

# 비디오 객체 정의에 의한 동적 증강 정보 저작

## Authoring of Dynamic Information in Augmented Reality Using Video Object Definition

남양희, 이서진

이화여자대학교 디지털미디어학부

Yang-Hee Nam(yanghee@ewha.ac.kr), Seo-Jin Lee(moon0329@ewhain.net)

### 요약

증강현실에서 움직이는 가상 모델을 증강물로 활용하려면 대개 별도의 모델링이나 애니메이션 도구 사용이 필요하며 이는 매우 전문적이고 복잡한 작업이다. 본 논문에서는 그러한 과정 없이 동적 가상 객체를 증강 환경에 삽입할 수 있는 방법으로 비디오 객체 기반 저작 방법을 제시한다. 그래프컷과 그로컷을 통합 적용하여 대상물의 초기 영역을 분리한 후, 스냅컷을 이용해 프레임간 경계선 변화를 자동 추적하여 동작이 담긴 연속 프레임을 실사 비디오에 증강하는 방식이다. 실험결과로, 몇 단계의 메뉴 선택과 경계선 오류 정정 스케치만으로 특정 비디오 객체 컷아웃 및 증강 객체 저작이 가능함을 보였다.

■ 중심어 : | 증강현실 | 저작도구 | 비디오 컷아웃 | 그래프컷 | 그로컷 | 스냅컷 |

### Abstract

It is generally required to use modeling or animation tools for inserting dynamic objects into augmented reality, and this process demands high expertise and complexity. This paper proposes a video object based authoring method that enables augmentation with dynamic video objects without such process. Integrated grab-cut and grow-cut method strips initial area of video target off the existing video clips, and snap-cut method is then applied to track objects' boundaries over frames so as to augment real world with continuous motion frames. Experiment shows video cut-out and authoring results achieved by only a few menu selections and simple correcting sketch.

■ keyword : | Augmented Reality | Authoring Tool | Video Cut-Out | Grap-Cut | Grow-Cut | Snap-Cut |

## 1. 서론

증강현실(Augmented Reality, AR)은 2008년 이후 줄곧 세계적으로 유망한 10대 기술 중 하나로 선정되면서, 스마트폰 사용의 일반화로 인해 급성장하는 분야이다. 현실세계를 인식하고 그에 맞는 가상의 정보를 덧

입혀 보여주는 증강현실 기술은 여러 위치 기반 서비스에 이용되고 있으나 증강의 형태가 대개 2차원 레이어의 중첩 형태로 비교적 단순로운 편이다. 이는 증강 콘텐츠의 저작 및 설정 단계에서 실물 공간과 가상 객체 간 관계 정의의 어려움 때문에 대개 단순한 오버레이 유형의 가상정보 증강을 지향하기 때문이다.

\* 본 연구는 2012년 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구입니다.

(NRF-2012S1A5B6034337)

접수번호 : #130429-004

접수일자 : 2013년 04월 29일

심사완료일 : 2013년 05월 08일

교신저자 : 남양희, e-mail : yanghee@ewha.ac.kr

이런 문제를 개선하고자 실제 3차원적 증강을 목표로 하는 증강현실 콘텐츠를 제작할 수 있는 저작도구들도 활발히 연구되고 있으나, 대개 프로그래밍 지식을 요구하는 스크립트 언어 기반이어서 프로그래밍 지식이 없는 서비스 개발자나 도메인 전문가들이 사용하기는 어렵다[1][2].

한편, 컴퓨터 관련 지식이 깊지 않은 일반 사용자들이 보다 간단한 상호작용적 조작을 통해 증강 정보를 설정할 수 있도록 하는 저작도구들도 일부 연구되고 있는데, 3차원 모델이나 동적 객체는 미리 애니메이션 및 모델링 도구 등을 통해 구축된 파일을 불러와서 배치하는 방식을 사용한다. 그러나 이와 같은 동적 객체 제작 방식은 모델링 도구에 대한 전문적 지식이 필요하며, 시간과 노력이 소요되는 어려운 작업이다[3][4].

