

경기장 모델을 이용한 가상 영상 삽입 시스템의 설계 및 구현

유 성⁰ / 한 송이, 이 성환
(주)버추얼미디어⁰ / 고려대학교 컴퓨터학과
syoo@virtualmedia.co.kr, {sihan, swlee}@image.korea.ac.kr

Design and Implementation of a Virtual Image Insertion System with a Sports Field Model

Seong Yoo⁰ / Song-Yi Han, Seong-Whan Lee
Virtual Media, Inc.⁰ / Dept. of Computer Science and Engineering, Korea University

요 약

본 논문에서 제안하는 가상 영상 삽입 시스템은 카메라의 조작이나 시스템 운영자의 개입 없이 모든 처리 과정이 자동으로 진행된다. 이를 위해 시스템은 경기장 좌표계를 정의하고 삽입할 영상의 크기와 위치를 정하는 과정, 경기장의 특징점들을 추출하는 과정, 경기장 좌표계와 참조 영상의 특징점들로부터 투영 변환 파라미터를 추출하는 과정, 실제 동영상에서 삽입 위치를 찾고 추적하여 가상 영상을 삽입하는 과정을 거치게 된다. 본 논문에서 제안한 시스템의 성능을 검증하기 위해 방송용 NTSC 비디오 데이터를 대상으로 실험을 하였으며 그 결과 각 모듈들과 시스템이 효율적임을 입증하였다.

1. 서 론

디지털 방송 시대의 도래와 함께 비디오 편집이나 특수 효과, 영상 합성 등에 대한 다양한 애플리케이션 수요가 잇따르고 있다. 최근에는 경기장 내에 가상 영상을 삽입하는 부가서비스가 시도됨에 따라 다양한 가상 영상 삽입과 교체에 대한 연구가 활발히 시도되고 있다.

보다 사실적이고 자연스러운 가상 영상 삽입을 위해 Astle 등은 센서 포인트에서 빛의 양의 변화를 추적하는 방법을 사용하였고[1], Rosser 등은 특수 카메라나 센서 등에 의존하지 않고 컴퓨터비전을 기반으로 가상 영상 삽입 위치 결정 방법을 제안하였다[2]. 또한 Medioni 등은 동영상 내의 임의의 객체를 가상 영상으로 교체하는 시스템을 구현하였다[3][4]. 국내에서는 영상 내의 카메라 동작 분석을 통해 삽입 위치를 찾아서 가상 광고를 삽입하는 방법이 개발되었다[5].

본 논문에서 제안하는 가상 영상 삽입 시스템은 경기장 모델을 이용함으로써 정확한 파라미터에 근거한 영상 변형을 통해 보다 정확하고 사실적인 가상 영상을 생성할 수 있을 뿐만 아니라 시스템 운영자가 가상 영상을 수동으로 변형시켜 생성해야 하는 불편함을 해소하였다. 또한 정확하면서도 빠른 알고리즘을 제안하여 온라인 상에서의 고속 처리를 가능하게 하였다.

2. 시스템 설계

본 논문에서 제안하는 시스템은 온라인 상에서의 운영자의 조작을 최소화하기 위해 오프라인 처리부와 온라인 처리부로 나누어져 있다. 오프라인 처리부에서는 시스템 운영자가 경기장 모델을 정의하고 삽입될 영상의 크기를 선택한 다음, 경기 시작 전 카메라로부터 참조 영상을

받아 등록해 준다. 그러면 시스템이 자동으로 특징점을 추출하여 그 부분의 패턴 영상을 등록해 주는 한편, 투영 변환 파라미터를 추출하여 실제 삽입될 영상으로 변형시켜 준다. 그림 1과 같이 온라인 처리부에서는 실제 동영상 데이터에서 시스템이 장면전환과 객체 겹침 문제를 고려하면서 삽입 영역을 자동으로 찾아내어 변형된 가상 영상을 삽입하게 된다.

