

비전 기반의 모바일 로봇 제어 시스템

장재식* 김은이** 장상수* 김향준*

* 경북대학교 컴퓨터공학과

** 건국대학교 인터넷 미디어 공학부

Vision-based Mobile Robot Control System

Jae Sik Chang* Eun Yi Kim** Sang Su Jang* Hang Joon Kim*

* Dept. of Computer Engineering, Kyungpook National University

** School of Internet and Multimedia, Konkuk University

요약

본 논문은 손 모양 인식을 이용한 비전 기반의 보행 로봇 제어 시스템을 제안한다. 손의 모양을 인식하기 위해서 움직이는 카메라 영상으로부터 정확한 손의 경계선을 추출하고 추적하는 일이 선행 되어야 한다. 따라서 본 논문에서는 민 시프트 방법을 사용한 활성 윤곽선 모델 기반의 추적 방법을 제안한다. 제안된 시스템은 손 추출기, 손 추적기, 손 모양 인식기 그리고 로봇 제어기, 4개의 모듈로 구성 된다. 손 추출기는 영상에서 미리 정의된 손의 모양을 가지는 피부색 영역을 추출한다. 추출된 손의 추적은 활성 윤곽선 모델과 민 시프트 방법을 사용하여 실행된다. 그 후 Hue moments를 사용하여 추적된 손의 모양을 인식한다. 제안된 방법을 평가하기 위해서 본 논문에서는 2족 보행 로봇 KHR-1에 제안된 방법을 적용 한다.

1. 서론

본 논문에서는 손의 모양 인식을 통한 비전 기반의 모바일 로봇 제어 시스템을 제안 한다. 손의 모양을 인식하기 위해서 움직이는 카메라 영상으로부터 정확한 손의 경계선을 추출하고 추적하는 일이 선행 되어야 한다. 최근 활성 윤곽선 모델은 곡선의 효율적인 표현능력과 유연한 특징 때문에 객체의 추적에 많이 사용되고 있다 [1]. 하지만 이전 프레임에서의 객체 경계선을 초기 커브로 사용할 경우, 활성 윤곽선 모델의 객체 추적 결과는 초기 커브의 위치에 매우 민감하다. 움직이는 카메라 영상의 경우, 손이 움직이고 있지 않더라도 영상에서 손의 위치는 변화한다. 이런 손의 위치 변화는 손의 추적 과정에서 손과 거리가 먼 초기 커브를 초래한다. 이런 부정확한 초기 커브는 손의 경계선으로 수렴하기 위해서 보다 많은 계산 량이 필요하고 손의 경계선과 유사한 특징을 가진 잡음과 출몰을 생성한다. 더욱이 움직임이 큰 손의 경우 추적 중에 잃어버리는 경우도 발생한다.

따라서 본 논문에서는, 이러한 활성 윤곽선 모델의 초기 커브에 대한 민감성을 최소화하기 위해서, 민 시프트 방법을 사용한 활성 윤곽선 모델 기반의 추적 방법을 제안한다. 제안된 시스템은 손 추출기, 손 추적기, 손 모양 인식기 그리고 로봇 제어기, 4 개의 모듈로 구성 된다.

손 추출기는 영상에서 미리 정의된 손의 모양을 가지는 피부색 영역을 추출한다. 추출된 손의 추적은 활성 윤곽선 모델과 민 시프트 방법을 사용하여 실행된다. 그 후 Hue moments를 사용하여 추적된 손의 모양을 인식한다. 제안된 방법을 평가하기 위해서 본 논문에서는 2족 보행 로봇 KHR-1에 제안된 방법을 적용 한다.

2. 시스템 개요

그림 1은 제안된 모바일 로봇 제어 시스템을 보여주고 있다. 사용자는 데이터 글로버와 같은 별도의 장비 없이 카메라 앞에서 미리 정의된 손 모양으로 로봇을 제어 할 수 있다.

