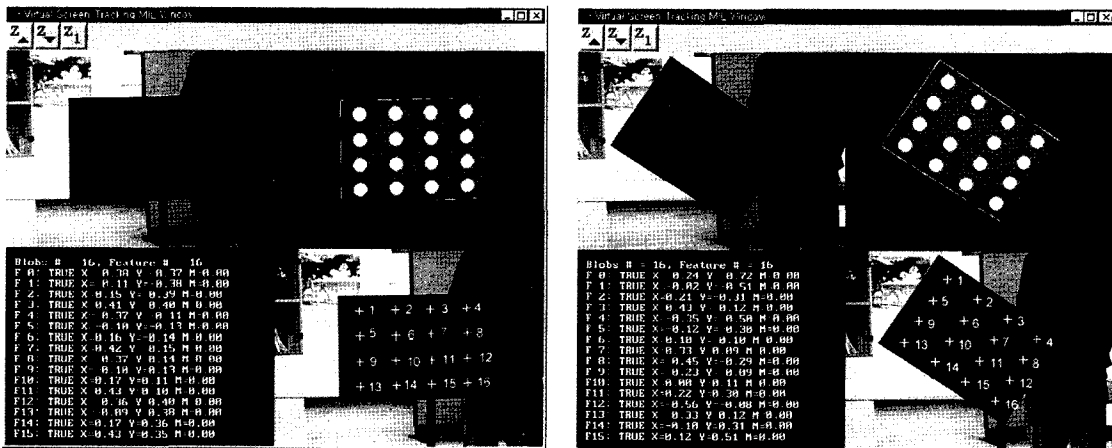


그림 5. 위치추정을 이용한 특징점 개체확인 및 인덱싱
Fig. 5. Identification and indexing of the features using position estimation

패널의 크기는 제한이 없으나, 가로 세로의 비율은 일반적인 화면 비율인 4:3 에서 16:9 사이로 하는 것이 적당하다.

카메라를 통해 입력된 영상은 실시간 처리를 위하여 1/4 크기로 축소되며, 이 과정에서 평균샘플링(average sampling)을 하여 정보의 손실을 줄임으로써 정확한 추적이 가능하도록 하였다.

축소된 영상은 RGB(Red, Green, Blue) 영상인데, 정확하고 안정적인 특징점의 인식을 위하여 HLS(Hue, Lightness, Saturation) 영상으로 변환하여 사용하였다. 변환된 HLS 영상에 적절한 임계값을 적용하여 이진 영상으로 만든다.

이진 영상에 블럽분석을 적용하여 원형 특징점 후보들을 찾는다. 영상 잡음에 의한 것과 일부가 가려진 원형 특징점들

을 제거하기 위하여 블럽 영역이 다른 블럽에 비해 상대적으로 아주 작거나 원형의 성질이 적은 블럽들은 제외시킨다.

이렇게 찾아진 원의 중심점을 구하여 특징점의 위치로 사용하는데, 그림 4에 이 방법으로 특징점의 위치를 찾은 결과를 두 가지 보이고 있다. 각 그림에서 좌측 상단 부분이 입력 영상이고, 우측 상단 부분은 패널과 원형 패턴의 이진영상을 조합하여 표시한 것이며, 우측 하단이 블럽분석에 의해 찾아진 원형 블럽의 중심 위치를 표시한 것이다. 이 그림에서 밝은(노란색) 십자 표시 부분이 제대로 찾아진 원형 블럽의 중심 위치로서 다음 단계에서 사용된다.

다음 단계는 어느 블럽이 어떤 특징점에 해당하는지 개체확인(identification)을 하여야 한다. 개체확인을 위하여

이전 프레임들의 분석 정보를 이용한다. 즉, 이전 프레임에서의 특징점들의 위치와 속도 등을 가지고 현재 프레임의 특징점들의 위치를 예측한 후 가장 가까운 블럽을 그 특징점으로 인덱싱(indexing)한다. 카메라와 패널이 등가속도 운동을 한다고 가정하여 식 (1)에 따라 위치 예측을 한다.

$$X_{i+1} = 3X_i - 3X_{i-1} + X_{i-2} \quad (1)$$

X_{i+1} : (i+1) 프레임에서의 예측 위치
 X_i, X_{i-1}, X_{i-2} : i, (i-1), (i-2) 프레임에서의 위치

그림 5에 위치 추정에 의한 특징점들의 개체확인 인덱싱 결과의 한 가지 예를 보이고 있는데, 원형 패턴 옆에 각 고유번호를 표시하였다.