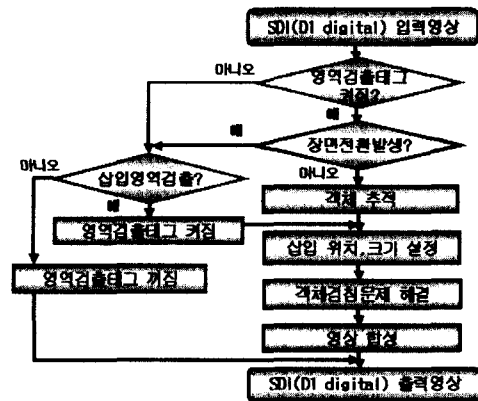


그림 1 온라인 처리부 흐름도

3. 가상 영상 삽입 알고리즘

3.1 특징점 추출

본 논문에서는 경기장 필드 영역과 외곽선 추출하여 제한된 영역 내에서 선분 및 특징점들을 추출한다. HSV 컬러 공간에서 양자화 및 Opening, Closing 연산자를 순차적으로 적용하고 최소제곱법을 이용하여 외곽의 방정

식을 얻어 필드영역의 추출 결과를 토대로 허프 변환(Hough Transform)을 이용하여 효율적으로 경기장 내의 특징점을 추출할 수 있게 된다.

3.2 가상 영상의 기하학적인 변형

본 논문에서는 경기장 좌표계와 영상 좌표계의 관계를 정립하여 보다 사실적이고 자연스러운 가상 영상의 삽입을 구현하였다. 이를 위해 우선 경기장 모델을 정의하여 좌표계를 설정하고 영상 좌표계의 대응점들로부터 변환 파라미터를 추출한 다음 그 파라미터들에 따라 삽입될 가상 영상을 워핑시켰다.

3.2.1 2차원 객체의 변형

일반적으로 스포츠 경기장은 2차원으로 구성된 고정된 규격의 필드 영역을 가지므로 경기장 좌표계는 2차원 좌표계로 정의할 수 있다. 결국 2차원 객체를 삽입하고자 할 때 투영 변환(Perspective Transform)에 관한 파라미터는 3x3의 행렬 C로 정의된다. 그리고 변환 전 점의 원래 좌표를 (x,y)라고 하고 그 점의 변환 후 좌표를 (x',y')라고 하면 방정식은 식 (1)과 같다.

$$\begin{matrix} x' \\ y' \end{matrix} = \frac{\begin{matrix} c_{11}x + c_{12}y + c_{13} \\ c_{21}x + c_{22}y + c_{23} \end{matrix}}{\begin{matrix} c_{31}x + c_{32}y + c_{33} \end{matrix}} \quad \begin{matrix} y' \\ x' \end{matrix} = \frac{\begin{matrix} c_{21}x + c_{22}y + c_{23} \\ c_{11}x + c_{12}y + c_{13} \end{matrix}}{\begin{matrix} c_{31}x + c_{32}y + c_{33} \end{matrix}} \quad (1)$$

여기서 c₃₃=1로 고정된 상수이며 따라서 위의 방정식을 풀고 적절한 파라미터 값인 행렬 C를 구하기 위해서는 4쌍 이상의 대응점을 필요로 한다.

3.2.2 3차원 객체의 변형

3차원 좌표로 구성된 경기장 안에 3차원 객체를 삽입할 수도 있도록, 이 관계를 카메라 보정(Camera Calibration)의 원리를 이용하여 정의하였다. 추출하고자 하는 변환 파라미터는 동차 좌표계(Homogeneous Coordinates)를 이용하여 식 (2)와 같이 3x4 행렬 C로 표현할 수 있다. 식 (2)에서 평면 영상 좌표 (x',y')는 동차 좌표 (hx, hy, h)로 나타낼 수 있으므로 이를 이용하면 식 (2)로부터 식 (3)와 같은 두 개의 선형 방정식을 얻을 수 있다.

$$\begin{bmatrix} x \\ y \\ h \end{bmatrix} = C \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2)$$

$$\begin{matrix} c_{11}X + c_{12}Y + c_{13}Z + c_{14} - c_{31}x X - c_{32}x Y - c_{33}x Z - c_{34}x = 0 \\ c_{21}X + c_{22}Y + c_{23}Z + c_{24} - c_{31}y X - c_{32}y Y - c_{33}y Z - c_{34}y = 0 \end{matrix} \quad (3)$$

유도된 식 (3)에서 12개의 미지 계수를 구하기 위해서는 알려진 좌표 (X_i, Y_i, Z_i), i = 1,2,..., m (m은 6 이상의 정수)을 갖는 실제점을 구하고, 이에 대응하는 영상점 (x_i, y_i), i = 1,2,..., m을 대입하여 선형 방정식의 최적해를 구한다.