본 논문에서는 프로그래밍 지식이 없는 일반 서비스 개발자나 도메인 전문가들이 동영상 형태로 이미 존재하는 비디오 파일이나 촬영된 비디오로부터 관심 객체를 정의하여 증강현실의 가상 객체로 활용할 수 있게 하는 저작 방식을 제안한다. 즉, 모델링 도구나 애니메이션 편집 도구를 이용하는 복잡한 제작방법이 아닌, 사용자 중심의 간단한 상호작용에 의한 저작 방식을 통해 동적인 증강 객체의 설정을 가능하게 하는 것이다.

## II. 증강현실 콘텐츠 저작

증강현실 저작도구는 크게 개발자용 저작도구[1][2]와 중급 및 일반 사용자를 위한 도구[3][4]로 구분된다. 개발자용 도구는 대개 프로그래밍 지식이 상당히 요구되는 개발 환경이라 할 수 있고, 중급 및 일반 사용자 도구들은 스크립트형 언어로 원하는 콘텐츠를 제작하거나, 기본 조작단위 컴포넌트들을 결합하여 콘텐츠를 구축한다. 그러나, 컴포넌트들을 메뉴 형식으로 로드하고 프로그래밍 언어로 작성된 매개변수들을 저장 보관하는 등의 일련의 저작 방식이 초보 사용자들에게는 쉽지 않다.

저작 내용은 객체 증강이 발생하는 조건의 정의와, 해당 조건에 부합하는 맥락에서 증강될 가상 객체를 그

림, 텍스트, 사운드, 3D 모델, 애니메이션 및 비디오 등으로 정의하는 것으로 구성된다. 이 때, 가장 간단한 증강 형태는 [그림 1](a)와 같이 비디오 레이어의 중첩을 통한 오버레이 유형 증강이며, 이는 GPS 등의 위치 신호를 기반으로 정해진 이미지 또는 비디오를 적당한 투명도를 두어 증강하는 것이어서 저작방법도 간단하고 실시간 재현기술 역시 간단하여 모바일 기기에서 상용화된 앱 형태로 종종 접할 수 있다.



(a) 비디오 오버레이형 증강기 (b) 3차원 객체 증강  
그림 1. 증강현실의 객체 증강 유형

두 번째 유형은 [그림 1](b)와 같이 실물 환경에 3차원적 객체를 증강하는 것인데, 이 때 가상 객체가 실사 환경의 객체에 대해 특정의 공간적 상관관계를 지니지 않는다면 오버레이 유형과 크게 다를 것이 없다. 그러나, 특정 공간관계를 만족시켜야 한다면 삽입되는 가상 객체와 실사 환경 간의 배치 관계 정의가 중요한 문제이며 실물-가상 간 통일된 좌표 수립이 필요하고, 이를 위해 실사 환경 비디오로부터의 3차원 공간 좌표 복원에 많은 연구들이 집중되어왔다[5][6]. 그러나 실사 환경의 3차원 복원이 이루어진다고 하더라도 증강될 가상 객체의 3차원 형상 정의가 쉽지 않으므로 간단한 큐브나 구를 제외하면 3차원 모델링 전문 도구를 별도로 사용해야 하며, 만들어진 모델 파일을 증강현실 저작환경 내로 가져와 3차원 배치를 수행하게 된다.

또한, 일반적으로 증강현실이 현실세계의 자연스러운 시간의 흐름을 반영하므로 가상 객체 역시 시간에 따른 움직임이나 변형의 정의가 요구되기도 하는데, 가상환경에서 동적 객체를 제작하는 것은 쉽지 않다. 기존 연구에서는 별도의 애니메이션 전문 도구에서 미리 구축한 3차원 애니메이션 파일을 증강현실 저작도구로 불러오거나[1-4], 저작 도구 내에서 스크립트 언어를 기반으로 동적인 움직임을 지정하는 방법[8]이 있다.

