Visualising emotions with linearly parameterized facial expressions

Paramoji

November 14, 2023

Abstract

The key challenge of affective computing is to translate subjective emotional experiences into measurable data. Most recent advances in that field have relied on facial expressions as indicators for inner emotional states. Our current understanding of these expressions is categorical, i.e. there are some subjective feelings that are understood to be indicated by specific faces. There is no agreed measure for their distance and no clear rule for how emotions can combine or do exclude. The computational processing of emotions therefore mostly relies on compositional approaches using intependent dimensions, like valence and arousal. However, these methods lack a consistent mapping to a facial expression that turned out so crucial for the identification of emotional states.

This paper seeks to solve this problem by introducing a consistent facial visualisation for a five dimensional model. A graphical representation as a comic style facial expression is provided in examples, explanations and code. It visualises emotions in a comically exaggerated style, similar to the successful emojis used in electronic communication. All graphical parameters depend linearly on the input dimensions (valence, arousal, dominance, contempt and control). This might be a rough and crude method. It might not even be a good one, but it is at least something linear in the space of emotions. Empirical results from crowd workers confirm that the encoded emotional information can be recorgnized intuitively and without further training.

Introduction

Humans have largely embraced the use of emoticons or emojis with facial expressions to show emotions when communicating over a digital channel. There is no doubt that these small faces fill a gap in informal communication that is hard to match with written words. A vast amount of human generated content filled with such emotional cues could have provided a unique opportunity for humans and computers to exchange emotional information, as envisioned in the field of Affective Computing[1]. However, this is not happening. There is no known application where humans and computers interact with the meaningful use of emoji or other facial expressions. Quite to the contrary, the lacking ability of showing appropriate emotional indicators hampers the field of robotics, as the missing empathy makes robots seem less intelligent and less trustworthy[2, 3]. Visualising an appropriate emotion with an appropriate facial expression seems to be a very difficult task for current algorithms. What makes this emotional content so intuitive for humans, but yet so hard to process for computers?

The natural starting point for the analysis of digitized facial expressions are of course the standardized emojis. Many attempts have been made to measure (or embed) these little icons with various axes of emotional intensity[4, 5]. Unfortunately, there is a lack of consensus among humans on how to evaluate different facial icons on a unified scale. Cross cultural variations significantly impact how emojis are ranked on a single sentiment axis[6]. Things get more complicated when system borders are crossed. The Unicode Consortium defines only rough graphical gidelines for many of the emotionally expressive symbols. What is defined as a "grinning face" was found to have at least 17 different graphical versions on various software platforms leading to a severe misunderstanding on even basic attributes of sentiment ranging from positive to negative[7]. In an anatomy based approach the facial features in the emojis were studied[8]. Inferring muscle activity and causing emotions proved difficult in the light graphical abstractions and tiny rendering differences. The bottom line is that standardized emojis are not suitable to encode nuanced and precisely defined emotions.

Outside the world of standardized emojis there are other sets of emoticons, i.e. graphical facial expressions, to fill the need for fine grained scales of emotions. Research has found them useful for the communication of qualias, such as pain and taste. Two effective scales of faces were designed to capture levels of perceived pain within children and adult veterans[9, 10]. It was found that visual scales of facial expressions produced more stable and consistent results in self-reported pain and greatly improved treatment. Likert scales are often used in surveys to assess agreement. Various designs of smiley faces were tested and shown to impact engagement in children[11]. A wider range of emotional experiences can be expressed if additional dimensions are allowed. An emoji grid spans two dimensions, valence and arousal, with selected emojis along the axes[12]. It was found useful for reporting food elicited emotions. Despite the additional depth of two dimensions the grid still lacks space for many emotions. In short, these custom sets are consistenty defined and outperform words on their specific domain. However, their expressive range is limited and does not nearly cover the variety of Unicode faces.

For a greater range of facial expressions there is active research into facial visualisation techniques. This includes 2D graphics[13, 14] as well as physical 3D models[15]. These techniques seek for realism in the displayed face and are loosing on the expressiveness of the exaggerated emojis. Additionally, they mimic certain facial expressions without properly explaining their meaning in the digital realm. Another –almost ancient– technique for facial visualisation uses line graphics driven by high dimensional data[16]. Chernoff faces provide unique facial features controlled by 18 parameters covering the sizes of eyes, mouth and ears as well as angles and distances between them. Unfortunately, this does not include any of the subtle activations that are necessary for the display of deep emotions. Chernoff faces come closest at being a predecessor, or next best solution, to the one I present in this paper. Both share two key features: they use line art and an exaggerated comic style to visualise data and neither has yet produced a success story in any practical application. In the following sections I will do my best to let latter change.

Compositionality of Emotions

This purpose of this article is to propose a visualisation technique for a compositional model of emotions. This approach differs from other visualisation methods that treat emotions as holistic, i.e. that represent specific emotional states in words, graphics or pictures, without regard for similarity or fine grained nuances between categories. The most common holistic visualisation is found in emoticons and emojis with emotional content. I want to argue that their undoubted success in our daily communication practices is due to their strong showing of comically exaggerated facial features. I want to create a similar visualisation for a compositional emotional model. This would open their interal topology to computational processes and hence make them more computable.

The idea of expressing a human emotion along independent dimensions dates back to Willhelm Wundt, who in 1896 derived the first compositional emotional model through self-inspection[17]. In a subsequent review his three dimensions of feeling were translated as pleasantness, excitement and tension[18]. The first quantitative experiment to objectify these insights were performed by Charles E. Osgood in 1957, during the era when punch card computing lead to the increased availability of calculation power. Osgood analyzed the associations that human participants would see between nouns and adjectives[19]. By performing factor analysis on a large data set he could show that three scales of meaning would explain more than 95% of their reported association strength. He called these dimensions pleasantness (good-bad), potency (strong-weak) and activity (fast-slow). While this result is not specific to emotions it is explaining the most fundamental structure of our language, which in turn is driven by our emotions.

