IFSCC 2025 full paper (IFSCC2025-1041)

"Proposal for a new method to realize comprehensive evaluation of facial color."

Kaoru Yogosawa [1,*], Ko Ohira [2], Yuki Matsuo [1], Takahito Nakai [2] and Hideki Nishiura [2]

[1] TOA Inc., Yokohama, Japan; [2] TOA Inc., Osaka, Japan

1. Introduction

Humans obtain a significant amount of information from faces during interpersonal interactions. Numerous studies have examined how people form impressions of others. For example, it has been shown that impressions can be formed remarkably quickly when observing another person's face [1]. Facial impressions are influenced by various factors, including facial features such as the eyes and nose, expressions, color, and other visual cues. Among these, facial color is considered to play a particularly important role. Research on facial skin color has been conducted across a wide range of disciplines, including studies on the key pigments that determine skin tone, such as melanin and hemoglobin, as well as on how facial color is perceived [2-4]. However, it is still not fully understood how facial color contributes to facial impressions. In considering racial diversity, exploring new values related to facial color is essential, and gaining a deeper understanding of the relationship between facial color and facial impressions is particularly important.

Humans perceive visible light in the wavelength range of 380 to 780 nm. When light from a source strikes an object, part of it is absorbed, and the rest is reflected. The reflected light enters the eye and is processed by the brain to produce the perception of color. Thus, facial color is influenced not only by the absorption characteristics of the object, such as melanin and hemoglobin, but also by various factors, including the spectral distribution of the light source and elements of ophthalmological optics of observers dependent on age and environment [5]. Furthermore, facial color is affected by gloss and shading, which are caused by the three-dimensional structure of facial features like the nose and cheeks. Therefore, in studies of facial color, it is important to consider whole and spatial color information of the face. To achieve this, we employed hyperspectral imaging (HSI) to obtain spectral data at each pixel across a wide range. HSI provides comprehensive spectral information, offering significantly richer chromatic content than RGB data, and thereby enabling more detailed color analysis.

In analyzing the detailed color information obtained by HSI, the influence of individual differences in facial shape could pose a challenge. To reduce this influence, we employed the technique of the average face. Since its inception in 1878, the average face has been a long-utilized method [6, 7]. It is widely recognized that average faces emphasize the features common to the individual faces used in their construction. By using the average face, it becomes possible to minimize the influence of individual differences in facial shape, thereby enabling, a more focused investigation on color attributes becomes possible. Based on the above, we employed this technique.

In this study, we present a novel method utilizing an average face constructed from HSI data. This approach enables evaluation independent of individual differences in facial shape and allows for the quantification of facial features common to the faces constituting the average face. Furthermore, we report an investigation conducted using this method to explore the spectral characteristics of the face that contribute to facial impressions.

2. Materials and Methods

2.1. Participants

The participants were 10 Japanese women, aged between 23 and 30. To observe changes attributable to color, participants of a younger age were selected before the appearance of pronounced age-related facial changes such as wrinkles and sagging. All participants were informed of the purpose of the study. They received detailed information about the procedures and gave written informed consent before enrolment.

2.2. Acquisition of HSI data

Hyperspectral imaging (HSI) data of individual faces were acquired using a 2D spectroradiometer (Topcon Technohouse Corp., SR-5100HWS) under illumination from two artificial solar lights (Seric, XC-100AF) in a dark room. HSI data were measured in the visible light range (390-780 nm) with a 10 nm interval. The hyperspectral images were captured at a spatial resolution of 0.135 mm, resulting in images of 1260 × 1560 pixels. Spectral data were recorded at each pixel, enabling the acquisition of spatially resolved spectral information across the entire face. Before measurement, the participants washed their faces with a face cleanser and rested for 10 minutes in a measurement room.

2.3. Generating average faces from HSI data

The 76 characteristic morphological landmark points were obtained automatically from the face images of all participants. The coordinates of the landmarks of the average face were determined by averaging the coordinates of the corresponding points for each image. HSI data of each participant were affinely transformed using the Delaunay triangulation method based on the coordinates of morphological landmark points (**Figure 1**). To obtain the RGB image of the average face, the affinely transformed HSI data were converted to RGB values according to the color matching functions.