전자의 방법은 애니메이션이나 3차원 모델링 전문

도구를 사용할 줄 알아야 하며, 후자의 경우는 일일이 객체의 부위마다 모션을 지정해야 하는 번거로움과 스크립팅 지식이 요구되며, 모션 캡처 데이터가 있는 모션만 적용 가능한 경우가 많다. 결과적으로, 이 두 방법 모두 움직임의 구축하기 위한 시간과 노력, 전문성이 복합적으로 요구되는 어려운 작업을 거쳐야 한다.

이와 같이 가상환경에서 완전한 3차원의 형상과 움직임으로 동적 객체를 저작하는 것은 비용과 노력, 요구되는 전문성의 수준이 꽤 높다. 컴퓨터에 대한 전문성이 없는 일반 교사들의 사용을 고려하여 기존 증강현실 저작도구들을 분석한 최근 연구[9]에서도 거의 모든 저작도구가 프로그래밍이나 3D 모델링 도구 사용 능력을 요구하는 것으로 나타났다.

반면, 3차원 객체는 아니지만 일반인이 쉽게 사용할 수 있는 방식으로 비디오에서 동적 객체를 추출하는 방법[10][11]들이 비디오 편집 분야에서 연구되어 왔는데, 이러한 방법을 이용하면 2차원 비디오의 연속 프레임으로부터 동적 객체를 추출해낼 수 있으므로, 이에 기반하여 증강 객체 저작을 고려할 수 있다.

비디오 기반의 증강객체 저작을 위한 기본 요구사항은 첫째, 비디오 프레임 중 한 곳(시작 프레임)에서 실사 공간에 삽입할 대상물, 즉, 관심 객체를 쉽게 정의할 수 있을 것, 둘째, 정의된 대상물(비디오 객체)을 연속 프레임에서 자동 추적할 것, 셋째, 분리된 동적 객체의 연속 프레임을 실사 공간의 3차원 좌표계 안에 삽입하고 움직임이 재현되도록 해야 하는 것이다.

다음 장에서는 그러한 방법의 하나로서, 간단한 사용자 상호작용을 통한 동적 비디오 객체의 추출과 실사 공간 내 2½ 차원 객체로서의 삽입 기술을 제시한다. 이는 [그림 1]에서 보였던 두 가지 증강 방식, 즉, 비디오 오버레이와 완전한 3차원 객체 증강의 중간 유형이라 할 수 있다.

### III. 비디오 기반 동적 객체 정의 알고리즘

본 논문에서 제안하는 동적 증강 객체 정의 방법은 사용자가 간단히 정의한 관심영역을 기반으로 비디오

영역을 분할하여 관심 객체를 추출하는 단계와, 분할된 객체 영역을 연속 비디오 프레임에서 자동 추적하는 단계, 그리고 추출된 연속 비디오 프레임 객체를 실사 환경 속에 3차원적으로 삽입하는 단계로 나누어 상세 방법을 각 절에서 설명한다.

#### 1. 관심 영역 기반 객체 분할

비디오에서 원하는 객체가 등장한 영상 시퀀스의 첫 프레임에 대해서는 관심 객체의 정확한 추출이 중요한데, 사용자가 일일이 객체 아웃라인을 정의하는 경우 잘못된 픽셀을 객체 범위에 포함하게 될 수 있다. 오류 픽셀의 존재 여부가 이후 연속 프레임의 추적 결과에 영향을 주기 때문에 최대한 정확한 객체 추출을 할 수 있는 객체 분할 방법이 필요하다.

객체가 포함된 프레임 하나만 놓고 볼 때는 정적인 이미지 분할 문제와 동일한데, 이 문제에 대해서 제시되어 온 여러 방법 가운데 먼저, 간단한 사용자 상호작용을 시작으로 전경과 배경의 분할을 반복적으로 개선하여 좋은 분할 성능을 보이는 것으로 알려져 있는 그랩컷(GrabCut)[13]을 적용해보았다. 그러나, 실험 결과 초기 최적화 결과에서 잘 못 분할된 영역의 경우 사용자의 추가적 전경/배경 브러싱을 여러 차례 거쳐야 하는 경우가 많음을 관찰하였다.