3.3 삽입 영역 탐지 및 추적

본 논문에서는 동영상 내에서 삽입 영역을 탐지하고 추적하기 위해 상관성에 의한 정합을 이용하였다. 상관

성의 의한 정합이란 한 영상과 다른 영상 내의 일정한 패턴과의 유사성을 수치로 나타내어 영상을 해석하고자 하는 방법이다. 또한 영상을 보다 고차원적으로 해석하고 객체 겹침이 발생하여 영상에서 사라진 특징점들을 복원하기 위하여 기하학적인 해상 방법을 이용하였다.

3.3.1 상관성에 의한 정합

온라인 상에서 시스템이 고속으로 삽입 위치를 찾고 계속해서 삽입 위치를 유지하기 위해서는 오프라인 상에서 등록했던 특징점들에 대한 패턴 영상을 이용하게 된다. 우선 이 패턴에 대한 영상들을 하나의 형판이라고 가정하고 동영상의 각 프레임들을 전체 영상이라고 가정하면 그 전체 영상의 일부에 해당하는 윈도우와 형판 사이의 상관성을 식 (5)와 같은 상관 계수(Correlation Coefficient)로 나타낼 수 있다. 이 때 식 (4)와 같이 정의하고 T̄는 형판 내의 점들의 평균값을 나타내며 윈도우 내의 점들의 평균값을 나타낸다.

$$\begin{matrix} T(x',y') = T(x',y') - \bar{T} \\ I'(x+x',y+y') = I(x+x',y+y') - \bar{I}(x,y) \end{matrix} \quad (4)$$

그러나 이러한 상관 계수는 윈도우와 형판 내 점들의 값의 변화에 민감하다는 단점을 가지므로 본 논문에서는 식 (6)의 정규화된 상관 계수(Normalized Correlation Coefficient)를 사용하였다.

$$\begin{matrix} R(x,y) = \sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x',y') \bar{I}(x+x',y+y') \\ \bar{R}(x,y) = \frac{\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x',y') \bar{I}(x+x',y+y')}{\sqrt{\sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} T(x',y')^2 \sum_{y'=0}^{h-1} \sum_{x'=0}^{w-1} \bar{I}(x+x',y+y')^2}} \end{matrix} \quad (5), (6)$$

3.3.2 가려진 특징점 복원

특징점이 선수에 의해 가려지는 객체 겹침 문제(Occlusion Problem)를 해결하기 위해 본 논문에서는 객체의 효율적 표현과 정합을 이용한 기하학적 해상[6]의 개념을 이용하였다. k개의 필요한 기준점을 k-basis라 하고 그 기준에 대해 나머지 모든 점들의 상대 좌표를 계산하여 해시 테이블에 저장한다. 그리고 이를 이용하여 객체 겹침 문제가 발생한 특징점들의 좌표를 복원한다.

4. 실험 결과 및 분석

4.1 실험 환경

본 논문에서는 방송용 장비와의 연동을 통해 실제 방송국과 유사한 환경을 구축하여 실험하였다. 우선 입력 데이터로서 SONY DigiBeta VCR을 통해 재생되는 다양한 포맷(SDI, Component, Composite)의 동영상을 데이터로 사용하였고 입력된 비디오 스트림의 프레임 단위 캡처 및 실시간 영상 합성을 위해 Matrox사의 Digisuite 보드를 이용하였다. 또한 Pentium III 700Mhz 듀얼 프로세서와 SCSI 하드디스크를 탑재한 IBM PC를 서버로 채

택하였고 서버에서 처리된 동영상 데이터는 NTSC 방송용 모니터를 통해 출력하도록 하였다.

4.2 실험 결과

그림 2는 오프라인 상에서 등록되어 변형된 가상 영상을 실제로 온라인 상에서 프레임으로 삽입했을 경우에 대한 결과로 가상 영상의 기하학적인 변형, 주변 환경과의 조화, 선수와의 겹침 문제가 해결된 것을 보여준다.

