

논문 2004-41SP-2-5

# 모양공간 모델을 이용한 영상분할 알고리즘

(An Image Segmentation Algorithm using the Shape Space Model)

김 대 희\*, 안 충 현\*, 호 요 성\*\*

(Daehee Kim, Chunghyun Ahn, and Yo-Sung Ho)

## 요 약

MPEG-4 표준에서는 객체 단위의 부호화를 수행하기 위해 자연영상으로부터 비디오 객체를 분리하는 영상분할(segmentation) 기술이 필요하다. 영상분할 방법은 크게 자동 영상분할(automatic segmentation)과 반자동 영상분할(semi-automatic segmentation)의 두 부류로 나눌 수 있다. 지금까지 개발된 대부분의 자동 영상분할 방법은 비디오 객체의 명확한 수학적 모델을 제시하기 곤란하며 한 화면에서 개별 객체를 추출하기 어렵기 때문에 그 성능에 한계가 있다. 본 논문에서는 이러한 문제점을 극복하기 위해 active contour 알고리즘을 이용한 반자동 영상분할 알고리즘을 제안한다. 초기 곡선으로부터 변화 가능한 모든 곡선의 집합을 모양공간으로 정의하고 그 공간을 선형공간이라고 가정하면, 모양공간(shape space)은 모양 행렬에 의해 행(column) 공간과 남은 빈(left null) 공간으로 나뉘어진다. 본 논문에서 제안하는 알고리즘은 행공간의 모양공간 벡터를 이용하여 초기 곡선으로부터 영상의 특징점까지의 변화를 기술하고 동적 그래프 검색 알고리즘을 이용하여 객체의 세밀한 부분을 묘사한다. 모양 행렬과 객체의 윤곽을 추정하기 위한 SUSAN 연산자의 사용으로 제안한 알고리즘은 저수준 영상처리로부터 생성되는 불필요한 특징점을 무시할 수 있다. 또한, 모양 행렬의 사용으로 생긴 제약은 동적 그래프 검색 알고리즘으로 보상한다.

## Abstract

Since the MPEG-4 visual standard enables content-based functionalities, it is necessary to extract video objects from video sequences. Segmentation algorithms can largely be classified into two different categories: automatic segmentation and user-assisted segmentation. In this paper, we propose a new user-assisted image segmentation method based on the active contour. If we define a shape space as a set of all possible variations from the initial curve and we assume that the shape space is linear, it can be decomposed into the column space and the left null space of the shape matrix. In the proposed method, the shape space vector in the column space describes changes from the initial curve to the imaginary feature curve, and a dynamic graph search algorithm describes the detailed shape of the object in the left null space. Since we employ the shape matrix and the SUSAN operator to outline object boundaries, the proposed algorithm can ignore unwanted feature points generated by low-level image processing operations and is, therefore, applicable to images of complex background. We can also compensate for limitations of the shape matrix with a dynamic graph search algorithm.

**Keywords:** Image Segmentation, Active Contour, Shape Space, Dynamic Graph Search

## I. 서 론

최근 멀티미디어 통신 서비스가 확대됨에 따라 MPEG-4 표준의 영상압축 부호화 방식들과는 달리, MPEG-4 표준에서는 VOP(video object plane)라는 개념

을 도입하여 객체기반 영상 부호화를 시도하고 있다<sup>[1]</sup>. 따라서 우선 주어진 화면에서 이를 구성하는 객체들을 분리해야 내야 한다. 이때 영상 객체의 모양에 관한 정보는 객체기반 부호화에서 중요한 내용이다. 즉, 입력되는 영상 시퀀스의 각 화면은 임의의 모양을 갖는 영역으로 분할되고, 분할된 각 영역은 의미있는 비디오 객체를 나타낸다.

MPEG-4 표준에서 권장하는 영상분할 방법은 시공간(spatio-temporal) 영상분할 알고리즘이다<sup>[1]</sup>. 이 방법은 움직임 검출 마스크(change detection mask, CDM)과 형태학적(morphological) 접근 방법을 사용한다. 즉, 비디오

\* 정회원, 한국전자통신연구원 방송시스템연구그룹  
(Broadcasting System Research Group, Telecommunications Research Institute)

\*\* 정회원, 광주과학기술원 정보통신공학과  
(Department of Information and Communications, Kwangju Institute of Science and Technology)

접수일자: 2003년5월6일, 수정완료일: 2004년3월1일

객체의 정확한 경계는 형태학적인 방법을 이용하여 검출하고, 비디오 객체의 영역은 CDM을 이용하여 찾는다. 그러나 이 방법은 단일 프레임에서 복수의 비디오 객체를 구분하기 곤란하며, 비디오 객체에 대한 정의 자체가 너무 주관적이어서 정확한 수학적 모델을 제시할 수 없기 때문에 MPEG-4 응용에 실제 사용하기에 부적절하다<sup>[2]</sup>.

만약 사용자가 초기 화면에서 비디오 객체에 대한 정보를 부여하고 이를 적절히 활용한다면, 이후에 연속되는 화면에서 보다 정확한 영상분할의 결과를 효과적으로 얻을 수 있다<sup>[2,3]</sup>. 이처럼 사용자의 도움을 이용하는 반자동 영상분할 방법이 MPEG-4 부호화 방법에서 VOP를 생성하는 데 보다 실질적인 대안이 될 수 있다. 본 논문에서는 active contour 알고리즘을 이용하여 사용자의 정보를 입력받아 이를 영상분할 과정에 활용한다. 즉, active contour를 이용하여 초기 객체를 정의하고, 정의된 객체로부터 연속되는 프레임에서 객체의 윤곽선을 추적하여 동영상의 객체를 연속적으로 추출한다.

또한, active contour 알고리즘을 일반적인 영상에 적용하기 위해 초기 곡선으로 모든 변화 가능한 곡선의 집합을 모양공간(shape space)이라고 정의한다. 이 모양공간은 모양 행렬(shape matrix)에 의하여 두 개의 직교 공간으로 나누어진다. 여기서 모양 행렬은 몇 개의 매개변수로 초기 곡선의 변화를 기술한다. 그러므로, 모양공간은 모양 행렬에 의해 정의된 공간과 몇 개의 매개변수로 곡선의 변화를 기술할 수 없는 공간으로 구성된다. 본 논문에서는 먼저 몇 개의 매개변수로 곡선의 변화를 기술하고, 모양 행렬에서 오는 제약을 보상하기 위해 동적 그래프 검색 알고리즘을 사용한다.

본 논문의 구성은 다음과 같다. 먼저, II장에서는 일반적인 active contour에 대한 개념을 소개하고, 기존의 방법들의 문제점을 살펴본다. III장에서는 모양공간 모델을 이용한 active contour 알고리즘을 제시하고, IV장에서는 동적 그래프 알고리즘을 이용하여 6개의 매개변수에 바탕을 둔 모양공간 모델의 제약점을 보상 방법을 제시한다. V장에서는 본 논문에서 제시된 영상분할 알고리즘에 대한 모의 실험 결과를 보이고, VI장에서 결론을 맺는다.

