

HD EMG 부착 시 따르는 position과 impedance 변화의 보정방법 연구

개별연구생 이해민
소 속 전기및전자공학부
학 번 20190533
이메일 byj4565@kaist.ac.kr

담당 교수님 박형순
사수님 조성현

● 주제 설명

HD EMG 특성 상 매번 붙이고 떼는 과정을 반복해야하고 이 때마다 항상 같은 위치와 상태로 붙이는 것은 불가능하다. 즉, 매번 HD EMG의 위치와 impedance, 신호의 세기 또는 패턴 등이 같을 수 없기 때문에 이를 보정해줄 필요가 있다.

따라서 다양한 환경에서의 data를 먼저 수집한 후에, 해당 data를 바탕으로 당일의 EMG는 어느 위치에 부착된 것으로 추정되며 따라서 그날 입력되는 신호들은 어떻게 해석하는 것이 적합한지를 판단하고 이에 맞게 손가락이 움직일 수 있도록 해주는 알고리즘을 연구할 예정이다.

전체 일정	1주차	관련 논문 찾아보고 연구 분야에 대한 지식 쌓기
	2주차	실험 프로토콜 작성하기 (이후에 상황에 맞게 조정 가능)
	3주차	TMSI 센서 사용법을 익힌 후 비교 대상이 될 기본 data 많이 수집하기
	4주차	데이터를 관찰하여 보정방법 개발하기
	5주차	개발한 보정방법으로 결과 도출하기
	6주차	지속적으로 보완하여 결과 개선하기
	7주차	전체 결과 정리 및 보고서 작성

1. 개별연구 일일 탐구일지

작성일	0주차	2020.07.10.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 탐구일지 양식 만들기 <input type="checkbox"/> 논문 2개 읽어보기 <input checked="" type="checkbox"/> 뇌졸중 환자분 글러브 실험 보조하기			

가. Recommended keyword

- high density EMG
- surface EMG
- Muscle-Computer Interface
- Gesture recognition
- identification
- stroke
- classification

나. Reference 1 (사수님 추천)

Title : [Advancing Muscle-Computer Interfaces with High-Density Electromyography](#)

HD EMG가 미세한 손 제스처를 식별할 수 있다는 가능성을 확인했다.

다. Reference 2

Title : [Making Muscle-Computer Interfaces More Practical](#)

여기서는 HD EMG가 아니라 muscle-sensing armband를 이용했다. 따라서 각도에 따른 placement는 신경 쓸 필요가 없고, 앞뒤의 placement만 신경 쓰면 된다. 해당 논문의 주요 내용은 data를 낸 후에 바로 측정했을 때(within-session), re-donning을 한 후에 측정했을 때, 다른 날에 측정했을 때의 accuracy를 비교했다. 확실히 후자로 갈수록 accuracy가 감소했다. 해당 논문에서는 화면을 누르는 경우와 finger snap을 하는 두 가지 경우에 대해 실험했는데, finger snap은 다른 날에 측정했을 때에도 8명 중 5명의 accuracy가 95%를 넘었기에 muscle-sensing armband가 실생활에서 쉽게 탈부착되면서 이용할 수 있을 것이라는 가능성을 제시했다. 하지만 화면을 누르는 실험은 finger snap에 비해 정확도가 현저히 떨어졌는데, 이는 누르는 손가락을 제외한 다른 손가락이나 팔의 움직임과 위치는 각자 자유롭기 때문에 그 자유로운 부위를 어떻게 하고 있는냐에 따라서 사람들 사이에 차이점이 발생하므로 오차가 컸을 것이라 예측한다. 내가 실험할 당시에라도 최대한 배경 요소를 통일해주는 것이 중요할 것 같다.

라. 뇌졸중 환자분 글러브 실험 보조하기

환자분께서 원하시는 대로 손을 움직이지 못하시고 말씀도 어눌하신 것 같다. 학생들에게 사탕을 주시는 것을 좋아하셔서 가방에 사탕을 많이 넣고 다니신다. 1시간 반의 실험 동안 자신은 손에 힘을 주는 경우가 없었고 예찬이 오빠가 컴퓨터로 글러브의 모터를 조절해서 원하는 position을 만들었다.

글러브가 손에만 있기 때문에 손목을 잡아주지는 못했다. 그런데 환자분은 손목에 힘을 주지 못하는 것 같았고 손목에 계속 무리가 가는지 힘들어하셨다. 가능하다면 손가락을 모터로 움직이는 것처럼 손목도 조종하거나 이것이 힘들면 받쳐주는 것이라도 해주면 좋을 것 같다. 이에 대해 사수님께 여쭙보았더니 목표는 실제로 손목뿐만 아니라 팔꿈치, 어깨까지도 control 가능하게 하는 것이라고 하셨다!

마. Reference 3 (사수님 추천)

Title : [High-Density Myoelectric Pattern Recognition Toward Improved Stroke Rehabilitation](#)

2. 개별연구 일일 탐구일지

작성일	1주차	2020.07.13.	작성자	이혜민
일일 목표	<input type="checkbox"/> Reference 1 마저 읽고 주요 부분 정리하기 <input type="checkbox"/> Reference 3 읽고 주요 부분 정리하기			

가. Reference 1 (사수님 추천)

Title : [Advancing Muscle-Computer Interfaces with High-Density Electromyography](#)

1. Abstract

Finger gesture를 EMG를 이용하여 구별하는 것을 목표로 한다. 192개의 electrode를 가지고 있는 HD EMG를 이용했고 안쪽 팔뚝(upper forearm)에 부착했다.

흠??

Within-session scenario에서 평균 90%의 정확도를 보였고 이는 감지하기 힘든 많은 gesture를 감지할 수 있을 것이라는 가능성을 보여준다.

많은 수의 Electrode가 있을수록 performance가 좋아짐을 보였다.

Cross-session 감지는 electrode의 position이 session에서 session으로 바뀌는 것에 영향을 받았다.

Session들 사이에서 electrode가 shift되는 것을 감지하는 2가지 방법을 제시할 것이다. 이는 조금의 calibration data를 이용하여 기준이 되는 baseline system에 비교함으로써 진행된다. Shift compensation을 해주고 나니 59%에서 75%로 accuracy가 증가했다.

-> 해당 부분에서 어떻게 shift compensation을 부여했는지 참고해서 나도 그렇게 해보자.

2. Introduction

Gesture로 computer에 input할 수 있다면 smart glasses나 watch를 조종하는 등 다양한 분야에 응용될 수 있을 것이다. 손에 바로 input device를 부착한다면 일상생활이 불편할 것이므로 간접적으로 정보를 얻는 방안이 많이 연구되었다. body-worn cameras, wrist-worn depth cameras, tendon의 움직임 감지하는 방법, EMG 등이 그 예시다. Finger에 관여하는 대부분의 muscles가 forearm에 위치하므로 그곳에 EMG를 부착하고 힘줄의 변화까지 종합하여 data를 얻어낼 수 있을 것이다.

3. Related Work

Array-like electrode는 단 몇 개의 electrode를 부착하는 것에 비해서 하나하나의 electrode를 정확한 position에 부착할 필요가 덜하다. 이는 re-donning을 자주 해야 하는 wearable device에게 득이 되는 부분이다. EMG를 이용해서 gesture를 discriminate한 previous research를 소개한다.

4. Background

Thumb, pinkie, index finger에 관여하는 근육은 distal에 존재한다. 하지만 thumb를 제외한 나머지 손가락에 관여하는 주요 근육은 proximal에 존재하므로 EMG를 forearm에 부착하면 thumb를 제외한 손가락들의 activity를 capture할 수 있을 것이다.

5. Methodology

6. Experimental setup

8*24의 EMG 사용. Cable의 길이로 인한 noise를 줄이기 위해서 pre-amplifier를 모든 cable에 부착했다. **bipolar recording을 이용하기 때문에 각 column에서 하나의 electrode는 meaningless한 data를 나타내고, 따라서 192개 중 168개의 channel만 usable data를 제공한다.**

실험 결과 168 channel까지는 필요없고 20에서 80개의 channel이 이상적이다. -> 오 왜지? 그러면 해

당 20에서 80개의 channel은 어떻게 부착할 때 효과적인가? 뒤에 설명되어있을 실험을 자세히 보자.
2048Hz의 속력으로 기록됐고 amplifier의 gain은 1000이었다.

