Objective Evaluation of Impression of Faces with Various Female Hairstyles Using Field of Visual Perception

Naoyuki AWANO, Member and Kana MOROHOSHI, Nonmember

SUMMARY
Most people are concerned about their appearance, and the easiest way to change the appearance is to change the hairstyle. However, except for professional hairstylists, it is difficult to objectively judge which hairstyle suits them. Currently, oval faces are generally said to be the ideal facial shape in terms of suitability to various hairstyles. Meanwhile, field of visual perception (FVP), proposed recently in the field of cognitive science, has attracted attention as a model to represent the visual perception phenomenon. Moreover, a computation model for digital images has been proposed, and it is expected to be used in quantitative evaluation of sensitivity and suitability called “kansei.” Quantitative evaluation of “goodness of patterns” and “strength of impressions” by evaluating distributions of the field has been reported. However, it is unknown whether the evaluation method can be generalized for use in various subjects, because it has been applied only to some research subjects, such as characters, text, and simple graphics. In this study, for the first time, we apply FVP to facial images with various hairstyles and verify whether it has the potential of evaluating impressions of female faces. Specifically, we verify whether the impressions of facial images that combine various facial shapes and female hairstyles can be represented using FVP. We prepare many combinational images of facial shapes and hairstyles and conduct a psychological experiment to evaluate their impressions. Moreover, we compute the FVP of each image and propose a novel evaluation method by analyzing the distributions. The conventional and proposed evaluation values correlated to the psychological evaluation values after normalization, and demonstrated the effectiveness of the FVP as an image feature quantity to evaluate faces.

key words: hairstyle, facial shape, field of visual perception, kansei, objective evaluation

1. Introduction
Most people are concerned about their appearance. While there are many means to change looks, changing the hairstyle is one of the easiest and most effective means. However, it is difficult to objectively judge the type of hairstyle that suits them. Currently, many professional hair stylists have propounded theories about hairstyles based on their senses and generally mention that the ideal facial shape is oval in terms of suitability to various hairstyles [1]; however, these theories have no scientific evidence and hence are difficult to evaluate.

Meanwhile, in the last ten years, impressions of many objects have been evaluated in kansei engineering. Many studies have employed statistical approaches such as the semantic differential method or the magnitude estimation method [2]–[5]. In these methods, images are displayed and the impressions of subjects in the images are evaluated by conducting a subjective questionnaire survey. Therefore, these methods cannot quantitatively and automatically evaluate objective impressions from images using only computers.

In cognitive science, visual information processing has attracted attention. In particular, field of visual perception (FVP) has been proposed by Yokose as a model to represent the visual perception phenomenon [6]. He hypothesized that there exists a field around the objects to be perceived, as shown in Fig. 1. The existence of the field has been proven via physiological and psychological experiments. Moreover, Nagaishi has proposed a computational model for calculation of FVP in digital images, and it is expected to be used in engineering applications [7]. In particular, he proposed the methods for evaluating impressions and sensitivities [8] and showed that evaluating the distributions of FVP is equivalent to evaluating “goodness of patterns” and “strength of impressions.” However, it is unknown whether the evaluation method can be generalized to a wide variety of subjects because conventional studies have applied it only to some research subjects, such as characters, text, and simple figures.

In this study, we apply FVP for the first time to facial images with various hairstyles and verify whether it has the potential to evaluate impressions of female faces. In this paper, we verify whether the goodness of the impressions of facial images that are obtained by combining various facial shapes and female hairstyles can be represented using FVP. Specifically, we prepare many combinational images of facial shapes and hairstyles and conduct a psychological experiment to evaluate impressions. Moreover, because conventional evaluation methods using FVP are not used for evaluating face, we further propose a dedicated evaluation method. Thereafter, we demonstrate the effectiveness...
of FVP by comparing the conventional and proposed evaluation values with the evaluation values of the psychological experiments.

In Sect. 2, we introduce FVP and its evaluation method. In Sect. 3, we describe psychological experiments and calculations using the novel FVP evaluation method. In Sect. 4, we show the experimental results and discuss the effectiveness. Finally, we state our conclusions in Sect. 5.

