Data Documentation

Date: 2014-11-23

### Description of files, and basic info about data

Pre processed data file with cleaned and formatted data from all participants: *wholedata.csv*

Summary information about the participants is in the file *participants.csv*

There are 27 participants, 10 controls and 17 synaesthetes.

10 controls and 10 synaesthetes did both rounds.

There are 7872 datapoints (without the system and practice trials).

Data structure is:

* **m** : month name
* **z** : front/back dimension - distance from Kinect device in mm
* **X** : left/right - negative numbers on the participant's left, 0 in front of Kinect, positive numbers on the participant's right side
* **y**: up/down - positive up / negative down / 0 - Kinect level
* **outlier** : coded as 0/1 where 1 means outlier
* **condition** : straight, head turn, body turn - coded as str/hd/tr
* **task** : month, crazy, mini
* **type**: control, synaesthete
* **participant** : participant code
* **rounds**: if participant did both rounds or only one, coded as - monthsOnly, crazyOnly, both

### Outliers

**System outliers**

Outliers related to system failures (failing to record hand position properly) were removed. All points with distance from the Kinect greater than 3000mm were treated as outliers - mostly meaning that the system failed to measure hand position and measured a distance to lab room wall behind the participant. Also all trials with 0 distance were removed - as system coded 0 when failed to measure the distance at all.

All system outlier trials are listed in the file *systemOutliers.csv*

There were 221 system outliers, which is 2.8074% of all data points.

**IMPORTANT: For Co1 months head condition all datapoints for September are outliers. This results in missing value for this month/condition in all subsequent analyses!**

**Standard deviation outliers**

Normal outliers were computed without "mini" task, because for each condition in mini there is only one datapoint per month/participant/condition.

For each datapoint we applied the following formula: - We computed average position for participant/task/condition (i.e all data points for Co2/month/hd) - measure of center of distribution in 3d space - Then we computed average distance of datapoints from that position (measure of spread in 3d space) - we used formula for Euclidean distance in 3d space. - Then we compared the distance of each data point to that average distance - All points that were more than 2.5 times away from the center than the average distance were considered outliers

List of those data points is in the file *advOutliers.csv*

There were 8 of them.

All these datapoints were were removed from further analysis.

### Data processing - basic summary statistics

Following summary statistics were computed:

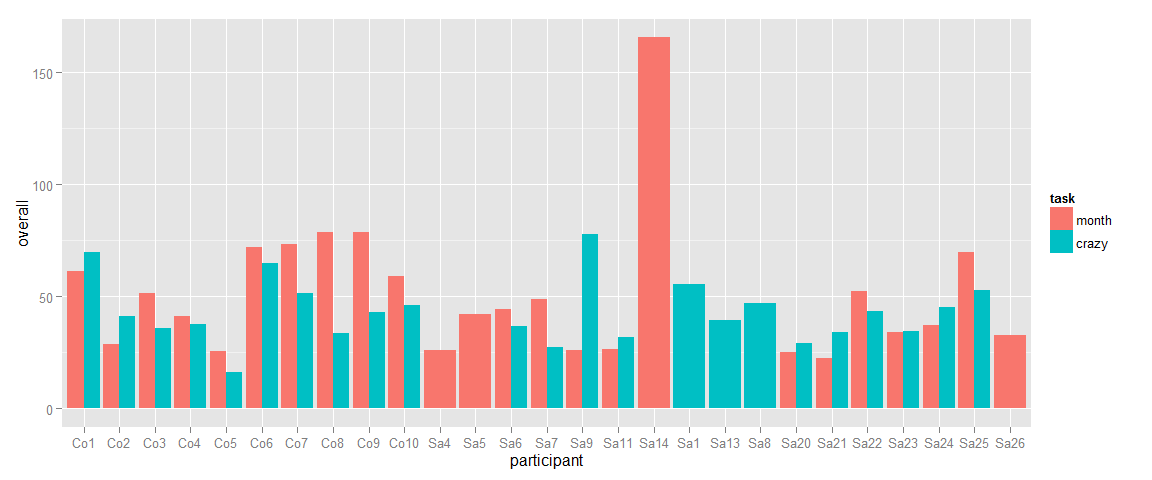
* mean positon for each month/condition/task/participant : data in file *means.csv*
* median positon for each month/condition/task/participant : data in file *medians.csv*
* standard deviation for each month/condition/task/participant : data in file *sds.csv*
* number of valid datapoints (after outlier removal) for each month/condition/task/participant : data in file *lengths.csv*
* average standard deviation (average of 12 separate month SD's) for each condition/task/participant : data in file *avofsds.csv*
* standard deviation of standard deviations (SD of 12 separate month SD's) for each condition/task/participant : data in file *sdofsds.csv*

Then, pooled SD's were computed for each participant/condition/task, by taking weighted average of each month SD's. The mini task was excluded, because there was only one data point per condition/month. Pooled SD data is in the file: *pooledsds.csv*

