



Reference 1 Summary

Advancing Muscle-Computer Interfaces with High-Density Electromyography

In this paper we present our results on using electromyographic (EMG) sensor arrays for finger gesture recognition. Sensing muscle activity allows to capture finger motion without placing sensors directly at the hand or fingers and thus may be used to build unobtrusive

 <https://dl.acm.org/doi/pdf/10.1145/2702123.2702501>



Association for
Computing Machinery

ABSTRACT

Finger gesture를 EMG를 이용하여 구별하는 것을 목표로 한다. 이는 직접적으로 손에 place하지 않아도 되므로 unobtrusive하다. 192개의 electrode를 가지고 있는 HD EMG를 이용했고 안쪽 팔뚝(upper forearm)에 부착했다. 우리는 27개의 gestures를 구별하기 위해 naive Bayes classifier 를 사용했고 gesture recognition의 baseline system 에 대해 자세하게 다루겠다. 5명의 subject를 대상으로 각각 5 session을 진행하여 총 25 session의 data를 얻었다.

Within-session에서 classify했을 때는 평균 90%의 정확도를 보였고 이는 구별하기 힘든 gesture를 감지할 수 있을 것이라는 가능성을 보여준다. 여기서 electrodes의 수에 따른 recognition performance도 분석하였는데 많은 수의 electrode가 있을수록 performance가 좋아졌다.

Cross-session의 경우 session이 달라질 때마다 변화하는 electrodes의 position이 accuracy에 영향을 미쳤다. 따라서 이를 calibrate하기 위한 2가지 method를 제시했으며 해당 method들을 적용한 경우와 적용하지 않은 baseline system을 비교하였다. 그 결과 cross-session case의 accuracy가 59%에서 75%로 증가하였다.

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 움직임을 감지하는 방법, EMG 등이 그 예시이다. Finger에 관여하는 대부분의 muscles가 forearm에 위치하므로 그곳에 EMG를 부착하여 손가락과 관련된 muscle activity를 감지할 수 있을 것이다.

Related Work


Array-like electrode는 단 몇 개의 electrode를 부착하는 것에 비해서 하나하나의 electrode를 정확한 position에 부착할 필요가 덜하다. 이는 자주 taken on and off 해야 하는 wearable device에

게 장점으로 작용한다. 따라서 array-like EMG를 활용한 연구를 중점적으로 소개한다.

- Array-like EMG는 아니지만 각 electrode의 data를 분석하는 과정이 자세하게 설명되어 있다.

Gesture-based control and EMG decomposition - IEEE Journals & Magazine

IEEE Xplore, delivering full text access to the world's highest quality technical literature in engineering and technology. | IEEE Xplore

 <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1643841>

BACKGROUND

Distal에는 thumb, pinkie, index finger에 관여하는 근육이 존재한다. 하지만 thumb을 제외한 나머지 손가락에 관여하는 주요 근육은 proximal에 존재하므로 EMG를 forearm에 부착하면 thumb을 제외한 손가락들의 activity를 capture할 수 있을 것이다.

Electromyography

Surface EMG는 몇가지 challenges를 동반한다.

- Positioning of the sensors
: Mainly anatomy reason
- inter- and intrapersonal differences in the signals
: [3]에서 recognition performance와 BMI가 negative correlation 관계에 있음을 밝혔다.

METHODOLOGY

Experimental setup과 classification pipeline, shift compensation의 2가지 방법에 대해 소개한다.

Experimental setup

8*24의 EMG 사용. Cable의 길이로 인한 noise를 줄이기 위해서 pre-amplifier를 모든 cable에 부착했다.

Bipolar recording을 이용한다. 이는 amplifier가 연이은 channel의 전압차를 이용하는 것으로, 따라서 각 column에서 하나의 electrode는 meaningless한 data를 나타낼 것이며, 총 192개의 channel 중 168개만 usable data를 제공한다.

실험 결과 channel이 168개까지는 필요없고 20에서 80개면 적절하다.

