

논문97-2-2-07

비디오 객체 생성을 위한 자동 영상 분할 방법

최재각*, 김문철*, 이명호*, 안치득*, 김성대**

An Automatic Segmentation Method for Video Object Plane Generation

Jae-Gark Choi*, Mun-Churl Kim*, Myoung-Ho Lee*, Chieteuk Ahn*, and Seong-Dae Kim**

요 약

본 논문은 MPEG-4와 같이 객체 및 내용 기반 영상 부호화에 필요한 동영상의 자동 영역 분할 알고리즘을 제안한다. 통계적 가설 검증(statistical hypothesis test)을 사용하여 영상 시퀀스내에 포함된 비디오 객체들(video objects)을 움직임 물체(moving objects)와 배경(background)으로 자동 분할하는 새로운 영상 분할 알고리즘을 제안한다. 기존 방법들이 두개의 연속된 영상을 사용하는 반면에, 제안된 방법은 3개의 연속된 영상을 사용하여, 2개의 차영상의 평균값을 비교하여 가설검증을 행함으로써 잡음에 강한 특성을 나타낸다. 그리고 제안된 방법은 기존 방법와는 달리 참분산(true variance)을 사전에 알고 있을 필요가 없는 장점을 갖고 있다[18]. 또한 시간정보만을 이용한 변화 검출 방법의 문제점인 불규칙하고 부정확한 영역의 경계를 공간정보를 이용하여 보정하는 새로운 방법을 제안한다. 실험 결과에서 주어진 것처럼 제안된 시공간정보를 이용한 영상 분할 알고리즘이 시각적으로 의미 있는 분할 결과를 제공함을 알 수 있고, 정확한 영역 경계를 추출할 수 있기 때문에 MPEG-4와 같은 객체 기반 영상 부호화에 적용할 경우에 영역 경계에서 상당히 우수한 재생 화질을 얻을 수 있다.

Abstract

The new video coding standard MPEG-4 is enabling content-based functionalities. It requires a prior decomposition of sequences into video object planes (VOP's) so that each VOP represents moving objects. This paper addresses an image segmentation method for separating moving objects from still background (non-moving area) in video sequences using a statistical hypothesis test. In the proposed method, three consecutive image frames are exploited and a hypothesis testing is performed by comparing two means from two consecutive difference images, which results in a T-test. This hypothesis test yields a change detection mask that indicates moving areas (foreground) and non-moving areas (background). Moreover, an effective method for extracting

I. 서 론

현재 표준화가 진행중인 MPEG-4[1][7]는 장면(scene)을 분리된 비디오 객체들(video objects: VO)의 관계로 표현함으로써 영상의 내용에 기반을 둔 다양한 기능 및 부호화(content-based coding and functionalities)를 추구하고 있다. 이러한 비디오 객체 단위 부호화는 부호화된 비트 스트림상에서의 비디오 정보의 접근, 조작, 편집(access, manipulation and editing)을 용이하게 하여 다양

* 한국전자통신연구원

Electronics and Telecommunications Research Institute

** 한국과학기술원 전기 및 전자공학과

Dept. of Electrical Engineering, KAIST

* 본 연구는 정보통신부 출연 "MPEG-4 기반 멀티미디어 코덱 기술 개발" 로 수정되었음

한 대화형 멀티미디어(interactive multimedia) 서비스를 가능케 한다. 또한 동영상의 객체를 단위로 하여 부호화가 행하여지기 때문에 특히, 영역의 경계 부근에서 우수한 재생 화질을 얻을 수 있다. 따라서 비디오 객체 생성을 위한 동영상의 자동 영상 분할[13][14][15]은 MPEG-4와 같은 내용기반 영상 부호화 및 대화형 멀티미디어 서비스에서 반드시 필요한 도구이다.

일반적으로 영상분할이란 정지영상 또는 동영상을 주어진 척도하에서 균일한 영역(homogeneous regions)으로 나누는 과정으로 정의할 수 있다. 그러나 MPEG-4에서 말하는 비디오 객체란 밝기 또는 컬러가 유사한 영역이나 움직임 모様が 유사한 영역이 아니라 의미론적인(semantic) 측면에서 의미있는(meaningful) 영상내의 물체를 말한다. 보통 MPEG-4에서의 비디오 객체 생성을 위한 영상분할은 영상 시퀀스의 각 프레임의 움직임을 배경으로 나누어서 이진화된 마스크(binary mask)를 생성하는 것을 말한다. 즉, 배경(background: default VO) 및 움직임 물체들(foreground moving objects: VOs)로 분리하는 것을 말한다.

현재까지 연속된 영상에서 배경 및 움직임 물체의 분할을 위한 중요한 단서(cue)로서는 연속된 영상사이의 밝기차(intensity difference between successive frames)가 사용되어 왔다[2][3][4][5]. 이것은 물체의 움직임이 연속된 영상에서의 밝기의 변화를 수반하기 때문이다. 연속된 영상의 밝기차를 이용한 기존의 영상분할 방법들[2][3]은 먼저 차영상(difference image)을 구한 뒤, 각 화소를 중심으로 국부적인 관측창(local observation window), W 를 만들고 이 관측창에 대한 차영상의 자승합(sum of squared values of difference pel values), 또는 차영상의 절대합(sum of absolute difference pel values)을 구하여 이를 경험적으로 사전에 정의한 문턱값(heuristically defined threshold)과 비교하여 변화영역(changed region)과 비변화영역(unchanged region)으로 영상을 분할하였다. 또한 참고문헌 [4][5]에서는 단순히 사전에 정의된 문턱값을 사용하는 대신에 식 (1)처럼 차영상을 참분산(true variance: σ)으로 나눈 분산 추정치(variance estimate: \bar{d}_i^2)를 검정 통계량(test statistic)으로 사용하여 변화영역과 비변화영역으로 분리하였다.

$$\bar{d}_i^2 = \sum_{k \in W_i} (d_k / \sigma)^2 \quad (1)$$

여기서 d_k 는 관측창내의 밝기차를 말하고 i 는 관측창의 중심을 나타낸다. 식 (1)의 검정 통계량은 χ^2 분포(Chi-Square distribution)를 가지며, 신뢰구간(confidence interval)은 주어진 유의수준(significance level)에 대해 이론적으로 결정되어 진다. 그러나 이러한 방법은 신뢰하한값(lower confidence limit)과 신뢰상한값(upper confidence limit)과 비교할 때, 가정되는 검정 통계량을 계산하기 위해 참분산(true variance)을 사전에 알아야 한다.