Summary of pooled SD data: "overall"" column is average of all the axes

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| condition | type | task | pooled\_z | pooled\_X | pooled\_y | overall |
| hd | control | crazy | 47.35 | 41.93 | 35.21 | 41.49 |
| str | control | crazy | 54.05 | 46.58 | 35.07 | 45.23 |
| tr | control | crazy | 45.37 | 43.70 | 36.06 | 41.71 |
| hd | synaesth | crazy | 45.73 | 53.48 | 32.93 | 44.04 |
| str | synaesth | crazy | 45.22 | 37.89 | 36.99 | 40.03 |
| tr | synaesth | crazy | 44.77 | 47.66 | 34.23 | 42.22 |
| hd | control | month | 59.81 | 71.67 | 46.30 | 59.26 |
| str | control | month | 57.16 | 52.29 | 42.23 | 50.56 |
| tr | control | month | 70.01 | 50.66 | 41.81 | 54.16 |
| hd | synaesth | month | 42.38 | 45.24 | 30.16 | 39.26 |
| str | synaesth | month | 43.38 | 36.40 | 32.85 | 37.54 |
| tr | synaesth | month | 46.58 | 36.62 | 29.38 | 37.53 |

Plot of pooled SD for each participant, averaged over conditions, and stacked by the task:



### Computing differences between conditions: straight vs. head, straight vs. trunk

Here the list of ANOVAS or t-tests.

Differences between mean position of each month/axis/condition/task/participant are in files:

*differences\_z\_month\_str\_hd.csv* *differences\_X\_month\_str\_hd.csv* *differences\_y\_month\_str\_hd.csv* *differences\_z\_month\_str\_tr.csv* *differences\_X\_month\_str\_tr.csv* *differences\_y\_month\_str\_tr.csv* *differences\_z\_crazy\_str\_hd.csv* *differences\_X\_crazy\_str\_hd.csv* *differences\_y\_crazy\_str\_hd.csv* *differences\_z\_crazy\_str\_tr.csv* *differences\_X\_crazy\_str\_tr.csv* *differences\_y\_crazy\_str\_tr.csv*

Results of t-tests and effect sizes are in the file *wholeDifferenceTests.csv*

There are 141 test results.

92 have p.value<0.05

77 have p.value<0.01

56 have p.value<0.001

### Model shift

We computed ideal expected distribution of data for each participant assuming given reference frame. Three possible reference frames were modelled - head centered, trunk centered and room centered.

If participant was using room RF, data points for each month in all conditions are expected to be placed in the same positons.

If head centered reference frame was used, participant's plot is expected to be shifted 45 deg to the right (from the plot in straight condition) in head turn trials, and remain close to straight plot in trunk turn trials.

If trunk centered RF was used, participant's plot is expected to be shifted 45 deg to the right in trunk turn trials, and remain close to straight plot in head turn trials.

NB: Distances between model and actual positions were computed on xz plane only.

## 45 deg expected position - adjustments

Adjustments to the above described models can be made - some systematic noise can arise because of the difference between center of rotaion of the chair and the participant head - for trunk turn trials, and the difference between center of rotaton of the head, and participant eyes/face in head turn trials)

Therefore, **if participant was using head centered RF**: in head turn trials, ideal predicted plot should be corrected for the distance between the eyes and center of rotation of the head.

In trunk turn trials, because of the distance between chair rotation point and participant's head, the ideal plot would not be expected at the same place as in straight condition, but offset to the left and forward.

**If participant was using trunk centered RF**: ideal predicted plot has to be corrected for trunk turn trials, taking into account the distance between chair rotation axis and participant trunk. In head turn trials, the plot should be at the same place as in straight condition.

## final model fit data

Adjusted and unadjusted model fit data were compared - using t-test for unadjusted best-fit and adjusted best-fit. There were no significant differences, therefore unadjusted algorithm was used to compute and analyze model fit

## main table (preliminary description)

All reference frame measures for each participant are listed in the file *mainTable.csv* The variables are (separately for month and horse tasks):

* head,trunk and room model fit
* ratio of best model fit / next to best (reversed to show how many times best fit is better than next to best fit)
* angle of regression line - difference between angle in straight and head condition
* angle of regression line - difference between angle in straight and trunk condition
* difference between those two differences

