

객체기반 비디오 편집 시스템을 위한 불확실 영역기반 사용자 지원 비디오 객체 분할 기법

유홍연[†], 홍성훈[‡]

요 약

본 논문에서는 객체기반 비디오 부호화 또는 멀티미디어 편집을 위한 반자동 비디오 객체 분할방식을 제안한다. 반자동 객체분할은 사용자 지원에 의한 분할 방식으로, 비디오 시퀀스의 초기 프레임에서 사용자가 관심객체의 경계를 표시하고 이후의 영상 프레임의 객체를 배경으로부터 연속적으로 분리해 낸다. 제안된 방식은 부분적으로 사용자 조력에 의한 프레임내 분할과 완전 자동에 의한 프레임간 분할 처리과정으로 구성되는데, 영상 전체에 대해 연산을 수행하는 기존 방식과는 달리 객체 경계가 존재하는 영상영역 부분에서만 연산을 수행한다. 프레임내 분할은 사용자가 관심객체의 경계를 지정하고, 이 경계 주위 화소들의 유사성을 이용한 후처리에 의해 정확한 초기 객체를 구한다. 프레임간 분할에서는 이전 프레임에서 추출한 객체의 경계 정보에 근거하여 시간적 유사성을 구한 후 경계와 영역 추적에 의해 연속적으로 동영상 객체를 추출한다. 실험결과로부터 제안된 방식은 비디오 편집, 객체기반 비디오 압축 및 인덱싱 등의 멀티미디어 응용에 사용 가능할 정도로 안정되고 정확한 객체추출을 수행함을 확인하였다. 이 결과를 바탕으로 다수의 편리한 기능을 포함한 비디오 편집시스템을 개발하였다.

Uncertain Region Based User-Assisted Segmentation Technique for Object-Based Video Editing System

Hong-Yeon Yu[†], Sung-Hoon Hong[‡]

ABSTRACT

In this paper, we propose a semi-automatic segmentation method which can be used to generate video object plane (VOP) for object based coding scheme and multimedia authoring environment. Semi-automatic segmentation can be considered as a user-assisted segmentation technique. A user can initially mark objects of interest around the object boundaries and then the selected objects are continuously separated from the unselected areas through time evolution in the image sequences. The proposed segmentation method consists of two processing steps: partially manual intra-frame segmentation and fully automatic inter-frame segmentation. The intra-frame segmentation incorporates user-assistance to define the meaningful complete visual object of interest to be segmentation and decides precise object boundary. The inter-frame segmentation involves boundary and region tracking to obtain temporal coherence of moving object based on the object boundary information of previous frame. The proposed method shows stable and efficient results that could be suitable for many digital video applications such as multimedia contents authoring, content based coding and indexing. Based on this result, we have developed objects based video editing system with several convenient editing functions.

Key words: Moving Object Segmentation(이동객체분할), Semi-automatic Segmentation(반자동객체분할), User-assistance(사용자조력), Object Tracking(객체추적), Video Editing(비디오 편집)

* 교신저자(Corresponding Author): 유홍연, 주소: 광주광역시 북구 용봉동 300(500-757), 전화: 062)530-0820, FAX: 062)530-1759, E-mail: keister@vip.chonnam.ac.kr
접수일: 2005년 6월 14일, 원료일: 2006년 3월 29일

[†] 전남대학교 전자정보통신공학과 박사과정

[‡] 정희원, 전남대학교 전자컴퓨터정보통신공학부 조교수
(E-mail: hsh@chonnam.ac.kr)

* 본 연구는 2003년도 전남대학교 학술연구비 지원으로 수행되었음.

1. 서 론

최근 방송, 통신, 컴퓨터 기술을 융합한 디지털 양방향 멀티미디어 방송서비스가 가시화되면서 기존의 MPEG-1, MPEG-2 등의 부호화 방식에서 벗어나 내용기반 부호화 방식인 MPEG-4에 대한 관심이 높아지고 있다. 국제표준인 MPEG-4는 기존의 프레임 단위의 부호화 방식과 다르게 VOP(video object plane)라는 개념을 이용하여 동영상내에 존재하는 움직이는 비디오 객체나, 관심있는 비디오 객체의 형태를 기술할 수 있는 객체분할 과정을 요구한다. 이러한 객체분할 과정은 MPEG-4 뿐만 아니라, MPEG-7 등에서 사용자 질의어에 의한 동영상내의 특정 내용을 검색하기 위해서 객체에 대한 형태를 정의하고 이를 추적 분할하는데 필요한 중요한 과정이다[1,2]. 또한 DVR(digital video recorder)과 같이 움직임을 갖는 물체를 검출 추적하여 기록하는 감시시스템 및 특정 상표의 노출방지나 초상권 보호를 요구하는 방송용 프로그램 제작에서 선행되는 요구사항이다.

일반적으로 동영상내에서 배경으로부터 객체를 분할하기 위해 공간상(spatial)에 존재하는 밝기값(intensity)정보나, 시간축(temporal)의 움직임(motion) 정보를 척도로 사용자의 개입유무에 따라 자동 및 반자동으로 객체분할을 수행한다[3-7].

자동 객체분할은 이전 프레임과의 밝기값 차이를 가우시안 분포로 모델링하여 임계값을 설정하고 움직임이 발생하는 영역에 대한 변화검출마스크(change detection mask)를 생성하여 분할한다[8-9]. 그러나 객체 경계에서 밝기 변화로 기인한 잡음 성분의 영향으로 신뢰성 있는 분할 결과를 제공하기 어렵다. 따라서 최근 사용자가 객체분할 과정에 일부 개입하는 반자동 객체분할에 관한 연구가 활발히 이루어지고 있다.

반자동 객체분할은 사용자가 입력영상 초기에 원하는 객체를 정의하고 공간정보를 이용한 프레임내(intra-frame) 분할과, 이전 프레임에서 추출된 객체의 움직임 정보를 이용하여 자동으로 연속영상을 추적 분할하는 프레임간(inter-frame) 분할 과정으로 구성된다. 밝기값을 분할 척도로 사용하는 프레임내 객체분할 과정은 두 영역의 경계선(boundary)에서 밝기값의 불연속성을 이용한 에지기반(edge based)

[10,11]과 밝기값의 유사성을 이용한 영역기반(region based) 기법[11-13]으로 분류할 수 있다. 에지기반 알고리즘들은 영상의 경계선에서 밝기값이 급격하게 변화하는 영역을 찾는 기법이다. 그러나 잡음 및 불규칙적인 형태를 가지는 에지정보를 다양한 에지 연산자들에 의해 찾는 것은 어려운 일이며, 경계가 불분명한 영상에서 객체를 분할하기에는 많은 문제점을 가지고 있다. 한편, 영역기반 방법은 유사한 화소들의 집단으로 구성하고 객체를 분할한다. 그러나 잡음이나 객체 경계에서의 번짐에 민감한 특성을 가지고 많은 문제점을 가지고 있다. 따라서 최근에는 수리형태학적인 모폴로지(morphology) 개념을 이용하여 영상을 단순화하고 워터쉐드(watershed) 알고리즘을 적용하여 유사한 영역으로 분할과 병합을 수행함으로써 객체를 분할한다. 그러나 워터쉐드 알고리즘은 영상의 기울기 정보를 국부 최소값(local minima)으로 정의하고 담수영역(catchment basin)을 생성 분할하기 때문에 배경과 객체의 경계가 불명확하거나 급격하게 변화하는 영역이 많을 경우 과분할(over-segment) 현상을 초래하고, 복잡한 국부적 영역들의 병합 등 후 처리 과정이 필요하게 된다[13-15]. 한편, 프레임간 객체분할 과정에서는 이동하는 객체에 대해 움직임 정보를 분할 추정하고 객체를 분할하게 된다. 일반적으로 움직임 정보 추출 방법에는 이전 프레임에서 탐색영역을 설정하고 현재 프레임과 유사한 영역을 찾아내는 블록매칭(block matching) 기법이 있다[16]. 그러나 영상전체에 대한 움직임 추정을 수행하는 블록매칭 기법은 많은 계산량을 요구하고, 배경과 객체의 경계선에서 잡음이 많거나 밝기값이 유사한 경우 부정확한 움직임 벡터를 추정하게 되며, 이는 연속된 프레임의 부정확 추정 벡터 누적으로 신뢰성 있는 객체분할 결과를 제공할 수 없다.

