얼굴의 특이점 검출 및 실시간 추적을 이용한 e-Book 제어

김현우 · 박주용 · 이정직 · 윤영로

연세대학교 대학원 의공학부

Unconstrained e-Book Control Program by Detecting Facial Characteristic Point and Tracking in Real-time

Hyun-woo Kim, Joo-yong Park, Jeong-jick Lee and Young-ro Yoon

Department of Biomedical Engineering, Graduate School of Yonsei University

(Manuscript received 7 December 2013; revised 8 March 2014; accepted 15 April 2014)

Abstract: This study is about e-Book program based on human-computer interaction(HCI) system for physically handicapped person. By acquiring background knowledge of HCI, we know that if we use vision-based interface we can replace current computer input devices by extracting any characteristic point and tracing it. We decided between-eyes as a characteristic point by analyzing facial input image using webcam. But because of three-dimensional structure of glasses, the person who is wearing glasses wasn't suitable for tracing between-eyes. So we changed characteristic point to the bridge of the nose after detecting between-eyes. By using this technique, we could trace rotation of head in real-time regardless of glasses. To test this program's usefulness, we conducted an experiment to analyze the test result on actual application. Consequently, we got 96.5% rate of success for controlling e-Book under proper condition by analyzing the test result of 20 subjects.

Key words: real-time tracking, circle-frequency filter, YCbCr color model, template matching, human-computer interaction(HCI), e-Book

1. 서 론

장애인가가 증가하며 장애인의 사회참여 분위기가 조성되고 있다. 이에 따라 장애인을 위한 정책, 시설 개선 등 장애인의 전반적인 생활환경이 빠르게 변화하고 있으며 그들의 삶의 질 개선을 위한 기술 요구 또한 증가하고 있다[1]. 하지만 일부 지체장애인은 타인의 도움 없이 이를 사용하는

데 있어 불편함을 겪고 있기 때문에 불편함을 해소할 수 있는 기술이 필요하다.

이를 위한 대표적인 기술이 인간과 컴퓨터의 상호작용(human-computer interaction, HCI) 시스템이다. 이는 음성, 생체신호, 시각 등의 인터페이스를 활용하여 인간의 제한된 능력을 향상시키거나, 인간이 컴퓨터를 편리하게 사용할 수 있도록 하는 기술이다. 이 중 시각 인터페이스는 측정 장소 또는 주위 조명 등에 쉽게 영향을 받지만, 카메라와 같이 구축성을 요구하지 않는 장비를 사용하므로 측정이 간단하다[2].

관련된 연구의 예로는 적외선 카메라를 사용하여 눈을 검출하고 이를 실시간으로 추적하거나[3], 눈과 입 모두를 검출하고 눈의 응시 방향과 입 모양을 인식하여 마우스와 키보드를 대체하는 연구 사례가 있었다[4-6]. 또한 일반 카메라를 사용하여 눈을 검출하고 이를 실시간으로 추적하여 얼굴의 회전동작을 추적하는 연구 사례도 있었다[7]. 하지만 상기의 연구들은 고가의 적외선 카메라를 사용하거나 다수

Corresponding Author : Young Ro Yoon
220-842, Medical Industry Techno Tower #201, Yonsei University, Maeji, Heungup, Wonju, Gangwon, Republic of Korea.

TEL: +82-33-760-2809 / FAX: +82-33-763-1953

E-mail : yoon@yonsei.ac.kr

본 연구는 교육부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임 (No.2011H1B8A2003304).

This research was financially supported by the Ministry of Education (MOE) and National Research Foundation of Korea(NRF) through the Human Resource Training Project for Regional Innovation (No. 2011H1B8A2003304).

그림 1. e-Book 제어 프로그램 알고리즘 순서도.

Fig. 1. Algorithm flow chart of e-Book control program.

의 일반 카메라를 사용하는 경우가 있었다. 이 외에도 검출 및 추적의 정확성을 나타내는 지표를 제시하지 않았다는 단점이 있다.

본 연구에서는 이와 같은 연구의 한계점을 보완하고자 손이 불편한 지체장애인의 편리한 독서 환경을 조성함으로써 실질적으로 삶의 질 향상에 도움 줄 수 있는 e-Book 제어 프로그램을 개발하였다. 개발된 프로그램은 적외선 카메라를 사용하지 않고 하나의 웹캠만으로 사용자의 얼굴 회전 동작을 인식하여 그 동작에 따라서 e-Book을 제어할 수 있

도록 설계하였다. 또한 프로그램의 객관적 평가를 위한 지표를 마련하고자 실험 프로토콜을 정의하고 제어 성공률의 분석을 통해 그 효용성을 검증하고자 한다.

