

얼굴 특징 추적을 이용한 인터페이스 구현

신윤희[○] 강신국^{**} 김은이^{*}

건국대학교 인터넷미디어공학과 시각정보처리 연구실^{*}

서울대학교 컴퓨터공학과^{**}

(ninharsa, eykim) @konkuk.ac.kr, skkang@dglab.suu.ac.kr

Interface Implementation using Facial Feature Tracking

Yunhee Shin[○], Sin Kuk Kang^{**}, Eun Yi Kim^{*}

^{*}Visual Information Processing Lab, Dept. of Internet and Multimedia Eng., Konkuk Univ., Seoul, Korea

^{**}Dept. of Computer Eng., Seoul National Univ., Seoul, Korea

요약

본 논문은 얼굴 특징 추적을 이용한 새로운 인터페이스를 제안한다. 눈의 움직임만으로 구현된 기존의 시스템은 마우스 클릭 이벤트에 걸리는 waiting time으로 인해 속도 개선이 필요했다. 이를 위해서 본 논문에서는 눈의 움직임 뿐 아니라 입의 움직임도 인식하여 사용자의 요구를 처리할 수 있는 시스템을 개발한다. 제안된 시스템은 얼굴 검출 모듈, 눈 검출 모듈, 입 검출 모듈, 얼굴 특징 추적 모듈, 마우스 제어 모듈의 5 가지 모듈로 구성되어 있다. 먼저, 피부색 모델과 연결 성분 분석을 이용하여 얼굴을 검출하고 신경망 기반의 분류기와 에지 검출기를 이용하여 검출된 얼굴 영역에서 눈과 입을 찾는다. 이후 프레임에서는 mean-shift 알고리즘과 템플릿 매칭을 이용하여 눈과 입이 정확하게 추적되어 눈의 움직임으로 마우스의 포인터를 움직이고 입의 움직임으로 메뉴나 아이콘을 클릭하게 된다.

제안된 시스템의 효율성을 검증하기 위해서 웹 브라우저의 인터페이스로 활용하였다. 25명의 사용자에게 대해 실험한 결과는 제안된 시스템이 보다 편리하고 친숙한 인터페이스로 활용될 수 있다는 것을 보여주었다.

1. 서론

컴퓨터 및 PDA와 같은 장치와 게임과 같은 응용 프로그램에서 입력수단 또는 제어 장치로 사용되는 얼굴 특징의 움직임을 이용한 인터페이스가 많은 관심을 받고 있으며 이와 관련된 시스템들이 많이 개발되고 있다[1]. 이러한 시스템들이 실질적으로 사용되기 위해서는 가능한 적은 시스템 자원을 이용하여 실시간으로 얼굴 특징을 찾아내고 추적할 수 있어야 한다.

본 논문에서는 눈과 입의 검출과 추적을 이용하여 PC 기반의 Human-Computer Interface(HCI) 시스템을 개발하였다. 전체적인 시스템의 구성은 그림 1에서 보인다. 제안된 시스템은 카메라를 이용해 컴퓨터 앞에 앉은 사용자의 실시간 영상을 받아 화면에 출력하면 사용자는 눈을 움직임으로써 마우스 포인터를 움직이고 입의 움직임으로 클릭을 가능하게 한다. 이를 위해 제안된 시스템은 얼굴 검출 모듈, 눈 검출 모듈, 입 검출 모듈, 얼굴 특징 추적 모듈, 마우스 제어 모듈로 구성된다. 즉, 피부색 정보를 이용하여 검출된 얼굴 영역에서 신경망 기반의 텍스처 분류기를 사용해 눈을 찾고 검출된 눈의 위치에 기반 하여 에지 검출기를 이용해 입을 찾는다. 이후 눈과 입을 정확하게 추적하기 위해 mean-shift 알고리즘과 템플릿 매칭을 이용하여 눈과 입을 추적한다. 추적된 결과에 기반 하여 마우스의 이동이나 클릭과 같은 작동이 구현 된다.