In 1980 two publications refined the naming of these three dimensions into their currently prevalent form: valence, arousal and dominance. The first two dimensions were part of the Circumplex model of affect [20]. Valence describes the emotion along the axis of positive, pleasure and approval vs negative, misery and disapproval. Arousal describes the axis from aroused, agitated and active to calm, sleepy and inactive. Recent neurological evidence provides further assurance that these two dimensions actually exist and correlate to measurable brain activities[21]. The name of the third dimension was settled by the pleasure-arousal-dominance model [22]. A replication of Osgood's results with more data on modern electronic equipment confirmed the three dimensions as "analogues" with only slight changes in naming. In this model strength, anger, power and boldness dominate over anxiety, infatuation, fear and loneliness.

During the 1970s Paul Ekman et al. published a series of research that clearly linked facial expressions to basic emotions that form as instinctive reactions and are not subject to culture[23, 24]. He developed the facial action coding system (FACS) to capture facial muscle activations and derive the emotional cause. Among his biggest achievements was the discovery of contempt, an instinctive reaction to the feeling of superiority, that occurs in possible combination to positive and negative emotions [25]. Ekman himself argued that emotions are a combination of 7 disctinct basic affective states: happiness, anger, fear, sadness, surprise, disgust and contempt. Some of them tend to cooccur naturally, such as happiness and surprise or anger and disgust. Others have opposing muscle movements, but might coincide in cases of active deceipt or modulation.

Various attempts have been made to unify these two approaches and explain the combinability and similarity of Ekman's basic emotions through a composition of more fundamental effects[26]. The lack of clusters in observed facial muscle activations seem to suggest that there is an undiscovered pattern at play[27]. Experiments have been made to map activation patterns back to the original dimensions of meaning[28, 29, 30], but fall short of a complete theory that embeds all 7 basic emotions. This article may contribute to this debate, not by analyzing faces, but by generating them. Instead of rating existing faces on known emotional dimensions, I want to generate faces from linear graphical dimensions. This approach might not be entirely faithful to all levels of decypherable detail, but it is expressive and linear and as such adds a minimum of model intrinsic distortions.

The parametric emoji

The proposed parametric emoji (paramoji) is a graphical icon that displays a human emotion from a set of continuous parameters. The meaning of these parameters follows the established systematics. However, it will value graphical consistency over language. It is easy to understand the definition of each dimensions in isolation. In combination, it quickly becomes unclear how to interpret their textual description. To solve this problem the visualisation in the form of a parametric emoji, a comic style face icon, provides a unique and valid facial expression for each choice of parameters. It is the graphics that defines the emotion, not the prose. The face is made to be generic, with next to no references to physiological properties, like age, gender or race. All graphical parameters depended linearly on the emotional dimensions.

It would, of course, have been preferable to construct a scale that is linear in the expressed emotional intensity. However, we would hardly be able to agree on a method for its measurement. Linearity in graphics is not just the second best option. It is objective, as it is easy to check that the generating code essentially consists of a matrix vector multiplication followed by standard plotting commands. It is at least something linear to which other systems can be calibrated. Finally, linearity creates the benefit that any detected smooth transition from one state to another would also show smoothly and without artefacts when animated. A midpoint in parameter space is guaranteed to also be the midpoint in its graphical representation and vice versa.

Valence and arousal

Figure 1 shows the chosen visualisation for the first two dimensions. The graphics is defined through the control points of Bezier lines with some rotation of the eyes' outlines. All coordinates and angles are mapped linearly from the two input dimensions. In this (still incomplete) model arousal controls eye and mouth sizes and valence controls curvature.

There is an obvious zero point for arousal when eyes and mouth stay totally shut. On the other end of the spectrum there is no physical limit on how wide a comic eye can open. Therefore, arousal values can only be compared against each other without an objective center. For valence the natural center is a flat mouth and flat eyes. Eyes and mouth bend upwards for pleasure. This is consistent with a respective muscle movements and was already part of early emoticons as `^_^` and `:)`. Eyes bend downwards for misery. This is derived from a drooping upper eyelid and the foreshortened view when the head is lowered.

Looking at facial expressions one can find the predicted emotions on the perimeter of the domain. Starting on the top in clockwise direction we find: happy and excited, alarmed, distressed, miserable, depressed, sleepy and content, The original publication postulates many more emotions that are supposed to sit tightly packed on a planar space. However, fear vs anger vs tension vs disgust: these are fundamentally different emotions that can hardly be seen in close proximity to each other. More dimensions are needed to separate them.

Dominance

The dominance dimension, previously known as potency, differentiates the emotions on the axis from dominant to submissive. Closely related adjectives are strong, hard, brave, tenacious, angry on one extreme versus weak, soft, cowardly, yielding, fearful on the other [19, 22]. The most typical scenario under which people would show these extremes is after a competetive sportive event that is either won or lost, especially when physical exhaustion suppresses valence cues. An analysis of respective photographs revealed a number of significant

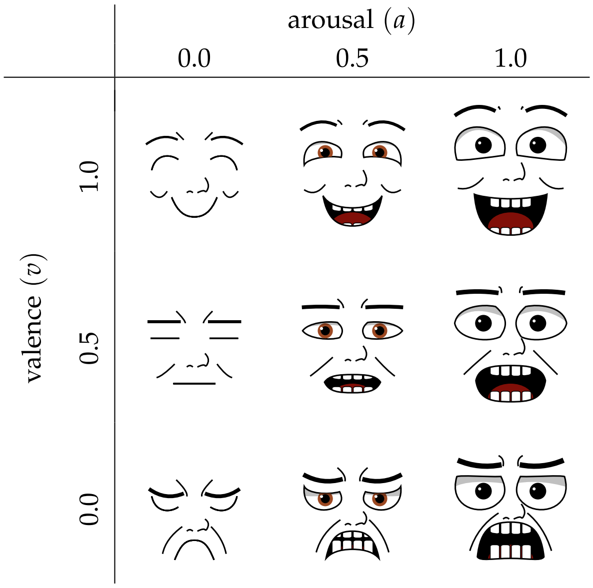


Figure 1: *Facial expressions created with three different values for valence (v) and arousal (a). All displacements of graphical control points depend linearly on these two dimensions and can be freely interpolated.*

features [28]. Interestingly, most effects are symmetric opposites when plotted in a reductionist style. Most salient are the "narrowed and lowered brows" in victorious conditions (AU4) and the "inner brow raise" upon defeat (AU1). Corraborating evidence for this indicator is drawn from a study in which randomly generated faces were rated along subjective assessments of trustworthyness and dominance [31].