가) Gesture set

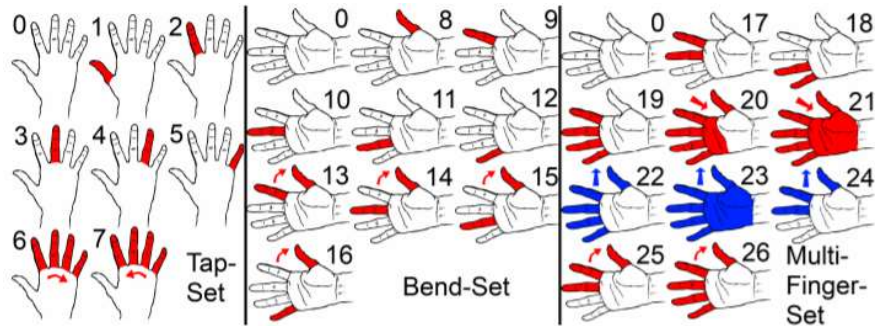


Figure 4. Iconic illustrations of the performed gestures. The gestures were divided into three sets: the tap-set, the bend-set and the multi-finger-set. The idle gesture (0) is contained in all three sets.

- 각 set of gesture에 대해 idle gesture도 알맞게 설정해야 한다.

7. Classification Pipeline

Standard classification scheme 중 Bayes classifier를 이용해서 27가지의 classes를 구분했다.

Electromyogram의 RMS는 muscle에서 생성되는 force와 연관이 커서 muscle의 activity를 측정하는데 일반적으로 사용된다.

Baseline normalization은 측정값에서 idle gesture의 평균 activity를 모두 빼주는 것이다. 이는 RMS value를 구한 후에 진행한다.

가) Signal Preprocessing

Powerline과 cable의 movement로 인한 noise는 둘째 치더라도 electrode가 직접적으로 부착되어 있는 skin의 movement는 치명적이다. 그런데 분석결과 이로 인한 artifact는 high amplitude, low frequency한 성질을 가지고 있다고 한다. 이와 함께 dc offset과 high-frequency noise의 영향도 줄이기 위해서 [8]과 [13]의 연구결과를 바탕으로 20-400Hz의 pass-band를 가지고 있는 fourth order butterworth band-pass filter를 이용했다.

나) Segmentation, Feature Extraction, Naive Bayes classifier

정확히 이해되지는 않지만, window를 나눠서 분석하고 그로써 해당 gesture에서 active된 muscle을 판별해내는 것 같다.

8. Estimation of Electrode Displacement

가) Estimation of ulna position

Ulna가 있는 쪽은 특별히 muscle activity가 low하다는 특징을 이용해서 이곳을 기준으로 한다. Y축과 ulna가 평행하다는 가정하에 x축에 대한 shift만 고려한다.

나) Estimation of center of main muscle activity

Main flexor와 extensor muscles가 있는 쪽은 특별히 muscle activity가 high하다는 특징을 이용해서 이곳을 기준으로 한다.

9. Experiment Procedure

팔꿈치에서 2-4cm 떨어진 곳에 부착됐다. 팔뚝의 바깥쪽에는 손가락에 관여하는 muscle이 없기에 바깥쪽에는 붙이지 않았다. 5명의 subject를 대상으로 5번의 session을 진행했는데, 각 session은 다른 날에

진행됐기에 조금씩 다른 위치에 부착됐을 것이며 이는 더 현실적인 실험을 위해서라고 언급되어 있다. -> 이에 대해 어떻게 보완할 것인지 discussion에 설명되어있는지 확인해보자.

3초 동안 행동을 하도록 지시되었으며 simulation interface를 어떻게 구성했는지 자세히 설명하고 있다. 갑작스러운 손의 변화는 몇몇 electrode를 떨어지게 만들 수 있고 이는 다시 부착시키는 데에 불편함이 있고 정확도를 떨어뜨리기 때문에 gesture 순서를 일정하게 유지했다고 한다.

결과에서 extensor는 왼쪽에서 flexor는 오른쪽에 나타난다.

근데 5번의 session을 한거는 알겠는데, 이거를 판단하기 위한 근거 자료로 사용한 것 아닌가? 아니면 조금은 때서 잘 classification하는지 performance test용으로 사용하는건가? -> 앞의 4개의 session은 train에, 마지막 session은 test에 이용했다.

10. Results and Analysis

가) Withiin-session classification

실험 결과를 보면서 비교하고 있다. 그리고 오차가 특별히 생기는 부분은 왜 오차가 생겼는지 분석하고 있다. 대부분 정확도가 아주 높다.

자신들은 mutual information을 최대한 많이 얻기 위해서 최대한 가능한 수의 electrode를 사용했으며 추정 당시 kernel density estimation을 이용했다. 참고하는 electrode의 수를 줄여가면서 다시 분석해본 결과 위에서 말한 20-80개가 가장 이상적이라는 결과가 나왔다. 그런데 이때 어떤 규칙으로 electrode를 선택했는지 나와있지 않다. 추측하건데 적은 수의 electroide를 선택하더라도 전체적으로 퍼져있도록 선택했을 것 같다. 따라서 이는 나의 상황처럼 팔뚝 전체가 아닌 일부분만 참고해야되는 상황에는 또 다르게 작용할 것으로 추측된다.

다음	<input type="checkbox"/> Reference 1 마무리하기
목표	<input type="checkbox"/> Reference 3 읽기

3. 개별연구 일일 탐구일지

작성일	1주차	2020.07.14.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Reference 1 마무리 <input type="checkbox"/> Reference 1 복습 및 전체 요약 재작성하기			

가. 사수님과 회의

- Reference 2에서 finger snap처럼 힘이 많이 들어가는 동작은 re-donning을 하더라도 오차율이 현저히 적었다. 따라서 나도 힘이 많이 들어가는 동작인 '주먹 쥐기'를 대상으로 실험하기로 했다.
- Calibration method는 reference 1의 방법을 일단 시도해본다. calibration을 했을 때와 하지 않았을 때 주먹 쥐기를 했는지 안 했는지 판단하고 이의 accuracy를 비교하는 것이 최종 목표이다.
- Python의 Matplotlib로 reference 1의 data를 plot 해보자.

나. Reference 1

1. Results and Analysis

가) Shift compensation

각 subject에 대해서 shift compensation을 진행한다. 이때 leave-one-out cross-validation을 이용하였는데, 이는 4개의 session으로 train을 한 다음에 해당 data를 바탕으로 나머지 하나의 session을 evaluation하되, calibration도 함께 적용한다.

Calibration gestures도 classified 되어 있기 때문에 accuracy가 증가되는 역할로 작용할 수 있다. -> 왜??

Calibration을 하지 않은 baseline system에 비해 shift compensation을 한 경우에는 accuracy가 60%에서 75%로 올랐다. Within-session의 결과와는 현저히 떨어지는 결과이지만 이는 다른 날에 행해졌기 때문에 어쩔 수 없다. 이를 더 보완하고 싶다면 더 많은 calibration 과정이 필요할 것이다. 신기하게도 11번 동작은 calibration을 했을 때 accuracy가 1 증가했다.

(1) Calibration gesture

어떤 gesture가 가장 shift를 판단하기에 좋은지 연구했다. Ulna method의 경우는 손을 짹 펴는 동작인 22번과 23번이 classification accuracy가 가장 높았다. GMM method는 11, 14, 23, 24가 전체적으로 accuracy가 높았고 그 중 11이 가장 높았다.

2. Restricted gesture sets

Gestures	WS	CS
11,17,21,23,24	98.8	94.6
1,2,3,11,17,18,21,23,24,25	97.0	92.3
1,2,3,7,9,10,11,13,14,17,18,21,23,24,25	94.2	83.9

Table 5. Results for the restricted gesture sets given as mean accuracy for the within-session (WS) and cross-session (CS) case.

Gesture의 종류를 제한해서 그 안에서만 classification을 해보니 accuracy가 크게 증가했다. 각 gesture set에서 2개는 calibration을 적용시켰다. 이는 어차피 실제 HCI에서 26가지의 gesture를 모두 이용하긴 어려울 것이므로 현실적인 방안이다.

