
FPGA와 OpenCV를 이용한 눈동자 모션인식을 통한 의사소통 시스템

이희빈 · 허승원 · 이승준 · 유운섭

한경대학교 전기전자제어공학과, IITC

Implementation to human-computer interface system with motion tracking using OpenCV and FPGA

Hee Bin Lee · Seung Won Heo · Seung Jun Lee · Yun Seop Yu

Dept. of Electrical, Electronic, Control Eng. and IITC, Hankyong National University

E-mail : lkb1721@naver.com

요 약

본 논문은 FPGA를 이용해서 몸을 움직이지 못하는 루게릭병 환자들을 위해 동공을 추적하여 의사소통을 할 수 있게 해주는 시스템을 소개한다. OpenCV를 이용한 얼굴 및 눈동자 추적과 DE1-SoC보드를 이용하여 눈의 움직임 검출과 문자를 출력한다. 본 논문에서는 웹캠을 사용하고 눈동자를 추적하고 눈동자의 좌표 값에 따라 눈동자의 움직임을 파악하고 사용자의 의도에 맞는 문자를 출력한다. 비교적 개발비용이 저렴하고 업데이트 가능한 FPGA를 사용하고, 블루투스를 이용하여 핸드폰으로 쉽게 문자를 출력할 수 있는 시스템을 제안한다.

ABSTRACT

This paper introduces a system that enables pupillary tracing and communication with patients with amyotrophic lateral sclerosis (ALS) who can not move free. Face and pupil are tracked using OpenCV, and eye movements are detected using DE1-SoC board. We use the webcam, track the pupil, identify the pupil's movement according to the pupil coordinate value, and select the character according to the user's intention. We propose a system that can use relatively low development cost and FPGA can be reusable, and can select a text easily to mobile phone by using Bluetooth.

키워드

FPGA, OpenCV, SoC, HDL

1. 서 론

루게릭병이란 운동세포만 선택적으로 사멸되어 몸을 움직이지 못하게 되는 병이다. 일 년에 10만 명당 약 1~2명에게서 발병하는 것으로 알려져 있다. 몸을 못 움직이고 의사소통을 못하는 루게릭병 환자들도 의사소통이라도 할 수 있게 된다면 환자들도 더 나은 삶을 살 수 있게 될 것이다.

눈동자를 추적해 문자를 출력하는 방법에는 두 가지 방법이 있다. 첫 번째로 눈동자가 어느 시간대에 움직였는지에 따라 문자가 결정되는 방법이고 두 번째는 눈동자가 움직인 방향에 의해 문자를 결정하는 방법이다. 첫 번째 방법은

시간이 오래 걸리지만 사용자의 눈의 피로를 덜어줄 수 있고 사용자가 움직임이 더욱 제한되었을 경우에 하나의 입력 신호로 선택할 수 있고, 두 번째 방법은 눈이 피로하지만 문자를 빠르고 효율적으로 출력할 수 있다.

본 논문에서는 저비용 고효율인 시스템을 만들기 위하여 SoC(system-on-chip)칩을 사용하고, 눈동자를 쉽게 추적할 수 있는 적외선 카메라를 사용하지 않아 추가적인 비용을 최소화한다. SoC는 한번 설계해 놓으면 생산 및 조립 비용이 타 하드웨어보다 저렴하다. 본 논문에서는 SoC 기반의 눈동자 움직임에 따른 의사소통 시스템 구현에 대해서 소개한다.

II. 본 론

SoC에 기반의 눈동자 모션 인식에 따른 의사소통 시스템[1-3]에 사용되는 장비는 그림 1-3과 같다. 눈동자 모션인식 시스템을 제어할 FPGA 및 Dual-core ARM Cortex-A9 프로세서가 장착된 DE1-SoC[4], 사용자의 얼굴 및 눈을 촬영할 카메라인 C922 PRO Stream Webcam, 보드 - 휴대폰 간 UART 형식으로 문자를 송수신하기 위한 블루투스 모듈 HC-06이 있다.