## II. Active Contour

Active contour의 모양은 내부 에너지, 외부 에너지, 제한 조건의 세 항목에 의해 조절된다. 외부 에너지는 active contour 영상의 특징을 추적하도록 도와주고, 내부 에너지는 곡선의 부드러운 변화를 보장한다. 사용자의 개

입은 제한 조건을 통해 허용된다. Active contour에 대한 수식은 다음과 같다<sup>[4]</sup>.

$$E_{snake}^* = \int_0^1 E_{snake}(\mathbf{r}(s)) ds \\ = \int_0^1 [E_{int}(\mathbf{r}(s)) + E_{ext}(\mathbf{r}(s)) + E_{con}(\mathbf{r}(s))] ds \quad (1)$$

식 (1)의 첫번째 항은 내부 에너지를, 두번째 항은 외부 에너지를, 그리고 마지막 항은 제한 조건을 각각 의미한다. Active contour 알고리즘의 목표는 식 (1)의 에너지 함수를 주어진 영역에서 국부적으로 최소화하는 곡선  $\mathbf{r}(s)$ 를 구하는 것이다. 여기서 내부 에너지는 곡선의 구부러지는 정도를 나타내며, 식 (2)와 같이 표현된다.

$$E_{int} = (\alpha(s)|\mathbf{r}_s(s)|^2 + \beta(s)|\mathbf{r}_{ss}(s)|^2) / 2 \quad (2)$$

식 (2)의 첫번째 항은 곡선의 1차 미분값을 이용하므로 곡선의 불연속적인 곳에서 큰 값을 갖는다. 두번째 항은 곡선이 급격하게 휘어지는 지점에서 큰 값을 갖는다. 그러므로 각 점에서의  $\alpha(s)$ 와  $\beta(s)$ 의 값은 그 점에서의 객체의 윤곽선 특성을 잘 반영해야 한다. 이에 비해, 외부 에너지는 영상에 의해서 주어지는 에너지로서, 일반적으로 영상 객체의 윤곽선을 반영할 수 있도록 정의되어야 한다.

식 (1)을 최소화하는 곡선  $\mathbf{r}(s)$ 를 구하는 것은 Euler 방정식을 이용한 variational calculus를 구하는 문제로 간주할 수 있다. Kass는 재귀적인 접근으로 이것을 구현하였으나<sup>[4]</sup>, 이 방법은 수치적으로 불안정한 방법으로 밝혀졌다. 이러한 문제를 해결하기 위해 dynamic programming 방법을 이용한 접근 방법이 많이 사용되고 있다<sup>[5]</sup>. 그러나 이 방법 또한 연산량이 많다는 단점을 가지고 있다. 한편, 이를 개선하기 위한 고속 active contour 알고리즘이 제안되었으나, 이 방법도 초기 곡선의 제어점들의 간격을 정교하게 설정해야 하는 단점을 가지고 있다<sup>[6]</sup>.

## III. 모양공간 모델

### 1. B-Spline Snake

식 (1)에서  $E(s)$ 의 실질적인 계산은 디지털 영역에서 수행되므로 연속적인 함수의 곡선도 역시 이산화되어야 한다. 따라서 이산화된 내부에너지는 식 (3)과 같다.

$$E_{int} = \alpha_i |\mathbf{r}_i - \mathbf{r}_{i-1}|^2 + \beta_i |\mathbf{r}_{i+1} - 2\mathbf{r}_i + \mathbf{r}_{i-1}|^2 \quad (3)$$

만약 외부에너지 없이 식 (3)만 사용하면, 곡선은 하나의 점으로 수렴한다. 이런 문제를 해결하기 위해 식 (3)

과 다른 변형들이 사용될 수 있다. 그러나 유한 점들의 차이를 이용한 식 (3)은 그 점들 사이에서의 곡선에 대한 정보를 완벽하게 제공하지 못한다. 따라서 연속적인 곡선  $r(s)$ 를 완벽하게 복원할 수 있는 제어점  $r(s_i)$ 을 이용하는 방안이 고려되어야 한다. 다항식의 spline으로  $r(s)$ 를 형성하면 곡선의 부드러운 변화를 얻을 수 있다. 그러나 spline 곡선은 영상의 특성을 고려하지 않기 때문에 spline 곡선은 영상의 구석점(corner point) 위를 지나가지 않을 수도 있다. 이런 문제는 다중 연결점 또는 다중 제어점을 이용하여 spline 곡선의 부드러운 변화를 조절하는 NURBS 곡선을 사용하여 해결할 수 있다<sup>[7]</sup>.

식 (1)에서 매개변수를 이용한 곡선  $r(s)=(x(s), y(s))$ 은  $s$ 의 함수이다. 차수  $d$ 의 B-spline의 기저함수(basis function)는 연결 점에서 서로 연결된  $d$  개의 다항식 곡선 조각으로 구성된 함수이다. 복잡한 모양은 높은 차수의 다항식으로 근사화될 수 있다. 일반적으로 다항식의 차수는 3차(quadratic)나 4차(cubic)가 많이 사용된다.

B-spline 함수  $x(s)$ 는  $N_B$  개의 기저함수  $B_n(s)$ 의 가중합(weighted sum)으로 나타낼 수 있다. 다항식의 차수를 3으로 설정함으로써 곡선은 연속적인 기울기를 갖는다. 따라서, 이렇게 구성된 spline 함수는 내부에너지의 요구 조건을 자연스럽게 충족시킨다.

일반적으로 spline 함수는 식 (4)와 같이 표현된다.

$$x(s) = \sum_{n=0}^{N_B-1} x_n B_n(s) \tag{4}$$

여기서  $x_n$ 은 기저함수  $B_n(s)$ 의 가중치이다. 식 (4)는 행렬을 이용하여 식 (5)처럼 나타낼 수 있다.

$$x(s) = \mathbf{B}(s)^T \mathbf{Q}_x \tag{5}$$

여기서  $\mathbf{B}(s)$ 와  $\mathbf{Q}_x$ 의 의미는 다음과 같다.

$$\mathbf{B}(s) = [B_0(s), B_1(s), K, B_{N_B-1}(s)]^T \tag{6}$$

$$\mathbf{Q}_x = [x_0, x_1, K, x_{N_B-1}]^T \tag{7}$$

매개함수를 이용한 곡선  $r(s)$ 은 행렬을 이용하여 식 (8)처럼 표현된다.

$$\mathbf{r}(s) = \mathbf{U}(s)\mathbf{Q} \tag{8}$$

여기서  $\mathbf{U}(s)$ 와  $\mathbf{Q}$ 는 식 (9)의 의미를 가지고 있다.

$$\mathbf{U}(s) = \mathbf{I}_2 \otimes \mathbf{B}(s)^T = \begin{pmatrix} \mathbf{B}(s)^T & 0 \\ 0 & \mathbf{B}(s)^T \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_x \\ \mathbf{Q}_y \end{pmatrix} \tag{9}$$

식 (9)에서 연산자  $\otimes$ 는 Kronecker product를 의미한다. 제어점 벡터  $\mathbf{Q}$ 는 제어점들의  $x$  좌표 값들로 구성된  $\mathbf{Q}_x$ 와  $y$  좌표 값들로 구성된  $\mathbf{Q}_y$ 로 이루어진 열벡터이다.

