

논문 2005-42CI-1-9

활성 윤곽선 모델을 이용한 얼굴 경계선 추출

(Facial Boundary Detection using an Active Contour Model)

장 재 식*, 김 은 이**, 김 항 준***

(Jae Sik Chang, Eun Yi Kim, and Hang Joon Kim)

요 약

본 논문에서는 복잡한 환경에서 정확한 얼굴영역의 경계를 추출하기 위한 활성 윤곽선 모델(Active Contour Model)을 제안한다. 제안된 모델에서 윤곽선은 레벨 함수 ϕ 의 제로 레벨 집합으로 표현되고, 레벨 집합의 편미분 방정식을 통해 진화된다. 이 때, 제안된 모델에서는 윤곽선의 진화와 종료를 위해 2차원 가우시안 모델로 표현되는 피부색 정보를 이용한다. 이를 통해 잡음 및 다양한 포즈를 가지는 복잡한 영상에서도 정확한 얼굴 경계선을 얻을 수 있는 강건한 추출 방법이 구현된다. 제안된 방법의 유효성을 평가하기 위해서 다양한 영상에 대해서 실험이 이루어졌으며, 그 결과를 geodesic 활성 윤곽선 모델의 결과와 비교하였다. 실험 결과는 제안된 방법의 보다 나은 성능을 보여준다.

Abstract

This paper presents an active contour model for extracting accurate facial regions in complex environments. In the model, a contour is represented by a zero level set of level function ϕ , and evolved via level set partial differential equations. Then, unlike general active contours, skin color information that is represented by 2D Gaussian model is used for evolving and stopping a curve, which allows the proposed method to be robust to noise and varying pose. To assess the effectiveness of the proposed method, it was tested with several natural scenes, and the results were compared with those of geodesic active contours. Experimental results demonstrate the superior performance of the proposed method.

Keywords : Active contour model, Face detection, Level sets, Skin color model.

I. 서 론

개인의 사생활 보호, 안전에 대한 필요성의 급증으로 인해, 최근 사용자의 생체 특징을 이용한 생체 인증 시스템의 연구가 활발히 진행되고 있다. 이들 시스템에서는 개인 식별을 위해 지문, 홍채 그리고 얼굴과 같은 생체 특징들을 사용하고 있으며, 이들 중 얼굴을 이용한 시스템은 사용자에게 강제성을 띄지 않는 비 접촉 시스템으로 많은 관심을 받고 있다. 실제 시스템에의 적용

을 위해서는 포즈변화 및 복잡한 환경에서도 동작이 가능해야 되고, 이를 위해서는 잡음 및 포즈에 강건한 얼굴 추출 방법의 개발이 반드시 선행되어야 한다.

지금까지 많은 얼굴 추출 방법들이 개발되었으며, 이들은 크게 지식 기반의 방법, 특징 방법, 템플릿 매칭 방법, 그리고 모양 기반의 방법으로 분류될 수 있다^[1]. 이들 중에서 템플릿 매칭 방법은 단순성으로 인해 가장 널리 사용되고 있다. 이 방법에서는 얼굴 특징과 그들 간의 관계를 표현할 수 있는 템플릿들을 구한 후, 입력 영상과 템플릿의 비교를 통해 입력영상에서 얼굴 영역을 찾는다^[1]. 이 템플릿 매칭은 정면 얼굴을 포함하는 영상에서는 매우 효과적이지만 그렇지 않은 경우에는 좋은 결과를 생성하지 못한다. 이러한 한계성을 극복하기 위해서 가변 템플릿이 제안되었다^[1]. 가변 템플릿 방법에서는 입력 영상과 템플릿의 비교를 위해서 사전에 정의된 템플릿을 변형한 후 이 변형된 템플릿과 입력영

* 학생회원, *** 정회원, 경북대학교 컴퓨터공학과
(Dept. of Computer Eng. at Kyungpook National Univ.)

** 정회원, 건국대학교 인터넷미디어공학부-교신전자
(Dept. of Internet and Multimedia Eng., Konkuk Univ.)

※ 본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 육성·지원사업의 연구결과로 수행됨.
접수일자: 2004년9월27일, 수정완료일: 2005년1월10일

상을 비교한다. 이러한 동작은 적합한 템플릿이 발견될 때 까지 반복적으로 수행 된다^[1]. 이 방법은 좋은 성능을 제공하지만, 영상의 잡음에 민감하고 연산 비용이 많이 든다는 단점이 있다.

이러한 문제점들을 해결하기 위한 대안으로 최근 활성 윤곽선 모델이 객체 추출의 해결 방법으로 많이 사용되고 있다^[2, 3, 4]. 이 모델은 초기 경계선에서 시작하여 관심 있는 객체에게 잘 대응될 때까지 계속해서 진화한다. 이 모델은 객체 추출 뿐 아니라 고정되지 않은 객체의 추적에도 그 유용성이 입증 되었다^[2, 3, 4].

활성 윤곽선 모델은 SNAKES라는 이름으로 Kass등에 의해 처음 제안되었다^[2]. 이 방법은 여러 개의 maker particles를 이용하여 윤곽선을 표현하기 때문에 움직이는 객체를 추출하고 추적하는 데는 한계가 있다.

이를 위해 레벨 셋 방법^[5]에 기반을 둔 geometric 활성 윤곽선 모델이 제안되었다^[3]. 이 방법은 제로 레벨 집합으로 표현되는 폐곡선의 진화를 종료하기 위한 조건으로 영상의 밝기 기울기를 사용한다. Geometric 활성 윤곽선 모델은 아래의 미분 방정식으로 표현된다.

$$\frac{\partial \phi}{\partial t} = g(|\nabla I|) \cdot |\nabla \phi| \cdot (k + v) \quad (1)$$

$$k = \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) = \frac{\phi_{xx}\phi_y^2 - 2\phi_x\phi_y\phi_{xy} + \phi_{yy}\phi_x^2}{(\phi_x^2 + \phi_y^2)^{3/2}} \quad (2)$$

식(1)에서, v 는 상수이고 I 는 밝기 영상이다. 그리고 $g(x) = (1+x^2)^{-1}$ 는 에지 정지 함수이다. 이 모델은 에지 검출기와 유사하게 동작한다. 즉, 추출하고자 하는 객체를 둘러싼 초기 곡선은 자신의 법선 방향으로 $g(|\nabla I|) \cdot (k+v)$ 의 속도로 진화하다가, 큰 기울기 값을 가진 객체의 경계선 영역에서 멈춘다^[3, 4]. 실제 영상에서 객체의 경계는 작은 기울기 값을 가지는 경우가 많으며, 또한 영상에 잡음이 많은 경우 잡음영역은 객체 영역보다 더 큰 기울기 값을 가지는 경우가 많다. 이들 경우에서 이 모델은 정확한 객체의 경계선을 추출할 수 없다.