2. Field of Visual Perception (FVP)

2.1 Overview

Yokose had hypothesized the presence of an induction field around objects in vision similar to an electrostatic field. Thereafter, the field was discovered by his psychological experiments[6]. The adequateness of the existence has been proven by physiological and psychological experiments[9], [10]. For line segments shown in Fig. 2 (a), FVP can be calculated by calculating the potential energy around them. For instance, the potential energy $M_p$ of a point $p$ can be calculated by Eq. (1).

$$M_p = \frac{1}{N} \sum_{i=1}^{N} \int ds f_i(s)$$  \hspace{1cm} (1)

where $N$ is the number of line segments, $s$ is a point on a line segment, and $f_i(s)$ is distance function from $p$ to $s$. Note that $s$ is a point on visible portions as shown in Fig. 2 (b).

Nagaishi has proposed a computational model of FVP in digital images based on above theory [7]. FVP in digital images applies only for binary images that consist of black pixels as objects and white pixels as background. The potential energy $M_p$ of a white pixel is computed from the contour pixels of objects using Eq. (2).

$$M_p = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{d_i}$$  \hspace{1cm} (2)

where $n$ is number of visible black pixels from a pixel $p$, and $d$ is the Euclidean distance from $p$ to visible black pixels. Figure 1 (b) shows the visualization result of assigning high-intensity values to high potential values, and Fig. 1 (c) shows the visualization result of connecting equipotential values like contour lines. From these figures, you can see that the pixels close to the objects have high potential values.

2.2 Kansei Evaluation

The commonly used image feature values cannot be used for evaluating kansei features such as impressions. Hence, researchers are trying to use FVP as an image feature for evaluating kansei. Nagaishi proposed an evaluation method for the quality, ease of reading, and beauty of handwritten characters [11]. Onaga evaluated the design of arch bridge [12], and Miyoshi evaluated spacing for character arrangement [13], [14]. Moreover, Nagaishi considered generalization of these methods and obtained good results [8].

2.2.1 Goodness of Patterns

In conventional studies, it has been hypothesized that a simple and low-energy state in which the shape of the equipoitential lines approach a circle is good [8]. In general, the goodness of patterns $C$ is calculated using Eq. (3).

$$C(m) = \frac{\text{len}(L_m)^2}{S(L_m)}$$  \hspace{1cm} (3)

where $L_m$ is a potential line with a potential energy $m$, $\text{len}(\cdot)$ is the length, and $S(\cdot)$ is the internal area. Note that there are cases where there are multiple equipotential lines with $m$, as shown in Fig. 3. In this case, $\text{len}(\cdot)$ and $S(\cdot)$ are calculated as follows:

$$\text{len}(L_m) = \sum \text{len}(l_i)$$  \hspace{1cm} (4)

$$S(L_m) = \sum S(l_i)$$  \hspace{1cm} (5)

where $L_m = \{l_1, l_2, \ldots\}$. The parameter $C$ is called “complexity”; the lower the value of $C$, the better the pattern. The mean value or the transition of the complexities are used when multiple equipotential lines are evaluated.

2.2.2 Strength of Impressions

The parameter strength of impressions is equivalent to the strength of potential energy [15]. Moreover, since FVP can
be interpreted as Coulomb potential [7], the whole field can be considered as potential energy. From these findings, the strength of impressions is defined as potential energy $E$, as shown in Eq. (6).

$$E = \int mS(L_m) \, dm$$ (6)

3. Method of Verification

In this paper, we verify whether the goodness of the impressions of facial images that are obtained by combining various facial shapes and female hairstyles can be represented using FVP and confirm whether FVP has the potential to evaluate impressions of faces with various hairstyles. We first prepare drawings of various facial images. Next, we conduct a psychological experiment using the facial images. Moreover, we propose an evaluation method for faces by improving on the conventional complexity measure. Finally, we compare the results of a psychological experiment with those of quantitative evaluation to verify whether the results are similar.