3. Conclusion and Future work

전체 요약, 미래에 더 발전된 EMG가 나오길 바란다.

4. 개별연구 일일 탐구일지

작성일	1주차	2020.07.15.	작성자	이혜민
일일 목표	<input type="checkbox"/> Reference 1 전체 요약 완료하기 <input type="checkbox"/> Reference 3 간단하게 요약 <input type="checkbox"/> Reference 1의 Data가 어떤 형식인지 확인하고 plot 해보기			

- 팔의 360도 data를 모두 구해놓은 다음에 current data는 어디에 locate하는지 찾는 방향보다는, 360도 중에서 어느 부분에 붙이는 것이 판별하기 쉬운지를 먼저 찾아내고, 최대한 그 부분에 붙이려고 노력하는 방향으로 가는 것이 좋을 것 같다. 최대한 그 부분에 붙이려고 해도 오차가 생길 것이니 그에 대한 calibration을 연구한다면 ref 1에서의 calibration과 방향이 일치한다.
- 만약 전체 360도 중에 현재 어디일까?를 판별하는 문제라면 ref 1에서 ulna와 main flexor/extensor를 참고한다는 아이디어만 가지고 오고, calibrate라기 보다는 search의 관점에서 연구해야한다. -> 둘 중 어느 방향으로 할지 논문을 다 읽고 정해보자.

5. 개별연구 일일 탐구일지

작성일	1주차	2020.07.16.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Reference 1 전체 요약 완료하기 <input checked="" type="checkbox"/> Reference 3 간단하게 요약			

- 완성된 번역/요약본은 notion에 있음

6. 개별연구 일일 탐구일지

작성일	1주차	2020.07.17.	작성자	이혜민
일일 목표	<input type="checkbox"/> Classifier 구현할 수 있는 library 구글링해서 이해하기			

- 사수님과의 회의
 - LDA 또는 SVM classifier 이용 추천
 - python “classifier 이름” 치면 library와 함께 다양한 예제 나오니까 그 예제를 이해해가면서 ref1의 data 입력해보기
 - ref 1 data가 너무 많으니까 1GB 안으로 추려서 이용하기
 - .mat file을 python에서 불러오는 것도 library 있으니까 구글링 해보기
 - 전체적인 흐름
 1. data load
 2. feature extraction
 3. numpy or array로 해당 feature를 X에 넣기
 4. Y는 label
 5. X와 Y를 classifier에 대입하기
 6. accuracy와 confusion matrix 결과 도출
 - classifier를 적용할 때 내가 조절할 수 있는 parameter들이 다양하게 있는데 그것들을 최대한 *일반적*

이제 조절해서 accuracy가 높게 나오는 parameter를 찾아야 한다.

- 우리의 목표는 보정을 하는 것이지만, 이 과정을 먼저 해야지 보정 결과를 비교할 수 있다

● Investigating Ref 1 data set

Form	Direc		Direc		.mat		array		array		array		array
Type	Subject		session		gesture		try		time domain		channel		value of sample
Amo	1	>	1	>	0	>	0	>	0	>	0	>	0
unt
	5		5		26		9/29				191		6143

● Useful bookmarks

- [How to open .mat from python](#)
- [How to open multiple .mat file in python](#)
- [scikit classification examples](#)
- [Butterworth Band-pass filter by python](#)
- [scipy.org : manual for scipy lib](#)
- [Digital signal Analysis](#)
- [LDA in python 1](#)
- [LDA in python 2](#)

7. 개별연구 일일 탐구일지

작성일	2주차	2020.07.20.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Git 사용법 익히기 <input checked="" type="checkbox"/> Data 불러와서 github에 업로드 하기			

● Useful bookmarks : Git

- [Git 매뉴얼 블로그](#)
- [Git book](#)

● Butterworth band-pass filter (20-400Hz)

- $T = 0.000488 \text{ sec/sample}$
- $f_s(\text{sampling rate}) = 1/T = 2048 \text{ samples/sec}$
- Nyquist frequency = $f_s/2$

다음 목표	<input type="checkbox"/> Band-pass filter 구현 완료하기 <input type="checkbox"/> feature extraction
----------	--

8. 개별연구 일일 탐구일지

작성일	2주차	2020.07.21.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Band-pass filter 구현 완료하기 <input type="checkbox"/> segmentation			

● Useful bookmarks : Git

- [Vscode git : Pull request](#)

● Structure

- windows_per_channel
 dimension : $40 * 150$
 type : np.ndarray
 row : window
 column : samples for each window

- Investigating Ref 1 data set : **After segmentaion (2)** : Computing RMS for each window
 RMS value로 구성된 새로운 ndarray를 만들자.

9. 개별연구 일일 탐구일지

작성일	2주차	2020.07.22.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> EMG, EEG 환자분 실험 보조하기 <input type="checkbox"/> segmentation 완료			

● Structure

- RMSwindows_per_channel

dimension : 40 * 1

type : np.array

row : RMS value of a window of that channel

column : *

- RMS_one_try

dimension : # of windows(40) * # of meaningful channels(168)

type : np.array

row : windows

column : channels

Type	gestures		tries		windows		window		RMS value of that channel in that window
name	RMS_gestures	>	RMS_tries_for_gesture	>	RMS_one_try	>	RMSwindows_per_channel	>	
len	row : 11		row : 10		row : 40		1-dimension : 168		

● Baseline normalization

논문에서는 각 gesture set 별로 normalization을 해주었지만, 나는 특정 gesture만 선택했기 때문에 2,4,7,8,11,13,19,25,26,30th idle gesture data 전체에 대해 average를 한 번 계산해서 모든 gesture에 적용시키겠다.

10. 개별연구 일일 탐구일지

작성일	2주차	2020.07.23.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 교수님과 미팅 <input checked="" type="checkbox"/> 프로그래밍 타임라인 만들기 <input checked="" type="checkbox"/> 주간 목표 다시 세우기			

전체 일정	1주차	관련 논문 찾아보고 연구 분야에 대한 지식 쌓기
	2주차	Ref 1의 전체 과정 구현해보기
	3주차	Ref 1의 전체 과정 구현해보기
	4주차	Ref 1의 전체 과정 구현해보기 및 TMSI 센서로 data 수집하기
	5주차	우리 data에 ref 1의 calibration 적용해보기
	6주차	우리 data에 ref 1의 calibration 적용해보기 및 보정법 개선안 제시
	7주차	보정법 개선안 제시 및 보고서 작성

Programming Timeline

# Order	category	To-do	progress	Start	Due	Done	Done
1	Signal Preprocessing	Apply butterworth band-pass filter	Done			2주차	Jul 21, 2020
2	Segmentation Data processing	Divide continuous data into 150 samples window	Done			2주차	Jul 21, 2020
3	Segmentation Data processing	Discard useless data : 192ch → 168ch	Done			2주차	Jul 22, 2020
4	Segmentation Data processing	Compute RMS for each channel	Done			2주차	Jul 22, 2020
5	Segmentation Data processing	Perform baseline normalization	Done	Jul 23, 2020	2주차	2주차	Jul 23, 2020
6	Segmentation Data processing	Check whether each window is represented by a 168-dimensional vector of RMS values	Done	Jul 23, 2020	2주차	2주차	Jul 23, 2020
7	Segmentation Data processing	Apply spatial order 3 1-dimensional median filter on the vector to compensate local artifacts			3주차		
8	Segmentation Determine whether ACTIVE	Compute average of the summarized RMS values per window → threshold			3주차		
9	Segmentation Determine whether ACTIVE	If the sum of RMS vector elements of one window is greater than the threshold, it's ACTIVE			3주차		
10	Segmentation Determine whether ACTIVE	If the predecessor and successor is active, it's ACTIVE			3주차		
11	Feature Extraction	compute RMS for each channel on all windows → feature (of each channel)	Not understood		3주차		
12	Feature Extraction	Normalize the mean RMS over all channels	Not understood		3주차		
13	Feature Extraction	Result : 168 * N dimensional feature RMS vector. With RMS is length normalized			3주차		
14	Naive Bayes classifier	Model the feature distribution by kernel density estimation with Gaussian kernel function			3주차		
15	Naive Bayes classifier	Apply naive Bayes classifier for each 27 classes			3주차		