On the lower part of the face opposing indicator pairs are the "relaxed mouth" with a dropping jaw (AU26) against the "lip funneler" (AU22) [28]. The jaw position is understood to cause the show of teeth when combined with parted lips and the funneler is indicated with a sharper mouth contour. The "raised cheeks" and "raised chin" are is plotted by shifting the entire lower part of the graphics. Only the "opened mouth" is not included. This article suggests that the effect is not due to dominance, but an overwhelmed reaction combined with arousal as represented by the upcoming control dimension.

Figure 2: *Three facial expressions plotted with high (1.0) and low (0.0) dominance. Although impossible to plot them all, the parameters can be combined arbitrarily.*

Figure 2 shows three example paramojis with lowest (0.0) and highest (1.0) values of dominance. Previous faces in figure 1 are implicitely assumed to be at medium dominance (0.5). The visual effects are obviously symmetric and linear. A rotation of the entire eyes makes for the comically exaggerated appearance. The rotation angle is linearly dependent on the input. The figure shows only extreme values for dominance, but all intermediate values can be interpolated for more natural looks.

The three parameters, valence, arousal and dominance span a unit cube domain. Four corners on the low valence side correspond to misery and disapproval. They already resemble some possible versions of Ekman's basic emotions: fear, anger, sadness and disgust. For positive valence we find pleasure and approval. Two universal emotions sit on the positive end of the cube: happiness and surprise. The expressions in the corners of the positive side are probably not universal. My uninformed attempt at naming them lead to simper, embarrassed, anticipating and a forced smile.

Contempt

The fourth dimension shall be "contempt", derived from a discovery that this feeling of superiority is not a distinct emotion, but combines differently with positive and negative feelings [25]. It is described by the adjectives scornful, haughty, smug, vain, or disdainful. New findings also suggest that "annoyance" is an adequate label for this emotion[32]. Contempt is a feeling of moral superiority that varies based on the perceived social status. Contempt against inferiors, the feeling of power can be a pleasurable experience. Contempt against a superior can feel misarable. Because the expression depends on other dimensions it is a dimension itself, rather than a specific spot in parameters space.

The only universally confirmed expression of contempt is the "unilateral lip-tightening expression"[25]. Tension in the cheek can also press the eyes into squeezing. A squinting eye is not in itself a signal for contempt, but a graphical choice to emphasizes its asymmetric nature and often occurs in coincidence. A subtle side eye adds a trace of annoyance. All effects are one-sided and limited to the left side of the face, an arbitrary choice, because there is no known emotional meaning to chirality.

v	0.9	0.3	0.4
a	0.2	0.1	0.7
d	0.7	0.8	0.4

Figure 3: *Random paramojis with and without contempt. The contempt dimension produces valid facial expressions when combined with any values in the previously defined dimensions.*

Figure 3 shows three emotions on two extreme contempt levels. At least graphically, contempt mixes perfectly with the previous dimensions. The squinting eye is created by an anisotropic scaling operation. Although the scaling factor is linear it does combine with rotation and translation of the scaled object. Technically this leads to nonlinear transformation path of the spline controls. In case you want to place doubt on the claimed trait of linearity then you should start your criticism here. The combinations of parameters produce sometimes comical, but always valid facial expressions, as long as they are in the unit cube.

Control

The control dimension is a less prominent attribute, first introduced by Osgood in 1966. It is the axis along the adjective pairs intentional-unintentional, deliberate-impulsive, controlled-uncontrolled [33]. Despite its scarce occurance in psychological literature, at least one study followed up on this terminology and listed control as the first of 32 measured dimensions[34], defining it as the "aspect of emotion which distinguishes stances taken and initiated by the subject from those elicited by the environment."

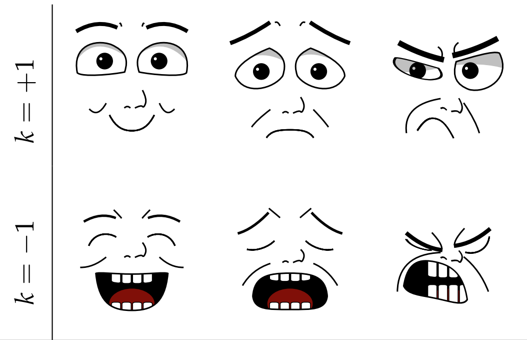
The concept of an intentionally initiated emotion might sound like an oxymoron at first. However, the show of emotions always has a conscious component that controls how much of our emotions are shown. This feeling of restraint is certainly an affective state, although cultural triggers may vary. The question about its instinctive nature should not be driven too far. It has a dinstinct facial expression, therefore it is included in this discussion.

The control dimension as it is used in this article ranges from controlled, modulated, suppressed  to uncontrolled, overwhelmed, expressed . The visualisation follows an hypothesis raised by Ekman and Friesen[35]. They wrote: "when a person is controlling what is shown on his face, more effort is focused on managing what occurs in and around his mouth and lips than in the area of the eyes/lids or brows/forehead." Although there might be many more detailed observations, it is certainly good enough as a first approximation to the subject matter.

The major and easiest to control aspect of an emotion is its intensity. In visual terms, this is arousal (a), because it enlarges eyes and mouth. Psychologically, any show of arousal or surprise would just draw attention and give away information about the precise moment of a state change. The graphical impact caused by a can be separated into arousal shown in the eye region (a_1) and arousal shown in the mouth region (a_2). The amount of excerted control (k) can be computed as the difference between these two. For a controlled and modulated expression k is positive. For an uncontrolled, overwhelmed expression k is negative. The dimensions a_1 and a_2 are mathemtically preferred, because they are defined on the Carthesian domain $[0,1]^5$. The dimensions a and k have, at least theoretically, an emotional interpretation, but produce valid visualisations only on a transformed domain.