그러나 실제 문제에 있어서는 참분산을 알 수 없고 그 값 또한 영상의 종류에 의존적인 문제점을 지닌다.

본 논문에서는 통계적 가설 검증(statistical hypothesis test)을 사용하여 연속된 영상에서의 밝기값의 변화를 검출하는 새로운 방법을 제안한다[14][15]. 기존 방법들 [2][3][4][5]이 두개의 연속된 영상을 사용하는 반면에, 제안된 방법은 3개의 연속된 영상을 사용하여 두 개의 차영상을 구한 뒤, 두 차영상에 대한 관측창의 평균값을 비교함으로써 밝기의 변화 유무를 검출하기 때문에 잡음에 강한 특성을 나타낸다. 그리고 두 관측창의 평균값을 비교하기 위해 통계적 가설 검증을 사용하며, 이 검증에 사용된 검정 통계량은 근사적인 t 분포를 따르게 된다. 또한 제안된 방법은 참고문헌[4][5]와는 달리 참분산을 사전에 알 필요가 없는 장점이 있다. 이렇게 차영상을 이용하여 영상을 변화영역과 비변화영역으로 구분하는 것을 변화검출(change detection)이라 하고, 그 결과 분할된 이진 마스크를 변화 검출 마스크(change detection mask)라 한다. 이렇게 구하여진 변화 검출 마스크의 변화영역에는 움직임 물체만 포함되어 있는 것이 아니라 드러난 배경(uncovered background) 및 감춰질 배경(background to be covered)이 있고, 또한 매우 불규칙적이고 부정확한 경계를 나타내게 된다. 따라서 공간 정보를 이용하여 움직임 물체의 정확한 경계를 보정(boundary correction)하는 새로운 방법이 제안되었다[10][11][12][13].

본 논문의 구성은 다음과 같다. 제2장에서는 제안된 비디오 객체 생성을 위한 자동영상분할 알고리즘에 대해 설명하고, 제3장에서는 제안된 영상분할 방식의 성능 평가를 위한 모의실험 결과를 제시한다. 마지막으로 제4장에서 결론을 맺는다.

II. 제안된 비디오 객체 분할 알고리즘

1. 개요

본 절에서는 공간 정보와 시간 정보, 즉 시공간 정보(spatio-temporal information)를 이용하여 비디오 객체를 자동 분할하는 알고리즘에 대해서 설명한다. 제안된 영상 분할 알고리즘은 다음과 같이 3단계로 구성된다.

- 가. 통계적 가설 검증을 이용한 변화 검출
- 나. 드러난 배경 및 감춰질 배경의 제거를 통한 초기 움직임 물체 마스크 검출
- 다. 공간 정보를 이용한 경계 보정

그림 1에 제안된 비디오 객체의 자동영상분할 알고리즘의 구성도를 나타내었다. 맨 첫 단계가 연속된 영상에서의 밝기의 변화 검출 단계이며, 통계적 가설 검증을 통하여 각 화소별로 시간적 변화 유무를 판정하게 된다. 그 결과

로 연속된 영상에서 시간적으로 변화가 있는 변화영역과 변화가 없는 비변화영역으로 분할된다. 이렇게 검출된 결과를 변화 검출 마스크(change detection mask : CDM)라 한다. 둘째 단계로서 앞 단계에서 구하여진 변화영역으로부터 드러난 배경과 감춰질 배경을 제거하여 움직임 물체를 추출하게 된다. 이렇게 구해진 결과를 초기 움직임 물체 마스크 (initial object mask : IOM)라 한다. 마지막으로 움직임 물체의 정확한 경계를 분할하기 위해 둘째 단계에서 구하여진 움직임 물체의 경계를 공간정보를 이용하여 보정하게 된다. 따라서 최종적인 움직임 물체 마스크(object mask : OM)가 구하여 진다.

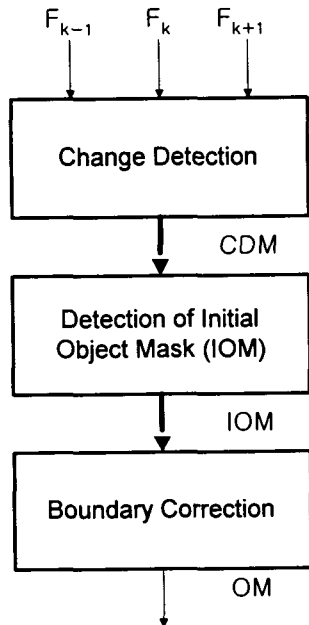


그림 1. 제안된 비디오 객체 분할 알고리즘의 구성도
Fig.1 The block diagram of the proposed algorithm for video object generation

2. 통계적 가설 검증을 이용한 변화 검출

시간적 변화 검출의 목적은 연속된 영상 사이에 밝기가 변화한 변화영역과 변화하지 않는 비변화영역을 구분하는 것이다. 물체의 움직임은 일반적으로 연속된 영상 사이에 밝기의 변화를 수반하기 때문에 연속된 영상의 차영상은 움직임 물체를 추출하는데 중요한 단서를 제공한다. 차영상을 이용한 기존의 방법들[2][3][4][5]이 두개의 연속된 영상의 차영상을 이용한 변화 판정법인데 반하여, 제안된 방법은 세 개의 연속된 영상으로부터 2개의 차영상을 구한 뒤, 두 차영상의 통계적 특성으로부터 변화를 판정함으로써 잡음에 강한 특성을 갖도록 하였다.

