Title

Hand Gesture Segmentation Method using a Wrist Worn Wearable Device

ABSTRACT

Objective: We introduce a hand gesture segmentation method using a wrist worn wearable device which can recognize simple clenching and unclenching a fist gesture.

Background: There are many types of smart watch and fitness band in the market. And most of them already adopt a gesture interaction to provide ease of use and increase user experience. However, there are many cases in which the malfunction is difficult to distinguish between the user's gesture commands and user's daily life motion. It is needed to develop a simple and clear gesture segmentation method to improve the gesture interaction function and user experience.

Method: At first, to segment a gesture we defined making a fist (start of gesture command) and releasing a fist gesture (end of gesture command) as segmentation gesture. Making a fist and releasing gestures are simple and intuitive. And we also designed a single gesture consists of a set of making a fist gesture, a command gesture, and releasing a fist gesture in order. To detect segmentation gestures on the bottom side of the wrist we used a wrist strap which an array of infrared sensors (emitters and receivers) are mounted on it. When user make a fist and release gesture, this changes the shape of the bottom side of the wrist and the reflected amount of the infrared light that was detected by the receiver sensor are also changes simultaneously.

Results: An experiment was conducted in order to evaluate gesture segmentation recognition. 12 participants took part in the experiment: 10 males, and 2 females with an average age of 38. The recognition rates of the segmentation gesture, holding a fist and unclenching the fist, are 99.58% and 100% respectively.

Conclusion: Through the experiment, we have evaluated segmentation gesture and its usability. The experimental results show a potential for our suggested segmentation method in the future.

Application: This can be adapted to user interface for fashion apparel such as smart watch and wrist bands.

Keywords

Gesture segmentation, Wearable computer, Adaptive threshold, Smart apparel.

1. Introduction

웨어러블 컴퓨터는 섬유/의류/패션 산업과 밀접한 관계를 가지고 있으며, ICT 와 타 산업 간의 융합 연구가 활발하게 진행되고 있는 분야이다. 다양한 형태와 기능의 스마트 워치와 헬스케어를 위한 손목 밴드가 출시되고 있다. 이 제품들은 단순한 IT 기기가 아니라 패션과 IT 가 융합된 새로운 시장을 형성하고 있으며, 대부분의 제품들은 제스처 인터페이스를 제공하고 있다. 하지만 사용자가 의도하지 않은 일상생활에서의 동작이 제스처 명령으로 인식되거나 사용자가 의도하는 제스처 명령을 인식하지 못해서 시스템이 오동작하는 경우가 자주 발생한다. 시스템이 안정적으로 동작하고 보다 더 나은 사용자 경험 (User Experience)을 제공하기 위해서는 일상생활에서 발생하는 제스처 노이즈와 사용자가 의도한 제스처 명령을 시스템이 정확하게 인지할 수 있도록 하고, 사용자의 입장에서도 쉽게 구별하고 사용할 수 있도록 하는 방법이 필요하다.

제스처는 가속도 센서, 자이로 센서 등과 같은 관성센서로 그 궤적을 인식할 수 있으나 일반적인 행동인지 디바이스와 의 인터페이스를 위한 의도한 제스처 인지를 구분하기 어렵다. 사용자의 연속된 동작에서 의도한 제스처를 구별하는 기술을 제스처 세그멘테이션 (gesture segmentation)이라고 한다. 제스처 세그멘테이션에 관한 선행 연구로는 카메라가 장착된 시스템에서 사용자의 동작을 촬영하여 영상처리 기술로 사용자의 제스처를 구분하는 연구 (Kahol et al., 2003; Alon et al., 2009; Kulkarni and Lokhande, 2010)와 신체에 여러 개의 관성센서를 부착하고 사용자의 동작으로부터 제스처를 구분하는 연구(Junker et al., 2008)가 있다. 이들 연구는 웨어러블 시스템과 같이 적은 용량의 메모리, 낮은 컴퓨팅 파워를 제공하는 임베디드 시스템에 적용하기에는 한계가 있다. 본 논문에서는 손목 밴드와 같은 단순한 웨어러블 디바이스를 이용하여 핸드 제스처 명령을 구별하기 위한 방법을 제안한다.

