

顔のクラスタリングと印象評価結果の可視化

孫 芸珂[†] 伊藤 貴之[‡]

[†]お茶の水女子大学理学部情報科学科 〒112-8610 東京都文京区大塚 2-1-1

E-mail: [†]{sun.yike, itot}@is.ocha.ac.jp

あらまし 顔画像の印象は顔の肌の色調や形状などの特徴に深く関係している。本研究では肌の色調や顔の形状に注目して、顔画像の肌領域を抽出した画像について、閲覧者がもつ印象との関係を可視化する。具体的には、色調や形状などの特徴量に基づく肌領域のクラスタリング結果と、顔画像の印象評価結果の相関を可視化している。我々の研究では、機械学習モデルで生成した架空の人物の顔画像 90 枚から、RGB 色空間での領域分割により肌領域だけを抽出した画像を生成する。顔画像の特徴量には肌の色調や質感、顔の形状などに対応したものを算出可能であるが、本報告では現段階で肌の色調と顔の形状を特徴量として実装している。具体的には、「肌領域の色分布」「顔の縦横の長さ」の各々を入力として、K-means 法によって画像をクラスタリングする。さらに、肌画像に対して SD 法による 5 段階評価を実施している。これらの結果を統合して、顔画像のクラスタリング結果と印象評価結果の相関を、我々が以前に発表した可視化ツール「平安京ビュー」を用いて可視化する。

キーワード 顔画像、クラスタリング、印象

1. はじめに

SNS の普及とともに、自分の顔画像をインターネットにアップロードする機会が増えている。我々は、顔画像の特徴量とその印象にはどのような相関があるかを調査する研究に着手している。具体的には、大量の顔画像群から肌領域を抽出して特徴量を算出するソフトウェアを開発すると同時に、顔画像に関する被験者回答形式の印象評価を実施し、顔画像の特徴量と印象評価結果の相関を分析している。

顔自体の特徴を比較するためには、同じ環境で撮影した大規模な顔画像データセットが必要であるが、これを入手することは簡単ではない。そこで本研究では、AI（機械学習による生成モデル）が生成する架空の人物の顔画像を多数用意する。また、画像に写っている服と髪形の要因を排除するために、RGB 色空間で顔画像の肌領域を抽出し、肌以外の領域を黒く塗りつぶした画像を使用する。一方で、大規模な顔画像から特徴量と印象評価の関係を探す必要があるが、画像が多くなると特徴が近い顔画像を探すことが困難である。そこで、我々は K-means 法を用いて特徴量によりクラスタリングすることで顔画像を分類する。クラスタリングをした結果と印象評価の結果の関係を可視化する閲覧ソフトウェアを開発している。

本報告の構成は以下の通りである。2 章では顔画像の印象と特徴量の関係、顔画像の特徴量と印象評価を分析ツールに関する関連研究について、3 章は提案手法について述べる。4 章では可視化実行例と考察について、そして 5 章ではまとめと今後の課題について述

べる。

2. 関連研究

2.1 顔画像の印象と特徴量の関係

本章では顔画像の印象に影響する要因について報告された内容を簡単に紹介する。

鄧ら[1]は日本人と中国人の好みの肌色の違いについて調査した。この調査では、49 種類の LAB 値を適用することで、同一人物の顔画像の肌領域の色を 49 種類に変化させて、49 枚の画像を生成した。被験者の中国人 60 人と日本人 40 人を対象として生成した各 49 枚の画像の中から、最も好みの肌色の顔画像を選び、それに対して 19 項目について印象評価を実施した。その結果、日本人は明度が低く彩度が高い肌色を好む傾向があった。このような肌色に対して因子分析を適用した結果、「積極的な」「活発的な」「陽気な」「健康的な」の 4 つの形容詞が第一因子となった。これに対して、中国人は明度が高く彩度が低い肌色を好む傾向があった。このような肌色に対して因子分析を適用した結果、「知的な」「真面目な」「感じの良い」「上品な」の 4 つの形容詞が第一因子となった。この研究からも、顔画像の印象評価は肌の色に深く関係していることが示唆される。また鈴木[2]は、400 人の女性(20-50 代)の顔写真について、印象を形容する 34 個の印象評定語と、顔の形態を形容する 40 個の形態印象評定語を選定し、上述の写真についてそれぞれの印象を調査した。その結果として、印象因子と形態印象因子の間に有意な相関があることを示した。この実験では顔画像からも信頼性の高い結果を得ることが確認されていること