## 2. 모양 행렬

초기 곡선이 목표곡선으로 수렴해 갈 때, 곡선의 변화를 기술하기 위해서 제어점 벡터의 차원보다 작은 차원을 갖는 모양공간을 이용하는 것이 효과적이다. 모양공간은  $N_X \ll N_Q$  차원의 벡터들로 구성되며, 모양공간은 기본 곡선에서 가능한 곡선의 변형들의 집합으로 정의된다. 이러한 모양공간을 선형 공간으로 가정하면 모양공간은 행공간  $R(\mathbf{W})$ 와 남은 빈(left null) 공간  $N(\mathbf{W}^T)$ 으로 나누어진다.  $R(\mathbf{W})$ 는  $N(\mathbf{W}^T)$ 와 서로 직교(orthogonal)한다<sup>[8]</sup>.

$R(\mathbf{W})$ 는  $N_X$  요소로 형성된 벡터들의 집합이다. 모양의 굴곡을 유지하려면, 제어점의 변이를 낮은 차원의 모양공간으로 제한하는 것은 타당하다. 이때 제한되지 않은 제어점 벡터  $\mathbf{Q}$ 는 불안정한 active contour를 생성시킬 수 있다<sup>[9]</sup>.  $R(\mathbf{W})$ 에서 곡선의 변화는 모양공간 벡터  $\mathbf{X}$ 의 제어점 벡터  $\mathbf{Q}$ 로의 선형 사상으로 나타낼 수 있다.

$$\mathbf{Q} - \mathbf{Q}_0 = \mathbf{W}\mathbf{X} \tag{10}$$

여기서  $\mathbf{W}$ 는  $N_Q \times N_X$  의 모양 행렬이며,  $\mathbf{Q}_0$ 는 초기 곡선의 제어점 벡터이다. 모양공간 벡터  $\mathbf{X}$ 는 초기 곡선으로부터의 변화를 묘사한다.

본 논문에서는 6개의 변수를 가진 affine 모델( $N_X = 6$ )을 사용하여 곡선의 변화를 기술한다. 이러한 변화는 초기 곡선의 제어점 벡터  $\mathbf{Q}_0$ 와 모양행렬  $\mathbf{W}$ 에 의해서 식 (10)으로 나타낼 수 있다.

$$\mathbf{W} = \begin{pmatrix} 1 & 0 & \mathbf{Q}_{x0} & 0 & 0 & \mathbf{Q}_{y0} \\ 0 & 1 & 0 & \mathbf{Q}_{y0} & \mathbf{Q}_{x0} & 0 \end{pmatrix} \tag{11}$$

여기서 모양행렬  $\mathbf{W}$ 의 처음 두 열은 수평 또는 수직 방

향으로의 곡선의 이동을 나타낸다. 행렬  $\mathbf{W}$ 의 각 열은 모양행렬  $\mathbf{W}$ 에 의해서 생성되는 행공간  $R(\mathbf{W})$ 를 구성하는 기저벡터 (basis vector)를 형성한다. 그러나, 각 기저벡터가 다른 기저벡터와 반드시 직교할 필요는 없다.

### 3. 공간 $R(\mathbf{W})$ 로의 투사

객체의 경계를 찾기 위해서 추정된 곡선  $\mathbf{r}(s)$ 와 목표곡선  $\mathbf{r}_d(s)$  사이의 모양 차이를 나타내는 식 (12)와 같은 곡선의 왜곡 정도를 정의하였다.

$$\|\mathbf{r}(f(s)) - \mathbf{r}_d(s)\|^2 = \frac{1}{L} \int_0^L |\mathbf{r}(f(s)) - \mathbf{r}_d(s)|^2 ds \quad (12)$$

여기서  $L$ 은  $s$ 축 위에서 전체 길이이다. 그러므로, 추정된 곡선이 목표곡선과 정확히 같다면 식 (10)은 이론적으로 0의 값을 가질 것이다. 식 (12)에서  $f(s)$ 는 두 곡선의 비교되는 점을 정합시키기 위해 사용된 매개변수  $s$ 의 조 절함수이다. 두 곡선이 유사하다면,  $\|\mathbf{r}(s) - \mathbf{r}_d(s)\|$ 는  $[\mathbf{r}(s) - \mathbf{r}_d(s)] \cdot \mathbf{n}(s)$ 의 형태로 근사화할 수 있다<sup>[9]</sup>. 식 (12)의 적분을 합으로 바꾸고 모양의 차이를 측정하기 위해 법선 벡터(normal vector)  $\mathbf{n}(s)$ 를 사용한다면, 최소화하는 비용함수는 식 (13)으로 표현된다.

$$\|\mathbf{r} - \mathbf{r}_d\|^2 \approx \frac{1}{N} \sum_{i=1}^N [(\mathbf{r}_d(s_i) - \mathbf{r}(s_i)) \cdot \mathbf{n}(s_i)]^2 \quad (13)$$

$N$ 은  $s$ 축에서 똑같은 간격으로 표본화된 점들의 수이다. 식 (13)을 초기 곡선  $\mathbf{r}_0$ 와 모양공간 벡터로 표현하면 식 (14)로 재정의된다.

$$\|\mathbf{r} - \mathbf{r}_d\|^2 \approx \frac{1}{N} \sum_{i=1}^N [(\mathbf{r}_d(s_i) - \mathbf{r}_0(s_i)) \cdot \mathbf{n}(s_i) - \mathbf{n}^T \mathbf{U}(s_i) \mathbf{W}(\mathbf{X} - \mathbf{X}_0)]^2 \quad (14)$$

식 (14)를 최소화하고 추정된 모양공간 벡터를  $\mathbf{X}^*$ 를 구하기 위해 최소자승법(least squares method)을 사용한다. 즉, 최소값  $\mathbf{X}^*$ 는  $\partial \|\mathbf{r} - \mathbf{r}_d\|^2 / \partial \mathbf{X} = 0$ 을 만족할 때 식 (15)를 이용하여 구한다.

$$\mathbf{X}^* = \left( \sum_{i=1}^N \rho_i \mathbf{W}^T \mathbf{U}^T \mathbf{n} \mathbf{n}^T \mathbf{U} \mathbf{W} \right)^{-1}$$

$$\times \left( \sum_{j=1}^N \rho_j \mathbf{W} \mathbf{U}^T \mathbf{n} (\mathbf{r}_d - \mathbf{r}_0)^T \mathbf{n} \right) \quad (15)$$

여기서  $\rho_i$ 는 곡선의 표본들이 등간격으로 추출되지 않을 수 있으므로 이를 보상하기 위한 가중치이다.  $\mathbf{X}^*$ 를 구하면 식 (10)을 이용하여 최적의 제어점 벡터  $\mathbf{Q}$ 를 얻는다. 보다 정확한 결과를 얻기 위하여 구해진  $\mathbf{r}(s)$ 를 다시 초기 곡선  $\mathbf{r}_0(s)$ 로 설정하여 최소화 알고리즘을 반복 수행한다.