제안된 시스템에서 모바일 로봇은 사용자의 오른손 모양에 의

해서 제어된다. 그림 2는 미리 정의된 가지 손의 모양과 그에 대응되는 명령들을 보여준다.

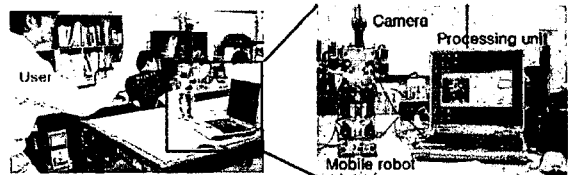


그림 1. 시스템 환경

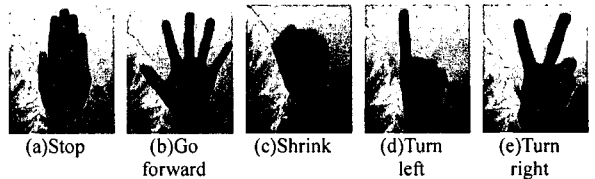


그림 2. 미리 정의된 손의 모양과 명령들

제안된 시스템에서 사용자가 로봇 제어를 시작 하려면 그림 2의 (a) 모양의 명령을 로봇에게 내려야 한다. 로봇의 손 추출기가 입력되는 영상열의 매 프레임에서 그림 2의 (a) 모양의 피부색 영역을 찾고, 그 영역이 찾지면 그 영역을 손이라고 간주하게 된다. 그리고 손 추적기가 손 영역을 추적하기 시작한다. 그리고 매 프레임에서 추출되거나 추적된 손 영역의 모양을 인식하고 연속적으로 3 프레임 이상 같은 모양이 인식이 되면 로봇 컨트롤러에게 그 모양에 해당하는 명령을 보내게 된다. "stop" 명령은 로봇의 동작을 멈추고 선 자세를 유지하게 하는 명령이다. 그리고 나머지 명령들은 "stop" 명령이 입력될 때까지 대응되는 동작을 실행하게 된다.

3. 모바일 로봇 제어 시스템

제안된 시스템은 손 추출기, 손 추적기, 손 모양 인식기 그

리고 로봇 제어기 네개 의 모듈로 구성된다. 본 논문에서는 로
봇 제어기로 상용 프로그램인 HeartToHeart1.0 을 사용하였다.
따라서 이 장에서는 앞의 3 모듈에 대해 설명한다.

3.1 손 추출기

제안된 시스템에서는 손 추출을 위해서 색채 칼라 공간에서
2D 가우시안 분포로 표현된 피부색 정보를 사용 한다. RGB 공
간에서 칼라 표현(color representation)은 색상 뿐 아니라 밝기
(brightness)도 포함한다. 따라서 RGB 는 피부색으로 픽셀들을
검출하는 것에는 적합하지 않다. 밝기는 칼라 픽셀의 세 구성
요소 (R, G, B)를 intensity 로 나눔으로써 제거될 수 있다. 이 공
간은 크로매틱(chromatic) 색상 공간으로 알려져 있고, 이 공간
의 정규화(normalize)된 색 벡터(color vector)인 두 개의 요소 (r,
g)로 구성된다.

피부색 모델은 1000 개의 표본 영상으로부터 구하였다. 피부색
모델은 아래와 같은 평균과 편차를 가지는 2 차원 가우시안 모
델로 근사화 될 수 있다.

$$\mu = (\bar{r}, \bar{g}) = (117.588, 79.064),$$

$$\Sigma = \begin{bmatrix} \sigma_r^2 & \rho_{X,Y}\sigma_r\sigma_g \\ \rho_{X,Y}\sigma_r\sigma_g & \sigma_g^2 \end{bmatrix} = \begin{bmatrix} 24.132 & -10.085 \\ -10.085 & 8.748 \end{bmatrix}$$

제안된 방법은 피부색 모델을 사용하여 영상에서 피부색 화소
들을 추출하고 모폴로지 연산을 통하여 잡음을 제거한다. 추출
된 화소를 중 일정 크기의 연결 성분들을 후보 영역으로 생성
하고 이 후보 영역들의 모양을 인식하여 그 모양이 "stop"인
영역을 손으로 추출한다. 이때 손 모양을 인식하기 위해서는
손 모양 인식기에서 사용하는 Hue 의 불변 모멘트를 사용한다.