3.1 Constraints

Faces have a great variety because they consist of many elements such as hairstyles, hair colors, and facial shapes, each of which can vary widely. As a practical matter, we cannot evaluate impressions of all faces; therefore, we set the following constraints to faces.

a) Japanese females
Constraining the nationality to Japanese, we conduct our questionnaire experiments only on Japanese people.

b) Black hair
Although many females dye their hair, Japanese generally have black hair; hence, we constrain the hair color to black.

c) No facial parts
Our objective is to verify whether the impressions of facial images that are obtained by combining various facial shapes and hairstyles can be represented using FVP. We use images with no facial parts because if facial parts are included, there is a possibility that the participants would evaluate whether a hairstyle suits a specific face rather than evaluating whether a hairstyle that suits a specific facial shape.

Additionally, even with excluding the variety in colors and facial parts, there are many varieties of hairstyles and facial shapes. Hence, we set the following constraints.

i) Five types of facial shapes
As typical Japanese facial shapes, we use oval, round, long, square, and inverted triangle, as shown in Fig. 4 based on [1].

ii) Three types of hair lengths
Female hairstyles are generally classified by length. As typical hair lengths, we use short hair, medium hair, and long hair.

iii) Three types of bangs
In hairstyles, bangs as well as lengths are important. We only use the popular bangs “Full,” “Side,” and “Center,” with their definitions as follows:

**Full:** Thick and straight-cut hair that cover the forehead.

**Side:** Side parting.

**Center:** Center parting or exposed forehead.

Moreover, to simplify the experiment and to equalize the conditions, we set constraints that ears are hidden by hair, thickness of hair is the same as much as possible, and the jaw is not hidden by hair. Furthermore, as hair on the back of a head is hidden by the neck or body, using only a head would look unnatural to participants in an evaluation, particularly when long hair is present. On the other hand, inclusion of a body introduces the possibility that the face might not be adequately evaluated. Thus, in this study we attach only necklines with equal sizes to each head and define the hair as being in front of a body. We drew 45 pictures under these constraints and classified them to nine categories as shown in Table 1.

3.2 Psychological Experiment

We evaluate the facial impressions of females using images shown in Table 1. For subjective evaluation in a psychological experiment, we employ a paired comparison method. There are two main methods of paired comparison: Thurstone’s method and Scheffe’s method [16], [17]. In Thurstone’s method, two objects are compared and a judgement is made on which is the better of the two. In Scheffe’s method, two objects are compared and a judgement is made...
<table>
<thead>
<tr>
<th>Facial images used in our experiment.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oval</strong></td>
</tr>
<tr>
<td>Short hair (Pull)</td>
</tr>
<tr>
<td>Short hair (Side)</td>
</tr>
<tr>
<td>Short hair (Center)</td>
</tr>
<tr>
<td>Medium hair (Pull)</td>
</tr>
<tr>
<td>Medium hair (Side)</td>
</tr>
<tr>
<td>Medium hair (Center)</td>
</tr>
<tr>
<td>Long hair (Pull)</td>
</tr>
<tr>
<td>Long hair (Side)</td>
</tr>
<tr>
<td>Long hair (Center)</td>
</tr>
</tbody>
</table>
on which is the better of the two and by what degree. Since Scheffe’s method involves evaluation of the degree of impressions, the burden imposed on the participants is larger.

In this study, since we use 45 images, the number of the combinations evaluated by a paired comparison method is \(45 \times 44 / 2 = 990\), which is intractable. Our objective is to verify whether FVP can evaluate the impressions of facial images that obtained by combining various facial shapes and hairstyles. Therefore, we conduct the experiment for each of the nine categories shown in Table 1. Accordingly, the number becomes \(5 \times 9 = 90\); this number is still large. Hence, we employ Thurstone’s method. Two images are displayed next to each other, and the participants choose the image with a better impression as a female.

For a more efficient and accurate experiment, we built a questionnaire system. In the system, two images belonging to the same category are displayed next to each other, and the participants click on the image they believe to have a better impression. The next set of images are automatically displayed after the participants click on one of the images. The next set of images are randomly selected from the remaining combinations. The previous images are faded out and the new images are faded in, with gray images displayed during the transition to prevent display of residual images of the previous images.