제안된 변화 검출 알고리즘은 다음과 같다. 그림 2에서 D_{k-1} 을 연속된 두 영상 즉, 이전 프레임(previous frame) F_{k-1} 과 현재 프레임(current frame) F_k 사이의 차영상이라고 하고, D_k 를 연속된 두 영상 즉, 현재 프레임 F_k 과 다음 프레임(next frame) F_{k+1} 사이의 차영상으로 정의한다. 이때 연속된 두 영상 사이의 차영상은 일반적으로 정규분포(normal distribution)로 모형화된다[3]. 차영상 D_{k-1} 의 임의의 위치 i 를 중심으로한 관측창 W_1 상에 존재하는 확률변수의 집합과 차영상 D_k 의 같은 공간 위치 i 에서의 관측창 W_2 상의 확률변수의 집합을 관측창의 평균 관점에서 비교한다. 즉, 두개의 평균을 비교함으로써 고려중인 화소가 변화영역에 속하는지, 비변화영역에 속하는지를 결정하게 된다. 여기서 두 관측창 W_1 과 W_2 상에 존재하는 차신호는 동일한 독립 확률변수라 가정한다. 평균의 비교는 두 관측창의 평균 추정치(mean estimates) \bar{X}_1 과 \bar{X}_2 사이의 차 성분을 구함으로써 얻어진다. 평균 비교를 위한 가설은 식 (2)처럼 가정할 수 있다.

$$\begin{aligned} H_0: \mu_1 &= \mu_2 \\ H_1: \mu_1 &\neq \mu_2 \end{aligned} \quad (2)$$

여기서 μ_1 및 μ_2 는 각각 관측창 W_1 과 W_2 의 참평균(true means)이고, H_0 및 H_1 은 각각 귀무가설(null hypothesis) 및 대립가설(alternative hypothesis)이다.

두 관측창에 대한 평균 추정치의 차 즉, 확률변수

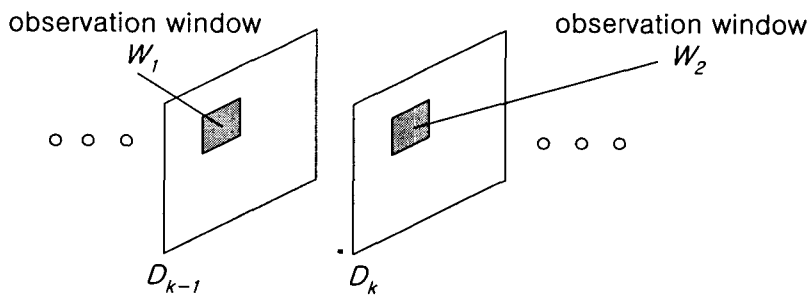


그림 2. 두 연속된 차영상에 대한 관측창
Fig. 2. Observation windows sliding over two consecutive difference images

$\bar{X}_1 - \bar{X}_2$ 는 $\mu_1 - \mu_2$ 의 unbiased estimator이고 평균이 $\mu_1 - \mu_2$ 이고, 분산이 $\sigma_1^2/n_1 + \sigma_2^2/n_2$ 인 정규 분포를 갖는다. 여기서 σ_1 및 σ_2 는 각각 관측창 W_1 과 W_2 의 참분산(true variance)이고, n_1 및 n_2 는 각각 관측창 W_1 과 W_2 내의 화소 개수이다. 따라서 확률변수(random variable) $\bar{X}_1 - \bar{X}_2$ 는 식 (3)과 같이 표준정규분포(standard normal distribution)로 표현할 수 있다.

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \quad (3)$$

그리고 귀무가설 즉, $\mu_1 = \mu_2$ 일 때, 검정 통계량 Z 는 식 (4)과 같이 표현된다.

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \quad (4)$$

그러나 이 검정 통계량 Z 를 위해서는 참분산을 사전에 알고 있어야 한다. 그러나 실제의 문제에 있어서는 두 참분산은 일반적으로 알 수 없기 때문에, 대신에 두 관측창 W_1 과 W_2 의 분산 추정치(the variance estimators) $s_{n_1}^2$ 와 $s_{n_2}^2$ 을 대신 사용한다. 따라서 식 (4)의 검정 통계량 Z 는 식 (5)와 같은 검정통계량 U 로 표현된다.

$$U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \quad (5)$$

여기서 확률변수인 검정통계량 U 는 표준정규분포를 가지는 확률변수 Z 로부터 근사화된 t 분포(t distribution)를 가지게 된다. 그러나 자유도(number of degree of freedom)를 관측창 W_1 과 W_2 에 있는 데이터로부터 추정하여야 한다. 따라서 자유도를 근사화하는 Smith-Satterthwaite 정리[17]로부터 자유도 df_U 는 식 (6)처럼 주어진다.

$$df_U \approx \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)} \quad (6)$$

여기서 만일 df_U 가 정수가 아닌 경우라면 가장 가까운 정수로 반올림한다. 따라서 귀무가설하에서 유의수준 α 가 주어지면, $\mu_1 - \mu_2$ 상의 신뢰구간에 대한 상한치와 하한치는 식 (7)에 의하여 얻어지게 된다.

$$P[-t_{\alpha/2} < U < t_{\alpha/2}] = 1 - \alpha \quad (7)$$

여기서 $t_{\alpha/2}$ 는 t 분포함수를 임의의 값에서부터 무한대

까지 적분하였을 때 그 값이 $\alpha/2$ 가 되는 값이다. 즉 $-t_{\alpha/2}$ 는 신뢰하한이고 $t_{\alpha/2}$ 는 신뢰상한이 된다. 따라서 $U < -t_{\alpha/2}$ 또는 $U > t_{\alpha/2}$ 이면 귀무가설이 거절되고(reject) 대립가설이 받아들여(accept) 진다. 그 역의 경우, 귀무가설이 받아들여 지고, 대립가설이 거절된다. 즉, 고려 중인 화소 i 에서 계산된 검정 통계량이 $t_{\alpha/2}$ 보다 크거나 또는 $-t_{\alpha/2}$ 보다 작으면 그 화소는 변화영역에 속하는 화소로 결정하고, 그렇지 않으면 그 화소는 비변화영역에 속하는 화소로 판정한다. 이렇게 구하여진 결과를 변화 검출 마스크(CDM)라고 한다. 따라서 제안된 방법은 기존방법[3]과는 달리 검정 통계량을 계산하는 과정에서 참분산을 필요로 하지않는 장점이 있다.

추정된 변화 검출 마스크는 잡음등의 결과로 매우 불규칙한 경계를 나타내게 되는데 이를 메디안 필터(median filter)로 제거하게 된다. 또한 매우 작은 변화영역이 나타나고 역으로 변화영역 내의 작은 구멍들이 나타나게 되는데 이를 제거하게 된다.