Equation (1) defines how the dimension can be converted. The complexity of the equations is commensurate with the psychological depth of this theoretical model. In the end its the graphics that counts, not the numbers. An empirical adjustment based on real human responses to the generated graphics is presented in a later section.

$$\begin{aligned} a_1 &= a + k/2 \\ a_2 &= a - k/2 \\ k &= a_1 - a_2 \end{aligned} \tag{1}$$



v	0.9	0.3	0.1
a	0.5	0.5	0.5
d	0.5	0.0	0.8
c	0.0	0.0	0.7

Figure 4: *Random emotions with minimal and maximal values of control (k) and a* = 0.5*. This dimension reflects that humans tend to excert conscious control over the mouth more than the eyes.*

Figure 4 shows the visual effect of the control dimension on some random random faces. It is interesting that even in the extremely overwhelmed versions one can still differentiate between positive  and negative  valence. This is in contrast to real faces, where extreme pain and extreme joy cannot be distinguished from the facial expression without context information or body language[36]. Emoji style visualisation are not just exaggerated versions of existing facial expressions. They can also encode emotional information that go beyond what real faces show.

Empirical evidence

In the previous section the graphical choices were presented for a parametric emoji that visualises emotions from a five dimensional state. While efforts were made to provide a rational basis for any graphical choices, it's important to acknowledge that many of these decisions were influenced by the author's subjective judgment. To establish the credibility of the final outcome, empirical evidence becomes imperative. This section will lay out the experiment that I made, how I hired untrained participants for annotation work and how I extracted the result from raw data. In particular, I want to show that this theoretical derivation is pretty close to human generated annotations.

The experiment

Figure 5 shows the user interface that test participants had to navigate for the annotation of emotional dimensions. Two faces were randomly drawn from a uniform distribution. The users' task was to order them on the assigned emotional dimension. One had to be chosen as more intense and the the other consistently as less intense. Alternatively, both faces could be labeled as equal. When the two choices matched participants could continue to the next pair. No additional explanation was given after following a link to this emotion classificion task.

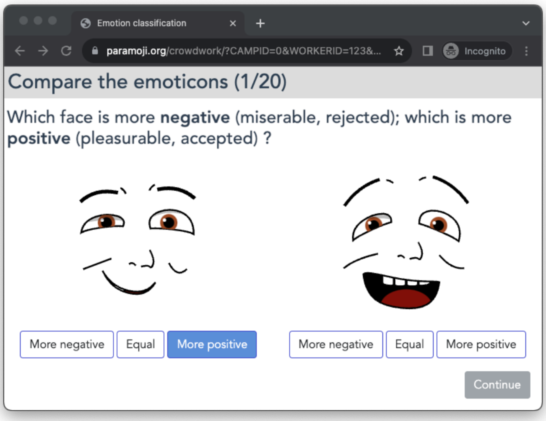


Figure 5: *User interface presented to untrained annotators. Random pairs of faces had to be ordered on one randomly selected emotional dimension.*

Each participant was assigned to one emotional dimension, to avoid any cross influence caused by switching contexts. Each dimension was described by adjectives for their negative and positive extreme. Synonyms were given to avoid miscommunication. The following list shows the five possible tasks descriptions to which participants were assigned.

1. Which face is more **negative** (miserable, rejected); which is more **positive** (pleasurable, accepted) ?
2. Which face is more **calm** (sleepy, relaxed, unimpressed); which is more **aroused** (surprised, panicking, facing the unexpected) ?
3. Which face is more **submissive** (in doubt, passive, insecure); which is more **dominant** (focused, strong, activated)?
4. Which face is more **uncontrolled** (overwhelmed, expressed); which is more **controlled** (pretended, modulated, suppressed) ?
5. Which face is more **humble** (moderate, respectful); which is more **contemptuous** (self-loving, vain, disdainful) ?

Participation

Test participants were hired from MicroWorkers.com, a global market place for crowdsourced work. A deep analysis of the workers' anatomy has been made previously and found to be still valid[37]. The overall work force has a global bias towards India, Bangladesh and the Philippines with some participation from the Western countries.

420 workers took part in this experiment and produced 13190 comparisons. The pay was 1 Cent per entry, paid for batches of size 20 or 30. Some workers submitted two batches without extra reward. The average time spent on one comparison was 10 seconds, the median 5 seconds. It is likely that some workers have seen previous versions of the experiment, but never got any feedback on their performance. The published data are the first results for a textual comparison task.

Regression results

The data were collected as five parameters for the Carthesian model and grouped by the five dimensions that were rated. Additionally, a label was stored to indicate whether it was seen as more, less or equally intense, yielding the rating numbers 1, 0 and 0.5. Since all pairs were chosen fully at random both sides of the comparison could be treated as independent samples. The average rating value of a position in five-dimensional parameter space can be interpreted as the fraction of paramojis that it supercedes, i.e. its quantile on the measured emotional scale.

A linear regression of the reported rating numbers onto the Carthesian dimensions reveals how much the report is swayed by each theoretical dimension. Table table 1 shows the results. The 1-norm column evaluates biggest possible change of the predicted quantile in two opposite corners  and . In absence of doubt or disagreement the least intense position should yield 0 and the most intense should yield 1. As we can see the reported change in valence can be explained through the graphical transformations[1]. It is also interesting that just over half the effect on valence originates from the iconic curvatures controlled by v, which many icon sets depend on solely. Pleasure is clearly more than a curved mouth.

The next thing to observe in table 1 is that most rows are dominated by the v column. With exception of dominance there is a high degree of colinearity between all rows. An aroused or contemptuous face was likely also seen as negative. Control was seen as positive. The describing adjectives were not understood as independent. Since we are interested in orthogonal dimensions, we must map these results onto independent axes.

	v	a_1	a_2	d	c	1-norm
Valence (v^e)	0.54	0.07	-0.06	-0.21	-0.16	1.03
Arousal	-0.29	0.01	0.12	0.16	0.06	0.63
Dominance	-0.05	0.01	-0.01	0.19	0.01	0.29
Control	0.28	0.10	-0.16	-0.10	-0.10	0.75
Contempt	-0.24	0.02	0.03	0.20	0.08	0.58

Table 1: *Regression results of the likelihood of an emotional attribute being reported as more intense onto the Carthesian visualisation parameters.*

Orthogonalization

Apparently, respondents did not understand the meaning of the questions as orthogonal. A better question to ask would have been along the line of "who looks more aroused, without considering the effect of negativity?" However such questions would become quite confusing when adding more dimensions.