이런 문제점들을 해결하기 위해서, Chan등은 곡선의 진화 종료 조건으로 기울기 정보 대신 영역의 밝기 정보를 이용할 것을 제안했다^[4]. 이 연구에서는 입력 영상 I 가 일정한 밝기를 가지는 두 영역으로 구성되었으며, 추출하고자 하는 객체는 밝기 I_1 를 가지는 영역이라고 가정하였다. 이 때, 이 모델은 다음과 같이 표현된다.

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \cdot [\underbrace{\mu k - v}_{(A)} - \underbrace{\lambda_1(I - c_1)^2}_{(B)} + \lambda_2(I - c_2)^2] \quad (3)$$

이때 k 는 곡률이다. v, μ, λ_1 , 그리고 λ_2 는 상수들이다. c_1 and c_2 는 각각 객체의 내부와 외부의 밝기의 평균값이다. 식 (3)에서 (A)와 (B)항은 각각 내부영역과 외부영역에서의 밝기의 분산 값을 나타낸다. 여기서 초기 곡선은 (A)와 (B)의 값의 차이가 최소가 될 때 까지 (B) - (A)의 속도로 법선 방향으로 진화한다. 이 모델은 약한 경계선과 끊어진 경계선을 추출하는데 적합하다^[4]. 하지만, 일정한 밝기 값을 가지는 두 영역으로 구성된 간단한 영상에서만 잘 동작한다는 제약을 가진다.

본 논문에서는 활성 윤곽선 모델을 사용하여 얼굴 영역의 경계선을 추출하는 방법을 제안한다. 얼굴 경계선 추출은 Bayesian framework에서의 에너지 최소와 문제로 정형화 되고, 이 에너지는 활성 윤곽선 모델에 의해 최적화된다. 본 논문에서 사용되는 활성 윤곽선 모델은 자동적인 위상 변화를 가능하게 해주는 레벨 셋 방법을 사용해 표현 되고 진화된다. 이 때, 곡선의 진화종료 조건을 위해 2차원 가우시안 모델로 표현되는 피부 색 정보를 사용하였다. 피부색 정보의 사용으로 제안된 방법은 잡음에 강하고 초기 곡선의 위치에 관계없이 정확한 객체 추출 결과를 제공할 수 있다.

제안된 방법의 장점은 다음과 같다. 1)잡음과 얼굴의 다양한 방향과 포즈 변화에 강건한 얼굴 경계선 추출 방법이다. 2)얼굴 경계선 추출뿐만 아니라 추출된 얼굴의 정확한 추적에 적용이 가능한 방법이다. 따라서 제안된 방법은 얼굴 추적시스템으로의 확장이 용이하다. 이러한 장점들은 실험 결과를 통해 증명되었다.

본 논문의 구성은 다음과 같다. II장에서는 피부 색 모델에 대해 서술하였다. III장에서는 얼굴 경계선 추출 문제를 어떻게 에너지 최소화 문제로 표현하였는지 설명하고, IV장에서 최소화 기법에 대해 설명한다. V장에서는 실험 결과를 보여주고, VI장에서 결론을 맺는다.

II. 피부색 모델

모든 사람들은 각각 다른 피부 색조를 가지고 있다. 더욱이 다양한 조명 상황에서는 같은 사람의 피부의 색상도 다르게 관측 될 수 있다^[6]. 이런 현상은 피부의 색상차이보다는 밝기 차이에 의해 발생한다^[6]. RGB공간에서 칼라 표현은 색상 뿐 아니라 밝기도 포함하기 때문에, 피부색을 표현하는데 적합하지 않다. 피부색의 조사를 위해 입력영상은 다음의 (r, g)로 표현되는 크로매틱 색상 공간으로 변환된다.

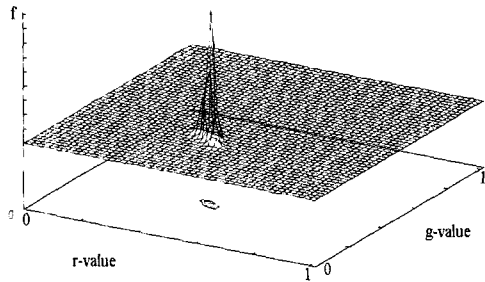


그림 1. 사람 얼굴의 색상 분포
Fig. 1. Color distribution of human faces.

표 1. 2차원 가우시안 파라미터
Table 1. Actual 2D-Gaussian parameters.

Parameters	Values
\bar{r}	117.588
\bar{g}	79.064
σ_r^2	24.132
$\rho_{X,Y} \sigma_r \sigma_g$	-10.085
$\rho_{X,Y} \sigma_r \sigma_g$	-10.085
σ_g^2	8.748

$$r = \frac{R}{R+G+B} \text{ and } g = \frac{G}{R+G+B}. \quad (4)$$

그림 1은 200장의 실험 영상으로부터 획득한 사람 얼굴의 색조 공간상에서의 칼라 분포를 보여주고 있다.

그림 1에서 보듯이, 피부색분포는 색조공간상에서 좁은 영역에 밀집되어져 있고 2차원 가우시안 분포로 근사화 될 수 있다. 따라서 피부색 모델은 아래와 같은 평균과 편차를 가지는 2차원 가우시안 모델 $N(m, \Sigma^2)$ 로 근사화 된다. 표 1은 실험 영상으로부터 획득한 사람 얼굴 칼라의 평균과 편차를 보여주고 있다.

$$m = (\bar{r}, \bar{g}) \text{ where } \bar{r} = \frac{1}{N} \sum_{i=1}^N r_i, \text{ and } \bar{g} = \frac{1}{N} \sum_{i=1}^N g_i.$$

$$\Sigma = \begin{bmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{bmatrix}. \quad (5)$$