2. Gesture Segmentation Definition

본 연구에서는 사용자가 명령 제스처를 취하려는 의도를 시스템에 전달하는 세그맨테이션 제스처로 주먹을 쥐는 동작(Start of gesture, SOG)과 펴는 동작(End of gesture, EOG)을 제안한다. 세그멘테이션 제스처는 시스템에서도 인식하기 쉽고 사용자도 쉽게 사용할 수 있는 동작이어야 한다. 주먹을 쥐고 펴는 동작은 사용자 입장에서는 쉽고 직관적인 동작이라 할 수 있다. 이렇게 정의된 세그멘테이션 제스처를 기반으로 기기를 제어하기 위한 하나의 제스처 명령은 (1) 주먹을 쥐고, (2) 명령 제스처, (3) 주먹을 펴고, 이렇게 연속 세 동작으로 정의한다. 하나의 제스처 명령 구성을 Figure 1 에 나타내었다. 이러한 구성으로 세그멘테이션 제스처, SOG, EOG 사이에 원하는 명령 제스처를 배치하여 제스처 명령을 만드는 것이 가능하다.

예를 들어 '왼쪽'이라는 제스처 명령은 (1) 주먹을 쥐는 동작, (2) 주먹을 쥔 채 팔을 오른쪽에서 왼쪽으로 움직이는 동작, (3) 주먹을 펴는 동작으로 정의 할 수 있다.

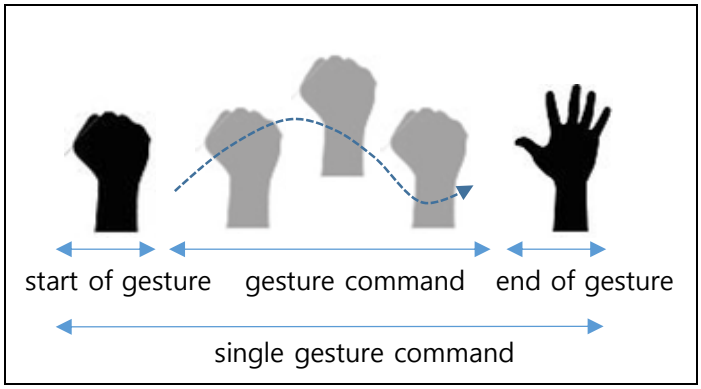

Figure 1. Configuration of a gesture command

3. Prototype of Wrist Worn Wearable Device

제스처 세그맨테이션을 인식하는 알고리즘을 구현하고 시험하기 위해 손목에 착용하는 밴드 형태의 제스처 인식 장치를 개발하였다. 본 연구에서 개발한 인식 장치는 Texas Instrument 사의 16bit MCU 인 MSP430F5436A 를 사용하여 아주 단순한 알고리즘으로 동작하는 저연산, 저전력 구조로 개발하였다.

Figure 2 와 같이 손목에 착용할 수 있는 스트랩에 적외선(Infrared, IR) 발광 및 수광 센서를 배열로 장착하였다. 동작 원리는 다음과 같다. 사용자가 주먹을 쥐고 펴는 동작을 취할 때 손가락과 연결된 힘줄 및 근육변화에 따라 손목 아래 부분의 모양에 변화가 생기게 되고, 이로 인하여 적외선 수광부에 수신되는 광량이 변하게 된다. 이러한 광량의 변화를 인지하여 주먹을 쥐고 펴는 동작을 인식한다.

