

충돌음 자동생성을 위한 영상/음향 동기화 기술 개발

김재우, 김용완, 김현빈
한국전자통신연구원 휴먼컴퓨팅연구부
대전시 유성구 가정동 161
woo@seri.re.kr

Image/Sound Synchronization for Simulating Sound from Colliding Objects

Jaewoo Kim, Yongwan Kim, Hyunbin Kim
Human Computing Department, ETRI
1 Kajong Dong, Yusong Gu, Taejon
woo@seri.re.kr

요약

멀티미디어 기술의 발달로 그래픽을 포함한 영상 데이터의 사용이 급증하고 있다. 멀티미디어 및 가상현실 응용 시스템은 그래픽을 포함한 영상과 음향이 동기화 되어 제공될 경우 사용자의 현장감과 몰입감을 증대시킬 수 있다.

본 논문에서는 가정용 비디오를 통하여 입력되는 일련의 영상데이터로부터 관심있는 물체의 운동을 추적함으로써 물체가 발생시키는 충돌음을 공간적으로 동기화 하여 생성하는 이미지 기반 영상/음향 동기화 기술의 개발에 관하여 논의한다.

대용량의 영상데이터에 대하여 실시간으로 음향을 동기화시키는 것은 현재 기술로 불가능하며, 본 논문에서는 영상과 음향의 공간적 동기화 기술에 대하여 기술한다.

1. 서론

최근 영상산업의 발달로 영상데이터의 처리가 필

수적인 요소기술로 급부상하고 있다. 그래픽 및 영상에 관한 연구는 활발히 진행되고 있는데 반하여 음향에 관한 연구는 상대적으로 미비하며, 특히 영상과 음향을 동기화하여 통합적으로 관리하는 방법론에 관한 연구는 거의 전무한 실정이다.

동영상의 경우 초당 30 프레임의 대용량 영상 데이터에 대한 음향효과 삽입 작업을 수작업으로 해야하는 번거로움이 있다. 본 연구에서는 동영상 데이터 처리를 위한 시스템으로써, 비디오 카메라를 통하여 입력되는 동영상 데이터로부터 관심있는 물체의 운동을 추적, 해석함으로써 물체의 충돌에 상응하는 음향효과를 동기화하여 자동생성하는 시스템에 관하여 기술한다. 본 시스템은 동영상 데이터에 음향효과를 첨가하는 과정에서 충돌음의 효과를 거리감, 공간감, 방향감을 갖는 입체음향 기술을 적용한 입체음으로 생성함으로써 고품질의 음향효과를 자동으로 생성할 수 있다.