특징점의 위치로부터 투영변환(perspective transformation) 매개변수를 얻기 위하여 선형 방정식을 적용한다. 변환 매개변수는 4x4 행렬이지만, 실제로 필요한 것은 8개뿐이다. 따라서, 4개의 특징점의 위치 (x, y)로부터 변환 매개변수를 구할 수 있다. 이 경우 4개의 점은 동일 선상의 점이 아니어야 한다. 사용하는 선형 방정식은 식 (5)와 같으며 다음과 같이 유도된다.

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} = \begin{bmatrix} P_1 & P_4 & P_7 \\ P_2 & P_5 & P_8 \\ P_3 & P_6 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (2)$$

$$x'_i = \frac{X_i}{Z_i} \quad y'_i = \frac{Y_i}{Z_i} \quad (3)$$

$$x'_i = \frac{P_1 * x_i + P_4 * y_i + P_7}{P_3 * x_i + P_6 * y_i + 1} \quad y'_i = \frac{P_2 * x_i + P_5 * y_i + P_8}{P_3 * x_i + P_6 * y_i + 1} \quad (4)$$

$$\begin{bmatrix} x_1 & 0 & -x_1 x'_1 & y_1 & 0 & -y_1 x'_1 & 1 & 0 \\ 0 & x_1 & -x_1 y'_1 & 0 & y_1 & -y_1 y'_1 & 0 & 1 \\ x_2 & 0 & -x_2 x'_2 & y_2 & 0 & -y_2 x'_2 & 1 & 0 \\ 0 & x_2 & -x_2 y'_2 & 0 & y_2 & -y_2 y'_2 & 0 & 1 \\ x_3 & 0 & -x_3 x'_3 & y_3 & 0 & -y_3 x'_3 & 1 & 0 \\ 0 & x_3 & -x_3 y'_3 & 0 & y_3 & -y_3 y'_3 & 0 & 1 \\ x_4 & 0 & -x_4 x'_4 & y_4 & 0 & -y_4 x'_4 & 1 & 0 \\ 0 & x_4 & -x_4 y'_4 & 0 & y_4 & -y_4 y'_4 & 0 & 1 \\ \dots & & & & & & & \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \\ P_6 \\ P_7 \\ P_8 \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \\ \dots \end{bmatrix} \quad (5)$$

(x, y) : 특징점의 위치(world coordinate)
(x', y') : 특징점의 투영변환 위치(image coordinate)
(P) : 투영변환 매개변수

찾아진 특징점이 4개보다 많은 경우에는 고유치 분해(singular value decomposition)라는 수학적 방법을 이용한다. 이 방법은 선형 방정식에서 최적 해를 찾는 방법으로 안정적인 결과를 얻을 수 있다^[4].

그림 4와 그림 5의 각 그림에서 우측 하단 부분의 결과 그림을 보면, 4개 이상의 특징점의 위치로부터 고유치 분해

박성춘 외 : 색상패턴 추적을 이용한 실시간 증강영상 시스템

에 의한 선형방정식을 적용시켜 얻은 매개변수로 패널의 위치와 모양을 계산하여 빨간색 사각형으로 표시하여 보여 주고 있다.

그림 4의 좌측 그림을 보면 패널의 일부분이 가려져 있어도 성공적인 추적이 이루어짐을 볼 수 있다. 가려져 보이지 않는 특징점의 위치는 얻어진 매개변수로부터 역으로 계산할 수 있으며, 다음 프레임에서 개체확인 인덱싱 과정에서 다시 사용된다. 그림에서는 빨간색 십자 표시로 나타내었다.

그림 4의 우측 그림과 같이 패널이 거의 가려지거나 화면 바깥으로 이동하여 특징점을 4개 이상 찾지 못하면 정확한 추적이 이루어지지 않지만, 이 경우 이전 프레임의 위치와 동일하게 간주하여 처리하므로 유사한 위치에서 패널이 보이는 부분이 다시 많아지게 되면 계속해서 정상적인 추적을 할 수 있다.