이에 보정 단계에서는 셀룰러 오토마타를 통한 멀티 레이블 분할법인 그로컷(GrowCut)[14]을 통합적으로 적용하였다. 그로컷의 적용은 전경과 배경 분리의 정확도를 높이는 외에도, 비디오 장면에서 추출하고자 하는 비디오 객체가 둘 이상 존재하며 이들이 프레임 상에서 서로 분리된 움직임을 보이는 경우에도 관심 객체 분할이 가능하므로 여러 비디오 객체의 동시 추출을 통한 증강 개체로의 활용을 도울 수 있는 방법이다.

이 두 방법은 모두 기존 연구에서는 정적인 이미지의 편집을 주목적으로 하여 사용되어온 방법들이지만 비디오의 초기 프레임 객체 분할에 이들을 통합 적용함으로써 관심 객체 정의를 위한 사용자 상호작용은 최소화하면서 분할 정확도를 높이고자 하였다.

이상과 같은 두 방법을 활용하여 비디오 프레임으로부터 관심 영역 정의를 통해 객체 영역을 분할하는 대

략적인 방법은 [그림 2]와 같다. 즉, 그래픽스에 의한 대략적 분할 단계와 그로켓에 의한 상세 보정 분할의 두 단계로 이루어진다.

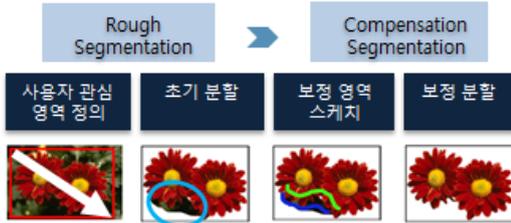


그림 2. 관심 영역 기반 객체 분할 과정

대략적 분할 단계에서는 먼저 사용자가 관심 영역을 정의하는데, 이 때 관심 영역(Region of Interest, ROI) 이란 2차원 비디오 영상 속 객체의 실루엣을 추출하기 위해 객체 주변을 대략적으로 감싼 영역을 의미한다. [그림 2]의 첫 단계 그림 표시와 같이 해당 객체를 포함하는 사각형 영역을 마우스로 드래깅하는 손쉬운 사용자 조작을 통해 대략적 관심 영역을 정의한다.

이렇게 선택된 사각형 영역 내부에 대해 [그림 3]과 같이 전경(TF), 배경(TB) 그리고 알 수 없는 영역(TU)으로 구성된 3가지 영역의 맵(trimap)을 형성하고 전경과 배경, 각각에 대한 칼라 분포의 가우시안 확률 모델(Gaussian Mixture Model, GMM)을 구축한다.

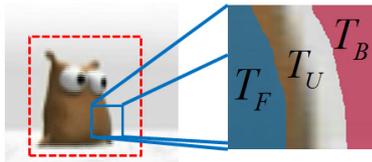


그림 3. 이미지 영역 분할 과정의 영역 맵(trimap)

사용자가 대략적으로 스케치한 영역은 모두 배경(TB)으로 간주하여 배경 픽셀  $n$ 의 레이블  $L_n=0$ , 나머지는 모두 알 수 없는 영역(TU)으로  $L_n=1$ 로 초기화한 후, 픽셀들의 색상 값( $z$ )으로부터 컬러분포를 나타내는 가우시안 확률 모델(Gaussian Mixture Model; GMM) 매개변수  $\theta$ 를 학습시킨다. GMM은  $K$ 개 컴포넌트(대개  $K=5$ )로 표현되며 이 컴포넌트 변수인  $k$ 를 TU 영

역에 속하는 픽셀마다 구하게 된다. 그런 후, 식 (1)의 에너지  $E$  (Gibbs energy)[13]가 최소화될 때까지 위 과정을 반복하면서 TU에 속하는  $L_n$ 의 레이블을 점차 결정하여 영역을 분할하게 된다.