II. 연구 방법

본 연구는 그림 1 과 같은 순서도를 기반으로 알고리즘을 구현하였다. 알고리즘은 이미지 전처리, 눈 사이 후보점 추출, 동공 영역 검출, 영상 추적의 과정으로 나뉜다.

1. 이미지 전처리

Circle-frequency filter는 명암 정보를 입력 값으로 가지는 입력 영상에 대한 이산 푸리에 변환(discrete fourier transform; DFT)의 일종이다. 이산 푸리에 변환의 특성 중 하나로 입력 신호가 코사인파(cosine wave)의 형태일 때 스펙트럴 파워(spectral power)가 커지는 특징이 있는데 이를 이용해 얼굴의 특징점을 추출할 수 있다. 이 필터를 적용하기 위한 선행 과정으로 그레이스케일 변환을 하고 이동평균필터를 적용한다[8]. 본 연구에서는 그림 2(a)와 같은 영상에 선행 과정을 적용한 결과 그림 2(b)와 같은 영상을 얻었다. 영상의 크기는 352×288 픽셀이며, 이동평균필터의 크기는 5×5 픽셀이다. Circle-frequency filter 적용을 위한 반경 r 은 24픽셀로 설정하여 그림 2(b)에 적용한 결과 그림 2(c)와 같은 영상을 얻었다.

그림 2. 영상신호처리 과정의 영상 예 (a) 원본 RGB 영상 (b) 그레이스케일 변환 및 이동평균필터 적용 영상 (c) circle-frequency filter 적용 영상 (d) 선별된 1차 눈 사이 후보점 (e) 선별된 2차 눈 사이 후보점 (f) 동공 영역 검출 및 눈 사이 중심점 추출 영상.

Fig. 2. Example of image signal processing (a) original RGB image (b) application of grayscale and moving average filter (c) application of circle-frequency filter (d) first candidate point (e) second candidate point (f) detection of pupil area and center point of between eyes.

2. 눈 사이 후보점 추출

1차 눈 사이 후보점을 찾기 위해 스펙트럴 파워가 최대인 점을 찾았다. 이를 위해서 영상을 16×16 픽셀의 정사각형 영역으로 구분하여 총 22×18 개의 블록으로 나누었고 각 블록에서 최댓값을 가지는 픽셀을 1차 눈 사이 후보점으로 선별하였다[8]. 선별된 1차 눈 사이 후보점은 그림 2(d)와 같이 많은 수가 검출되었다. 이로 인해 연산에 많은 시간이 소요될 뿐 아니라 추후에 동공을 검출하는 과정에서 머리카락이 검출되는 문제가 발생할 수 있다. 따라서 병렬적인 영상처리 과정인 YCbCr 색상 변환과 YCbCr 피부 모델 적용을 통해[9] 피부 영역에 해당하는 후보군만을 선별해서 후보군의 수를 줄이는 과정을 거쳤다. YCbCr 피부 모델을 적용하기 위해 1차 눈 사이 후보점을 선별할 때와 같은 크기의 블록 단위(16×16 픽셀)로 구획을 나누었다. 그리고 각 블록에서 절반 이상의 픽셀이 YCbCr 피부 모델에 포함될 때 그 블록을 피부 영역으로 판별하여 그림 2(e)와 같이 피부 영역에 해당하는 2차 눈 사이 후보점을 선별하였다.

3. 동공 영역 검출

2차 눈 사이 후보점 중 단 하나의 눈 사이 중심점을 찾기 위해서 그레이스케일 영상에서 동공 추정 영역의 임계 값을 설정하여 비교하는 방법으로 동공을 검출하였다[10]. 웹캠을 통해 입력되는 영상의 통계적 수치를 분석한 결과 눈 사이 중심점과 동공의 거리가 $40 \sim 45$ 픽셀 범위인 것을 확인하였다. 따라서 2차 눈 사이 후보점에서 수평 양방향에 대해 각각 45픽셀 떨어진 점을 중심으로, 40×20 픽셀 크기의 동공 검출 영역을 선택하였다. 다수의 예비 실험을 통하여 동공 검출을 위한 임계 값을 결정하고(임계 값 = 10) 임계 값보다 작은 픽셀을 선택 영역에서 검출하였다. 그림 2(f)와 같이 양쪽 모두에서 동공 영역을 검출하는 2차 눈 사이 후보점을 찾아 이를 눈 사이 중심점으로 결정하였다.