3. 드러난 배경 및 감춰질 배경의 제거를 통한 초기 움직임 물체 마스크 검출

변화 검출의 결과로서 얻어진 변화 검출 마스크는 변화 영역과 비변화영역으로 구성된다. 그러나 변화영역은 그림 3과 같은 세 영역으로 구성되기 때문에 변화영역으로부터 드러난 배경 및 감춰질 배경을 제거하여 움직임 물체를 추출하여야 한다[2].

- 가. 현재 프레임 F_k 상에 존재하는 움직임 물체(moving object : MO)
- 나. 현재 프레임 F_k 상에 드러난 배경(uncovered background for the current frame : UB)
- 다. 다음 프레임 F_{k+1} 상에 감춰질 배경(background to be covered for the next frame : BC)

움직임 물체를 추출하기 위해서는 먼저 변화영역으로부터 현재 프레임상에서 드러난 배경을 제거하게 된다. 그런 다음, 그 결과 마스크로부터 다시 다음 프레임상에 감춰질 배경을 제거한다. 그 이유는 2절에서 추출된 변화영역이 연속된 세 프레임(이전, 현재, 다음 프레임)으로부터 추출된 결과이기 때문이다.

즉 이전 프레임과 현재 프레임의 차영상으로부터 기인된 변화영역중 현재 프레임에서 드러난 배경이 되는 부분을 제거해야하고 또한 현재 프레임과 다음 프레임의 차영상으로부터 기인된 변화영역중 다음 프레임에서 감춰질 배경이 되는 부분을 제거하여야 한다. 감춰질 배경을 제거하는 작업은 두 개의 연속된 프레임을 사용하는 기존방법에는 필요치 않다.

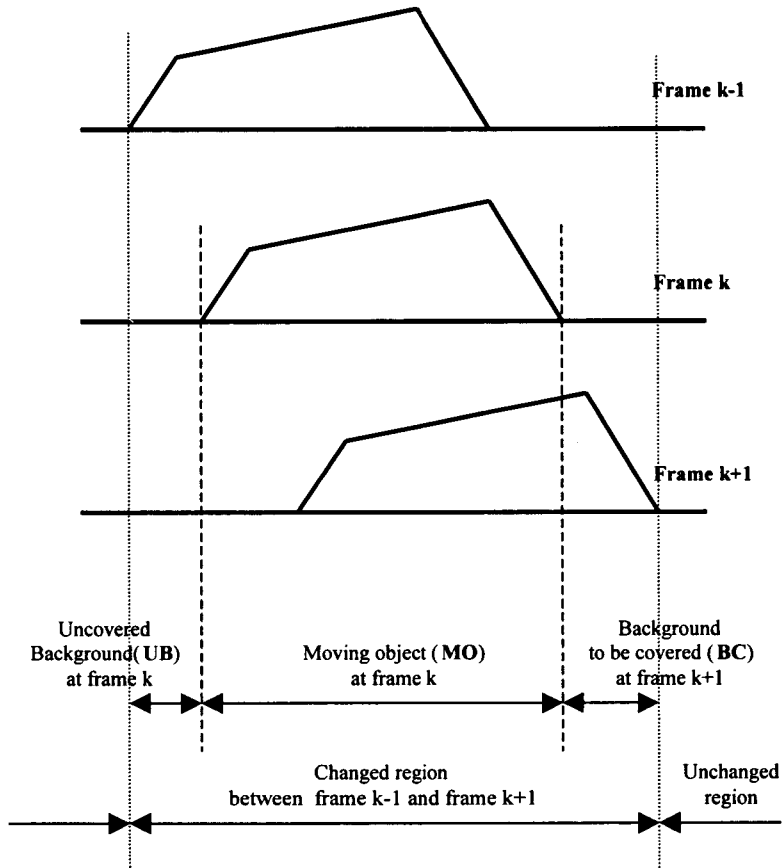


그림 3. 변화영역을 움직임 물체, 드러난 배경, 감춰질 배경으로 분리
 Fig. 3. The separation of changed regions into moving objects, uncovered background and background to be covered.

현재 프레임상의 드러난 배경을 제거하기 위해서는 먼저 변화영역에 대하여 현재 프레임과 과거 프레임사이의 움직임 벡터장을 추정하여야 하며, 이를 위하여 식 (8)과 같은 6개의 파라미터로 표현되는 affine 움직임 모형을 사용하여 미분추정법(differential method)[16]으로 움직임을 추정한다.

$$\begin{aligned} mv_x &= x - x' = a_1 + a_2x + a_3y \\ mv_y &= y - y' = a_4 + a_5x + a_6y \end{aligned} \quad (8)$$

여기서 (x', y') 은 움직임 전 화소의 좌표이며, (x, y) 는 움직임 후의 대응좌표이다, 그리고 (mv_x, mv_y) 는 화소 (x, y) 에서의 움직임 벡터를 나타낸다.

변화영역의 추정된 affine 움직임 모형 (a_1, \dots, a_6) 으로 부터 화소의 위치 (x, y) 를 대입하면 변화영역 화소의 움직임 벡터 (mv_x, mv_y) 를 계산할 수 있다. 그런 뒤, 현재 프레임상의 변화영역내의 모든 화소를 움직임 벡터를 통하여 과거 프레임상의 대응 화소로 연결할 수 있다. 이때

현재 프레임상의 변화영역에 속하는 화소가 과거 프레임상의 비변화영역에 대응되면 그 화소는 현재 프레임상에 드러난 배경으로 판정한다. 그 역으로 변화영역으로 대응되면 움직임 물체에 속한 것으로 판정한다.

다음 프레임상에 감춰질 배경을 분리하는 과정도 위와 동일하게 처리될 수 있다. 먼저 앞 단계에서 구한 움직임 물체 영역에 대해 현재 프레임과 다음 프레임사이의 움직임 벡터장을 구하고 이를 이용하여 현재 프레임의 움직임 물체 영역에 대한 다음 프레임상의 대응점을 구하여 그 화소가 비변화영역에 속하면 그 화소는 감춰질 배경으로 판정하고 그 역이면 움직임 물체에 속한 것으로 판정한다. 따라서 위의 과정을 통하여 초기 움직임 물체 마스크(IOM)를 구하게 된다.