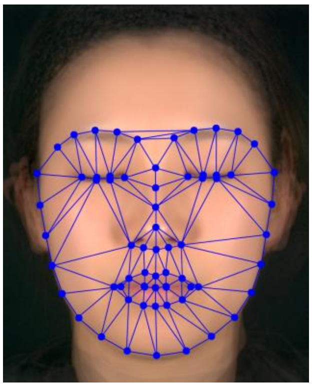


Figure 1. An example image of Delaunay triangulation.

Next, to evaluate facial impressions, we generated a large number of average faces. We repeatedly created an average face by selecting three out of ten participants. Consequently, 120 different average faces were generated. When creating the average faces for the three individuals, HSI data were transformed using the average landmark coordinates obtained from the images of ten individuals.
All procedures for creating the average face from HSI data were carried out using Python 3.12.

2.4. Scoring of facial impressions

Three experts trained in the evaluation of facial impressions assigned scores for the stimulus images using a visual analog scale. Based on the previous study [8], six impression evaluation terms were established: “Elegant”, “Clear”, “Bright”, “Warm”, “Lively”, and “Healthy”. All scorings were performed on the same display (15.6-inch) and conducted in a dark room, with the distance between the evaluator and the monitor standardized to 60 cm. The 120 images of the average faces were randomly assigned to three sets, each consisting of 40 images. The evaluations were conducted with sufficient breaks between each set. The images within each set were presented in a randomized sequence. Each image remained continuously displayed until the evaluator completed their response. A schematic representation of the evaluation interface is shown in **Figure 2**. The interface was primarily presented in Japanese, with impression terms additionally displayed in English.

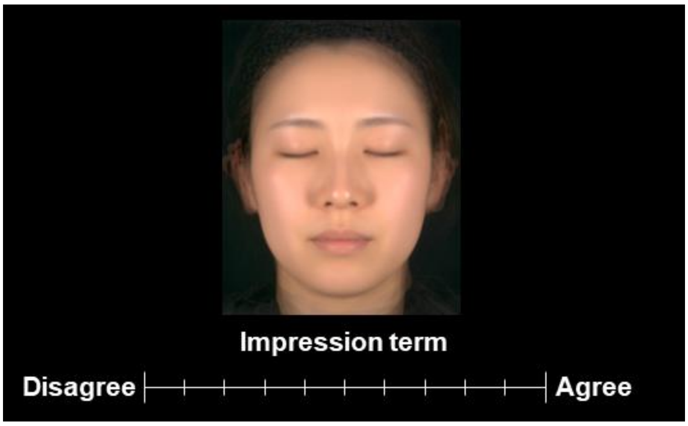
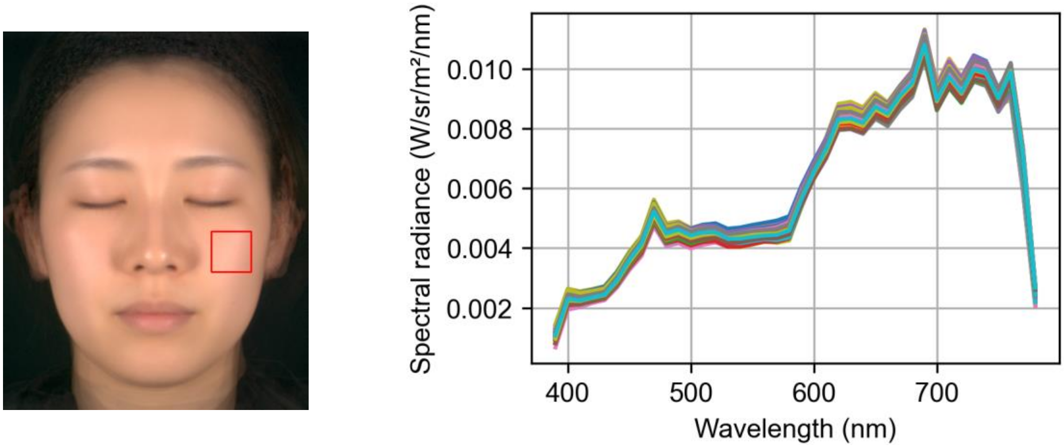


Figure 2. Schematic diagram of the monitor display used during the evaluation.