Kim[1]의 프레임내 객체분할은 수리형태학적 모폴로지 필터를 적용한 기울기 영상 전체에 워터쉐드 알고리즘을 적용하고 분할된 영역들을 사용자가 정의한 불명확한 영역내에서 배경과 객체로 분류한다. 그리고 프레임간 객체분할 과정에서는 사용자의 선택에 따라서 rigid 혹은 non-rigid 객체로 구분하고, affine 움직임 모델 추정을 통한 영역을 추적 혹은 SAD(sum of weighted absolute difference)을 척도로 움직임 벡터를 추정 분할하게 된다. 그러나 프레

임내 객체분할과정에서 전체 영상에 대한 워터쉐드를 적용함으로써 객체 윤곽선이 배경과 혹은 객체내부와 유사하면 부정확한 영역을 생성하고, 프레임간 분할과정에서는 동영상에서 객체가 rigid 혹은 non-rigid인지를 사용자가 판별하여 객체를 분할하는 문제점을 가지고 있으며, affine 모델을 적용함으로써 많은 계산량을 요구하게 된다.

그림 1은 기존 방법[1]으로 잘못 추출된 영상의 객체분할 결과를 보여주고 있다. 그림 (a)에서는 배경과 객체의 윤곽선의 구분이 명확하지 않아 객체의 내부 영역으로 배경이 설정된 경우이고, (b)는 객체의 움직임이 크고 전역 움직임으로 인해 부정확한 추정 벡터가 누적되어 객체의 경계에서 배경 혹은 내부로 윤곽선이 번진 결과를 보여준다.

본 논문에서는 기존의 반자동 객체분할[1] 구조로 객체분할을 수행한다. 그러나 워터쉐드 알고리즘을 사용한 프레임내 객체분할 과정에서 잡음 성분으로 발생되는 과분할 문제와 영역병합을 통한 후처리를 사용자가 정의한 불명확한 영역에서만 처리하고, 프레임간 분할과정에서 객체의 형상변화나 전역 움직임이 존재하는 영상을 이전 프레임에서 추출된 객체 윤곽선만을 중심으로 움직임을 추정 분할한 후, 워터쉐드 알고리즘으로 보정함으로써 신뢰성 있는 객체를 추출할 수 있는 방법을 제안한다. 그리고 제안된 반자동 객체분할 기법을 적용한 실시간 동영상 편집 기기를 구현함으로써 제안된 반자동 객체분할 기법의 효율성을 보인다.

본 논문의 구성은 다음과 같다. 2장에서 제안된 반자동 객체분할 기법을 기술하고 3장에서 제안된 객체분할 알고리즘을 기반으로 구현한 실시간 객체분할 및 합성을 수행하는 편집 시스템에 대해 설명한다. 4장에서 제안된 객체분할 알고리즘과 실시간 영상 편집 시스템의 실험결과를 보이고, 5장의 결론을 맺는다.

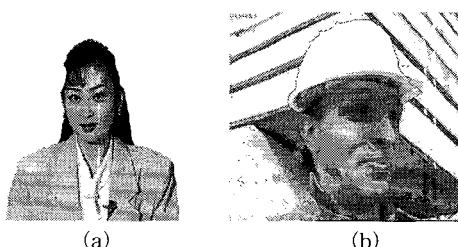


그림 1. 기존 객체분할 결과: (a) 초기 객체분할 결과, (b) 추적 객체분할 결과

2. 제안된 반자동 객체분할

2.1. 사용자 개입에 의한 프레임내 분할

프레임내 객체분할과정은 동영상에서 객체 영역을 사용자의 조력으로 정의하는 불명확한 영역 설정 단계, 수리형태학적 모폴로지 필터와 워터쉐드 알고리즘[14]을 적용한 불명확한 영역내의 분할단계, 분할된 영역으로부터 정확한 객체의 윤곽선을 추출하는 영역분류 단계로 구성된다.

그림 2는 프레임내 객체분할 과정을 도시한 것으로, 사용자 조력단계에서는 사용자가 입력 영상으로부터 관심있는 객체의 윤곽선을 사용자가 정의하고 불명확한 영역으로 설정한다. 영상 단순화 단계에서는 입력 영상에서 불명확한 영역으로 설정된 곳을 수리형태학적 모폴로지 필터를 적용하여 객체에 대한 형태는 유지하면서 기생하는 잡음성분 등을 제거하는 단순화(simplification)를 한다. 그리고 마커 생성 단계에서 단순화된 불명확한 영역내에 국부 최소값을 찾는 마커(seed)를 설정하고, 국부적 워터쉐드 단계에서 계층적인 큐를 구성하여 영역을 분할한다. 마지막으로 영역분류 단계에서는 지역적 레이블이 할당된 독립적인 영역들을 전역 레이블을 기준으로 병합하고 정확한 객체의 윤곽선을 추출한다.

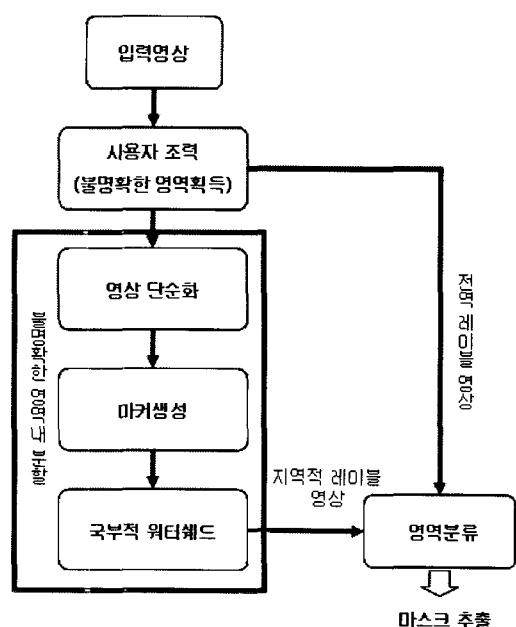


그림 2. 프레임내 객체분할 과정

2.1.1. 사용자 개입에 의한 초기 영역 지정

초기 입력되는 동영상에서 사용자가 요구하는 객체를 자동으로 수학적 모델링을 하거나 분할할 수 있는 척도를 정의하기는 불가능하다. 따라서 처음 입력되는 영상에서 관심있는 객체영역을 사용자가 정의해야만 한다.

본 논문에서는 사용자의 직관적인 편리함을 위해 원하는 객체 윤곽선의 임의의 제어점을 선택하고 이를 곡선 혹은 직선으로 표현하는 B-Spline과 Polygon, 그리고 연속으로 추적하여 그려주는 free draw 의 3가지 조력 방식을 구현했다. 연속추적 방식은 영상내의 의미있는 객체를 정의하기위해 객체의 경계선을 연속으로 추적하여 초기객체의 경계선을 표현해주는 반면, B-Spline과 polygn 방식은 객체의 특징점으로 대표되는 임의의 제어점을 선택하여 제어점 사이의 곡률값에 따라 곡선을 생성하고 초기 객체의 경계선을 정의한다[17]. 이는 사용자의 관점에서 연속추적 방식보다 직관적인 편리성을 제공하여 사용자의 개입을 최소화 하고 간편하게 초기 객체의 경계선을 정의할 수 있도록 도와준다.

사용자 개입에 의해 정의된 객체의 경계선 영역을 불명확한 영역인 '0'으로 할당한다. 그리고 불명확한 영역을 중심으로 객체내부 영역은 '1'로 할당하고, 외부 영역은 '2'로 할당한 전역 레이블 영상을 생성한다.

