University of Tartu
Institute of Computer Science
Course “Introduction to Computational Neuroscience”
Roberts Mencis

PREDICTION OF FINGER FLEXION FROM ELECTROCORTICOGRAPHY DATA

Abstract

This project aims to predict finger movements from electrocorticography (ECoG) data based on datasets from BCI Competition IV, which took place in Berlin at 2008. The official winner of competition achieved 0.46 correlation between predicted and actual finger flexion data. The goal of this project is to perform in-depth analysis of data, discover relevant relations within data, build prediction model and hopefully achieve comparable or better result than current competition winner.

Introduction

Experimental data was obtained from three epilepsy patients who had electrode grid (see figure 1) placed on the surface of the brain to localize seizures (Miller & Schalk 2008).

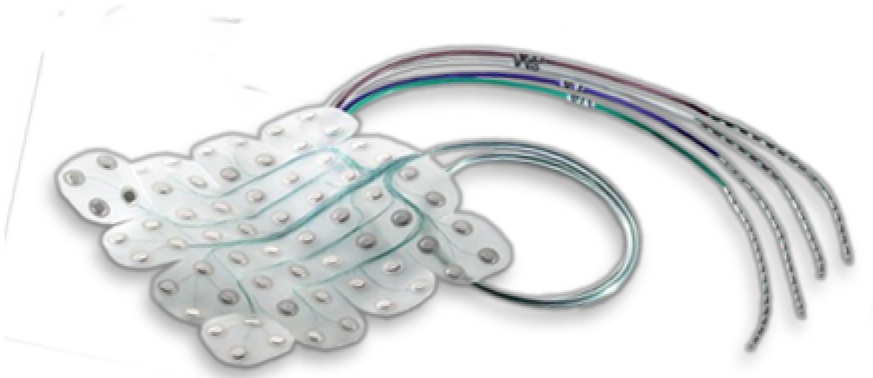


Figure 1: Electrode grid for electrocorticography
(Source - http://www.blackrockmicro.com/content.aspx?id=41)

Subjects were wearing data glove from Fifth Dimension Technologies (see figure 2), which makes it possible to register movements of all fingers (Miller & Schalk 2008).

Figure 2: Dataglove from Fifth Dimension Technologies (Source - http://www.5dt.com/products/pdataglove5u.html)

Duration of experiment was 10 minutes for each subject. Every 2 seconds random cue was presented on computer monitor indicating which finger to move. Subject then moved required finger several times. Each cue was followed by 2 seconds rest interval. See figure 3 (Miller & Schalk 2008).

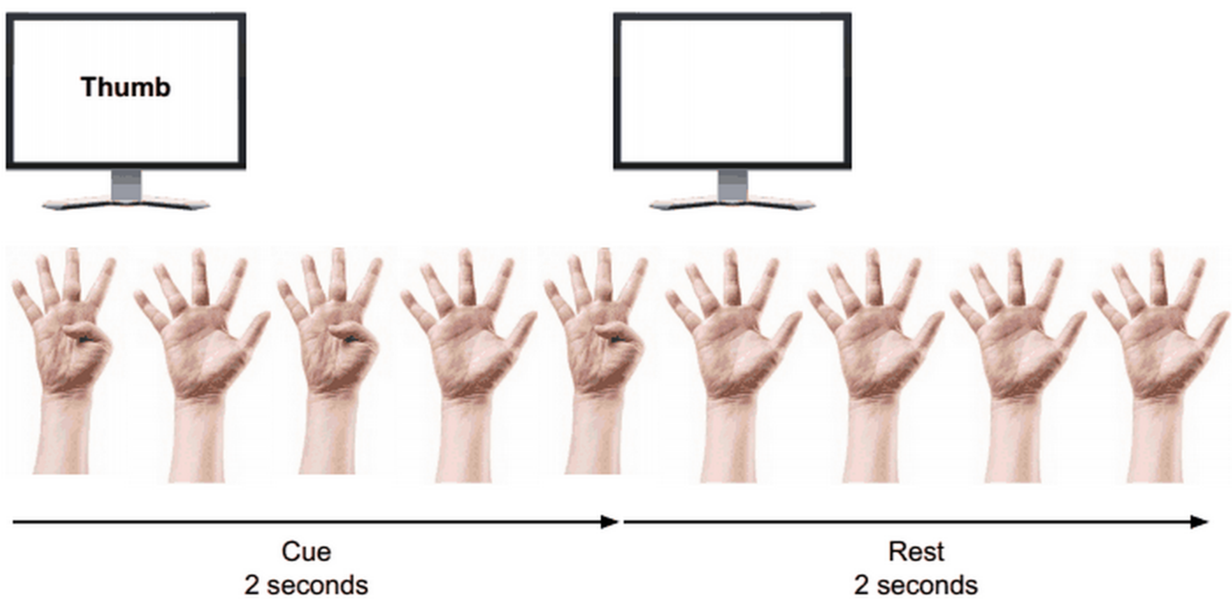


Figure 3: Cue/Rest sequence

Dataset is divided in two parts - training data and test data (see figure 4). Training data includes first 400 seconds of ECoG data and finger flexion data. Test data to be used for prediction contains last 200 seconds of ECoG data. Actual finger flexion data for last 200 seconds were not included in original dataset and become available only after end of competition (Miller & Schalk 2008).

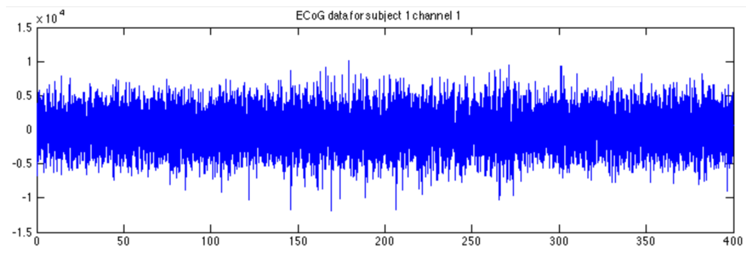
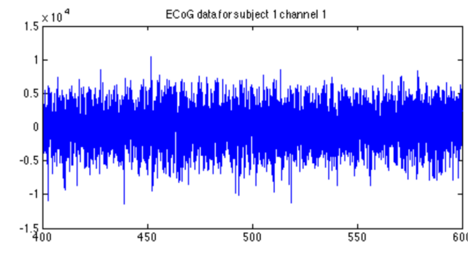
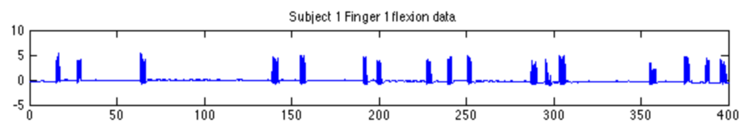
Training data (400 seconds)	Test data (200 seconds)
	
	???

Figure 4: Training and test data

Neural basis of finger movement

Brodmann area 4 (primary motor cortex) is area of human brain responsible for planning and executing movements. It is located in back section of frontal lobe (see figure 5).

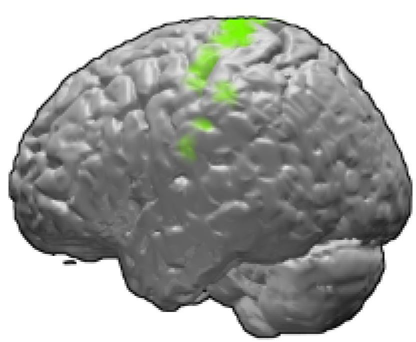


Figure 5: Brodmann area 4 (Source http://en.wikipedia.org/wiki/File:Ba4.png)

Studies using fMRI indicate that arms and hands generally follow somatotopic arrangement similarly to other body parts. Fingers are represented in overlapping manner, however, each finger may have cortical hot spot. These hot spots are arranged in somatotopic order. (Meier et al 2008). There is also indications that finger representation in cortex follow core-surround organisation (Strother el at 2012, Meier et al 2008).
Studies indicate that cortical hot spots of each finger is separated by very small distance (few millimeters). According to competition documentation “The electrodes had a diameter of 4 mm (2.3mm exposed), 1 cm inter-electrode distance” (Miller & Schalk 2008). It means that distance

between electrodes are generally much larger that distance between finger representations in brain.

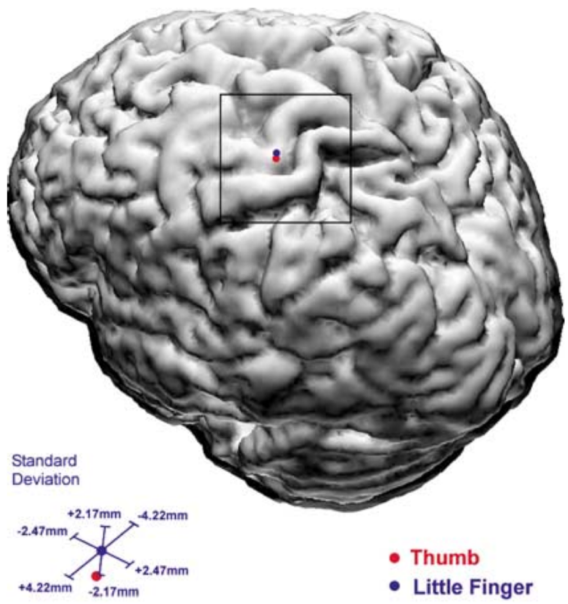


Figure 6: Location of thumb and little finger in primary motor cortex (Beisteiner et al 2004)

General sequence of neural events during each presentation of cue (name of finger) is shown on following diagram (see figure 7):

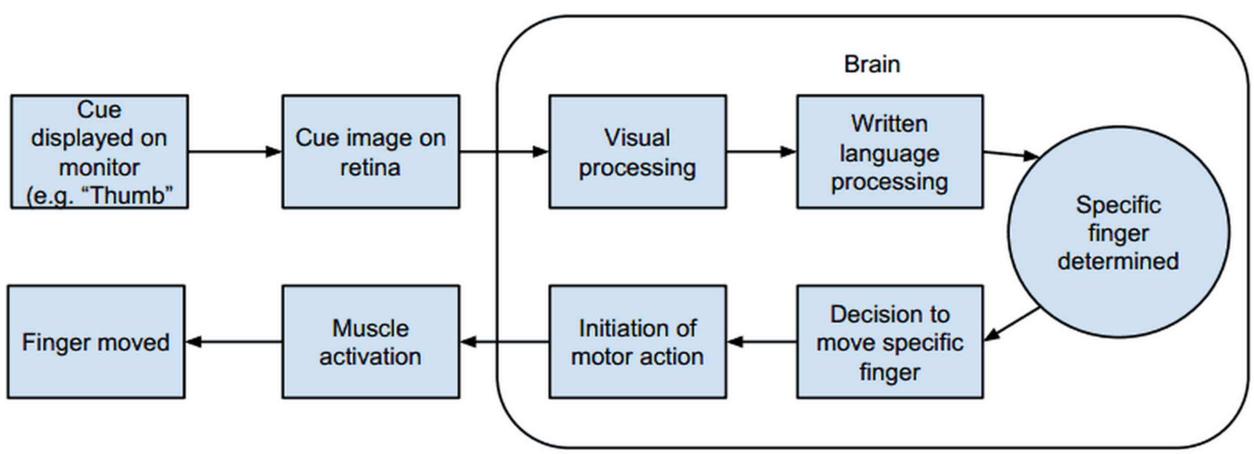


Figure 7: General sequence of sensory-motor loop

Each stage of sensory and motor processing might introduce it own delays.
It should be taken into account that during each cue phase:

- Each finger flexion includes 2 actions - flexing finger and unflexing finger, which uses different muscles and probably have separate neural circuits.
- Corresponding finger is flexed and unflexed 3-5 times.