2048Hz의 속력으로 기록됐고 amplifier의 gain은 1000이었다.

Gesture set

Gesture들은 각 손가락들이 extension되고 flexion도 될 수 있도록 구성되었으며 HCI에서 활용될 수 있을 gesture도 포함되었다. 최종적으로 Tapping gestures, Bending gestures, Multi-finger gestures로 구성되고 총 27가지의 gesture가 존재한다.

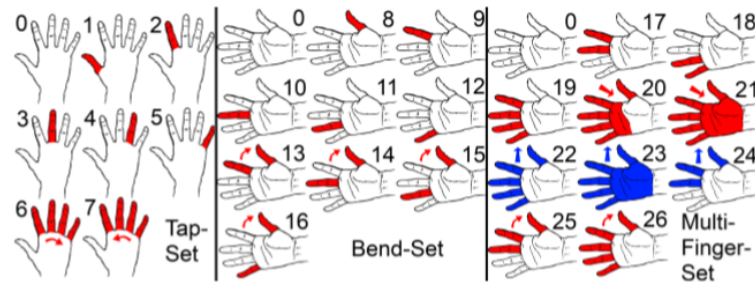


Figure 4. Iconic illustrations of the performed gestures. The gestures were divided into three sets: the tap-set, the bend-set and the multi-finger-set. The idle gesture (0) is contained in all three sets.



각 gestures에 대해 idle gesture도 알맞게 설정해야 한다.

Classification Pipeline

우리는 naive Bayes classifier라는 일반적인 classification scheme을 이용해서 classification을 했다.

전체적인 classification 과정은 다음과 같다. Signal preprocessing, segmentation, feature extraction, applying to naive Bayes classifier이다. 먼저 얻어낸 data에서 artifact를 줄이기 위해 preprocessing 작업을 거친다. 이후의 segmentation step에서는 해당 gesture를 취할 때 activate되는 muscle의 part가 어딘지 판별해낸다. 이 과정을 거치면 data가 분할된 상태가 될 것이고 여기서 feature를 extract하는 것이 세 번째 과정이다. 마지막으로 extract한 feature를 naive Bayes classifier에 apply하여 27가지의 class 중 어떤 class, 즉 어떤 gesture 인지 판별해내면 된다.

본격적으로 설명하기 전에 RMS와 baseline normalization에 대해 설명하겠다.

- EMG의 RMS value는 muscle에서 생성되는 force와 연관이 커서 muscle의 activity를 나타내는데에 일반적으로 사용된다.
- Baseline normalization은 영점을 보정해주는 것과 같다. Baseline normalization을 적용할 RMS value에서 average idle gesture의 RMS value를 subtract 해주는 것이다. 이때 average idle gesture는 각 session의 각 gesture set (tapping, bending, multi-finger) 마다의 평균을 말한다. 전체 idle gesture data를 대상으로 average를 구하는 것이 아니기 때문에 훨씬 정확한 baseline을 잡을 수 있을 것이다.

이제 각각의 단계에 대해 알아보자.

1. 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.

2. Segmentation

Subject들은 각 gesture를 3초 동안 시행하도록 안내받았다. 따라서 해당 3s window 안에서 실제로 gesture를 취하고 있는 부분을 찾아서 그 곳을 segment로 지정해야한다. 우리는 gesture를 취하고 있는 상태를 active 하다고 할 것이며 가장 길게 active한 상태가 연속되는 구간을 segment로 정할 것이다.

먼저 3s window를 non-overlap한 여러개의 window로 segmentation하자. 이때 하나의 window는 150 sample, 즉 73.2ms의 구간을 포함하도록 하고 마지막에 남은 window는 무시한다. 지정된 window들 안에서 각 channel 별로 data의 RMS value를 구하면 각 window는 168개의 RMS value를 element로 가지는 168-dimensional vector로 표현될 수 있다.

