Import - Edge angle analysis - experimental data

Lisa Schunk

2021-02-04 16:28:44

# Goal of the script

This script imports and merges all single CSV-files generated with the ‘edge angle method’. The data derives from 3D models of artefacts from three different experiments: Initial experiment, arificial VS natural experiment, tool function experiment. The data always contain the 3D models from before, after 50, 250, 1000 and 2000 strokes.

The script will:

1. Read in the original CSV-files
2. Combine the data from all samples into one
3. Write an XLSX-file and save an R object ready for further analysis in R

dir\_in <- "analysis/raw\_data"  
dir\_out <- "analysis/derived\_data/"

Raw data must be located in “analysis/raw\_data”.  
Formatted data will be saved in “analysis/derived\_data/”. The knit directory for this script is the project directory.

# Load packages

library(tidyverse)

Warning: package 'ggplot2' was built under R version 4.0.3

Warning: package 'readr' was built under R version 4.0.3

Warning: package 'dplyr' was built under R version 4.0.3

Warning: package 'forcats' was built under R version 4.0.3

library(R.utils)  
library(openxlsx)

Warning: package 'openxlsx' was built under R version 4.0.3

library(tools)

# List all files and get names of the files

# List all CSV files in dir\_in  
CSV\_files <- list.files(dir\_in, pattern = "\\.csv$", recursive = TRUE, full.names = TRUE)

# Merge all files and format the data

# Create a list   
data\_final <- vector(mode = "list", length = length(CSV\_files))   
names(data\_final) <- basename(CSV\_files)  
  
# For each sample  
for (s in seq\_along(data\_final)) {  
   
   
 # Gets name of the experiment from path names   
 exp <- dirname(dirname(dirname(CSV\_files[s]))) %>%  
 basename()  
   
 # read the data files  
 data\_final[[s]] <- read.csv(CSV\_files[s]) %>%   
 mutate(experiment = exp) %>%   
 select(experiment, everything()) %>%   
 rename(Angle\_number = angel\_number, Distance\_origin =  
 dist.to.origin.on.curve..mm., Segment =  
 segment.on.section..mm.,Three\_point =  
 angle.1..3.points...degree., Two\_lines =  
 angle.2..2.constructed.lines...degree., Best\_fit =  
 angle.3..2.BestFit.lines...degree.)  
}  
  
# rbind all files   
data\_final2 <- do.call(rbind, data\_final)  
# adds indices as row names   
row.names(data\_final2) <- 1:nrow(data\_final2)  
  
# split column section  
underscore\_split <- strsplit(data\_final2[["section"]], "\_")  
underscore\_bind <- do.call(rbind, underscore\_split)  
minus\_split <- strsplit(underscore\_bind[,1], "-")  
minus\_bind <- do.call(rbind, minus\_split)  
sample\_ID <- paste(minus\_bind[,1], minus\_bind[,2], sep="-")  
  
data\_final3 <- data\_final2 %>%   
 mutate(ID = sample\_ID, strokes = minus\_bind[,3], edge = underscore\_bind[,2],   
 sec = underscore\_bind[,4])  
  
# extracts the raw material based on the ID  
data\_final3[grep("FLT", data\_final3[["ID"]]), "Raw.material"] <- "flint"  
data\_final3[grep("LYDIT", data\_final3[["ID"]]), "Raw.material"] <- "lydite"  
data\_final3[["Raw.material"]] <- factor(data\_final3[["Raw.material"]])  
  
data\_final3[["Raw.material"]] <- factor(data\_final3[["Raw.material"]])  
  
# adds the contact/worked material  
data\_final3[grep("LYDIT4-1", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("LYDIT4-4", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("LYDIT4-6", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("LYDIT4-2", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("LYDIT4-3", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("LYDIT4-8", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("LYDIT4-5", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("LYDIT4-7", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("LYDIT4-12", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("LYDIT4-9", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
data\_final3[grep("LYDIT4-10", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
data\_final3[grep("LYDIT4-11", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
  
data\_final3[grep("FLT4-4", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("FLT4-8", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("FLT4-9", data\_final3[["ID"]]), "Contact.material"] <- "pork skin"  
data\_final3[grep("FLT4-5", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("FLT4-7", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("FLT4-10", data\_final3[["ID"]]), "Contact.material"] <- "bone plate"  
data\_final3[grep("FLT4-15", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("FLT4-14", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("FLT4-6", data\_final3[["ID"]]), "Contact.material"] <- "bos scapula"  
data\_final3[grep("FLT4-11", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
data\_final3[grep("FLT4-12", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
data\_final3[grep("FLT4-13", data\_final3[["ID"]]), "Contact.material"] <- "skin pad"  
  
data\_final3[grep("initial\_experiment", data\_final3[["experiment"]]), "Contact.material"] <-  
 "bone plate"  
data\_final3[grep("tool\_function-experiment\_cutting", data\_final3[["experiment"]]),  
 "Contact.material"] <- "bone plate"  
data\_final3[grep("tool\_function-experiment\_carving", data\_final3[["experiment"]]),  
 "Contact.material"] <- "bone plate"  
  
data\_final3[["experiment"]] <- factor(data\_final3[["experiment"]])  
  