The Gram-Schmidt method is a standard procedure to calculate such an orthogonalization mathematically. Equation (2) shows the calculation. The vectors v are the linear dependent input shown in the rows of table 1. The orthogonalized output u is seen in the rows of table 2.

$$\mathrm{u}_i = \mathrm{v}_i - \sum_{j=1}^{i-1} \frac{\langle \mathrm{v}_i, \mathrm{u}_j \rangle}{\langle \mathrm{u}_j, \mathrm{u}_j \rangle} \mathrm{u}_j \tag{2}$$

The Gram-Schmidt method depends on an order of the input and yields slightly different results when rearranged. Some thought should be put into this choice to ensure correct interpretation of the results. The first dimension should clearly be the valence direction. It is the one most decisively identified by participants. It is also the most commonly agreed dimension in literature with some research using it as the only measured variable. Table 2 shows the results starting with valence and then ordered by decreasing 1-norm of the remaining vector.

[1] The value for valence is greater than the valid probability range from 0 (always labeled less intense) to 1 (always labeled more intense). This is because the parameters were estimated based on random points within the domain. The corner points were not among the queried samples and thus the results here are extrapolated values

	v	a_1	a_2	d	c	1-norm
Valence (v^e)	0.54	0.07	-0.06	-0.21	-0.16	1.03
Dominance (d^e)	0.05	0.03	-0.02	0.15	-0.02	0.27
Control (k^e)	-0.03	0.05	-0.12	-0.02	-0.01	0.24
Arousal (a^e)	-0.02	0.06	0.04	-0.00	-0.03	0.15
Contempt (c^e)	0.01	0.01	0.00	-0.00	0.03	0.05

Table 2: *Orthogonalized results. This table shows the independent impact of each dimension after applying Gram-Schmidt orthogonalization.*

Orthogonalized results

The orthogonalization started with the emperical measurement for valence (v^e). Being the first vector it was not changed by the process. Figure 6 shows the result as a graphical transformation, which can be interpreted to sway the perceived sentiment from negative to positiv most efficiently, i.e. with minimal graphical displacements. For the visualisation it is applied to a neutral face , which is chosen at the parameters $(0.5, 0.5, 0.5, 0.5, 0.25)$. This is a purely arbitrary choice. It just so happens to be in the center of the domain with a bias to the non-contemptuous symmetric expression. Any other parameters could have been chosen as the center, because the emotional effect of the graphical transformations were measured across the entire domain and not at a single point, or along a single axis. From the chosen center all parameters are moved in proportion to vector v^e until they reach the boundaries of the drawable space in the unit cube.

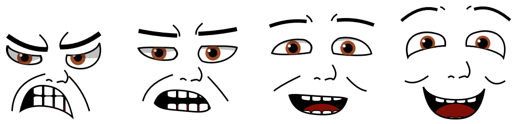


Figure 6: *Emperical data suggest that this is the most effective graphical change to sway reported impressions from negative to positive.*

The second strongest orthogonal dimension is dominance (d^e). This finding is consistent with other research that found valence and dominance as the two main driver for facial assessment [31]. Figure 7 shows the faces on the dominance axis after removing all the colinear projections. Dominance is almost naturally understood as an independent dimension. Little correction was necessary, but the mouth on the dominant side shows a light smile, which would be positive. However, the frowning eyebrows are considered negative. According to this analysis the effects cancel each other such that there is no linear measurable valence change on this scale.

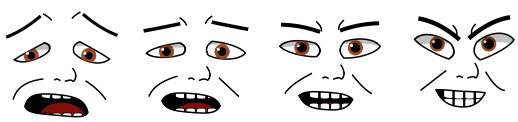


Figure 7: *Emperical evidence for dominance confirms the eye brow angle as the driving factor.*

The third strongest dimension is control (k^e) in figure 8. Although control is correlated to positive sentiment in the raw data, the orthogonalization process seems to slightly overcompensate. Anyways, the main driver is confirmed as the difference between the opening of eyes and mouth.

The fourth empirically confirmed dimension is arousal (a^e) in figure 9. After freeing the reported numbers from the effects of negativity and dominance, the expected synchroneous opening of eyes and mouth emerges. This effect adds up to 0.15 to the probability of being more aroused than a random face. This is a low value given its status as the number two in most models. Apparently there is some limit to the extent to which a static image can convey the impression of arousal. This dimension would be expected to coincide with sudden movements and rapid changes that cannot be shown in a printable text. Nevertheless, the effect is measurable

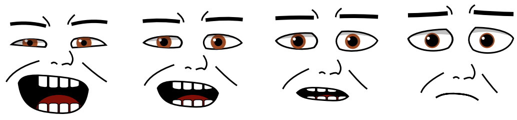


Figure 8: *Emperical reports for control show the scale from throwing a tantrum to the proverbial biting of one's tongue.*

and follows expectation.

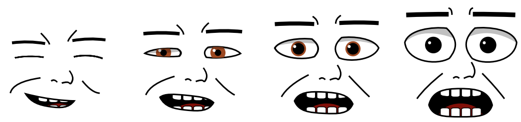


Figure 9: *Empirical results confirm that synchroneously opening eyes and mouth make a face look aroused.*

The final dimension is basically whatever remains as independent from previous observations. Figure 10 visualises this orthogonal remnant. This confirms our expectations of a contemptuous look, although it only explains 0.05 of the probability. Despite being barely measurable we have no reason to doubt the correctness of this result. The asymmetry is a unique facial expression and well documented, though possibly not widely understood. In addition to the asymmetry in mouth and eyes there is an even further enlarged right eye. It can thus be assumed that the asymmetry in the eyes helps in making this emotion more recorgnizable.