4. 공간 정보를 이용한 경계 보정

3절에서 구하여진 초기 움직임 물체 마스크는 시간정보 즉, 차영상을 이용한 분할(temporal segmentation) 결과이

기 때문에 매우 불규칙적이고 부정확한 경계를 나타내게 된다[9]. 그 이유는 움직임 경계(motion boundary)가 인간이 시각적으로 느끼는 경계 즉, 밝기상의 에지(edge)와 일치하지 않기 때문이다. 따라서 본절에서는 이러한 움직임 물체의 부정확한 경계를 보정하기 위해 공간 정보를 이용한 방법을 제안한다. 제안된 방법은 2단계로 구성된다.

정확한 영역의 경계를 찾기 위해서는 먼저 초기 움직임 물체 마스크(IOM)의 경계 부근을 따라 일정한 크기의 대역(band)을 불확실한 영역으로 형성한다. 그 이유는 IOM이 정확한 물체의 경계를 표시하지는 못하지만 움직임 물체의 위치를 근사적으로 나타내기 때문이다. 따라서 불확실한 영역 내에 정확한 움직임 물체의 경계가 존재한다고 가정할 수 있다[11]. 불확실한 영역은 이진 영상인 IOM을 수축과 확장(shrinking and blowing)함으로써 얻어진다. 여기서 수축과 확장을 위해 모포로지컬 팽창 및 침식(morphological dilation and erosion)을 사용하였다. 그리고 그 크기는 영상의 해상도 및 종류에 따라 가변적이나 실험에 사용한 QCIF 영상에 대해서는 5x5로 하였다.

불확실한 영역을 형성한 뒤, 불확실한 영역에 속하는 화소를 주위의 두 영역 즉 움직임 물체 영역과 배경 영역 중 하나로 할당하게 된다. 이를 위하여 분수령 알고리즘(watershed algorithm)을 사용하였다. 분수령 알고리즘은 기준영역(seeds)으로 부터 시작하여 불확실한 영역에 속하는 화소들을 주위의 영역중 하나로 할당함으로써 영역을 확장하는 영역확장법이다. 분수령 알고리즘은 주로 지형학 분야에서 연구된 것으로, 영상 화소의 밝기값의 크기를 고도(altitude)로 생각함으로써 영상처리에 응용되었다. 이는 담수지역을 구분하는 분수령을 찾음으로써, 각각의 국부 최소치에 연관된 영역을 분할해 내는 알고리즘이다. 이제까지 분수령을 계산하기 위한 다양한 알고리즘이 제안되었으나, 가장 효과적인 것은 immersion simulation 기법[6][8]으로서, 이것은 계층적 큐(queue)에 의하여 구현된다.

Immersion simulation은 이중 순서의 특성을 가진다. 즉, 유사성이 큰 화소가 유사성이 적은 화소보다 먼저 병합되며, 또 유사성이 같을 때에는 먼저 병합된 쪽에 가까운 위치의 화소가 먼 위치의 화소보다 먼저 병합된다. 이 성질을 이용하여 immersion simulation을 계층적 큐를 사용하여 구현할 수 있다. 계층적 큐는 우선권(priority)이 다른 큐들의 집합이다. 이 각각의 큐들은 FIFO(first-in-first-out)의 데이터 구조로 되어 있으며, 큐에 의해서 처리되는 원소들은 화소의 위치이다. 이 계층적 큐도 이중 순서의 특징을 가진다. 즉, 우선권이 높은 큐에 있는 원소들이 먼저 처리되며, 같은 큐내에서는 먼저 들어간 원소가 먼저 처리된다.

계층적 큐가 immersion simulation을 구현하는 과정은 큐 초기화 과정과 범람(flooding)과정의 두단계로 구성된다. 초기화는 모든 국부 최소치의 좌표를 큐에 입력하는 과정이다. 즉 마커(marker)의 이웃한 화소를 큐에 입력하

는 과정이다. 이때, 그들의 유사성 척도의 역수로서 우선권을 계산한다. 그러므로 우선권에 따라 계층적 큐중 하나의 큐에 마커에 이웃한 화소의 위치를 입력하게 된다. 범람과정은 큐에서 화소들을 꺼내는 것으로 구성된다. 만약 그 화소가 아직 마커에 속해 있지 않다면 그 화소의 주위에는 적어도 하나 이상의 이웃한 마커가 존재한다. 그러므로, 이웃한 모든 마커 영역을 검사하여 시공간 결합 척도가 가장 가까운 기준영역(marker)으로 할당한다. 그런 후, 이 화소에 접해 있는 화소들 중에서 아직 기준영역에 할당되어 있지 않은 화소들을 큐에 입력한다. 즉, 큐에 입력되는 화소들은 주위의 기준영역의 화소를 접하고 있는 화소들이다. 이때 역시 유사성 척도를 계산한 후, 유사성 척도에 따라 계층적 큐중 하나에 입력한다. 이 범람 과정을 모든 화소들이 영역에 할당될 때까지 계속적으로 수행한다. 따라서 최종적으로 움직임 물체와 배경으로 이진화된 움직임 물체 마스크(OM)가 구해진다.

III. 모의실험 결과 및 고찰

본 장에서는 제안된 비디오 객체 분할 알고리즘을 이용하여 모의실험한 결과를 제시하고, 그 결과를 분석하였다.

모의실험에 사용된 테스트 시퀀스는 "MOTHER_DAUGHTOR" 영상(그림 4 (a)-(c))과 "HALL_MONIATOR" 영상(그림 5 (a)-(c))으로 QCIF(176x144)의 크기를 가지며 8 비트로 양자화되어 있다. 그리고 제안된 분할 알고리즘의 각 단계별 역할 및 중요성을 보이기 위해 각 단계에 대한 결과를 각각 그림 4 및 그림 5에 나타내었다.