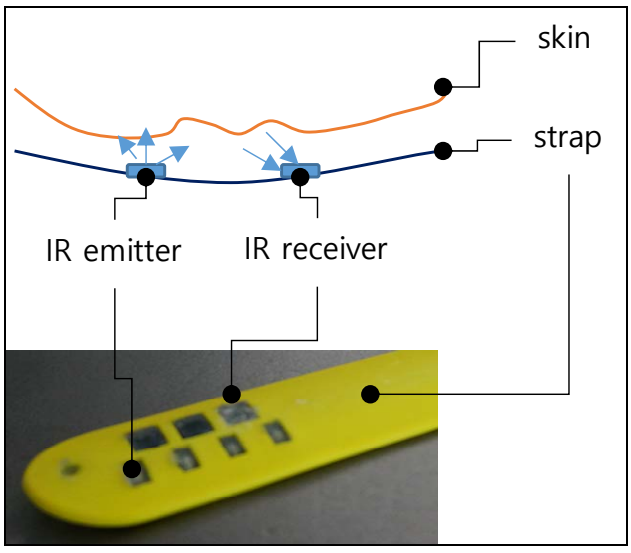

Figure 2. The architecture of the system and wrist strap

Figure 3 에 사용자가 동작을 취할 때 변화하는 손목 모양에 따라 수광부에 수집된 반사광 양의 변화를 그래프로 나타내었다. x 축은 시간을 y 축은 수신된 적외선(IR)의 광량을 나타낸다. 각각의 선들은 적외선 수광 센서에서 수집된 데이터를 나타내며, 그래프의 위쪽에 손의 상태를 표시하였다.

스트랩에 3 개의 적외선 수광 센서가 장착되어 있기 때문에 3 채널의 데이터를 얻을 수 있으며, 주먹을 쥘 때 스트랩과 손목 아래 부분의 피부 사이에 틈이 생겨서 수광부에 전달되는 빛의 양이 증가하기 때문에 수신되는 광량도 많아져 그래프가 올라가고, 주먹을 펼 때는 그 반대로 빛의 양이 줄어들어 그래프가 내려가는 것을 볼 수 있다.

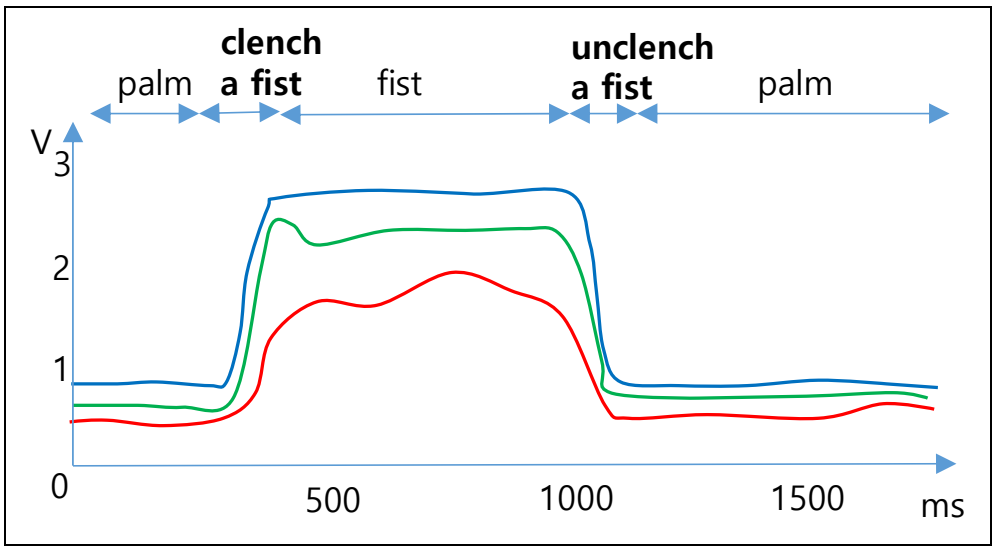

Figure 3. Output of the sensor

본 연구에서 구현한 제스처 인식기 프로토타입을 Figure 4 에 나타내었다. 그림에서와 같이 손을 편 상태에서는 LED 하나만 켜지고, 주먹을 쥔 상태가 인식 되면 LED 가 원형으로 모두 켜지도록 해서 실험 시에 직관적으로 결과를 알아보기 쉽게 구현했다.