We conducted the psychological experiment on 112 Japanese participants (52 males and 60 females, ages 10–60 years) after obtaining an informed consent from them.

3.3 Evaluation of Distributions of FVP for Faces

A face that evokes a good impression is labelled with various adjectives, such as cute, beautiful, and elegant. Our evaluation does not have a bias to a specific adjective. As mentioned in Sect. 2.2.1, the complexity obtained using Eq. (3) based on FVP evaluates the goodness of patterns. The complexity is calculated based on circularity; however, the human facial shape is generally not an exact circle. Moreover, the ideal facial shape is said to be oval. We therefore change the complexity to include evaluation of ellipses because the goodness of impression of a face might be based on its similarity to an ellipse.

First, an approximate ellipse \(\tilde{l}_j\) of equipotential line \(l_j\) of \(L_m\) is calculated based on [18]. If there are multiple equipotential lines in \(L_m\), approximate ellipses are calculated for each \(l_j\). Then, we calculate the ellipses under the condition \(\text{len}(l_j) \geq th\) because an ellipse cannot be calculated if its length is too short.

Next, to produce values that are larger for more complicated results (in a manner similar to the conventional complexity \(C\)) we calculate the error between \(l_j\) and \(\tilde{l}_j\) using the evaluation function \(Er\), as follows:

\[
Er(l_j, \tilde{l}_j) = \frac{Pr(\tilde{l}_j) + e(l_j, \tilde{l}_j)}{Pr(l_j) - e(l_j, \tilde{l}_j)} \quad (7)
\]

\[
e(l_j, \tilde{l}_j) = \text{len}(l_j) - Pr(l_j) \quad (8)
\]

where \(Pr(l_j)\) is the perimeter of \(l_j\), and is calculated using the following equation:

\[
Pr(l_j) = 4a \int_0^2 \sqrt{1 - k^2 \sin^2 t} dt \quad (9)
\]

where \(k = \sqrt{1 - \frac{b^2}{a^2}}\) is the eccentricity of an ellipse, \(2a\) is the major axis, and \(2b\) is the minor axis. Then, for the case that \(L_m\) contains multiple \(l_j\), their \(l_j\) are multiplied in order to evaluate them equally. Hence the proposed error function \(Er(L_m)\) is defined as follows:

\[
Er(L_m) = \sqrt{\frac{1}{|L_m|} \sum_{j=1}^{|L_m|} Er(l_j)} \quad (10)
\]

Furthermore, we consider the size of each equipotential line because the error is based on a ratio. Specifically, the equipotential lines distant from the face are normally long but those that are close are short. Thus, as we focus on regions near the face when seeing it, the error is divided by the length, and the proposed complexity \(C'\) is finally defined as Eq. (11):

\[
C'(m) = \frac{Er(L_m)}{\text{len}(L_m)} \quad (11)
\]

In other words, this equation gives the error per unit length when each equipotential line is approximated by an ellipse.

4. Experimental Results

4.1 Experimental Settings

We conducted experiments to verify whether FVP can be used to evaluate impressions of female faces. It is necessary to set sufficient margin around each facial image shown in Table 1 for evaluating the distributions. Accordingly, in this study, we prepared white images as background and set their resolution to \(768 \times 768\), and each of the facial images was placed in the center of the white images with enough margin; The size of the face is half of the resolution. Moreover, in this study, \(th\) is set based on 5% of the length of one side of an image.

Moreover, the evaluation values—psychological evaluation value, conventional complexity \(C\), and proposed complexity \(C'\)—all have different dimensions and cannot be compared directly. Hence, we convert each of them to a standard score \(Z_i\) [19] as follows:

\[
Z_i = \frac{10 \cdot \frac{x_i - \mu}{\sigma}} + 50 \quad (12)
\]

where \(x_i\) is a value of data set \(X\), \(\mu\) is the average value, and \(\sigma\) is the standard deviation.