수학적으로 모양공간 벡터  $\mathbf{X}^*$ 는 초기 곡선으로부터의 변화를 공간  $R(\mathbf{W})$  위로 투사하여 표현한 것이다.  $R(\mathbf{W})$ 는 너무나 제한적인 공간이기 때문에 active contour가 임의의 모양 변화를 표현하는데 부적절하다. 따라서 이러한 문제를 보완하기 위하여 공간  $N(\mathbf{W}^T)$ 를 고려하며, 이에 대해서는 IV장에서 기술한다.

### 4. 목표곡선 추정

Active contour 알고리즘의 목적은 사용자에게 의해 주어진 초기 곡선을 영상 특징점에 정합시키는 것이기 때문에, 식 (12)의  $\mathbf{r}_d(s)$ 를 주어진 영상으로부터 추정해야 한다. 기존의 active contour 알고리즘은 배경이 없는 영상에서 개발되었기 때문에 복잡한 배경이 있는 영상에는 효과적인 성능을 얻지 못한다. 본 절에서는 복잡한 배경을 갖는 영상으로부터 목표곡선  $\mathbf{r}_d(s)$ 를 추정하는 방법을 제시한다.

기존 active contour 알고리즘처럼 단순한 윤곽선 검출기를 사용한다면 복잡한 배경을 갖는 영상에서 만족스러운 영상분할 결과를 얻기 힘들다. 본 논문에서는 SUSAN(smallest univalue segment assimilating nucleus) 연산자를 사용하여 영상의 특징점을 추출하였고 추출된 특징점을 이용하여 목표곡선  $\mathbf{r}_d(s)$ 를 정의하였다. SUSAN 윤곽선 추출기는 형태학적(morphological) 윤곽선 추출기 또는 Canny 윤곽선 추출기에 비해 연산량이 적으면서도 복잡한 영상에 대하여 좋은 성능을 가진다.

SUSAN 윤곽선 추출기의 동작은 3단계로 구성된다<sup>[10]</sup>. 첫번째 단계는 원형 마스크의 중심을 검사하는 화소의 위치로 이동시킨다. 두번째 단계는 식 (16)에 주어진 식을 이용하여 마스크의 중심과 유사한 화소의 숫자를 구한다.

$$c(\mathbf{p}, \mathbf{p}_0) = \exp \left( - \left| \frac{I(\mathbf{p}) - I(\mathbf{p}_0)}{t} \right|^6 \right) \quad (16)$$

$$n(\mathbf{p}_0) = \sum_{\mathbf{p}} c(\mathbf{p}, \mathbf{p}_0) \quad (17)$$

여기서  $\mathbf{p}$ 는 원형 마스크 안에 있는 임의 화소의 위치 벡터이고,  $\mathbf{p}_0$ 는 마스크 중심의 위치 벡터이다. 마스크 중심과 비슷한 화소의 숫자를 USAN 영역이라고 정의한다. 그리고 식 (16)의 연산 속도를 개선하기 위하여  $I(\mathbf{p}) - I(\mathbf{p}_0)$ 의 모든 값에 대하여 식(16)의 연산 결과를 lookup table로 만든다. 마지막 단계는 영상으로부터 윤곽선을 검출하기 위하여 기하학적 임계값에서 USAN 영역을 빼는 것이다.

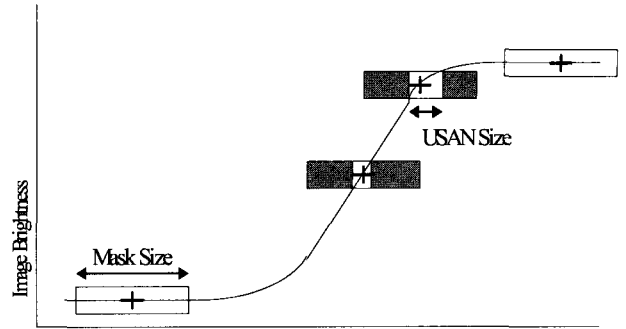
$$R(\mathbf{p}_0) = \begin{cases} g - n(\mathbf{p}_0) & \text{if } \mathbf{p}_0 < g \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

여기서  $R(\mathbf{p}_0)$ 는 영상의 윤곽선을 나타내고,  $g$ 는 기하학적 임계값이다. 만약 잡음이 없을 때 윤곽선을 검출하고자 한다면, 임계값  $g$ 는 필요가 없다. 그러나 최적의 잡음 제거를 위한 임계값은  $3n_{max}/4$ 이다<sup>[11]</sup>.

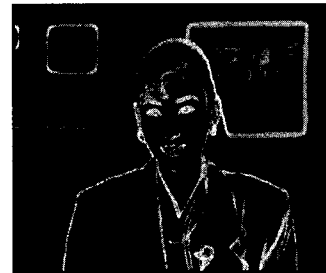
그림 1(a)는 SUSAN 연산자의 개념을 1차원 공간에 나타낸 것이다. 그림 1(a)에서 곡선은 영상의 밝기 변화를 나타내고, 흰색 영역은 마스크의 중심과 유사한 화소를 갖는 지역을 나타낸다. 그림 1(a)에서 보는 것처럼, 영상이 윤곽선의 중심에 있을 때 가장 적은 흰색 영역을 갖는다. 따라서 윤곽선의 중심에서  $R(\mathbf{p}_0)$ 는 제일 큰 값을 가진다. 그림 1(b)는 SUSAN 연산자를 영상에 적용하여 얻은 결과를 보여주고 있다.

가상의 목표곡선  $\mathbf{r}_d(s)$ 를 추출하기 위해 SUSAN 알고리즘으로부터 얻어진 영상을 사용하고 가상의 목표곡선  $\mathbf{r}_d(s)$ 는 특징점 위에 있다고 가정한다. Active contour 알고리즘의 목적은 초기 곡선을 객체의 특징점에 정합시키는 것이므로 초기 곡선의 주변의 영역에서 특징점을 찾는다. 그러므로 주어진 초기 곡선을 따라 일정 길이의 법선 벡터를 가지고 검색 영역을 설정한다. 즉, 식 (13)에 따라 초기 곡선의 표본화된 법선 위에서 특징점을 추출한다. 곡선  $\mathbf{r}(s)$ 의 표본화된 점을 따라서 법선들이 생성되면, 가상 목표곡선  $\mathbf{r}_d(s)$ 를 따라  $\mathbf{r}_d(s_i)$ 의 시퀀스를 구할 수 있다.

