



# Measurement of JND Thresholds and Riemannian Geometry in Facial Expression Space

Runa Sumiya<sup>1</sup>, Reiner Lenz<sup>2</sup>, and Jinhui Chao<sup>1</sup>(✉)

<sup>1</sup> Department of Information and System Engineering,  
Chuo University, Tokyo, Japan

rsumiya603@gmail.com, jchao@ise.chuo-u.ac.jp

<sup>2</sup> Linköping University, SE-58183 Linköping, Sweden  
reiner.lenz@liu.se

**Abstract.** Currently the most popular approach to facial expression analysis uses categorical representations of expressions based on labels like sad, happy and angry. Subtle expression variations require however a quantitative and continuous representation. Besides, today's subjective expression spaces, built by semantic differential level scores and reduced to low dimensional continuous spaces using MDS or PCA have no direct correspondence with the physical stimuli or the expression images. On the other hand, the spaces used in engineering are based on purely physical stimuli or images which can hardly be called expression spaces. Even in models incorporating spacial structure, the geometry of the expression space received little attention and is usually assumed to be Euclidean. The aim of this paper is to build an expression space which is directly connected with the physical stimuli or the expression images. At the same time, it has to incorporate the subjective characteristics of expression perception. We use methods from psychophysics to build an expression space based on the physical stimuli or expression image space equipped with JND or discrimination threshold data. The construction follows the approach used in color science where the MacAdam ellipsoids provide for every color a metric tensor in a Riemannian space. We show that the discrimination thresholds indicate that the space is not Non-Euclidean. We will also illustrate the intrinsic geometrical structure of the expression spaces for several observers obtained from two large image databases of face expressions.

**Keywords:** Facial expressions · Emotions · Categorical theory  
Dimensional theory · JND thresholds · Riemannian geometry

## 1 Introduction

Face expressions and emotions are now mainly represented by categories such as basic expressions and emotions. In particular, most expression recognition systems are based on features defined by discrete levels of Action Units (AUs) to classify an arbitrary expression image into some basic category [1, 2].

Two theories known in expression perception researches in psychology e.g. [20] are the categorical theory by Ekman [4] which assumes the existence of universally invariant basic expressions among different races and cultures which provide a model of expression perception by classification into these discrete categories described by language labels. The dimensional theory by Schlosberg [5], Russell and Bullock [6] seeks geometric relationships between expressions by arranging the basic expressions into a circular order in a psychological space obtained from multidimensional scaling (MDS) [7]. The dimensional theory could be regarded as an extension rather than a denial of the categorical theory. On the other hand, it was also suggested that there is a natural way to describe emotions or expressions as a continuous distribution in the psychological space.

Current representations of expressions often use verbal labels of the categorical names thus providing a qualitative characterization only. Facial expression images or their features are reduced to certain low dimensional space by e.g. principal component analysis (PCA) or MDS and are then divided into basic categories based on AU matching. The abundance of subtle and delicate, but unnamed expressions and their variations are hard to describe without a quantitative representation. An attempt to eliminate possible linguistic effects (due to the labeling of the expressions using natural language) is due to Russell who used preschool children in their experiments. In the following we assume that the inner structure and the relationships between expressions are also important therefore quantitative descriptions are required. Such relationships, including similarity, can be described intuitively by spatial or geometric properties in an expression space.

Engineering approaches to facial expression recognition have mainly been of the categorical type, trying to classify facial expressions. Methods using more detailed and complicated facial features trying to provide a more detailed analysis have, however, received more interest recently [2]. Combinations of basic expressions are used to describe compound expressions as sub-categories [3]. These approaches provide more accurate, however still discrete descriptions, of expressions. On the other hand, it seemed that a quantitative representation in the form of coordinates of a continuous space could unify the categorical and dimensional theories since the categories can be regarded as domain decomposition of the continuous expression space, just as color categorization is obtained in a color space [11].