から、本研究ではこの研究の評価項目を参考にして、印象評価の項目を選択した。

2.2 顔画像の特徴量と印象評価の可視化

顔画像に対しての印象評価は複合的な分析手法が必要なので、一目で分析結果を理解できるような可視化手法の適用が有用である。また、特徴量と印象評価の関係を探すには多くの顔画像が必要であるが、画像が多くなればなるほど共通した特徴を有する顔画像をくまなく探すことが困難になる。そこで、本研究ではクラスタリングによって画像を分類し、可視化手法を用いて各クラスタを構成する顔画像を目視確認する。本研究では「平安京ビュー」[3]を可視化手法として適用した。「平安京ビュー」は、階層型データの葉ノードを長方形のアイコンで、枝ノードを長方形の枠で表現し、階層構造を2次元の長方形群の入れ子構造で表現し、これらをできるだけ小さい画面空間に配置することで、階層型データ全体を一画面にする。

3. 提案手法

本章では提案手法の処理手順について論じる。まず3.1節にてデータセットを定義し、続いて3.2節で肌領域の抽出、3.3節で印象評価の実施手順、3.4節で顔画像の特徴量によるクラスタリング、そして3.5節でクラスタリング結果の可視化について説明する。

3.1 データセット

本研究のような顔画像の印象評価には、良好かつ統一的な照明条件で撮影した多数の顔画像が必要であるが、このようなデータセットの入手は簡単ではない。実在する人物の顔画像には肖像権を伴うことから、オープンな大規模な顔画像データセット自体があまり多くない。実在するデータセットにしても、雑多な照明条件で撮影された画像が収録されていることが多い。そこで代替手段として本研究では、GANという機械学習技術を用いて生成された架空の顔画像[4]を使用している。このような手段で用意した顔画像は実在する人物ではないことから、肖像権などの問題は発生しない。また撮影条件や人物属性を統一した形で顔画像を生成できることから、本研究のような印象評価にも向いている。

本研究が採用した顔画像生成システムでは、性別、人種、表情などをオプションとして設定できる。現段階の本研究では、性別を女性、人種をアジア人、エモーショナルをニュートラルと設定して生成した顔画像90枚を使用している。また、図1(1)のような斜めから撮影した画像や顔が部分的にはみ出している画像の使用を避けて、原則として図1(2)のような正面画像かつ顔全体が写っている画像を選んでいる。

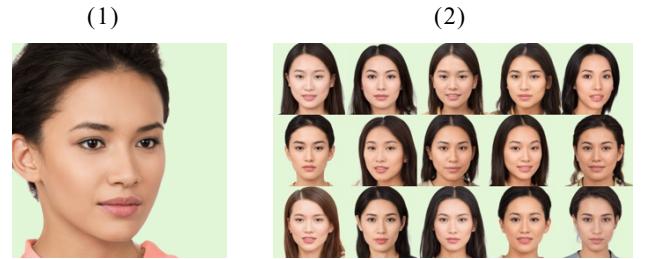


図 1 (1) 斜めを向いた画像. (2) 正面画像.

3.2 肌領域の抽出

本手法ではまずGANで生成した顔画像群の各々から肌領域を抽出する。本研究ではShinら[5]による手法を採用している。Shinらは9種類の色空間(RGB, YCbCr, HISなど)を用いた肌検出の性能を比較し、RGB色空間での肌検出の性能が一番よいという実験結果を示している。我々もYCbCr色空間とRGB色空間で肌領域を抽出して比較した。その結果図2に示すように、RGB色空間で抽出した肌領域のほうが欠損を生じにくいことを確認した。以上により本研究では、RGB色空間で肌領域を抽出することにした。また、背景に残っている髪の毛や服はPhotoshopで黒く塗りつぶした。