III. 문제 정의

1. 사후 확률 최대화 추정

$S = \{s \mid 1 \leq s \leq M_1 \times M_2\}$ 는 2차원상의 공간 평면을 나타낸다. 이 때, 입력영상 $Y = \{Y_s\}$ 는 공간 S 상에서 정의되는 랜덤 필드이고, 각 요소 Y_s 는 크로매틱 색상 공간상에서 색상 벡터 값을 가지는 확률 변수이다. 본 논

문에서 입력 영상 Y 는 얼굴 영역 R 과 배경 영역 R^c 로 구성되어 있다고 가정한다. $X = \{X_s \mid X_s \in A\}$ 를 S 상에서 정의되는 레이블의 배열이라 하자. 이 때, X_s 는 집합 $A = \{\lambda, \lambda^c\}$ 에서 하나의 값을 가지는 확률 변수이다. λ 와 λ^c 은 각각 입력영상에서 얼굴 영역과 배경 영역을 나타내는 레이블이다. $\Gamma = \{\eta_s\}$ 을 공간 S 에 대한 주변 시스템이라고 정의할 때, η_s 는 화소 s 에 이웃한 화소들의 집합을 나타낸다. 본 논문에서는 X 는 다음의 특징을 만족하기 때문에 마르코프 랜덤 필드 (Markov random field: MRF)로 모델화 된다^[7, 8].

$$P(X_s = \omega_s \mid X_k = \omega_k, s \neq k) = P(X_s = \omega_s \mid X_k = \omega_k, k \in \eta_s). \quad (6)$$

ω 를 X 의 실현이라고 할 때, 얼굴 영역 추출은 관측 영상 y 로부터 레이블의 배열 ω 를 추정하는 문제로 표현된다. 본 논문에서는 최적화 척도로 사후 확률 최대화 추정 (maximum a posteriori: MAP) 을 사용한다. 따라서 얼굴 영역을 검출하기 위해서는 입력영상 y 에 대해 다음의 사후확률을 최대화 시키는 ω 를 추정해야 한다.

$$\omega^* = \arg \max_{\omega} P(X = \omega \mid Y = y) = \arg \max_{\omega} \frac{P(y|\omega)P(\omega)}{P(y)}. \quad (7)$$

확률 $P(y)$ 는 상수 값을 가지기 때문에, 위의 식은 아래와 같이 표현 될 수 있다:

$$\omega^* = \arg \max_{\omega} P(y|\omega)P(\omega). \quad (8)$$

식 (8)에서 사후 확률은 사전 확률 함수 $P(\omega)$ 와 우도 함수 $P(y|\omega)$ 로 나누어진다. 여기서 사전 확률은 임의 화소의 레이블은 이웃한 화소의 레이블에 영향을 받는다는 가정을 내포하고, 우도 함수는 추출하고자 하는 영역의 칼라 분포가 일정하다는 가정을 내포하고 있다.

사전확률과 우도함수를 실제로 다루기 쉬운 형태로 표현하기 위해서, 본 논문에서는 다음과 같이 가정한다.

(1) 두 개의 화소 s 와 k 가 서로 다른 값을 가질 때, 임의의 s 와 k 에서의 확률 변수 y_s 와 y_k 는 주어진 ω 상에서 서로 독립적이다. 이는 아래의 식으로 표현 된다^[9].

$$P(y_s, y_k \mid \omega) = P(y_s \mid \omega)P(y_k \mid \omega). \quad (9)$$

(2) 식 (9)의 분포함수들은 ω 상에서 s 의 값에 대한 함수이다^[9]. x_{ω} 가 λ 의 지시함수라면, 임의의 s 에 대해,

$$P(y_s | \omega) = P(y_s | x_{\omega}(s)) \quad (10)$$

다시 말해서, ω 상에서 한 점 s 의 의존성은 s 의 레이블의 의존성에 의해서만 영향을 받는다.

2. 사전확률

사전확률 $P(\omega)$ 는 아래의 Gibbs 분포를 갖는다.

$$P(\omega) = \exp(-U(\omega)) = \exp\left[-\sum_{c \in C} S_c(\omega)\right]. \quad (11)$$

여기서 C 는 모든 가능한 클릭들의 집합을 나타낸다. $U(\omega)$ 는 모든 클릭들에 대한 포텐셜 함수 $S_c(\omega)$ 의 합으로, 에너지 함수라 불린다. c 에서의 모든 화소들이 같은 레이블 값을 가지면 $S_c(\omega)$ 는 $-a$ 의 값을 가지고, 그렇지 않은 경우 a 의 값을 가진다.

ρ_s 를 화소 s 를 포함하는 모든 클릭들의 집합이라고 하자. 집합 C 는 모든 화소들에 대한 ρ_s 의 합과 같기 때문에, $U(\omega)$ 는 모든 화소들에 대한 지역 포텐셜 함수, $S_c(\omega_s)$ 들의 합으로 표현될 수 있다. 또한 입력 영상 Y 는 얼굴 영역 R 과 배경 영역 R^c 로 구성되어 있으므로, 에너지 함수는 다음과 같이 단순화 될 수 있다:

$$\begin{aligned} U(\omega) &= \sum_{s \in S} \sum_{c \in \rho_s} S_c(\omega_s) = \sum_{s \in R \cup R^c} \sum_{c \in \rho_s} S_c(\omega_s) \\ &= \sum_{s \in R} \sum_{c \in \rho_s} S_c(\omega_s) + \sum_{s \in R^c} \sum_{c \in \rho_s} S_c(\omega_s). \end{aligned}$$

따라서 $P(\omega)$ 는 아래의 식과 같이 MRF상의 확률변수의 지역 사전 확률들, $P(\omega_s)$ 의 곱으로 표현된다.

$$\begin{aligned} P(\omega) &= \exp\left[-\sum_{s \in R} \sum_{c \in \rho_s} S_c(\omega_s)\right] \\ &\quad \times \exp\left[-\sum_{s \in R^c} \sum_{c \in \rho_s} S_c(\omega_s)\right] \quad (12) \\ &= \prod_{s \in R} P(\omega_s) \times \prod_{s \in R^c} P(\omega_s). \end{aligned}$$

3. 우도 함수

우도 함수 $P(y|\omega)$ 는 가정 (1)에 의해 모든 픽셀 상의 지역 우도 함수 $P(y_s|\omega_s)$ 의 곱으로 표현된다. 그리고 지역 우도 함수는 가정 (2)에 의해 자신의 레이블 상태에 따라 결정되므로, $P(y|\omega)$ 는 아래의 식과 같이 표현된다.