그림 4 및 그림 5의 (a)-(c)는 실험에 사용된 3개의 연속된 프레임에 나타낸다. 그리고 식 (5)의 제안된 검정 통계량 U 를 계산한 후, 유의 수준 $\alpha = 10^{-3}$ 에 대한 신뢰상한값 및 신뢰 하한값과 비교하였다. 검정 통계량 U 를 계산하기 위하여 사용된 관측치의 크기는 7x7로 하였다. 연속된 영상에서의 밝기의 변화를 검출한 변화 검출 마스크의 결과를 그림 4(d) 및 그림 5(d)에 각각 표시하였다. 그림에서 알 수 있듯이 물체의 움직임에 의하여 연속된 영상에서 밝기의 변화가 나타나고 이에 따라 검출된 변화 마스크가 움직임 물체를 잘 포함하고 있음을 알 수 있다. 그리고 그림 5(d)에서 홀(hall)을 걷는 사람의 오른쪽 다리의 오른쪽에 홀 바닥이 움직임 물체로 나타나 있는데 이것은 사람이 걷는 동안 그림자의 변화로 인해 발생한 것이다.

제안된 방법의 변화 검출은 연속된 세 개의 영상을 사용하기 때문에 이전 프레임과 현재 프레임 사이의 움직임 물체로 인하여 드러난 배경이 발생하고, 또 현재 프레임과 다음 프레임 사이의 물체의 움직임으로 인하여 감춰질 배경이 발생하게 된다. 따라서 변화 검출 마스크로부터 드러난 배경 및 감춰질 배경을 제거하게 된다. 따라서 그림 4(e)-(f) 및 그림 5(e)-(f)에 제안된 방법에 따라 추출된

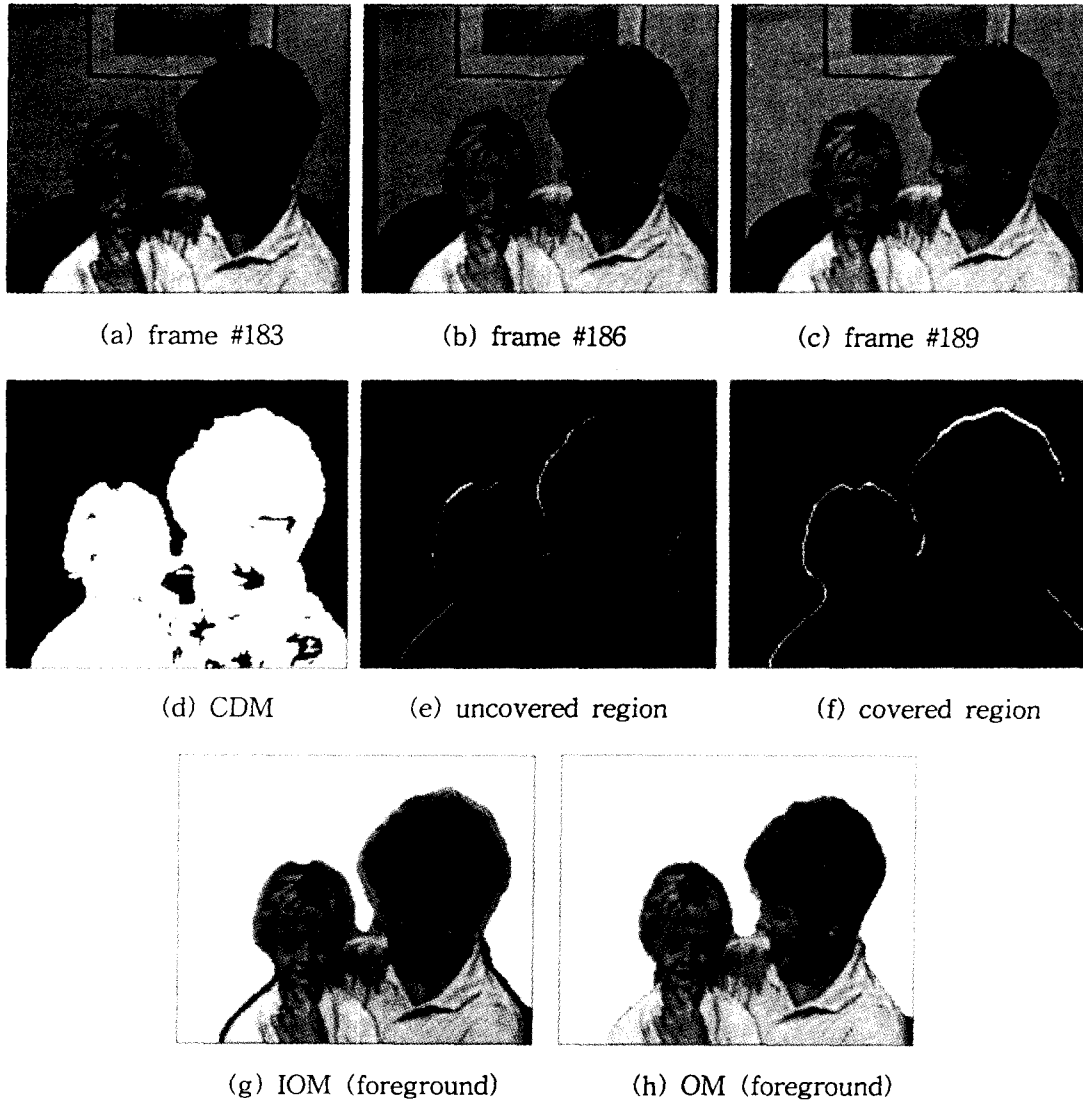


그림 4. MOTHER_DAUGHTOR 영상에 대한 실험결과
 Fig. 4. Segmentation of moving objects in MOTHER_DAUGHTOR image sequences

드러난 배경 및 감춰질 배경을 표시하였다. 그리고 변화영역으로부터 드러난 배경 및 감춰질 배경을 제거한 후에 구하여진 초기 움직임 물체 마스크(IOM)를 그림 4(g) 및 그림 5(g)에 표시하였다. 그림에서도 알 수 있지만 실제 움직임 물체의 경계보다 분할된 움직임 물체의 경계가 조금 크게 나타남을 알 수 있다. 즉, 시간 정보만을 이용한 변화검출 결과만으로는 정확한 영역의 경계를 찾지 못함을 알 수 있다. 따라서 공간 정보를 이용하여 움직임 물체의 경계를 보정한 결과의 움직임 물체 마스크(OM)를 그림 5(h) 및 그림 6(h)에 표시하였다. 공간 정보를 함께

이용함으로써 움직임 물체의 정확한 경계를 찾음을 알 수 있다. 즉, 시간정보를 이용하여 움직임 물체의 근사적인 위치(location)를 추정하고 추출된 초기 움직임 물체 마스크의 경계 부근에 대해 공간정보를 이용함으로써 정확한 움직임 물체 경계를 추출하게 된다.