$$\min_{\{L_n : n \in TU\}} \min_k E(L, k, \theta, z) \quad (1)$$

이와 같은 방법으로 대략적 영역 분할 결과 분할이 제대로 되는 경우도 있지만, 이미지의 전경과 배경의 칼라가 유사할 경우 [그림 4]와 같이 제대로 분할되지 않은 경우가 발생한다. 청색으로 칠해진 부분은 객체 영역(전경)으로 분리된 영역이며, 나머지 객체로 선택되지 않은 부분인데, 한쪽 눈의 경계가 배경색과 유사하여 객체 분할이 제대로 되지 않은 예이다. 이러한 오류를 보완하기 위해 사용자의 간단한 스케치를 거쳐 그로켓에 의한 2차 보정 분할을 수행한다.



(a) 1차 분할 오류 영역



(b) 보정 분할을 위한 사용자 스케치와 분할 결과

그림 4. 1차 분할 오류와 보정 분할

2차 분할은 전경과 배경 영역 각각의 일부를 사용자가 [그림 4](b)의 왼쪽과 같이 마우스로 간단히 스케치하는 상호작용을 통해 이루어진다. 초록색 스케치는 배경 영역을, 붉은 색은 전경 영역을 간단히 지정한 것이다. 이 때 스케치를 통해 지정된 분리 영역의 픽셀 값을 기초로 선택된 영역에 이웃하는 픽셀들의 칼라 값과의 유사도를 계산하여 선택된 객체인지(1) 아닌지(0) 재차 레이블링(labeling)한다. 이 과정은 1차 분할을 거친 후 사용자가 지정한 매우 적은 영역에 대해서 행해지

로 전역적인 분할 계산에 비해 계산 시간이 절약된다.

## 2. 지역 윈도우에 의한 경계선 자동 추적

관심 객체가 등장하는 시작 프레임에서 1절과 같이 관심 영역의 분할을 통한 객체 정의가 이루어지면, 이후 연속되는 프레임에서 해당 객체를 추적하여 지속적인 객체 분할을 수행해야 한다. 이를 위한 알고리즘은 전역적 장면의 색상 통계를 사용하는 방법들 대신, 복잡하고 다양하게 구성된 장면들 속에서 원하는 객체를 잘라낼 수 있게 하는 지역적 윈도우 기반의 비디오 스탬프 알고리즘을 참조하였다[15].



그림 5. 지역 윈도우 기반 객체 추정

즉, 1절의 방법으로 추출된 관심 객체의 경계선을 따라 [그림 5]의 맨 왼쪽 상단과 같이 일정 간격으로 윈도우(30x30 크기)  $W_t$ 를 생성한 후, 윈도우 각각에서 칼라 분포, 형태, 그리고 움직임(motion) 정보를 이용하여 객체 영역을 분석한다. 이 때, 지역적 색상 GMM 모델에 기반한 픽셀  $x$ 의 전경 확률( $p_c(x)$ )과, 해당 색상 모델이 전경과 배경을 구분하게 하는 정도를 나타내는 색상 모델의 신뢰값( $f_c$ ), 그리고 지역적 형태 확률 모델( $M_s$ )로 구성된 윈도우 분류기를 사용한다. 형태 확률 모델은 앞 절에서 이미 구해진 분할 결과를 기반으로 한 분할 마스크( $L^t(x)$ )를 사전 확률로 사용하고, 객체 경계선으로부터의 거리에 비례하는 형태 신뢰도 마스크( $f_s(x)$ )를 두는데, 칼라 확률과 형태 신뢰도 마스크, 형태 사전 확률 모델의 곱(교집합)이 최종적으로 지역 윈도우  $k$ 내에서 픽셀  $x$ 가 객체 안에 포함될 확률

( $p_F^k(x)$ )을 표현하게 된다.