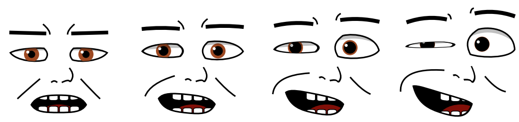


Figure 10: *The contempt dimension is the remnant that cannot be explained by previous dimensions. There is no doubt; this asymmetric look is indeed contempt.*

Transformation

Formula (3) summarizes the transformation from empirical coordinates into Cartesian ones for plotting. They are ordered by decreasing effect on the reported emotional impression. Each entry changes the impact of its emotional component without affecting the previous ones. The function maps the default point $(0.5, 0.5, 0.5, 0.5, 0.25)$ onto itself.

$$\begin{pmatrix} v \\ a_1 \\ a_2 \\ d \\ c \end{pmatrix} = \begin{pmatrix} +1.00 & +0.32 & -0.25 & -0.24 & +0.22 \\ +0.13 & +0.18 & +0.44 & +1.00 & +0.48 \\ -0.11 & -0.13 & -1.00 & +0.56 & +0.07 \\ -0.39 & +1.00 & -0.14 & -0.07 & -0.04 \\ -0.30 & -0.10 & -0.09 & -0.47 & +1.00 \end{pmatrix} \begin{pmatrix} v^e \\ d^e \\ k^e \\ a^e \\ c^e \end{pmatrix} + \begin{pmatrix} +0.03 \\ -0.50 \\ +0.83 \\ +0.31 \\ +0.48 \end{pmatrix} \tag{3}$$

The entries in the matrix are scaled to map each axis through the default point into a valid line within the unit cube. This ensures that each individual axis can always be visualised with the appended plotting function. Off-axis points in the empirical domain may be mapped outside the unit cube of the Cartesian model. This is not such a big problem. As long as a_1 and a_2 are positive and keep the eyes at least as open as shut, the face can be extrapolated and recorgnized with some acceptance for artistic abstractions. Figure 11 shows some results.

Paramoji

Algorithmic emotions

Before concluding this article I want to spend a few thoughts on a possible reinterpretation of the five dimensions. All existing descriptions of emotions tended to antropomophize the entity that experiences them. It is unclear if emotions will ever be experienced by anything other than us humans or our close relatives. The first application of the presented visualisation technique would surely be to use detected human emotions, aggregate them and then redisplay them. However, there is an interesting case where these data are genuinely generated by some algorithm. Regardless of whether technical systems can really feel or just show these emotions, it doesn't really matter, they are still theirs. We will have to talk about them, even before we will fail to distinguished their natures. Therefore I want to provide a possible interpretation of the emotional dimensions such that they may apply to biological and technical systems alike.

Dimension	High value	Low value
Valence	System is operative. Tasks are performed at healthy rate.	System state is critical. Full or partial shutdown seems imminent.
Dominance	System provides instructions and may act autonomously.	System expects instructions, will ask for fine grained permissions before acting.
Control	The system response is restrained by rules ensuring politeness and safety.	The system response is amplified by rules enabling entertainment and emergency reactions.
Arousal	Sensor data indicates upcoming changes. Computational resources are allocated on their analysis.	System state is constant. No change is expected.
Contempt	System input is selfcontradictory or violates fundamental rules of conduct.	No unexpected signs of intellectual underperformance detected on the peripheries.

Table 3: *Possible reinterpretation of the emotional dimensions to include biological and technical systems alike.*

Conclusion

This article introduced a parametric facial expression that can be parametrized to display a wide variety of emotions. Unlike other graphical methods, each parameter has a distinct emotional interpretation. Empirical evidence has been provided that the emotional teint of each axis can be extracted by untrained test participants. Their rating of the resulting faces along the scales of valence, arousal, dominance, contempt und control reproduced the input parameters with minor adjustments.

It has become apparent that little facial expressions can add an emotional note that is hard to match in written text. Emojis have filled that role in human to human communication. They owe their immediate success largely to facial expressions with strong emotional cues. The presented visualisations are similarly suited for the exchange of emotional information. They synthesize a comparable range of facial expressions, hence called, parametric emoji or paramoji for short. However, they are much more suitable for computer processing, as questions for similarity, intensity and sentiment are straight forward to resolve. Stasticial measures like averages are supported in a meaningful manner.

It is known, that cultural norms demand varying degrees of self control with regards to the display of emotions in public. Consequently, facial expressions are interpreted according to these expections. When mapping emotions to adjectives language skills and educational background also play a role. A method was presented to adjust the emotion parameters to a specific target audience, crowd workers in this case. Only linear adjustments were considered, but there is reason to believe that non-linear effects are small. The graphical transformation introduced by each of the five dimensions is extremely smooth and all its parameters are linear. When animated, all transitions always look smooth and straight, making them ideal for visualisations in real time or for replay.