그리고 local artifacts를 compensate 해주기 위해서 해당 vector들에 대해 spatial order 3 1-dimensional median filter를 적용한다.

우리는 해당 window의 vector element의 전체 합이 threshold를 넘으면 해당 window가 active하다고 할 것이며 해당 window가 앞의 조건을 만족하지 않더라도 predecessor와 successor가 모두 active하면 해당 window가 active하다고 할 것이다. 여기서 threshold는 각 window들의 vector element의 전체 합의 average로 결정한다. 이 과정을 통해 각 window는 active함의 여부가 결정되었고, 이제 가장 길게 연속된 active sequence를 gesture segment로 선택하면 된다.

3. Feature Extraction

발견한 segment는 같은 길이의 N개의 windows로 구성되어 있을 것이다.

We compute the RMS as feature for each channel on all windows and normalize the mean RMS over all channels.

→ 우리는 전체 window의 각 channel의 feature를 RMS value로 계산한다. 그리고 전체 channel에 대해 mean RMS를 구해서 normalize한다.

그 결과 각 window 당 168 dimensional feature vector가 만들어질 것이고, 이는 168*N dimensional feature vector로 볼 수 있다. 이때 RMS가 length normalized 되었기 때문에 각 gesture 별로 다른 길이의 activity segments를 가지고 있더라도 상관이 없다.

4. Naive Bayes classifier

우리는 naive Bayes classifier로 classification을 할 것이다. 이때 feature distribution은 Gaussian kernel function을 이용하여 kernel density estimation으로 모델링 되었다. 우리 실

험은 independence assumption이 성립되지 않지만 이와 같은 조건 하에서도 naive Bayes classifier는 잘 작동하므로 괜찮다.

Estimation of Electrode Displacement

2개의 session 사이에서 electrode의 displacement가 어떻게 차이 나는지 비교하는 방법으로 우리는 2가지를 사용하였다. 첫 번째는 ulna 위에서는 항상 muscle activity가 low하다는 성질을, 두 번째는 Main flexor와 extensor 위에서는 항상 muscle activity가 high하다는 성질을 이용한 것이다. 이를 각각 gesture에 대해서 분석한 후에 평균을 내서 최종 coordinate를 결정했다.

두 방법을 통해서 shift를 판별했다면, data를 그 만큼 shift 시킨다. 그런데 이때 정확하게 inter-electrode distance인 10mm씩 shift 되는 경우는 드물기 때문에 bicubic interpolation을 이용하여 interpolate한 값을 이용했다. 그리고 이렇게 shift하게 되면 border 부분에는 unknown data가 분포하는 부분이 발생하게 되므로, 여기는 marginal한 value로 채웠다.

1. Estimation of ulna position

Ulna의 바로 위에는 muscle이 아래 존재하지 않아서 매우 낮은 activity가 감지된다. 이때 ulna가 y축에 평행하다는 가정하에 x축에 대한 shift만 고려할 것이다.

Estimation은 다음과 같은 순서로 행해진다.

1. Preprocessing과 data normalization을 한다.
2. 전체 x의 범위에서 middle 부분에 비중을 두기 위해서 높거나 낮은 x value를 가진다면 penalty를 부여한다.
: Preprocessing된 data는 [0,1]에 분포하기 때문에, penalty function $p(x)$ 는 다음과 같이 정의할 수 있다. $N(12.5, 10^2)$ 의 density function을 $\varphi(x)$ 라 할때,