이렇게 얻어진 투영변환 매개변수는 디지털 영상효과 모듈의 제어에 사용되며, 결과적으로 가상스크린에 들어갈 영상을 정확하게 투영변환 시킨다.

IV. 운용 소프트웨어 개발

K-비전 운용 소프트웨어는 윈도우NT 운영체제 상에서 Microsoft사의 Visual C++TM, Matrox사의 MIL(Matrox Imaging Library)^{TM[6]}, DigiSuite SDK(Software Development Kit)^{TM[7]} 등을 이용하여 개발하였다. 그림 6에 사용자 인터페이스를 보이고 있다.

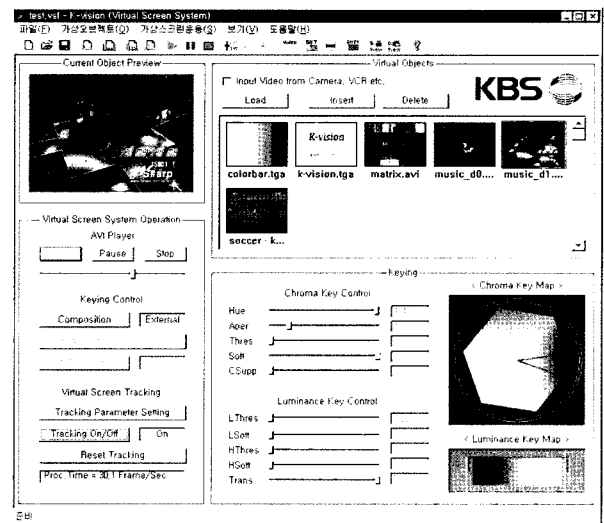


그림 6. K-비전 소프트웨어 사용자 인터페이스
Fig. 6. The user interface of K-vision software

스크린에 들어갈 영상으로는 사전에 M-JPEG 포맷으로 녹화 저장한 AVI 동영상 파일이나 그래픽 이미지 또는 입력 비디오 영상을 선택하여 나타낼 수 있는데, 화면의 우측 상단부분이 사전에 녹화 저장한 AVI 동영상 파일이나 그래픽 이미지를 선택할 수 있는 오브젝트 선택 창이고, 선택된 것은 좌측 상단의 현재 오브젝트 미리보기 창에 나타난다. AVI 동영상상이 선택되었을 경우에 재생, 잠시멈춤, 재생 끝내기 등을 이용하여 WAV 포맷의 오디오와 함께 동화상 AVI 파일을 실시간에 재생한다.

운용 소프트웨어는 그림의 우측 하단 부분에서 크로마키와 루미넌스키(luminance key)를 이용하여 합성할 수 있도록 몇 가지 매개변수를 조절하도록 되어 있다. 특히, 크로마키는 스크린의 두 가지 색상을 포함하도록 색상, 범위, 임계값 등의 매개변수를 조절하여 키 범위를 지정할 수 있도록 하였다.

그리고, 스크린과 특징점들의 정확한 추적을 위하여 HLS 범위, 특징점 블럽의 원형성 및 영역의 크기 등에 대

한 임계값 매개변수를 조절할 수 있도록 하였다.

스크린용 비디오를 지정하고, 크로마키와 추적 매개변수를 조절한 후에 실시간 추적을 실시하면, 스크린 패널을 추적하여 합성된 결과 영상을 볼 수 있다.

V. 방송 활용

K-비전은 2000년 제16대 국회의원 선거개표방송에서 처음으로 사용되었다. 개표방송을 위하여 프로토타입 시스템 두 대를 설치하여 사용하였는데, 관심지역의 투표자조사결과, 지도를 통해 알아보는 전국/ 시도별 당선자수, 정당별 당선자수의 분석표 등을 보여주는 용도로 활용하였다.

K-비전은 시사 다큐멘터리 프로그램인 '피플 세상속으로'에서 오프닝과 클로징 시에 프로그램 타이틀을 넣거나 방송 아이템이 바뀌는 부분에서 진행자가 아이템의 내용을



그림 7. K-비전 시스템의 방송 활용 예
Fig. 7. The application examples of K-vision system

이야기할 때 해당 비디오가 표시되는 용도로 사용되었다.