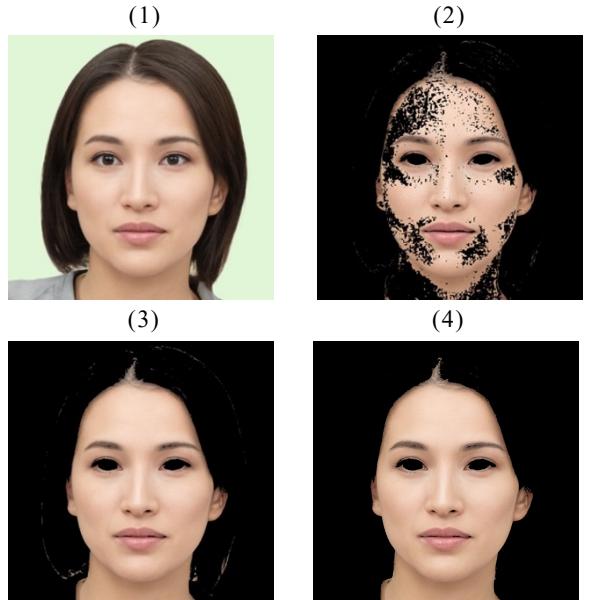
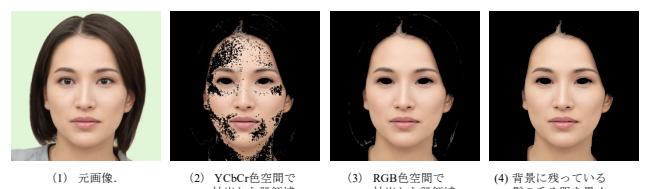


図 2 (1)元画像.(2)YCbCr色空間で抽出した肌領域. (3)RGB色空間で抽出した肌領域. (4)背景に残っている髪の毛や服を黒く塗りつぶした画像.



(1) 元画像. (2) YCbCr色空間で抽出した肌領域. (3) RGB色空間で抽出した肌領域. (4) 背景に残っている髪の毛や服を黒く塗りつぶした画像.

3.3 印象評価

本研究では印象評価という手段で肌の印象を数値化する。具体的には、SD (semantic differential)法[6]にもとづいて何種類かの形容詞を用意し、その各々に対する適合度を閲覧者に5段階評価してもらい、これをリッカート尺度として扱う。

このような被験者実験では、被験者に過剰な負荷を与えないことが重要である。具体的には、

[要件 1] 被験者に回答させる評価項目の総数が多くなりすぎない。

[要件 2] 被験者に閲覧させる画像の枚数が多くなりすぎない。

という点に注意する必要がある。

要件 1のために我々は、肌の専門家と議論して肌画像の評価項目を選定した。専門家の知見に沿って肌の「特徴系」と「雰囲気系」の2種類から評価項目を検討した結果、「明るさ」「赤み」「黄色味」「白さ」「若さ」「健康さ」「親しみやすさ」の合計7項目に絞った。そ

の内で「明るさ」といった肌の特徴系の評価項目は反対語をもたない。このような評価項目では反対語ではなく、「明るさを感じる～明るさを感じない」という形式で評価項目を設定した。表 1に評価項目をまとめる。

要件 2を満たすために我々は、各被験者に提示する画像の枚数を制限することにした。このとき、同一被験者に提示される画像が類似したものばかりであるよりは、バラエティに富んでいたほうが回答しやすいと考えられる。そこで本手法では、顔画像の色を配慮して画像を9つのグループに分割し、各グループに属する画像を1枚ずつ評価させることにした。被験者は20代女性20名で、1人あたり9枚の画像を提示して評価してもらった。

ここで我々は、図 3に示すような評価用画面を開発した。この画面上で、各グループからランダムに選択された顔画像を1枚ずつ提示し、提示された顔画像に対してその印象を5段階のリッカート尺度で評価させた。

表 1 印象評価項目.

特徴系 :
明るさを感じる ⇔ 感じない
肌の赤みを感じる ⇔ 赤みを感じない
肌の黄色味を感じる ⇔ 黄色味を感じない
肌の白さを感じる ⇔ 白さを感じない
雰囲気系 :
若々しい ⇔ 老けた
健康的な ⇔ 不健康的な
親しみやすい ⇔ 親しみにくい



5/9

肌の明るさを感じる ○1 ○2 ○3 ○4 ○5 明るさを感じない
肌の赤みを感じる ○1 ○2 ○3 ○4 ○5 赤みを感じない
肌の黄色味を感じる ○1 ○2 ○3 ○4 ○5 黄色味を感じない
肌の白さを感じ ○1 ○2 ○3 ○4 ○5 白さを感じない
若々しい ○1 ○2 ○3 ○4 ○5 老けた
健康的な ○1 ○2 ○3 ○4 ○5 不健康的な
親しみやすい ○1 ○2 ○3 ○4 ○5 親しみにくい

next

図 3 印象評価用画面.