$$\begin{aligned} P(y|\omega) &= \prod_{s \in S} P(y_s | \omega_s) \\ &= \prod_{s \in R} P(y_s | \omega_s = \lambda) \times \prod_{s \in R^c} P(y_s | \omega_s = \lambda^c). \end{aligned}$$

앞 장에서, 얼굴색은 2차원 가우시안 모델로 표현 될 수 있다는 것을 보였다. 따라서 지역 우도 함수 $P(y_s|\omega_s)$ 는 다음과 같은 2차원 가우시안 분포를 가진다.

$$P(y_s | \omega_s = \lambda) \propto N(m, \Sigma^2). \quad (13)$$

얼굴 영역과는 달리 배경영역의 색은 색 공간상에서 작은 영역에 밀집 되어 있지 않다. 오히려 복잡한 배경에서의 색은 전체 색 공간상에 고르게 분포되어진다. 이에 따라 배경 영역에 속하는 화소에서의 지역 우도 함수 $P(y_s|\omega_s = \lambda^c)$ 는 상수 값 V 에 비해 한다.

지역 우도 함수 $P(y_s|\omega_s)$ 는 다음과 같이 표현된다.

$$P(y_s | \omega_s) \propto \begin{cases} \left(2\pi\sigma_r\sigma_g\sqrt{1-\rho^2}\right)^{-1} \exp(-F_s) & , \omega_s = \lambda \\ V & , \omega_s = \lambda^c \end{cases} \quad (14)$$

여기서 F 는 아래와 같다:

$$F_s = \frac{1}{2(1-\rho^2)} \left[\frac{(r_s - \bar{r})^2}{\sigma_r^2} + \frac{(g_s - \bar{g})^2}{\sigma_g^2} - 2\rho \frac{(r_s - \bar{r})(g_s - \bar{g})}{\sigma_r\sigma_g} \right]$$

4. 사후 에너지 함수

임의의 랜덤 필드상의 사전 확률과 우도 함수는 각각 필드상의 지역 확률 함수들의 곱으로 표현될 수 있기 때문에, MAP은 다음의 식으로 표현 될 수 있다.

$$\begin{aligned} \omega^* &= \arg \max_{\omega} \prod_{s \in R} [P(y_s | \omega_s = \lambda) P(\omega_s)] \\ &\quad \times \prod_{s \in R^c} [P(y_s | \omega_s = \lambda^c) P(\omega_s)]. \end{aligned} \quad (15)$$

식 (15)에 로그 함수를 취하면 사후 확률 최대화 추정은 다음과 같이 다시 쓰여 질수 있다.

$$\begin{aligned} \omega^* &= \arg \max_{\omega} [\sum_{R} \log(P(y_s | \omega_s = \lambda) P(\omega_s))] \\ &\quad \times [\sum_{R^c} \log(P(y_s | \omega_s = \lambda^c) P(\omega_s))]. \end{aligned} \quad (16)$$

식 (16)에 앞에서 얻어진 사전 확률과 우도 함수를 대입하면 MAP는 아래 함수의 최소화 문제로 표현된다.

$$\omega^* = \arg \min_{\omega} E(\omega)$$

이 때 사후 에너지 함수, $E(\omega)$ 는 다음과 같이 정의된다.

$$\begin{aligned} E(\omega) &= \sum_{R} \left[F_s + \sum_{c \in C} S_c(\omega_s = \lambda) \right] \\ &\quad + \sum_{R^c} \left[V + \sum_{c \in C} S_c(\omega_s = \lambda^c) \right]. \end{aligned} \quad (17)$$

결과적으로 얼굴 경계선 추출 문제는 사후 에너지 함수의 최소화 문제로 표현 된다.

IV. 사후 에너지 함수의 최적화

본 논문의 목표는 정확한 얼굴 경계선을 찾는 것이다. $f(p): [0,1] \rightarrow \mathcal{R}^2$ 를 얼굴 영역 R 의 경계선을 나타내는 폐곡선이라고 하자. f 에 둘러싸인 영역 $R_{\vec{\gamma}}$ 은 얼굴 영역 R 을 나타내고 그 여집합 $R^c_{\vec{\gamma}}$ 는 배경 영역 R^c 를 나타낸다. 따라서 사후 에너지 함수는 아래와 같이 곡선 $E(\vec{\gamma})$ 의 함수로 표현될 수 있다.

$$E(\omega) = \sum_{R_{\vec{\gamma}}} [F_{\vec{\gamma}(p)} + \sum_{c \in C} S_c(\omega_{\vec{\gamma}(p)} = \lambda)] + \sum_{R^c_{\vec{\gamma}}} [V + \sum_{c \in C} S_c(\omega_{\vec{\gamma}(p)} = \lambda^c)] \quad (18)$$

사후 에너지 함수를 최소화하기 위해서, 본 논문에서는 $\vec{\gamma}$ 에 대한 steepest descent 방법을 사용한다. 이 방법에서 곡선 $\vec{\gamma}$ 상의 어떤 점 $\vec{\gamma}(p)$ 에 대한 모션 방정식은 다음과 같이 표현 된다:

$$\frac{d\vec{\gamma}(p)}{dt} = - \frac{\partial E(\vec{\gamma})}{\partial \vec{\gamma}(p)} \quad (19)$$

이 식의 오른쪽 항은 에너지의 일차 도함수 이다^[9, 10].

따라서 한 점 $\vec{\gamma}(p)$ 에 대한 모션 방정식은 다음과 같이 정의 된다:

$$\frac{d\vec{\gamma}(p)}{dt} = \underbrace{[-F_{\vec{\gamma}(p)} - 9]\vec{n}(\vec{\gamma}(p))}_{term1} + \underbrace{[-\sum_{c \in C} S_c(\omega_{\vec{\gamma}(p)} = \lambda) - \sum_{c \in C} S_c(\omega_{\vec{\gamma}(p)} = \lambda^c)]\vec{n}(\vec{\gamma}(p))}_{term2} \quad (20)$$

식 (20)에서 $\vec{n}(x)$ 은 점 x 에서 $R_{\vec{\gamma}}$ 의 바깥 방향으로 향하는 $\vec{\gamma}$ 의 단위법선 벡터이다.