IV. 결 론

본 논문에서는 효과적인 비디오 객체 생성을 위한 동영상

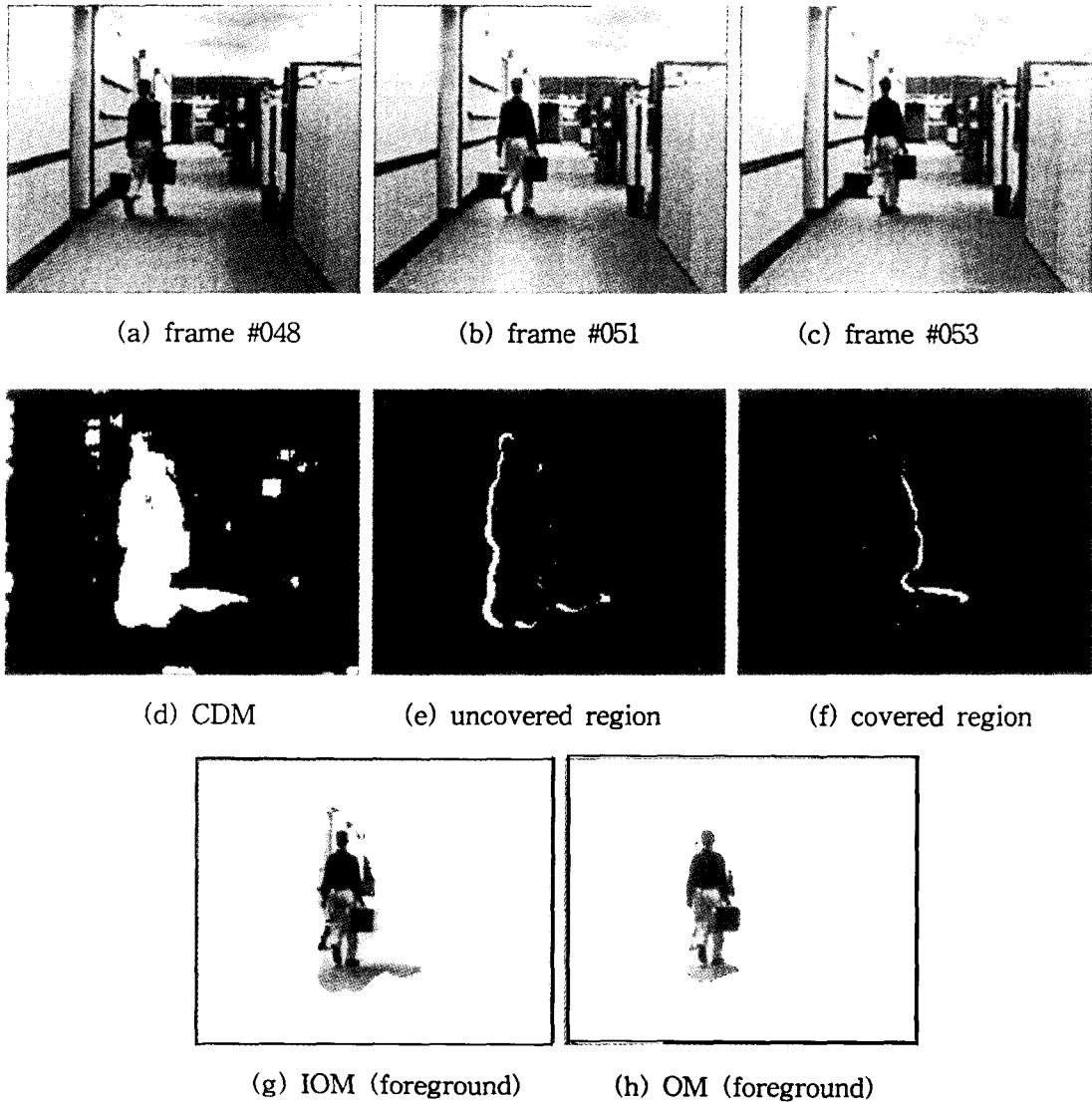


그림 5. Hall_MONITOR 영상에 대한 실험결과
 Fig. 5. Segmentation of moving objects in HALL_MONITOR image sequences

상의 자동 영상 분할 알고리즘에 대하여 기술하였다. 먼저 통계적 가설 검증을 사용하여 연속된 영상에서의 밝기값의 변화를 검출하는 새로운 방법을 제안하였다. 기존 방법들이 두개의 연속된 영상을 사용하는 반면에, 제안된 방법은 3개의 연속된 영상을 사용함으로써 잡음에 강한 특성을 나타내었다. 또한 제안된 방법은 기존방법과는 달리 참 분산을 사전에 알 필요가 없는 장점이 있다. 그리고 시간 정보만을 이용한 변화 검출 과정의 문제점인 불규칙하고 부정확한 영역의 경계를 보정하는 새로운 방법을 제안하였다. 실험 결과에서 주어진 것처럼 제안된 시공간 영상 분할 알고리즘이 시각적으로 의미있는 분할 결과를 제공

함을 알 수 있고, 정확한 영역 경계를 결정하기 때문에 MPEG-4 비디오 부화화과 같이 객체 기반 영상 부호화에 적용할 경우에 영역 경계에서 특히 우수한 재생(reconstruction) 화질을 얻을 수 있다. 또한 제안된 방식은 사전 지식을 필요로 하지 않는다는 점에서 자동적인 영상 분할 (automatic segmentation)을 가능케 한다.

앞으로의 연구는 연속된 프레임의 분할된 결과들 사이에 유사성(segmentation similarity)이 유지되도록 하기 위한 time-recursive 분할에 대한 연구가 진행될 계획이다.

이러한 연속된 프레임 분할 결과들 사이의 유사성은 현재 프레임에 대한 움직임 마스크를 결정할 때, 이전 프레

임들에서 분할된 움직임 마스크를 시간적으로 축적 (temporal accumulation)함으로써 이루어질 수 있다.