3.4 顔画像の特徴量によるクラスタリング

顔画像の特徴量として、肌の色調や質感、顔の形状などに対応する値を算出することが可能である。現段階の実装では、「肌領域の色分布」「顔の縦横の長さ」の2種類に関する特徴量を算出している。

本研究では、3.2節で示した処理により肌領域だけを抽出した全ての肌画像に対して、肌領域の色分布と顔の縦横の長さを特徴量として、顔画像群にクラスタリングを適用する。

3.4.1. 肌領域の色分布によるクラスタリング

以下の手順で画像の色分布を示すヒストグラムを生成し、これを多次元ベクトルとみなして次元削減を適用する。

- (1) RGB 色空間を $32 \times 32 \times 32 = 32768$ 個の立方体領域に分割し、各領域に該当する画素数を集計し、この結果を 32768 次元ベクトルとする。
- (2) 全ての画像において 0 である次元を削除し、それ以外の次元だけを残したベクトルを生成する。我々の実験ではこの処理により次元数が 500 程度に削減されている。
- (3) (2)のベクトルに PCA (主成分分析) を適用し、その主成分を選ぶ。我々の実験では第 1～第 5 主成分を選んで 5 次元ベクトルを生成している。
続いて、(3)で得た 5 次元ベクトルを各画像の特徴量として、K-means 法によるクラスタリングを実行する。K-means 法は重心をクラスタの代表点として、以下の値

$$\sum_{i=1}^k \sum_{x \in C_i} (d(\mathbf{x}, \mathbf{c}_i))^2$$

の評価関数を最小化するように K 個のクラスタを生成する。

3.4.1.1. 最適なクラスタ数の確定

本研究は最適なクラスタ数を確定するために、3.2 節の処理手順(3)で得た 5 次元ベクトルのうちの第 1 ~ 第 3 主成分に対して、エルボー法とシルエット分析[7]を適用した。エルボー法は SSE(クラスタ内誤差平方和)から判明する歪みが最大となる k の値を最適なクラスタ数として導き出す方法である。シルエット分析は、クラスタ内の凝集度とクラスタ間の乖離度を計算することで、クラスタリング結果の品質を評価する手法の一つである。シルエット分析から得たシルエット係数は範囲 [-1, 1] に属し、値が 1 に近いほど性能が高い。我々のデータにおける評価結果の例を図 4 に示す。この例では、歪みが最大となる k の値は $k=4$ である。さらに、 $k=4$ の時にシルエット係数も比較的大きいことがわかる。以上により、 $k=4$ が最適なクラスタ数と判断できる。

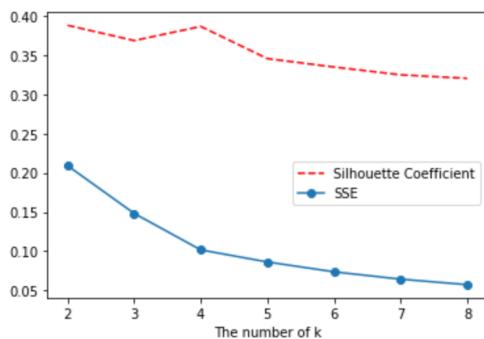


図 4 クラスタ数によるシルエット係数と SSE 値の変化。

3.4.2. 顔の縦横の長さによるクラスタリング

以下の手順で顔の縦横の長さを算出する。

- (1) dlib というオープンソースの機械学習ライブラリを使用し、図 5 のように顔の輪郭を表す 81 個の特徴点を検出する。
- (2) 特徴点の座標値の最小値・最大値から、顔の縦横の長さを計算する。
- (3) 縦横の長さを 2 次元の特徴ベクトルにする。

続いて、(3)で得た 2 次元ベクトルを各画像の特徴量として、K-means 法によるクラスタリングを実行する。



図 5 顔の輪郭を表す 81 個の特徴点。

3.5 クラスタリング結果の可視化

3.4.1 節および 3.4.2 節に示した手順で得たデータを「平安京ビュー」を用いて可視化する。肌領域の色分布を特徴量として、90 枚の顔画像を 4 つのクラスタに分類した可視化画面を図 6 に示す。顔の縦横の長さを特徴量として、90 枚の顔画像を 7 つのクラスタに分類した可視化画面を図 7 に示す。

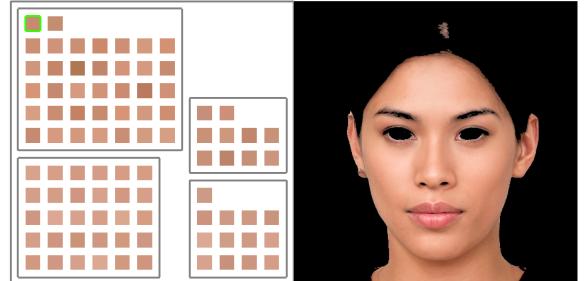


図 6 肌領域の色分布にもとづくクラスタリング結果の可視化画面。左側の枠がクラスタに、小正方形が画像に対応する。小正方形を選択すると、それに対応する顔画像が右側に表示される。