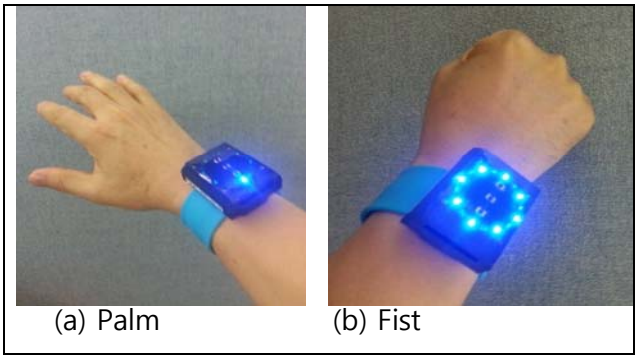
(a) Palm (b) Fist

Figure 4. Implementations of the gesture segmentation system

4. Algorithm

알고리즘의 주 목적은 사용자의 손의 상태가 주먹을 쥔 상태인지 편 상태인지를 인식하는 것이다. 이러한 목적을 달성하기 위해서는 스트랩의 센서 데이터들의 특성으로 볼 때 두 상태(state)를 판단할 threshold 를 결정하는 문제로 귀결된다. 본 논문에서는 상태를 인식하기 위하여 Otsu's threshold(Otsu, 1975) 알고리즘을 적용하였다. 알고리즘 실행 전에 각 채널 별 state 에 따른 적외선 데이터를 수집하고, 각 채널의 클래스 내 데이터의 거리(intra-class distance)를 최소로 하면서 동시에 클래스간 거리(inter-class distance)를 최대로 할 수 있는 최적의 threshold 값을 계산한다. 그런 상태에서 threshold 보다 큰 데이터가 들어오면 주먹을 쥔 상태로 분류하고 그렇지 않으면 손을 편 상태로 분류한다.

하지만 이 알고리즘을 실제로 적용하기에는 2 가지 문제점이 있다. 첫 번째 문제는 손목에 착용한 스트랩이 헐거운 경우에 사용자가 제스처를 취할 때 마다 센서의 위치가 달라 데이터가 정확하 측정되지 못하고 노이즈가 발생하는 경우이다. 두 번째 문제는 오버피팅(overfitting)의 경우이다. 즉, 사용자 마다 다른 특성이 존재하기 때문에 사전에 수집된 데이터를 가지고 threshold 값을 설정하는 Otsu's threshold 알고리즘을 서로 다른 사용자에게 적용할 수가 없다.

이러한 문제를 해결하기 위해서 사전에 수집된 데이터 기반으로 결정된 threshold 를 사용하지 않고 시간에 따라 입력된 신호로부터 threshold 를 결정하는 adaptive threshold 알고리즘을 적용하였으며, 사용자 맞춤 적용을 위해 켈리브레리션 과정을 두었다.

시스템에 저장된 default threshold 값이 없거나, 인식결과가 일치하지 않는 경우 켈리브레이션을 수행한다. 켈리브레이션은 초기 threshold 값을 설정하는 단계로 사용자마다 수광센서에 수신되는 빛의 양이 다르기 때문에 필요한 단계이다. 사용자에게 주먹을 쥐고 펴는 것을 3 회 이상 반복하게 하여 주먹을 쥔 상태와 편 상태의 데이터를 각각 수 초간 수집한 뒤, 각 클래스의 평균 값을 초기 threshold 값으로 설정한다. 여기서 threshold 는 세 개의 채널 별로 수행하며, 결과적으로 세 개의 초기 threshold 를 계산한다. Adaptive threshold 적용 과정은 다음과 같다.