$$p(x) = 1 - \frac{\varphi(x)}{\max(\varphi(x))}$$

3. Watershed algorithm을 이용해서 모든 watershed를 찾는다.
4. Dijkstra's algorithm을 이용하여 lowest한 한 watershed를 찾고 그 곳을 ulna라고 판단한다.

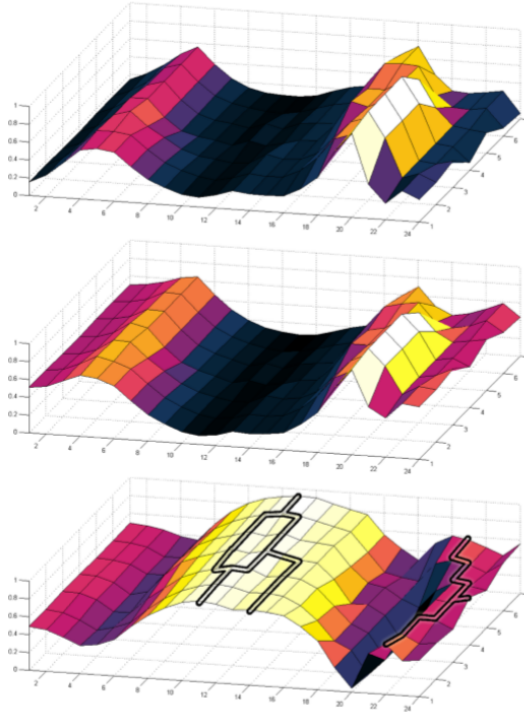


Figure 5. Example steps of identifying the ulna: 1) preprocessed normalized data, 2) data after application of the penalty function, low and high x values receive penalty, 3) found watersheds on the inverted data.

2. Estimation of center of main muscle activity

Flexor와 extensor 위에서는 muscle activity가 high하며 각각으로 인한 muscle activity는 독립된 normal distribution을 따른다고 간주한다. 따라서 각 normal distribution의 center와 shape를 찾는 것이 목표이고 이에 GMM(Gaussian Mixture Model)을 이용한다.

μ_1, μ_2 의 initial value는 각각 (1,4), (24,4)로 지정하여 각각이 GMM의 왼쪽과 오른쪽으로 수렴하도록 support 해준다. Model fitting은 Expectation-Maximization algorithm을 이용한 ML estimation을 이용한다. 각각을 이용해서 shift된 정도를 찾아내고 둘의 결과를 평균 낸 값을 최종 shift된 정도로 결정한다.

EXPERIMENT

5명의 subject를 대상으로 모두 다른 날에 각각 5 session 씩 진행하였다. 각 session 별로 idle gestures는 총 30회, 다른 gestures는 10회씩 recording하여 최종적으로 7250개의 data를 확보했다.

Experiment Procedure

EMG는 팔꿈치에서 2-4cm 떨어진 곳에 부착했다. 팔뚝의 바깥쪽에는 손가락에 관여하는 muscle이 없기에 바깥쪽에는 붙이지 않았다. 각 session이 모두 다른 날에 진행됐기에 조금씩 다른 위치에 부착됐을 것이며 이는 더 현실적인 실험이었다고 해석할 수 있다.

Subject를 위한 simulation interface를 어떻게 구성했는지 자세히 설명하고 있다. Gesutre는 항상 tap-set, bend-set, multi-finger-set의 순서로 진행됐는데, 이는 갑작스러운 손의 변화를 막기 위해서였다. 손이 갑작스럽게 변하면 더 강한 skin deformation을 발생시킨다. 이는 몇몇 electrode를 떨어뜨릴 수도 있고 그렇게 된다면 다시 부착하는 것도 불편할뿐만 아니라 다시 부착하는 과정에서 condition이 변화하기 때문에 정확도 또한 떨어뜨릴 것이다. 하지만 같은 gesture set 안에서는 크게 문제가 없기 때문에 Figure 4의 순서와는 다르게 배치하였다.