$\mathbf{r}(s_i)$ 는 부동 소수점의 좌표를 가지므로, 법선 벡터도 부동 소수점 요소로 구성된다. 따라서  $\mathbf{r}_d(s_i)$ 의 밝기는 양방향 선형보간법에 의해서 구해지고,  $\mathbf{r}_d(s_i)$ 는 각 법선에서 최대값을 갖는 점으로 설정된다.



(a)



(b)

그림 1. SUSAN 알고리즘  
Fig. 1. SUSAN algorithm

#### IV. 동적 그래프 검색

III장에서는 초기 곡선의 변화를 공간  $R(\mathbf{W})$ 에 투사하였다. 여기서 저수준(low level) 영상처리로부터 발생하는 불필요한 특징점들을 무시할 수 있고, 곡선의 전체적인 변화를 기술할 수 있다. 그러나 모양 행렬로 구성된 공간이 단지 6개의 기저벡터(basis vector)로 구성되어 있기 때문에 곡선의 상세한 변화를 기술하기는 어렵다.

이 장에서는 정합된 곡선을 보상하여 최종 객체의 윤곽선을 구하기 위하여 공간  $R(\mathbf{W})$ 에서 표현하지 못한 공간  $N(\mathbf{W}^T)$ 에서의 변화를 기술한다. 그러나 객체의 정의가 매우 주관적이어서 유한 차원의 공간으로 객체를 묘사할 수 없기 때문에 모양공간의 실질적인 차원을 알 수 없다. 그러므로 공간  $N(\mathbf{W}^T)$ 에서의 변화를 기술하기 위해 일반적인 선형대수적 접근을 사용하지 않는다. III장에서 얻은 정합된 곡선 위의 특징점을 공간  $R(\mathbf{W})$ 에서 공간  $N(\mathbf{W}^T)$ 의 정보를 이용하지 않고 정확하게 묘사된 특징점으로 간주한다. 그 밖의 곡선 위의 다른 점들은 공간  $R(\mathbf{W})$ 로의 투사를 통해 왜곡되었다고 간주한다. 여기서 공간  $R(\mathbf{W})$ 에서 정확하게 묘사된 특징점들의 시퀀스를 근원점(seed points)의 시퀀스로 정의한다. 최종 객체 윤곽선을 추정하기 위해, 근원점 시퀀스를 사용하여 III장에서 정합된 곡선의 왜곡된 점들을 수정한다.

객체의 윤곽선은 그래프 검색(graph search) 문제로 전환될 수 있는데, 그 최종 목적은 시작점과 최종점 사이의

최적 경로를 찾는 것이다. 여기서 주요 개념은 각 화소는 그 주변 화소들로 방향성과 가중치를 가지고 연결되어 있는 노드(node)로 간주되어, 하나의 영상은 이러한 노드로 구성된 그래프로 간주되는 것이다<sup>[12]</sup>. 먼저 근원점 시퀀스의 첫번째 점과 두번째 점을 각각 시작점과 최종점으로 설정한다. 그리고 그 두 점간의 최적 경로가 검색되면, 최종점은 새로운 시작점으로 설정되고 시퀀스의 세번째 점이 새로운 최종점이 된다. 이 과정은 근원점 시퀀스의 마지막 점까지 반복적으로 수행된다. 최적 경로는 시작점에서 목표점까지의 최소 누적 비용을 갖는 경로로 정의된다.

최소비용 경로는 객체의 윤곽선에 대응되므로, 윤곽선이 극명한 곳에서의 화소는 낮은 국부 비용(local cost)을 나타내며, 평탄한 지역의 화소는 높은 비용을 가진다. 다양한 윤곽선의 특징들을 국부비용함수에 반영하기 위하여, 정적비용(static cost) 함수와 동적비용(dynamic cost) 함수로 국부 함수를 정의하고, 전체 비용은 정적비용 함수와 동적비용 함수의 가중합으로 정의한다.

### 1. 정적 비용 함수

정적비용 함수는 Laplacian 영 교차점(zero crossing)과 경사도(gradient)의 크기(magnitude)를 이용하여 식 (19)와 같이 정의한다.

$$f(x, y) = \alpha_z \Psi_z(x, y) + \alpha_G \Psi_G(x, y) \quad (19)$$

여기서  $\alpha_z$  와  $\alpha_G$ 는 각 비용함수의 가중치이다.  $\Psi_z$  와  $\Psi_G$ 는 각각 Laplacian 영 교차점과 경사도의 크기로부터 얻어진 비용함수이다. 다중 커널(kernel)의 사용이 인간 시각 특성을 보다 잘 반영하므로<sup>[13]</sup>, 각 비용함수 요소는 다중 커널을 사용하여 정의하였다. 각 커널은 2차원 가우스 확률분포 함수의 서로 다른 표준 편차를 가지고 생성된다.  $\Psi_z$ 와  $\Psi_G$ 의 정의에 다중 커널을 사용하였으므로 각 요소는 해당되는 커널의 가중합으로 정의된다.

LOG(Laplacian of Gaussian) 커널  $\Phi(x, y)$ 는 다음과 같이 정의된다.

$$\Phi(x, y) = \frac{-1}{\pi\sigma^4} \left[ 1 - \frac{x^2 + y^2}{2\sigma^2} \right] \exp \left[ -\frac{x^2 + y^2}{2\sigma^2} \right] \quad (20)$$

여기서 커널 중심의 음수 영역은 반지름  $\sqrt{2}\sigma$  크기를 갖는 원판이므로<sup>[14]</sup>,  $1/\sqrt{2}$ 의 배수로  $\sigma$ 를 선택하면, 정수 크기의 커널 마스크를 구할 수 있다.

### 2. 동적 비용 함수

객체의 윤곽선 가운데 일부는 주변의 강한 윤곽선에

비해 다소 희미한 윤곽선을 가질 수 있다. 주변의 강한 윤곽선은 상대적으로 낮은 비용함수 값을 취하므로 최적 경로는 원하는 윤곽선이 아닌 강한 윤곽선으로 이동할 것이다. 그러므로 이러한 문제를 해결하기 위해 경사도 크기(gradient magnitude) 비용함수는 동적으로 할당되어야 한다.

근원점 시퀀스의 이전 구간에서 얻어진 영상분할 결과에 해당하는 최적 경로로부터 동적비용 함수를 정의할 수 있다. 윤곽선 특성의 점진적인 변화를 수용하기 위해 동적비용 함수를 근원점 구간이 변화될 때마다 갱신한다. 동적비용 함수는 이전 최적 경로에서 비용함수의 히스토그램으로부터 구해진다<sup>[15]</sup>. 높은 경사도 크기를 갖는 화소가 객체의 윤곽선 주변에 있을지라도 이전의 최적 경로가 높은 경사도 값을 가지는 화소를 통과한 적이 없다면, 경사도 크기 한 동적비용 함수는 경험하지 않은 경사도 값의 화소에서 최대값을 갖는다.