16	Estimation of Electrode Displacement Ulna position	Apply penalty function to favor the region in the middle of the array's x range			4주차
17	Estimation of Electrode Displacement Ulna position	Apply Watershed algorithm in order to find possible paths	Not understood		4주차
18	Estimation of Electrode Displacement Ulna position	Apply Dijkstra's algorithm to choose the lowest cost path			4주차
19	Estimation of Electrode Displacement Center of main muscle activity	Apply Gaussian Mixture Model (GMM)			4주차
20	Estimation of Electrode Displacement Center of main muscle activity	Take mean of two estimation shift			4주차

+ New

COUNT 24

11. 개별연구 일일 탐구일지

작성일	2주차	2020.07.24.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 랩 세미나 참석 <input checked="" type="checkbox"/> Baseline normalization 완료			

12. 개별연구 일일 탐구일지

작성일	3주차	2020.07.28.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Segmentation : data processing 완료하기			

- 원래 내일 마무리하려 했지만, 오늘 분량이 빨리 끝나는 바람에 segmentation까지 완료함
- 하지만 debugging이 필요해 보이기에 이때까지 한 코딩이 모두 제대로 작동하는지 디버깅을 완료하는 것을 내일까지의 목표로 삼자.
- Baseline normalization 전까지 디버깅 완료.

13. 개별연구 일일 탐구일지

작성일	3주차	2020.07.31.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 현재까지의 코드 디버깅 완료하기			

- 사수님께 질문 : “apply a spatial order 3 1-dimensional median filter on the vector”라고 했는데, 전체 array를 median filter에 적용시키면 가장자리에 있는 모든 gesture와 try들의 값이 0이 됩니다. 그러면 각 gesture의 한 try의 channel 별로 median filter를 적용해야 되는 것 아닌가요?

14. 개별연구 일일 탐구일지

작성일	3주차	2020.08.01.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Segmentation 완료 <input type="checkbox"/> Feature extraction 완료			

- 사수님께 질문 : Feature extraction에 적혀있는 앞 두 문장을 내가 잘 해석한 것이라면 segmentation에서 이미 한 활동인데 그게 맞는건가요? 그러면 normalization만 하면 끝인건가요?
- Feature Extraction The found segment is partitioned in N windows of equal length. We compute the RMS as feature for each channel on all windows and normalize the mean RMS over all channels. This results in a 168 dimensional feature vector per window. We concatenate the obtained RMS vectors resulting in a 168·N dimensional feature vector. Since the RMS is length normalized, it is irrelevant if the found activity segments do not have the same length.

15. 개별연구 일일 탐구일지

작성일	3주차	2020.08.02.	작성자	이혜민
일일 목표	<input type="checkbox"/> Feature extraction 완료			

- 사수님께 질문 : RMS도 알고 mean도 알고 normalization도 아는데, normalizing the mean RMS는 뭔가요?
- 내 깃허브 주소
- https://github.com/Hyedora/2020_Summer_Individual_study.git

- <https://github.com/Hyedora>

# Order	category	To-do	progress	Start	Due	Done	Done
1	Signal Preprocessing	Apply butterworth band-pass filter	Done			2주차	Jul 21, 2020
2	Segmentation Data processing	Divide continuous data into 150 samples window	Done			2주차	Jul 21, 2020
3	Segmentation Data processing	Discard useless data : 192ch → 168ch	Done			2주차	Jul 22, 2020
4	Segmentation Data processing	Compute RMS for each channel	Done			2주차	Jul 22, 2020
5	Segmentation Data processing	Perform baseline normalization	Done	Jul 23, 2020	2주차	2주차	Jul 24, 2020
6	Segmentation Data processing	Check whether each window is represented by a 168-dimensional vector of RMS values	Done	Jul 23, 2020	2주차	2주차	Jul 24, 2020
7	Segmentation Data processing	Apply spatial order 3 1-dimensional median filter on the vector to compensate local artifacts	Done	Jul 28, 2020	3주차	3주차	Jul 28, 2020
8	Segmentation Determine whether ACTIVE	Compute average of the summarized RMS values per window → threshold	Done	Jul 28, 2020	3주차	3주차	Jul 28, 2020
9	Segmentation Determine whether ACTIVE	If the sum of RMS vector elements of one window is greater than the threshold, it's ACTIVE	Done	Jul 28, 2020	3주차	3주차	Jul 28, 2020
10	Segmentation Determine whether ACTIVE	If the predecessor and successor is active, it's ACTIVE	Done	Jul 28, 2020	3주차	3주차	Jul 28, 2020
	Debugging	Check whether it's well operating until now	Done	Jul 28, 2020	3주차	3주차	Aug 1, 2020
	Segmentation	Select the longest contiguous sequence of active windows → gesture segment	Done	Jul 30, 2020	3주차	3주차	Aug 1, 2020
11	Feature Extraction	compute RMS for each channel on all windows → feature (of each channel)	Not understood		3주차		
12	Feature Extraction	Normalize the mean RMS over all channels	Not understood		3주차		
13	Feature Extraction	Result : 168 * N dimensional feature RMS vector. With RMS is length normalized			3주차		
14	Naive Bayes classifier	Model the feature distribution by kernel density estimation with Gaussian kernel function			3주차		
15	Naive Bayes classifier	Apply naive Bayes classifier for each 27 classes			3주차		

16. 개별연구 일일 탐구일지

작성일	4주차	2020.08.03.	작성자	이혜민
일일 목표	<input type="checkbox"/> 질문 답변 내용 코드에 적용 <input type="checkbox"/> feature extraction, classifier 구현			

- Data를 processing하는 단계이기에 data 전체를 for문으로 탐색하는 과정이 매우 반복적으로 일어난다. 그래서 이를 조금 더 효율적으로 하고자 처음에 RMS_gesture array를 구현하는 과정에서 base normalization과 median filter 적용까지 함께 할 수 있도록 수정해보는 중이다.

17. 개별연구 일일 탐구일지

작성일	4주차	2020.08.04.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Median filter 수정 <input checked="" type="checkbox"/> feature extraction			

- Active segment 판별까지 완료

```
PS C:\Users\byj45\OneDrive\4 KAIST\20 Summer Individual study\code> & C:/Users/byj45/AppData/Local/Programs/Python/Python37/python.exe "c:/Users/byj45/OneDrive/4 KAIST/20 Summer Individual study/code/Building_ref1.py"
0번째 gesture의 각 try의 segment 길이들 : 6 21 17 6 24 11 9 16 13 5
1번째 gesture의 각 try의 segment 길이들 : 3 12 12 5 7 11 11 18 9 10
2번째 gesture의 각 try의 segment 길이들 : 6 5 7 6 5 6 7 16 6 6
3번째 gesture의 각 try의 segment 길이들 : 6 5 5 5 5 4 4 6 5 5
```

- Mean normalization

https://en.wikipedia.org/wiki/Feature_scaling#Mean_normalization

- Debugging 시작

Active 판별 이후에 [0 2 0], [0 3 0], [0 6 0], [0 7 0], [0 8 0]의 window가 모두 0의 값만 가진다.

-> threshold가 음수여서 active를 판별할 때 0인 것들도 active로 판별됨.

근데 RMS값인데 왜 음수인 값이 생기는거지?

-> base normalization

=> idle gesture에서만 나오는 예외인거니까, 그냥 전부 0인 window는 임의로 active에서 제외시켜도 될 것 같다.

```
0번째 gesture의 각 try의 segment 길이들 : 6 21 16 5 24 11 7 15 11 5
1번째 gesture의 각 try의 segment 길이들 : 3 12 12 5 7 11 11 18 9 10
2번째 gesture의 각 try의 segment 길이들 : 6 5 7 6 5 6 7 16 6 6
3번째 gesture의 각 try의 segment 길이들 : 6 5 5 5 5 4 4 6 5 5
```

18. 개별연구 일일 탐구일지

작성일	4주차	2020.08.05.	작성자	이혜민
일일 목표	<input type="checkbox"/> Naive Bayes classifier 적용하기 <input type="checkbox"/> Calibration : Ulna position 시작하기			

- Naive Bayes classifier 적용하는 중. 논문에서는 Leave-one-out cross-validation (LOOCV) 방식을

쓰고 있는데, 나는 data 용량을 줄이기 위해서 한 subject의 한 session에서 4개의 gesture만 이용하고 있기 때문에 LOOCV를 이용하지 못한다. LOOCV를 적용하고 싶으면 code를 훨씬 확장시켜야 한다. 그래서 일단 기본적으로 제공하는 predict method를 이용하겠다.