Code

The source code of the paramoji may look compressed, but it is actually very easy to check for linearity. A sparse matrix contains all values that define the transformation from input to drawing parameters. There is a row for each free parameter in the drawing instruction and up to 6 columns for each entry in the transformed input vector $(1, 2v-1, a_1, a_2, 2d-1, c)$. The dot product is computed between each matrix row and the input vector, which yields a values for each question mark in the template string. The actual template might be difficult to deconstruct, but for the scope of this discussion it is sufficent to understand that it is an SVG format with a well documented limited feature set confined to numeric coordinates and affine transformations [38]. Hence, the linearity of the transformation can be confirmed without going into further detail. Both, the JavaScript code and the SVG instructions, can be interpreted by any contemporary web browser.[2]

```
function paramoji(v, a1, a2, d, c) {
 const matrix=[[,,,,,.4],[70,-2,,-8,-1],[-9,-1,-2],[37,-5,,,2,4],[68,-1,-2,-7,-3,-4],[29,-2,,-3,2,-17]
  ,[76,-1,,-8,,-15],[25,-4,,-6,2,-10],[77,-10,,-2],[65,4,,,-2,-.5],[68,-1,-2,-7,-3],[71,2,,3,-2],[76,-
  1,,-8],[75,4,,6,-2],[77,-10,,-2],[80,5,,6,-6],[66,2,,-5,-1],[14,1,,9,-2],[,-12,,-2],[10,1,,11,-3],[,
  -2,,11],[6,,,6,-3],[,1,,12],[,1,,12],[-10,-1,,-11,3],[,-2,,11],[-14,-1,,-9,2],[,-12,,-2],[-10,-1,,-
  11,3],[,-2,,-11],[-6,,,-6,3],[,1,,-12],[,1,,-12],[10,1,,11,-3],[,-2,,-11],[14,1,,9,-2],[,-12,,-2],[,
  ,,,,.3],[50,,,,,-12],[81,2,,,-3,-2],[51,-2,-10],[,22,,,-26],[,,-11],[-8,,-8],[,,-4],[-8,,-8],[,,2],[
  -8,,-3],[,,9],[-4,,1],[,-6,13],[5,,12],[,-9,11],[11,,8],[5,,12],[,-9,-11],[-4,,5],[,-6,-13],[,,-11],
  [,,,,26],[1,,,,,-.8],[,-22],[,,-11],[8,,8],[,,-4],[8,,8],[,,2],[8,,3],[,,9],[4,,-1],[,-6,13],[-5,,-
  12],[,-9,11],[-11,,-8],[-5,,-12],[,-9,-11],[4,,-5],[,-6,-13],[,,-11],[1,,,,,1],[,,,,-1,-4],[,-2.5,,,
  -1.5],[.5,,5],[-36,,,,2],[5,1,-4,,8],[2,-9,-17,,5],[-34,1,-10,,-6,3],[2,-9,-17,,4,15],[59,2,-1,,-6],
  [40,-7,-19,,12],[70,1,2,,-5],[40,-12,-21,,3],[81,1,5,,-4],[40,-4,-19,,-3],[2,-.5,1,,.5],[62,,-4,,-2]
  ,[33,-5,-16,,7],[56,2,,,-3],[39,-7,-22,,9],[56,1,,,-1],[41,-8,-20,,10],[,,,,,4]]
 const template='<svg width="100%" height="100%" viewBox="0 0 100 100"><defs><clipPath id="clip-eyes'+
  '"><use href="#eye-l"/><use href="#eye-r"/></clipPath><clipPath id="clip-lips"><use href="#lips"/>'+
  '</clipPath></defs><path transform="matrix(1,?,0,1,53,?)" d="M-4,0q-2,-2 -4,0M-1,0Q3,-2 4,-1T6,-3 '+
  '4,?" fill="none" stroke="black"/><path d="M?,?Q?,? ?,? M?,?Q?,? ?,?" fill="none" stroke="black"/>'+
  '<g clip-path="url(#clip-lips)"><rect height="100" width="100"/><ellipse cx="50" cy="91" rx="15" r'+
  'y="10" fill="#800f08"/><g transform="translate(0,?)"><use href="#tooth" x="-14"/><use href="#toot'+
  'h" x="-7"/><use href="#tooth"/><use href="#tooth" x="7"/></g><g transform="translate(0,?)"><use h'+
  'ref="#tooth" transform="matrix(1,.14,0,1,-14,-8)"/><use href="#tooth" x="-7"/><rect id="tooth" x='+
  '"50.5" rx="2" ry="1" height="15" width="6" fill="white" stroke="black" stroke-width=".5"/><use hr'+
  'ef="#tooth" transform="matrix(1,-.14,0,1,7,7)"/></g></g><path id="lips" d="M?,?C?,? ?,? 0,?S?,? ?'+
  ',?C?,? ?,? 0,?S?,? ?,?Z" transform="matrix(1,?,0,1,?,?)" fill="none" stroke="black"/><g transform'+
  '="translate(68,?)"><path id="eye-r" transform="rotate(?)" d="M-8,?C?,? ?,? ?,? ?,? ?,? ?,0 ?,? ?,'+
  '? -8,?Z" fill="white" stroke="black"/><path id="eye-l" transform="translate(-36,0) rotate(?) scal'+
  'e(1,?) rotate(?)" d="M8,?C?,? ?,? ?,? ?,? ?,? ?,0 ?,? ?,? 8,?Z" fill="white" stroke="black" strok'+
  'e-width="?"/><g clip-path="url(#clip-eyes)"><g id="lens" transform="translate(?,?)"><circle r="5"'+
  ' fill="#9d4922"/><circle r="?"/><circle r="1" cx="-2.5" cy="-1.5" fill="white"/></g><use href="#l'+
  'ens" x="?"/><path d="M-70,-30H30L20,0Q?,? -18,-2 ?,? -56,0Z" opacity=".25"/></g></g><g id="brow-r'+
  '"><path d="M?,?Q?,? ?,?" fill="none" stroke="black" stroke-width="?"/><path d="M?,?Q?,? ?,?" fill'+
  '="none" stroke="black"/></g><use href="#brow-r" transform="matrix(-1,0,0,1,100,?)"/></svg>'
 const dotprod = (X,Y) => X.reduce((a, b, i) => a + b*(Y[i] || 0), 0)
 const vec = [1, 2*v-1, a1, a2, 2*d-1, c]
 let values = matrix.map(row => dotprod(vec, row))
 return template.replace(/\?/g, () => values.shift())
}
```

Conflict of interest statement

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

[2] A demo of this code is currently deployed at `https://paramoji.org`

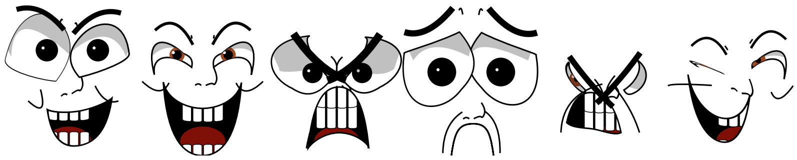


Figure 11: *Extrapolations of the facial features outside the valid range: recognizable, but increasingly grotesque.*

References

[1] Rosalind W. Picard. *Affective Computing*. MIT Press, Cambridge, MA, 1997.

[2] Christiana Tsiourti, Astrid Weiss, Katarzyna Wac, and Markus Vincze. Multimodal integration of emotional signals from voice, body, and context: Effects of (in)congruence on emotion recognition and attitudes towards robots. *International Journal of Social Robotics*, pages 1–19, 2019.

[3] Stephanie Hui-Wen Chuah and Joanne Yu. The future of service: The power of emotion in human-robot interaction. *Journal of Retailing and Consumer Services*, 61:102551, 2021.