4.2 Results and Discussion

4.2.1 Psychological Evaluation

The results of the psychological evaluation are shown in
4.2.2 Evaluation Method and Range

Although the method for evaluating complexity varies in conventional studies, we employed the mean value of each of the complexities, as in [11], to conduct our comprehensive evaluation. Accordingly, it was necessary to define the evaluation range \([r_l, r_u]\) for potential energy.

In this study, the lower limit \(r_l\) and the upper limit \(r_u\) of the range were defined through supplemental investigations. We set \(r_l = 0.005\), which is the minimum value not crossing the frame of the image. We also choose \(r_u\) to minimize the error between the psychological and complexity evaluations. Figure 6 shows the mean absolute errors. The value of \(r_u\) with the lowest error differs for different hairstyles; however, those values are roughly similar. Accordingly, we define \(r_u = 0.045\) for \(C\) and \(r_u = 0.030\) for \(C'\). Both values of \(r_u\) are so low that it is possible that using a large \(r_u\) would result in evaluating more detailed parts of the shape rather than just the facial shape.

4.2.3 Overall Results

The error of \(C\) is larger than \(C'\) for all hairstyles, as shown in Fig. 6. This indicates that \(C'\) is better than \(C\) as a quantitative evaluation value. Additionally, the correlation diagrams are shown in Fig. 7, with the plotted points representing respective hairstyles.

We used a significance level of .01 in this study. Defining \(R(x)\) is the correlation coefficient of \(x\), we obtained \(R(C) = .575\) (p-value: \(.359 \times 10^{-4} < .01\)) and \(R(C') = .949\) (p-value: \(.375 \times 10^{-22} < .01\)). This indicates both complexities are correlated to psychological evaluation values. We validated the difference between \(R(C)\) and \(R(C')\) and evaluated whether the difference is sufficiently large; we determined that \(R(C')\) is larger than \(R(C)\) with a p-value calculated by Meng-Rosenthal-Rubin method [20] of \(.982 \times 10^{-10} < .01\). This indicates there is a significant difference, suggesting that the proposed \(C'\) can perform significantly better than conventional \(C\) in evaluating impressions of faces.

Figure 7 shows the standard errors at a confidence interval of 90% as gray lines. In the graph for complexity measure \(C'\), six hairstyles fall outside of the lines, namely, “Center” bangs for all round and long facial shapes. This is considered to be caused by a trend peculiar to the Japanese. According to Muta [21], the aspect ratios of shapes preferred by Japanese are close to \(1:1\) or \(1:1.414\) (the silver ratio), whereas those preferred by people of other nationalities are approximately \(1:1.618\) (the golden ratio). Accordingly, too long a face might make a poor impression on Japanese people. When actually seeing the images of the bangs “Center” with long faces in Table 1, since a large part of the forehead is visible, it might give an impression that the faces are too long and hence not as attractive as the round faces. Therefore, the psychological evaluation values of “Center” tend to differ slightly from those of the other bangs “Full” and “Side,” and this is considered to be a factor for the difference in values between psychological evaluation and the
proposed complexity.
Moreover, both correlation coefficients are positive, indicating that “goodness of impressions” is correlated to conditions in which the equipotential lines are complex. As mentioned above, conventional complexity indicates “good patterns” when the value is low. The trend here is in contrast with the conventional trend and indicates that “good impression” is different from “good pattern.” This is because the potential energy depends upon the number of visible pixels \( n \), as shown in Eq. (2). In other words, places where \( n \) differs significantly from surrounding values of \( n \) are where the shape of the equipotential line becomes distorted. For instance, in the case of a rectangle there are two background pixel conditions: one side is visible, or two sides are visible. Equipotential lines become distorted at locations where \( n \) differs significantly close to their borders. In the case of a circle, by contrast, the number of visible pixels, \( n \), remains approximately constant and therefore the equipotential lines are smooth. On the other hand, for an ellipse or egg shape the shape of the contour smoothly and constantly changes, leading to small, constant changes in \( n \). This is reflected in middle- to high-frequency changes in the shape of the equipotential lines.
To illustrate this, Fig. 8 shows a clothoid curve, in which the curvature changes linearly with curve length. The equipotential lines around the linear part are smooth but around the curves become relatively rougher. This is because pixels around curves with changing curvature change in terms of \( n \). Thus, because oval face do not have more linear parts than other facial shapes, the equipotential lines actually seen for an oval face become relatively complex (Fig. 9).