2.5. PLS regression analysis

Partial Least Squares (PLS) regression analysis was conducted to explore the relationship between scores of each impression and spectral radiance ($W/(sr/m^2/nm)$) of each wavelength included in HSI data. The analysis was performed using averaged spectral data within the left cheek region, which was commonly defined across all average faces based on the coordinates of morphological facial landmarks (**Figure 3**). Before analysis, the spectral data were preprocessed using the Standard Normal Variate (SNV) transformation, which standardizes each spectrum by subtracting its mean and dividing by its standard deviation. Prior to model training, the dataset consisting of 120 samples was randomly split into a training set consisting of 90 samples and a test set consisting of 30 samples to evaluate the generalizability of the model. The number of principal components was determined by performing 3-fold cross-validation within the training set, selecting the value that maximized the average of the coefficient of determination (R^2) score on the validation folds. The performance of the PLS model was evaluated using R^2 and the Root Mean Square Error (RMSE) on the test set. The regression coefficients across the wavelength were analyzed to identify the specific wavelength regions that contributed most significantly to the perception of facial impressions. This analytical approach allowed us to identify the spectral characteristics that influence perceived facial impressions. PLS analysis was conducted using Python 3.12.



(**a**) (**b**)

Figure 3. Spectral data for PLS regression analysis. (a) Analysis region on the left cheek; (b) Spectral radiance of the left cheek from the 120 average face images.

2.6. Simulation of enhancing "Bright" impression

We modulated the spectral radiance by using the regression coefficients for "Bright" impression obtained from the PLS regression analysis. The modulated spectral radiance ($L_{mod,\lambda}$) was calculated for each wavelength (λ) using the following equation:

$$L_{mod,\lambda} = L_{\lambda} \times (1 + \beta_{\lambda} \times \alpha)$$

where L_{λ} denotes the original spectral radiance of the average face, β_{λ} represents the regression coefficient for wavelength λ obtained through PLS regression analysis for "Bright", and α is a scaling factor. The modulated spectral radiance was then converted into an RGB image.

3. Results

3.1. The average face from HSI data

We created the average face image from the HSI data of all 10 participants. Furthermore, 120 distinct average face images were constructed by combining subsets of three individuals from 10 participants. **Figure 4** shows the average face image constructed from all 10 individual participants.

Figure 4. The average face image of 10 individual participants.

3.2. The score of facial impression in subjective evaluation.

Three evaluators assigned scores based on six impression terms, which were determined with reference to a previous study [8]. As an example, **Figure 5** illustrates the distribution of scores for the impression of “Elegant”. The results are presented as the mean of the evaluations provided by the three evaluators.

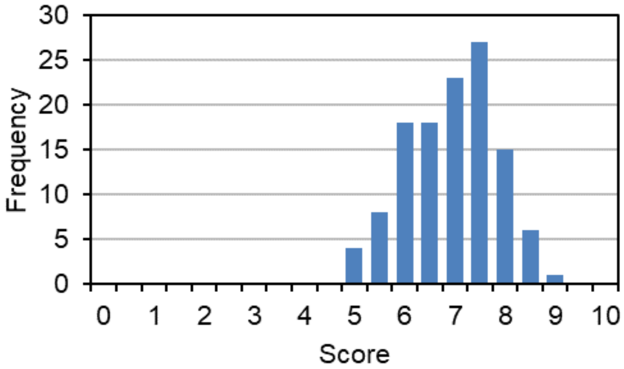


Figure 5. Mean score for the impression of “Elegant”.

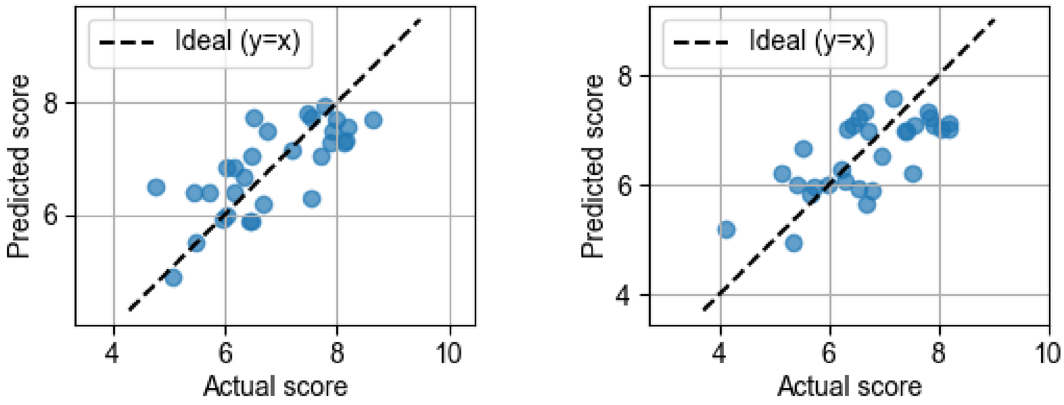
3.3. Establishment of a model for estimating facial impression

PLS regression analysis was conducted to predict scores of six impression terms based on spectral radiance data. The predictive performance of each model was evaluated using R^2 and RMSE. These metrics and the optimal number of components of each model are shown in **Table 1.**

Table 1. Summary of PLS regression models for predicting facial impression terms from spectral radiance data.