4. 영상 추적

검출된 눈 사이 중심점을 실시간으로 추적하기 위해 그 점을 중심으로 하는 24×24 픽셀의 영상을 템플릿으로 생성하고 저장하였다. 그리고 템플릿 매칭법(template matching method)을 활용하여 매 프레임마다 템플릿을 갱신하고 추적하는 과정을 반복하였다[11]. 하지만 안경 착용자가 본 프로그램을 사용할 경우 눈 사이 중심점 영역에 포함되는 안경의 입체적 구조 때문에 템플릿 추적에 실패하는 경우를 확인하였다. 이를 보완하기 위해 저장할 템플릿이 안경을 포함하지 않는 영역이 되도록 템플릿 생성 위치를 변경하였다. 따라서 입력되는 영상의 통계적 수치 분석을 통해 검출된 눈 사이 중심점으로부터 25픽셀 하단에 위치한 콧등 영역을 템플릿 영상으로 설정하였다.

그림 3. 실험 프로토콜(효율성 검사 실험).
Fig. 3. Protocol of experiment.

III. 실험 방법

프로그램의 적절한 사용 조건을 확인하기 위해 얼굴 회전 속도 및 화면과의 거리를 결정하는 선행 실험을 하였다. 본 연구에서는 얼굴의 회전이 1회 완료되었을 때를 1틱으로 설정하고, 속도 단위는 분당 틱 수(ticks per minute; TPM)를 사용한다. 또한 얼굴 회전 동작에 의해 페이지 전환이 올바르게 된 경우를 제어가 성공한 경우로 보았다. 이와 같은 선행 실험 결과 90 TPM 이하의 회전 속도와 $45 \text{ cm} \sim 70 \text{ cm}$ 의 거리 조건에서 제어가 모두 성공했다. 이를 바탕으로 60 TPM의 속도, 50 cm의 거리를 사용 조건으로 설정하여 실제 사용 환경에서의 효율성을 검증하기 위한 효율성 검사 실험을 그림 3과 같은 프로토콜로 진행하였다. 피험자는 본 프로그램 사용에 영향을 주지 않도록 경추 관련 질환이 없는 일반 성인남녀 20명(남성 10명, 여성 10명, 23 ± 2 세; 안경 착용자 2명, 미착용자 18명)을 선정하였다. 본 실험에 앞서 피험자에게 프로그램의 동작 원리를 설명을 통해 충분히 숙지시켰고, 90 TPM 이하의 적정 회전 속도 범위에서 수직 네 방향에 대한 훈련을 5회 시행하였다. 훈련 완료 후 회전 속도를 피험자 임의로 설정하고 좌우 왕복 20회, 상하 왕복 20회를 시행하여 제어 성공률을 분석하였다. 위의 실험은 얼굴에 빛이 고르게 비치는 실내조명 조건 하에서 진행하였다.

IV. 결 과

1. 구현 결과

프로그램은 다음과 같은 환경에서 개발되었다. 사용한 웹캠

- ① 템플릿화 버튼
 - 템플릿 생성 및 추적
 - 템플릿 표시 영역에 템플릿 표시
- ② e-Book 열기 버튼
 - PDF 형식의 파일 불러옴
- ③ 웹캠 영상 영역
 - 30frames/sec 로 입력 영상 출력
- ④ 좌표 표시 영역
 - 기준 템플릿과 유동 템플릿의 중심점 좌표 표시
- ⑤ 템플릿 표시 영역
 - 기준 템플릿 : 붉은색 사각형
 - 유동 템플릿 : 푸른색 사각형
- ⑥ e-Book 출력 영역
 - 파일명 표시 및 e-Book 출력
 - 현재페이지 / 마지막 페이지 표시

그림 4. 구현된 e-Book 제어 프로그램 인터페이스.

Fig. 4. Interface of e-Book control program.

그림 5. 실험 결과(피험자 n = 20)

Fig. 5. Result of experiment(subjects n = 20).

은 1300만 화소의 노트북 내장형 Vimicro UVC PC Camera이고, 화면에 표시되는 영상의 크기는 352×288 픽셀이다. 개발 언어는 Visual studio 2010 C#이다. 개발된 e-Book 운용 프로그램은 그림 4와 같은 인터페이스로 구성된다.