또한, 오락 프로그램인 '뮤직뱅크'에서는 가요 순위를 소개하는 그래픽 애니메이션 동영상을 보여주거나 오프닝에서 진행자간 대화를 특수한 효과로 보여주는 데 사용하기도 하였다.

K-버전은 모든 타입의 카메라와 사용 가능하며, 특별한 부가 장치가 필요하지 않다. 센서를 부착하지 않아도 되고, 카메라 매개변수의 사전 측정과 조율 작업인 캘리브레이션 과정 또한 필요하지 않다. 이 시스템은 일반 스튜디오나 가상 스튜디오에 모두 적용될 수 있다.

그림 7에서 실제 방송에서 사용한 몇 가지 예를 보이고 있다.

VI. 결 론

본 논문에서 가상스크린 시스템인 K-버전에 대해 소개하였다. 가상스크린 시스템은 증강현실 기술을 응용한 것으로 사용자가 들고 움직이는 패널에 동영상이나 그래픽 영상 등을 실시간으로 합성하여 보여주는데, 패널에 보여지는 모든 영상은 카메라의 움직임과 패널의 움직임에 따라 정확하게 입혀진다. K-버전은 윈도우NT를 탑재한 PC를 기반으로 구성하였고, 원형의 색상패턴을 갖는 패널을 사용하는데, 영상에서 패널의 움직임을 추적하기 위하여 이진영상

화, 블러분석, 특징추적과 같은 영상처리 기술을 이용한다. 특징점을 구별하기 위하여 위치 추정을 사용하였으며, 특징점 위치로부터 투영변환 매개변수를 얻기 위한 선형방정식을 사용하고, 고유치 분해를 이용하여 안정적인 결과를 얻는다. 이 시스템은 일반 스튜디오나 가상 스튜디오에 모두 적용될 수 있으며, 선거개표방송, 다큐멘터리, 오락 프로그램 등 생방송 프로그램에서 활용하였다.

참 고 문 헌

- [1] Ronald T. Azuma, "A Survey of Augmented Reality," Presence : Teleoperators and Virtual Environments, Vol. 6, No.4, Aug. 1997.
- [2] <http://www.symah-vision.fr/vhtml/epsis.htm>
- [3] <http://www.orad.co.il/virsets/index.htm>
- [4] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, "Numerical Recipes in C," Cambridge University Press, 2nd Ed., pp. 59-70, 1992.
- [5] Emanuele Trucco and Alessandro Verri, "Introductory Techniques for 3-D Computer Vision," Prentice Hall, 1998.
- [6] "Matrox Imaging Library Version 6.0, User Guide & Command Reference," Matrox Electronic Systems Ltd., 1999
- [7] "DigiSuite SDK Reference Manual," Matrox Electronic Systems Ltd., 1999.

저 자 소 개

박 성 춘



- 1988년 : 서울대학교 컴퓨터공학과 학사
- 1990년 : 한국과학기술원 전산학과 석사
- 1990년~현재 : 한국방송 기술연구소 연구원
- 주관심분야 : 컴퓨터 그래픽스, 영상처리 및 컴퓨터비전, 디지털 콘텐츠 처리

남 승 진



- 1989년 : 연세대학교 전자공학과 학사
- 1991년 : 연세대학교 공과대학원 전자공학과 석사
- 1991년~현재 : 한국방송 기술연구소 연구원
- 주관심분야 : 가상현실



오 주 현

- 1997년 : 부산대학교 전자공학과 학사
- 1999년 : 부산대학교 대학원 전자공학과 석사
- 1999년~현재 : 한국방송 기술연구소 연구원
- 주관심분야 : 영상처리, 가상현실, 콘텐츠 제작 시스템



박 창 섭

- 1980년 : 한양대학교 전자공학과 학사
- 1982년 : 한양대학교 산업대학원 석사
- 2001년 : 한국과학기술원 정보 및 통신공학과 박사
- 1981년~현재 : 한국방송 기술연구소 차장(선임연구원)
- 주관심분야 : 영상신호처리, 가상현실, 컴퓨터비전, 입체 TV