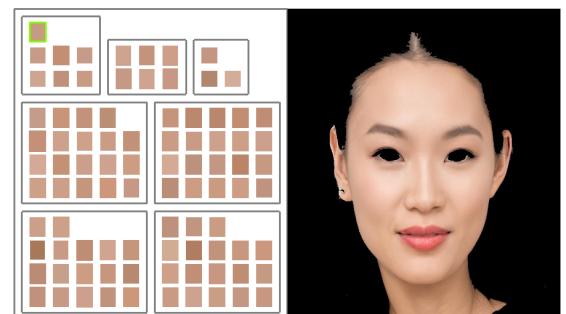


図 7 顔の縦横の長さにもとづくクラスタリング結果の可視化画面。左のアイコンの形状の違いを強調するため、顔の縦横比が 1.2 より小さかったらアイコンを正方形にしている。さもなければ、アイコンの縦横比を顔の縦横比に等しくする。

4. 可視化実行例

本章ではクラスタリングと印象評価の結果の可視化の実行例を示す。

本報告で示す可視化手法では、クラスタを構成する各顔画像の印象評価の平均値を枠の色で表現している。印象評価の平均値と色の関係を表 2 に示す・また、枠に色をつけた例を図 8(1)と図 8(2)に示す。

図 8(1)に示した例では、3.3 節で示した評価項目のうち「明るさ」に関する印象評価結果を反映した結果である。この可視化結果から、左上のクラスタには「明るさを感じない」と評価された顔画像が多いことがわかる。言い換えれば、このクラスタリング結果と「明るさ」に関する印象評価には一定の相関があることが示唆される。

続いて、「明るさ」に関する印象評価結果に対して、最も肌の明るさを感じない平均値を有するクラスタ1と、最も肌の明るさを感じる平均値を有するクラスタ4を対象として、図9(1)(2)に示す各5枚の画像を無作為抽出した。クラスタ1には肌の黄色味を感じられ、健康的と思われる顔画像が集まっている。これに対してクラスタ4では、肌の明るさや赤みを感じられる顔画像が集まっていることが確認できる。さらに、10枚の顔画像の輝度値のヒストグラムを比較した。結果はクラスタ1の画像の輝度値はクラスタ4より低いところに集中していることがわかった。従って、健康的に見える顔画像の特徴としては「肌の黄色味を感じる」「画像の輝度が低い」という傾向があることが示唆される。

図8(2)の例は、3.3節で示した評価項目のうち「若さ」に関する印象評価結果を反映した結果である。この可視化結果から、「若さ」に関する印象評価結果に対して、7つのクラスタの中で得点が一番低い、すなわち平均的に一番「若々しさを感じる」と評価されたクラスタ3には、正方形に近い形状の顔が多いことを認める。これが正しいことを確認するために、クラスタごとの顔の縦横の比率を比較した。図10で示したように、クラスタ3は平均的に顔の縦横比の比率が低いことがわかった。「若々しく見える顔」は比較的に顔の形状は丸く、小さい傾向があることが示唆される。