19. 개별연구 일일 탐구일지

작성일	4주차	2020.08.05.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Naive Bayes classifier 적용하기 <input checked="" type="checkbox"/> Confusion matrix plot 하기			

- Data 수가 적어서 그런지 accuracy가 많이 떨어진다. 조절할 수 있는 parameter를 조절해서 최대한 높여보자. 안되면 LOOCV를 적용할지도..
- Classifier에 넣는 input data를 다르게 가공할 아이디어를 사수님께서 제안하셨다. 해당 방법은 github의 README.md에 서술해놓았다. 두 번째 방법도 거의 다 구현함!

20. 개별연구 일일 탐구일지

작성일	4주차	2020.08.05.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Method 2 구현하기 <input type="checkbox"/> 이때까지 진행사항 정리하기			

- Classifying까지 구현 완료!! Part 1 CLEAR

Method2까지 구현을 다 했다. Method1에 비해 data 수가 증가해서 training을 심도있게 가능해졌고 그 결과 accuracy가 크게 증가했다. 아래에 예시 result를 정리했다.

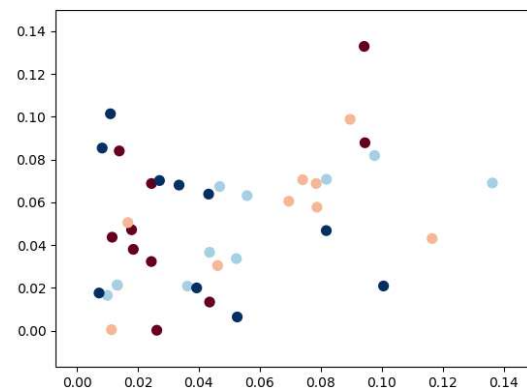
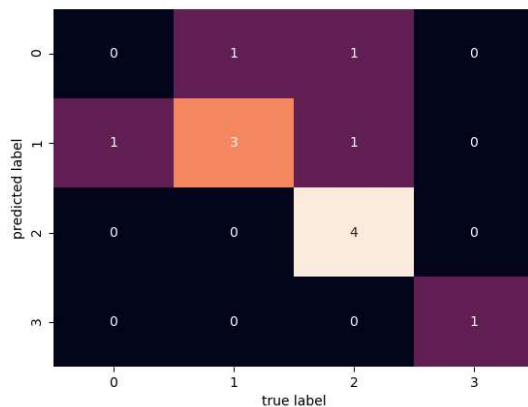


그림 8 . test_ratio=0.3, method1 N=3, confusion matrix

그림 9 . test_ratio=0.3, method1 N=3, scattered data

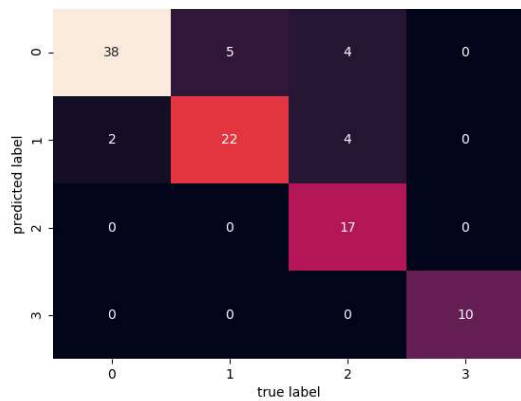


그림 10 . test_ratio=0.3, method2, confusion matrix

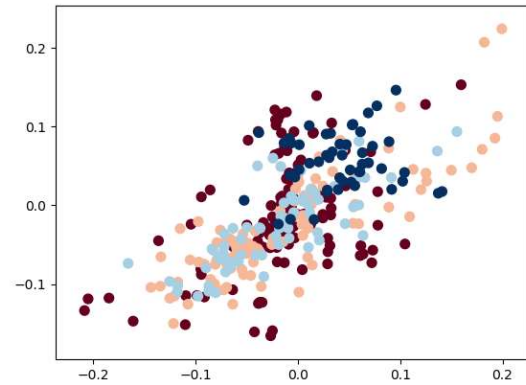


그림 11 . test_ratio=0.3, method2, scattered data

● To-do

- ▶ Readme 정리하기 : method1과 2에 대한 설명 추가, 전체적인 요약글 추가, 프로그램 타임라인 추가
- ▶ Scattered data를 그릴 때 사용하는 function의 parameter 제대로 알아보고 시각적으로도 gesture 별로 data가 구분되도록 plot하기
- ▶ 현재는 4가지의 gesture로만 실행해봤지만, gesture의 수도 증가시켜서 결과 얻어보기
- ▶ 앞으로의 계획 점검하고 진행상황에 맞게 수정하기

21. 개별연구 일일 탐구일지

작성일	5주차	2020.08.10.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Markdown 사용법 공부 <input type="checkbox"/> ReadMe 작성하기			

- ReadMe에 넣을 illustration을 그리고 markdown 사용법을 익힘. Method 1과 2에 대한 설명 추가 중.
- 참고링크
- [Markdown 사용법](#)

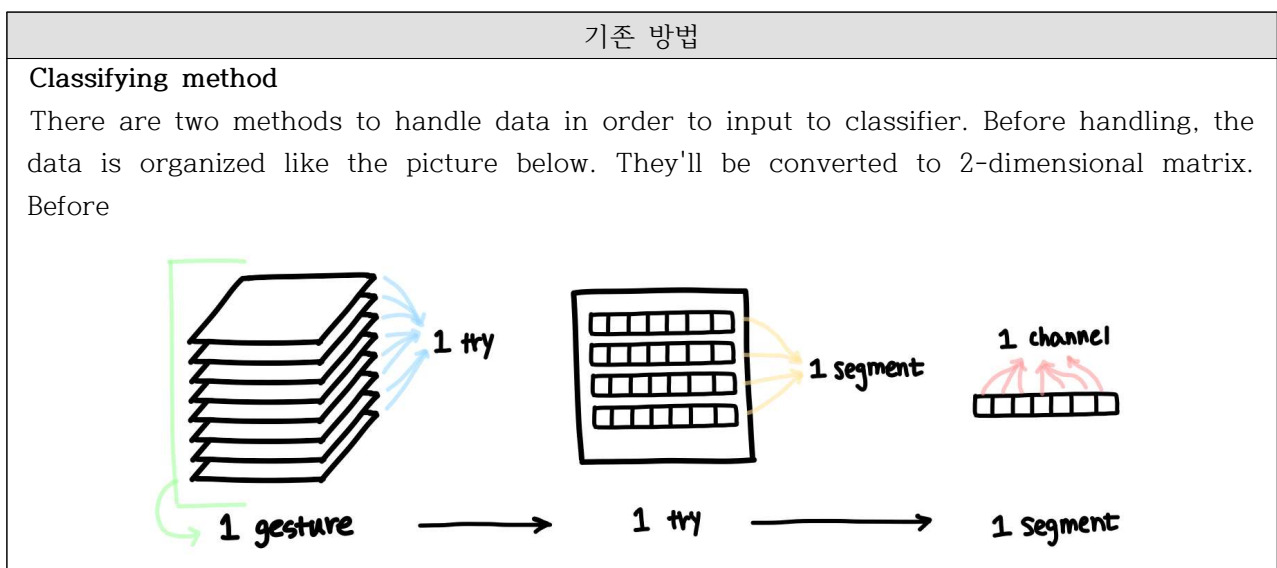
22. 개별연구 일일 탐구일지

작성일	5주차	2020.08.11.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> ReadMe 작성하고 github 정리 <input checked="" type="checkbox"/> 앞으로의 할 일 정리하기			

- 앞으로의 할 일
 1. 아래에 설명된 대로 data processing 코드 수정하기
 2. 다른 session의 gesture 0,1,2,3 data도 함께 넣고 train해서 result 얻기
 3. ref1 방법으로 calibration 구현하기
 4. calibration 적용한 result와 2에서의 result 비교하기
 5. 우리 EMG에 맞게 발전시킨 calibration 방법 구상, 적용

● Data processing

Data processing은 크게 2가지 단계로 나눌 수 있다. Active filter를 거치기 전과 후이다. active한 segment를 찾은 후에는 mean normalization을 하고 classifier에 넣을 수 있도록 2-dimensional matrix를 구성한다. 나는 아래의 method 1과 같은 방식으로 논문을 해석해서 코딩했지만 정확도가 너무 떨어졌고, 사수님이 제안해주신 method 2로 구현한 결과 정확도가 매우 높았다. 그런데 오늘 사수님과 앞으로의 계획을 정리하다가 내가 논문을 잘못 이해한 것이라는 결과가 나왔다.. method 1과 2를 적절히 조합하면 실제 논문에서 한 방법이 나온다. 그렇게 코드를 수정해야 한다.