## V. 실험 결과 및 분석

본 논문에서는 30Hz의 프레임율을 갖는 CIF(352×288) 동영상을 사용하여 제안된 영상분할 알고리즘의 동작을 실험하였다. 그림 2는 동적비용 함수의 효과를 보여준다. IV장에서 기술한 것처럼, 원하는 객체 윤곽선의 일부가 약한 경사도 크기를 가질 경우, 최적 경로는 그림 2(a)와 같이 원하는 객체의 윤곽선이 아닌 곳을 지난다. 그림 2에서 원하는 객체의 윤곽선은 뺨에 위치하지만, 가까운 곳에 있는 입술 보다 윤곽선의 경사도 낮은 값이므로 최적 경로가 입술 위를 지나고 있다. 이러한 문제를 해결하기 위해 IV장의 동적비용 함수를 고려한다. 그림 2(b)는 최적 경로가 이전 구간과 유사한 윤곽선을 어떻게 지나는지 보여주고 있다. 그림 2(b)에서 어깨의 백색 선은 오로지 정적비용 함수만을 사용하여 구했으나, 뺨의 백색 선은 어깨 부분으로부터 점진적으로 변해 온 동적비용 함수를 이용하여 얻었다.

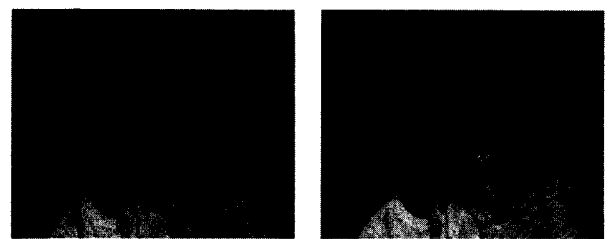


그림 2. 동적 비용 함수의 효과  
Fig. 2. Effect of dynamic cost

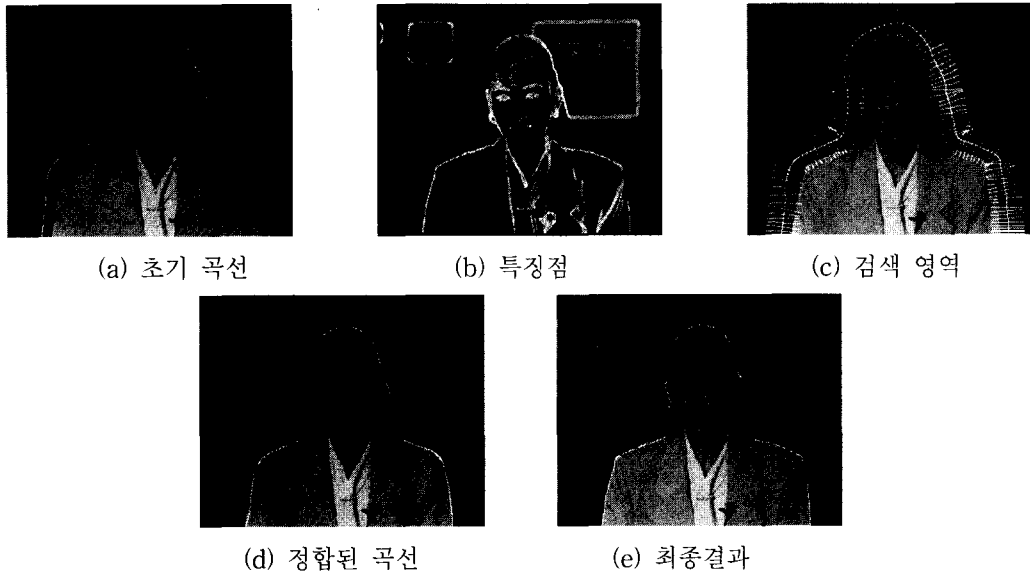


그림 3. AKIYO 영상에 대한 결과  
Fig. 3. Results for AKIYO image

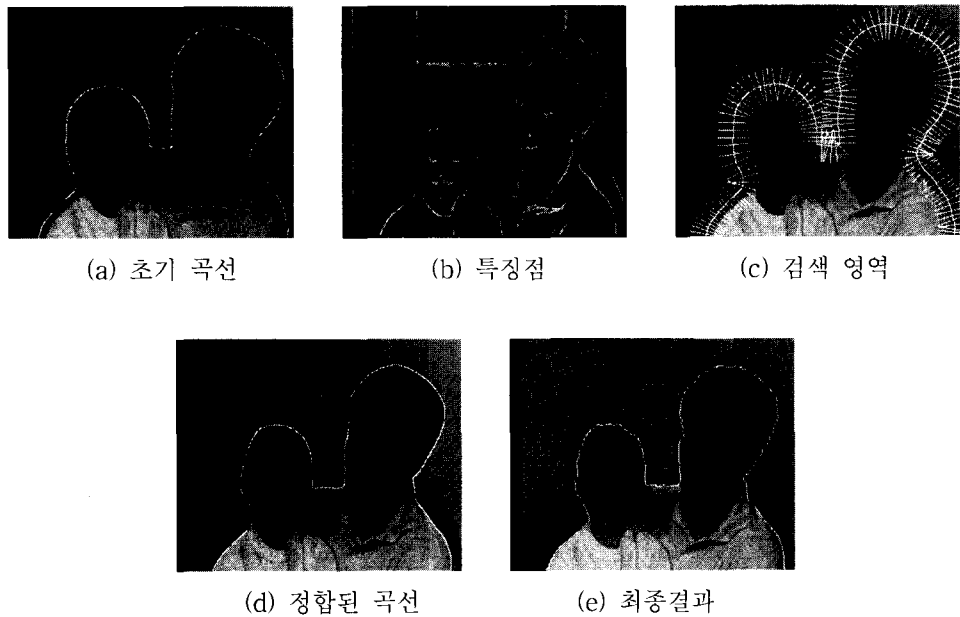


그림 4. MOTHER AND DAUGHTER 영상에 대한 결과  
Fig. 4. Results for MOTHER AND DAUGHTER image

그림 3은 제안된 방법을 AKIYO 영상에 적용하여 얻은 실험결과이다. 그림 3(a)는 사용자가 마우스와 같은 입력장치를 통하여 입력한 초기 곡선이다. 그림 3(b)는 SUSAN 연산자를 사용하여 얻은 경사도 영상이다. 기하학적 임계값에서 USAN 영역의 크기와의 차이를 표시하기 때문에 윤곽선 부분에서 가장 밝은 색을 보인다. 그림 3(c)에서 백색선 위의 점들은 초기 곡선을 따라서 만들어진 법선 벡터를 통해 얻어진다. 그림 3(c)에서 백색선 위