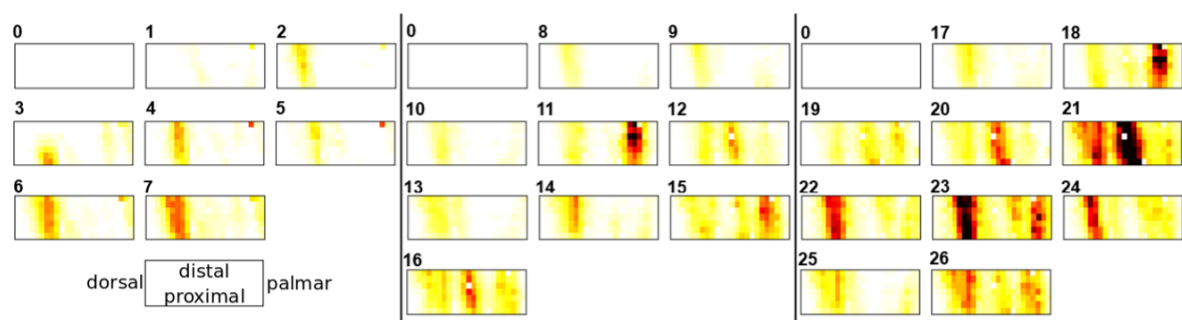


Figure 6. Average muscle activity per gesture for one selected session. The position of the graphs resembles the order in Figure 4. The color map encodes the average RMS value over the ten repetitions per gesture after segmentation and baseline normalization. Each pixel represents on channel. Single isolated channels with very high or low RMS values are defective channels which were not compensated by filtering and baseline normalization.

Figure 6는 하나의 session 결과이다. 10번씩 행해진 각 gesture의 average RMS value를 색으로 나타내고 있다. 각 pixel은 각 electrode를 나타내며 주로 오른쪽에서 보이는 activity는 flexor로 인한 것이고 왼쪽에서 보이는 activity는 extensor로 인한 것이다.

Results and Analysis

먼저 같은 session 안에서 classification을 시행해보고, 다른 session들도 포함해서 분석한 경우를 살펴 보겠다. 마지막으로 gesture의 종류를 줄여서 제한된 gesture 안에서 분석하고자 한다.

Within-session classification

LOOCV(leave-one-out cross-validation)을 이용하여 각 session 별로 classification을 실행했다. 그리고 모든 session의 performance를 평균 내어 최종 accuracy를 결정했다. LOOCV를 할 때 window를 1개만 한 경우가 Table 2.이고 window가 3개인 경우가 Table 3.이다.

Sub	S1	S2	S3	S4	S5	Avg	Std
A	83.7	90.0	95.9	94.4	84.4	89.7	5.6
B	89.6	94.4	90.0	88.5	88.2	90.1	2.5
C	87.4	91.5	90.4	88.5	83.3	88.2	3.2
D	90.4	95.6	82.6	90.4	88.2	89.4	4.7
E	93.3	88.9	94.1	92.6	88.2	91.4	2.7
						89.8	3.7

Table 2. Accuracies of the session-wise cross-validations with $N = 1$. Each row corresponds to one subject and the columns denoted by S1 to S5 correspond with session 1 to session 5.

Sub	S1	S2	S3	S4	S5	Avg	Std
A	85.2	90.7	95.2	94.8	90.7	91.3	4.0
B	83.7	92.2	94.1	88.9	90.0	89.8	3.9
C	88.9	93.3	92.6	92.2	88.9	91.2	2.1
D	87.8	92.2	87.0	85.4	84.8	87.4	2.9
E	91.5	89.6	96.3	94.1	90.4	92.4	2.8
						90.4	3.2

Table 3. Accuracies of the session-wise cross-validations with $N = 3$. Each row corresponds to one subject and the columns denoted by S1 to S5 correspond with session 1 to session 5.

세로축은 subject, 가로축은 session을 나타내며, N을 변화시키며 실험해본 결과 N=3일 때 가장 accuracy가 높았다.