2.1.2. 영상 단순화 및 마커 추출

영상내의 존재하는 객체의 형태는 일반적으로 평탄하지 않으며, 조명의 변화나 잡음 성분의 영향으로 국부 최소값을 기반으로 영역을 분할하는 워터쉐드 알고리즘을 적용하게 되면, 원하지 않는 영역까지 분할하는 과분할 문제점을 가지고 있다. 따라서 앞서 획득된 불명확한 영역에서 객체 경계선을 유지하고 평탄한 영역은 더욱 부드럽게 만들어주는 영상 단순화를 수행함으로써 국부 최소값을 줄이고 과분할 현상을 방지한다.

본 논문에서는 5×5 구조 요소(structuring element)로 수리형태학적 모폴로지 필터인 팽창(dilation)과 침식(erossion) 연산자의 조합으로, 주위에 비해 밝기 및 색상값이 큰 미세한 부분들을 제거(opening)하고, 주위에 비해 밝기 및 색상값이 작은 미세한 부분들 채우(closing)는 단순화과정을 불명확한 영역에서만 수행한다. 그리고 워터쉐드 알고리

즘 적용을 위한 국부 최소값을 식 1의 (a)와 (b)의 차이를 이용하여 고도값을 갖는 기울기 영상을 생성한다.

- (a) Erosion: $\varepsilon_n(f)(x,y) = \min\{f(x+x_0, y+y_0), (x_0, y_0) \in M_n\}$
- (b) Dilatation: $\delta_n(f)(x,y) = \max\{f(x-x_0, y-y_0), (x_0, y_0) \in M_n\}$
- (c) MG(Y): $g(f)(x, y) = \delta_1(f)(x, y) - \varepsilon_1(f)(x, y)$
- (d) MG(YC): $g(x, y) = \max(g_r(x, y), g_c(x, y))$

(1)

여기서 $f(x,y)$ 는 불명확한 영역내의 화소이고, M_n 은 크기가 n 인 구조요소(structuring element)이다. 그러나 식 1의 (c)에서와 같이 밝기값만을 가진 기울기 영상을 구하면, 배경과 객체 사이에 밝기값의 변화가 적을 경우 정확한 영역을 분할할 수 없는 문제점이 있다. 따라서 식 1의 (d)와 같이 밝기와 색상값을 동시에 고려한 기울기 영상을 생성한다.

그림 3은 불명확한 영역에만 단순화 필터를 수행하고, 워터쉐드 마커인 국부 최소값이 존재하는 영상으로 나타낸 것이다. (a)는 입력영상을 나타내고 있으며, (b)는 불명확한 영역에서 워터쉐드의 마커로 사용할 기울기 영상을 추출한 결과를 보여 주고 있다.

2.1.3. 워터쉐드를 이용한 영역 분할

워터쉐드 알고리즘은 영상의 화소값을 고도로 하는 2차원 지형으로 간주하고, 기울기 영상으로부터 획득된 국부적 최소값을 기반으로 유사한 화소값을 가지는 동질영역들로 분할한다. 기존의 워터쉐드 알고리즘 수행과정은 국부적 최소값으로부터 계층적 큐를 구성하고, 주위의 유사한 화소값을 채워 나가면서 최종적으로 만나는 곳을 영역사이의 경계로 구분하는 알고리즘이다[18]. 그러나 일반적인 워터쉐드 알고리즘은 입력 영상 전체에 대해 국부 최소값을 기반으로 영역을 분할하게 되므로 불필요한 영역 병



그림 3. 국부 최소값을 갖는 마커 생성: (a) 입력영상, (b) 기울기 영상

합 등의 많은 연산을 필요로 하는 문제점을 가지고 있다. 또한 전체 영상에 대한 워터쉐드 알고리즘을 적용할 경우 객체의 윤곽선과 배경이 유사하면, 정확한 객체를 분할하기 어렵다. 따라서 본 논문에서는 단순화된 영상으로부터 획득된 기울기 값을 갖는 불명확한 영역에서만 국부적으로 워터쉐드 알고리즘을 적용하여 영역분할을 한다.

제안된 워터쉐드 알고리즘은 그림 4와 같이 사용자가 정의한 불명확한 영역에서 객체의 실제 윤곽선을 중심으로 좌우가 낮아지는 지역에서 가장 높은 기울기 값을 갖는 윤곽선을 찾음으로써 정확한 객체를 분할한다. 제안된 워터쉐드 알고리즘은 다음과 같이 크게 2단계로 구성된다.

우선 첫 단계에서는 사용자의 조력에 의해 세 영역으로 구분된 마스크영상의 불명확한 영역 중에 객체 영역과 인접한 화소와 배경영역과 인접한 화소지점들을 찾고, 찾아진 화소가 포함된 두 선의 위치를 현재영상에서 구해진 기울기 영상에 투영하여 기울기 값이 최소인 화소를 찾는다. 이러한 화소를 그림 4에서 검은 공으로 나타내었는데, 이들이 마커가 된다. 그림 4를 보면 검은 공으로 표시된 각각의 마커들은 영역을 처음 구성하는 시작화소가 되고, 각 영역을 구분하는 독립적인 국부 최소값이 할당된다. 그리고 마커들의 주위 화소를 우선적인 화소로 설정하고, 기울기값이 작은 순서대로 큐에 집어넣게 된다.

두 번째 단계에서는 마커 주변의 화소가 전부 큐에 넣어졌으면 기울기값 순서대로 화소를 하나씩 꺼내어서 인접한 거리가 최소인 마커영역에 포함시키게 된다. 각각의 영역에 화소가 포함되는 진행과정은 화소의 기울기 값에 따라 순차적으로 일어나게 되며, 마커 주변의 화소들이 영역에 전부 포함되었다면 그 화소들의 주위 화소들을 큐에 집어넣은 후 앞서 수행

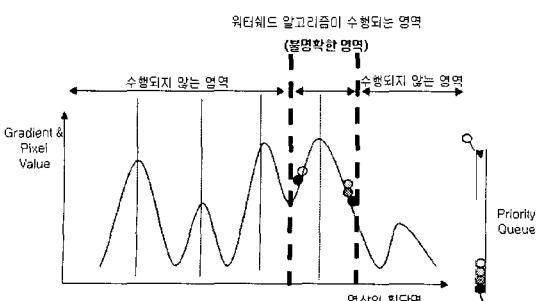


그림 4. 불명확한 영역내의 워터쉐드 알고리즘

했던 과정을 반복하게 된다. 이러한 영역확장이 기울기 단계를 기반으로 계속적으로 수행되다가 다른 국부 최소값을 갖는 화소를 만나면 확장을 멈추고 워터쉐드 경계선을 형성하게 되며, 불명확한 영역은 초기 화소로부터 확장된 개별적인 분할영역이 실제 객체의 과선을 기점으로 좌우로 분할되게 된다. 그리고 이렇게 분할된 영역에 지역적 레이블을 할당한다.

그림 5-(a)는 기존의 전체 영상에 워터쉐드 알고리즘을 적용한 결과로써 객체의 왼쪽 머리카락 영역이 배경과 동일한 영역으로 분할되어 있어 그림 1-(a)와 같이 객체영역을 파고 들어가는 현상이 나타난다. 이와반면, 제안된 워터쉐드 알고리즘은 그림 5-(b)와 같이 불명확한 영역에서만 국부 최소값을 기반으로 영역을 분할하였으므로 영역분류 단계에서 정확한 객체 영역과 배경영역을 분할할 수 있다.