본 논문에서 곡선은 함수 $\phi: \mathcal{R}^2 \rightarrow \mathcal{R}$ 의 제로 레벨 집합으로 함축적으로 표현된다. 또한 곡선의 안쪽 영역은 $\phi > 0$ 의 집합으로 표현 된다^[3, 4, 5]. 따라서 식 (20)은 아래의 레벨 집합의 진화 식으로 표현 된다^[3, 4, 5]:

$$\frac{d\phi(s)}{dt} = \underbrace{\{(-F_s) - V\}}_{term1} \|\nabla \phi\| + \underbrace{[-\sum_{c \in C} S_c(\lambda) + (-\sum_{c \in C} S_c(\lambda^c))]}_{term2} \|\nabla \phi\| \quad (21)$$

레벨 함수 $\phi(s)$ 를 초기화하기 위해서, 유클리디언 거리를 사용한다. 영상의 각 픽셀과 초기 곡선의 거리를 그 픽셀의 레벨 값으로 할당 한다. 레벨 집합의 진화 식은 제로 레벨 값을 가진 화소의 이웃들을 대상으로 간단한 forward difference를 사용하여 구현된다. n 이 반복 파라미터라면, 모션 방정식의 forward difference는 다음과 같이 정의된다.

$$\frac{\phi^{n+1}(s) - \phi^n(s)}{dt} = [-F_s - V] \|\nabla \phi\| + [-\sum_{c \in C} S_c(\lambda) - \sum_{c \in C} S_c(\lambda^c)] \|\nabla \phi\| \quad (22)$$

여기서 기울기의 놈(norm)은 소벨 연산자(Sobel operator)로 구현된다.

매번의 반복수행 후에, 진화된 경계선-제로 레벨 집합-은 재구성되고, 레이블 필드는 갱신된다. 레벨 값 제로가 교차하는 위치를 경계선으로 재구성하였다^[11]. 그 다음 화소 s 의 레벨 값이 양수라면 s 의 레이블은 λ 으로 그렇지 않다면 λ^c 으로 설정된다.

반복 수행 후 경계선 안쪽의 화소의 수 변화가 임계치를 넘으면 반복수행을 종료한다.

V. 실험결과

제안된 방법의 효율성을 입증하기 위해서 다양한 포즈와 각도의 얼굴을 포함하는 영상과 동영상에 대해 실험이 이루어졌다. 이 때 얻어진 실험결과는 geodesic 활성 윤곽선 모델의 결과와 비교 및 평가 되었다^[12].

Geodesic 활성 윤곽선 모델에서 윤곽선은 반복적으로 수축하다가 큰 기울기 값을 가진 객체의 경계선 근처에서 멈춘다. 이 모델은 곡선을 표현하고 진화시키기 위해서 레벨 셋을 사용하였기 때문에 복잡한 물체의 경계선이나 위상변화를 효율적으로 다룰 수 있다.

1. 평가함수

제안된 모델과 geodesic 활성 윤곽선 모델의 성능을 정량적으로 평가하기 위해서 Chamfer distance를 사용한다. 이는 두 윤곽간선간의 정합 평가 기준으로 많이 사용되어 왔다^[13, 14]. Chamfer distance를 구하기 위해서 입력 영상에서 수동으로 얼굴 영역의 정확한 경계선, 즉 ground truth를 추출한다. 이 후 추출된 ground truth 와 각 모델을 이용하여 얻어진 얼굴 경계선간의 Chamfer distance를 계산한다.

두 경계선간의 Chamfer distance는 다음과 같다.

$$C(F, G) = \frac{1}{3} \sqrt{\frac{1}{n} \sum v_i^2} \quad (23)$$

이 때 F 와 G 는 각각 제안된 방법과 사람의 손으로 추출한 얼굴 영역 경계선 픽셀들의 집합이다. 식 (23)에서 v_i 는 F 의 각 픽셀로부터 G 상의 가장 가까운 픽셀까지의 거리를 나타내고, n 은 곡선상의 화소의 수이다. 거리 값 v_i 는 [14]에서 설명되어져 있다. 즉, Chamfer distance는 두 경계선상에서 가장 가까운 점들 간의 거리의 평균을 의미한다.

2. 단일 얼굴 영상에서의 경계선 검출

그림 2는 제안된 방법과 geodesic 활성 윤곽선 모델을 이용하여 추출한 얼굴 영역들을 보여준다. 실험 영상은 TV 뉴스에서 얻어졌으며, 이 영상들은 정면을 응시하는 얼굴 영역 하나를 포함한다. 두 방법은 주어진 얼굴 영역 근처의 초기 곡선에서 시작하여 정확한 얼굴 영역의 경계선을 찾을 때 까지 곡선을 진화 시킨다. 그림 2(a)는 입력 영상 상의 초기 곡선을, 그림 2(b)와 (c)는 각각의 방법을 이용하여 추출된 얼굴 영역의 경계선들을 보여준다. 마지막으로 그림 2(d)는 두 방법의 Chamfer distance를 보여준다. 시각적으로 비교했을 때, 제안된 방법은 geodesic 모델보다 더 정확한 결과를 보여준다. 제안된 방법은 정확한 얼굴 영역의 경계선을 검출한 반면, geodesic 모델은 대략적인 경계선을 검출하였다. 뿐만 아니라, 그림 2(d)에서 보이듯이 제안된 모델의 Chamfer distance는 geodesic 윤곽선 모델보다 급격히 감소된다. 이러한 성능 차이는 geodesic 모델이 에지 검출기로 동작하기 때문에 얼굴 근처에 두드러진 에지를 포함한 영역들을 모두 추출하기 때문이다. 따라서 에지 정보만으로 정확한 얼굴을 추출하기가 어렵다.

그림 3은 얼굴 측면을 포함한 영상에 대한 실험 결과를 보여준다. 이 실험에서 초기 곡선은 그림 3(a)와 같이 주어졌다. 그림 3(b)와 (c)에는 각각 제안된 방법과 geodesic 모델에 의해서 추출된 얼굴 경계선을 보여준다. 그 때 두 방법의 Chamfer distance가 그림 3(d)에서 보인다. 제안된 방법은 피부의 색상 정보를 이용하여 윤곽선을 진화하기 때문에 얼굴의 방향에 관계없이 제안된 방법은 얼굴 경계선을 정확히 추출한다.

그림 2와 3은 제안된 모델이 얼굴의 관점에 영향을 받지 않는 얼굴 영역 검출 방법이라는 것을 보여준다.