2. 영상/음향 동기화 기술

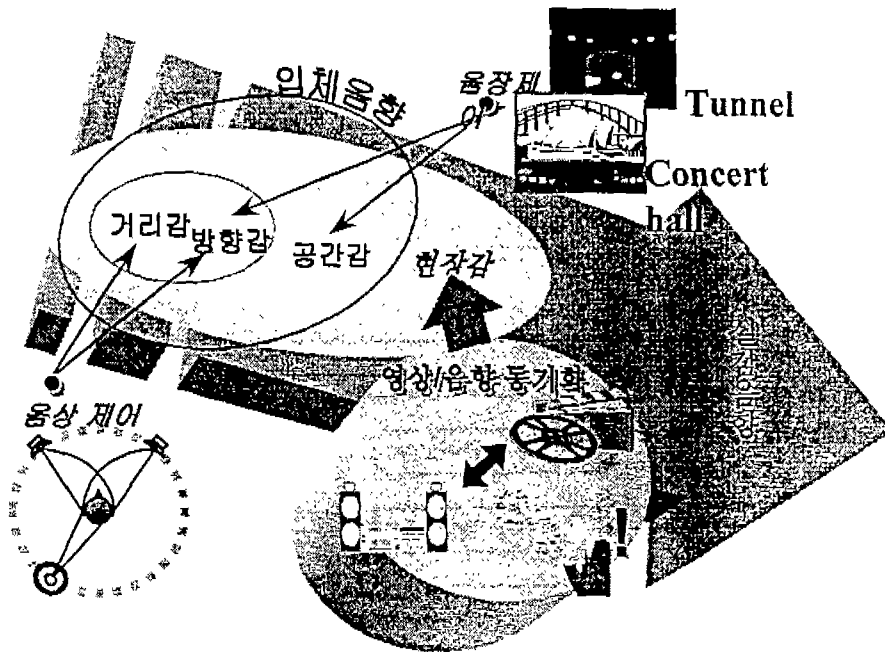


그림 1. 이미지 기반 영상/음향 동기화 기술

멀티미디어 또는 가상현실 등의 분야에서 사용한 종래의 음향효과는 모노 또는 스테레오음으로 사용자에게 거리감, 공간감, 방향감등을 전달할 수 없었으나 입체음향 기술을 적용할 경우 사용자에게 보다 현실감 있는 음향을 제공할 수 있어 현장감과 몰입감을 증대할 수 있다. 이와같은 입체음향 기술은 영상과 동기화 되어 구현되었을 때 효과가 극대화 된다. 그림 1은 입체음향 기술과 통합된 영상/음향 동기화 기술의 구성을 보여준다.

현재까지 그래픽 영상에 대한 영상/음향 동기화 기술에 대한 연구는 부분적으로 진행되어 왔으나 이미지 데이터로 구성된 동영상에 대한 영상/음향 동기화 기술에 대한 연구는 전무한 실정이며 이를 이미지 기반 영상/음향 동기화 기술이라 정의한다. 이와같은 기술이 적용될 경우 앞으로 TV 방송, 영화 등의 영상산업 분야에 획기적인 기술 발전을 이룰 수 있을 것으로 기대한다. 본 논문에서 제안한 기술이 실용화 될 경우 종래의 수작업에 의존하던 영상에 음향효과를 삽입하는 작업의 단점을 부분적으로 개선하여 자동으로 이루어 질 수 있게 해줄 것이다. 이미지 기반 영상/음향 동기화 기술은 동영상 데이터에 대한 물체 추적 기술, 입체음향 기술, 그리고 두가지 기술 사이의

인터페이스로 구성된다.

3. 이미지 기반 영상/음향 동기화 기술 개발

본 논문에서 개발한 이미지 기반 영상/음향 동기화 시스템은 PC에 기반을 둔 시스템으로서, 운동하는 물체의 위치를 파악하여 그 운동 궤적을 구하고, 이를 해석하여 충돌을 감지함으로써 충돌 위치와 주변 환경에 상응하는 입체음향의 제공한다.

상기 목적을 달성하기 위하여 본 발명은, 가정용 비디오 카메라를 이용하여 촬영한 비디오 데이터를 영상 획득장치를 통하여 입력한 후, 획득한 일련의 영상 데이터들에 대하여 전처리 과정을 거친 후, 차영상을 구하여 물체 위치의 후보영역을 구하고, 일련의 후보영역들에 대하여 역학모델을 통하여 물체 위치를 보정하고 검증하여 일련의 위치 정보를 얻는다. 상기의 과정을 통하여 구한 일련의 위치정보들로부터 추출한 운동궤적을 이용하여 물체의 충돌을 감지한 후 감지된 충돌에 대하여 공간의 특성과 물체의 방향에 기반하여 이에 상응하는 효과음을 정의하고 이를 사전에 측정된 HRTF DB 및 음장 파라미터를 사용

