Analyses of face reader data in case of autistic persons in frame of the MAUGLI project

PITLIK, László, MY-X team[[1]](#footnote-1)[[2]](#footnote-2)

Motto: Science is what we understand well enough to explain to a computer. Art is everything else we do. (KNUTH - [http://miau.gau.hu/miau2009/index\_tki.php3?\_filterText0=\*knuth](http://miau.gau.hu/miau2009/index_tki.php3?_filterText0=*knuth))

Abstract: This paper is trying to demonstrate how robot-eyes work. Robot-eyes should be capable of seeing in multidimensional (big data) spaces – not only in case of digitalized drawings but also in non-graphical contents like face reader information units. Being able to see by a robot means that abstract terms like saturation, exposure concerning drawings, emotions, faces, etc. can be derived based on measurable phenomena like colors, movements, shapes, etc. The re-constructed human terms support hermeneutics/diagnostics/simulations – e.g. in case of autistic persons. The same analytical logic will be involved into the interpretations of the interactions of sporting teams or teachers/students. The interpretation of autistic persons based on face reader equipment is a specific challenge because the initializing of emotions are calibrated for non-autistic persons. Changes, similarities can be interpreted however in arbitrary constellations.

Keywords: artificial intelligence, robot eye, similarity analysis, consistence, emotion

# Acknowledgement

The data asset for this paper comes from the MAUGLI-project in Hungary[[3]](#footnote-3), where masterminds are Mr. SZANI (Ferenc) and Mr. NAGY (Zoltán). The measurements were made by ETRESEARCH[[4]](#footnote-4).

# Introduction

The robot-eye-concept[[5]](#footnote-5) is an interpretation logic, where the objective is: creation of algorithms being capable of seeing in different ways like human eyes (eg. pictures through thermo-cameras vs. photos with natural colors).

Face reading makes possible to estimate emotions like:

1. Neutral
2. Happy
3. Sad
4. Angry
5. Surprised
6. Scared
7. Disgusted

The detailed log in a face reading process contains all the emotional classifier outputs. Each emotion is expressed as a value between 0 and 1, indicating the intensity of the emotion. '0' means that the emotion is not visible in the facial expression, '1' means that the emotion is fully present. When no face is found or the face cannot be modelled (e.g. because the quality of the video is poor), a record with FIND\_FAILED (it means no object could be identified to measure) or FIT\_FAILED (object is given, but measuring is not possible) is added.

Similarity analysis is a complex methodology[[6]](#footnote-6) to support intuition-generation like deriving artificial terms and/or price/ratio-optima and/or forecast, etc. Similarities can be defined between arbitrary objects based on appropriate level of meta-data.

Autism is a complex phenomenon, which is hard to interpret with standard/biological human intuition processes. In order to be capable of interpreting autistic behavior in a better way, it is necessary to be able to sense (see/hear/touch/etc.) as wide range as possible. The extension of the range of sensations means: each combinatorial possible variation of interpretation of measurable signs should be generable.

This paper is a kind of follow-up action for a lot of analytical activities based on similarities. Through chains of similarity analyses for creation of arbitrary abstractions (terms) and through flexible defining OAMs[[7]](#footnote-7) it is possible to extend the possibilities of statistical descriptions. Based on new terms a new world is born like the term variant for snow in case of Eskimos or the same variability of terms for sand in desert populations compared to the simple words like snow and sand used in “standard” environments. Further analogies can be: hunting without eyes but based on ultrasound or reading without eyes but with fingers (c.f. braille alphabet).

# Measurements

There are two simple measurements: two persons have been measured with the same face reader unit in different environment (it means during watching different videos with different lengths and contents). The persons were sitting in swivel chairs during the face reading processes. Therefore it was trivial to lose the initial signs or even the object itself. As it can be seen: the quality and the quantity of the available data asset as such is not yet good enough for statistic robust interpretations. So, the objective of this paper is to demonstrate possibilities on the field of term creations and hermeneutics based on data from face readers.

The measurement delivered two charts with 1+7 columns: the time-co-ordinates + the above listed 7 emotions. The charts had 4526 and 5485 records for 5 minutes and 08 seconds vs. 6 minutes and 05 seconds. Measurements were executed in different time intervals. The average sequence is ca. 0.07 second. The two anonymized persons will be described as P(A) and P(B), where figures about P(A) will have a background color of grey and this person had the shorter time interval. The figures for P(B) do not have any colored background.

# Results

P(A) produced correct measurements in 27% of all records. P(B) has the same value higher: 59%. The differences in concentration capacity can be caused both through the personality and the content of the video - the environment (temperature, etc. during the measurements) was the same.

