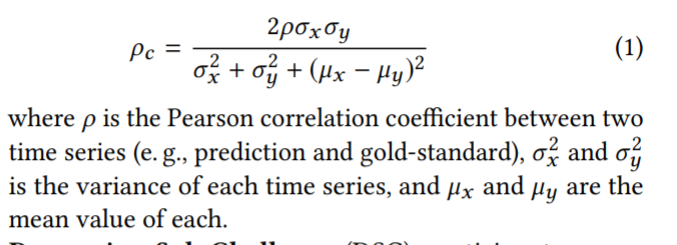
AVEC 2017: Real-life Depression and Affect Recognition

Concordance Correlation Coefficient (CCC): evaluates agreement between two time series by scaling their correlation coefficient with their mean square distance.



# Baseline Features

Different acoustic feature sets included, all based on extended version of Geneva Minimalistic Acoustic Parameter Set (eGeMAPS). It is an expert knowledge based feature set consisting of 23 acoustic LLDs extracted over a time frame.

LLDs only capture very local info in time so need segment level representation of features.

Bag of audio words (BoAW) – originally for text. LLDs over a certain segment are first quantised using codebook of “audio words”, then a histogram of audio words occurring in the segment is created.

Video features:

* Face orientation in degrees (3 features, pitch, yaw and roll)
* Pixel co-ords for 10 eye points (for x and y so 20 features in total)
* Pixel co-ords for 49 facial landmarks (x and y so 98 in total)

Segment level representation – bag of video words (BoVW) computed from normalised facial features.