의 점들은  $r(s_i)$ 에 대한  $r_d(s_i)$ 의 후보점들이다. 그림 3(d)는 반복 수행된 최소자승법을 통해서 얻어진 정합된 곡선이다. 그림 3(e)는 동적 그래프 검색 알고리즘을 사용하여 공간  $N(W^T)$ 의 손실을 보상한, 최종적인 영상분할의 결과이다. 그림 3(d)에서는 6개의 매개변수 근사화와 B-spline 곡선의 매끄러운 연결 특성으로 인해 AKIYO의 귀와 팔과 어깨의 연결부분에서 정합된 곡선이 실제 객체의 윤곽선에서 벗어났다. 그러나 그림 3(e)의 마지막

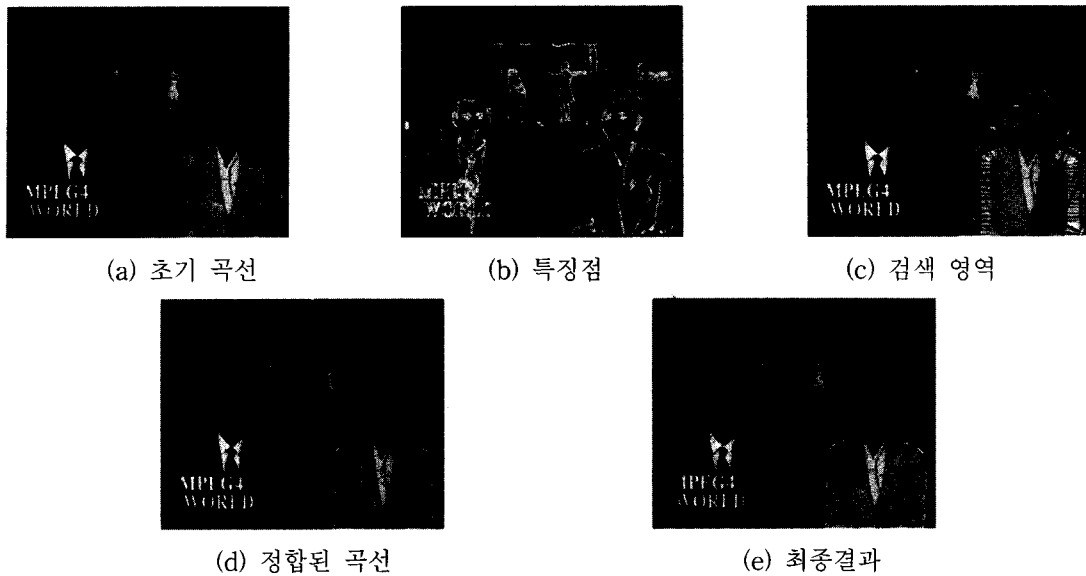


그림 5. NEWS 영상에 대한 결과  
 Fig. 5. Results for NEWS image

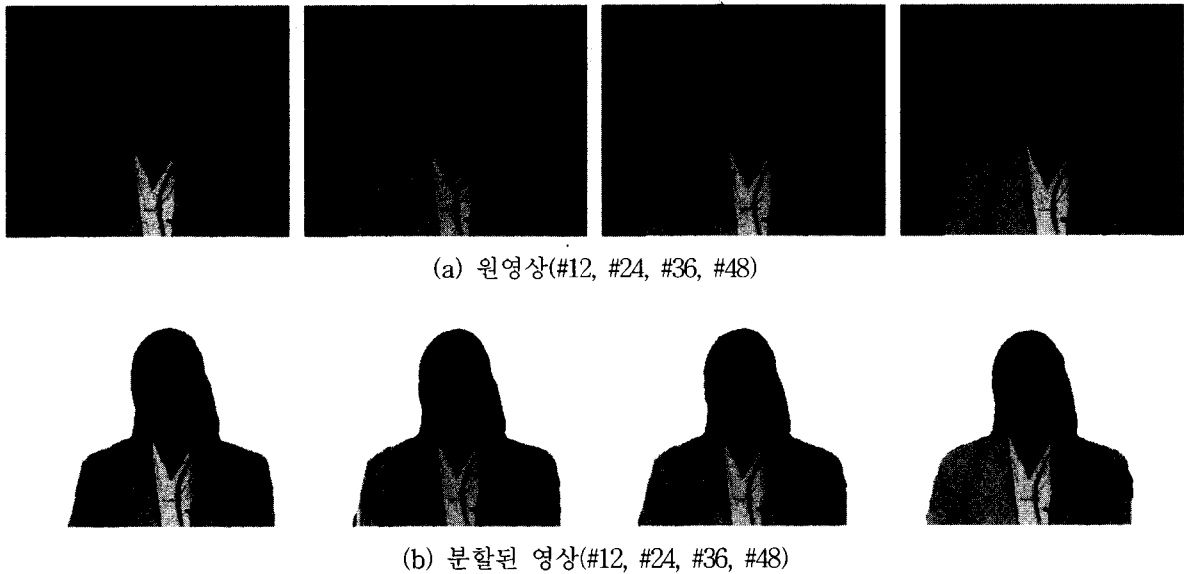


그림 6. AKIYO 시퀀스에 대한 영상분할 결과  
 Fig. 6. Segmentation results for AKIYO sequence

곡선은 객체의 윤곽선을 보다 정확하게 추정하였다.

그림 4는 MOTHER AND DAUGHTER 영상에 대한 영상분할 결과이다. 이 영상은 AKIYO 영상보다 복잡한 모양을 갖는다. 그림 4(e)는 객체의 머리카락 부분에서 그림 4(d)보다 정확한 객체의 윤곽선을 표현하고 있다.

일반적으로 기존의 active contour 알고리즘은 단순한 배경(homogeneous background)를 갖는 영상을 위해 설

계되었기 때문에, 다소 복잡한 배경을 갖는 객체에 대해서는 부적절한 결과를 보였다. 그림 5에서 보듯이, 본 논문에서 제안한 알고리즘은 복잡한 배경의 객체에 대해서도 만족할 만한 결과를 얻는다. 그림 5는 AKIYO 영상과 MOTHER AND DAUGHTER 영상보다 복잡한 배경을 갖는 NEWS 영상에 대한 영상분할 결과이다.