[4] Petra Kralj Novak, Jasmina Smailović, Borut Sluban, and Igor Mozetič. Sentiment of emojis. *PLOS ONE*, 10(12):1–22, 12 2015.

[5] Francesco Barbieri, Francesco Ronzano, and Horacio Saggion. What does this emoji mean? a vector space skip-gram model for Twitter emojis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3967–3972, Portorož, Slovenia, May 2016. European Language Resources Association (ELRA).

[6] Andrey Krekhov, Katharina Emmerich, Johannes Fuchs, and Jens Harald Krueger. Interpolating happiness: Understanding the intensity gradations of face emojis across cultures. New York, NY, USA, 2022. Association for Computing Machinery.

[7] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. "blissfully happy" or "ready to fight": Varying interpretations of emoji. 03 2016.

[8] Jennifer M. B. Fugate and Courtny L. Franco. Implications for emotion: Using anatomically based facial coding to compare emoji faces across platforms. *Frontiers in Psychology*, 12, 2021.

[9] D. L. Wong and C. M. Baker. Pain in children: Comparison of assessment scales. *Pediatric Nursing*, 14:9–17, 1988.

[10] Thomas H. Nassif, Amanda Hull, Stephanie Brooks Holliday, Patrick Sullivan, and Friedhelm Sandbrink. Concurrent Validity of the Defense and Veterans Pain Rating Scale in VA Outpatients. *Pain Medicine*, 16(11):2152–2161, 11 2015.

[11] Lynne Hall, Colette Hume, and Sarah Tazzyman. Five degrees of happiness: Effective smiley face likert scales for evaluating with children. pages 311–321, 06 2016.

[12] Alexander Toet, Daisuke Kaneko, Shota Ushiama, Sofie Hoving, Inge de Kruijf, Anne-Marie Brouwer, Victor Kallen, and Jan B. F. van Erp. Emojigrid: A 2d pictorial scale for the assessment of food elicited emotions. *Frontiers in Psychology*, 9, 2018.

[13] Anna Wojdel and Léon Rothkrantz. Parametric generation of facial expressions based on facs. *Computer Graphics Forum*, 24:743 – 757, 12 2005.

[14] Etienne Roesch, Lucas Tamarit, Lionel Reveret, Didier Grandjean, David Sander, and Klaus Scherer. Facsgen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units. *Journal of Nonverbal Behavior*, 35:1–16, 03 2011.

[15] Zanwar Faraj, Mert Selamet, Carlos Morales, Patricio Torres, Maimuna Hossain, Boyuan Chen, and Hod Lipson. Facially expressive humanoid robotic face. *HardwareX*, 9:e00117, 2021.

[16] Herman Chernoff. The use of faces to represent points in k-dimensional space graphically. *Journal of the American Statistical Association*, 68(342):361–368, 1973.

[17] Wilhelm Wundt. *Grundriss der Psychologie*. W. Engelmann, 1 edition, 1896.

[18] E. B. Titchener. The tridimensional theory of feeling. *The American Journal of Psychology*, 19(2):213–231, 1908.

[19] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum. *The Measurement of Meaning*. Illini Books, IB47. University of Illinois Press, 1957.

[20] James Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 12 1980.

[21] Jonathan Posner, James A. Russell, and Bradley S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(03), sep 2005.

[22] Albert Mehrabian. Basic dimensions for a general psychological theory : implications for personality, social, environmental, and developmental studies. 1980.

[23] Paul Ekman and W V Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17 2:124–9, 1971.

[24] Paul Ekman. Universals and cultural differences in facial expressions of emotions. *Nebraska Symposium on Motivation*, pages 207–282, 1972.

[25] Paul Ekman and Karl Heider. The universality of a contempt expression: A replication. *Motivation and Emotion*, pages 303–308, 1988.

[26] Craig A. Smith, Heather S. Scott, and George Mandler. *A Componential Approach to the meaning of facial expressions*, page 229–254. Studies in Emotion and Social Interaction. Cambridge University Press, 1997.

[27] Klaus Scherer and Heiner Ellgring. Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? *Emotion (Washington, D.C.)*, 7:113–30, 03 2007.

[28] Federica Cavicchio, Svetlana Dachkovsky, Livnat Leemor, Simone Shamay-Tsoory, and Wendy Sandler. Compositionality in the language of emotion. *PLoS One*, 2018.

[29] Hanzhong Zhang, Jibin Yin, Xiangliang Zhang, and Francesco Gringoli. The study of a five-dimensional emotional model for facial emotion recognition. *Mob. Inf. Syst.*, 2020, jan 2020.

[30] Feng Zhou, Shu Kong, Charless C. Fowlkes, Tao Chen, and Baiying Lei. Fine-grained facial expression analysis using dimensional emotion model. *Neurocomputing*, 392:38–49, 2020.

[31] Nikolaas N. Oosterhof and Alexander Todorov. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, 105(32):11087–11092, 2008.

[32] Nancy Alvarado. New findings on the contempt expression. *Cognition and Emotion*, 10(4):379–408, 1996.

[33] Charles E. Osgood. Dimensionality of the semantic space for communication via facial expressions. *Scandinavian journal of psychology*, 7 1, 1966.

[34] Nico H. Frijda. Recognition of emotion. volume 4 of *Advances in Experimental Social Psychology*, pages 167–223. Academic Press, 1969.

[35] Paul Ekman and Wallace V. Friesen. *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall, Oxford, 1975.

[36] Sofia et al. Wenzler. Beyond pleasure and pain: Facial expression ambiguity in adults and children during intense situations. *Emotion (Washington, D.C.)*, 16,6:807–14, 2016.

[37] Matthias Hirth, Tobias Hoßfeld, and Phuoc Tran-Gia. Anatomy of a crowdsourcing platform-using the example of microworkers.com. In *2011 Fifth international conference on innovative mobile and internet services in ubiquitous computing*, pages 322–329. IEEE, 2011.

[38] Dean Jackson and Craig Northway. Scalable vector graphics (svg) full 1.2 specification. World Wide Web Consortium, Working Draft WD-SVG12-20050413, April 2005.