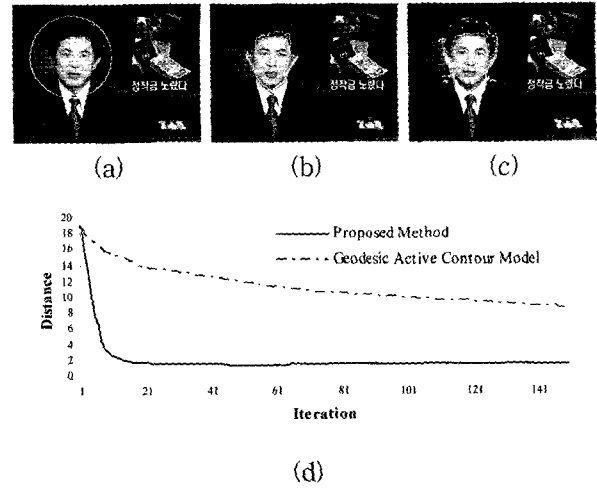


그림 2. 정면 얼굴 영역의 경계선 추출. (a) 초기 곡선; (b) 제안된 방법으로 추출한 얼굴 경계선; (c) geodesic 활성 윤곽선으로 추출한 얼굴 경계선; (d) 두 방법의 Chamfer 거리 비교

Fig. 2. Boundary detection of a front-view facial region. (a) initial curve; (b) facial boundary detected by proposed method; (c) facial boundary detected by geodesic active contours; (d) Chamfer distances in two methods.

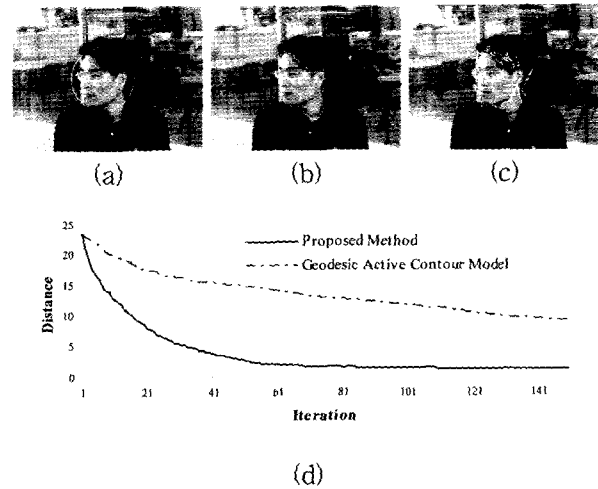


그림 3. 정면 얼굴 영역의 경계선 추출. (a) 초기 곡선; (b) 제안된 방법으로 추출한 얼굴 경계선; (c) geodesic 활성 윤곽선으로 추출한 얼굴 경계선; (d) 두 방법의 Chamfer 거리 비교

Fig. 3. Boundary detection of a side-view facial region. (a) initial curve; (b) facial boundary detected by proposed method; (c) facial boundary detected by geodesic active contours; (d) Chamfer distances in two methods.

3. 여러 개의 얼굴 영역 경계선 추출

유연한 물체의 경계를 추출하고 위상변화를 자동으로 다루기 위해서 제안된 모델은 레벨 셋을 이용하여 윤곽선을 표현하고 진화한다. 제안된 방법의 유효성을

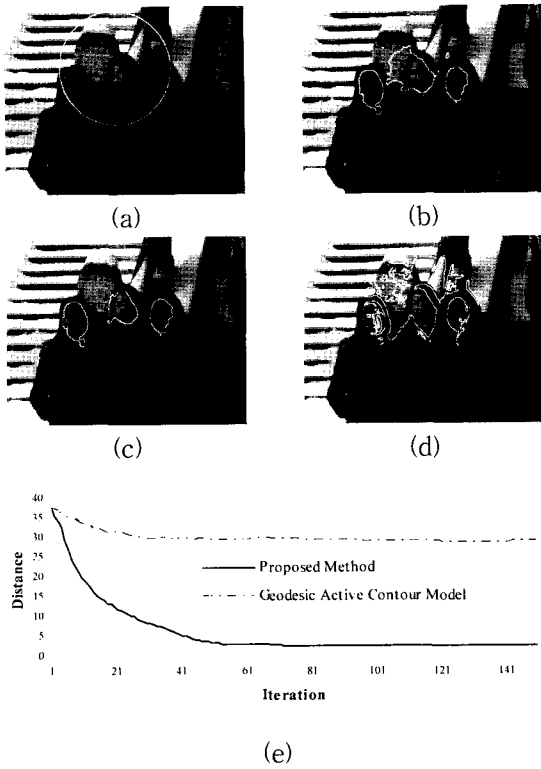


그림 4. 여러개의 얼굴을 포함한 영상에서 얼굴 경계선 추출. (a) 초기 곡선; (b) 진화하는 곡선; (c) 수렴한 곡선, 제안된 방법으로 추출한 얼굴 경계선; (d) geodesic 활성 윤곽선으로 추출한 얼굴 경계선; (e) 두 방법의 Chamfer 거리 비교

Fig. 4. Facial Boundary detection in an image including multiple faces. (a) initial curve; (b) evolving curve; (c) converged curve, facial boundary detected by geodesic active contours; (d) facial boundary detected by geodesic active contours (e) Chamfer distance in two methods .

보여주기 위해서 다양한 포즈를 가진 다수의 얼굴을 포함하는 영상에 대하여 실험을 하였다. 이러한 실험의 대표적인 결과가 그림 4에 주어졌다. 그림 4(a)는 초기 곡선을, (c)와 (d)는 각각 두 방법의 결과를 보여주고 있다. 그림 4(a) - (c)에서 보여주듯이, 제안된 방법에서 초기 곡선은 진화과정 중 점차 여러 개의 부분으로 분리되고(그림 4(b)), 각 분리된 곡선은 3개의 얼굴의 경계선으로 수렴된다. 반면, geodesic 모델은 정확한 얼굴 영역을 추출하지 못한다.

이 실험은 제안된 방법이 얼굴의 포즈나 수에 관계없이 정확한 얼굴 영역을 검출한다는 것은 보여준다.