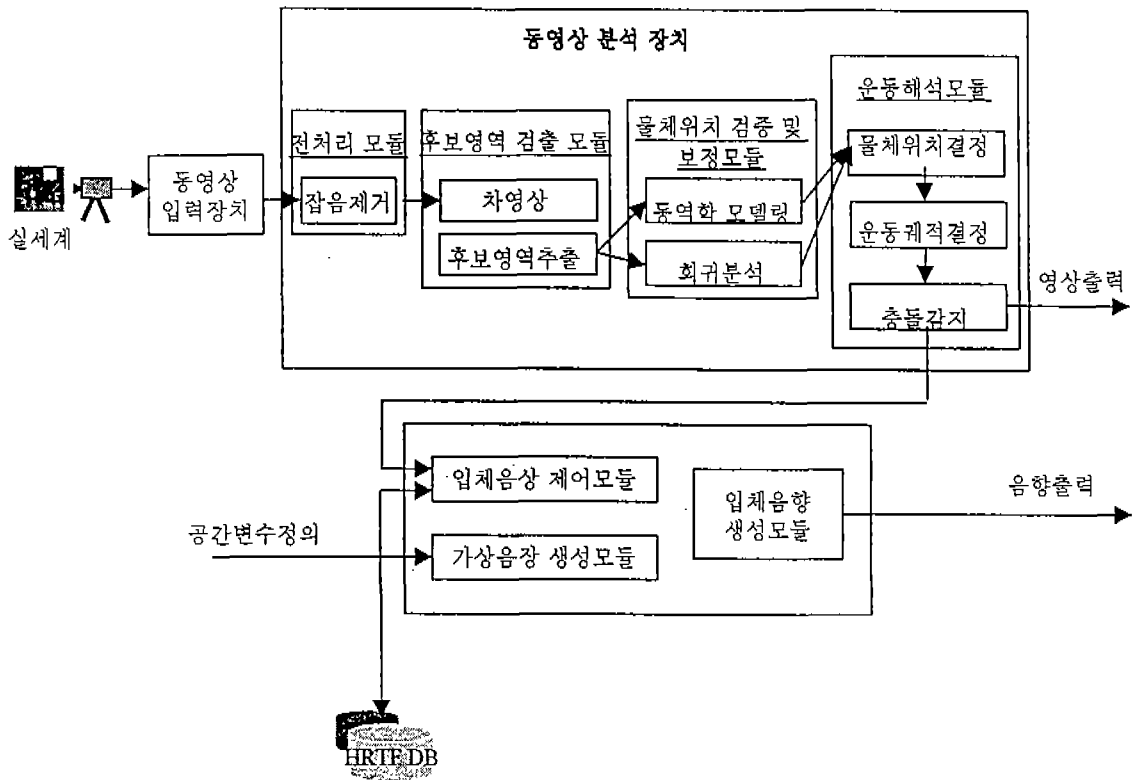


그림 2. 이미지 기반 영상/음향 동기화 시스템 구성도

하여 입체음향을 생성하여 영상과 동기화한다. 그림 2는 본 시스템의 전체 구성을 나타낸다. 그림 2에서 보는 바와 같이, 본 시스템은 크게 동영상 분석장치와 입체음향 생성장치로 구분된다. 동영상 분석장치는 실세계에서 운동하는 물체의 운동을 분석하여 추출된 정보를 입체음향 생성장치로 전달하게 되고 입체음향 생성장치에서는 물체가 운동하고 있는 주변의 공간적 특성과 물체의 동작에 동기화된 입체음향을 생성하여 영상과 음향을 동기화 한다. 그림 3에서 보는 바와 같이 본 연구에서는 자연 배경상에서 운동하는 탁구공을 대상으로 물체의 충돌탐지를 통한 충돌음을 구현하는 시스템을 개발하였다.

3.1 전처리 과정

실세계에서 운동하는 물체를 가정용 비디오 카메라로 입력하고 이를 동영상 입력장치를 통하여 컴퓨터 저장장치에 저장한다. 저장된 일련의 영상들은 전처리 모듈을 통하여 잡음제거 과정을 통하여 영상내에 삽입되어 있는 잡음을 제거한

다.