시간  $t$ 에서의 지역적 객체 분류가 이상과 같이 지역 윈도우 분류기를 통해 계산되는 한편, 프레임 간의 지역 윈도우 전과를 통해 장면 간 모션을 추정한다. 이 때 연속 프레임  $I_t, I_{t+1}$  간에 SIFT 특징점들[16]의 정합으로 전역적 어핀 변형(affine transform)  $A$ 를 계산하여 커다란 강체 모션을 추정하며, 광류(optical flow) 추정에 의해 지역적 변형을 파악한다. 광류 추정은 객체 일부가 가려지는 경우 부정확한 결과를 내는 것으로 알려져 있지만, 여기서는 이미지  $I_t$ 에 대해 위에서 구한 어핀 변형  $A$ 를 적용하여 계산된  $I'_{t+1}$ 과 실제의 다음 프레임 이미지  $I_{t+1}$  간에 특정 경계 픽셀(contour pixel) 위치 중심의 평균적 플로우 벡터를 구함으로써 비교적 안정적인 광류 추정이 가능하다.

즉, 어핀 변형으로 구해진  $I'_{t+1}$ 의 경계선  $C'_{t+1}$ 상의 각 점  $x'$ 에 대해,  $x'$ 를 중심으로 하는 윈도우  $N'_{x'}$ 와  $C'_{t+1}$ 의 내부 영역(객체로 추정된 부분) 간 교집합 안에서의 영역 평균 광류 벡터  $\overline{v_x}$ 를 구하여, 식(2)와 같이 교정된 새 위치  $x^*$ 를 얻는다.

$$x^* = x' + \overline{v_x} \tag{2}$$

이와 같이 윈도우가 다음 프레임으로 전파될 때마다 앞서 언급한 지역적 윈도우 분류기의 색상 모델 및 형태 모델을 갱신하는데, 여기서 얻은 경계선의 교정 위치들이 구성하는 새로운 객체 추정 영역값은 형태 마스크  $L^t(x)$ 를 갱신하는 데 사용된다. 즉, 객체가 등장하는 첫 프레임의  $L^1(x)$ 는 앞 절에서 구한 객체 영역 분할값을 마스크로 사용하지만, 이후 프레임에 대해서는 마스크가 정의되어 있지 않으므로 본 절에서 구한 모션 추정을 적용함으로써 갱신된 이전 마스크  $L^{t+1}(x)$ 를 만들게 된다.

이상과 같이, 지역적 윈도우 기반 분류 결과에 덧붙여 장면 간 전역적 움직임과 지역적 객체 움직임의 전과 영역을 통합적으로 고려함으로써 연속 프레임의 객체를 추적한다.

### 3. 비디오 객체의 3차원 실물 공간 삽입

2차원 비디오의 시퀀스로부터 추출한 동적 비디오 객체를 증강 객체로 사용하기 위해 실물 비디오로부터 PTAM[5]을 이용하여 복원한 3차원 공간 맵의 좌표에 따라 동적 객체를 삽입한다. 실물 공간의 3차원 복원 자체는 본 논문의 주제와 무관하여, PTAM 이외의 다양한 3차원 복원 방법을 사용해도 무방하다.

비디오 객체의 3차원 실공간 삽입은 (A) 실공간의 3차원 복원시 구해져 있는 3차원 좌표 공간에 위치, 방향 설정하여 배치하고, (B) 증강현실 프레임들과 비디오 객체의 연속 시퀀스 플레이 시간의 매핑을 정하는 과정으로 구성된다. 이 때, 비디오 객체는 볼륨을 갖지 않으므로 [그림 6]과 같이 2<sup>1</sup>/<sub>2</sub> 차원의 3차원 내 평면 객체로서 삽입된다.

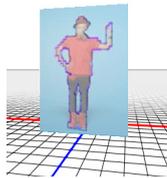


그림 6. 2<sup>1</sup>/<sub>2</sub> 차원 비디오 객체의 3차원 공간 삽입

## IV. 구현 및 실험 결과

본 논문의 구현시스템은 윈도우 XP Professional 2.66GHz, 3.25GB RAM, NVIDIA GeForce GTS 250인 PC에서 구현되었다. 카메라는 초당 30프레임 처리 속도의 HD 센서가 장착된 웹캠을 사용하였고, Visual C++ 2008 환경에서 영상처리 라이브러리 OpenCV, OpenGL를 사용하여 구현되었으며, 실험영상의 전체 해상도는 640X480이다.