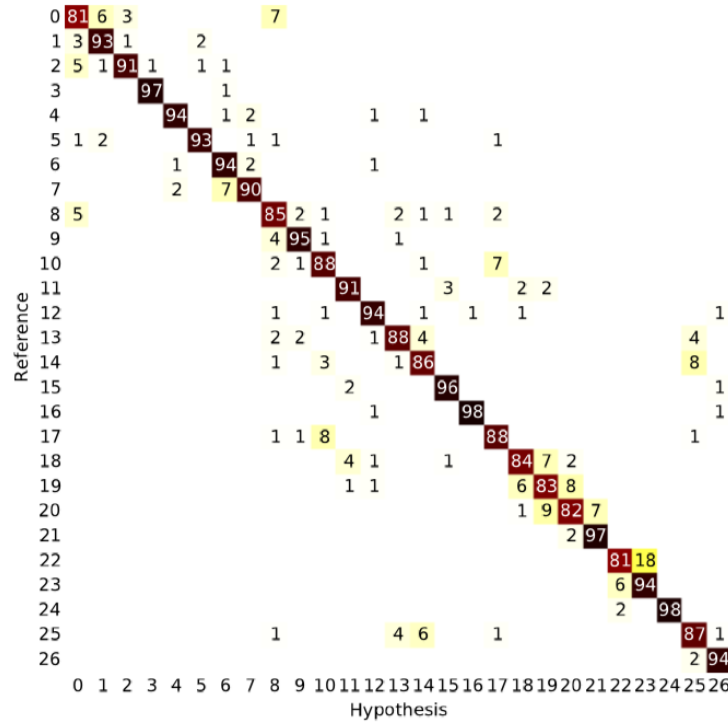


Figure 8. Accumulated confusion matrix for the within-session evaluation over all subjects and sessions. The numbers give the percentage of how often the pair of reference and hypothesis appeared in the classification results. Numbers were rounded and zeros were omitted.

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

Electrodes의 수에 따른 performance의 변화를 분석하기 위해서 실제로 반영하는 electrode의 수를 조절하여 실험해보았다.

The actual electrodes that we used were chosen by maximizing mutual information with the labels on a per session basis.

→ 분석에 반영하는 electrode는 session 당 label을 이용하여 mutual information을 최대화 하는 방향으로 선택했다.

As feature selection method we compute the mutual information of each of the channels with the label vector.

→ Feature를 선택하는 방법은 label vector를 이용하여 각 channel의 mutual information을 계산함으로써 계산했다.

Mutual information is computed using kernel density estimation.

→ 여기서 mutual information은 kernel density estimation을 이용하여 계산되었다.

This results in different feature sets for each session.

그 결과 Figure 9와 같은 결과가 나왔고 20-80개가 가장 이상적이라고 해석할 수 있다.

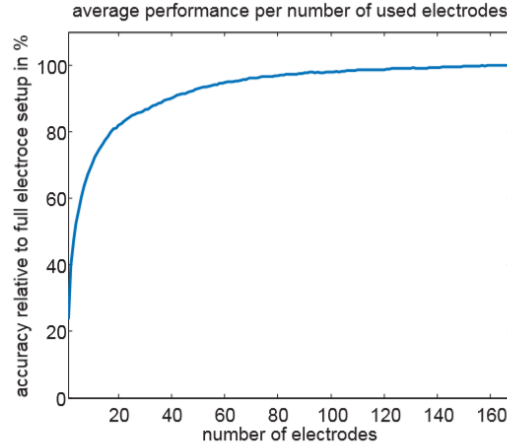


Figure 9. Average classification accuracy by number of used electrodes relative to full setup of 168 channels. Electrodes were selected by highest mutual information with the labels on a per session basis.

Shift compensation

각 subject에 대해서 shift compensation을 진행했고 이때 session에 대해서 leave-one-out cross-validation을 이용하였다. 그런데 calibration gestures도 classified 되어 있기 때문에 classified되지 않은 독립적인 자료들보다 accuracy가 증가되어 있을 것이다. 하지만 우리는 이가 결과에 큰 영향을 미치지 않으리라 믿는다.