3.2 후보영역 검출

상기 잡음제거 과정을 통하여 잡음이 제거된 일련의 영상에 대하여 후보영역 검출 모듈은 각각 시간 축에 대한 두 개의 영상을 쌍으로 차영상을 구한 후 이를 바탕으로 물체 운동에 의해 변화된 정보를 이용하여 물체 위치의 후보영역을 검출한다. 후보영역의 검출에 있어서 본 연구에서 사용한 탁구공의 경우 물체의 크기가 상대적으로 작고 물체의 검출에 있어서 유용하게 사용될 수 있는 단서가 적당치 않아 차영상에 대한 임계치 결정을 통하여 얻은 이진영상을 단서로 사용하였다.

3.3 물체위치 검증 및 보정

물체위치 검증 및 보정 모듈은 상기 후보영역 검출모듈에서 검출한 후보영역들에 대하여 동역학 모델 또는 회귀분석 기법을 이용하여 물체의 위치를 검증하고 이를 보정한다.

3.4 운동 해석

운동해석 모듈은 상기 물체위치 검증 및 보정 모듈의 결과를 기반으로 각 프레임상에서의 물체의 위치를 결정하고 이를 바탕으로 물체의 운동 궤적을 구한 후 충돌을 감지한다. 충돌의 감지는 역학모델에 위배되는 급격한 운동방향의 전환을 단서로 결정한다.

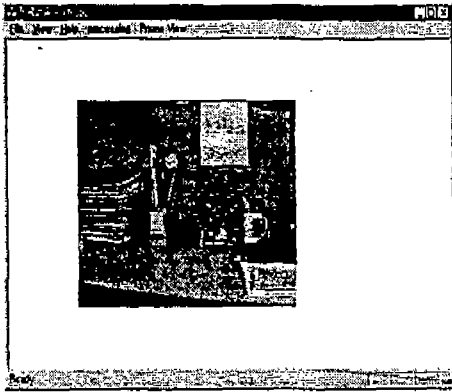


그림 3. 이미지 기반 영상/음향 동기화 시스템

3.5 입체음향 생성

이상의 과정을 통하여 해석된 결과들은 매개변수로 입체음향 생성장치로 전달된다. 입체음향 생성장치에서는 상기의 동영상 분석장치에서 전달된 매개변수를 입력하여 입체 음향 생성 모듈에서 해당 HRTF DB를 참조하여 위치음을 생성하고, 사용자의 요구에 따라 입력되는 공간변수의 정의에 따라 가상 음장 생성 모듈에서 음장감을 구현한다. 이와 같이 생성된 입체음향은 영상과 함께 동기화 영상은 모니터로, 입체음향은 스피커를 통하여 재생된다.

4. 결론

본 논문에서는 이미지 기반 영상/음향 동기화 기술을 정의하고 실세계에서 운동하는 단순물체의 운동을 해석하여 그 충돌음을 공간적으로 동기화 하는 시스템의 개발에 관하여 논의하였다. 음상정위 및 음장생성의 결과물은 영상과 동기화 되어 구현될 때 그 효과를 극대화 할 수 있으며 사용자에게 마치 현장에서 보고 듣는 것과 유사

한 경험을 제공할 수 있어 멀티미디어 및 가상현실 응용분야에 적용될 경우 시스템의 성능을 배가시킬 수 있다.

참고문헌

- [1] Tapio Takala et. al. "Using Physically-Based Models and Genetic Algorithms for Functional Composition of Sound Signals, Synchronized to Animated Motion", '93 International Computer Music Conference, 180-184, Sep.1993.
- [2] Suneil Mishra et. al. "Mapping Motion to Sound and Music in Computer Animation and VE" Invited Paper, Pacific Graphics, Vol 1, Aug. 95.
- [3] Bill Gardner and Keith Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone", MIT Media Lab, Technical Report #280, May, 1994.
- [4] Durand R, Begault, "3D SOUND for Virtual Reality and Multimedia", Academic Press, Inc., 1994.
- [5] Jens Blauert, "Spatial Hearing", (translated by John Allen), The MIT Press, 1983.
- [6] 강성훈, 강경옥, "입체음향", 기전연구사, 1997.