2.1.4. 워터쉐드 영역분류

불명확한 영역내에서 지역적 영역으로 분할된 영역들을 객체 혹은 배경 영역으로 분류하기 위하여, 그림 6과 같이 지역적 레이블 영상을 영상에 사용자가 정의한 불명확한 영역을 투영하여 객체의 내부 영역과

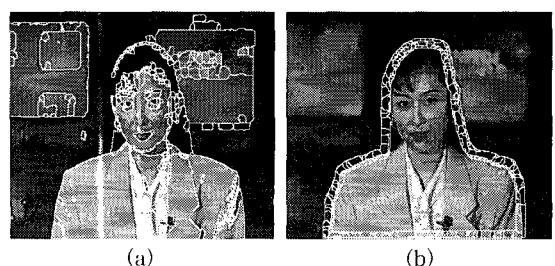


그림 5. 제안된 워터쉐드 영역분할 결과 비교: (a) 기존 워터쉐드, (b) 제안된 워터쉐드

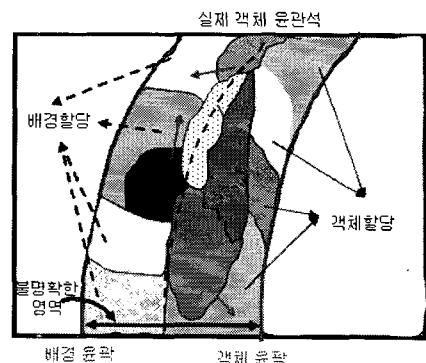


그림 6. 워터쉐드 영역 분류

만나는 라인을 in boundary로 설정하고 배경 영역과 만나는 라인을 out boundary로 설정한다. 그리고 첫 번째로, 설정된 라인을 중심으로 불명확한 영역내에 존재하는 영역들 중 in boundary 라인과 만나는 영역을 객체영역으로 설정하고, out boundary 라인과 만나는 영역들은 배경영역으로 할당한다. 그리고 in boundary 라인 혹은 out boundary 라인 중 어느 곳에도 걸쳐있지 않는 미결정 영역들은 in boundary와 out boundary를 통해 결정된 영역과의 절대평균값 차이가 최소가 되는 곳으로 영역을 할당함으로써 지역적 영역들을 분류하여 객체를 분할한다.

2.2. 객체 추적에 의한 프레임간 분할

프레임내 객체분할에서는 사용자의 조력에 의해 객체를 추출하는 반면, 프레임간 객체분할에서는 이전 프레임에서 추출된 객체 정보를 이용하여 연속된 프레임에서 객체추적(object tracking)을 함으로써 자동으로 현재 프레임의 객체를 추출한다[19].

연속된 프레임내의 객체를 추출하기 위해서는 무엇보다도 정확한 움직임을 찾아내는 것이 중요하며, optical flow나 블록매칭 등의 기법을 적용한 분할 방법 등이 있다[20]. 본 논문에서는 배경과 객체의 경계사이의 잡음으로 인한 추정 오류와 계산량을 줄일 수 있도록, 이전 프레임에서 추정된 객체의 윤곽선 정보에 대한 블록매칭 알고리즘을 적용함으로써 움직임을 추정한다.

그림 7은 제안된 프레임간 객체 분할과정을 나타낸 것으로, 이전 프레임에서 추출된 객체(O^{t-1})의 윤곽선 정보를 중심으로 BMA(block matching algorithm)을 이용하여 현재 프레임의 움직임 객체를 추출한다. 그러나 BMA를 통한 현재 프레임 객체의 움직임 벡터를 추출하면, 움직임이 발생하는 객체의 배경과 윤곽선 주위의 잡음 성분 및 객체 영역의 움직임으로 부정확한 영역이 발생하게 된다. 따라서 추정된 움직임 벡터장에 메디안 필터와 양선형 보간을 수행하여 부정확한 움직임 영역을 제거하고 부드러운 객체 영역을 추출한다. 그리고 추출된 현재 프레임 객체 마스크의 윤곽선을 중심으로 수리형태학적 모폴로지 필터를 적용하여 실제 객체 윤곽선이 존재할 불확실한 영역을 설정하고, 불확실한 영역 내에서만 워터쉐드 알고리즘을 적용하여 정확한 객체의 윤곽선 추출을 위한 보정을 수행한다.

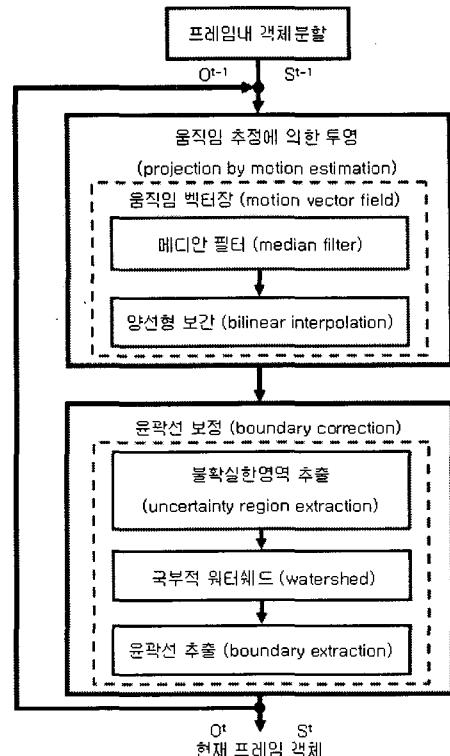


그림 7. 프레임간 객체분할 과정

2.2.1. 움직임 추정

이전 프레임의 객체로부터 현재 프레임의 객체를 추정하기 위해서는 배경과 객체사이의 잡음 성분으로 기인하는 추정오류를 최소화하는 블록단위의 움직임 추정을 한다. 그러나 이전 프레임 전체에 대한 블록단위의 움직임 추정을 하거나, affine 움직임 파라미터를 추출하여 영역단위의 추적 분할을 수행하게 되면, 실시간으로 객체를 분할하는 응용 측면에서 효율성이 떨어지게 된다. 따라서 본 논문에서는 이전 프레임에서 추출된 객체의 윤곽선 정보만을 이용한 움직임 추정을 수행한다.

그림 8은 이전 프레임인 S^{t-1} 의 객체의 윤곽선 정보로부터 블록매칭 알고리즘을 적용하여 현재 프레임의 객체인 O^t 를 움직임 추정에 의해 추출하는 과정을 나타내고 있다. 현재 프레임의 객체 O^t 를 추출하기 위하여 이전 프레임 객체인 O^{t-1} 의 윤곽선을 중심 화소로 탐색 영역을 설정한다. 그리고 clock wise 방향으로 S^{t-1} 와 S^t 의 모든 윤곽선내 화소에 대한 블록단위의 MAD(mean absolute difference)가 최소화 되는 값을 움직임 벡터로 설정한다.

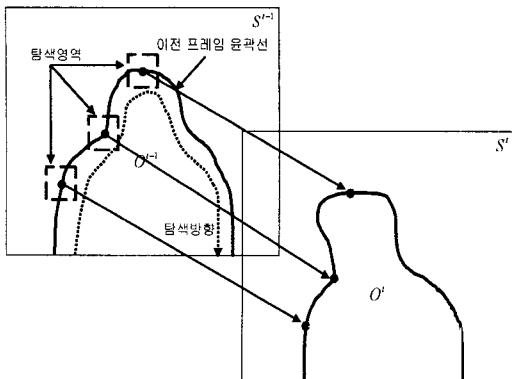


그림 8. 움직임 추정 영역 설정과 추적

2.2.2. 윤곽선 보정

현재 프레임 객체를 추출하기 위하여 블록매칭 알고리즘을 적용하고 이전 프레임 객체의 움직임 정보를 추정 분할한다. 그러나 배경과 객체의 화소값이 유사하거나 객체의 움직임이 크고 배경이 복잡하여 전역 움직임을 가지고 있는 경우 부정확한 움직임을 추정하게 된다. 또한 이전 프레임의 움직임 정보를 잘못 추적하게 되면 연속된 프레임에 추정된 에러가 계속 누적되어 객체 윤곽선이 배경으로 번지는 현상이 발생하게 된다. 따라서 추정된 움직임 정보에 의해 추출된 현재 프레임 객체 윤곽선을 보정함으로써 정확한 객체를 추출한다.