Impression Terms	Number of Components	R^2	RMSE
Elegant	5	0.3640	0.7979
Clear	5	0.4712	0.7303
Bright	6	0.5395	0.6960
Warm	2	-0.0362	0.6375
Lively	6	0.2072	0.7143
Healthy	8	0.1376	0.8027

Among the six impressions, the model for “Bright” exhibited a relatively high predictive accuracy, with an R^2 value of 0.5395. The model for “Clear” showed moderate predictive performance, with R^2 values of 0.4712. In contrast, the predictive models for other impressions, especially “Warm,” “Lively,” and “Healthy”, exhibited low predictive performance. **Figure 6** illustrates the relationships between the actual and predicted scores for “Bright” and “Clear”. The scatter plots indicate a good model fit for “Bright” and “Clear” with data points distributed closely around the ideal line. For the “Bright” impression, which exhibited the best predictive accuracy, the 120 facial images were mapped based on the actual evaluation scores (**Figure 7**).



(**a**) Bright (**b**) Clear

Figure 6. Results of prediction in impression scores for the test set data using PLS regression. (a) Bright; (b) Clear

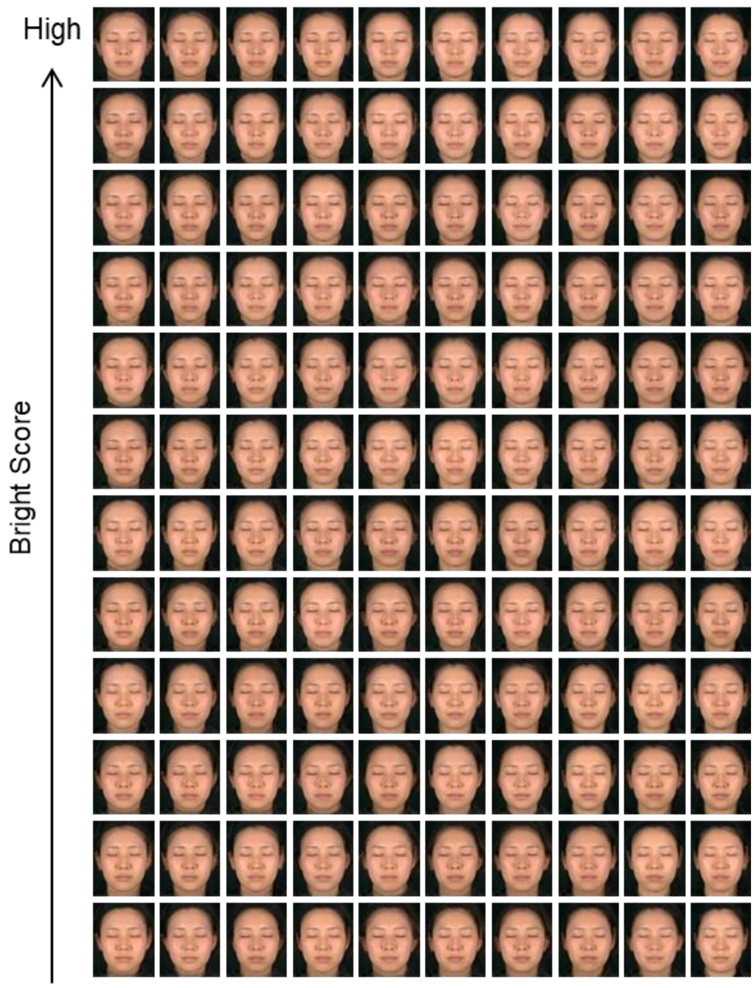
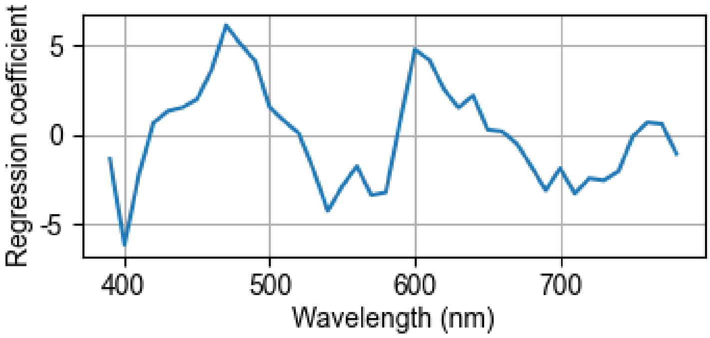


