

멀티미디어 응용을 위한 카메라 앵커 처리 기술

노연^o 신규용 임근휘 안병찬
조민희 윤호상 은성배[†] 윤현수

한국과학기술원 전산학과

[†]한남대학교 정보통신공학과

Camera Anchor for Multimedia Hyperpresentation

Yeon No^o Kyuyong Shin Geunhwi Lim Byungchan Ahn
Minhee Cho Hosang Yoon Seongbae Eun[†] Hyunsoo Yoon

Department of Computer Science

Korea Advanced Institute of Science and Technology

[†]Department of Information Communication Engineering
Han-Nam University

요약문

대화형 멀티미디어 응용에 많이 이용되는 하이퍼미디어 시스템은 사용자가 관심을 끄는 정보에 연결된 링크를 이용하여 다양한 정보를 쉽게 획득하도록 지원한다. 이렇게 링크가 부착된 지점을 앵커라 하는데, 기존의 앵커에 대한 연구는 텍스트나 이미지 등의 정적미디어와 VOD (Video On Demand)에서와 같이 이미 만들어져 저장된 동영상에 한정되어 있다. 본 논문에서는 카메라로부터 실시간으로 생성되는 동영상에 앵커를 설정하고 처리하는 방법에 대해 제안하고, 제안된 방법으로 구현한 카메라 앵커 처리 시스템 프로토타입에 대해 소개한다.

1 서론

최근 들어 대화형 멀티미디어 정보 시스템의 사용이 급증하고 있다. 이러한 대화형 멀티미디어 정보 시스템들은 사용자와의 상호 작용을 주고 받을 수 있는 다양한 기능을 갖추며, 저장된 정보들을 다양한 방식으로 이용할 수 있도록 구조화되어 있다. 특히, 하이퍼미디어 시스템과 같은 경우, 사용자의 관심을 끄는 정보마다 직접 링크를 연결하여 다양한 정보를 순서에 구애받지 않고 쉽게 접근하도록 지원한다. 하이퍼미디어에서 링크가 부착된 지점을 앵커(anchor)라 한다. 기존의 텍스트나 이미지 등의 정적미디어에서는 앵커를 다는 것이 비교적 쉬우나, 음성이나 동영상과 같은 동적 미디어에 앵커를 다는 것은 여러가지 고려되어야 할 사항을 갖고 있다.

동영상에서 앵커를 처리하는 문제에 관한 기존의 연구를 살펴 보면, 앵커의 형태와 사람의 반응과의 관계를 조명한 연구 [1], Hypercard에서의 sticky button 문제 [2], 음성 앵커를 다는 연구 [3], 동영상에서 앵커의 형태와 이의 처리에 관한 실험 [4], 동영상에서 앵커를 시간축으로 다는 dynamic links [5]와 active anchor [6]의 예, 동영상에서 surfing [7]하는 문제 등이 있다. 그러나, 동영상에 앵커를 다는 기존의 연구는 대부분 VOD 서비스를 위한 MPEG 화일과 같이 이미 만들어진 동영상물 대상으로 하고 있으

며, 카메라에 의해 실시간으로 생성되는 동영상에 앵커를 다는 문제에 관해서는 거의 연구된 바가 없다.

이미 만들어진 동영상의 경우에는 대상 객체가 각 프레임 상에서 어디에 위치하는지를 미리 알 수 있기 때문에 앵커의 설정 및 처리과정이 용이하다. 즉, 각 프레임마다 앵커의 위치를 미리 설정해 놓고 이와 함께 앵커가 선택되었을 때 취해야 할 행동(action)을 정의하면 된다. 그에 반하여 카메라에 의해 실시간으로 생성되는 동영상에서는 대상 객체가 실제 환경(physical environment) 내의 3차원 좌표로 표현되어진다. 따라서, 대상 객체가 카메라로 포착되는 영상 내에서 어떠한 2차원 좌표를 가지는지를 구해야 한다. 또한 대상 객체는 실제로는 점이 아니라 모양과 크기를 가지는 물체이므로 3차원 객체 모델링도 필요하다.