비디오 객체의 윤곽선을 검출하고 그 윤곽선을 이후의



연속된 프레임에서 추적하기 위하여, 첫 프레임에서 정의된 비디오 객체의 윤곽선 정보는 다음 프레임으로 투사되고 투사된 윤곽선 정보는 다음 프레임을 위한 active contour 알고리즘의 초기 곡선으로 사용된다. 그림 6은 AKIYO 영상 시퀀스에 대한 영상분할의 결과이다. 첫 프레임에서 비디오 객체가 정의되면 정합된 곡선의 제어점 벡터  $Q$ 가 다음 프레임으로 전달된다. 여기서 그래프 검색 알고리즘에서 얻어진 최종 곡선 대신에 정합된 곡선의 제어점 벡터  $Q$ 를 사용하여 다음 프레임을 위한 초기 곡선을 설정하는 것은 최종 곡선보다 제어점 벡터  $Q$ 가 조작하기 용이하기 때문이다. 다음 프레임을 위한 초기 곡선이 설정되면, 제안된 알고리즘을 사용하여 비디오 객체를 추출한다. 이 과정을 영상 시퀀스의 끝까지 반복적으로 수행한다. 그림 6은 마지막 영상분할 결과를 매 12 프레임마다 표시한 것이다.

일반적으로 자동 영상분할 알고리즘은 20~50 프레임 동안 과도응답(transient response)을 보인다<sup>[16]</sup>. 과도응답시간은 비디오 객체를 찾기 위해 사용되므로, 과도응답 시간 동안의 영상분할의 결과는 별로 의미가 없다. 이러한 과도응답은 시공간(spatio-temporal) 영상분할 알고리즘이 통계적 가설검증 방법을 사용하기 때문에 오보 확률(false alarm probability)과 누락 확률(missing probability)에 의해서 발생한다. 그림 6에서 제안된 사용자 개입 알고리즘은 과도응답을 생성하지 않는다.

## VI. 결 론

본 논문에서는 영상 시퀀스에서 비디오 객체를 추출하기 위하여 active contour를 이용한, 새로운 사용자 개입 영상분할 알고리즘을 제안하였다. 제안한 알고리즘은 복잡한 배경으로부터 객체의 윤곽선을 추출하기 위하여 SUSAN 연산자를 사용하였다. 또한 곡선의 변화를 묘사하기 위해 모양공간 벡터를 사용했기 때문에 저수준 영상 처리에서 발생할 수 있는 잘못 추출된 특징점을 효과적으로 무시할 수 있다는 장점이 있다. 제안한 알고리즘은 6 변수 affine 모델의 제약을 보상하기 위해 동적 그래프검색 알고리즘을 사용했으며, 동적 그래프검색 알고리즘은 모양 행렬로 기술할 수 없는 곡선의 상세한 부분을 기술할 수 있다.

## 참 고 문 헌

- [1] ISO/IEC FDIS 14496-2: "Information technology - generic coding of audio-visual objects, Part 2: visual," ISO/IEC JTC1/SC29/WG11, Oct. 1998.
- [2] 김대희, 김민호, 호요성, "MPEG-4 동영상 부호화를 위한 영상 객체 추출 기법," 대한전자공학회지, 제26권, 제7호, pp. 714-723, 1999년 7월.
- [3] C. Gu and M.C. Lee, "Semiautomatic segmentation and tracking of semantic video objects," IEEE Trans. Circuit and System for Video Technology, vol. 8, no. 5, pp. 572-584, Sept. 1998.
- [4] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active contour models," Proceedings of First International Conference on Computer Vision, pp. 259~269, 1987.
- [5] A. Amimi, T. Weymouth and R.C. Jain, "Using dynamic programming for solving variational problems in vision," IEEE Trans. Patt. Anal. Mach. Intel., vol. 12, no. 9, pp. 855-867, Sept. 1990.
- [6] D.J. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," CVGIP:Image Understanding, vol. 55, no. 1, pp. 14-26, Jan. 1992.
- [7] J.D. Foley, A. Dam, S.K. Feiner, J.F. Hughes and R.L. Phillips, Introduction to Computer Graphics, Addison-Wesley, New York, 1995.
- [8] G. Strang, Linear Algebra and Its Applications, 3rd edn., Harcourt Brace Jovanovich, 1988.
- [9] A. Blake and M. Isard, Active Contours, Springer, London, 1998.
- [10] S.M. Smith, "Flexible filter neighborhood designation," Proc. 13th Int. Conf. on Pattern Recognition, vol.1, pp. 206-212, 1996.
- [11] S. M. Smith, "SUSAN nonlinear noise reduction," <http://www.fmrib.ox.ac.uk/fsl/susan>
- [12] E. N. Mortensen and W. A. Barrett, "Interactive segmentation with intelligent scissors," Graphical Models and Image Processing, pp. 349-384, 1998.
- [13] D. Marr and E. Hildreth, "Theory of edge detection," Proc. R. Soc. Lond., B 270, pp. 187-217, 1980.
- [14] R. M. Haralick and L. G. Shapiro, Computer and Robot Vision, Addison-Wesley Publishing, 1992.
- [15] E. N. Mortensen and W. A. Barrett, "Intelligent scissors for image compositions," Proc. of the ACM SIGGRAPH 95: Computer Graphics and Interactive Techniques, pp. 191-198, 1995.
- [16] D. Kim and Y. S. Ho, "A user-assisted segmentation algorithm using B-Spline curves," Proceedings of SPIE Visual Communications and Image Processing, pp. 734-744, Jan. 2001.

## 저 자 소 개



김 대 희(정회원)

1995년 2월 서울시립대학교 제어계  
측공학과 졸업(학사)

1997년 2월 광주과학기술원 정보통  
신공학과 졸업(석사)

2003년 2월 광주과학기술원 정보통  
신공학과 졸업(박사)

2003년 2월~현재 한국전자통신연  
구원 방송시스템연구그룹 연구원

<주관심분야: 영상신호 처리 및 압축, 멀티미디어  
통신, 비디오 객체 분할, 3 차원 그래픽>



호 요 성(정회원)

1981년 2월 서울대학교 전자공학과  
졸업 (학사)

1983년 2월 서울대학교 전자공학과  
졸업 (석사)

1983년 3월~1995년 9월 한국전자  
통신연구소 선임연구원

1989년 12월 University of  
California, Santa Barbara, Dept. of Electrical and  
Computer Eng. 졸업 (박사)

1990년 1월~1993년 5월 미국 Philips 연구소 Senior  
Research Member

1995년 9~현재 광주과학기술원 정보통신공학과 교  
수.

<주관심분야: 디지털 신호처리, 영상신호처리 및 압  
축, 초저속 영상통신, 디지털 TV와 고선명 TV 방  
식, 멀티미디어 통신, 디지털 워터마킹, 3차원 실감  
방송 기술>



안 충 현(정회원)

1985년 2월 인하대학교 해양학과  
졸업(학사)

1989년 2월 인하대학교 해양학과  
졸업(석사)

1995년 3월 지바대학교 자연과학연  
구과 졸업(박사)

1995년 3월~1995년 12월 지바대학

교 정보공학과 조수,

1996년 1월~현재 한국전자통신연구원 방송시스템  
연구그룹 3DTV시스템연구팀장.

<주관심분야: 원격탐사, 영상처리, 컴퓨터비전>