表2 評価結果と色の関係。

色	評価結果xの範囲
■	$x \leq 2$
■	$2 < x \leq 3$
■	$3 < x \leq 4$
■	$4 < x \leq 5$

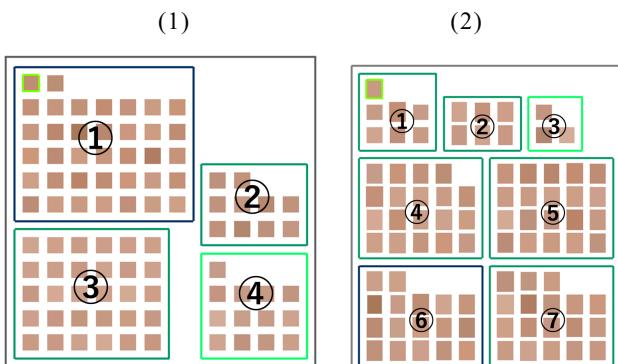


図8 (1)肌領域の色分布によるクラスタリングと印象評価結果の可視化画面。(2)顔の縦横比によるクラスタリングと印象評価結果の可視化画面。

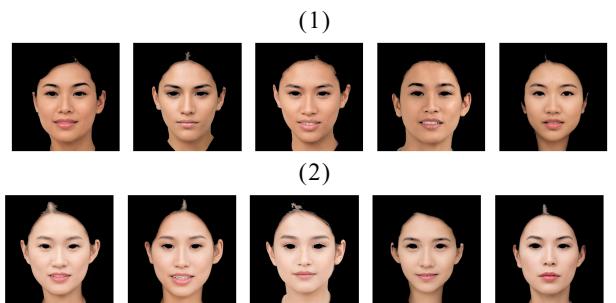


図9 (1)クラスタ1に属する画像5枚。(2)クラスタ4に属する画像5枚。

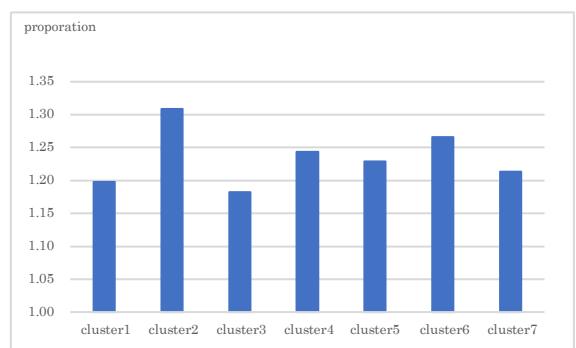


図10 クラスタ毎の顔の縦横の比率。

5.まとめと今後の課題

本報告では肌画像のクラスタリング結果と印象評価結果の相関を導くための一手法を提案し、現在開発しているソフトウェアについて詳細を述べた。また、可視化実行例から特定の印象評価項目と特定の顔画像の特徴量の相関を観察し、分析した。

現時点での実装では、「肌領域の色分布」「顔の縦横の長さ」の2種類だけが特徴量として適用されている。そこで今後は他の特徴量、例えば肌の質感、顔の部位別の色分布、顔の部位別の形状、といった特徴と印象評価との関係を可視化したい。さらに我々は、印象評価の項目間の相関についても調査中である。例えば図11に示すように、「明るさ」に対しての印象評価は「白さ」に対して正の相関を有するのに対して、「黄色味」に対して負の相関を有している。このような印象評価の項目間の関係をあわせて可視化するという点にも、今後取り組んでいきたい。

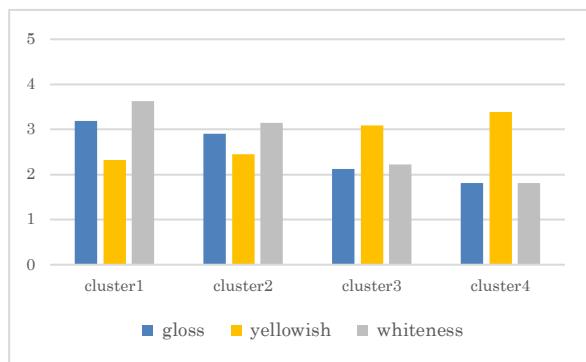


図 11 肌領域の色分布によるクラスタリングと印象評価の肌の「明るさ」、「黄色味」と「白さ」に対しての評価.

謝辞

印象評価項目の選定について多くのアドバイスをいただいた、資生堂グローバルイノベーションセンターの豊田成人様に感謝いたします。

参考文献

- [1] 鄧珮, 青木直和, 小林裕幸, “種々肌色から受ける印象の日・中比較とその決定要因”, 日本写真学会誌, Vol.76, No. 1, pp. 63-69, 2013.
- [2] 鈴木ゆりか, “顔の形態と印象の関係,化粧心理学: 化粧と心のサイエンス”, 資生堂ビューティーサイエンス研究所(編), フレグラッシュジャーナル社, pp. 124-133, 1993.
- [3] 伊藤貴之, 山口裕美, 小山田耕二, “長方形の入れ子構造による階層型データ視覚化手法の計算時間および画面占有面積の改善”, 可視化情報学会論文集, Vol. 26, No. 6, pp. 51-61, 2006.
- [4] <https://generated.photos/>
- [5] M. C. Shin, K. I. Chang, L. V. Tsap, “Does Colorspace Transformation Make Any Difference on Skin Detection? ”, IEEE Workshop on Applications of Computer Vision, pp. 275-279, 2002.
- [6] C. E. Osgood, G. J. Sugi, P. H. Tannenbaum, “The measurement of meaning, University of Illinois Press”, Urbana, 1957.
- [7] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis”, Journal of Computational and Applied Mathematics, 20:53–65, 1987.