- Occasionally other fingers are also moved.
- Sometimes subject forgets to stop flexing fingers after cue disappears from computer screen.

ECoG data

Electrocorticography (ECoG) data is obtained by placing grid of platinum electrodes (8x6 or 8x8) directly on surface of the brain and recording readings from these electrodes (Miller & Schalk 2008).

Subject	Number of channels	Probable electrode arrangement
1	62	8x8
2	48	8x6
3	64	8x8

Table 1: ECoG electrode arrangement

Training data (*train_data*) contains 400'000 samples at sampling rate 1000 Hz for each channel (first 400 seconds of ECoG readings). Test data (*test_data*) contains 200'000 samples at same sampling rate (last 200 seconds of ECoG readings) (Miller & Schalk 2008).

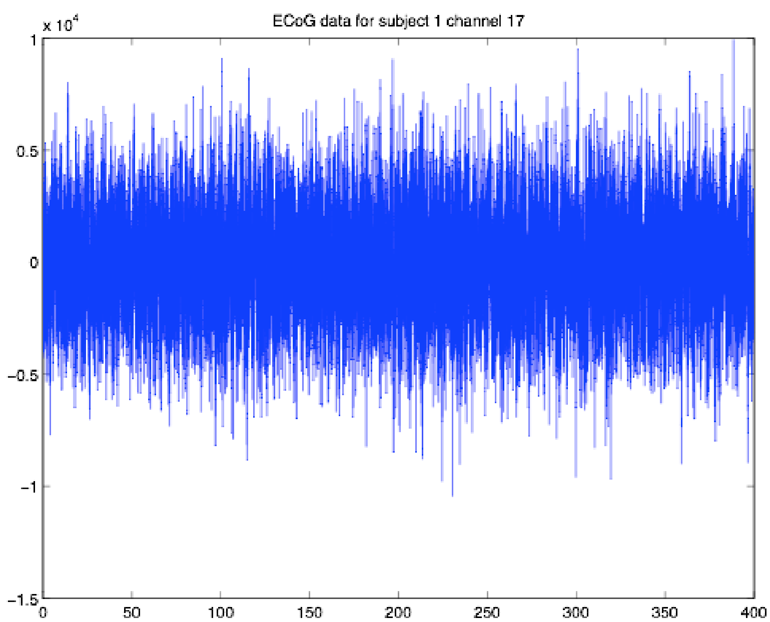


Figure 8: Sample of typical ECoG data (subject 1, channel 17, time 0-400s) (Matlab source code - ecog_plot.m)

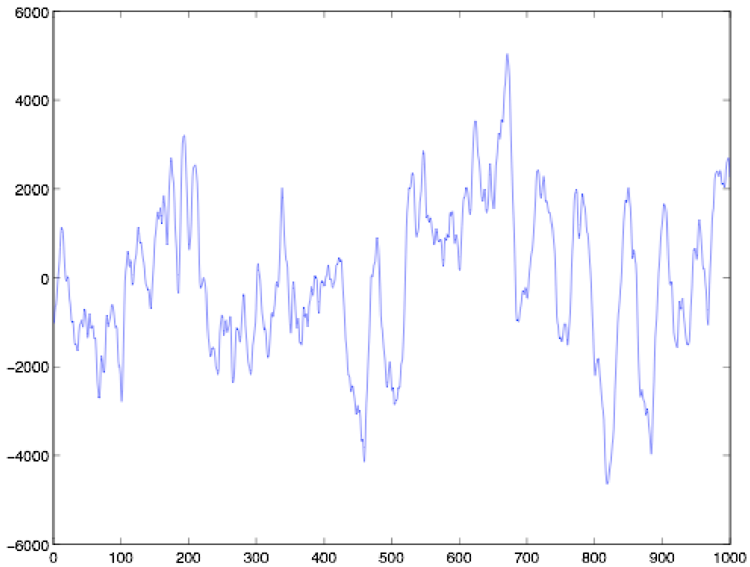
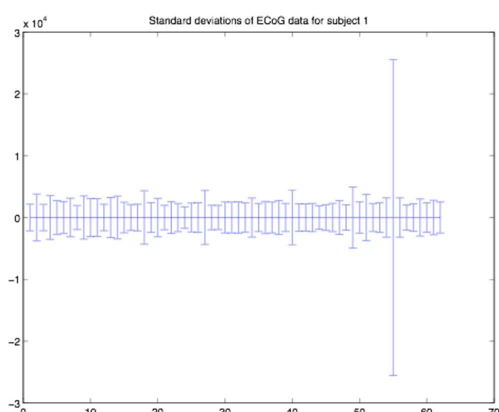
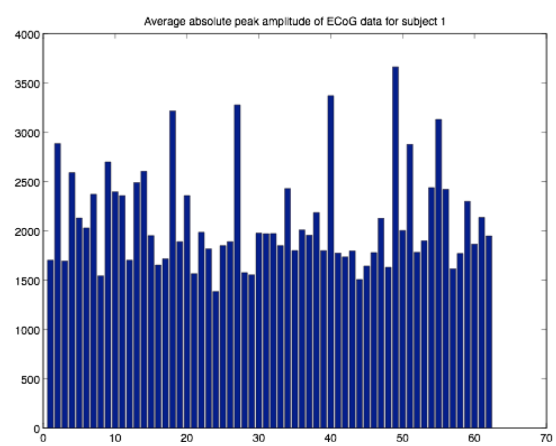


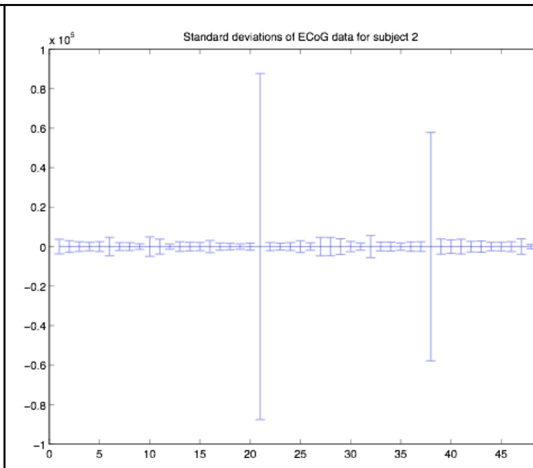
Figure 9: Sample of typical ECoG data in higher temporal resolution (subject 1, channel 17, time 123.000-123.999s)

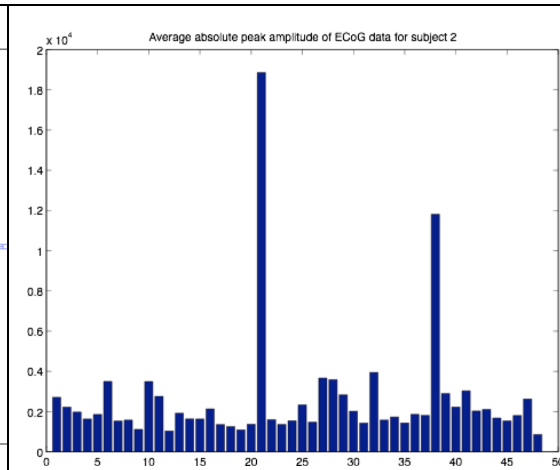
Properties of ECoG data

Standard deviations and average absolute peak amplitude were calculated for each channel of ECoG data:

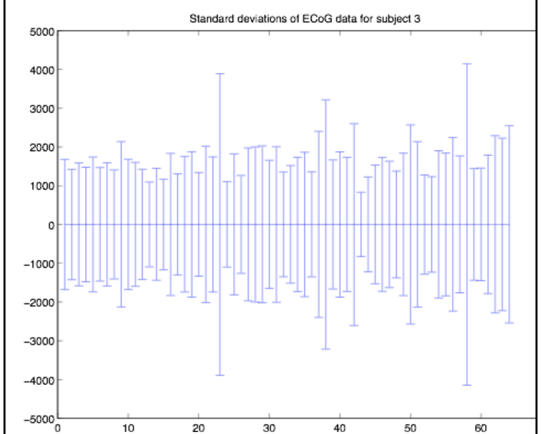
Subject	Standard deviations of ECoG data per channel	Average absolute peak amplitude of ECoG data per channel
#1		

#2





#3



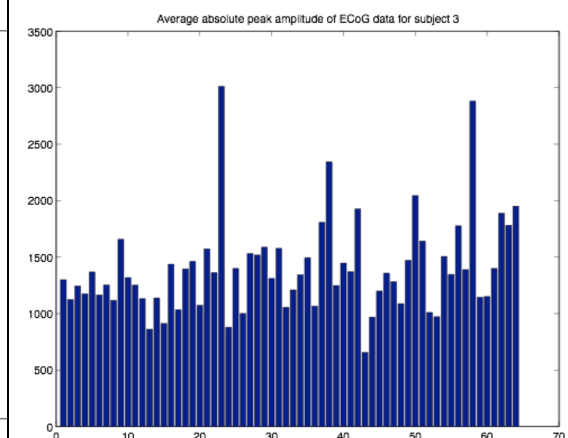
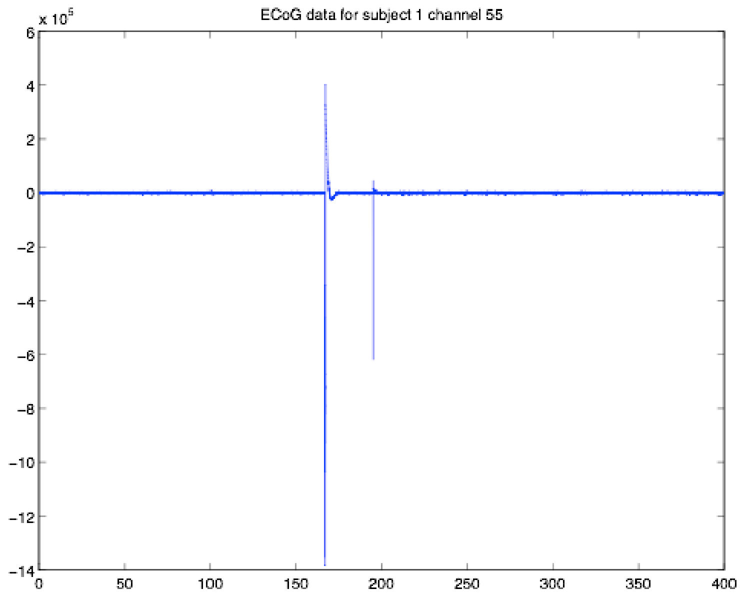
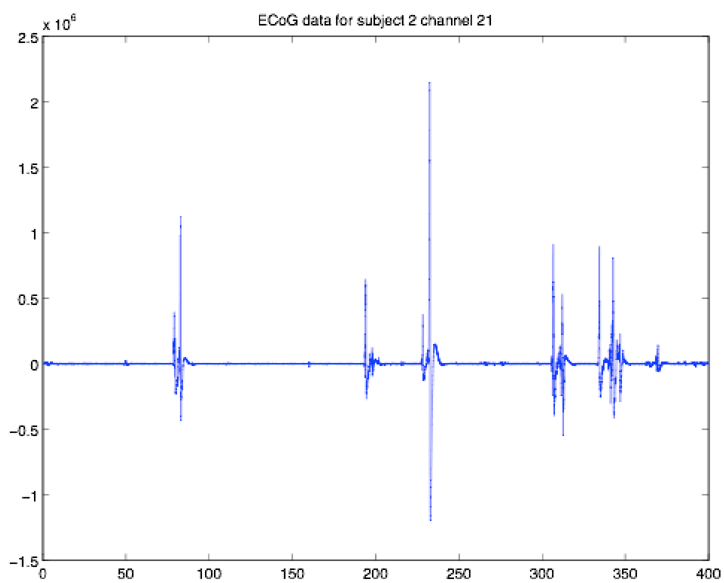