4.2.4 Evaluation of Variability
All the above results are based on standard scores obtained by Eq. (12) to unify the dimensions. As the scores are divided by their standard deviation (SD) in the equation, the variability of the values is not reflected in the results. Hence, SDs are shown in Fig. 10 for comparing the variabilities of the different values. The conventional and proposed complexity values are separately graphed because their dimensions are different. Since the values represent the variance between good and bad impression, high values indicate that the face looking good or bad will be clearly distinguished.

Compared with trend in the SD of the psychological evaluation values, the SD of complexity \( C \) shows a different trend, whereas the complexity \( C' \) shows a very similar trend. All evaluation values have different dimensions and it is difficult to compare them directly; however, it can be confirmed that the complexity \( C' \) is also effective regarding
variability. In conjunction with the results of the previous section, this suggests that FVP has great potential for evaluating facial impressions. This is in particular more evident when the proposed measure $C'$ is used than when $C$ is used. However, caution is required in interpretation of the results as the paired comparison method we employed is a relative evaluation. We separately conducted the psychological experiments in nine categories. Hence, it is difficult to distinguish which hairstyle is more likely to suit faces. It is therefore necessary to conduct additional experiments if we clarify that.

5. Conclusions

In this study, we verified an objective evaluation method for faces using FVP. Specifically, we verified whether the impressions of facial images that are obtained by combining various facial shapes and female hairstyles can be represented using FVP. We prepared various images combining several hairstyles and facial shapes and conducted psychological experiments. In addition, we proposed evaluation method based on the distributions of FVP for human faces. Then we compared the conventional and proposed evaluation values with the psychological evaluation values and evaluated the effectiveness of FVP.

The results showed that the conventional and proposed evaluation values correlated to the psychological evaluation values when both are normalized by their standard deviations; the proposed evaluation values strongly correlated to the psychological evaluation values. It was also shown that “goodness of impressions” does not equal conventional “goodness of patterns,” however, FVP has a great potential of being able to evaluate faces with various hairstyles.

In this study, the evaluation range of FVP was experimentally determined; however, in future work we will need to validate the experimentally determined range. Furthermore, as mentioned in the discussion, we cannot easily determine the hairstyle that is more likely suit a particular facial shape because our psychological experiment was separately conducted in several categories. Hence, we should still evaluate the differences of goodness between these categories. Another limitation of this study is that the research was limited to Japanese people. For generalizing the evaluation algorithm, we therefore also need to conduct experiments on different ethnic groups and determine on how different ethnic groups perceive hairstyles for faces outside of their own ethnic groups.

Acknowledgments

We would like to express the deepest appreciation to the barbershop WEATHER COCK (Mr. Katsuro Ikeda, Mr. Masahiro Ikeda, and Mrs. Mitsuko Ikeda) in Fuchu-shi, Tokyo.

References

[1] H. Shirasaka, Salon de tsukaeru hea dezain no kyoukasho [A hair design textbook that can be used at a salon], JOSEI MODE SHA Co., Ltd., 2012. (in Japanese)


Naoyuki Awano received the BIS, MIS, and DIS degrees from Osaka Institute of Technology, Osaka, Japan, in 2007, 2009, and 2012, respectively. He was a research associate at Osaka Institute of Technology in 2009–2013. He was an assistant professor at Seikei University in 2013–2018. Since 2018, he has been a lecturer at Osaka University of Economics. His research interests are in the areas of image processing and computer graphics.

Kana Morohoshi received the BST degrees from Seikei University, Tokyo, Japan, in 2017. She works at Tokio Marine & Nichido Systems Co., Ltd.