Method 1

This method follows the reference 1.

There are various number of segments in each try, since active time is different for every try. Therefore, we'll make some groups for segments first. Suppose that we set $N=n$ and there are M segments in some try. This means we want to make n groups with M segments. From 0th to $(M/n*1)-1$ th segments will grouped into one group. Similarly, from $(M/n*1)$ th to $(M/n*2)-1$ th segments will grouped, and so on. The remainings are ignored.

When grouping is completed, compute RMS for each channel in each group. This results n numbers of 168-dimensional vectors in each group.

If we flatten the n vectors and attach the tries vertically continuously, 2-dimensional matrix will be constructed.

Method 2

Just attach all the segments and tries vertically continuously, so that one segment is regarded as one data in classifier.

재해석한 방법

각 try에 대해서 active한 segment를 찾는 것까지는 이전과 동일하다. 하지만 해당 segment를 이용하는 것이 아니라 해당 segment가 분포해있는 index를 기억해야한다. 그리고 해당 index에 대응되는 raw data를 다시 이용한다.

$N=n$ 이고 active한 구간 segment 길이가 l 이라 하자. Raw data는 각 channel이 시간에 따른 data를 가지고 있을 것이다. 이를 시간에 대해 n 등분하여 그룹으로 묶는다. 그러면 $(l//n)$ 개씩 묶인 n 개의 그룹이 나오고 뒤에 남은 몇 개의 raw data는 버려질 것이다.

이제 각 channel의 각 group에 대해 RMS 값을 구한다. 그러면 168개의 channel에서 n 개씩 RMS 값이 생길 것이다. 이는 각 channel의 RMS값으로 구성된 168-dimensional vector가 n 개만 들어진다고 해석할 수 있다. 우리는 이 n 개의 vector 각각을 classifier에 data로 넣을 것이다. 즉, 해당 vector들을 2-dimensional matrix에 정리하여 classifier에 input한다. 결론적으로 한 try에 대해 n 개의 data를 extract 하게 되는 것이다.

23. 개별연구 일일 탐구일지

작성일	5주차	2020.08.14.	작성자	이혜민
일일 목표	□ data processing 코드 수정하기			

- 위에서 같은 index의 rare data를 이용한다고 했는데, 진짜 rare인지 pre-process가 된 data인지, baseline normalization까지 된 data인지, medfilter까지 처리한 data인지 논문을 읽고 확인해보자

24. 개별연구 일일 탐구일지

작성일	5주차	2020.08.15.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 최종적으로 어떤 data 이용하는지 확인하기 <input checked="" type="checkbox"/> data processing 코드 수정하기			

-> pre-proces만 된 data, 즉 butterworth band-pass filter 처리만 된 data를 이용한다.

- pre_processed_one_try

shape : (40, 168, 150)

shape[0] : windows

shape[1] : channels

shape[2] : pre-processed data

- pre_processed_gestures

shape : (4, 10, 40, 168, 150)

meaning : (gestures, tries, windows, channels, pre-processed data)

- i_ACTIVE_windows

shape : (4, 10, 2)

meaning : (gestures, tries, start index of ACTIVE window, length of ACTIVE segment)

- ACTIVE_pre_processed_gestures

shape : (4, 10, n, 168, 150)

meaning : (gestures, tries, ACTIVE windows, channels, pre-processed data)

- ACTIVE_N_gestures

shape : (4, 10, 168, n*150)

meaning : (gestures, tries, channels, all the pre-processed data in ACTIVE windows)

- ACTIVE_N_RMS_gestures

shape : (4, 10, 168, N)

meaning : (gestures, tries, channels, RMS)

or

shape : (4, 10, N, 168)

meaning : (gestures, tries, large windows, RMS)

- mean_normalized_RMS

shape : (4, 10, N, 168)

meaning : (gestures, tries, large windows, mean normalized RMS)

● 주의 : ndarray도 list처럼 function으로 parameter로 넘겨준 것과 main에서의 값이 연동된다.

● Code 수정 후 결과

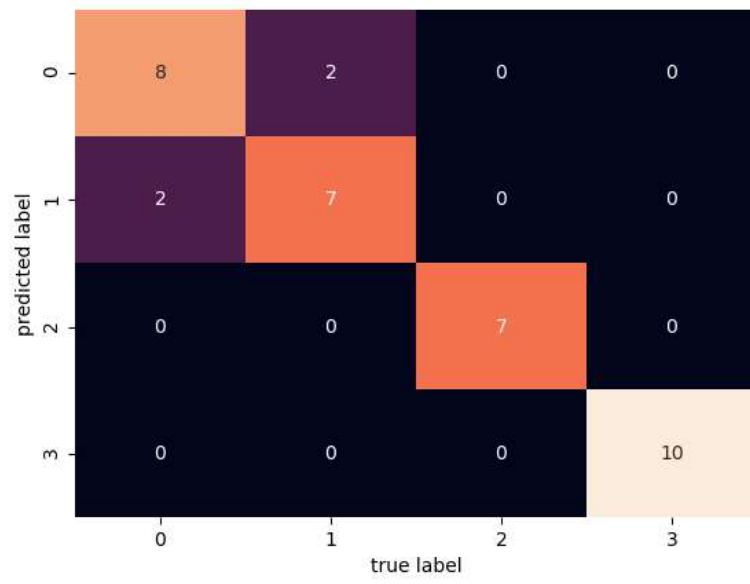


그림 13 . 20200815, test_ratio=0.3, N=3, confusion matrix

25. 개별연구 일일 탐구일지

작성일	6주차	2020.08.20.	작성자	이혜민
일일 목표	☑ 다른 session의 data도 input할 수 있게 코드 수정하기			

● 교수님과의 미팅 준비

- (7초-8초) 관련 논문 찾아보고, Ref 1의 classification 까지 재현
- 논문 ?

7초~8초	관련 논문 서치 Ref1의 classification 과정 재현	
8말	Ref1의 calibration 과정 습득 및 구현	
9초	우리 쪽 환경에 맞춰 알고리즘 수정 현재 idea : 1개의GMM을 기준으로main muscle activity estimation & position correction	
9말	1명 일반인 대상으로 한 HD EMG를 이용한 실험 (1차) HD EMG 데이터를 기존 코드에 적용	
10초		중간고사 준비
10말	---	중간고사(10/19-23)
11초	우리 쪽 환경에 맞춰 알고리즘 수정	
11말	결과 정리 및 피드백	
12초		기말고사 준비
12말	---	기말고사(12/14-18)
겨울방학	HD EMG를 이용한 실험 (2차) 더 많은 사람? 다른 프로토콜?	

● Result

Gesture : 0,1,2,3

Sessions : all the 5 sessions

Accuracy : 33% (흠..)