Figure 10: Standard deviations and average absolute peak amplitudes for all channels (Matlab source code - ecog_errorbar.m)

Plots of standard deviation and average absolute peak amplitude indicate that for each subject there are 1-2 channels which might contain outliers (e.g. erroneous measurements):

Subject	Channel	Plot of ECoG data
1	55	
2	21	

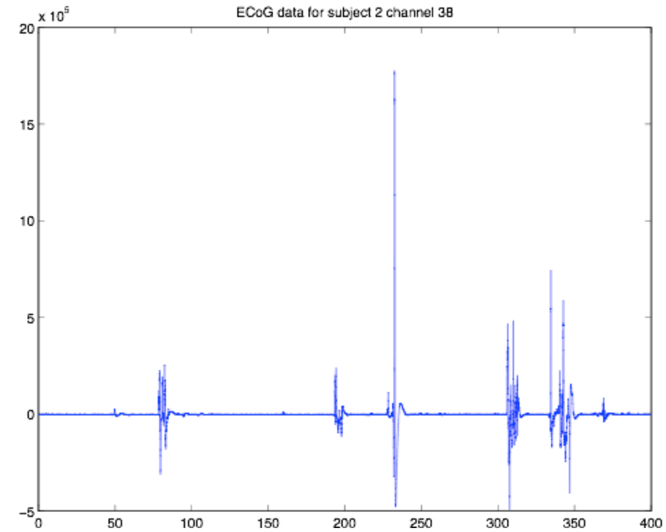
2	38	

Figure 11: Channels with possible outliers (Matlab source code - ecog_plot.m)

ECoG data of channel #55 for subject #1 and channels #21, #38 for subject #2 seem to contain outliers. These channels probably should be excluded then performing machine learning. Although some channels of subject #3 exhibit greater variability there seems to be no significant abnormalities in distribution of values for these channels.

Correlations among ECoG channels

Correlation analysis was performed to determine how related is data obtained from each channel.

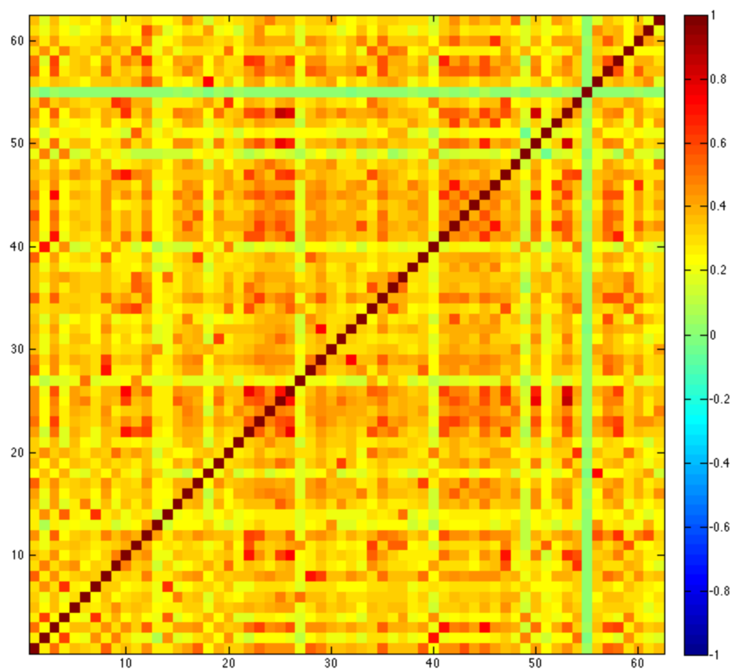


Figure 12: Correlations among channels for subject #1 (average correlation 0.3297, standard deviation 0.1262*) (Matlab source code - channel_corr.m)

* Note: Average correlation is calculated excluding equal channels (with correlation 1.00).

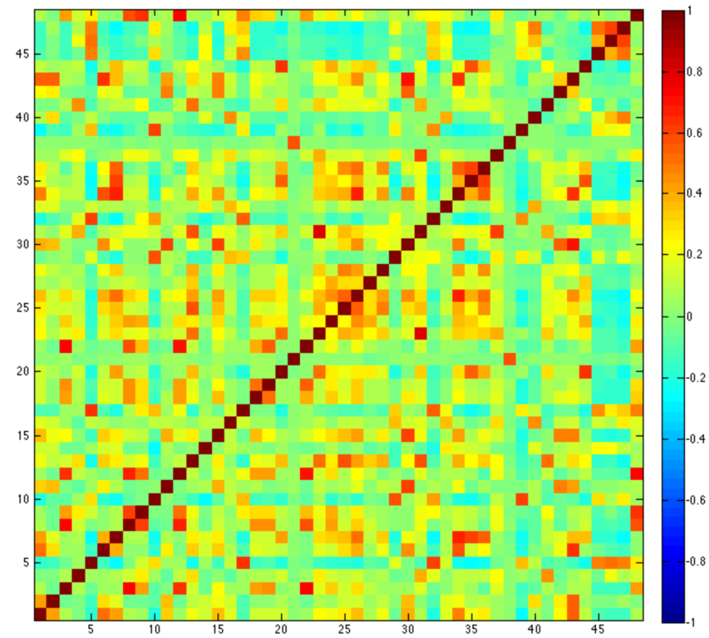


Figure 13: Correlations among channels for subject #2 (Average correlation 0.0924, standard deviation 0.1946) (Matlab source code - channel_corr.m)

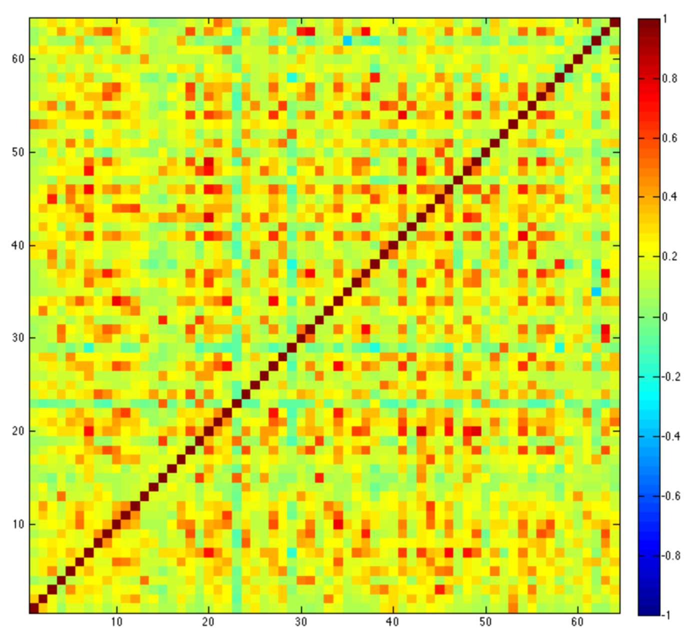


Figure 14: Correlations among channels for subject #3 (Average correlation 0.2106, standard

deviation 0.1530) (Matlab source code - channel_corr.m)

Correlation analysis shows that data obtained from each channel has low to average positive correlation with most other channels.

Finger flexion data

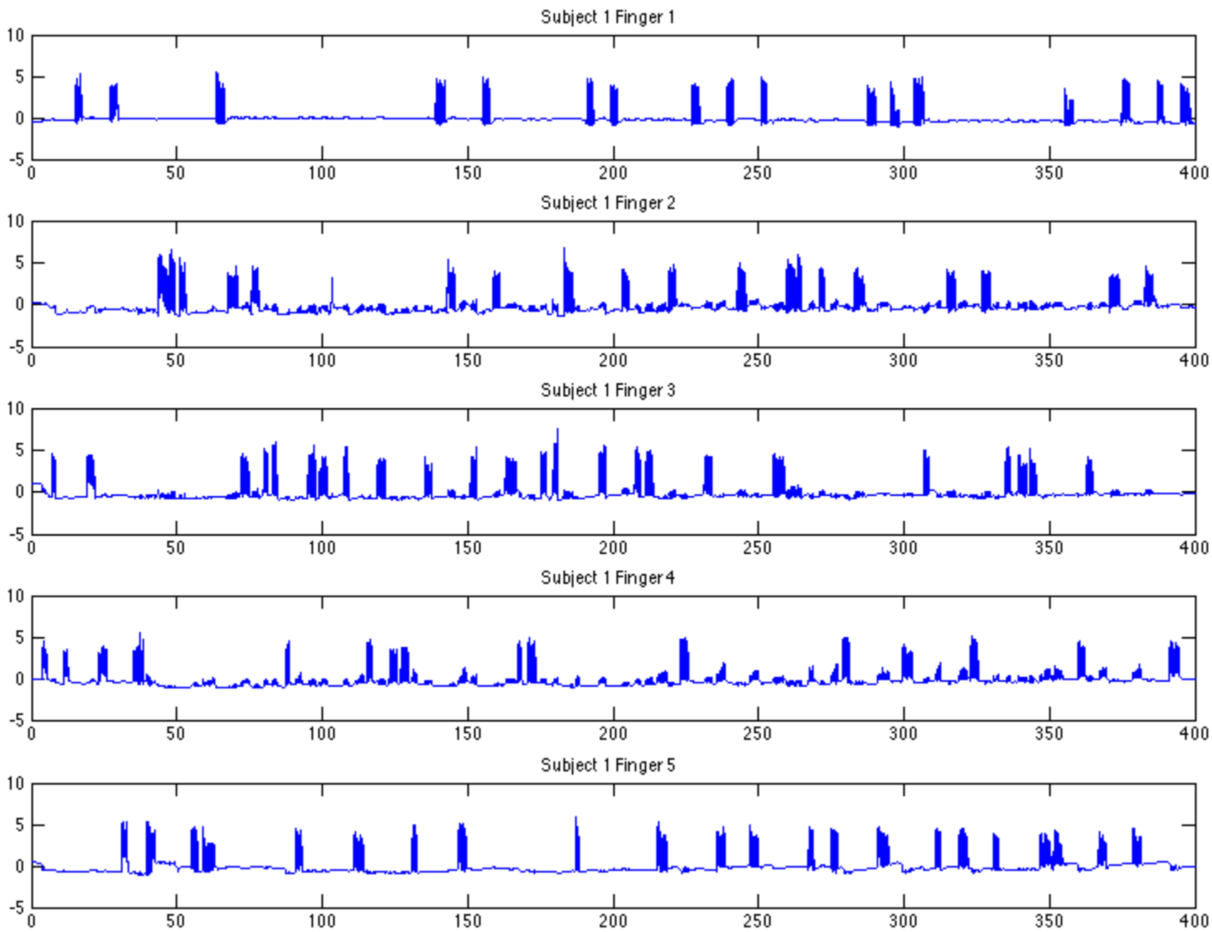


Figure 15: Subject #1 finger flexion data (Matlab source code - finger_flexion_plot.m)