그림 1. C922 PRO

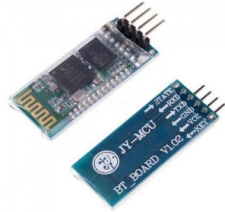


그림 2. HC-06



그림 3. DE1-SoC

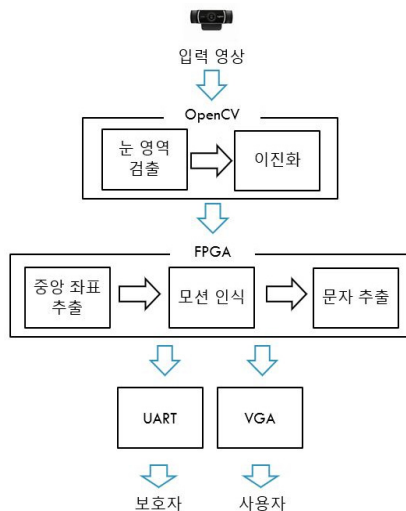


그림 4. 하드웨어 및 알고리즘 순서도

본 시스템 구성(그림 4)은 다음과 같다. 1) Linux 환경에서 OpenCV를 통한 눈 영역 검출 및 이진화

를 하고, 2) FPGA에서 눈동자의 좌표를 설정하고, 3) 눈동자의 좌표값에 기반하여 눈동자의 모션을 인식한다. 4) 모션 인식한 결과에 따른 문자를 선택하고, 5) UART형식으로 블루투스로 보호자 또는 사용자와 대화할 사람에게 휴대폰으로 문자를 보낸다. 또한 사용자 본인은 VGA를 통해 문자를 눈동자의 움직임으로 선택하게 된다.

1) 눈 영역 검출 및 이진화

눈 영역 검출은 쉽게 설계하기 위해서 OpenCV Library의 Haar feature-based cascade classifiers를 사용한다. 이미 많은 이미지들이 학습되어 있고, xml 파일을 통해서 쉽게 사용할 수 있다. 다만, 눈 영역 검출하는 데 Cortex-A9 기준으로 프레임이 낮기 때문에 눈 영역에 대한 범위 좌표 값을 구한 뒤, 그 범위 좌표에 있는 영상을 추출한다. 사용자는 환자이므로 거의 움직임이 없다. 따라서 실시간으로 눈 영역의 위치를 갱신할 필요는 없다. 다만 본 보고서를 실험할 때 사람이 가만히 있기는 어려우므로 수 초에 한 번씩 눈 영역의 범위 좌표를 갱신할 필요가 있다.

이진화는 FPGA에서 처리 속도 향상과 Linux에서 FPGA로 넘어가는 과정에서 적은 용량을 사용하고, 빠르게 전송하기 위한 것이다. 이진화는 특정한 임계값을 정하고 이미지의 화소 값을 임계값과 비교해 작으면 0 크면 1로 변환하는 작업이다. 이때 눈동자는 짙은 검은색 또는 갈색이기 때문에 임계값보다 낮고, 흰자는 핏줄이 보이더라도 임계값보다 높기 때문에 눈동자와 흰자가 뚜렷히 구별된다.



그림 5. 이진화된 눈 영역

그림 5와 같이 이진화된 눈 영상은 눈동자를 제외한 양 옆부분에도 검은 색이 나타난다. 따라서 정확한 눈동자 모션 인식을 위해 검은색 부분의 좌표들의 평균 값과 차이가 많이 나는 검은색 부분을 제외시킨다. 동공 부분이 하얀 색을 띄는 것은 밝은 조명 아래서 실행했기 때문이고, 실행 환경에 따라서 동공과 홍채를 구별할 수 있는지가 달라진다.

2) 눈동자 좌표값 설정 및 모션 인식

그림 5와 같이 눈 영역이 눈동자처럼 원으로 나오지 않기 때문에, 빨간색 선을 따라서 좌, 우 방향으로 모션인식을 한다. 좌, 우 모션 인식은 중앙 좌표값을 기준으로 눈동자 위치를 판별해야 한다.

입력받은 이진화된 눈 영역 중에서 중심점은 OpenCV의 눈 영역 검출을 시행할 때 흰자와 눈동자의 화소 차이로 눈 영역을 검출하므로 눈

영역 중심부와 눈동자의 중심부가 대략적으로 일치한다. 다만, 눈 영역 좌표 검출의 프레임이 낮고, 눈 영역은 실시간으로 전송하므로 차이가 있을 수 있다. 따라서 FPGA에서 다시 중앙 좌표를 인식해줄 필요가 있다. 이진화된 눈 영역의 중심부는 Low bit가 가장 밀집해 있는 곳이다. 따라서 모든 비트를 XOR 연산을 실행한 후 High bit의 좌표 값들의 평균으로 중심부를 구할 수 있다. 중심부 좌표가 특정 경계선을 기준으로 좌, 우로 갔으면 움직였다고 판단한다.

3) 문자 추출

시중에 나온 사지마비 환자를 위한 문자 선택 방법은 기본 쿼티 키보드나 천지인 등 여러 가지가 있다. 본 논문에서는 좌, 우 움직임으로 문자표 중에 하나를 선택해야 하므로 마우스처럼 버튼을 누르는 식으로는 구현할 수 없다. 2가지 입력으로 1가지 출력, 한글을 선택하기 위해서는 2차원 배열을 사용하는 것이 가장 효율적이다.

