

내추럴 인터랙션을 위한 사용자 얼굴 인식 기반 인터페이스 구현 방법

엄태영 진문섭 박종일¹⁾

한양대학교 전자컴퓨터통신공학과

{uty02, jinms}@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

A design method of interface base on user's face recognition for natural interaction

Uhm, Taeyoung Jin, Mun-Seop Park, Jong-II

Hanyang University

요약

얼굴을 인식하는 방법은 많은 연구가 이루어져 왔으며, 이를 통해 다양한 형태의 어플리케이션으로 보급되어 왔다. 특히, 사용자의 인식은 사용자와 시스템 간의 한 차원 높은 인터랙션을 제공하는 것이 가능하다. 본 논문에서는 사용자에게 내추럴한 인터랙션을 제공하기 위해 적외선 영상과 컬러영상을 동시에 입력 받아 추출된 휴먼 영역을 이용하여 사용자의 얼굴 인식을 수행하고 특정 사용자에게만 인터랙션을 제공하는 인터페이스 구현 방법을 제안한다. 이를 위해, 간단하게 제작된 시스템을 통해 같은 시점의 적외선과 컬러 영상을 추출하여 조명 변화에 강인하게 얼굴과 손 영역을 인식한다. 이는 얼굴 인식을 기반으로 사용자를 분별하여 서비스 하는 것이 가능하며, 다양한 어플리케이션으로 향상된 서비스를 제공할 것으로 기대한다.

1. 서론

20세기부터 사용자와 컴퓨터 간의 상호작용(Interaction)에 관한 연구들이 많이 이루어져 왔으며, 최근에는 이러한 인터랙션 연구들 중 사용자가 별도의 하드웨어를 직접 사용하지 않아도 되는 비전 기반의 인터페이스들이 주목을 받고 있다. 이러한 휴먼 인터페이스는 컨트롤러 없이 다수의 사람이 있는 환경에서 사용자를 판별하고 움직임을 추적하는 것을 필요로 한다.

본 논문에서는 컨트롤러 없는 자연스러운 인터랙션을 위해 같은 시점의 적외선 영상과 컬러 영상을 동시에 획득할 수 있는 비전 시스템을 이용하여 다수의 사람이 포함된 복잡한 환경에서도 정확하게 주요 휴먼 영역(카메라에서 가장 가까운 사람의 얼굴과 손 영역)을 추출하고, 그 영역에서 사용자의 얼굴을 적외선 영상과 컬러 영상을 이용하여 조명의 변화에도 강인하게 인식하여 추적하는 방법을 제안한다. 이는 사용자에게 고유의 접근 권한을 부여할 수 있을 뿐만 아니라 저가의 웹캠을 이용하여 제작된 시스템과 잘 알려진 비전 방법들을 이용하여 구성되었기 때문에 접근이 쉽고 다양한 디스플레이 기반 인터페이스에 적용이 용이하다.

2. 관련 연구

비전 기반 비접촉식 인터랙션 방법들은 적외선 영상과 컬러 영상을 사용하는 많은 연구들과 새로운 장비들을 적용함으로써 발전되어 왔다. 먼저 적외선 영상을 이용하는 방법들은 적외선 파장의 차이로 얼

굴 혹은 피부 영역을 추출하는 방법[1,2]과 적외선 열화상을 이용하여 객체를 추출하는 방법[3], 그리고 컬러와 적외선 영상을 적절하게 분류를 수행한 연구[4]가 있다.

하드웨어를 기반으로 한 방법들의 경우, 최근 ToF(Time of Flight)를 이용해 다수의 사물에 대해서 깊이 정보를 추출하여 3 차원 추적을 하기 위해 주로 사용되고 있다. 이러한 특수한 장치는 짧은 적외선 펄스파장을 보내어 사물에 반사되어 돌아오는 시간을 측정하여 깊이 정보를 얻는다. 이를 적용하여 질 높은 깊이 정보와 영역 추출이 가능한 방법[5,6]이다. 이 방법을 적용한 대표적인 사례가 마이크로소프트사 키넥트(kinect)이다[7]. 하지만 ToF(Time of Flight) 방법을 위해 적외선 프로젝터를 내장한 특수 장비가 필요 하여 일반적으로 사용하기에는 어려움이 있다.