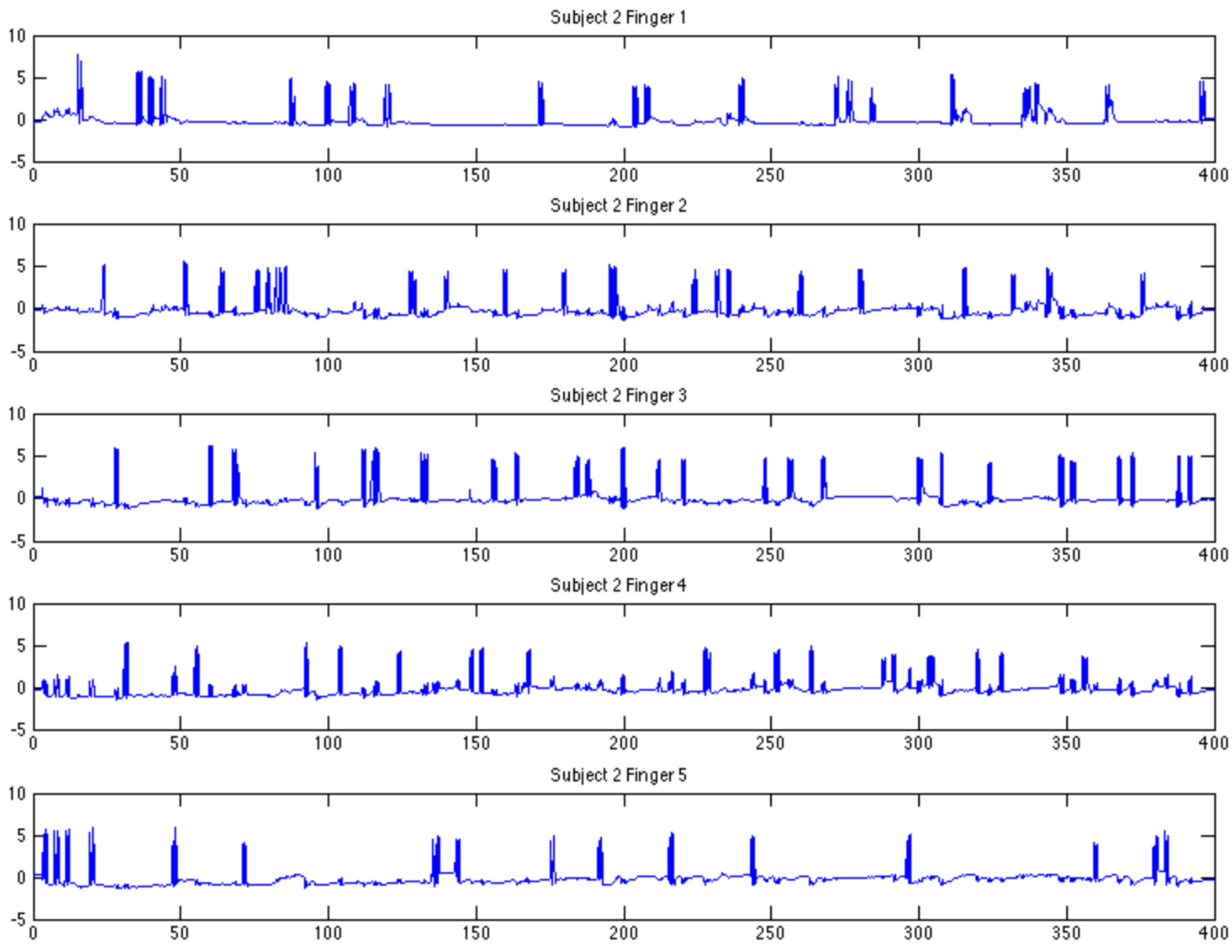


Figure 16: Subject #2 finger flexion data (Matlab source code - finger_flexion_plot.m)

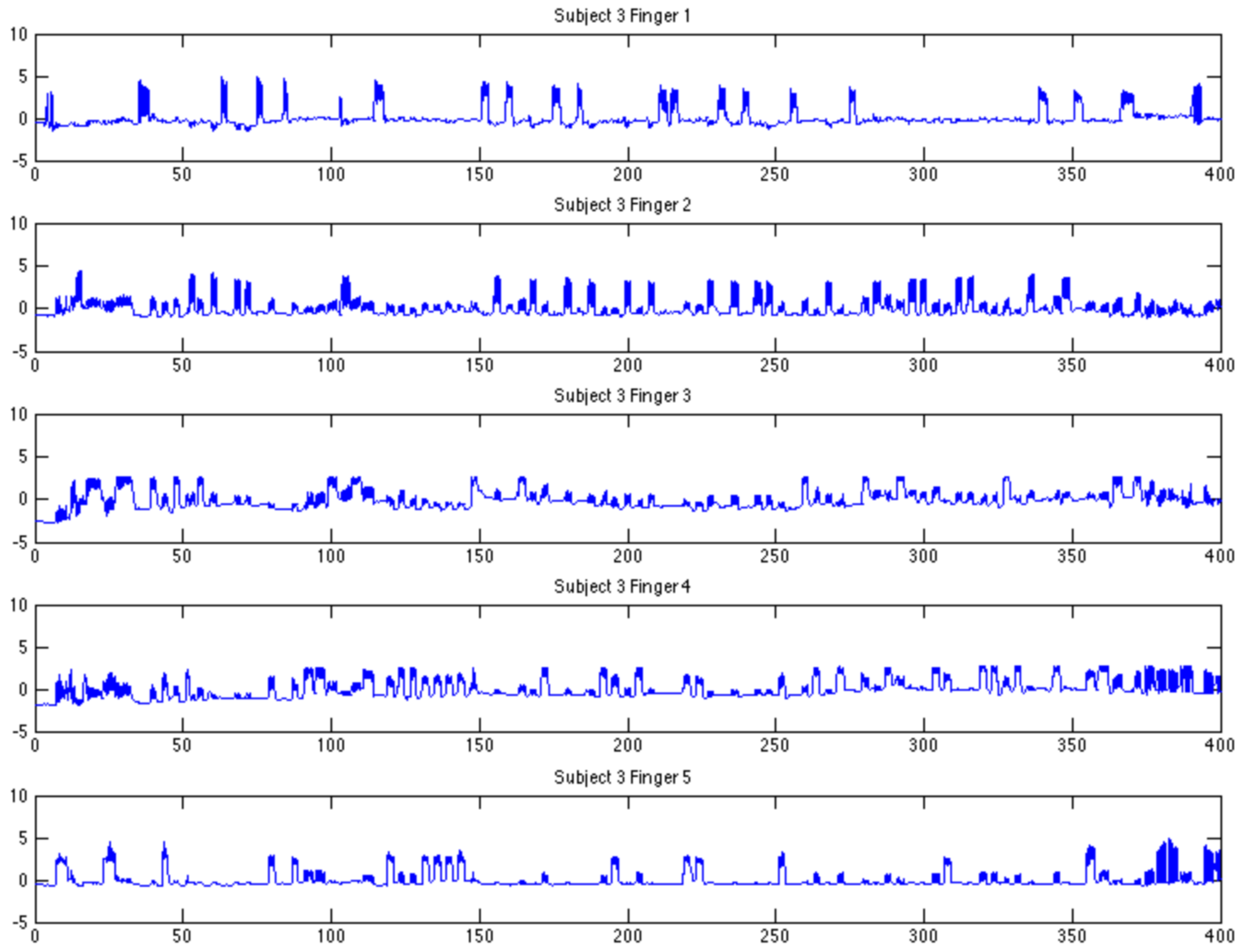


Figure 17: Subject #3 finger flexion data (Matlab source code - finger_flexion_plot.m)

According to documentation “During each cue, the subjects typically moved the requested finger 3-5 times” (Miller & Schalk 2008). It can indeed be seen in finger flexion data, looking at higher temporal resolution:

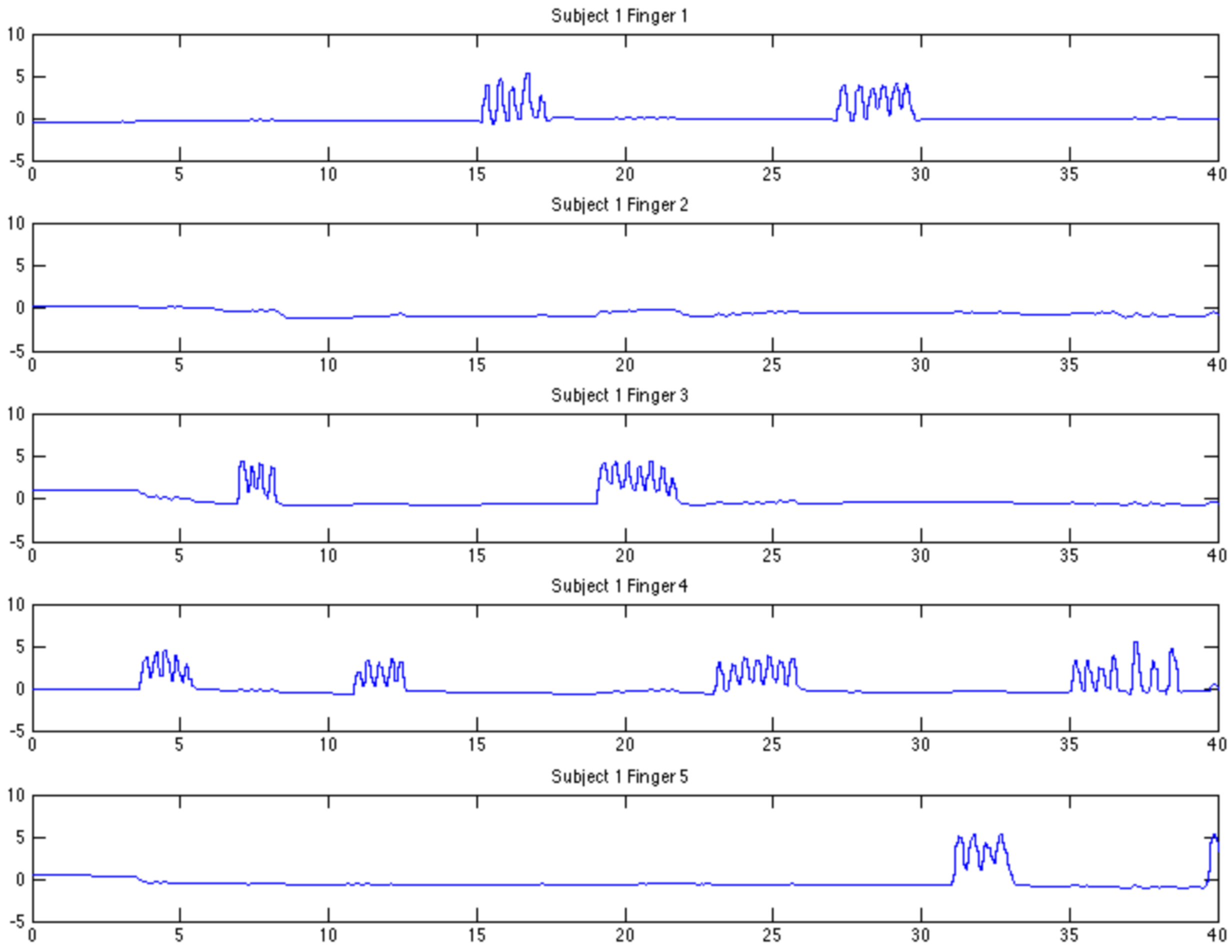


Figure 18: Subject #1 finger flexion data in higher temporal resolution (0-40s)

Cumulative finger flexion data

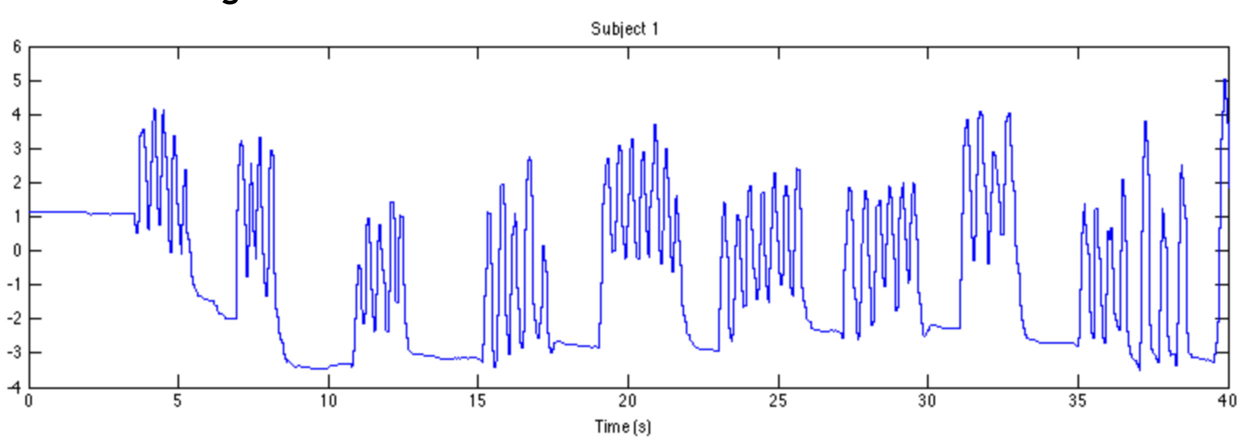


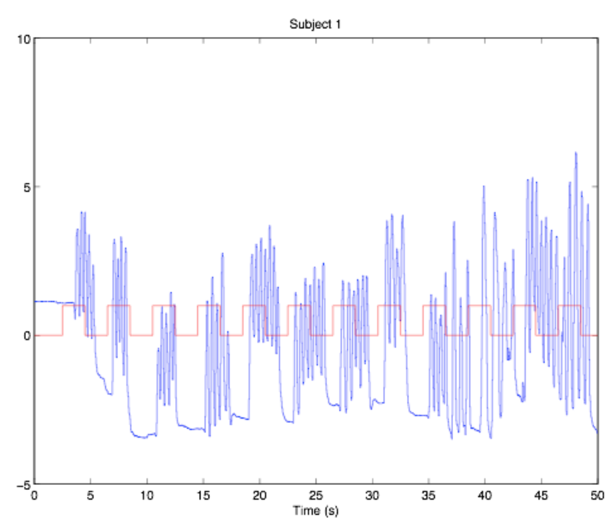
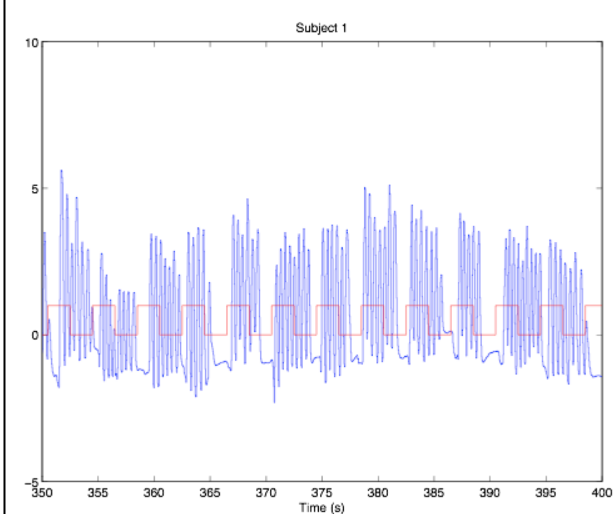
Figure 19: Cummulative finger flexion data (sum of all finger flexion data) for subject #1 (0-40s) (Matlab source code - all_finger_flexion_plot.m)

Cue and rest period

According to documentation “The subject were cued to move a particular finger by displaying the corresponding word (e.g. “thumb”) on a computer monitor place at the bedside. Each cue lasted two seconds and was followed by a two-second rest period during which the screen was blank.” (Miller & Schalk 2008)

By plotting cumulative finger flexion data with repeated cue-rest period (red curve) it can be seen that finger flexions indeed roughly follow alternating 2 second activation and 2 second rest pattern. In some cases fingers are being flexed even after cue is no more displayed (probably subject forgot to stop).

Note: To better match finger flexion data first cue is displayed at 2500 ms offset, but probably real offset is around 2000 ms (because it takes time for subject to understand cue and react by flexing fingers).

Cumulative finger flexion data (0-50s)	**Cumulative finger flexion data (350-400s)**
	

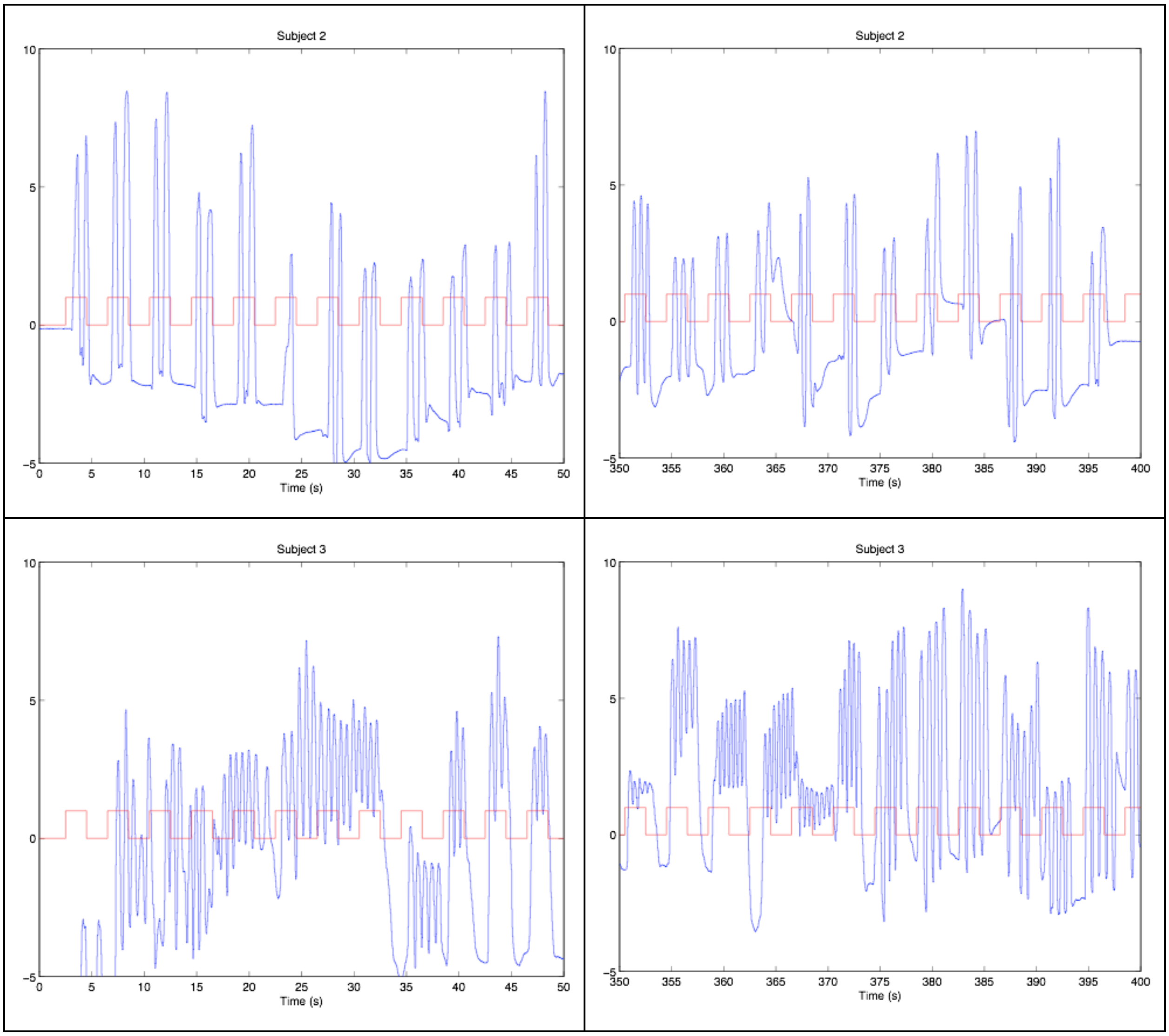


Figure 20: Cue and rest periods for all subjects (beginning and end of training data) (Matlab source code - all_finger_flexion_plot.m)

Known cue-rest period also means that predicting finger flexion more weight can be assigned to cue period and less weight to rest period.

