Emotion Disc and Emotion Squares: Tools to Explore the Facial Expression Space

Zsófia Ruttkay\(^1\), Han Noot\(^1\) and Paul ten Hagen\(^2\)

\(^1\)CWI, Amsterdam, The Netherlands
\(^2\)Epictoid, Amsterdam, The Netherlands
Zsófia.Ruttkay@cwi.nl, Han.Noot@cwi.nl, Paul.ten.Hagen@epictoid.nl

Abstract

In the paper we present two novel interactive tools, Emotion Disc and Emotion Squares, to explore the facial expression space. They map navigation in a two-dimensional circle, by the first tool, or in two two-dimensional squares, by the second tool, to the high-dimensional parameter space of facial expressions, by using a small number of predefined reference expressions. They can be used as exploration tools by researchers, or as control devices by end-users to put expressions on the face of embodied agents or avatars in applications like games, telepresence and education.

Keywords: facial animation, emotional expressions, principle component analysis

1. Introduction

There has been an increasing interest in generating expressive faces by computer [1] for applications like embodied conversational agents, telepresence and entertainment. The most widely used coding systems to describe facial expressions systems are FACS [2] with 46, and the MPEG-4 standard [3], with 66 low-level facial action parameters (FAPs). These systems allow model-independent, perception-based description of facial expressions.

It is a cumbersome task to produce the required effect, given in terms of dozens of facial expression parameters, by the deformation mechanisms available for the model. Moreover, relatively little empirical knowledge is available on the characterization of facial expressions other than the widely used six generic basic emotional ones — joy, surprise, fear, sadness, anger and disgust [4].

When facial animation is to be produced by hand, the usual way to escape from ‘getting lost in the parameter space’ is to use higher-level compound parameters like mouth width/opening or eyebrows raised, and/or intensity measures for complete facial expressions [5]. Blending of expressions, if supported at all, is done by blending parameter signals of different expressions or also of noise [6]. By these methods semantical questions (what is the result of the blending, is it a meaningful expression?) are not considered. In order to overcome these shortcomings, recently much work has been addressing the issue of discovering the coarticulation hidden in facial expressions, to serve as a basis for lower-dimensional description and control of facial expressions [7].

Our work is in the line of the latter approach. Namely, we provide tools for exploring the expression space — the space of real-valued, low-level control parameters — both to learn about its structure, the location of meaningful–meaningless–physically impossible regions, and to quickly produce a multitude of facial expressions.

2. Emotion Disc

2.1. Theoretical basis and related work

The design of Emotion Disc is based on a perception psychological experiment by Schlosberg [8], concluding that that emotional expressions are perceptually related in such a way that they can be arranged in a two-dimensional (2D) space along a circle. They used five of the six basic expressions by Ekman, sadness was replaced by contempt.
Figure 1: The trace of moving the mouse is indicated on Emotion Disc, with expressions on a schematic cartoon face, corresponding to some locations on the path.

In ComicChat [9], for selecting emotions of characters a wheel was used, with partly similar control function as our EmotionDisc. Eight emotions were arranged along a circle, but without any empirical or theoretical justification for the selection and arrangement of the emotions in 2D. Intensity of emotions changed along the radius. However, only a few discrete emotions and the according (full-body) appearance of the character could be chosen from. Hence the tool was not used to generate new expressions and explore all the possible ones, as in our case, but as an intuitive control device to select from a predefined set.

2.2. The device

Emotion Disc is a device which allows to generate and explore a continuum of facial expressions, assuming that initially the snapshots of the 6 basic expressions were designed and given in terms of facial expression parameters (see Figure 1). The six basic expressions are placed regularly around a circle, basically according to the arrangement gained by empirical experiment cited above. The neutral expression is in the middle of the circle. When moving the cursor on the disc, a new expression is computed based on bi-linear interpolation from two given expressions (see Figure 2). Movement of the cursor along a radius produces different intensities for the same expression.

The disc can be refined further, making non-linear intensity interpolation possible. Namely, one can define specific expressions for a number of M intensities for each emotion. The circle is divided into M rings, which are separated to segments of different emotions. The parameter vector
corresponding to a point P in a segment of the circle is computed by bi-linear interpolation from the parameters of the points defining the segment. This facility allows also that at different intensities different features are used (e.g. mouth opens only at intense surprise).

Finally, it is also possible to use the disc to generate exaggerated, unrealistic versions of the expressions, by linear extrapolation (beyond the normal extreme), or by using designed samples (see Figure 3).

2.3. Applications

Emotion Disc has been used with success by naïve users to reproduce the expressions on cartoon faces which they had seen on real photos [10].

Emotion Disc was also used as a control device: users of a virtual reality forum could put emotional expressions on their avatar’s face [11]. The device proved to be handy, and imposed hardly any extra cognitive load on the user.