2. 실험 결과

실제 프로그램 사용시의 효용성을 확인하는 실험 결과는 그림 5와 같다. 피험자의 개별 평균 데이터 분석 결과 최소 92.5% 이상의 제어 성공률을 보인다. 그리고 전체 피험자 20명에 대한 데이터 분석 결과 평균 96.5%의 제어 성공률을 확인할 수 있다.

V. 토 의

본 연구에서는 하나의 웹캠만을 사용하여 얼굴의 특징점

을 검출하고 이를 e-Book 프로그램 실시간 제어에 활용하였다. 실험 결과, 20명의 피험자를 대상으로 평균 96.5%의 높은 제어 성공률을 보였기 때문에 실제 적용에 효용성이 있다고 판단된다. 하지만 2차 눈 사이 후보점 선정을 위해 사용한 피부 영역 추출 모델이 동양인에게 한정된 피부 모델이었기 때문에 다양한 피부 모델 피험자에 대한 추가적인 실험이 필요하다.

개발된 프로그램 사용에 있어 조명 조건이 변할 경우 제어 성공률의 저하가 유발되므로 실험 시에는 한정된 조명 조건에서 실험하여 그 영향을 최소화하였다. 따라서 안면에서 발생하는 불균일한 조명분포 효과를 최소화시킬 수 있는 조명 정규화와 같은 영상처리 방법이 적용된다면[12,13] 구현한 프로그램이 효용성을 현재보다 더 갖출 수 있을 것이라 사료된다.

Reference

- [1] K.D. Kwon, "Statistics of disabled person", Employment Development Institute, Nov. 2012.
- [2] J. Varona, C. Manresa-Yee, F.J. parales, "Hands-free vision-based interface for computer accessibility", JNCA, vol. 31, pp. 357-374, 2008.
- [3] Y. Li, S. Wang, X. Ding, "Eye/eyes tracking based on a unified deformable template and particle filtering", Pattern Recognition Letters, vol. 31, pp. 1377-1387, 2010.
- [4] G.H. Shin and J.C. Chun, "Vision-based Multimodal Human Computer Interface based on Parallel Tracking of Eye and Hand Motion", International Conference on Convergence Information Technology, Gyeongju, Korea, Nov. 2007, pp. 2443-2448.
- [5] K. Arai and R. Mardiyanto, "Camera Mouse Including "Ctrl-Alt-Del" Key Operation Using Gaze, Blink, and Mouth Shape", IJACSA, vol. 4, no. 3, pp. 183-191, 2013.
- [6] K. Arai and R. Mardiyanto, "Eye-based Human Computer Interaction Allowing Phoning, Reading E-Book/E-Comic/E-Learning, Internet Browsing, and TV information Extraction", IJHCI, vol. 2, no. 12, pp. 26-32, 2011.
- [7] M. Nebete, M. Lokhande, K. Ahire, "Design an Eye Tracking Mouse", IJARCCCE, vol. 2, issue 2, pp. 1118-1121, Feb. 2013.
- [8] S. Kawato and J. Ohya, "Real-time of Nodding and Head-shaking by Detecting and Tracking the "Between Eyes", in Proc. 4th IEEE International Conference on Automatic Face and Gesture Recognition, Mar. 2000, pp. 40-45.
- [9] J.H. Jung, S.I. NA, J.H. Lee, M.C. Shin and D.S. Jung, "Robust Face Detection using Geometric Luminance Distribution Mask and Color Model under Illumination Variation", Proc. Of KIISE Fall Conference, vol. 32, no. 2, 2005, pp. 913-915.
- [10] A. Giachetti, "Matching techniques to compute image motion", Image and Vision Computing, vol. 18, pp. 247-260, 2000.
- [11] R. Stiefelhagen, J. Yang, and A. Waibel, "Tracking eyes and monitoring eye gaze", Proceedings of the workshop on perceptual user interfaces, Alberta, Canada, Oct. 1997, pp. 98-100.
- [12] S. Shan, W. Gao, B. Cao, D. Zhao, "Illumination Normalization for Robust Face Recognition Against Varying Light Conditions", IEEE International Workshop on Analysis and Modeling of Faces and Gestures, Nice, France, Oct. 2003, pp. 157-164.
- [13] W. Chen, M. Joo and S. Wu, "Illumination Compensation and Normalization for Robust Face Recognition Using Discrete Cosine Transform in Logarithm Domain", IEEE T SYST MAN CY B, vol. 36, no. 2, pp. 458-466, 2006.