Another problem with today's subjective expression spaces is that all of them were built from psychological evaluations such as the affect grid or semantic differential (SD) level scores. These data are reduced to low dimensional continuous spaces using MDS or PCA. As a result, these psychological expression spaces have no direct correspondence with the physical stimuli or the facial expression images. On the other hand, the spaces used in engineering are based on purely physical stimuli or facial images which contain no information of subjective perception. In this paper we use a psychophysical approach to build an expression space based on the physical stimuli or facial expression image space and equip it at every point with the expression JND discrimination thresholds to incorporate subjective characteristics in expression perception. This new expression space is

actually a Riemann space with a metric tensor defined by the JND thresholds. In fact, this approach is well known in color science where the color space is defined in the RGB stimuli space equipped with color JND thresholds known as MacAdam ellipses and ellipsoids [13, 14].

There is also an issue that even in researches used spatial representation of expressions, either in psychological or engineering usage, the geometry of the space received little attention. Dimensional theories, for example, tried to explain why the placement of basic expressions obtained from psychological rating or SD score are far from a perfect circle [9], while engineers used the Euclidean distance between feature vectors such as AU or other descriptors in the PCA subspaces as measure of similarity for classification and recognition. All these methods are based a tacit assumption that the expression space is a Euclidean space which, according to the above arguments, requires to be tested. In this paper, we introduce a framework which enables us to analyze intrinsic geometry in the space of facial expressions.

It is known that it is hard to compare two remote expressions and even harder to describe their difference quantitatively. This suggests that the global geometry of expression space is hard to investigate or properties to be measured. On the other hand, it is easy to compare two expressions close to each other or with a subtle difference and so to obtain objectively stable measurements. In fact, the discrimination threshold measurements are known in psychophysics to be one of the most fundamental ways for substantial understanding of phenomena [10]. Furthermore, the discrimination thresholds are known to define a metric tensor describing intrinsic geometry in a Riemann space [23].

Thus, in this paper we measured the JND discrimination thresholds of facial expressions and draw them as ellipsoids in a 3D expression space, by simultaneous comparison between two facial expressions avoiding the influence of language and category judgments. The data is produced by morphing between different expression images. The results show that all the discrimination threshold ellipsoids have very different shapes and sizes for different basic expressions. The measurements from different subjects show similar trends in these variations. Since the discrimination thresholds as ellipsoids are subjectively unit spheres which have the same size and shape everywhere, it is a strong evidence that the expression space is not a Euclidean space. Indeed, these discrimination thresholds define local metric tensor for every expressions so a natural outcome is that one obtains the expression space as a Riemann space. We will show a finer distribution of 23 JND thresholds ellipsoids from a single subject in 2D and 3D PCA spaces, which show a smooth transition of local geometry among different expression and certain distinct features of the Riemann space.

## 2 Discrimination Thresholds and Riemann Space

### 2.1 Definitions of Thresholds and a Riemann Space

It is known that there are two types of thresholds in psychophysics [10]. The stimulus thresholds are the minimal stimulus to invoke a sensation. The other

type are the discrimination thresholds which are defined as the just noticeable difference (JND) between two simultaneously presented stimuli.

It is also known that a Riemann space is a space  $S$  in which a bilinear function or an inner product  $(\cdot, \cdot)$  is defined at the tangent space  $T_x S$  for every point  $x \in S$  such that

$$(dx, dy) := dx^T G(x) dy, \quad \forall dx, dy \in T_x S$$

where the matrix  $G(x)$  is smoothly defined for every  $x \in S$ . The  $G(x)$  defines local geometry such as distances and angles around the point  $x$ . It is known as the (Riemann) metric tensor [23].

It is well known that a color space is a Riemann space defined by the metric tensor obtained from the MacAdam's ellipsoids and ellipses [13, 14]. Similarly, we will use the JND thresholds near an expression as the unit sphere in local distance which describes the perceptual difference between slight variations of the expression in the expression space. This information at every point in the space then defines the Riemann metric tensor therefore giving the expression space the structure of a Riemann space.

It is reported in [8] that tests along a morphing image sequence between basic expression images show a sharp boundary between basic expressions A and B. This result is then used as a major evident, so called as a paradox against dimensional theory. In fact, those reported thresholds could be thresholds in categorical perceptions which are results of discontinuity between categories when the subjectives have no other choices besides A and B. In our measurement of the JND thresholds, as will be mentioned later, we will take care to avoid categorical judgments and influence of verbal labels of expression categories.