The disc has been used to control expressions of 2D as well as 3D facial models. A disc can be made for any particular face, only the six basic expressions has to be designed for the model in question, and the corresponding parameters have to be provided. In case of MPEG-4 compatible faces, a common definition of the six basic expressions in terms of the MPEG-4 facial action parameters can be re-used. In our experiments, we used our in-house Animation Editor [12] to produce the input data for Emotion Disc, but any other animation tool could be used.

3. Emotion Squares

3.1. Theoretical basis

The design of the Emotion Squares is based on dimension reduction of the expression space by performing principle component analysis (PCA) [17]. An experienced animator produced 59 expression stills by using CharToon [12], a tool developed by us earlier to design 2D faces which can be animated. The animator himself categorized and labelled the expressions either as variants of the six basic expressions or as belonging to a group of ‘other’ expressions. The expressions were coded automatically according to the MPEG-4 standard, by using 15 facial action parameters (FAPs). (The choice of the 15 FAPs was done according to similarly coded tracked data, for comparison and discussion see [14].)

We investigated with PCA if our data points in the 15-dimensional expression space lie on some lower-dimensional subspace (hyperplane). We found that the first four components contribute significantly to the data, 83.6% altogether, while the contribution of the 5th and...
higher components can be neglected. Hence the original 15 dimensional data is approximated in a four-dimensional hyperplane. Several FAPs occur in multiple components. Based on the FAPs with the highest coefficients, the four components can be characterised as:

<table>
<thead>
<tr>
<th>Princ. Comp.</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Lowering of the eyebrows, raising lips</td>
</tr>
<tr>
<td>2</td>
<td>Squeezing mouth corners, raising inner brows, asymmetric squeeze of brows</td>
</tr>
<tr>
<td>3</td>
<td>Raising mouth corners</td>
</tr>
<tr>
<td>4</td>
<td>Raising inner brows, among others</td>
</tr>
</tbody>
</table>

3.2. The device

The Emotion Squares consist of two rectangles, the first in the sub-space spanned by the first two largest principle components, the second in the space spanned by the 3rd and 4th principle components. The rectangles are envelopes of the projection of the 59 data points. The user may identify a point in the four-dimensional space of the principle components, by moving the cursor in the two rectangles, and see the resulting expressions.

The projection of subsets of the 59 expressions, belonging to the same basic expression, can be used for orientation. When remaining in the neighbourhood of the points in the two rectangles corresponding to designed expressions of a known type, variants of that expression are produced. Further away from the convex hull of the known expressions, entirely new expressions can be hit upon (see Figure 4). While the first usage is interesting for animators to generate quickly a variety of facial expressions, the latter is of relevance when one is trying to identify regions of ‘meaningful’ and ‘meaningless’ facial expressions in the expression space, or wishes to generate unusual expressions, e.g., for cartoon faces.

3.3. Applications

Emotion Squares is currently being tested by animators. It still has to be proven if the navigation in the two squares is handy enough manually, and manageable mentally, also in comparison to the traditional slider-based controllers.

3.4. Discussion

Both devices were developed with similar motivations, namely:

- to map the high-dimensional expression space to a lower dimensional one;
- to generate a continuum of meaningful emotional expressions, on the basis of a few known expressions in the space, by navigating in the lower dimensional space.

It is interesting to compare the 2D arrangements of the devices. One might expect that the PCA reproduces the early empirical results on perception. However, this was not the case. First of all, the negative emotions did not form in 2D different clusters. Similar result was
reported for tracked data by others [13]. Secondly, the ‘neighbourhood relationship’ in the Euclidean parameter space did not suggest the pattern found by perception. This discrepancy can be explained in several ways. To begin with, the results of Schlosberg have been brought to discussion recently [15,16]. What delays confirmation of the results is the difficulty of gaining reliable empirical data, especially of the negative emotions. A more fundamental explanation of the discrepancy could be that humans do not use Euclidian distances to classify facial expressions: much more characteristics (also non-geometrical ones like colouration) are taken into account, and probably non-uniformly. More empirical research is needed to reveal the contribution of the different factors to the overall impression of the expression.

As of methods of generating new expressions from the few given ones, the two devices are complementary to some extent. With Emotion Disc, one gets high-dimensional data all the time, the samples are reproduced without loss, but only in a limited part of the expression space. With the Emotion Squares, one is restricted to a four-dimensional sub-space of the Expression Space. However, this sub-space can be explored completely, resulting in a bigger variety of new expressions.

While exploring with Emotion Disc, none of the blends turned out to be odd or impossible, unlike in the case of using Emotion Squares. The control parameters of Emotion Disc have clear semantic meaning for the user, which is not the case for Emotion Squares.

Both tools allow fast and easy generation of a continuum of expressions from a few sample ones, and can be coupled with any facial model. They can be used as exploration tools by researchers or as control devices by end-users of different applications.

Acknowledgement

Part of the work was done in the framework of the FASE project nr. STW CWI 66 4088. The PCA of the data was done by J. Hendrix, the facial expressions were designed by A. Lelievre.

References