한글은 자모음이 점과 선으로 구성되는 특수성 때문에 효율적으로 자판을 구성할 수 있다. 가장 대표적으로 천지인이 있고, 환자를 위해 천지인을 활용한 여러 자판들이 있다. 자판은 글자를 입력하는 부분과 삭제나 전송하는 부분으로 구성되며 4x4 또는 4x5의 배열로 되어 있다. 따라서 좌우 움직임으로 행과 열을 선택해 해당 성분을 선택하는 형식으로 문자를 추출할 수 있다.

4) 인터페이스 출력

선택된 문자를 출력하기 위해 UART 형식으로 출력한다. 수신부에서 다시 문자로 쉽게 변환하기 위해 한글을 유니코드로 변환한다. 유니코드는 2byte이기 때문에 1byte씩 나눠서 순차적으로 출력하고, 수신부에서 순차적으로 받아서 다시 한글로 변환한다. 선택된 문자를 유니코드로 변환하는 것은 SoC안의 ROM을 이용한다. 본 논문에서는

사용자가 보는 환경을 구성하기 위해 VGA를 사용한다. 여기서 사용자는 실시간으로 선택하고 싶은 문자 및 선택된 문자를 볼 수 있다. 선택한 문자 테이블은 ROM에 저장되며 이미지를 MATLAB으로 변환한 .mif 파일 형식으로 ROM의 초기값으로 저장된다. ROM에서 나오는 이미지 화소 값을 VGA의 타이밍 신호에 맞춰서 모니터에 출력한다.

III. 결 과

OpenCV로 얼굴 및 눈 영역을 검출할 때 Detection time이 0.38~0.45 초가 나온다. 따라서 실시간 눈 영역 검출하기에는 무리가 있지만, 한번 영역을 검출하고, 몸을 움직이지 않으면 눈 영역을 실시간으로 보낼 수 있다. 이진화된 영상으로 모션 인식하는 것은 프레임 지연 없이 실시간으로 가능했다. 다만, 실험 과정 중 메뉴판

을 보다가 눈이 의도치 않게 움직이므로 사용자에게 화면과 적당한 거리를 제공해야 할 것이다. 마찬가지로 OpenCV가 정면 얼굴을 검출하므로 카메라는 사용자의 정면에 놓아야 한다.



그림 6. 눈동자 위치 추출 결과 화면

그림 7에서 보여진 것처럼 VGA 포트가 연결된 모니터에서 6x5 표에서 선택된 행렬을 십자가로 교차해 색을 변환하는 것으로 어느 곳에 커서가 있는지 파악할 수 있다. 신호에 따라 행과 열이 선택 되는 것은 지연없이 실행되었다.

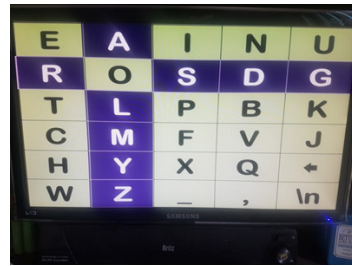


그림 7. 사용자용 VGA 모니터 화면

Linux device driver를 이용한 Linux to FPGA 전송 및 모션 인식은 구성 중에 있다.

IV. 결 론

본 요약에서는 OpenCV와 FPGA를 활용하여 눈 영역 검출 및 모션 인식을 통한 문자 출력을 하였다. 이 시스템은 OpenCV 라이브러리를 통해 쉽게 구현했고, FPGA를 통해서 저렴한 가격을 갖추면서도 재 프로그래밍이 가능한 시스템을 구현했다. 카메라와 SoC 등의 장비 외의 추가적인 장비 없이 가성비가 높은 시스템을 구현하였다. 사용자의 환경 특성상 추가적인 비용 지출이 힘든데, 이 부분에서 조금이나마 도움이 될 수 있을 것이라고 생각된다.

참고문헌

- [1] P. Salunkhe and A. R. Patil, "A device controlled using eye movement," in *Proc.*

-
- International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, p. 732.*
- [2] R. Ramesh and M. Rishikesh, "Eye ball movement to control computer screen," *J. Biosens. Bioelectron.* vol. 6, no. 3, pp. 1000181, 2015.
- [3] M. S. Narayan and W. P. Raghoji, "Enhanced cursor control using eye mouse," *International Journal of Advances in Electronics and Computer Science*, vol. 14, no. 1, pp.40-44, Mar. 2016.
- [4] Altera DE1-SoC User Manual [Internet], Available: <https://www.altera.com/support/training/university/boards.html#de1-soc>.