3.2 손 추적기

제안된 시스템에서는 손이 추출된 후 계산량을 줄이고 연속된
두 프레임 간에 손의 대응관계를 보장하기 위해서 매 프레임
손을 추출 하는 것 대신에 추출된 손을 추적한다. 손을 추적하
기 위해서 본 연수에서는 활성 윤곽선 모델을 사용한다. 최근
활성 윤곽선 모델은 곡선의 효율적인 표현능력과 유연한 특징
때문에 객체의 추적이 많이 사용되고 있다. 하지만 그 결과와
초기 커브의 위치에 매우 민감한 단점을 가지고 있다. 찾고자
하는 객체와 멀리 떨어진 초기 커브는 객체의 경계선으로 수렴
하기 위해서 보다 많은 계산량이 필요하고 객체의 경계선과 유
사한 특징을 가진 잡음과 출들을 생성한다. 더욱이 움직임이
큰 객체의 경우 추적 중에 잃어버리는 경우도 발생한다. 따라
서 본 연수에서는 초기 커브에 대한 민감성을 줄이기 위해 민
시프트 방법을 이용한 활성 윤곽선 모델 기반의 손 추적 방법
을 제안한다. 손 추적은 초기 커브 재위치 단계 와 커브 진화
단계 두 단계로 실행된다.

커브 재위치 단계 : 이 단계에서는 이전 프레임의 손의 경계
선으로 초기화된 초기 커브를 민 시프트 알고리즘을 이용하여
손 가까이 이동시킨다.

민 시프트 알고리즘은 손의 확률 분포
{P(I_{ij}|α_o)}_{i,j=1}, IW, IH(IW: image width, IH: image height)를 사용하
여 반복적으로 탐색창의 위치를 바꾼다. 손의 확률 분포는 영
상에서 한 화소(i,j)가 손의 일부인 확률을 나타낸다. 이때 α_o
는 확률 분포의 파라미터들이고 I 는 칼라 값이다. 본 연수에서
는 손의 확률 분포로 피부색 모델을 사용하였다. 탐색창의 새
로운 위치는 다음과 같이 간단히 계산된다. [2]:

$$m_x = M_{10}/M_{00} \quad \text{and} \quad m_y = M_{01}/M_{00} \quad (1)$$

, 이때 M_{ab} 는 (a + b)차 모멘트이고 다음과 같이 정의된다.

$$M_{ab}(W) = \sum_{i,j \in W} i^a j^b P(I_{ij} | \alpha_o).$$

손의 위치는 탐색창의 위치를 계산함으로써 연속적으로 구할
수 있다. 탐색창의 위치는 초기 커브의 바운딩 박스 B 로 초기
화 된다. 그리고 탐색창의 크기는 아래와 같이 각 프레임에서
손의 움직임 양에 비례한다.

$$W_{width(height)} = \max(\alpha(m_y^{(t-1)} - m_y^{(t-2)}) - B_{width(height)}, 0) + \beta B_{width(height)}$$

(t > 2), 이때 α 와 β 는 상수이고 t 는 프레임 인덱스이다.

커브 진화 단계 : 초기커브가 재 위치된 후, 커브는 정확한 손
의 경계선과 일치되도록 영역 경쟁 기반의 활성 윤곽선 모델을
이용하여 반복적으로 진화한다. 영역 경쟁은 영상 분할 방법으
로 Zhu 와 Yuille 에 의해서 제안되었다 [3]:

$$E[\Gamma, \{\alpha_i\}] = \sum_{i=1}^M \left\{ \frac{\mu}{2} \int_{R_i} ds - \log P(\{I_s : s \in R_i\} | \alpha_i) + \lambda \right\} \quad (2)$$

, 이때 Γ 은 영상에서 경계선이고, P(•)는 영역 R_i 의 특징 분
포이다. α_i 는 분포들의 파라미터들이고, M 은 영역들의 수이다.
s 는 영상 좌표계에서 하나의 위치이고 μ 와 λ 는 상수들이다.