이렇게 대부분의 인터페이스는 다수의 휴먼 영역에 대해서 특정한 사용자와의 인터랙션을 실행하는 것이 어렵다. 이 문제는 사실상 많은 인터랙션 영역들 중 강인하게 사용자 영역을 분류하여 인식할 수 있는 단서를 이용하면 해결이 가능하다. 그러므로 환경에 강인한 적외선 세기에 대한 영상들을 사용하여 컬러 정보와 함께 인식을 위한 단서로 사용하는 새로운 방법을 제시하고자 한다.

3. 얼굴 인식을 위한 사용자 영역 추출 방법

그림 1은 자연스러운 인터랙션을 위한 얼굴인식 기반 인터페이스의 전체적인 블록 다이어그램을 보여주고 있다. 상호작용 시스템은 크게 적외선 영상처리 부분과 컬러영상처리 부분으로 나눌 수 있다. 이는

1) 교신저자

동시에 이중의 영상을 획득하여 배경추출, 스킨영역 추출, 그리고 적외선 영역 추출 후, 최종 휴먼 영역을 결정하여 사용자 인식(얼굴 인식)을 수행하고 이를 기반으로 인터랙션 실시 여부를 결정함으로써 손 영역을 이용한 자연스러운 인터랙션을 수행한다.

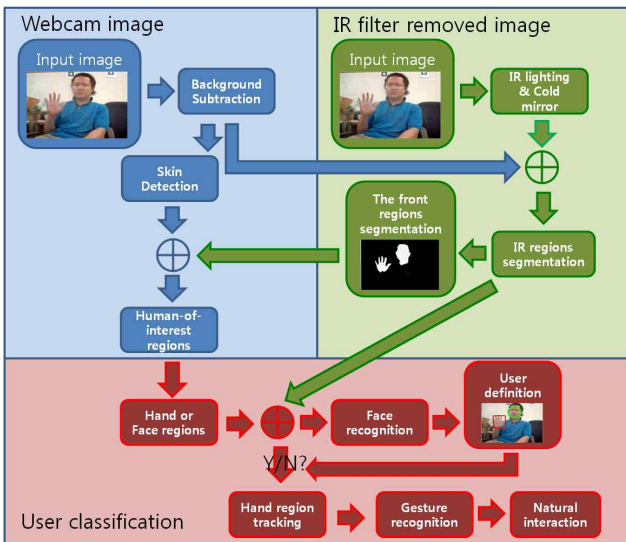


그림 1: 블록 다이어그램

가. 검증시스템

그림 2에서와 같이 휴먼 관심 영역을 추출하는 검증 시스템을 최적화하여 제작하였다. 이는 비전 센싱 파트를 보여준다. 그 구성은 컬러 영역 카메라와 적외선 영역 카메라, 콜드 미러, 그리고 동적인 적외선 빛을 비추기 위한 장치이다. 두 개의 카메라는 저가의 웹캠을 사용하였으며 적외선 영역만을 추출하기 위해서 두 개의 웹캠 중 하나의 적외선 필름을 제거하고 콜드미러를 사용하였다. 이러한 콜드미러를 사이에 두고 물리적으로 같은 거리에서 동일 시점이 되도록 카메라를 위치 시켰으며, 640x480 의 해상도로 30Hz로 영상을 취득하였다. 테스트 베드 시스템은 카메라에 들어오는 적외선 영역을 더욱 효과적으로 만들기 위해서 검은 아크릴 박스로 제작되었다.[8]

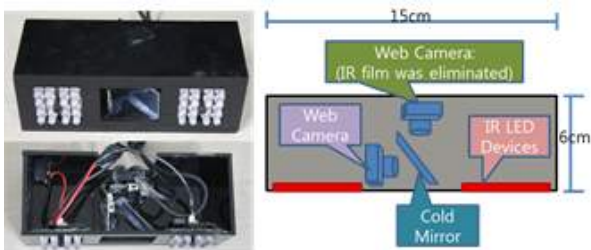


그림2: 검증 시스템

나. 적외선과 컬러 영상을 이용한 얼굴인식 방법

카메라 시스템을 통해 획득하는 영상은 그림 1.에서와 같이 적외선 영상과 컬러 영상을 동시에 같은 시점에서 추출한다. 이러한 특수 영상 기반 얼굴 인식 방법들은 대부분 적외선 열화상을 이용한 방법[9]이나 다수 시점의 컬러와 적외선 영상을 이용한 방법[10]이 주로 사용