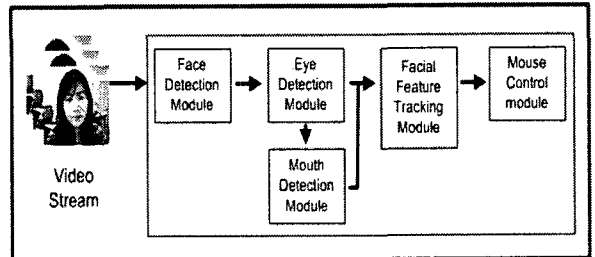


그림 1. 제안된 인터페이스의 구성

제안된 시스템의 효율성을 검증하기 위해서 웹 브라우저의 인터페이스로 활용하였다. 25명의 사용자에게 대해 제안된 시스템을 테스트한 결과는 제안된 시스템이 시간에 따라 변화는 조명변화에 강건하고 안경에 의해 반사되는 빛에 덜 민감하다는 것을 보여주었다. 실험 결과는 또한 제안된 시스템이 사람의 의사소통 수단으로 사용되는데 편리하고 친숙한 인터페이스로 활용될 수 있음을 보여주었다.

2. 얼굴 검출

얼굴 영역을 추출할 때 가장 일반적인 방법은 피부색 모델을 이용하여 얼굴 영역을 추출하는 것이다. 본 논문

에서는 chromatic color space에서 2차원 가우시안 모델로 표현되는 피부색 모델을 사용하여 입력 영상을 얼굴 영역과 비 얼굴 영역으로 나눈다[2].

피부색 모델로 검출된 얼굴 영역은 그림 2에서 보인다. 그림 2(a)는 원 영상이고 그림 2(b)는 추출된 얼굴 영역을 나타낸다. 그림에서 보이듯이, 이진영상에서 연결 성분 분석을 수행하여 작은 영역들은 모두 제거한 후, 영역 중에서 가장 큰 영역이 얼굴 영역으로 간주한다.



그림 2. 얼굴 영역 검출 예: (a) 원 영상, (b) 피부색모델로 검출된 얼굴 영역

3. 눈 검출

제안된 시스템은 다양한 환경에서 얼굴 영역을 눈 영역과 비눈 영역으로 자동적으로 구별하기 위해서 다층 퍼셉트론(multilayer perceptron, MLP)을 이용한 텍스처 분류기를 사용한다. 제안된 신경망은 입력층, 은닉층, 출력층으로 구성되어 있으며 각 계층에서의 노드들은 완전 연결되어 있다. 그림 3은 제안된 신경망 기반의 눈 검출 시스템의 개요를 보여준다.

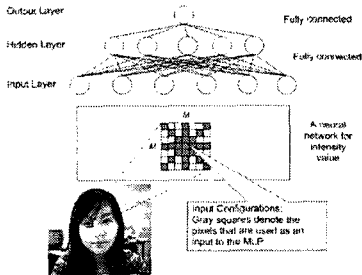


그림 3. 제안된 신경망 기반의 눈 검출 시스템

제안된 신경망은 입력 값으로 gray-level값을 사용하며 입력 벡터의 크기를 줄이기 위해 autoregressive 방법을 사용한다[3]. 이로 인해 결과적으로 신경망 분류기의 성능과 분류속도는 향상된다. 학습 후에 신경망의 출력은 0과 1사이의 실수 값으로 표현되며 만약 신경망의 출력이 임계치 값보다 크면 눈 영역으로 분류하고 그렇지 않으면 비눈 영역으로 분류한다.

그림 4는 신경망 분류기의 눈 검출 예를 보여준다. 그림 4(a)에서 파란색 영역이 얼굴 영역에서 검출된 눈 영역에 대응되는 화소들이다. 신경망 분류기로 검출된 결과는 많은 오 추측(false alarms)을 포함한다. 이를 제거하기 위해서 몇 가지 휴리스틱 정보를 이용하여 잘못된 분류된 영역들을 제거한다. 그림 4(b)에서 후처리 후의 최종적으로 검출된 눈 영역은 녹색으로 표시되어 있다.