[그림 7]은 비디오 불러오기 및 실행, 멈춤, 관심영역 지정 및 분할 각 과정을 단계별로 실행하고 결과를 저장할 수 있는 구현 화면인데, 장면 단위 분할 결과를 관찰하면서 첫 장면과의 많은 차이로 분할 오류가 커지는 프레임에서는 멈춤 버튼을 누르고 사용자가 보정 영역 스케치를 수행할 수 있게 하여 결과 재개선이 가능하도

록 설계되었다. [그림 8]은 비디오 객체의 분할 과정 실험 예이다.

비디오 동적 객체 분할 과정을 거쳐 저장 버튼을 누르면 각 프레임에서 객체 마스크 영역( $L^t(x) = 1$ 인 부분)을 제외한 평면 부분을 투명 처리한 객체 평면이 삽입되어, 증강현실의 실사 장면으로 입력되는 실시간 비디오의 복원된 3차원 좌표에 따른 시점 투영을 거쳐 실사와 함께 렌더링 된다.



그림 7. 비디오 동적 객체 분할 시스템



그림 8. 비디오 동적 객체 정의과정 실험





그림 9. 분할 비디오 객체가 삽입된 증강현실

동적 객체 정의 처리는 오프라인으로 이루어지므로 시간이 크게 중요하지는 않지만 연속 프레임의 동적 비디오 객체 추적의 실험 결과 각 단계별 소요시간이 평균적으로 [표 1]과 같았다. 지역 윈도우 배분 및 윈도우별 색상과 형태 모델 수립에 걸리는 시간과 두 프레임 간 SIFT 특징점 정합에 의한 어핀 변형 구하는 시간이 거의 비슷하게 소요되었고, 어핀 변형 적용에 의해 예측된 이미지와 실제 이미지 간 광류 계산 및 객체의 지역적 움직임에 의한 프로우 벡터 추정 시간을 통한 경계선 추출 시간이 가장 오래 소요되었다.

표 1. 동적 객체 정의 처리 시간

처리 과정	1프레임 당 처리 시간 (약 29 frame/sec)
전처리	약 0.1050 sec
두 프레임 간의 특징점 매칭 후 매트릭스 구하기	약 0.156 sec
경계선 추출	약 0.453 sec
총	약 0.714 sec

증강현실에 증강 객체로 삽입된 후의 렌더링 소요 시간은 사실상 비디오에 평면 하나가 삽입된 후, 비디오 객체의 프레임율과 증강 환경의 실시간 간 플레이 시간 관계 설정에 따라 해당 평면의 텍스처만 다음 비디오 프레임으로 변경하면 되므로 전체적인 프레임 율에 거의 영향을 주지 않는다.

2장에서 제시했던 비디오 기반 동적 증강물 저작의 요구사항에 따른 작업과정 단순화 결과는 아래 [표 2]와 같이 증강객체로 사용할 가상 객체의 모델링에 소요되는 사전 저작 부분의 시간 단축으로 설명된다.

표 2. 사용자의 동적 객체 정의 작업 과정

동적 증강객체 저작 과정	모델링 기반의 기존 저작방법	제안한 방법
사전 저작	모델링, 애니메이션 편집 (수 시간~수 일 소요)	비디오 선택, 메뉴 클릭 (수 초~수 분)
실세계 비디오의 3차원 좌표 복원 및 추적	증강현실 시점추적 및 복원 기법 사용 (PTAM 등)	기존과 동일
실세계 비디오에 증강객체 삽입, 배치	Layout 도구 사용	기존과 동일

## V. 결론 및 향후 연구

본 논문에서는 증강현실 여러 저작 기능 중에서 특히 동적 증강 객체 저작에 대해 비디오 기반의 쉬운 방법을 제안하였다. UCC(사용자 생성 비디오 콘텐츠) 등 쉽게 제작이 가능하고 구하기 쉬운 비디오 클립들을 증강현실 속의 가상 객체로 활용할 수 있게 해줌으로써, 복잡한 모델링 도구나 전문 지식이 필요한 애니메이션 편집도구를 사용하지 않아도 되는 장점이 있다.