The following results try to demonstrate comparable characteristics – without delivering any causalities. The figures stands for new terms being not visible in a direct way by any humans. The paper tries to find the appropriate magic of words to create texts about the graphical information units. Figures will always be presented in pairwise sequence to support visual comparisons.

## Most relevant emotions

The first figures show colored codes for emotion being the most robust detectable in time intervals, where the measurements were possible at all. (In each time interval each emotion had a kind of potential presence. The most relevant emotion is the emotion with the highest detection index value.)





Figure-1: Most relevant emotions (source: own presentation, where P(A) chart without background color and P(B) has a greyish background)

The following interpretation makes it possible to have a holistic picture about the interpretation potential of this first (basically statistical) figure based only on valid measurements without showing non-measurable time intervals. It can be reported:

* the ratio of emotions
* the most determinant/characteristic emotion
* the maximal and average value of the intensity of emotions (see emotion level – Y-axis)
* the homogeneity of colored codes
* the dynamic of the intensities of expression of emotions (increasing/decreasing of emotions)
* the trend of emotional intensities during the time
* …

About the two Persons:

* P(B) is more happy and have angry phases in higher frequency than P(A)
* P(A) is more sad and longer surprised than P(B)
* P(A) has higher intensity values than P(B)
* P(B) has high intensity phases with higher frequency
* …

In general: In case of the same video content and redundant/frequent measurements the habits of the persons can be described. If the autistic person is willing to estimate/declare the own emotions, then it can be assumed, that the measured sadness/happiness/etc. are not the real emotions. The self-evaluation should not (always) deliver realistic emotions. EEG-figures and/or other physiological indicators and/or eye tracking measurements can also support the declaration of the real psychical status. It would be a potential finding, if the interpretation of standard emotion could be calibrated for autistic persons in a more robust way.

## Detected emotions

The delivered estimations about the detected emotions are not values of probabilities. The sum of the emotion levels do not lead up to 1.000. The sum of the emotion levels is theoretically 7, but the emotions should have antagonisms – even in case of actors with very flexible face muscles. Therefore the realistic highest value should be lower (ca. 3). If the sum of the 7 emotion levels are given (in case of the really interpretable measurements), the following reports can be derived:







Figure-2: Sum of measured emotion levels (source: own presentation, where the last figure is a proportioned view of the figure in the middle to show impulsivity in the same time intervals)

* the trend/steepness of the saturation of the aggregated emotion levels (red lines)
* the amount of positive and/or negative peaks over given thresholds
* the maximal, minimal, average, etc. distances of peaks
* the most relevant cutting points[[8]](#footnote-8) and their positions, distances
* lengths of intervals (integrals) over/under the trend line
* …

About the two Persons:

* P(A) has a decreasing trend compared to P(B) with light increasing trend – it means a kind of temporal “burn out” in case of decreasing
* P(A) has a higher impulsivity of emotional changes in the aggregated level
* …

In general: The aggregation of emotion level is already a kind of new term-creation, where the appropriate words, phrases, sentences are not trivial to define/find for the human intuition processes.

## Invalid and valid phases

If the non-measurable intervals are also presented, the interpretations of Figure-2 lead to new characteristics, such as below:





Figure-3: Measurable and non-measurable phases (source: own presentation, where the values for non-measurable time units are substituted through the last valid measured values)

* lengths/ratios of (non-)measurable time intervals
* amounts of come-back-phases
* …

About the two Persons:

* P(A) has longer phases delivering non-measurable situations then P(B)
* …

## Only invalid phases

The barcode-like figures about the invalid (non-measurable) phases can be seen as a kind of fingerprint, with characteristics like below:





Figure-4: Invalid phases (source: own presentation, where phases with a lost object has a value of -14 and the phases with lost initializing has a value of +14 – quasi random choice)

* trend/amount of changes
* the lengths of homogeneous phases
* …

About the two Persons:

* P(A) can be characterized through longer (typical) non-measurable phases, while P(B) has more and more non-measurable phases but first the lost initializing has a more massive presence
* P(A)’s face can be lost for longer monotone time intervals than in case of P(B)
* …

In general:

* Figure-3 and Figure-4 have seemingly a massive statistical characteristics.
* Yet, the barcode (or fingerprint) can also be seen as a new term – needing new magic of words for communication of the potential hermeneutics.
* …

## Stability indices

A really new term can be derived e.g. if a lot of standard characteristics are calculated for aggregated time intervals in form of average values. Potential standard characteristics can be:

* the already known sum of emotion levels
* the maximum value of the 7 emotion levels
* the minimum value of the 7 emotion levels
* the standard deviation for the 7 emotion levels
* the maximum value of 6 emotion levels (where neutral emotion will be rejected)
* the minimum value of 6 emotion levels (where neutral emotion will be rejected)
* the standard deviation for 6 emotion levels (where neutral emotion will be rejected)

The neutral emotion level is the most dominant value, therefore a kind of partial exclusion leads to new characteristics.