본 연구에서는 이와 같이 카메라로부터 실시간으로 생성되는 영상에 앵커를 설정하는 것에 관해 연구한다. 이때, 앵커를 설정하고자 하는 객체는 고정되어 있는 것을 대상으로 한다. 본 연구에서는 카메라와 물체의 3차원 위치를 알아내는 방법, 카메라와 물체의 3차원 위치와 카메라 렌즈의 방향, 줌(zoom)의 정도로부터 카메라 영상 내에서 물체가 가지는 영역을 계산해 내고 이로부터 물체에 앵커를 설정하는 방법, 그리고 앵커 선택시 취해야 할 행동에 대해서 연구한다. 또한, 카메라 앵커 시스템의 프로토타입을 설계하고 구현하였다. 이러한 카메라 앵커 시스템은 흥미평

이나 원격 관람 시스템 등에 응용될 수 있다.

본 논문의 구성은 다음과 같다. 2절에서는 카메라 앵커 시스템에서 카메라로부터 실시간으로 생성되는 동영상에 앵커를 설정하고 처리하는 방법에 대해 살펴 보고, 3절에서는 본 연구에서 구현한 카메라 앵커 처리 시스템 프로토타입에 대해 소개한다. 마지막으로 5절에서 결론을 맺고, 앞으로의 연구 방향에 대해 논의한다.

2 카메라 앵커 시스템

카메라 앵커 처리는 카메라로부터 실시간으로 생성되는 동영상에서의 앵커 처리 방법이다. 실시간으로 생성되는 동영상에 대한 앵커 처리 방법은 대상 물체가 고정된 경우와 이동하는 경우로 나눌 수 있다. 대상 물체가 고정된 경우라 할지라도 사용자가 카메라를 제어함에 따라 사용자가 보는 영상은 계속 변화하게 된다. 이 경우에는 물체의 위치 정보를 사전에 얻을 수 있으므로 이를 앵커 처리에 이용할 수 있다. 그러나, 물체가 움직이는 경우에는 실시간으로 물체의 위치 정보를 획득하지 않는 한 위치 정보만을 이용해서 물체를 식별하는 것을 불가능하다. 따라서, 움직이는 물체의 식별을 위해서는 물체의 초기 위치 정보뿐만 아니라 부가적으로 패턴 인식 등의 이미지 처리 기술을 필요로 한다.

본 논문에서 제안하는 카메라 앵커 시스템은 고정된 물체를 대상으로 한다. 카메라 앵커 처리를 위해서는 먼저 대상 객체의 3차원 위치를 측정하여 데이터베이스에 저장하여야 하고, 이와 함께 카메라가 설치된 위치, 카메라 렌즈의 방향, 줌(zoom) 정도에 따라 카메라의 영상에서 대상 객체가 가지는 영역을 알아내어야 한다. 부가적으로 하나의 카메라가 볼 수 있는 시야는 제한되어 있으므로, 넓은 지역을 관찰하기 위해서는 카메라 간의 제어 이동 기술이 필요하다. 본 절에서는 카메라 앵커 시스템의 전체 구조와 대상 객체의 3차원 위치를 측정하는 방법, 영상에서 대상 객체가 보이는 위치 및 영역, 그리고 카메라 간 제어 이동 방법에 대하여 알아 본다.

2.1 카메라 앵커 시스템 전체 구조

카메라 앵커 시스템은 각각의 세부 지역을 담당하는 카메라를 제어하기 위한 카메라 제어기, 모든 카메라 제어기를 총괄하여 담당하는 중앙 서버(anchor server), 각각의 사용자를 위한 클라이언트의 크게 세 부분으로 나뉘어진다. 그림 1은 목표 시스템의 전체 구조를 나타낸다.