추출된 객체의 윤곽선 보정은 이전 프레임 객체의 움직임 정보로부터 투영된 현재 프레임의 객체 윤곽선을 중심으로 정확한 객체 윤곽선이 존재할 가능성이 있는 불명확한 영역을 설정하고, 불명확한 영역내에서 워터쉐드 알고리즘을 이용하여 윤곽선 보정을 한다.

불명확한 영역을 설정하기 위해서는 이전 프레임의 움직임 정보를 이용하여 생성된 현재 프레임의 객체 마스크에 모폴로지 필터인 팽창 연산자를 적용하여 생성된 마스크와 침식 연산자를 적용하여 생성된 마스크의 차이 영역을 불명확한 영역으로 할당한다. 그리고 불명확한 영역을 중심으로 프레임내 객체 분할과정에서와 같이 국부적 워터쉐드를 적용하고, 분할된 영역들의 분류를 통해 현재 프레임의 최종 객체를 획득한다.

그림 9는 모포로지 필터인 팽창과 침식 연산을 이용하여 실제 객체 윤곽선이 존재할 불명확한 영역으로 설정한 것을 보여 주고 있다.

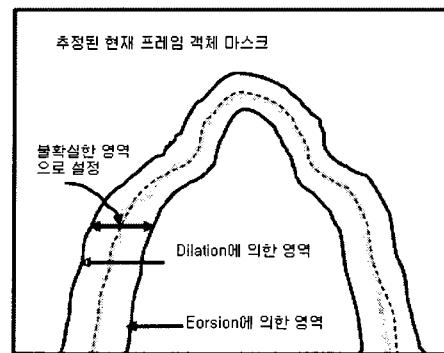


그림 9. 불명확한 영역 설정

3. 객체 기반 영상편집 시스템

최근 디지털 방송 프로그램 제작에 있어서 가상의 객체인 3D 영상을 실사 영상과 합성하거나, 기존의 실사 영상으로부터 원하는 영역을 분리하고 다른 디지털 영상 데이터와 합성을 수행하기 위해서 객체단위의 편집을 수행할 수 있는 시스템을 요구하고 있다. 이러한 편집 시스템은 무엇보다도 실시간성이 보장된 합성 기술이 필요하다. 또한 부호화 측면에서 정밀하게 추출된 객체가 있어야 한다. 따라서 본 논문에서는 실시간으로 추출된 객체와 2D 혹은 3D 영상과 화질의 열화 없이 정확하게 합성을 수행하고 편집을 할 수 있는 객체기반 디지털 영상 편집 시스템을 구현 하였다.

구현한 객체기반 편집 시스템은 디지털 영상이 입력되면 B-Spline, Polygon, Free draw 형태의 선택적인 사용자 조력으로 영상내에 존재하는 객체 특징에 따라 간편하게 영역을 정의하고 제안된 객체분할 알고리즘을 적용하여 원하는 객체영역을 분할한다. 그리고 분할된 객체를 전경(foreground)으로 설정하고 컴퓨터 저장매체와 같은 하드디스크에 저장되어 있는 2D 혹은 3D 디지털 영상 데이터와 프레임 단위의 합성을 수행한다. 또한 객체 정보를 보호할 목적으로 1단계에서 8단계의 양자화 레벨에 따라 추출된 객체의 수평 및 수직 방향으로 밝기값의 대표값을 설정하고 대표화소를 주변의 밝기값에 할당함으로 객체정보 은닉 기능을 수행한다. 마지막으로 합성된 디지털 영상데이터를 MS사의 표준 압축 비트스트림을 지원하는 미디어 플레이어 등에서 재생 가능하도록 국제표준 영상 부호화 방식인 MPEG를 이용하여 부호화를 수행한다.

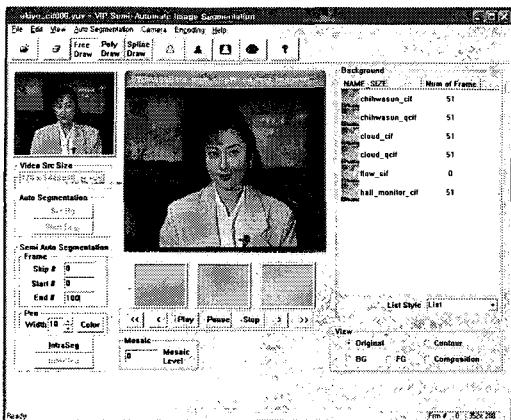


그림 10. 객체기반 편집 시스템

그림 10은 제안된 객체기반 편집시스템의 사용자 환경을 보여 주고 있다. 왼쪽의 창은 입력되는 영상을 보여주고, 중앙의 창은 현재 처리 되고 있는 결과 영상과 사용자가 선택한 조력 도구를 이용하여 원하는 객체 영역을 정의하는 작업 창을 제공한다. 그리고 오른쪽 상단의 영역은 기존의 하드디스크에 DB 형태로 저장되어 있는 2D 혹은 3D 그래픽 배경 및 실사 영상의 프레임 수, 사이즈 정보를 사용자에게 제공함으로써, 합성을 원하는 배경을 사용자의 요구에 따라 실시간으로 선택할 수 있는 편리한 환경을 제공한다. 중앙의 하단에 있는 세 개의 창은 왼쪽부터 각각 추출된 마스크, 객체, 배경 영상을 실시간으로 디스플레이 해줌으로써 객체 추출 상황을 모니터링 할 수 있다. 또한 중앙 하단의 입력값은 추출된 객체에 대한 사용자 객체정보 은닉 강도를 가변적으로 입력을 받아 객체정보를 제거한다.

4. 실험결과 및 고찰

본 논문에서는 비교적 움직임이 적은 “head & shoulder” 형태의 CIF(352x288)급 “Akiyo”와 “Mother & daughter”영상, 움직임이 비교적 크고 전역 움직임을 갖는 CIF급 “foreman”영상, 그리고 720x480 영상인 “driving” 영상을 제안된 객체분할 알고리즘을 적용하여 추적/분할한 결과를 보이고 분할된 객체정보를 실사 영상과 합성 등의 편집 실험을 하였다.

표 1은 pentium 2.8Ghz 컴퓨터에서 기준[1]의 전체 워터쉐드를 이용하는 방법과 제안된 국부적 워터쉐드 방법에 대해 지역적 영역의 수와 객체분할 시간

표 1. 전역 및 국부적 워터쉐드 수행 시간 비교

영상 (영상사이즈)	분할 영역 수 (영역 병합 대상)		분할시간 (frame/sec)	
	전역 워터쉐드	제안된 워터쉐드	전역 워터쉐드	제안된 워터쉐드
akiyo (352x288)	478	272	12.860	1.6881
mother & daughter (352x288)	547	478	13.250	1.781
foreman (352x288)	879	304	13.071	1.734
driving (720x480)	4850	246	29.501	6.078

을 비교한 것이다. 표 1과 같이 사용자가 정의한 영역 만을 중심으로 영상 단순화를 수행하고 기울기 영상을 구하여 제안된 국부적 워터쉐드 알고리즘을 적용 분할한 결과가 전체 영상에 대한 단순화 및 영상 기울기값을 마커로 하는 워터쉐드를 적용한 결과보다 영역병합 대상이 적고, 객체분할 수행 시간을 단축시킬 수 있었다. 그리고 무엇보다도 그림 15에서와 같이 객체의 겹침과 같은 형상변화 등에도 효과적으로 분할할 수 있었다.

그림 11은 “Akiyo”와 “Mother & daughter” 영상에 대한 프레임내 객체분할을 수행한 결과를 보여준다. 실험한 사용자 조력 방식은 B-spline을 이용하여 객체 영역을 정의하고 제안된 프레임내 객체분할 알고리즘을 적용하여 객체를 추출하였다. 실험결과 “akiyo”的 머리카락 부분의 윤곽선 일부가 배경과 유사하여 전체 영역에 워터쉐드 알고리즘을 적용하여 분할하는 방법[1]과 불명확한 영역에 대해 multi-valued 워터쉐드를 적용한 방법[6]은 그림 5-(a)의 결과에서 보듯이 영역을 잘못 분할하여 그림 1-(a)와 같이 부정확한 객체를 추출하게 된다. 그러나 제안된 방법은 불명확한 영역에서만 워터쉐드를 수행하여 정밀한 영역으로 분할하고 이를 영역을 병합함으로써 정확한 객체를 추출할 수 있었다.