	Baseline		GMM		Ulna	
N	1	3	1	3	1	3
A	43.5	60.6	73.9	83.9	69.9	82.9
B	58.1	73.0	65.4	77.2	66.2	76.4
C	55.4	62.3	58.3	71.0	56.8	69.1
D	40.5	53.3	53.9	66.5	59.0	71.3
E	27.9	45.3	67.4	78.4	63.5	76.4
Avg	45.1	58.9	63.8	75.4	63.1	75.2

Table 4. Accuracies of the cross-session evaluation. Each row corresponds to one subject and the columns correspond to the different shift compensation methods including the baseline with no shift compensation and the number of windows used (parameter N).

Table 4.는 각 calibration 방법에 따른 accuracy를 나타낸다. Baseline은 아무 calibration도 하지 않은 경우이다. 그런데 GMM의 경우 gesture 11의 accuracy를, Ulna의 경우 gesture 23의 accuracy를 나타낸 것이다. Ulna를 이용했을 때 전체 gesture에 대한 accuracy는 Figure 10.에 나타나있다.

Calibration을 하면 accuracy가 약 60%에서 75%로 증가한다. 하지만 within-session의 accuracy에는 도달하지 못한다. 이는 땀의 유무로 매일 달라지는 피부의 conductance와 그날의 electrode의 접착력이 얼마나 좋은지, 그리고 매일 조금씩 달라지는 gesture performance의 영향으로 볼 수 있다.

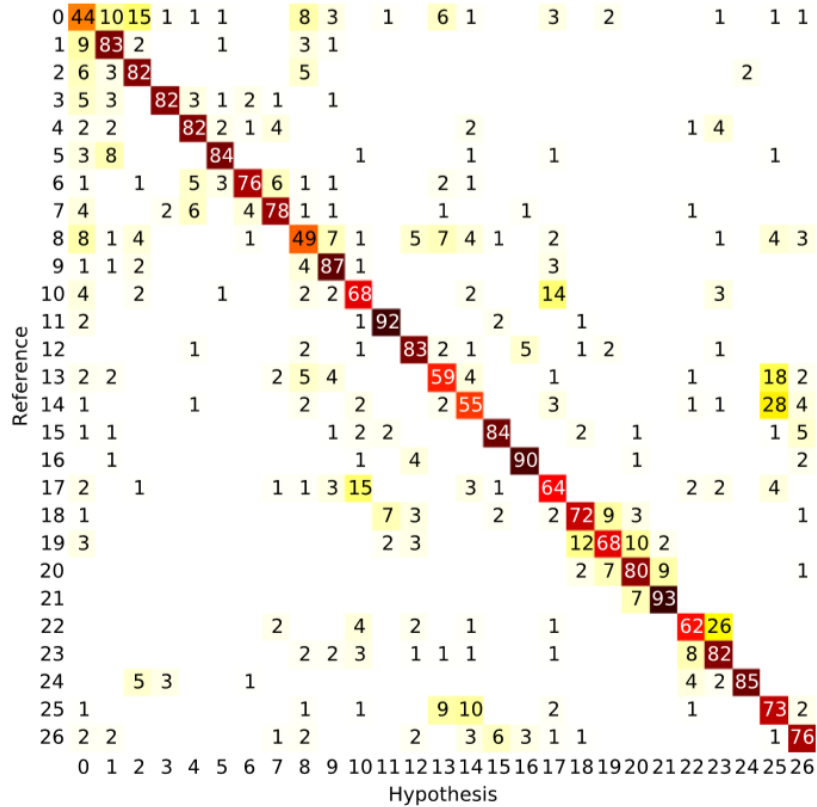


Figure 10. Accumulated confusion matrix for the cross-session evaluation over all subjects and sessions. The numbers give the percentage of how often the pair of reference and hypothesis appeared in the classification results. Numbers were rounded and zeros were omitted.

Figure 10.을 보면, within-session에서 주로 혼동됐던 gesture들은 그 정도가 훨씬 심해졌고, 다른 gesture들도 전체적으로 혼동이 생긴것을 알 수 있다.

• 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.

Restricted gesture sets

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.

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.

Conclusion and Future work

전체 요약, 미래에 더 발전된 EMG가 나오길 바란다.