하나의 카메라를 제어하기 위한 카메라 제어기는 중앙 서버로부터의 제어 명령에 따라 카메라 시선을 제어한다. 또한 카메라와 연결된 비디오 코덱(video codec)으로부터 생성된 동영상에 사용자가 볼 수 있도록 클라이언트들로 전송한다. 동영상 흐름의 정보량은 매우 크므로 모든 제어 정보가 중앙 서버를 거치는 것보다 달리 카메라 제어기로부터 직접 클라이언트로 전송된다. 하나의 지역을 여러 사용자가 시간대 별로 나뉘어 동시에 관찰할 수 있으므로 카메라 제어기는 여러 클라이언트와 동시에 연결될 수 있어야 한다. 중앙 서버는 모든 카메라의 시선 방향을 정하는 제어 명령을 생성하여 전달한다. 각각의 사용자의 클라이언트는 모두 중앙 서버에 연결되며 중앙 서버는 연결된 클라이언트들로부터의 요구를 받아들여 적절하게 카메라 제어기에 제어 명령을 전달한다. 사용자들의 요구를 충돌 없이 처리하기 위해서는 중앙 서버가 카메라의 소유권이 특정

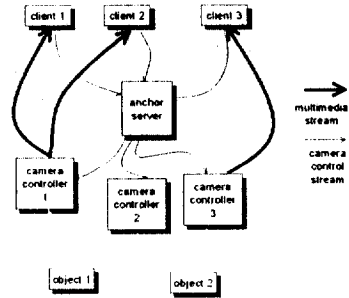


그림 1: 본 연구의 목표 시스템 전체 구조

사용자에게 속하도록 해야 하며, 또한 최대한 많은 사용자들의 요구를 만족시킬 수 있도록 설계되어야 한다. 클라이언트는 사용자에게 카메라 제어로 부터 전달된 동영상을 보여주며, 사용자가 카메라 시선을 움직이려 하는 것에 대한 요구를 중앙 서버에 전달한다.

2.2 객체의 위치 설정 방법

카메라 앵커를 처리하기 위해서 객체의 위치 정보를 표현하는 방법과 이에 따른 위치 측정 방법이 선행되어야 한다. 객체의 위치 측정 방법은 카메라 조절 변수 저장법과 삼각측량법을 이용한 3차원 위치 측정 방법을 생각해 볼 수 있다. 본 절에서는 각각의 방법과 그 장단점을 살펴 본다.

2.2.1 카메라 조절 변수 저장법

본 방법은 등록할 객체의 위치 정보를 렌즈 방향, 확대/축소 정보와 같은 카메라의 조절 변수로써 저장하는 것이다. 이것은 실제로 사용자가 카메라의 시선 이동을 통해 원하는 객체를 선택하는 경우에 객체의 선택은 하나의 카메라 조절 변수 설정과 대응하는 것에 착안한 것이다. 객체의 등록은 사용자가 보게 될 화면과 똑같은 화면을 관리자가 보는 상태에서 그때의 카메라 조절 변수로 이루어진다. 단 사용자가 반드시 화면의 중앙에 있는 객체를 선택하지는 않으므로 객체의 등록과 검색을 위해서는 선택된 객체가 화면 중앙으로 오게 하는 카메라 조절 변수 변환을 필요로 한다. 객체 등록 시에 저장되는 카메라 조절 변수는 P (Panning), T (Tilting), Z (Zooming) 값이며, 이 값을 바탕으로 카메라의 방향성(orientation)을 (P, T, Z) 로 정한다. 사용자가 카메라가 (P, T, Z) 에 있는 상태에서 화면의 (X, Y) 에 있는 객체를 선택하면, 본 방식은 이것을 화면의 중앙 (C_x, C_y) 로 옮긴 것에 해당하는 (P', T', Z') 으로 등록된 객체를 데이터베이스에서 찾게 된다. 이러한 방식은 객체의 실제 3차원 좌표를 정확하게 측정하기 어려운 경우에 효과적으로 사용할 수 있다. 반면, 카메라를 추가하는 경우에 하나의 카메라를 통해 입력된 객체를 다른 카메라에서는 어떻게 사용할 것인지가 어려운 문제를 갖는다.

2.2.2 삼각측량법을 이용한 3차원 위치 측정 방법

카메라 영상은 2차원 영상이므로 카메라 영상만을 이용하여 3차원 물체의 위치를 측정하는 것은 불가능하다. 그러