## 2.2 From Local Properties to Global Geometry

We hereafter consider a continuous space and its discrimination thresholds. It is known that it is easier for humans to judge relative and small differences than abstract levels of a stimulus. On the other hand, it is difficult to recover a global perception from a greater collection of local relative information. The celebrated Weber-Fechner law is an example showing how to obtain a global law in the whole space from local laws at every different point. In this case the logarithm function between the sensation and stimulus is obtained. Its success lies in the fact that the local thresholds are represented by an ODE which is fortunately integrable and one can obtain the global law in terms of elementary functions. In fact, the disagreement between the Weber-Fechner law and data comparing with Stevens' law is not due to this local to global integration strategy but to the error in the local Weber's law and Stevens' law can be obtained from integration using a correct local law. Unfortunately such closed forms are not always possible since an integration or solution of an arbitrary ODE, even when it exists, can rarely be expressed in a closed form in terms of elementary functions. On the other hand, nowadays one does not need such a closed form as long as the functional relationships are computable. e.g. a LUT will do almost the same if not better.

### 2.3 1D vs Higher Dimensions

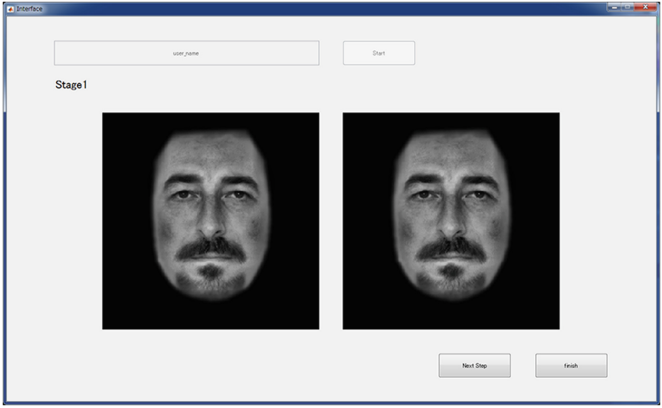
Another fact is that most theories in psychophysics were about stimulus-response of 1-dimensional data. This is understandable considering that the integration from local information to a global representation would become much harder when the stimuli and responses are high dimensional. In fact, it seemed there are few new additions after the Weber-Fechner and Stevens' laws. In particular, the only reports for discrimination thresholds researches in multi-dimensional case were those about color perception. The first 1D discrimination thresholds along straight lines in a color space were measured by Wright [12], which was followed by MacAdam [13] to measure discrimination thresholds in 2D or chromaticity plane, so-called MacAdam ellipses. The 3D discrimination thresholds which are ellipsoids in a color space were obtained by Brown-MacAdam [14]. One of major implications of these data is that the discrimination threshold show that a color space is not an Euclidean space but a distorted or curved space, which can be modeled as a Riemann space. In fact, colors lie on these discrimination thresholds, ellipses or ellipsoids, have the same subjective difference from the center colors, so they should form a unit circle or sphere and have the same size and shape everywhere. MacAdam ellipses show however various sizes and shapes for different center colors which then provide definite evidents that a color space is not a Euclidean space. The thresholds which define subjective distances depending on center colors then provide a Riemann metric for the color space [11]. Such intrinsic geometry is then used to apply the tools of Riemann geometry to find fruitful applications e.g. [21, 22].

In this research, we are aiming at an application of the above successful strategy to the perception of facial expressions.

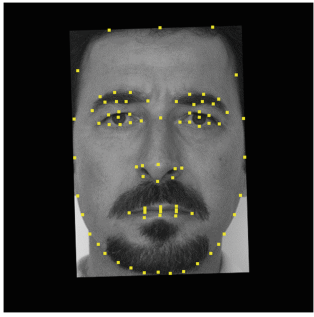
## 3 Measuring JND Discrimination Thresholds

We used images from two database of expression images: the Japanese Female Facial Expression (JAFPE) Database [15] and the Bosphorus Database [17–19]. The images contain seven basic expressions: Anger, Disgust, Fear, Sadness, Happiness, Surprise and Neutral and AU images from one person in each database. These images are used to create morphing image sequences combining all pairs among different expressions. This is done using the FUTON system (ATR Japan) [16] using the feature points shown in Fig. 2.