Correlations among fingers

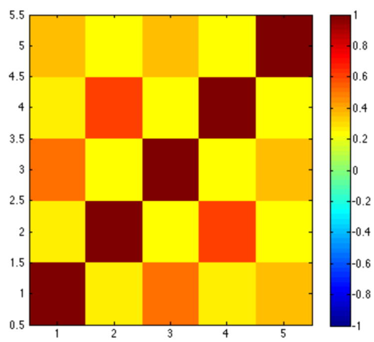
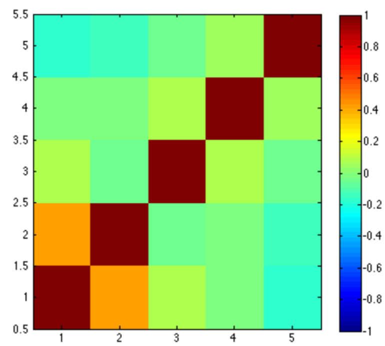
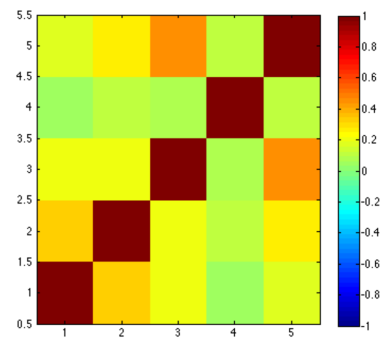
Subject #1	Subject #2	Subject #3
		
Average 0.3282 Standard deviation 0.1362	Average 0.0177 Standard deviation 0.1674	Average 0.2004 Standard deviation 0.1267

Figure 21: Correlations among fingers (Matlab source code - finger_flexion_correlations.m)

#	**Name**	**Average correlation for subject #1**	**Average correlation for subject #2**	**Average correlation for subject #3**	**Average correlation for all subjects**
1	Thumb	0.3478	0.0773	0.1855	0.2035
2	Index finger	0.3316	0.0581	0.2264	0.2053
3	Middle finger	0.3316	0.0195	0.2467	0.1992
4	Ring finger	0.3399	0.0187	0.0883	0.1489
5	Little finger	0.2903	-0.0851	0.2550	0.1534
	Average for all fingers	0.3282	0.0177	0.2004	

Table 2: Average correlations among fingers

Although in competition documentation it is said that finger #4 is excluded from evaluation due to high correlation with fingers #3 and #5 (Miller & Schalk 2008), this correlation analysis shows that there is nothing special about finger #4. Instead thumb and index fingers seem to be most correlated with other fingers.

Window size

There is tradeoff between window size and correlation between predicted and actual finger flexion data. If finger flexion data is averaged using window of certain size and then restored back to original size, some data is lost (see table 3):

Window size, ms	Correlation
100	0.983723
200	0.952892
500	0.855633
1000	0.798057
2000	0.655488

Table 3: Average correlation between window averaged finger flexion data restored to original size and actual data (Matlab source code - window_averaged_finger_corr_with_actual.m)

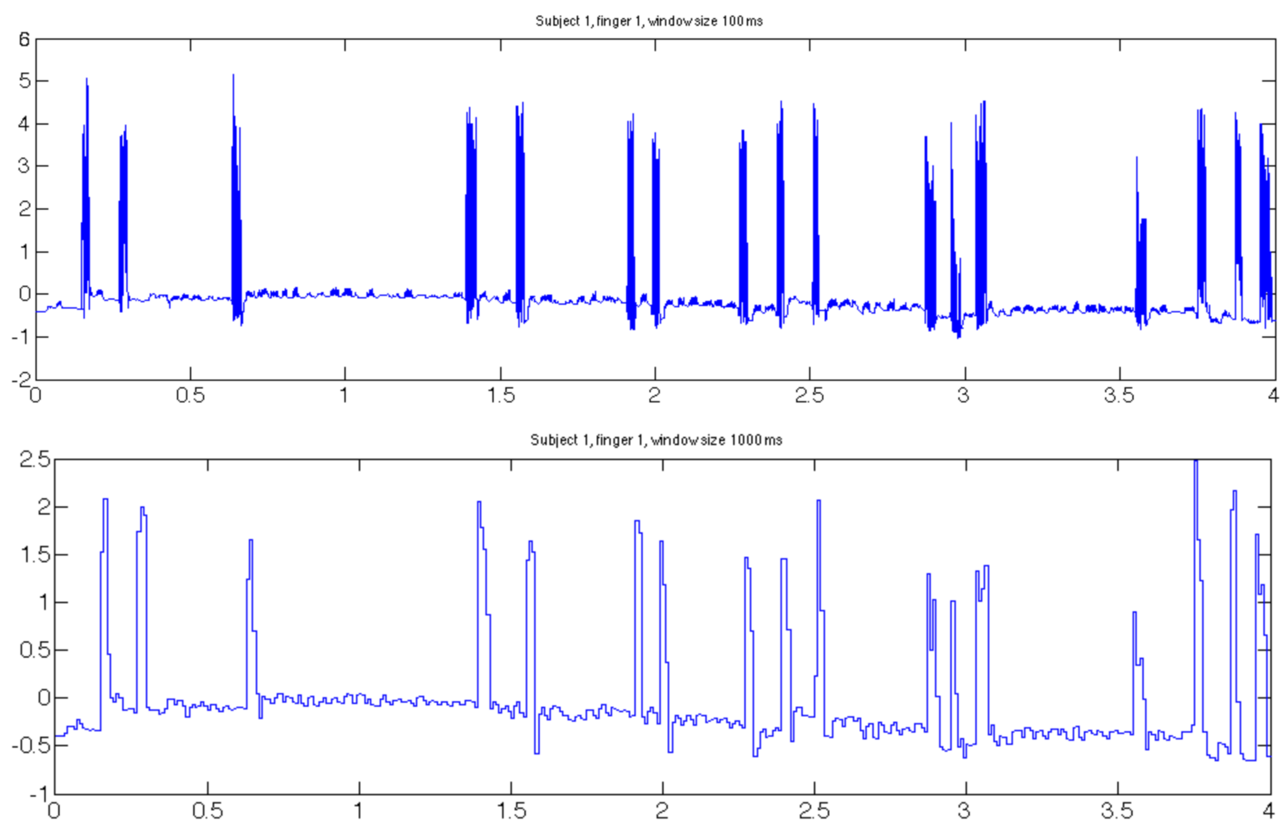


Figure 22: Averaged finger flexion data using window of size 100 ms (upper plot) and 1000 ms (lower plot). (Matlab source code - window_averaged_finger_corr_with_actual.m

It means that even perfect model will not give better result, if window of certain size is used. Correlation can be slightly increased if smoothing is used (with moving average).

Relations among ECoG data and finger flexion data

Average correlations among ECoG channels and fingers

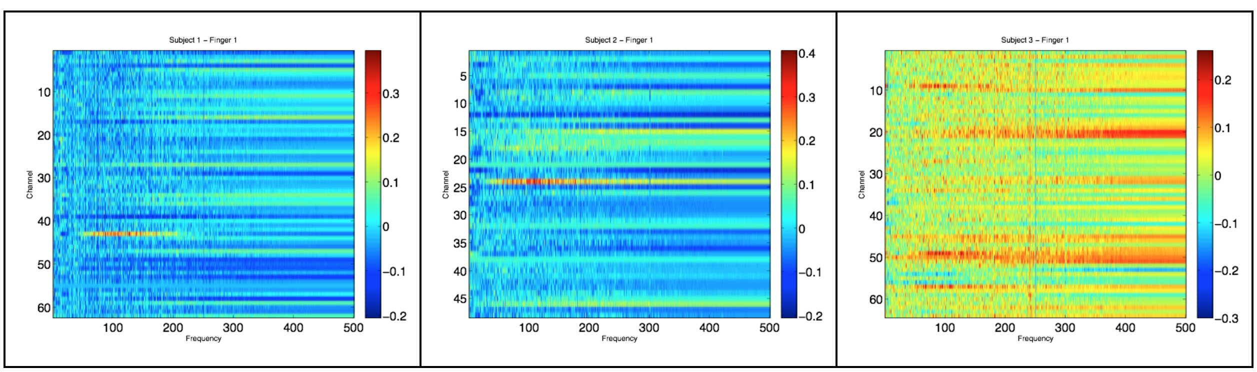
Subject	Average correlation among ECoG channels	Average correlation among fingers
1	0.3297	0.3282
2	0.0924	0.0177
3	0.2106	0.2003

Table 4: Correlations among ECoG channels and fingers

It seems that there is relation between average correlations of ECoG channels and finger flexion. Possible explanation - subject #2 is able to move each finger independently and it is also reflected in his/her ECoG measurements, while subject #1 and subject #2 tend to move fingers together (like most people do) and therefore their ECoG measurements are more correlated. From this perspective subject #2 could be better choice to explore how individual finger flexion appears in ECoG data.

Correlations among channels and frequencies

Correlations among ECoG FFT data for each channel and frequency and finger flexion data were obtained for each subject. Original ECoG data and finger flexion data were averaged using 1000 ms window (see Figure 23).