현재는 비디오 객체를 실사 공간 좌표계 안에 고정된 평면으로서 세워놓고 투명한 배경으로 렌더링되는 평면 속에 비디오의 동적 전개에 따라 객체 부분 텍스처 매핑만 변경하여 시작 위치에서 동작만 변경되는 한계점이 있으나, 향후 연구에서는 간단한 경로 지정 상호작용 등을 추가하여 평면 자체의 이동도 가능하게 함으로써 좀 더 다양한 동적 증강을 가능하게 할 것이다.

## 참고 문헌

- [1] <http://www.hitl.washington.edu/artoolkit/>
- [2] <http://www.artoolworks.com/community/osgart/>
- [3] J. Jung, J. Hong, S. Park, and H. S. Yang, "Smart phone as an Augmented Reality Authoring Tool via Multi-touch based 3D Interaction Method," VRCAI'12, pp.17-20, 2012.
- [4] 이정규, 이종원, "모바일 환경 증강현실 콘텐츠

현장 저작 인터페이스”, 한국콘텐츠학회논문지, 제10권, 제7호, pp.1-9, 2010.

[5] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," ISMAR'07, pp.225-234, 2007.

[6] R. A. Newcombe and A. J. Davison, "Live dense reconstruction with a single moving camera," CVPR, pp.1498-1505, 2010.

[7] N. Braun, "Storytelling in Collaborative Augmented Reality Environment," WSCG Short papers proceedings, 2003.

[8] M. G. Choi, K. H. Lee, and J. H. Lee, "Motion Patches," Journal of KISS: Computer Systems and Theory, Vol.33, No.1\_2, pp.119-127, 2006.

[9] N. Saforrudin, H. B. Zaman, and A. Ahmad, "Technical Skills in Developing Augmented Reality Application: Teachers' Readiness," Visual Informatics: Sustaining Research and Innovations-LNCS, Vol.7067, pp.360-370, 2011.

[10] X. Bai and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," Proc. of IEEE ICCV, pp.1-8, 2007.

[11] C. J. Armstrong, B. L. Price, and W. A. Barrett, "Interactive segmentation of image volumes with live surface," Computers and Graphics, Vol.31, No.2, pp.212-229, 2007.

[12] A. van den Hangel, R. Hill, B. Ward, and A. Dick, "In Situ Image-based Modeling," ISMAR 2009, pp.107-110, 2009.

[13] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut - interactive foreground extraction using iterated graph cut," Proceedings of ACM SIGGRAPH, pp.309-314, 2004.

[14] V. Vezhnevets and V. Konouchine, "GrowCut: Interactive Multi-Label N-D Image Segmentation By Cellular Automata," pp.150-156, proc. of Graphicon 2005.

[15] X. Bai, J. Wang, D. Simons, and G. Sapiro, "Video Snapcut: robust video object cutout using localized classifiers," ACM Transactions on Graphics, Vol.28, No.3, pp.70:1-70:11, 2009.

[16] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," I. J. of Computer Vision, Vol.60, No.2, pp.91-110, 2004.

저 자 소 개

남 양 희(Yanghee Nam)

정희원



- 1989년 2월 : 이화여자대학교 전자계산학과(이학사)
- 1991년 2월 : KAIST 전산학과(공학석사)
- 1997년 8월 : KAIST 전산학과(공학박사)

▪ 2002년 8월 ~ 현재 : 이화여자대학교 디지털미디어학부 부교수

<관심분야> : 인터랙티브 미디어, 증강현실, HCI

이 서 진(Seo-Jin Lee)

준희원



- 2003년 2월 : 한국기술교육대학교 디자인공학과 졸업(공학사)
- 2010년 8월 : 이화여자대학교 디지털미디어학부 미디어공학 졸업(공학석사)

<관심분야> : HCI, 모바일 미디어, 영상처리