본 논문에서 제안된 활성 윤곽선 모델의 주요한 장점 중 하나는 추적 메커니즘을 포함하고 있다는 것이다. 이를 증명하기 위해서 두 명 이상의 사람을 포함하는 동영상에서의 얼굴 영역 검출과 추적에 대한 실험이 이

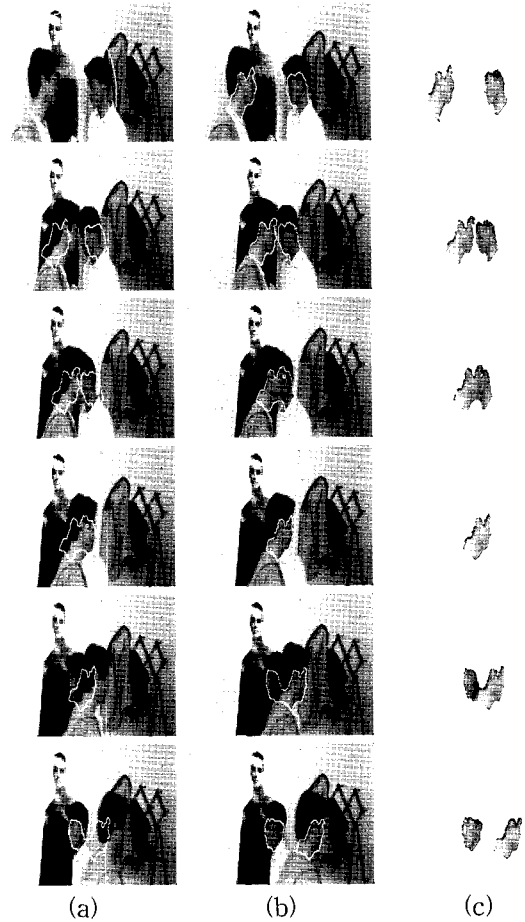


그림 5. 여러 개의 얼굴을 포함한 연속영상에서 얼굴 경계선의 추출 및 추적. (a) 초기 곡선; (b) 얼굴 경계선의 추출 및 추적; (c) 추출 및 추적된 영역.

Fig. 5. Facial boundary detection and tracking on the images including multiple facial regions. (a) initial curves; (b) results of facial boundary detection and tracking; (c) detected and tracked regions.

루어졌다. 전체 동영상에서 얼굴 영역들을 추적하기 위해 이전 프레임에서 추출된 얼굴 영역의 경계선을 그 다음 프레임의 초기 곡선으로 사용했다. 이후 그 곡선들은 얼굴 영역 추출에서와 같은 방법으로 진화한다. 그림 5는 두 개의 얼굴 영역을 포함하는 동영상에서의 얼굴 경계선 추출과 추적 결과를 보여준다. 제안된 모델에서 윤곽선은 양방향으로 진화하지만 geodesic 활성 윤곽선 모델에서 곡선은 한 방향으로만 진화한다. 따라서 geodesic 모델에서 초기 곡선의 전체는 추출하고자 하는 객체 경계선의 안쪽 또는 바깥쪽에 있어야만 하기 때문에, tracking을 위해서는 매 프레임마다 초기화 과정이 별도로 요구된다.

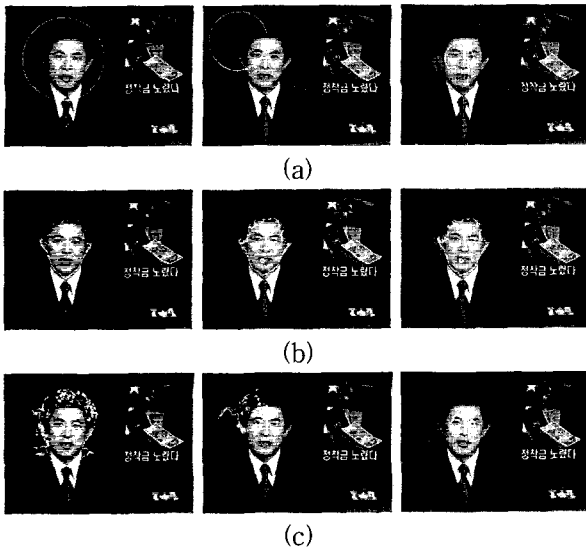


그림 6. 다양한 초기 곡선에서 얼굴 경계선 추출. (a) 다양한 초기 곡선; (b) 제안된 방법의 추출 결과; (c) Geodesic 활성화 윤곽선의 추출 결과
 Fig. 6. Facial boundary detection for various initial curves. (a) initial curves; (b) the results of the proposed method; (c) the results of the Geodesic AC.

표 2. 그림 6의 다양한 초기 곡선에서 Chamfer 거리
 Table 2. Chamfer distances on initial curves (Fig. 6).

	Fig. 6(a)	Fig. 6(b)	Fig. 6(c)
제안된 방법	1.37572	1.40122	1.41856
Geodesic AC	10.8484	13.3286	11.056

4. 다양한 환경에서의 성능 평가

활성 윤곽선 모델은 초기 곡선을 요구하는 방법이기 때문에 그 결과는 초기 곡선의 위치에 영향을 받는다. 일반적인 활성 윤곽선 모델은 한 방향으로 움직이면서 진화한다. 즉 폐곡선은 팽창하거나 또는 수축한다. 따라서 초기 곡선이 추출대상 객체의 내부 또는 외부에 위치하지 않을 경우 정확한 객체 추출을 보장하지 않는다. 제안된 방법에서 활성 윤곽선 모델은 양방향으로 움직이기 때문에 초기 곡선의 위치에 영향을 받지 않는다. 그림 6은 다양한 초기 곡선에서 대한 두 모델의 얼굴 영역 검출 결과를 보여준다. 실험 결과는 제안된 모델이 초기 값에 상관없이 항상 정확한 얼굴 영역을 추출함을 보여준다.

표 2는 제안된 방법과 geodesic 모델의 성능 비교를 보여준다. 제안된 방법에서는 각각의 초기 곡선에 대해 수렴된 곡선들의 Chamfer distance 간에는 큰 차이가 없다. 하지만 geodesic 활성화 윤곽선 모델의 경우 초기 곡선에 많은 영향을 받는다. 이 실험 결과는 제안된 방

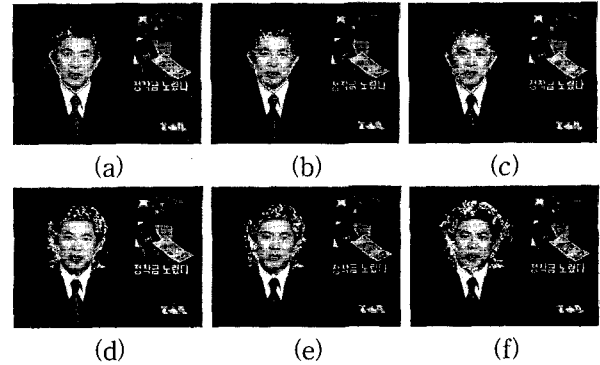


그림 7. 잡음 영상에서 얼굴 경계선 추출. (a)-(c) 5, 10 그리고 15% 가우시안 잡음이 첨가된 영상에서 제안된 방법의 결과, (d)-(f) 5, 10 그리고 15% 가우시안 잡음이 첨가된 영상에서 Geodesic 활성화 윤곽선의 결과.