# adds column about the task/movement   
data\_final3[grep("tool\_function-experiment\_cutting", data\_final3[["experiment"]]),   
 "Task"] <- "cutting"  
data\_final3[grep("tool\_function-experiment\_carving", data\_final3[["experiment"]]),   
 "Task"] <- "carving"  
data\_final3[grep("initial\_experiment", data\_final3[["experiment"]]), "Task"] <- "cutting"  
data\_final3[grep("'aVSn'-experiment", data\_final3[["experiment"]]), "Task"] <- "cutting"  
data\_final3[grep("LYDIT5-14", data\_final3[["ID"]]), "Task"] <- "scraping"  
  
data\_final3[["Task"]] <- factor(data\_final3[["Task"]])  
  
# adds column about the edge angle   
data\_final3[grep("'aVSn'-experiment", data\_final3[["experiment"]]), "Edge.angle"] <- "60°"  
  
data\_final3[grep("FLT8-1", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-2", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-3", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-2", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-3", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-4", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-4", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("FLT8-5", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("FLT8-6", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-5", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-6", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-7", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
  
data\_final3[grep("FLT8-10", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-11", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-12", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-8", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-9", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("LYDIT5-10", data\_final3[["ID"]]), "Edge.angle"] <- "45°"  
data\_final3[grep("FLT8-7", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("FLT8-8", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("FLT8-9", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-11", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-12", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-13", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
data\_final3[grep("LYDIT5-14", data\_final3[["ID"]]), "Edge.angle"] <- "35°"  
  
data\_final3[grep("FLT4-2", data\_final3[["ID"]]), "Edge.angle"] <- "40°"  
data\_final3[grep("FLT4-3", data\_final3[["ID"]]), "Edge.angle"] <- "60°"  
data\_final3[grep("LYDIT1-2", data\_final3[["ID"]]), "Edge.angle"] <- "40°"  
data\_final3[grep("LYDIT1-3", data\_final3[["ID"]]), "Edge.angle"] <- "40°"  
data\_final3[grep("LYDIT1-4", data\_final3[["ID"]]), "Edge.angle"] <- "40°"  
data\_final3[grep("LYDIT3-2", data\_final3[["ID"]]), "Edge.angle"] <- "60°"  
data\_final3[grep("LYDIT3-3", data\_final3[["ID"]]), "Edge.angle"] <- "60°"  
  
data\_final3[["Edge.angle"]] <- factor(data\_final3[["Edge.angle"]])  
  
  
# reorder columns  
data\_final3 <- data\_final3[c(2, 1, 9:16, 3:8)]

# Save data

## Format name of output file

file\_out <- "EdgeAngle\_experiment"

## Write to XLSX

write.xlsx(list(data = data\_final3), file = paste0(dir\_out, file\_out, ".xlsx"))

## Save R object

saveObject(data\_final, file = paste0(dir\_out, file\_out, ".Rbin"))

# sessionInfo() and RStudio version

sessionInfo()

R version 4.0.2 (2020-06-22)  
Platform: x86\_64-w64-mingw32/x64 (64-bit)  
Running under: Windows 10 x64 (build 19041)  
  
Matrix products: default  
  
locale:  
[1] LC\_COLLATE=German\_Germany.1252 LC\_CTYPE=German\_Germany.1252   
[3] LC\_MONETARY=German\_Germany.1252 LC\_NUMERIC=C   
[5] LC\_TIME=German\_Germany.1252   
  
attached base packages:  
[1] tools stats graphics grDevices utils datasets methods   
[8] base   
  
other attached packages:  
 [1] openxlsx\_4.2.3 R.utils\_2.10.1 R.oo\_1.24.0 R.methodsS3\_1.8.1  
 [5] forcats\_0.5.1 stringr\_1.4.0 dplyr\_1.0.3 purrr\_0.3.4   
 [9] readr\_1.4.0 tidyr\_1.1.2 tibble\_3.0.6 ggplot2\_3.3.3   
[13] tidyverse\_1.3.0   
  
loaded via a namespace (and not attached):  
 [1] tidyselect\_1.1.0 xfun\_0.20 haven\_2.3.1 colorspace\_2.0-0   
 [5] vctrs\_0.3.6 generics\_0.1.0 htmltools\_0.5.1.1 yaml\_2.2.1   
 [9] rlang\_0.4.10 pillar\_1.4.7 glue\_1.4.2 withr\_2.4.1   
[13] DBI\_1.1.1 dbplyr\_2.0.0 modelr\_0.1.8 readxl\_1.3.1   
[17] lifecycle\_0.2.0 munsell\_0.5.0 gtable\_0.3.0 cellranger\_1.1.0   
[21] zip\_2.1.1 rvest\_0.3.6 evaluate\_0.14 knitr\_1.31   
[25] broom\_0.7.4 Rcpp\_1.0.6 scales\_1.1.1 backports\_1.2.1   
[29] jsonlite\_1.7.2 fs\_1.5.0 hms\_1.0.0 digest\_0.6.27   
[33] stringi\_1.5.3 grid\_4.0.2 cli\_2.3.0 magrittr\_2.0.1   
[37] crayon\_1.4.0 pkgconfig\_2.0.3 ellipsis\_0.3.1 xml2\_1.3.2   
[41] reprex\_1.0.0 lubridate\_1.7.9.2 assertthat\_0.2.1 rmarkdown\_2.6   
[45] httr\_1.4.2 rstudioapi\_0.13 R6\_2.5.0 compiler\_4.0.2

RStudio version 1.3.1073.

END OF SCRIPT