그림 12와 13은 이전 프레임에서 추출된 객체 정보를 이용하여 연속된 프레임에 대해 자동으로 분할한 결과를 30번쨰, 90번쨰 프레임에 대해 왼쪽부터 입력 영상, 추출된 객체 영상, 추출된 배경 영상을 보여주고 있다. 이전 프레임에서 획득한 객체 정보로부터 현재 프레임의 객체정보를 추정하기 위하여 불

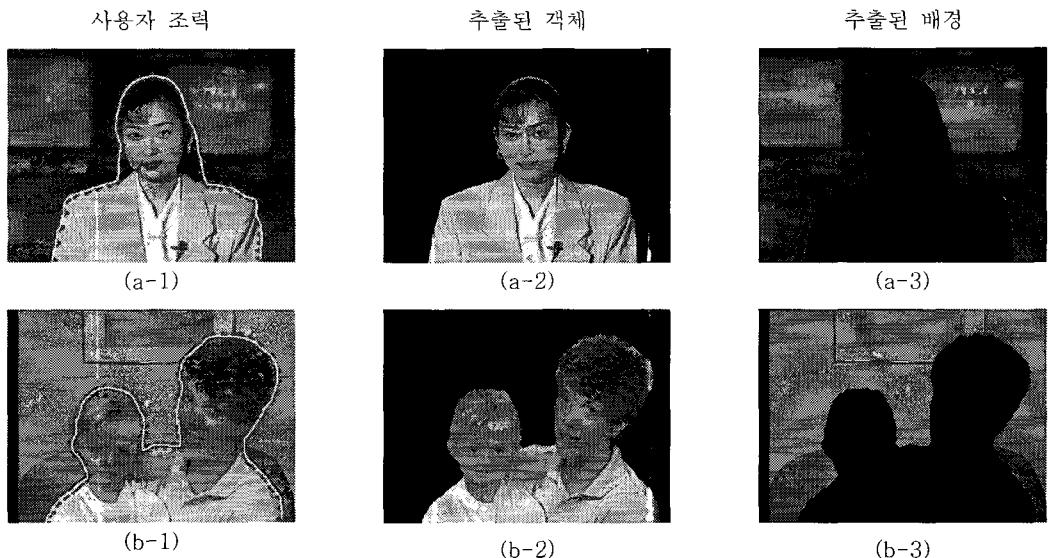


그림 11. 프레임내 객체분할 실험 결과(왼쪽부터, 사용자 개입에 의한 윤곽선 지정 영역을 나타내고, 그 다음은 추출된 객체를 나타내며 마지막은 배경을 나타낸다.)

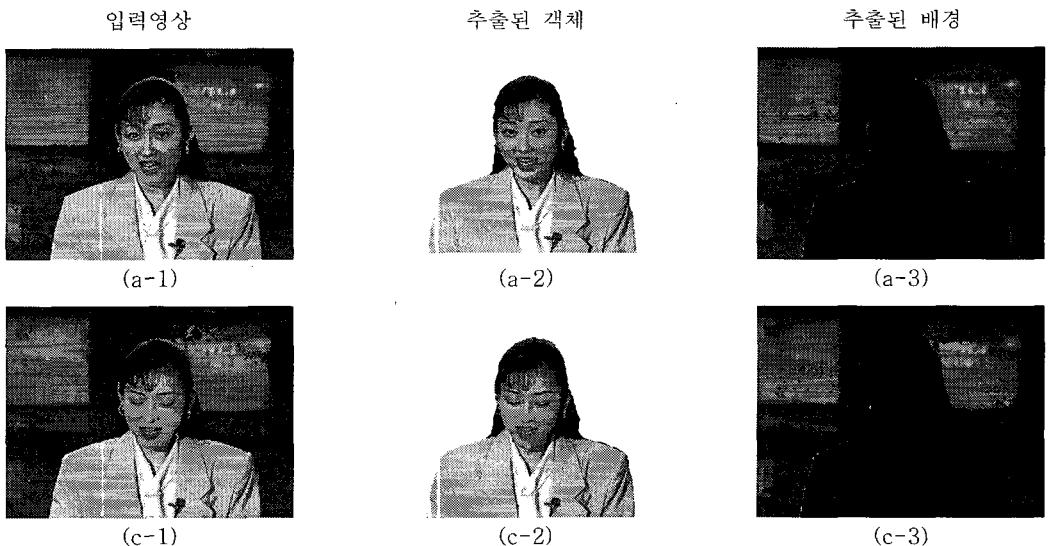


그림 12. "Akiyo" 프레임간 객체분할 실험 결과: (a) 30th frame, (b) 90th frame

록매칭 알고리즘을 적용하여 추정된 움직임 벡터를 현재 프레임에 투영하고 제안된 윤곽선 보정 알고리즘을 적용하여 추출하였다. 그리고 각각의 영상의 크기와 이동되는 객체의 특성을 고려하여 텁색영역을 8 화소로 설정하고, 블록의 사이즈를 6화소로 설정하여 블록매칭 알고리즘을 적용하였다.

그림 14는 10Hz의 프레임율을 갖는 CIF급 "foreman" 영상의 0번째 프레임에 대해서는 프레임내 객

체분할을 수행하고 이후의 프레임에 대해서는 추적 분할한 결과를 보여주고 있다. 실험영상의 객체는 움직임이 크고 카메라의 이동으로 전역 움직임이 발생하는 특징을 가지고 있어, 기존[1,6]의 움직임 파라미터를 이용한 영역 추적분할 방법보다 제안된 방법이, 움직임 추정오류로 인해 영역분할 결과가 객체 외부나 내부로 번지는 현상을 현격히 줄일 수 있어 신뢰성 있는 분할결과를 제공함을 알 수 있다.