그림 4. 눈 검출 예: (a) 신경망으로 분류된 결과(파란색 영역), (b) 후처리로 검출된 눈 영역(녹색 영역)

4. 입 검출

제안된 시스템은 매 프레임마다 에지 검출기를 이용하여 얼굴 영역에서 입 영역을 검출한다. 제안된 에지 검출기는 입력 영상의 gray-level값을 이용하여 얼굴 윤곽을 구하며 처리되는 시간을 줄이기 위해 본 논문에서는 전체 얼굴 영역이 아닌 검출된 눈의 위치에 기반을 둔 몇 가지 휴리스틱 정보를 이용하여 입의 범위를 구한다. 정해진 범위에서 에지 검출기를 이용하여 나타난 결과는 연결 성분 분석을 통해 작은 영역들은 제거되고 최종적으로 입 영역이 검출된다.

그림 5는 에지 검출기의 의한 입 검출 예를 보여준다. 그림 5(a)와 그림 5(b)는 에지 검출기에 의해 검출된 얼굴 윤곽 영역이며 그림 5(b)와 그림 5(d)는 각각 그림 5(a)와 그림 5(c)의 후처리 결과로 검출된 입 영역이다.

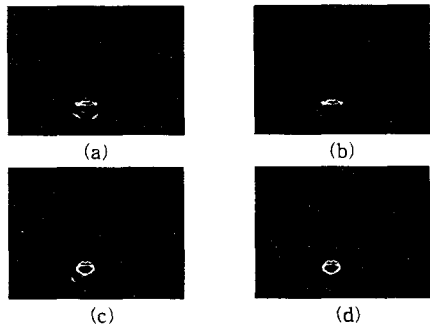


그림 5. 입 검출 예: (a) 에지 검출기로 검출된 윤곽(입을 다물었을 때), (b) (a)의 후처리 결과로 얻어진 입 영역, (c) 에지 검출기로 검출된 윤곽(입을 펼쳤을 때), (d) (c)의 후처리 결과로 얻어진 입 영역

5. 얼굴 특징 추적

본 논문에서는 mean-shift 알고리즘과 템플릿 매칭을 사용하여 눈과 입 영역을 추적한다. mean-shift 알고리즘은 정해진 범위 내의 탐색 윈도우를 스캔하여 눈이라고 짐작되는 가장 확률 높은 영역을 찾아 탐색 윈도우의 무게 중심을 옮김으로써 점차 눈이 있는 곳으로 이동하며 눈을 찾는다[4]. 템플릿 매칭은 입에 대한 표준 템플릿을 만든 후, 입력 영상과 비교하여 입을 검출하는 것이다[5].

그림 6은 얼굴 특징 추적의 결과를 보여주며 흰 영역이 찾아진 눈과 입 영역이다. 그림에서 보듯이 제안된

시스템은 얼굴 특징들의 영역을 정확하게 추적하고 있다.



그림 6. 얼굴 특징 추적 결과

6. 마우스 조작

컴퓨터는 PC카메라부터 얻어진 영상을 제안된 얼굴 특징 추적 시스템으로 처리함으로써 사용자의 눈과 입의 움직임을 마우스 움직임으로 전환한다.

제안된 시스템은 마우스의 초기 좌표를 첫 프레임의 눈의 중심으로 결정하며 이후 프레임에서는 추적된 눈의 위치에 기반을 두어 자동적으로 좌표가 계산된다. 각 프레임에서 눈 영역에 기반을 둔 좌표와 입의 크기는 윈도우 함수를 통해 시스템에 보내져 마우스의 이동과 클릭 이벤트를 수행한다. 마우스의 포인트는 눈의 움직임에 따라 이동하며 만약 입을 벌린 상태가 되면 좌표에 대응하는 메뉴나 아이콘을 클릭 하는 것으로 간주한다.