The test and a comparison expression image are shown simultaneously on a screen (Fig. 1). To avoid adaptation and prediction, a 2s interval (showing a neutral background) is inserted between sessions, the order of the images is randomized. The observer is instructed to compare the two expressions and answer if they are the “same” or “different” without interpreting them, and no hints on the labels of the expressions (emotions) are provided. We instruct the observer to ignore the type of the expression and to answer only if the images show identical expressions or not. By discouraging attempts to interpret the expressions we hope to minimize the influence of category perception during the process.



**Fig. 1.** Measurement environment

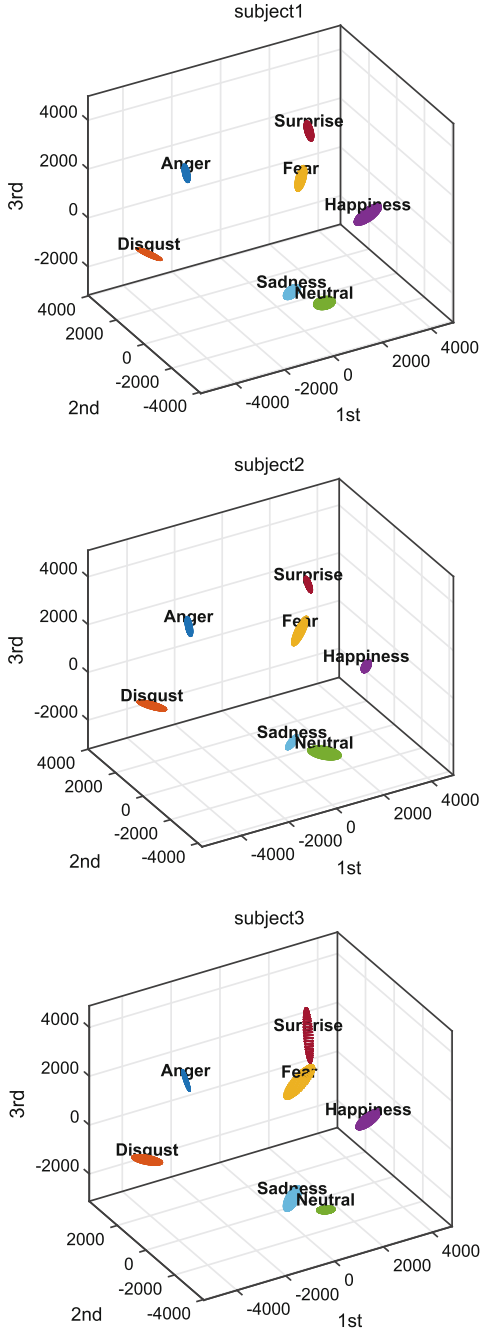


**Fig. 2.** Feature points

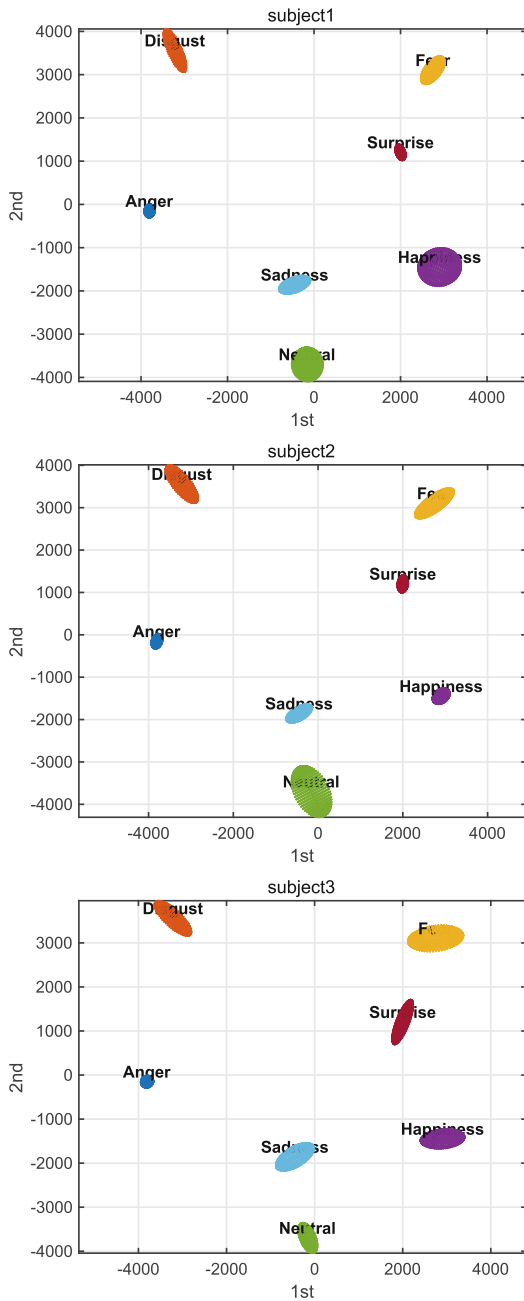
The observers are also informed in advance that the changes are about 1% so it is possible that they will not be able to discriminate the different images. The observers are also instructed to vote “different” when the decision takes more than 5 s, in order to avoid prediction, classification into emotion categories and to minimize the influence of other high-level cognition processes. There is a 5 min break after each session. Each comparison of two images was repeated three times, inserted among all image pairs in a random order.