Confusion matrix :

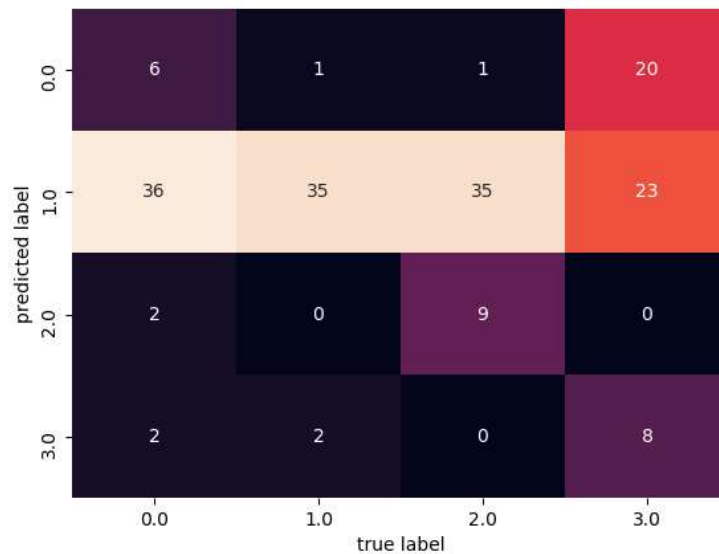


그림 14 . 20200820, test_ratio=0.3, N=3, every_5_sessions
confusion_matrix

- 사수님과 회의 결과, 우리 HD EMG를 이용한다면 두 군데의 main area와 ulna, 총 세 군데의 기준점을 모두 포함하게는 절대 붙일 수 없다. 그리고 신호가 발생하는 곳에 붙일 것이기 때문에 ulna 부근의 data는 input 받는 일이 없을 것이라고 판단했다. 따라서 두 번째 방법인 Estimation of center of main muscle activity를 이용할 것이며 이때도 우리 EMG는 하나의 main area만 한 번에 입력 받을 수 있는 크기이기 때문에 두 개의 gaussian model을 estimate하는 기존의 방법과는 다르게 하나의 gaussian model만 estimate하도록 할 것이다.

결론적으로 우리는 입력받은 데이터에서 main muscle activity가 일어나는 지점을 하나의 GMM으로 estimate하여 calibrate 할 것이다.

- Bookmark

[Plotting 3d data in 2d](#)

26. 개별연구 일일 탐구일지

작성일	6주차	2020.08.21.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> Data plot해보기 <input checked="" type="checkbox"/> 교수님과 미팅			

- 교수님 미팅

- 논문을 쓰거나 URP까지 하려면 calibration만으로는 부족. Application까지 생각해보자. 가을학기까지 작업하면서 괜찮은 주제 계속 생각해보기.
- 기계과 랩에서 URP 가능한지 알아보기 -> 문의 결과 다른 학과 교수님도 가능!

- Bookmark

[plotting 3d data](#)

[plot discrete 3d data in colormap](#)

- reshape 배열이 잘못되고 있는듯함. 내일 수정하자!
- 주의
 - 모든 session 폴더에 같은 gest7 넣어둠
 - random하게 뽑을 때 4번째 gesture만 뽑히도록 설정해놓음
 - plot_some_data 마지막에 check 있음

27. 개별연구 일일 탐구일지

작성일	6주차	2020.08.23.	작성자	이혜민
일일 목표	<input checked="" type="checkbox"/> 디버깅 해서 plotting 완료			

- Plotting 완료!!

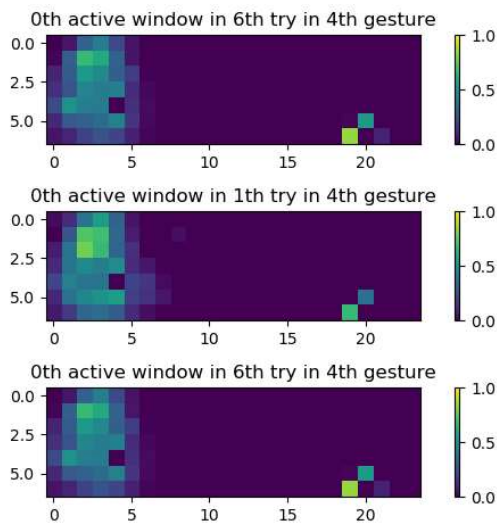


그림 15 . 20200823_data_gest7_method2

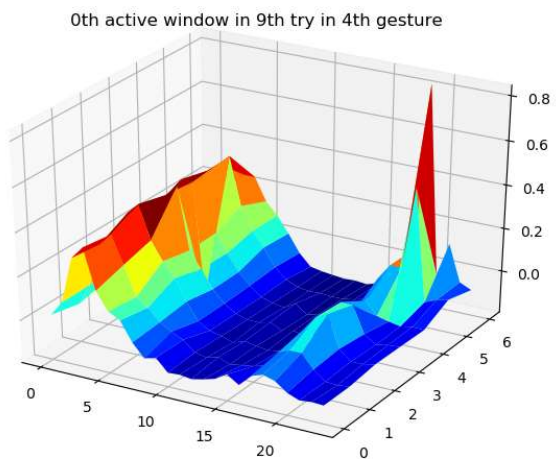
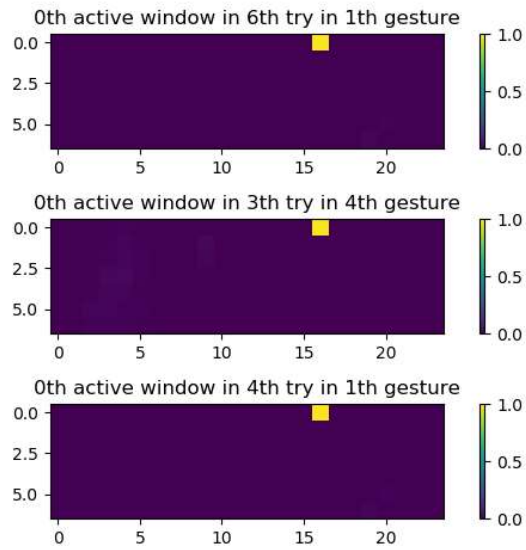


그림 16 . 20200823_data_gest7_method1

- Data plot 시 특이사항
 - 각 session 별로 임의의 data를 선택하여 plot한다. 무슨 gesture의 몇 번째 try의 어떤 window일 때의 data인지는 모두 random이다.
- 문제점

그 다음 session들의 모든 데이터가 아래와 같이 나옴. 디버깅 필요. 위에서 다른 session을 넣었을 때 accuracy가 급격히 감소한 이유로 추정됨.



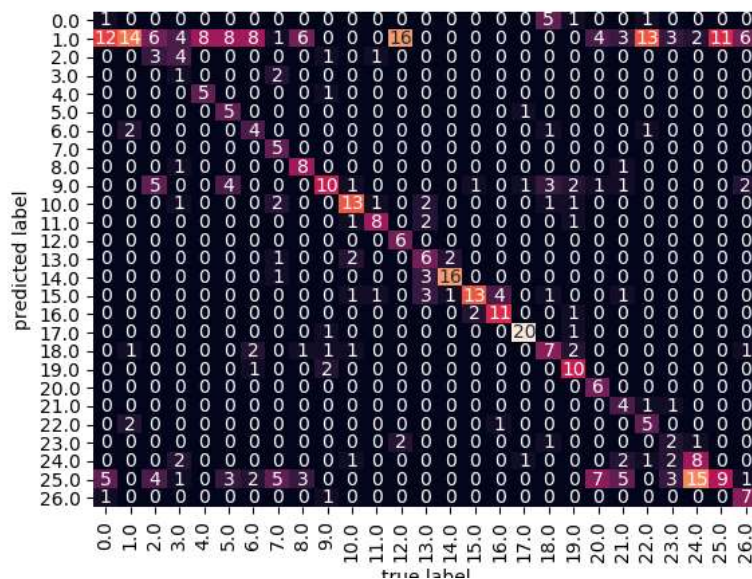
다음주 할 일	<input type="checkbox"/> 디버깅 : 첫 번째를 제외한 session들 data 처리 <input type="checkbox"/> Data plot 시 세로축이 아래로 갈수록 커지는 것이 아닌 위로 갈수록 커지도록 바꾸기 (논문에서의 channel 넘버링이 해당 순서임) <input type="checkbox"/> Calibration 구현
------------	--