되고 있다. 이러한 인식 방법들과 비교하여 같은 시점에서 들어오는 영상을 이용한 두 가지 영상 정보를 혼합하여 인식에 이용하는 것이 더욱 효율적일 수 있다. 특히, TV와 같은 대형 혹은 원거리 디스플레이에서의 손 영역을 이용한 자연스러운 인터랙션을 위해서 적외선 영상 기반 사용자 인식을 통해 사용자를 구별하는 인터페이스가 가능하다.

얼굴 영역은 Adaboost 알고리즘을 사용하여 인식한다. 이 알고리즘의 적용 후에는 추적되는 영역들 중 얼굴 영역은 손 영역과 분류될 수 있을 뿐만 아니라 사용자 분류가 가능 하다. 이러한 영역들을 인식하기 위한 얼굴 추출 문제를 적용하고 얼굴 인식 분류기를 위한 학습으로 Adaboost 학습을 조합하였다 [11]. 따라서 얼굴 영역과 그렇지 않은 영역을 결정하기 위한 분류기는 HaarLike 특징량을 입력으로 하는 층계형 분류기로 인식하였다.

3. 실험결과

그림 3에서는 얼굴 영역을 인식하는 과정을 보여준다. 왼쪽 위 이미지 (a)는 컬러 영역 카메라로부터 입력된 원본 영상을 나타내며, 오른쪽 위 이미지 (b)는 적외선 영상에서 Adaboost 알고리즘을 이용해 인식된 얼굴 영역을 보이며, 왼쪽 아래 이미지 (c)는 얼굴 인식 후 컬러 영상에서 휴먼 영역을 매칭하여 손 영역을 추적하는 것을 보인다. 마지막으로 오른쪽 아래 이미지 (d)는 (b)와 (c) 이미지들의 결과를 원본 영상에서 합성하여 보여준다.

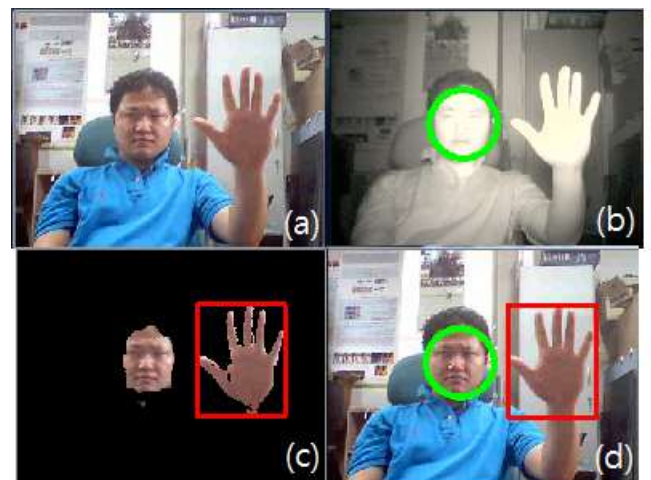


그림3: Adaboost 알고리즘에 의한 휴먼 관심 영역 인식, (a) 원본 영상, (b) 적외선 영상 기반 얼굴 인식 영역, (c) 컬러 영상과 영역 매칭 및 분류, (d) 최종 인식 및 트래킹 결과.

또한, 실질적으로 TV에 적합하도록 소형화한 시스템을 이용하여 간단한 어플리케이션에 적용하였다. 얼굴 인식 후 그림 4. (a)에서와 같이 손 영역을 트래킹 하고, 그림 4. (b)에서와 같이 TV에 적용하여 인터랙션을 수행하였다. 손 영역 트래킹은 사용자의 얼굴 영역으로부터 가장 가까운 손 영역을 추적하였으며, 그림 4. (c)와 (d)에서는 적용된 퍼즐 어플리케이션을 제시한다.