The images and threshold data are then reduced by PCA using the covariance matrix and MDS to the first 2D and 3D eigen subspaces. The discrimination ellipsoids are then fitted using a Gaussian or RFB function fitting method.

In the first experiment we used 7 basic expression images from JAFFE to produce 21 morphing sequences. 2086 images are projected by PCA to 2D and 3D spaces while 126 threshold points were used to fit 7 ellipsoids shown in Figs. 3, 4 and 5. They show that the obtained expression spaces are not Euclidean, since JND ellipsoids define subjective unit spheres. Besides, the personal

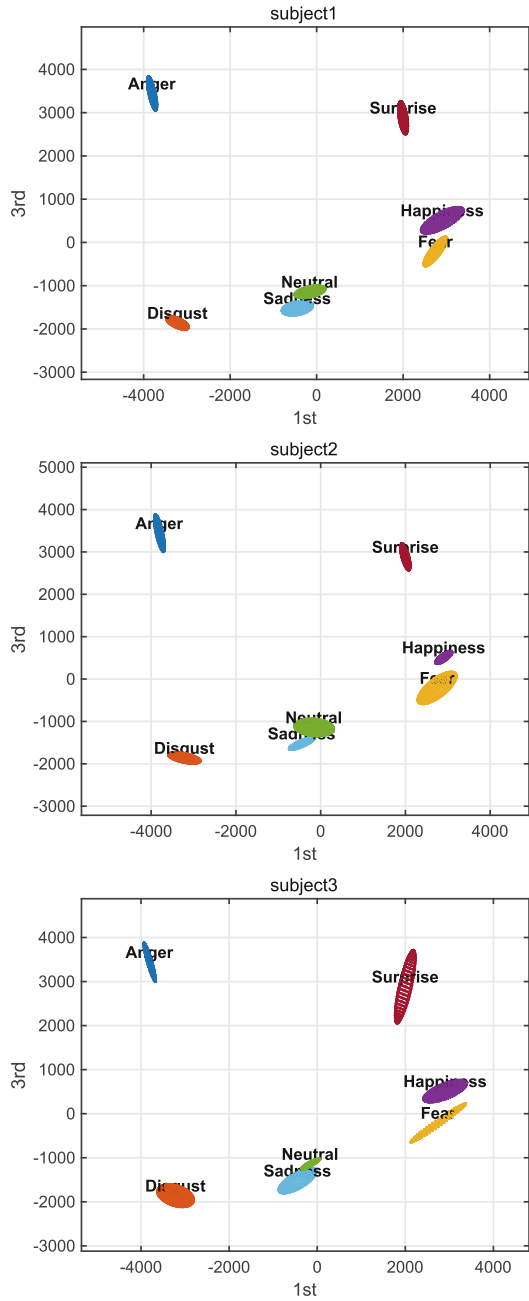


**Fig. 3.** Subject 1, 2, 3; JND thresholds in 3D PCA space

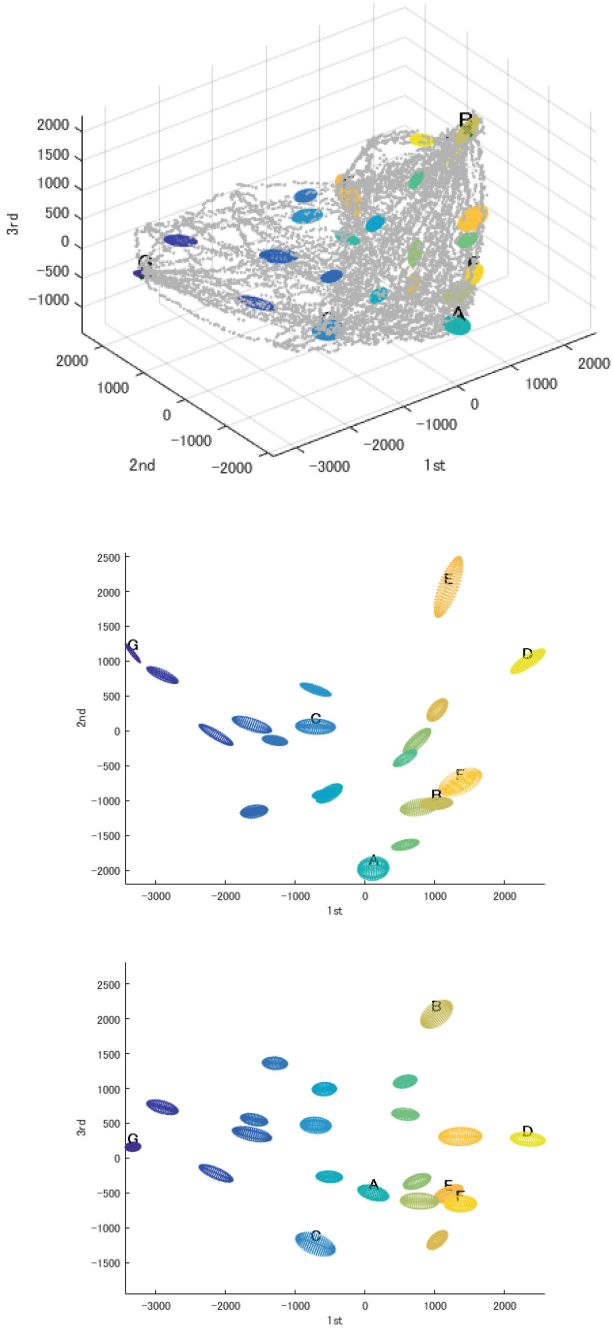


**Fig. 4.** Subject 1, 2, 3; JND thresholds in the 1st-2nd PCA space





**Fig. 5.** Subject 1, 2, 3; JND thresholds in the 1st-3rd PCA space



**Fig. 6.** 23 ellipsoids in the 3D, ellipses in 1st-2nd, 1st-3rd PCA spaces

variations between observers can be quite large. However, there are clear structural similarities among the JND thresholds distributions in the expression spaces of different observers.

In the second experiment we used 7 basic expressions and 21 AU images from the Bosphorus database to produce 616 morphing sequences. 57844 images were projected by PCA to 2D and 3D spaces, 2233 threshold points were used to fit 23 ellipsoids from a single observer. The results of 3D, 1st-2nd and 1st-3rd PCA projections are shown in Fig. 6. They indicate a smooth variation of shapes, directions and sizes of the ellipsoids and a global flow in the space, which indicate the metric tensor smoothly defined over the whole space and distinct features of the intrinsic geometry of the Riemann space.

## 4 Discussion, Conclusions and Future Work

The experiments require controlled conditions therefore are very labor intensive and time consuming. Further measurements are necessary to verify the results reported above and also to understand how the different steps in the experimental setup influence the final results. Besides, efficient approaches to produce natural and accurate morphing sequences of facial expressions are also important.

The facial image space is of very high dimensional, here we only illustrated JND sections in 2D and 3D subspaces of PCA space for easy visibility. JND thresholds in higher dimensional spaces can be obtained in the same way but with much more measurement data. Further work is therefore required to fix the expression space and estimate its dimension. Dimensionality reduction and other possibility of psychophysical spaces may also be considered. A problem to use features extracted from facial images is the difficulty to produce morphing sequences corresponding to trajectories of shifts in the feature space.

Applications are expected in theoretical modeling, performance improvement in facial expression analysis and recognition, etc. We are, however, confident that the Riemannian approach to expression analysis is necessary and that it will play a similarly important role as in the study of color perception.