1) 초기 threshold 값으로 손을 쥔 상태와 편 상태에서 측정된 IR 의 양을 채널 별로 각각 저장한다.
2) 각 채널의 저장된 데이터로부터 Otsu's threshold 알고리즘을 적용하여 inter-class distance 와 inner-class distance 를 이용하여 각 채널 별 threshold 를 계산한다.
3) inter-class distance 와 inner-class distance 를 이용하여 각 채널 센서 신호의 신뢰도(credibility)로 정의하고, 이 값을 이용하여 각 채널 별 weighting factor 를 결정한다.
4) Weighting factor 를 이용하여 각 채널 별 데이터와 threshold 를 weighted sum 하여 하나의 통합된 데이터와 threshold 를 계산한다. 이러한 방법으로 신호가 좋지 않은 채널의 weight 값을 낮추어 배제시킴으로써 인식성능을 개선할 수 있다.
5) 통합 데이터와 threshold 로부터 state 를 결정한다.

6) 새로운 데이터가 들어오면 2 개의 state 에 대한 데이터와 threshold 를 갱신한다.

본 연구에서 개발한 알고리즘을 손목형 제스처 인식기에 탑재하였으며 12Mhz 의 메인 클럭과 6KB 의 메모리 환경에서 동작함을 확인하였다.

5. Method

본 연구에서 구현한 손목형 장치의 제스처 세그멘테이션 인식 성능을 평가하기 위해서 12 명의 피시험자를 대상으로 사용자 시험을 수행하였다. 남자 10 명, 여자 2 명이 시험에 참가하였으며, 평균나이는 38 세이다. 사람들의 손목 둘레가 다르기 때문에 스트랩에 센서 위치를 5mm 간격으로 서로 다른 6 개의 스트랩을 준비하였다. 시험 진행자는 시험에 앞서 피시험자들의 손목 둘레를 측정하여, 6 개의 스트랩 중 피시험자의 손목 아래 부분에 센서가 가장 잘 위치 하는 것을 선택해서 시험하였다. 실제 시험을 시작하기 전에 피시험자에게 3 분간 세그멘테이션 제스처를 연습 할 수 있도록 하였다.

본 실험에서 진행자는 피시험자에게 세그멘테이션 제스처를 포함한 임의의 제스처를 20 회 수행하도록 하였다. SOG 와 EOG 의 성공 실패 여부만을 카운트하여 기록하였으며, SOG 와 EOG 사이의 제스처 명령은 무시 하였다.

4. Results

Table 1 에 시험 결과를 나타내었다. SOG, EOG 각각의 인식률은 99.58%, 100%로 측정되었다. 위치 변경 횟수는 3 분의 연습과정에서 인식이 잘 되지 않아 피시험자가 직접 장치의 착용 위치를 조정한 횟수이다.

Table 1. The result of user test for gesture segmentation

	SOG Number of successes	EOG Number of successes	Wrist circumference (cm)	Number of location change	Wear location (Left/Right hand)	Age	Sex
Tester 1	20	20	18	0	R	39	M
Tester 2	20	20	15.9	1	L	30	M
Tester 3	20	20	17.4	0	R	44	M
Tester 4	20	20	16.9	0	L	42	M
Tester 5	20	20	19	2	R	36	M
Tester 6	20	20	17	0	R	42	M
Tester 7	20	20	17	0	R	41	M

Tester 8	20	20	17	0	L	42	M
Tester 9	20	20	17	0	R	26	M
Tester 10	19	20	14.2	0	R	38	F
Tester 11	20	20	16	2	R	39	F
Tester 12	20	20	17.9	1	R	45	M
Recognition rate	99.58	100.00	-	-	-	38.67 Average	-