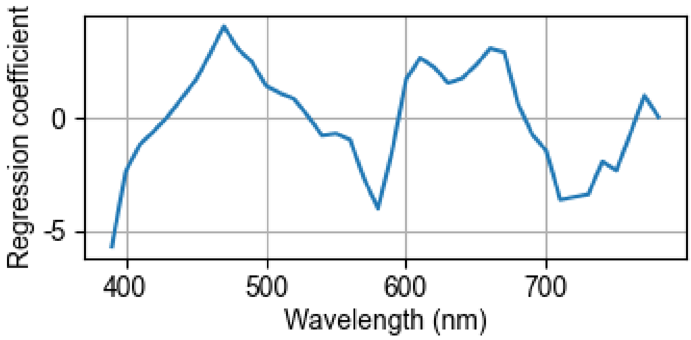
Figure 7. Mapping 120 average face images based on actual evaluation scores for "Bright" impression.

3.4. Weights of the wavelength in the perception of the impression

The results of the PLS regression analysis revealed that, among the six impression terms, “Bright” and “Clear” could be predicted from the spectral radiance data with relatively high accuracy. Using PLS regression models, regression coefficients were calculated to examine the spectral regions contributing to these perceptions. The results are shown in **Figure 8**. For the perception of “Bright”, the wavelength regions around 470 and 600 nm within the visible light range contributed in the positive direction, while the regions around 400 and 540 nm contributed in the negative direction. Similarly, for the perception of “Clear”, the wavelength regions around 470 nm and between 600–670 nm contributed positively, whereas the regions around 390 nm and 580 nm contributed negatively.



(**a**) Bright



(**b**) Clear

Figure 8. Regression coefficients of the predictive models for “Bright” and “Clear” impressions. (a) Bright; (b) Clear

3.5. Simulation of enhancing “Bright” impression

Based on the results obtained from the PLS regression analysis, the average face image enhanced “Bright” impression was generated by modulating the spectral radiance by regression coefficients. **Figure 9** shows the original image of the average face and the simulated image.

(**a**) (**b**)

Figure 9. An average face image is simulated to enhance the impression of “Bright”. (a) The original image; (b) The simulated image

4. Discussion

We developed a novel technique for creating an average face based on HSI data, enabling us to generate facial images with detailed spectral information. By utilizing 76 morphological landmarks for affine transformation and averaging spectral data across multiple participants, we were able to generate standardized facial stimuli while preserving detailed spectral information. This technique eliminates individual differences in facial shape, allowing us to focus specifically on the contribution of spectral characteristics to facial impressions. The average faces contain comprehensive spectral information at 10nm intervals, providing much richer color data than conventional RGB images. As a result of these advantages, it was possible to gain a deeper understanding of the relationship between spectral characteristics and facial impressions.

Utilizing the developed average face generation technique, we conducted impression evaluations. Based on a previous study, six impression terms (“Bright”, “Clear”, “Elegant”, “Warm”, “Lively”, and “Healthy”) were selected [8]. Among these six impressions, our PLS regression analysis revealed notable differences in predictability from spectral radiance (**Table 1**). The impressions of “Bright” and “Clear” yielded models with high predictive accuracy, suggesting that these perceptions are strongly influenced by the spectrum of specific wavelengths rather than simply overall reflectance magnitude (**Figure 6**).