에너지 E 를 최소화하기 위해서, 각 경계선 Γ 에 대해 steepest
descent 방법을 사용한다. 경계선 상의 어떤 점 에 대해서 아래
와 같은 결과를 얻을 수 있다:

$$\frac{dV}{dt} = - \frac{\delta E[\Gamma, \{\alpha_i\}]}{\delta \Gamma} \quad (3)$$

, 이때 오른쪽 항은 에너지 함수의 마이너스 미분함수이다.
함수를 미분함으로써 다음의 각 포인트에 대한 모션 방정식을
구할 수 있다:

$$\frac{dV}{dt} = \sum_{k \in Q_{\beta_i}} \left\{ -\frac{\mu}{2} k_{k(\beta)} \rho_{k(\beta)} + \log P(I_{\beta_i} | \alpha_k) \rho_{k(\beta)} \right\} \quad (4)$$

, 이때 Q_{β_i} = {k | V lies on Γ_k} 이다. 는 Γ_k 의 한 점 에서의 곡률
이고 는 같은점 에서의 단위법선 벡터이다.

활성 윤곽선은 아래의 식과같이 영상을 공동의 경계선을 가
지는 두 영역으로 분류하는 영역 경쟁의 특수한 예이다 :

$$\frac{dV}{dt} = -\mu k_{o(\beta)} \rho_{o(\beta)} + (\log P(I_{\beta_i} | \alpha_o) - \log P(I_{\beta_i} | \alpha_k)) \rho_{o(\beta)} \quad (5)$$

활성 윤곽선의 진화는 레벨 셋 방법을 사용하여 구현된다. 커
브 Γ 는 함수 u : R² → R₊ 의 재로 레벨 집합으로 함축적으로 표
현한다. 이때, 커브 Γ 의 내부는 u > 0 인 영역이다. 따라서 식
(7)은 아래의 레벨 셋 함수로 표현될 수 있다[8]:

$$\frac{du(s)}{dt} = -\mu k_s \|\nabla u\| + (\log P(I_s | \alpha_o) - \log P(I_s | \alpha_k)) \|\nabla u\| \quad (6)$$

3.3 손 모양 인식기

본 논문에서는 손의 모양을 표현하기 위해서 Hu 의 불변 모멘
트를 사용하였다.

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02}, \quad \phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} + \eta_{21})^2, \quad \phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\ \phi_5 &= (3\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{12}) - 3(\eta_{21} + \eta_{03})^2 \\ &\quad + (3\eta_{21} - \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) \\ \phi_6 &= (\eta_{30} - \eta_{02})(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \\ &\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) \end{aligned} \quad (7)$$

불변 모멘트는 천이, 회전, 크기 변화에 불변한 성질을 가지기 때문에 제안된 시스템에서 손의 모양을 표현하기에 적당하다 [4]. 식(7)은 불변 모멘트를 보여주고 있다.

손의 모양을 인식하기 위해서 현재 프레임에서 손과 템플릿의 7 개 불변모멘트들간의 유클리디언 거리를 사용하였다.

4. 실험 결과

제안된 제어 시스템은 MS Visual C++ 6.0 과 OpenCV beta3.1 으로 구현되었고 모바일 로봇에 부착된 카메라로부터 초당 3 프레임의 320×240 24-비트 칼라 영상을 캡처 하였다. 사용된 로봇은 17 개의 서보 모터로 구성된 KHR-1 이다.