5. Discussion

본 연구에서 제작한 손목 밴드형 웨어러블 제스처 장치의 개발단계에서 인식률을 측정했을 때 손목 아래 부분에 센서가 정확하게 위치하는 경우에 인식률이 높게 측정되었다. 이는 손목 아래 부분이 손을 쥐고 펼 때 모양 변화가 가장 크기 때문이다. 이러한 문제를 극복하기 위해서 본 장치에서는 3 개의 수광 센서를 배치하여 약간의 위치 변화에도 좋은 결과를 얻을 수 있도록 하였고, 인식 알고리즘에서도 신호가 양호한 채널에 가중치를 더 주도록 하여 극복하고자 노력하였다. 하지만 사람마다 손목의 굵기가 다르기 때문에 개발 단계에서도 개발자 간에 인식률 차이가 크게 나타났다. 따라서, 사용자 시험에서는 이런 문제를 해결하기 위해 길이에 따라 센서 위치가 다른 스트랩 6 개를 준비하여 개인차를 최대한 줄였다. 상용화를 위해서는 시계줄 길이를 조절하는 것과 같이 스트랩에서 센서의 위치를 조정할 수 있는 기구 개발이 필요하다.

또 다른 경우로는 일상생활에서도 주먹 쥐고 펴는 동작을 취할 수 있기 때문에 이러한 경우 오작동 가능성이 존재한다. 이러한 문제를 해결하기 위해서는 본 연구에서 제공하는 주먹을 쥐는 시간과 횟수를 체크 기능을 활용하여 일상생활에서 자주 발생하지 않는 동작 즉, 주먹을 빠르게 두 번 쥐거나 길게 쥐는 등의 동작으로 대체할 수 있을 것이다.

본 연구에서는 제스처 세그멘테이션 방법으로 주먹을 쥐고 펴는 동작이 얼마나 가능성이 있는지 그 인식률을 우선 측정해 보았다. 추후 연구에서는 일반적인 인간공학 연구에서 사용하는 실험 및 분석 방법을 적용하여 이러한 제스처 세그멘테이션 방법이 일반 사용자들에게 보편적으로 사용될 수 있는지에 대한 종합적인 연구를 수행할 계획이다.

6. Conclusion

본 논문에서는 스마트워치 또는 손목 밴드에 적용 가능한 핸드 제스처 세그멘테이션 방법을 제안하였다. 스트랩에 적외선 센서 발신부와 수신부를 배열한 손목 밴드형 프로토타입 시스템을 구현하고, 사용자 시험을 통하여 제안 기술의 적용가능성을 평가하였다. 시험 결과, 하드웨어 및 소프트웨어의 연산량을 고려해 볼 때 충분히 스마트워치 및 스마트 밴드와 같은 웨어러블 시스템에 적용 가능한 방법이라 할 수

있다. 향후 개인차를 극복할 수 있는 연구를 진행하여 보다 더 안정적으로 많은 사용자들이 사용할 수 있는 시스템을 개발할 계획이다. 이러한 첨단 기능을 전통적인 패션 액세서리인 시계, 팔찌, 손목 밴드 등에 구현하는 스마트 어패럴과 같은 패션 ICT 융합 제품들이 전통 패션 제품의 고부가가치화와 이를 통한 신시장 창출에 기여할 것으로 기대한다.

Acknowledgements

This work was supported by the ICT R&D program of MSIP/IITP. [B0101-15-0420, Development of Transformational and Slap-on Wearable Device and UI/UX].

References

Alon, J., Athitsos, V., Yuan, Q. and Sclaroff, S., A unified framework for gesture recognition and spatiotemporal gesture segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9), 1685-1699, 2009.

Junker, H., Amft, O., Lukowicz, P. and Tröster, G., Gesture spotting with body-worn inertial sensors to detect user activities, *Pattern Recognition*, 41(6), 2010-2024, 2008.

Kahol, K., Tripathi, P., Panchanathan, S. and Rikakis, T., Gesture segmentation in complex motion sequences, *Proceedings of the International Conference on Image Processing*, 105–108, 2003.

Kulkarni, V.S. and Lokhande, S.D., Appearance based recognition of american sign language using gesture segmentation, *International Journal on Computer Science and Engineering*, 2(3), 560-565, 2010.

Otsu, N., A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66, 1979.