Fig. 7. Facial boundary detections in noise images. (a)-(c) the results of the proposed method in images corrupted by 5, 10 and 15 % Gaussian noise, respectively, (d)-(f) the results of the Geodesic AC in images corrupted by 5, 10 and 15 % Gaussian noise, respectively.

표 3. 잡음 영상에서의 Chamfer 거리
 Table 3. Chamfer distances on noise images.

Gaussian Noise(%)	5%	10%	15%
제안된 방법	1.45503	1.90303	2.31946
Geodesic AC	7.97117	8.08237	12.296

법이 초기 곡선 값에 영향을 받지 않는 강건한 방법이라는 것을 증명한다.

제안된 방법에서는 인접한 화소간의 공간적인 상호관계를 표현하는데 효율적이고 잡음에 강한 MRF를 이용하여 영상을 모델링하였다. 잡음에의 강건함을 보여주기 위해서, 잡음 민감성 테스트를 수행 하였다. 그림 7은 잡음이 있는 영상에 대한 제안된 방법과 geodesic 활성화 윤곽선 모델의 실험 결과를 보여준다.

원래의 영상에 서로 다른 비율의 Gaussian noise가 첨가 되었다. Geodesic 모델의 결과와 시각적으로 비교했을 때 제안된 방법은 보다 정확한 얼굴 경계선을 제공한다. 표 3은 각각의 결과에 대한 Chamfer distance를 보여준다. 잡음이 증가할수록 제안된 방법에서의 Chamfer distance는 증가하였다. 하지만, geodesic 모델의 결과에 비하여 증가율이 현격히 낮다. Geodesic 모델이 9.45의 평균을 가지는 것에 비해 제안된 방법은 1.89의 평균을 보였다. 이 결과는 기존의 방법에 비해 제안된 방법이 잡음에 강하다는 것을 보여준다.

VI. 결 론

본 논문에서는 포즈, 방향 그리고 잡음에 관계없이 정확한 얼굴 경계선을 추출할 수 있는 활성 윤곽선 모델을 개발하였다. 잡음으로부터 강건하기 위해 MRF를 사용하여 영상을 모델링 하였으며, 최적화 척도로 MAP를 사용하였다. 따라서 얼굴 경계 추출은 에너지 최소화 문제로 정형화 되었다. 제안된 모델에서 곡선은 레벨 함수의 제로 레벨 집합으로 표현되고, 레벨 함수에 대한 편미분 방정식을 통해 진화 된다. 이 때, 곡선 진화와 정지를 위해서 2차원 가우시안 모델로 표현되는 피부색 정보를 이용하였다. 실험 결과는 제안된 방법이 잡음 및 얼굴의 포즈에 강건하고 초기 곡선에 영향을 받지 않는 방법임을 증명하였다. 비록 초기 값에 영향을 받지 않는 않지만, 제안된 방법은 초기 곡선을 입력으로 받는 반자동 방법이다. 따라서 초기 곡선을 자동으로 결정하는 방법에 대한 연구가 현재 진행 중에 있다.

참 고 문 헌

- [1] M. H. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 24, no 1, pp. 34-58, 2002.
- [2] M. Kass et al., "Snakes: Active contour models", *International Journal of Computer Vision*, Vol 1, pp 321-331, 1988.
- [3] V. Caselles et al., "A geometric model for active contours in image processes," *Numerische Mathematik*, Vol 66, pp 1-31, 1993.
- [4] T. F. Chan and A. Vese, "Active Contours Without Edges," *IEEE Transactions on Image Processing*, Vol 10, no. 2, pp. 266-277, 2001.
- [5] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed", *Journal of Computational Physics*, Vol. 79, pp. 12-49, 1998.
- [6] J. Yang and A. Waibel, "A real-time face tracker," in *Proc. of the Third IEEE Workshop on Applications of Computer Vision*, pp. 142-147, Sarasota, Florida, 1996.
- [7] E. Y. Kim S. W. Hwang, S. H. Park and H. J. Kim, "Spatiotemporal Segmentation Using Genetic Algorithms," *Pattern Recognition*, Vol. 34, no. 10, pp. 2063-2066, 2001.
- [8] E. Y. Kim, *Video segmentation using genetic algorithms*, PhD thesis, Kyungpook National University, DaeGu, Korea, 2001.
- [9] A. R. Mansouri, "Region Tracking via Level Set PDEs without Motion Computation," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 24, no. 7, pp. 947-961, 2002.
- [10] S. C. Zhu and Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Transactions on PAMI*, Vol 18, no. 9, pp. 884-900, 1996.
- [11] S. Kulkarni and B. N. Chatterji, "Accurate shape modeling with front propagation using adaptive level sets," *Pattern Recognition Letters*, Vol. 23, no. 13, pp. 1559-1568, 2002.
- [12] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours," In *Fifth International Conference on Computer Vision*, Boston, MA, 1995.
- [13] S. D. Fenster and J. R. Kender, "Sectorized Snakes: Evaluating Learned Energy Segmentations," *IEEE Transactions on PAMI*, Vol. 23, no. 9, pp. 1028-1034, 2002.
- [14] G. Borgefors, "Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm," *IEEE Transactions on PAMI*, Vol 10, no. 11, pp. 849-865, 1998.

저 자 소 개



장 재 식(학생회원)
2003년 경북대학교 컴퓨터공학과 석사 졸업
2004년~현재 경북대학교 컴퓨터공학과 박사과정.
<주관심분야: 컴퓨터 비전, 패턴 인식, 제스처 인식, 물체 추적>

김 은 이(정회원)
논문 2004-41C1-4-2 참조

김 향 준(정회원)
논문 2004-41CI-4-7 참조