28. 개별연구 일일 탐구일지

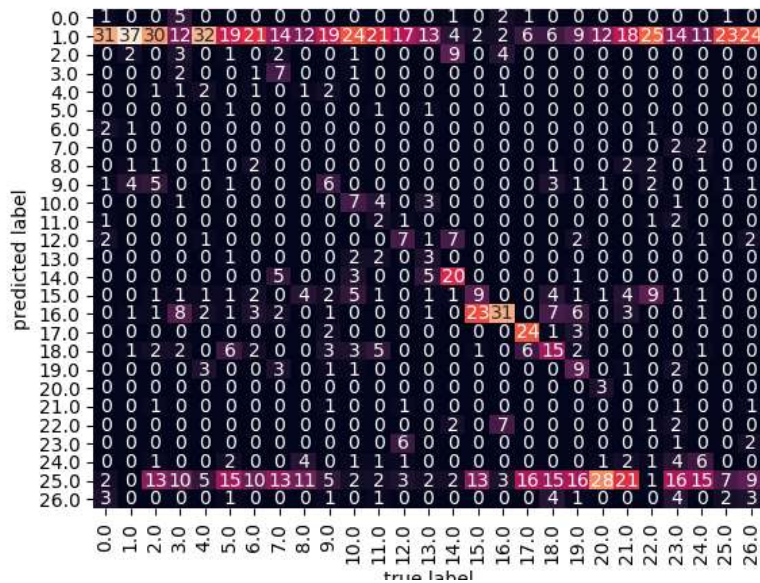
작성일	7주차	2020.08.25.	작성자	이혜민
일일 목표	<input type="checkbox"/> 디버깅 : 첫 번째를 제외한 session들 data 처리 <input type="checkbox"/> Data plot 시 세로축이 아래로 갈수록 커지는 것이 아닌 위로 갈수록 커지도록 바꾸기 (논문에서의 channel 넘버링이 해당 순서임)			

● 디버깅

- 0번째 session과 1번째 session의 data만 바꿈 -> 1번째 session만 제대로 나온 것을 봤을 때 프로그램 문제가 아니라 data 문제.
- 노란점이 나타나는 곳은 105index.
- 각 session만 실험해본 결과 (Accuracy)
session1: 77%, session2: 95%, session3: 97%, session4: 35%
- session 0,2 : 46% (gest0에 몰림)
- **session 2,3 : 93%**
- **session 1,2,3 : 71%**
- session 0,2,3 : 37% (...)
- session 0,1,2 : 40%
- session 1,2,3,4 : 52% (-> 6m 22s 걸림.. 전체적으로 너무 오래 걸린다.)
- session 0,1,2,3 : 32% (gest0,1에 predict 몰림)
- session 0,1,2,3,4 : 26% (gest1에 predict 몰림)
- session 0,1,2 : gest1,2,3,7 : 37%
- session 0,1,2,3,4 : gest2,7,11,20 : 36% (신호가 확연히 차이나는 것으로 해도 똑같다..)
- session 0,1 : gest all : 42% (아직 논문에 비하면 낮은 값이지만 적은 gesture를 이용했을 때 보다는 결과가 좋다. 대부분의 오류는 gesture 1로 잘못 predict하면서 일어난다.)



- session 0,1,2,3,4 : gest all : 16% (gest1 또는 25로 몰린다.)

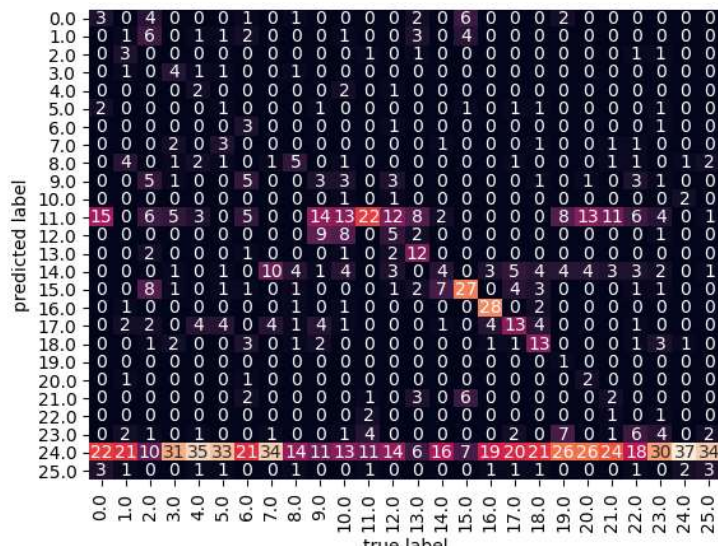


29. 개별연구 일일 탐구일지

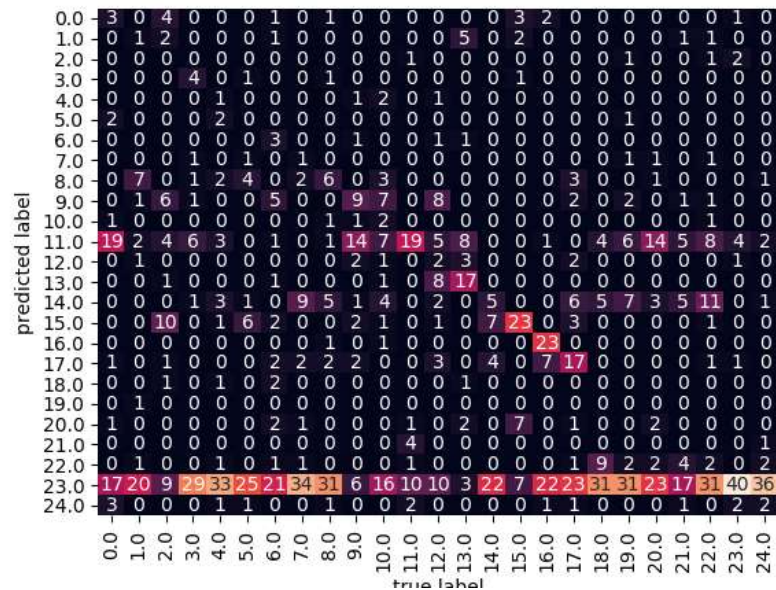
작성일	7주차	2020.08.26.	작성자	이혜민
일일 목표	□ 디버깅 : 여러 session 넣었을 때 classifying			

● 디버깅

- session 0,1,2,3,4 : gest All except 1 : 16% (그래프에서 24, 즉 gest25에 몰린다. 1과 25가 둘 다 없다면?)



- session 0,1,2,3,4 : gest All except 1,25 : 16% (23분 소요. 개선된 점 없음. gest 24로 몰린다..)



● 사수님과 체크해본 결과

코드에 문제가 있다. 데이터를 가공하는 과정에서 뭔가 잘못되었고 데이터가 이상하게 변해서 분류가 안 되는 것.....

한 단계 한 단계 plot 해가면서 어디가 문제인지 판별해내자..

이것이 힘들면 numpy를 잘 활용해서 다시 코드를 짜자..

* within-session이더라도 session1과 4는 제대로 작동하지 않는다. 그것을 고치는 방향으로 해보자.

할 일	<input type="checkbox"/> 디버깅 : 무언가 잘못되었다. 처음부터 확인하여 classify가 정상적으로 되도록 하자.
-----	---

30. 개별연구 일일 탐구일지

작성일	7주차	2020.08.27.	작성자	이혜민
일일 목표	<input type="checkbox"/> 디버깅 : 무언가 잘못되었다. 처음부터 확인하여 classify가 정상적으로 되도록 하자.			

- session 1,4 : gest 0,1,2,3,7 : 53%

- session 4 : gest 0,1,2,3,7 : 35%

● RMS computing하기 전까지 간단히 함!

- 간단히 하면서 plot해서 오류 찾는 중. 실행 시간도 가능하면 동시에 줄이자!

- 이전에는 window를 새로 만들었는데, 이번에는 기존의 .mat 파일이 가지고 있는 형식 그대로를 최대한 이용하고자 함. 그리고 numpy의 기능을 최대한 활용하여 for문을 줄여서 시간을 단축하는 것!

31. 개별연구 일일 탐구일지

작성일	7주차	2020.08.28.	작성자	이혜민
일일 목표	□ 디버깅 : 무언가 잘못되었다. 처음부터 확인하여 classify가 정상적으로 되도록 하자.			

[python numpy slicing](#)

● 수정된 이후의 variable shapes

- sessions

type : dict

keys : "ref1_subject1_session#"

values : gest#.mat data

- gestures

type : ndarray

shape : (5, 10, 192) <- 각각의 192에는 (6144,)의 ndarray가 들어있음

meaning : (gestures, tries, channels, raw data)