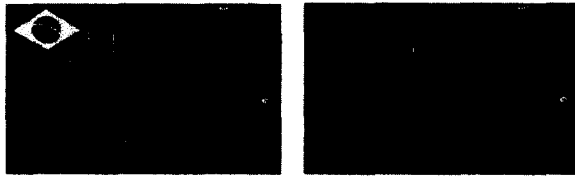
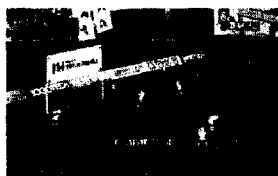


그림 2 가상 영상 삽입 예

그림 3의 (a)는 2차원 가상 영상을 바닥이 아닌 경기장 내의 다른 2차원 평면으로 삽입한 결과이고 (b)는 OpenGL을 이용하여 3차원 객체인 구를 모델링하여 동영상 프레임 내에 삽입한 결과이고 (c)는 기하학적 해싱을 이용한 삽입 위치 방법으로 복원된 삽입 위치에 따라 가상 영상의 일부만이 삽입된 결과이다.



(a)



(b)

(c)

그림 3 (a) 다른 평면으로 영상 삽입 (b) 3차원 가상 객체 삽입 (c) 기하학적 해싱을 이용한 영상 삽입

한편 표 1과 같이 Medioni가 제안한 시스템과 PVI 등의 업체에서 사용하고 있는 시스템을 대상으로 성능을 비교, 분석하였다. 첫 번째, 네 번째 항목이 자연스러운 삽입의 정도, 두 번째, 세 번째 항목이 시스템의 자동화 정도, 다섯 번째 항목이 시스템의 처리 속도를 나타내는 것으로, 그 결과 본 논문에서 제안한 시스템이 가상 영상의 삽입 시 자연스러운 삽입, 시스템 자동화, 시스템의 처리 속도 면에서 상당히 우수함을 입증할 수 있었다.

5. 결론 및 향후 연구 방향

본 논문에서는 스포츠 동영상에 가상 영상을 자연스럽게 삽입할 수 있는 자동화된 시스템을 설계하고 구현하

평가항목 \ 시스템	제안된 시스템	Medioni 시스템	PVI 시스템
삽입좌표계	실좌표계	영상좌표계	영상좌표계
삽입영상변형	자동	수동	수동
삽입영역 추출	컴퓨터비전	컴퓨터비전	특수 하드웨어
객체 겹침 해결	크로마키	없음	크로마키
처리속도	중	하	상

표 1 시스템들의 성능 평가

였다. 오프라인/온라인 처리부의 분리로 그 효율성을 높였으며, 삽입 위치의 추적 시에 일부 정보들이 가려질 경우에도 가상 영상을 삽입하는 것이 가능하게 하였다.

제안한 알고리즘들이 매우 만족할 만한 성능을 보임에도 불구하고 여전히 개선해야 할 문제들이 있다. 우선 경기장 모델을 사용하고자 할 경우 필드 안의 선분들이나 특징점이 존재해야 하므로 그렇지 않은 영역에서는 본 논문에서 제안하는 방법을 사용할 수 없다는 제약이 가진다. 또한 생중계되는 스포츠 동영상에 가상 영상을 삽입하고자 한다면 실시간으로 처리, 즉 초당 30 프레임 이상을 처리해야 한다. 그러나 현재의 하드웨어 환경에서는 10-15 프레임 밖에 처리할 수 없어 실시간 처리는 불가능하다. 앞으로의 위와 같은 문제들에 대한 연구가 지속적으로 이루어져야 할 것이다.

참고 문헌

- [1] B. Astle and S. Das, "System and Method of Real Time Insertions into Video Using Adaptive Occlusion with a Synthetic Reference Image", U.S. Patent #5,953,076, September, 1999.
- [2] R. Rosser, S. Das, Y. Tan and P. von Kaenel, "Pattern Recognition System Employing Unlike Templates to Detect Objects Having Distinctive Features in a Video Field", U.S. Patent #5,627,915, May, 1997.
- [3] G. Medioni, G. Guy, H. Rom and A. Fracois, "Real Time Billboard Substitution in a Video Stream", Proc. of the 10th Tyrrhenian International Workshop on Digital Communications, Multimedia Communications, Ischia, Italy, pp. 71-84, 1998.
- [4] G. Medioni, G. Guy and H. Rom "Video Processing System for Modifying a Zone in Successive Images", U.S. Patent #5,436,672, July, 1995.
- [5] 신 수원, 이 형석, 이 성환, "카메라 움직임을 고려한 자연스러운 가상 객체 삽입 방법 및 장치", 출원 번호 2000-49575, 2000년 12월.
- [6] Y. Lamdan and H. J. Wolfson, "Geometric Hashing: A General and Efficient Model-Based Recognition Scheme", Proc. of the International Conference on Computer Vision, pp. 238-249, December 1988.