The PLS regression coefficients for “Bright” impression revealed specific wavelength regions that contribute significantly to that perception (**Figure 8a**). We observed a prominent positive peak around 470nm, which is contained within the blue light region of the visible spectrum. This suggests that reflectance in this wavelength region positively contributes to the perception of brightness in the face. Another notable positive peak was identified around 600nm, in the orange light region. This suggests that healthy blood circulation and good complexion contribute positively to perceived brightness. This appears to support the findings of a previous study reporting that reddish skin appeared brighter than yellowish skin did when both had the same colorimetric lightness [3]. Interestingly, we also observed negative peaks at approximately 400nm and 540nm. These wavelength regions coincide with hemoglobin absorption bands. Lower reflectance in these regions indicates higher hemoglobin content, which may translate to a healthier complexion that enhances the perception of brightness. Furthermore, about the impression of “Clear”, specific wavelength regions that significantly contribute to its perception were also identified (**Figure 8b**). A prominent positive peak around 470 nm suggests that this region contributes not only to the perception of “Bright” but also to

"Clear". Another notable positive peak was broadly observed in the red region, between 600 and 670 nm. The finding that a combination of wavelengths in the blue and red regions positively contributes to the impression of "Clear" overlaps with insights from previous studies [8]. In addition, a negative peak was found at 580 nm. This suggests that increased reflectance in the yellow-green region may result in a duller appearance, thereby diminishing the impression of "Clear". These results collectively suggest that the perception of the impressions "Bright" and "Clear" is not determined by the overall level of reflectance, but rather by a specific balance of spectral characteristics.

In contrast, the impressions of "Warm", "Lively", and "Healthy" showed lower predictive accuracy in our PLS models (**Table 1**). This does not necessarily indicate that these impressions are unrelated to color information. Several factors may explain these findings. First, the average faces used in our study may have exhibited limited variation in these particular attributes, making it difficult to establish clear correlations. Second, impressions like "Warm" and "Healthy" might be influenced by melanin, which has broad absorption characteristics without distinctive spectral peaks, making it difficult to identify specific wavelength contributions through PLS analysis.

The differential predictability across impression terms raises important questions about the relative contributions of spectrum information versus facial morphology and expressions. While our approach controlled for differences of facial shape, we cannot rule out the possibility that impressions with lower spectral predictability may be more strongly influenced by variations in facial appearance or expressions than by spectrum characteristics.

Because our study involved female participants aged between 23 and 30 years, there may be limitations in generalizing the findings. It is also important to note that our findings are based on measurements under simulated sunlight conditions. The relationship between spectral characteristics and facial impressions may vary under different lighting conditions, as the spectral distribution of the light source is considered to significantly affect color perception. It would be interesting to investigate how these relationships change across other lighting environments in future research.

The innovative technique developed in this study for creating an average face with detailed spectral information has applications beyond understanding facial impressions. We consider that this technique has the potential to bring benefits to a variety of fields, including cosmetic science and dermatology. For instance, this approach could be valuable for evaluating the effects of skincare products on skin appearance. This may facilitate the development of cosmetic products that are more precisely tailored to diverse consumer needs. Such products are expected to contribute to enhancing consumer satisfaction.

5. Conclusion

In this study, we developed a new method for the evaluation of facial color by using comprehensive spectrum information and removing the influence of individual differences in the shape of the face. This approach provides a valuable tool for investigating the complex relationship between spectral properties and facial impressions. This technique demonstrated that the spectral radiance of specific wavelengths has a considerable influence on facial impressions of "Bright" and "Clear". This methodology is considered to be applicable for evaluating alterations in dermal characteristics induced by cosmetic products, thereby contributing to the development of products that achieve higher consumer satisfaction by meeting a wide range of consumer objectives.

References

[1] J Willis, A Todorov (2006) First impressions: making up your mind after a 100-ms exposure to a face. *Psychol Sci.*, 17(7): 592–598

[2] Igarashi T (2018) Resent research on color and chromophores of the skin: a short survey. *J. Color Sci. Assoc. Jpn.*, 42(2): 65-73.

[3] Yoshikawa H et al. (2012), Effect of Chromatic Components on Facial Skin Whiteness. *Color Research and Application*, 37(4): 281-291.

[4] Tsumura N et al. (2003). Image-based skin colour and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Transactions on Graphics*, 22(3): 770-779.

[5] KONICA MINOLTA SENSING INC. (2007) Precise color communication.

[6] F Galton (1878) Composite Portraits. *J Anthropolog Inst G B and Irel* 8: 132-142.

[7] Q Marin, et al. (2023) Comparative analysis of original face and skin-warped average face images for the scoring of skin attributes. *Skin Res Technol*. 29(4): e13324.

[8] Kakimoto R et al. (2022). Discovery of global common denominators of skin reflectance that enhance attractive impressions -Paradigm shift in benefits of foundations to bring out individual beauty-. *Proceedings of the IFSCC 2022 Congress.*