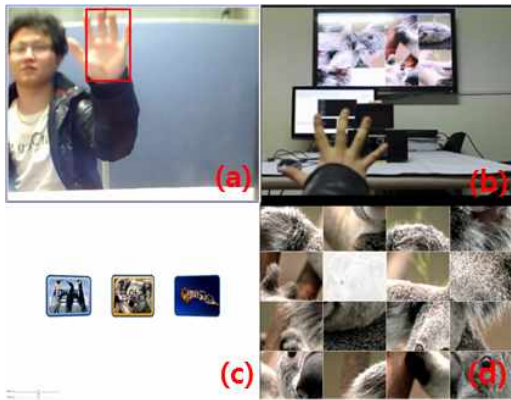


그림4: 자연스러운 인터렉션 예시, (a) 손 영역 추적 영상, (b) TV에 시스템 적용 후 인터렉션 예시, (c) 퍼즐 어플리케이션 메뉴, (d) 퍼즐 어플리케이션 동작 화면.

감사의 글

본 연구는 지식경제부, 문화체육관광부 및 한국산업기술평가관리원의 IT산업원천기술개발사업의 일환으로 수행하였음. [KI001820, 영상 및 비디오 콘텐츠를 위한 계산사진학 기술 개발]

4. 결론 및 향후 과제

본 논문에서는 컬러 영상에서 휴먼 영역(얼굴과 손 영역)을 추출하고 다양한 조명 환경에 강인한 적외선 영상에 기반한 얼굴 인식을 수행하여 상호작용에 이용한다. 이러한 자연스러운 인터랙션을 기반으로 TV와 같은 원거리 디스플레이에 간단한 어플리케이션 형태로 적용하였다. 이는 실내에서 다양한 조명 환경의 영향을 받지 않고 얼굴을 인식하여, 사용자에게 사용자 인증과 자연스러운 상호작용을 동시에 할 수 있도록 제공하였다.

향후에는 손 동작에 대한 인식을 고려하여 다양한 어플리케이션으로 확장하여 적용이 가능하도록 발전시키는 연구가 필요하다.

참고문헌

[1] J. Dowdall, I. Pavlidis, and G. Bebis "Face detection in the near-IR spectrum," *Image and Vision Computing* 21, pp. 565 - 578, 2003.

[2] Y. Suzuki, K. Yamamoto, K. Kato, M. Andoh, and S. Kojima "Skin detection by near infrared multi-band for driver support system," *Proceedings of ACCV*, pp. 722-731, 2006.

[3] K. Yasuda, T. Naemura, and H. Harashima, "Thermo-Key: human region segmentation from video," *IEEE Computer Graphics and Applications*, 24(1): 26-30, 2004.

[4] L. Kun, G. Leia, L. Huihui and C. Jingsong "Fusion of color and infrared video for moving human detection," *Chinese Journal of Aeronautics* 22, pp.75-80, 2009.

[5] E. Parvizi and Q. M. J. Wu "Multiple object tracking based on adaptive depth segmentation," *Canadian Conference on Computer and Robot Vision*, pp. 273-277, 2008.

[6] P. Breuer, C. Eckes, and S. Müller, "Hand gesture recognition with a novel IR time-of-flight range camera - A pilot study," *Lecture Notes in Computer Science* 4418, pp. 247-260, 2007.

[7] Microsoft Corp. Redmond WA. Kinect for Xbox 360.

[8] T. Uhm, H. Park, and J. Park, "Human-of-Interest Tracking System for Natural Interaction," *Proceedings of International Conference on Consumer Electronics (ICCE'11)*, pp. 623-624 Las Vegas, USA, Jan. 9-12, 2011

[9] Seong G. Kong, Jingu Heo, Besma R. Abidi, Joonki Paik, and Mongi A. Abidi, "Recent advances in visual and infrared face recognition - a review," *Computer Vision and Image Understanding* 97 (2005) 103-135.

[10] M. Grgic, K. Delac and S. Grgic, "SCface - surveillance cameras face database," Springer Science + Business Media, LLC 2009.

[11] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," *Proc. of IEEE ICIP*, vol. 1, pp. 900-903, 2002.