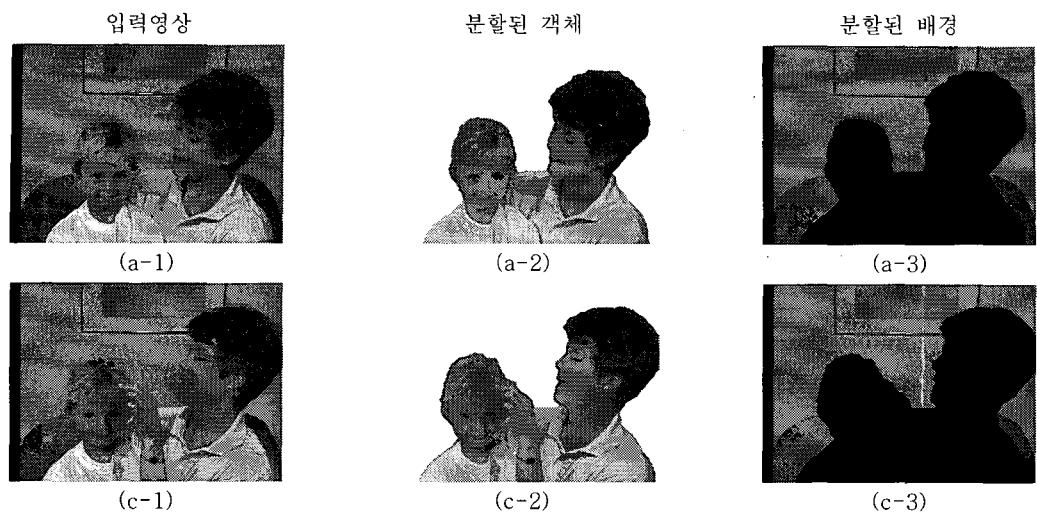


그림 13. "Mother&daughter" 프레임간 객체분할 실험 결과: (a) 30th frame, (b) 90th frame

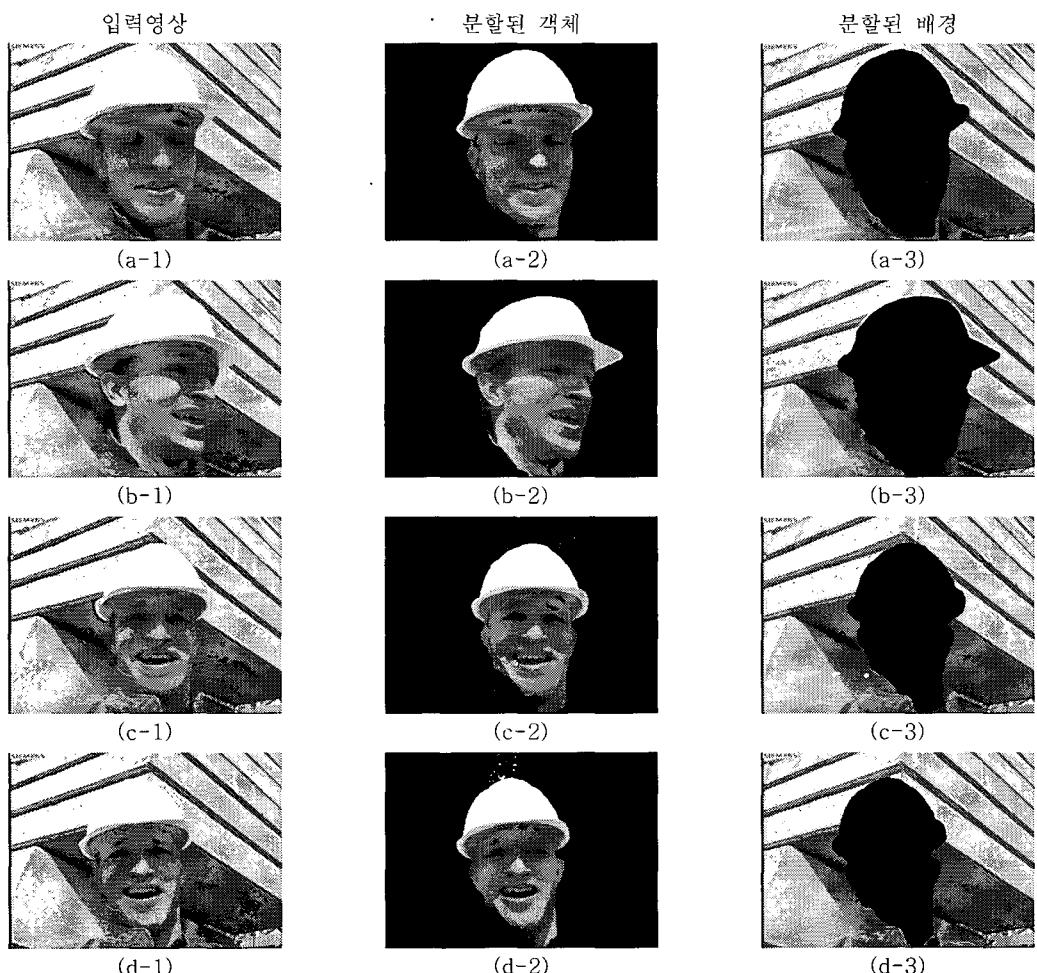


그림 14. "foreman" 프레임간 객체분할 실험 결과: (a) 0th frame, (b) 10th frame (c) 20th frame, (d) 30th frame

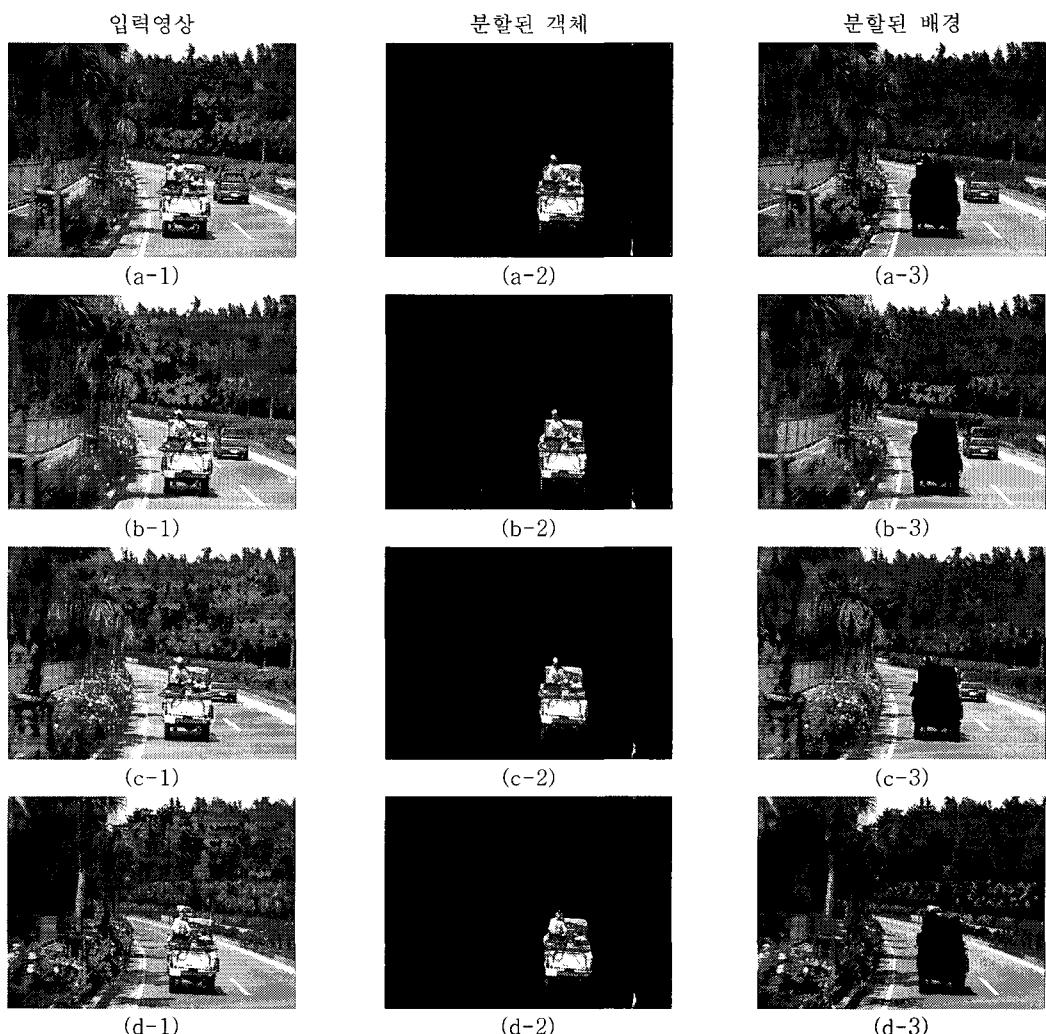


그림 15. “driving” 프레임간 객체분할 실험 결과: (a) 1th frame, (b) 6th frame (c) 24th frame, (d) 89th frame

그림 15는 720x480 사이즈의 “driving” 영상에 대해 추적 분할한 결과를 보여 주고 있다. 실험영상 역시 전역 움직임을 가지고 있고, 추출하고자 하는 객체의 오른쪽에 새로운 객체가 겹치는 현상을 가지는 영상이다. 그림 15-(b)에서 15-(d)의 추출결과와 같이 오른쪽에 나타난 객체가 점진적으로 추출하고자 하는 객체와 겹쳐도 제안된 방법이 정확한 추출 결과를 제공함을 알 수 있다. 이는 추적 대상이 되는 객체의 이전 프레임에서 추출된 객체의 윤곽선 정보만을 이용 움직임을 추정하고, 불명확한 영역을 설정하여 국부적인 워터쉐드 알고리즘을 적용하여 보정하기 때문이다.