참 고 문 헌

- [1] ISO/IEC JTC1/SC29/WG11, N1683, "Overview of the MPEG-4 standard," Sevilla, Spain, Feb., 1997.
- [2] Michael Hotter and Robert Thoma, "Image segmentation based on object oriented mapping parameter estimation," *Signal Processing*, vol. 15, pp. 315-334, 1988.
- [3] H. G. Musmann, M. Hotter, and J. Ostermann, "Object-oriented analysis-synthesis coding of moving images," *Signal Processing: Image Communication*, vol. 1, pp. 117-138, 1989.
- [4] T. Aach and A. Kaup, "Statistical model-based change detection in moving video," *Signal Processing*, vol. 31, pp. 165-180, 1993.
- [5] R. Mech and M. Wollborn, "A noise robust method for segmentation of moving objects in video sequences," *ICASSP '97*, Munich, Apr. 1997.
- [6] P. Salembier and M. Pardas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. on Image Processing*, vol. 3, no. 5, pp. 639-651, Sep. 1994.
- [7] ISO/IEC JTC1/SC29/WG11, "MPEG-4 Video Verification Model Version 8.0," *MPEG97/N1796*, Stockholm, Jul. 1997.
- [8] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based immersion simulations," *IEEE Trans. on PAMI*, vol.13, no.6, pp. 583-598, Jun. 1991.
- [9] Jae Gark Choi and Seong-Dae Kim, "Multi-stage segmentation of optical flow field," *Signal Processing*, vol. 54, pp. 109-118, Oct. 1996.
- [10] Jae Gark Choi, Si-Woong Lee, and Seong-Dae Kim, "Spatio-temporal video segmentation using a joint similarity measure," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 279-286, Apr. 1997.
- [11] Si-Woong Lee, Jae Gark Choi, and Seong Dae Kim, "Scene Segmentation Using a Combined Criterion of Motion and Intensity," *Optical Engineering*, vol. 36, no. 8, pp. 2346-2352, Aug. 1997.
- [12] J. G. Choi, S. W. Lee and S. D. Kim, "Segmentation and motion estimation of moving objects for object-oriented coding," *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, Detroit, MI, May 9-12, 1995, pp. 2431-2434.
- [13] Jae Gark Choi, Munchurl Kim, Myoung Ho Lee, Chietuek Ahn, S. Colonnese, U. Mascia, G. Russo, P. Talone, R. Mech, and M. Wollborn, "Combined algorithm of ETRI, FUB and UH on Core Experiment N2 for automatic segmentation algorithm of moving objects," *Contribution Document number MPEG97/M2383*, Stockholm Meeting of ISO/IEC JTC1/SC29/WG11, Jul. 1997.
- [14] Munchurl Kim, Jae Gark Choi, Myoung Ho Lee, and Chietuek Ahn, "Core Experiments N2: Test results on ETRI's combined module of temporal and spatial segmentations for automatic segmentation techniques," *Contribution Document number MPEG97/M2386*, Stockholm Meeting of ISO/IEC JTC1/SC29/WG11, Jul. 1997.
- [15] Jae Gark Choi, Munchurl Kim, Myoung Ho Lee, and Chietuek Ahn, "New ETRI results on core experiment N2 on automatic segmentation techniques," *Contribution Document number MPEG97/M2641*, Fribourg Meeting of ISO/IEC JTC1/SC29/WG11, Oct. 1997.
- [16] Jean-Luc Dugelay and Henri Sanson, "Differential methods for the identification of 2D and 3D motion models in image sequences," *Signal Processing: Image Communication*, vol. 7, pp. 105-127, 1995.
- [17] J. S. Milton and J. C. Arnold, "Introduction to Probability and Statistics," McGraw Hill, 1995.
- [18] Munchurl Kim, Jae Gark Choi, and Myoung Ho Lee, "Localizing moving objects in image sequences using a statistical hypothesis test," to appear in *International Conference on Computational Intelligence and Multimedia Applications (ICCIAMA)*, Churchill, Australia, Feb. 1998.

저 자 소 개



최 재 각

1961년 5월 9일생

1980년 3월 ~ 1984년 2월 경북대학교 전자공학과 졸업(공학사)

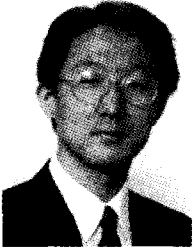
1985년 3월 ~ 1987년 2월 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

1992년 3월 ~ 1997년 8월 한국과학기술원 전기 및 전자공학과 졸업(공학박사)

현 재 한국전자통신연구원 선임연구원

주관심분야 : 영상처리, 컴퓨터비전, 영상 및 멀티미디어 통신, VLSI 구현 등

김 문 철



1966년 1월 25일생
1985년 3월 ~ 1989년 2월 경북대학교 전자공학과 졸업(공학사)
1990년 8월 ~ 1992년 12월 University of Florida(Master of Engineering)
1993년 1월 ~ 1996년 8월 University of Florida(Ph. D)
현 재 한국전자통신연구소 선임연구원
주관심분야 : Image Processing, Pappern, Recognition, Networks for Signal processing

이 명 호



1958년 6월 25일생
1983년 숭실대학교 전자공학과 졸업(학사)
1985년 숭실대학교 대학원 전자공학과 졸업(석사)
1996년 오사카대학(일본) 통신공학과 졸업(박사)
현 재 한국전자통신연구원 선임연구원
주관심분야 : 영상처리, 컴퓨터비전, 영상통신 및 멀티미디어통신 등

안 치 득



1956년 8월 15일생
1980년 2월 서울대학교 공과대학 전자공학과 졸업(학사)
1982년 2월 서울대학교 대학원 전자공학과 졸업(석사)
1991년 7월 미국 University of Florida 대학원 전기공학과 졸업(박사)
현 재 한국전자통신연구원 책임연구원
주관심분야 : 신호처리, 영상통신 등

김 성 대



1953년 12월 26일생
1977년 서울대학교 전자공학과 졸업(공학사)
1979년 한국과학기술원 전기 및 전자공학과 졸업(공학석사)
1983년 프랑스 INPT ENSEEIHT 졸업(공학박사)
현 재 한국과학기술원 및 전자공학과 교수
주관심분야 : 영상처리, 컴퓨터비전, 영상 및 멀티미디어 통신, VLSI 구현 등