**Acknowledgment.** The authors wish to thank Hajime Tasaki for discussions and cooperation in this research.

## References

1. Li, S.Z., Jain, A. (eds.): Handbook of Face Recognition. Springer, New York (2011). <https://doi.org/10.1007/978-0-85729-932-1>
2. Benitez-Quiroz, C.F., Srinivasan, R., Martinez, A.M.: Emotionet: an accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016)
3. Du, S., Tao, Y., Martinez, A.M.: Compound facial expressions of emotion. Proc. Nat. Acad. Sci. **111**(15), E1454–E1462 (2014)

4. Ekman, P., Friesen, W.V.: Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto (1978)
5. Schlosberg, H.: The description of facial expression in terms of two dimensions. *J. Exp. Psychol.* **44**(4), 229–237 (1952)
6. Russell, J.A., Bullock, M.: Multidimensional scaling of emotional facial expressions: similarity from preschoolers to adults. *J. Pers. Soc. Psychol.* **48**(5), 1290–1298 (1985)
7. Richardson, M.W.: Multidimensional psychophysics. *Psychol. Bull.* **35**(9), 659–660 (1938)
8. Calder, A.J., Young, A.W., Perrett, D.I., Rowland, D.: Categorical perception of morphed facial expressions. *Vis. Cogn.* **3**, 81–117 (1996)
9. Takehara, T., Ochiai, F., Suzuki, N.: Fractals in emotional facial expression recognition. *Fractals* **10**, 47–52 (2002)
10. Gescheider, G.A.: Psychophysics, The Fundamentals, 3rd edn. Lawrence Erlbaum Associates, Mahwah (1997)
11. Wyszecki, G., Stiles, W.S.: Color Science; Concepts and Methods, Quantitative Data and Formulae, 2nd edn. Wiley, New York (1982)
12. Wright, W.D., Pitt, F.H.G.: Hue-discrimination in normal color-vision. *Proc. Phys. Soc.* **46**(3), 459–473 (1942)
13. MacAdam, D.L.: Visual sensitivities to color differences in daylight. *J. Opt. Soc. Am.* **32**(5), 247–274 (1942)
14. Brown, W.R.J., MacAdam, D.L.: Visual sensitivities to combined chromaticity and luminance differences. *J. Opt. Soc. Am.* **39**(10), 808–834 (1949)
15. Lyons, M.J., Akamastu, S., Kamachi, M., Gyoba, J.: Coding facial expressions with Gabor wavelets. In: 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200–205 (1998)
16. Mukaida, S., Kamachi, M., Oda, M., Kato, T., Yoshikawa, S., Akamatsu, S., Chihara, K.: Facial image synthesis system: FUTON: evaluation as tools for cognitive research on face processing. *J. Inst. Electr. Inf. Commun. Eng.* **A 85**(10), 1126–1137 (2002)
17. Savran, A., Alyüz, N., Dibekliöglu, H., Çeliktutan, O., Gökberk, B., Sankur, B.L., Akarun, L.: Bosphorus database for 3D face analysis. In: The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde University, Denmark, 7–9 May 2008
18. Savran, A., Sankur, B.L., Bilge, M.T.: Comparative evaluation of 3D versus 2D modality for automatic detection of facial action units. *Pattern Recogn.* **45**(2), 767–782 (2012)
19. Savran, A., Sankur, B.L., Bilge, M.T.: Regression-based intensity estimation of facial action units. *Image Vis. Comput.* **30**(10), 774–784 (2012)
20. Young, A.W., Rowland, D., Calder, A.J., Etcoff, N.L., Seth, A., Perrett, D.I.: Facial expression megamix: test of dimensional and category accounts of emotion recognition. *Cognition* **63**(3), 271–313 (1997)
21. Mochizuki, R., Kojima, T., Lenz, R., Chao, J.: Color-weak compensation using local affine isometry based on discrimination threshold matching. *J. Opt. Soc. Am. A Optica Image Sci. Vis.* **32**(11), 2093–2103 (2015)
22. Oshima, S., Mochizuki, R., Lenz, R., Chao, J.: Modeling, measuring, and compensating color weak vision. *IEEE Trans. Image Process.* **25**(6), 2587–2600 (2016)
23. do Carmo, M.P.: Riemannian Geometry. Birkhäuser, Boston (2013)