The time intervals (records) with valid measurements can be aggregated to minimize random-like impulses. In case of P(A) there are 39 time intervals, while P(B) had 71 aggregated time units. In order to ensure comparable visual effects for the human eyes and human intuition processes, the characteristics of P(A) have 3 variants:

* full version with 71 time units after a learning phase with also 71 time units
* partial version with the first 39 time units from 71 units
* version with 39 time units after a learning phase with also 39 time units,

where a learning phase means pairwise[[9]](#footnote-9) similarity analyses to check whether each aggregated time interval can be seen as the same exposure level of emotions. The result is an estimation about a constant value (1.000.000) to show where the emotions were rel. overloaded and under loaded.









Figure-5: Modelled characteristics (source: MY-X FREE online engine – own presentation)

Reporting possibilities:

* steepness of linear trends
* waves (lengths, courses) of non-linear trends
* amplitudes of estimated stability of emotion levels
* …

About the two Persons:

* P(B) has more waves than P(A)
* P(B) starts with a position over linear trend
* …

In general:

* The two Persons can not be seen as emotionally relative neutral – it means: the aggregated time units for valid data deliver different stabilities (waves).
* The resonance according video-content can be visualized (like Hedonism index[[10]](#footnote-10)).
* …

# Data-visualization

The raw and or modelled terms can be visualized in a multi-dimensional frame of Rosling-animations[[11]](#footnote-11) (c.f. Gapminder): <http://miau.gau.hu/miau/233/maugli_rosling>

Examples:

* It is possible to follow parallelisms or antagonisms of time series of emotions concerning the expectations that similarities and controversial ratios can be detected in a rational way (leading to the conclusion that emotions are general valid categories):
	+ happiness vs. sadness
	+ happiness vs. angriness
	+ sadness vs. angriness
* It is also possible to animate time series for abstract terms like:
	+ aggregated emotion level (saturation) or
	+ stability of emotions.

# Conclusions

The robot-eye-concept is able to deliver more or less complex/abstract terms (views) based on statistical approaches and/or similarity analyses. The new terms need new verbalisms (magic of words) to interpret the new constellations of (big) data assets. There are unlimited possibilities to build new terms. The usefulness of the new terms can be evaluated if the data quantity and data quality make it possible to define specific OAM for production functions, estimations, simulations, anti-discriminative analyses. Similarity analyses support both creating new terms and learning from OAM involving the new terms. Data-visualization can support the minimizing of the magic of words in an effective way…

# References

…affected literatures are in the footnotes or directly in the text…

General background: <http://www.kre.hu/ebook/dmdocuments/ii_orszagos_alkalmazott_pszichologiai_phd_hallgatoi_konferencia/chap_09.html>

1. <http://my-x.hu/> [↑](#footnote-ref-1)
2. email: pitlik@miau.gau.hu / maugli@infoter.eu [↑](#footnote-ref-2)
3. <https://www.bettshow.com/bett-suppliers-videos/maugli-project-mood-film#/> [↑](#footnote-ref-3)
4. <http://www.etresearch.hu/index.php> [↑](#footnote-ref-4)
5. <http://miau.gau.hu/miau/228/robotszem2.docx> [↑](#footnote-ref-5)
6. <http://miau.gau.hu/myx-free/index.php3?x=e0> or <http://www.tess-project.eu/deliverables/TESS_wp4_d41_Database_of_models_that_relate_species_and_incomes_to_land-use_15_Nov_2010_IST.pdf#page=41> [↑](#footnote-ref-6)
7. object-attribute-matrix = OAM [↑](#footnote-ref-7)
8. <http://miau.gau.hu/miau/200/szakaszolas.doc> [↑](#footnote-ref-8)
9. Similarity analyses are capable of symmetry analyses to ensure a kind of consistence of results… [↑](#footnote-ref-9)
10. <http://miau.gau.hu/miau/kofop/apetit.docx> [↑](#footnote-ref-10)
11. <http://miau.gau.hu/miau2009/index.php3?x=e0&string=rosling> [↑](#footnote-ref-11)