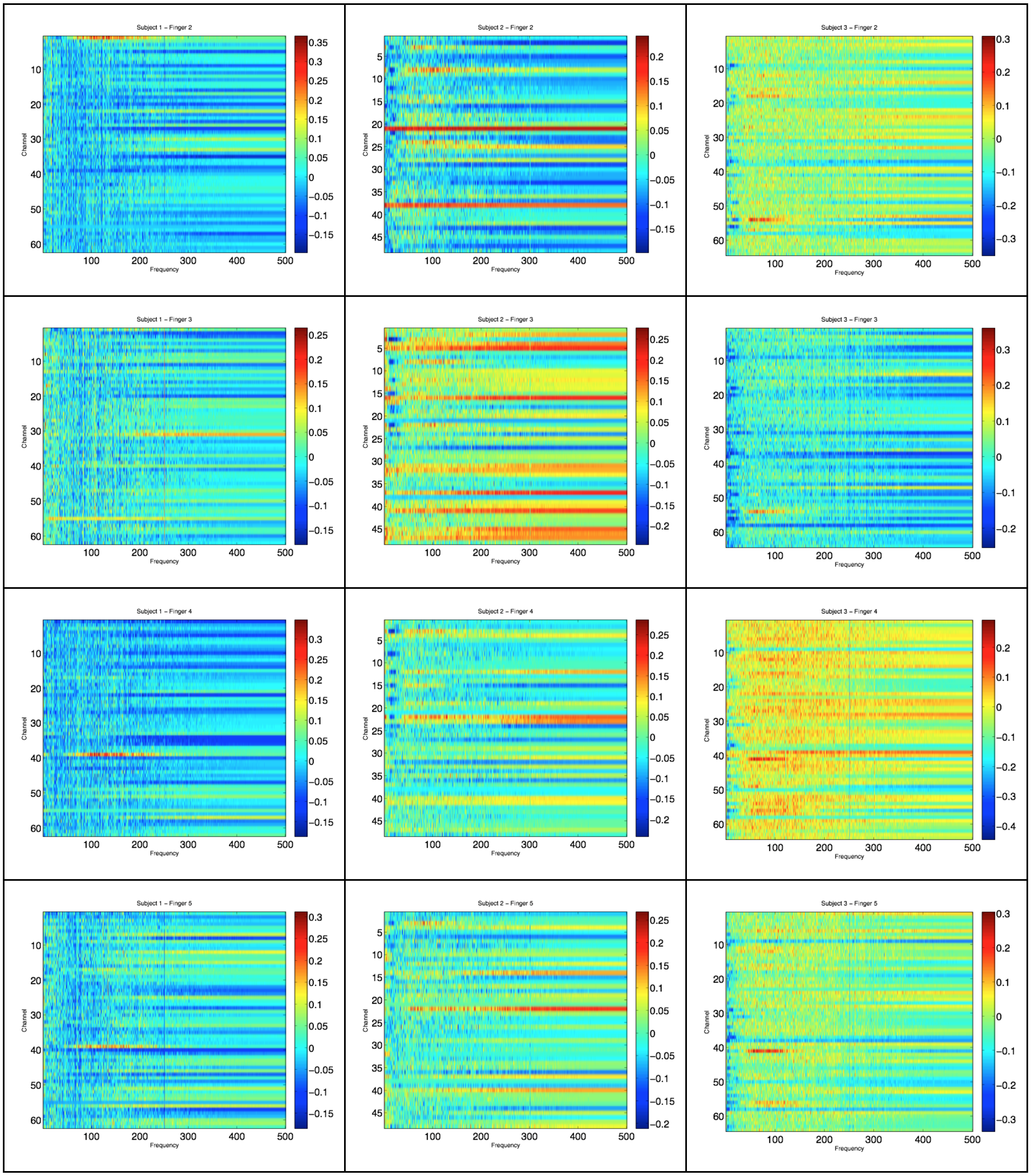


Figure 23: Correlations among ECoG FFT data and finger flexion data, window size 1000ms (Matlab source code - channel_freq_corr.m)

For most subject/finger combinations where is at least one channel correlated with finger flexion data (correlation around 0.2-0.4). Subject #1 has well defined specific channel and frequency range (50-200 Hz) for each finger (except finger #3). Subject #2 has more channels for each finger and frequency range is also wider (although previous analysis indicated that subject #2

can move his fingers more independently). Subject #3 has less pronounced correlation differences, but still for most fingers there are specific significant channels and frequency ranges.
Generally it means that flexion of specific finger is accompanied by increase in power for some frequency range in specific channel.
For different subjects different channels correspond to each finger. It can be explained by taking into account that ECoG electrode grid (8x8 for subjects #1 and #3, 8x6 for subject #2) is placed on slightly different regions of brain for each subject and brain structure/wiring can also differ for each subject.
The largest correlation obtained with this approach is for subject #2 - ECoG data from channel #24 and frequency 110 Hz has correlation 0.4058 with finger #1 flexion data. Other frequencies of this channel also has similar correlations (133 Hz - 0.4042, 88 Hz - 0.3788 etc.).
It should be mentioned that results are sensitive to window size. For example, changing window size from 1000ms to 500ms or 2000ms still gives the same channel #24 for subject #2 finger #1, but other frequencies are associated with the largest correlation.

Closer look at finger flexion and corresponding ECoG data

Subject #2 finger #1 has medium correlation (around 0.4) with ECoG data from channel #24 in frequency range from 50 to 200 Hz.

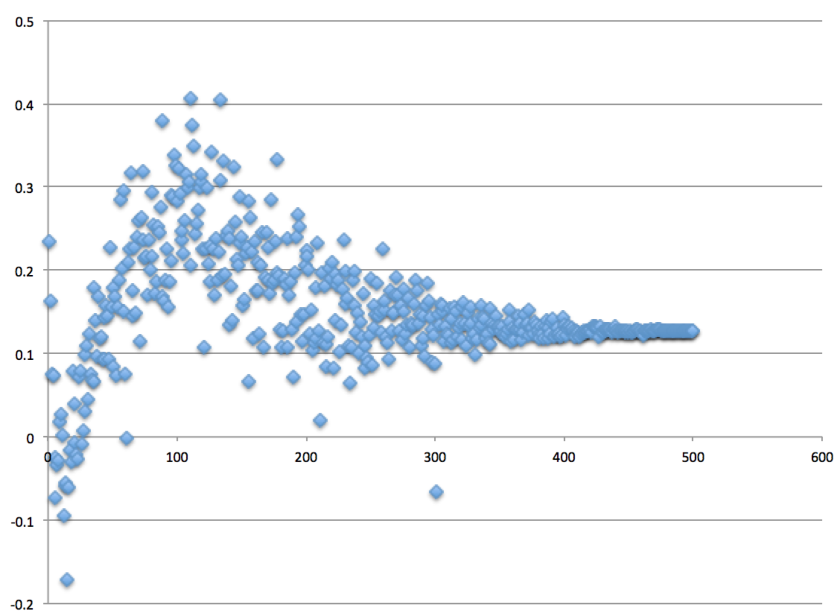


Figure 24: Distribution of frequencies (x-axis) versus correlation (y-axis) for subject #2, finger #1, channel #24

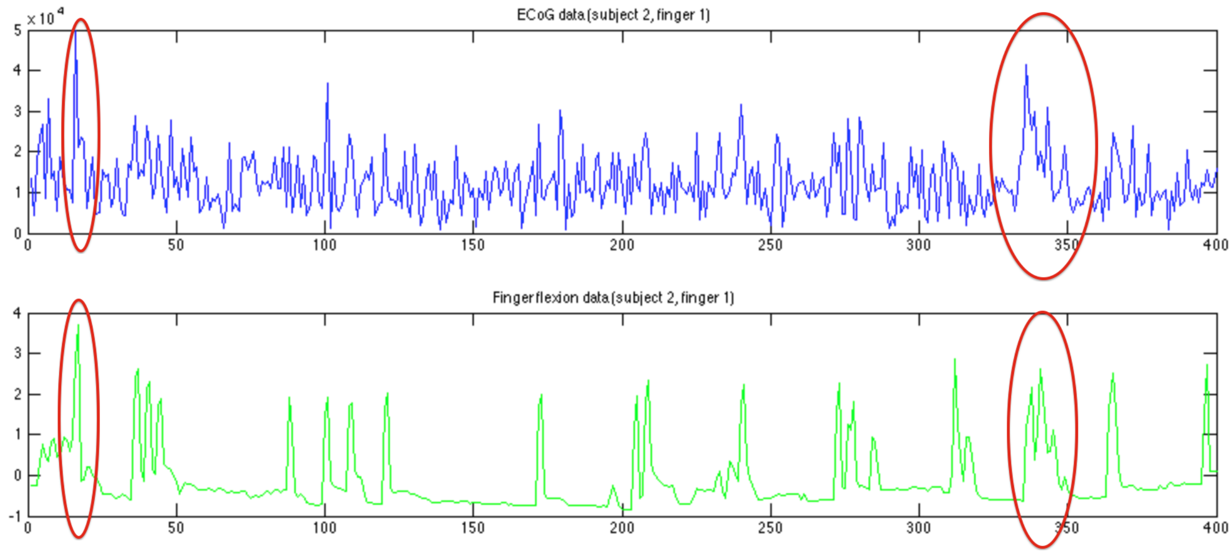


Figure 25: ECoG and finger flexion data: subject #2, finger #1, channel #24, window size 1000 ms, frequency 110 Hz, correlation 0.4058 (Matlab source - ecog_freq_finger_flexion_plot.m)

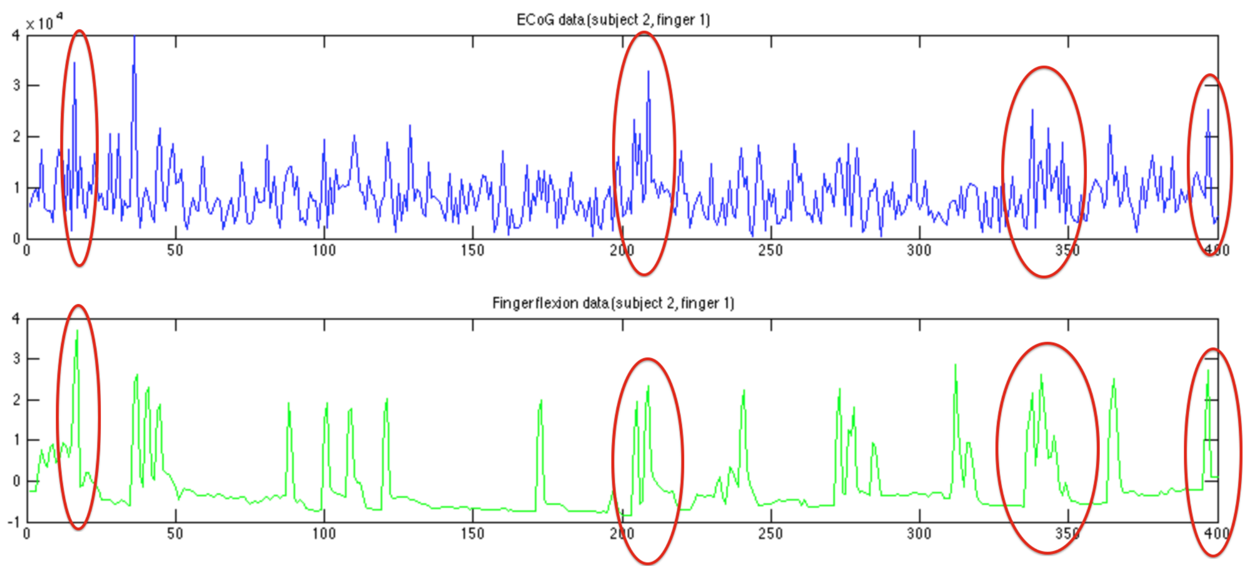


Figure 26: ECoG and finger flexion data: subject #2, finger #1, channel #24, window size 1000 ms, frequency 133 Hz, correlation 0.4042 (Matlab source - ecog_freq_finger_flexion_plot.m)

If best 20 frequencies are combined, much better correlation can be achieved:

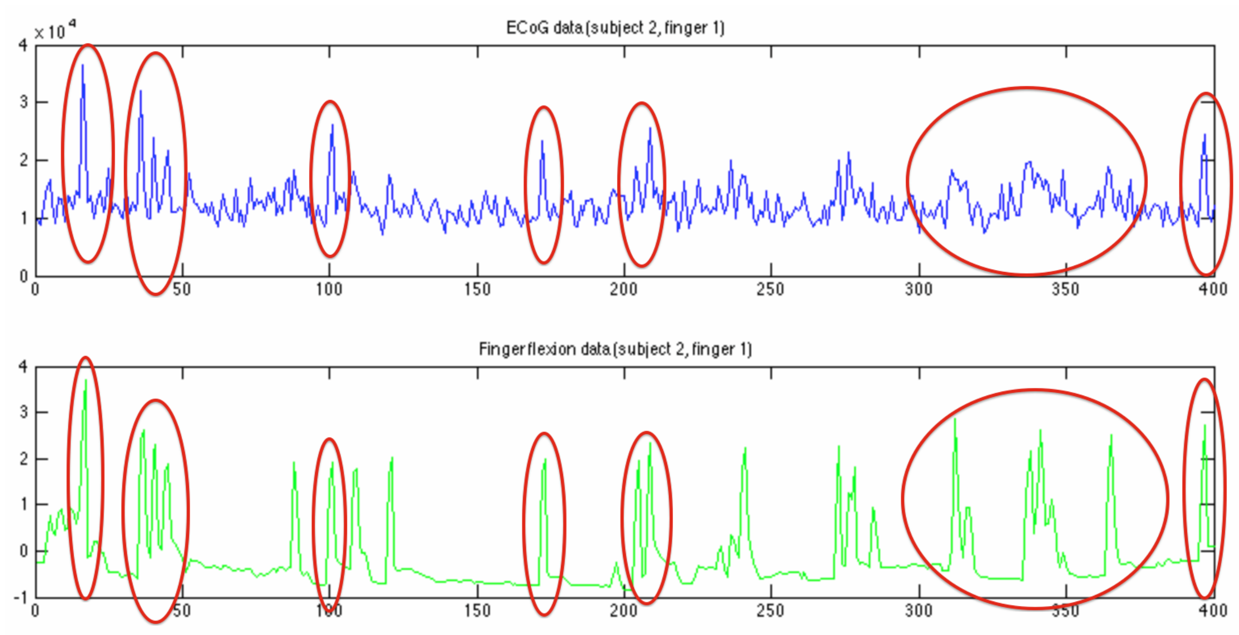


Figure 27: ECoG and finger flexion data: subject #2, finger #1, channel #24, window size 1000 ms, combined 20 best frequencies (110, 133, 88, 112, 113, 127, 98, 177, 136, 99, 144, 101, 73, 64, 119, 107, 134, 118, 109 Hz), correlation 0.6869 (Matlab source - ecog_freq_finger_flexion_plot.m)

Plotting ECoG frequency data versus finger flexion data it can be observed that there is clearly increase in activity for all frequencies (75-200 Hz) when finger is flexed (see figure 28 and figure 29).

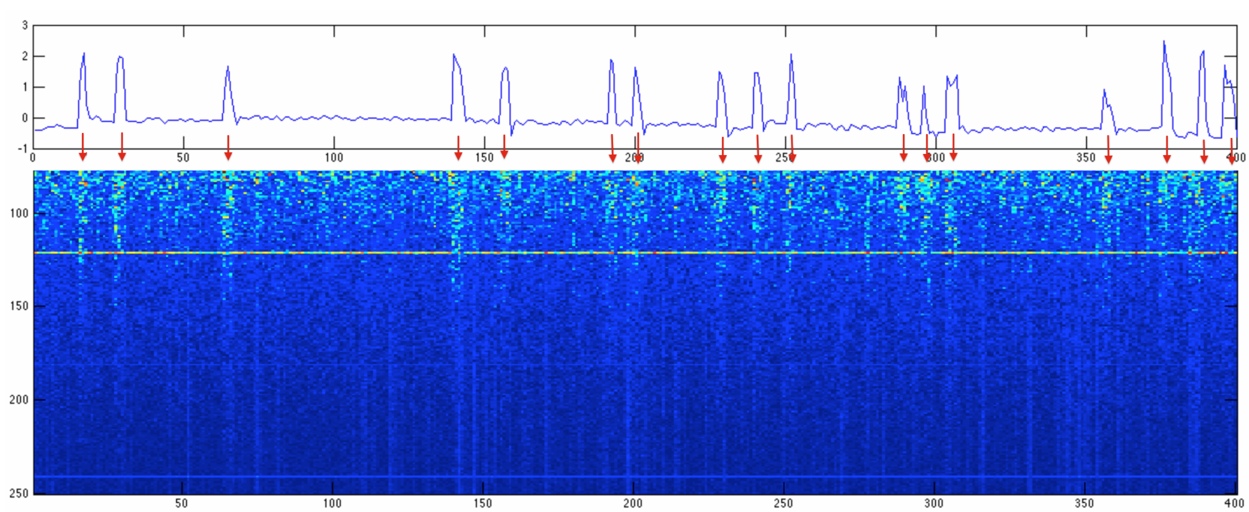


Figure 28: ECoG frequency activity and finger flexion data: subject #1, finger #1, channel #43, window size 1000ms

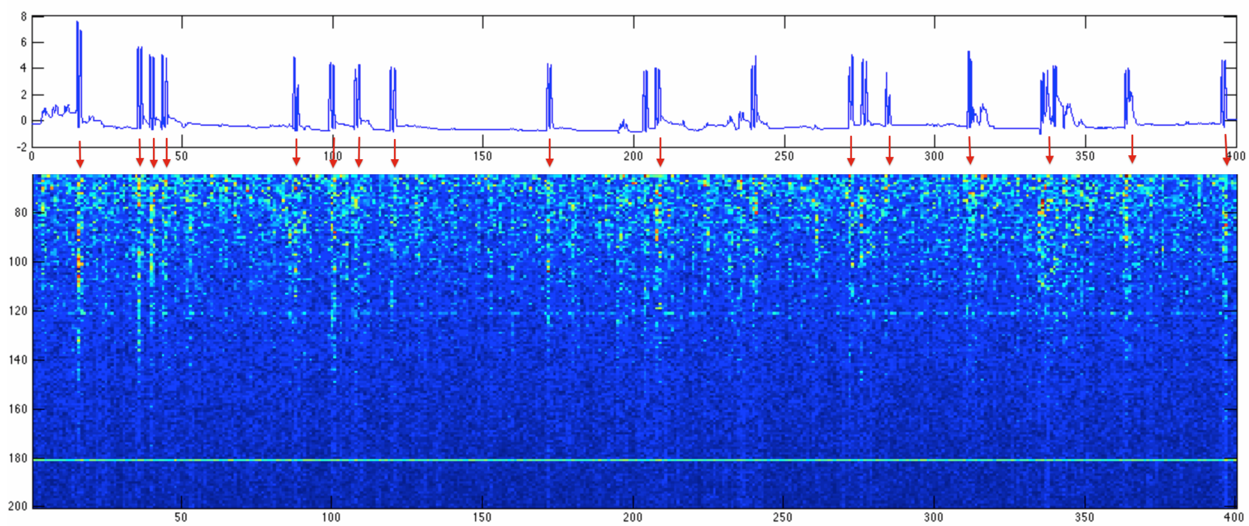


Figure 29: ECoG frequency activity and finger flexion data: subject #2, finger #1, channel #24, window size 1000ms

Prediction model

Following prediction model can be proposed, containing 3 stages:

- For each subject and finger calculate correlation among finger flexion training data and ECoG training data for each channel and frequency, using window of size 1000 ms. Store resulting correlations (table with columns: Channel, Frequency, Correlation) in separate file for each subject and finger (Matlab source code - channel_freq_corr.m).
- For each subject and finger identify N best channel-frequency pairs, which combined gives best correlation among finger flexion training data and ECoG training data (Matlab source code - identify_best_channel_freq.m). For example, for subject #1 and finger #1 combination of 47 channel-frequency pairs with highest correlations gives best correlation 0.692098 (on training data). See table 5 for results.
- For each subject and finger combine best channel-frequency pairs identified in previous steps to obtain predicted finger flexion data. Smooth predicted data using moving average. Calculate correlations among predicted and actual finger flexion test data for each subject-finger pair (except finger #4) to determine total average correlation. See table 6 for results. (Matlab source code - predict_by_best_channel_freq_2.m)

	Subject 1	Subject 2	Subject 2
Finger 1	47	37	41
Finger 2	40	2	48
Finger 3	4	23	49
Finger 4	21	20	44
Finger 5	32	13	49

Table 5: Number of channel-frequency pairs, which then combined gives best correlation for given subject-finger pair.

	Subject 1	Subject 2	Subject 3	Average
Finger 1	0.625197	0.512844	-0.062371	**0.358557**
Finger 2	0.620245	0.104221	0.351041	**0.358502**
Finger 3	-0.009404	-0.004246	0.466127	**0.150826**
Finger 4	0.540267	0.226995	0.481325	**0.416196**
Finger 5	0.292669	0.054888	0.630593	**0.326050**
Average (fingers 1,2,3,5)	**0.382177**	**0.166927**	**0.346348**	**0.298484**

Table 6: Correlations among actual and predicted test finger flexion data.

It seems that feature selection is important step in building model. Various other approaches were tried (without significant improvement):

- Use only channel-frequency pairs with correlation not less than best channel-frequency pair correlation multiplied by some ratio (e.g. 0.8).
- Determine minimal and maximal frequency for N best channel-frequency pairs and use all frequencies in this range.
- Determine average frequency for N best channel-frequency pairs and use all frequencies in 1-2 standard deviation range around average frequency (Matlab source code - predict_by_best_freq_range_5.m).

Different classification methods were also examined, but none give better results than simple combination of ECoG channel data:

- Linear regression
- Matlab “classify” function (Matlab source code - predict_by_classify_best_channel_4.m)

Conclusion

Performed data analysis clearly shows that for some fingers of subjects there are one or two corresponding ECoG channels, whose activity in certain frequency range could be used as reliable predictor of finger flexion. However result of prediction model is not very satisfactory and probably indicate that some other approach should be used for feature selection and machine learning. Assumption that finger flexion data can be directly derived from ECoG data might not be valid, because different degrees of finger flexion might correspond to same neural activity for muscle movement. Knowledge of experimental setup also could be used to improve predicted finger flexion data (for example, fact that there is 2 second cue-rest period and each finger is moved 3-5 times might be taken into account). Sufficient time for fine-tuning prediction model should also be reserved as some initially promising ideas might turn out dead-ends and developing alternative solutions take time.

References

- Beisteiner R., Gartus A., Erdler M., Mayer D., Lanzenberger R., Deecke L. (2004) Magnetoencephalography indicates ®nger motor somatotopy. European Journal of Neuroscience, Vol. 19, pp. 465-472, 2004
- Meier J., Aflalo T., Kastner S., Graziano M. (2008) Complex Organization of Human Primary Motor Cortex: A High-Resolution fMRI Study. J Neurophysiol. 2008;100:1800–1812. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2576195/
- Miller K.J., Schalk G. (2008) Prediction of Finger Flexion 4th Brain-Computer Interface Data Competition. Retrieved from http://bbci.de/competition/iv/desc_4.pdf
- Strother L., Medendorp P.,Coros A.M, Vilis T. (2012) Double representation of the wrist and elbow in human. European Journal of Neuroscience, Vol. 36, pp. 3291–3298, 2012. http://www.tutis.ca/PDFS/StrotherMedendorpEtAl_2012%20(2).pdf