그림 16은 추출된 객체를 사용자의 요구에 따라

실시간으로 2D 실사 영상과 합성된 결과와 추출된 실사 영상을 기준의 3D 객체와 합성된 결과 영상을 보여주고 있다. 그림 17은 사용자 입력 레벨에 따른 객체정보를 은닉한 결과를 보여주고 있다.

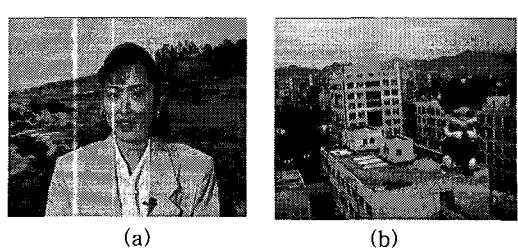


그림 16. 사용자 요구에 의한 영상합성 결과: (a) 2D 객체 합성, (b) 3D 객체 합성

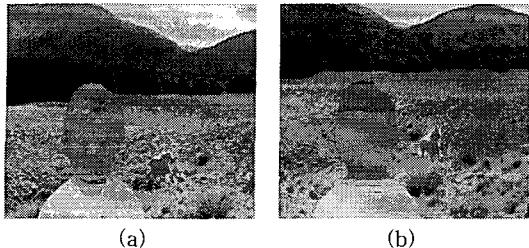


그림 17. 객체정보 은닉 결과: (a) 1레벨 은닉, (b) 8레벨 은닉

5. 결 론

본 논문에서는 효율적인 객체기반 편집 시스템을 위해서 사용자 개입 방식의 불명확한 영역만을 이용하는 객체분할기를 제안하였다. 제안된 객체분할기는 프레임내 객체분할과정에서 기존의 입력 영상 전체에 대해 워터쉐드 알고리즘을 적용하는 대신 사용자가 정의한 영역을 불명확한 영역에서만 객체분할을 수행한다. 또한 워터쉐드에 의해 분할된 불명확한 영역내의 지역적 영역들을 효과적으로 병합하기 위하여 불명확한 영역 외부와 내부 라인을 설정하고 이를 기반으로 정확한 영역을 분류함으로써 고속으로 정밀한 객체를 추출할 수 있었다. 프레임간 객체분할에서는 이전 프레임에서 추출된 객체 마스크의 윤곽선을 중심으로 움직임 추정을 수행하고, 움직임 보상된 객체에 대한 윤곽선 보정을 국부적인 워터쉐드 알고리즘을 적용하여 부정확한 움직임 정보에 의한 객체의 형상변화 등에 강인한 추적 분할을 할 수 있었다. 그리고 제안된 반자동 객체분할 알고리즘을 적용한 객체 기반 동영상 편집기는 실시간으로 영상을 합성할 때 시각적 열화없이 객체합성을 하고, 초상권 보호를 목적으로 하는 객체정보 은닉시 추출된 객체의 마스크 정보만을 이용함으로써 신뢰성 있는 초상권 보호를 기능을 확인할 수 있었다.

참 고 문 헌

- [1] Munchurl Kim, J.G. Jeon, J.S. Kwak, M.H. Lee, and C. Ahn, "Moving object segmentation in video sequence by user interaction and automatic object tracking," *Image and Vision computing*, Vol. 19, pp. 245-260, 2001.
- [2] B.S. manjunath, Philippe Salembier, and Thomomas Sikora, *Introduction to MPEG-7*, WILEY Ltd., 2002.
- [3] M. Kim, J.G. Choi, D. Kim, H. Lee, M.H. Lee, C. Ahn, and Y. Ho, "A VOP generation tool: Automatic segmentation of moving objects in image sequence based on spatio-temporal information," *IEEE Trans. CSVT*, Vol. 14, pp. 277-296, Feb. 1999.
- [4] T. Aach, A. Andre Kaup, and R. Rudolf Mester, "Statistical model-based change detection in moving video," *Signal Processing*, Vol. 31, No. 22, pp. 165-180, Mar. 1993.
- [5] A. Neri, S. Colonnese, G. Russo, and P. Talone, "Automatic moving object and background separation," *Signal Processing*, Vol. 66, No. 2, pp. 219-232, Apr. 1998.
- [6] Chuang Gu and Ming-Chieh Lee, "Semiautomatic Segmentation and Tracking of Semantic Video Objects," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, No. 5, pp. 572-584, Sep. 1998.
- [7] J.G. Choi, M.C. Kim, M.H. Lee, and C.D. Ahn, "A User Assisted Segmentation Method for Video Object Plane Generation," *IEEE Trans. on Image Processing*, pp. 881-898, 1996.
- [8] R. Mech and M. Wollborn, "A noise robust method for segmentation of moving objects in video sequence," *IEEE Int. Conf. Acoust., Speech, Signal Processing*, Vol. 4, pp. 2657-2660, Apr. 1997.
- [9] R. Mech and P. Gerken, "Automatic segmentation of moving objects(partial result of core experiment n2)," ISO/IEC JTC1/SC29/WG11 MPEG97/m1949, Bristol, U.K., Apr. 1997.
- [10] V. Torre and T. Poggio, "On edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, pp. 147-163, Mar. 1986.
- [11] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing 2nd Edition*, Prentice Hall, 2001.
- [12] S. Chen, W. Lin, and C. Chen, "Split-

- and-merge image segmentation based on localized feature analysis and statistical tests," *CVGIP: Graph. Models Image Processing*, Vol. 53, pp. 457-475, Sept. 1991.
- [13] Hai Gao, Wan-Chi Siu, and Chao-Huan Hou, "Improved Techniques for Automatic Image Segmentation," *IEEE Trans. on Circuits and systems for Video Technology*, Vol. 11, No. 12, Dec. 2001.
- [14] Shao-Yi Chien, Yu-Wen Huang, and Liang-Gee Chen, "Predictive Watershed: A Fast Watershed Algorithm," *IEEE Trans. on Circuits Syst.*, Vol. 13, No. 453-461, May 2003.
- [15] P. Salembier and Montse Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 639-651, 1994.
- [16] Thomas Meier and King N. Ngan, "Automatic Segmentation of Moving Objects for Video Object Plane Generation," *IEEE Trans. on CSVT*, pp 525-538, Vol. 8, No. 5, Sep. 1998.
- [17] Donald Hearn and M. Pauline Baker, *Computer Graphics - C version*, Prentice Hall, Inc., pp. 297-355, 1997.
- [18] Luc Vincent and Pierre Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 6, June 1991.
- [19] Hong Y. Yu, Sung-Hoon Hong, Mile M. Lee and Jae-Gark Choi, "A new user-assisted segmentation and tracking technique for an object-based video editing system," *Proc. SPIE Int. Soc. Opt. Eng.*, Vol. 5274, pp. 470-481, Dec. 2003.
- [20] J. L. Barron, D. J. Feelt and S. S. Beauchemin, "Performance of Optical Flow techniques," *Int'l J. Computer vision*, Vol 12, No. 1, pp.43-77, Feb. 1994.

유 흥 연



2001년 동신대학교 전기전자공학
과(공학사)
2004년 전남대학교 전자정보통신
공학과(공학석사)
현재 전남대학교 전자정보통신공
학과 박사과정

관심분야 : 멀티미디어 시스템, 의료영상 처리, 영상부호화

홍 성 훈



1988년 영남대학교 전자공학과
(공학사)
1991년 한국과학기술원 전기및전
자공학과(공학석사)
1999년 한국과학기술원 전기및전
자공학과(공학박사)
1991년 ~ 2000년 7월: LG전자
DTV연구소 책임연구원
2000년 7월 ~ 현재 전남대학교 전자컴퓨터정보통신공학
부 조교수
관심분야 : 영상 부호화, 의료영상 처리, 멀티미디어 시스
템, 영상처리용 임베디드 시스템