7. 실험결과

제안된 시스템의 효율성을 검증하기 위해서 웹 브라우저의 인터페이스로 활용하였다. 실험은 윈도우 XP기반의 펜티엄4 1.7GHz의 성능을 가지는 컴퓨터에서 수행하였다. PC카메라는 USB포트로 컴퓨터와 연결되며 320×240 영상을 초당 30 프레임의 속도로 컴퓨터에 공급한다.

표 1. 시스템의 실험 결과

	Correct	Incorrect
Correct	748 (99%)	2 (1%)
Incorrect	0 (0%)	750 (100%)

실험을 위해서 25명의 사람이 다양한 환경에서 제안된 시스템을 실험하였으며 실험내용은 각 사용자들이 웹 브라우저를 보면서 원하는 아이콘이나 메뉴를 클릭하는 것이다. 각 사용자는 실험 전 5분 동안 연습 시간을 가진 후 30번씩 테스트를 하였으며 결과는 표 1에서 보인다. 클릭과 동시에 눈을 움직이는 경우 메뉴가 클릭되지 않는 경우가 발생했지만 대부분 사용자들은 정확히 웹 브라우저의 메뉴를 클릭했다. 따라서 제안된 시스템은 99%의 정확도를 보인다.

표 2. 각 모듈의 처리 시간(ms)

Stage	Time
Face Detection	60
Eye Detection	110
Mouth Detection	47
Tracking	62
Average Time	78

표 2는 제안된 시스템의 각 모듈에 걸리는 처리 시간을 보여준다. 이 시스템에서 얼굴과 얼굴 특징 영역의 검출은 초기 프레임에서만 이루어지고 이후 프레임에서는 얼굴 특징만이 추적된다. 따라서 한 프레임의 평균 처리 시간은 78ms이며 초당 12 프레임을 처리할 수 있다.

결론적으로 실험결과는 제안된 시스템이 실시간 시스템에 적용되기에 충분하며 몸이 불편한 사람을 위한 인터페이스 및 여러 응용 프로그램에서 일반적인 사용자 인터페이스로 사용 될 수 있음을 보여준다.

8. 결론

본 논문에서는 얼굴 특징 추적을 이용하여 마우스를 움직이는 PC 기반의 인터페이스를 구현하였다. 제안된 시스템은 PC카메라로부터 얻어진 320×240 크기의 영상에서 얼굴 특징 영역을 초당 12 프레임의 속도로 추적하였다. 시스템의 효율성을 검증하기 위해 웹 브라우저의 인터페이스로 활용하였고 25명의 사람들이 테스트하였다. 실험 결과는 제안된 시스템이 실시간 시스템에 적용되기에 충분하며 보다 편리하고 친숙한 인터페이스로 활용될 수 있다는 것을 보여주었다.

참고문헌

[1] Kaufman, Arie E., Bandopadhyay, Amit., Shaviv, Bernard D.: An Eye Tracking Computer User Interface. IEEE 1993 Symposium on Research Frontiers in , 25-26, 1993.

[2] S. H. Park, E. Y. Kim, S. W. Hwang, Y. C. Lee, and H. J. Kim : Face Detection for Security System on the Internet. IEEE International Conference on Consumer Electronics, 276-277, 2001.

[3] Chan A. D. C., Englehart K., Hudgins B., and Lovely D.F. : Hidden Markov model classification of myoelectric signals in speech. IEEE Engineering in Medicine and Biology Magazine, vol. 21, no. 4, pp. 143-146, 2002.

[4] Gary R. Bradski and Vadim Pisarevsky : Intel's Computer Vision Library: applications in calibration, stereo segmentation, tracking, gesture, face and object recognition. IEEE Conference on Computer Vision and Pattern Recognition, 2, 796 -797, 2000.

[5] Lawrence, S., Giles, C.L., Ah Chung Tsoli; Back, A.D. : Face recognition: a convolutional neural-network approach. Neural Networks, IEEE Transactions on Volume 8, Issue 1, Page(s):98-113, 1997.