제안된 시스템은 사용자 4 명을 대상으로 총 529 프레임에 대해 평가하였다. 제안된 손 추적방법의 효율성을 보여주기 위해서 민시프트를 사용한 제안된 방법과 민시프트 없이 활성윤곽선모델만을 사용한 방법을 비교하였다. 그림 3 은 제안된 방법의 손 추적 결과를 보여준다.

그림 4 는 수렴속도를 이용한 두 방법의 정량적인 비교를 보여주고 있다. 제안된 방법에서는 초기 곡선이 민시프트를 이용하여 손 근처로 재위치 되어지기 때문에 곡선진화의 반복이 빨리 수렴한다. 따라서 제안된 방법이 활성윤곽선만 사용한 방법에 비해서 빠른 수렴 시간필요로 한다.

표 1 은 불변 모멘트를 사용한 손 모양 인식 결과를 보여주고 있다.

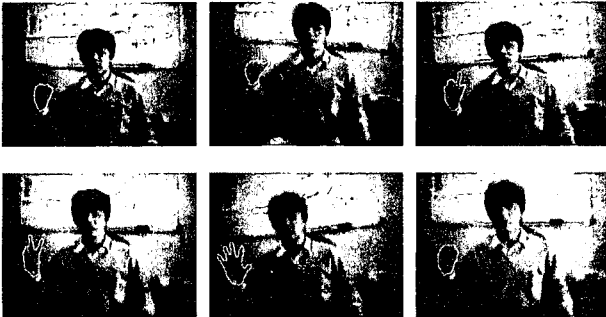


그림 3. 손 모양 추적 결과

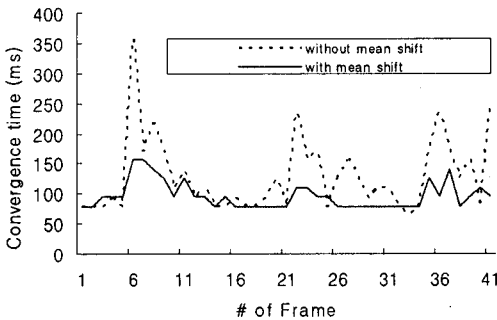


그림.4 두 방법간의 수렴 속도 비교

5 결론

본 논문은 손 모양 인식을 이용한 비전 기반의 보행로봇 제어 시스템을 제안한다. 제안된 시스템은 손 추출기, 손 추적기, 손 모양 인식기 그리고 로봇 제어기, 4 개의 모듈로 구성된다. 손 추출기는 영상에서 미리 정의된 손의 모양을 가지는 피부색 영역을 추출한다. 추출된 손의 추적은 활성 윤곽선 모델과 민 시프트 방법을 사용하여 실행된다. 그 후 Hue moments 를 사용하여 추적된 손의 모양을 인식한다. 제안된 방법을 평가하기 위해서 본 논문에서는 2족 보행 로봇 KHR-1 에 제안된 방법을 적용 한다. 제안된 시스템의 효율성을 보여 주기 위해서, 2족 보행로봇 KHR-1 을 대상으로 실행하였다.

표 1. 손 모양 인식 결과

		Shape recognized				
		Stop	Go forward	Shrink	Turn left	Turn right
Shape given	529					
Stop	107	101	2	4		
Go forward	103	2	101			
Shrink	112	2		109	1	
Turn left	106			1	103	2
Turn right	101		2		1	98

참고문헌

1. Freedman, D., and Zhang, T.: Active Contours for Tracking Distributions. IEEE Transactions on Image Processing. Vol. 13, No. 4 (2004) 518-526
2. Kim, K. I., Jung, K., and Kim, J. H.: Texture-Based Approach for Text Detection in Image Using Support Vector Machines and Continuously Adaptive Mean Shift Algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 25, No. 12 (2003) 1631-1639
3. Zhu, S. C., and Yuille, A.: Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 18, No 9 (1996) 884-900
4. Gonzalez, R. C., and Woods R. R., *Digital Image Processing*, Prentice Hall, New Jersey, 2002