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| participant | head-fit-horse | head-fit-month | trunk-fit-horse | trunk-fit-month | room-fit-horse | room-fit-month | ratio-crazy | ratio-month | angle-str-hd-crazy | angle-str-hd-month | angle-str-tr-crazy | angle-str-tr-month | angle-diff-of-diffs-crazy | angle-diff-of-diffs-month |
| Co1 | 240.31 | 261.57 | 215.47 | 238.07 | 140.82 | 116.45 | 1.53 | 2.04 | 0.46 | 5.380 | 13.770 | 2.680 | -13.31 | 2.700 |
| Co10 | 230.13 | 200.05 | 92.64 | 74.77 | 136.57 | 123.95 | 1.47 | 1.66 | 6.66 | 2.700 | 51.450 | 45.150 | -44.79 | -42.450 |
| Co2 | 84.48 | 144.29 | 433.28 | 477.90 | 271.10 | 309.57 | 3.21 | 2.15 | 31.10 | 61.980 | 5.100 | 9.529 | 26.00 | 52.451 |
| Co3 | 132.74 | 85.87 | 476.99 | 462.58 | 263.51 | 270.81 | 1.99 | 3.15 | 53.46 | 24.250 | 6.900 | 7.020 | 46.56 | 17.230 |
| Co4 | 118.63 | 158.00 | 221.49 | 183.29 | 121.04 | 53.50 | 1.02 | 2.95 | 21.07 | -2.006 | 21.690 | 4.550 | -0.62 | -6.556 |
| Co5 | 274.75 | 237.42 | 292.62 | 258.71 | 79.89 | 87.19 | 3.44 | 2.72 | -2.07 | 0.960 | 9.980 | 18.520 | -12.05 | -17.560 |
| Co6 | 237.77 | 309.81 | 205.36 | 135.68 | 115.14 | 194.94 | 1.78 | 1.44 | -18.27 | 1.370 | 9.020 | 44.660 | -27.29 | -43.290 |
| Co7 | 163.38 | 144.17 | 212.57 | 365.61 | 121.42 | 225.08 | 1.35 | 1.56 | 17.10 | 44.290 | 21.250 | 1.740 | -4.15 | 42.550 |
| Co8 | 245.49 | 150.22 | 383.78 | 445.94 | 253.54 | 239.20 | 1.03 | 1.59 | -48.06 | 54.230 | 7.440 | 9.330 | -55.50 | 44.900 |
| Co9 | 245.23 | 162.63 | 232.43 | 348.08 | 105.13 | 204.95 | 2.21 | 1.26 | -6.20 | 6.590 | 8.410 | 5.390 | -14.61 | 1.200 |
| Sa1 | 239.34 | NA | 186.83 | NA | 89.63 | NA | 2.08 | NA | -5.10 | NA | 17.600 | NA | -22.70 | NA |
| Sa11 | 174.65 | 184.28 | 156.22 | 65.50 | 67.03 | 116.91 | 2.33 | 1.78 | 3.10 | 10.030 | 16.740 | 43.820 | -13.64 | -33.790 |
| Sa13 | 110.10 | NA | 463.20 | NA | 293.57 | NA | 2.67 | NA | 65.90 | NA | 2.900 | NA | 63.00 | NA |
| Sa14 | NA | 246.50 | NA | 100.75 | NA | 218.88 | NA | 2.17 | NA | 19.860 | NA | 61.760 | NA | -41.900 |
| Sa20 | 238.71 | 285.98 | 108.60 | 96.39 | 141.76 | 165.68 | 1.31 | 1.72 | -7.50 | 1.676 | 57.100 | 15.670 | -64.60 | -13.994 |
| Sa21 | 99.86 | 121.12 | 197.60 | 227.20 | 104.36 | 125.08 | 1.05 | 1.03 | 28.80 | 36.280 | 24.400 | 25.320 | 4.40 | 10.960 |
| Sa22 | 161.85 | 125.25 | 309.81 | 284.87 | 162.33 | 180.45 | 1 | 1.44 | 29.79 | 52.850 | 8.450 | 9.850 | 21.34 | 43.000 |
| Sa23 | 161.74 | 69.23 | 187.52 | 274.43 | 42.79 | 204.07 | 3.78 | 2.95 | 0.70 | 42.030 | 6.700 | 7.330 | -6.00 | 34.700 |
| Sa24 | 215.23 | 193.52 | 92.14 | 125.16 | 112.82 | 180.80 | 1.22 | 1.44 | 11.90 | -14.700 | 49.555 | -5.600 | -37.66 | -9.100 |
| Sa25 | 387.64 | 350.61 | 150.88 | 177.78 | 288.65 | 289.29 | 1.91 | 1.63 | 4.25 | 4.295 | 78.110 | 55.765 | -73.86 | -51.470 |
| Sa26 | NA | 194.24 | NA | 65.28 | NA | 124.83 | NA | 1.91 | NA | 8.370 | NA | 51.200 | NA | -42.830 |
| Sa4 | NA | 73.27 | NA | 425.65 | NA | 243.56 | NA | 3.32 | NA | 51.550 | NA | 3.940 | NA | 47.610 |
| Sa5 | NA | 144.13 | NA | 149.82 | NA | 107.86 | NA | 1.34 | NA | 8.130 | NA | 27.990 | NA | -19.860 |
| Sa6 | 70.90 | 348.76 | 273.59 | 78.47 | 171.68 | 195.65 | 2.42 | 2.49 | 34.50 | -2.880 | 9.000 | 59.810 | 25.50 | -62.690 |
| Sa7 | 76.41 | 93.15 | 490.60 | 459.43 | 304.39 | 278.00 | 3.98 | 2.98 | 44.64 | 54.200 | -2.411 | 12.000 | 47.05 | 42.200 |
| Sa8 | 195.13 | NA | 196.72 | NA | 55.94 | NA | 3.49 | NA | 2.10 | NA | 13.700 | NA | -11.60 | NA |
| Sa9 | 107.49 | 171.89 | 181.84 | 163.96 | 151.93 | 46.49 | 1.41 | 3.53 | 29.35 | 5.520 | 21.950 | 7.630 | 7.40 | -2.110 |