

CNN特徴量を用いたスーツスタイルにおける印象推定モデルの構築

新海 公章[†] 谷 伊織^{††} 飛谷 謙介^{†††} 都賀 美有紀[†] 長田 典子[†]

[†] 関西学院大学大学院理工学研究科 〒669-1337 兵庫県三田市学園 2-1

^{††} 神戸大学情報基盤センター 〒657-0013 兵庫県神戸市灘区六甲台町 1-1

^{†††} 長崎県立大学情報システム学部 〒851-2195 長崎県西彼杵郡長与町まなび野 1-1-1

E-mail: [†]{k.shinkai,toga.m,nagata}@kwansei.ac.jp, ^{††}Iori_tani@penguin.kobe-u.ac.jp, ^{†††}tobitani@sun.ac.jp

あらまし オーダースーツのビスポーク（推薦）サービスを目的として、スーツのスタイル（質感と形状）から受ける印象の推定手法を提案する。3398枚のスーツ画像を収集し、心理実験により付与した25種類の主観印象値と、画像から得られるCNN特徴量（スタイル特徴、コンテンツ特徴）との関係性をRandom Forestによりモデル化する。9分割交差検証より未学習データの主観印象値とCNN特徴量から算出した推定印象値との相関係数を求めた結果、中程度の正の値を示し、提案手法の有効性を確認した。

キーワード ファッション, スーツ, CNN, スタイル特徴, コンテンツ特徴, 印象推定

Impression Estimation Model for Suit Styles Using CNN Features

Kimiaki SHINKAI[†], Iori TANI^{††}, Kensuke TOBITANI^{†††}, Miyuki TOGA[†], and Noriko NAGATA[†]

[†] School of Science and Technology, Kwansei Gakuin University 2-1 Gakuin, Sanda-shi, Hyogo, 669-1337 Japan

^{††} Information Infrastructure Center, Kobe University 1-1 Rokkodai-cho, Nada-ku, Kobe-shi, Hyogo, 657-0013 Japan

^{†††} Faculty of Information Systems, University of Nagasaki 1-1-1 Manabino, Nagayo-cho, Nishi-Sonogi-gun, Nagasaki, 851-2195 Japan

E-mail: [†]{k.shinkai,toga.m,nagata}@kwansei.ac.jp, ^{††}Iori_tani@penguin.kobe-u.ac.jp, ^{†††}tobitani@sun.ac.jp

Abstract In this study, propose a method for estimating the impressions evoked by the style (texture and shape) of a suit for the purpose of a bespoke suit (recommendation) service. We collected 3398 images of suits and used Random Forest to model the relationship between the subjective impressions of 25 words assigned by psychological experiments and CNN features (style features and content features) obtained from the images. The correlation coefficients between the subjective impression scores of the untrained data and the estimated impression scores calculated from the CNN features are obtained from the 9-fold cross-validation, and the results show moderately positive values, confirming the effectiveness of the proposed method.

Key words fashion, suit, CNN, style feature, content feature, impression estimation

1. はじめに

インターネットの発展による第4次産業革命の進行は、顧客の好みや用途に合わせたカスタマイズ需要に柔軟に 대응することを可能にした[1]。特にファッション分野では、人によってサイズや用途が異なるため、カスタマイズ製品は需要が高い。一方で、カスタマイズ製品は顧客の選択肢を増大させるため、購買行動において多くの決定を迫ることになり、多大な労力と時間を必要とさせる。そこで、意思決定を補助するためにユーザが興味を持つと思われる情報を提示するレコメンドシステムが注目されている。レコメンドシステムには購買履歴に基づく協調フィルタリングや製品の特徴の類似度に基づくコンテンツベ

スフィルタリングがある。しかし、協調フィルタリングには購買履歴がなければレコメンドできないという問題があり、コンテンツベースフィルタリングには似たような製品しかレコメンドされないという問題がある。そのため、個人の好みやニーズを推定する技術が望まれている。

個人の好みやニーズによってカスタマイズする製品の一例として、スーツの伝統的なオーダー方式である「ビスポーク」が挙げられる[2]。ビスポークとは「テーラーや販売員などが対話をしながら服を仕立てることが転じて、顧客の好みや要求に合わせて製造すること」を意味する。ビスポークでは専門知識を持った熟練度の高いテーラーや販売員が必要とされる。しかし、それ故に人材不足が懸念される。製造業における熟練技術

であるビスポークをデジタル化するシステムは、熟練度の低いビスポークのテーラーや販売員を補助することができるため求められている。

一方で、対象物に対して人が感じる感性については、最近、好みや美的良さや感情などの審美性 (aesthetic scores, aesthetic values) を CNN でモデル化する研究が盛んに行われている [3], [4]. しかし心理学, デザイン学, 感性工学などの多くの分野では、対象に対する評価者の反応と、対象そのものが備える性質とを明確に区別することが前提となっている [5]. 例えば「美しい」や「好き」といった審美性評価は、「派手な」や「スリムな」といった対象の具体的な性質に関する情報の処理を経て、より高次の感情的な反応に媒介されて形成される評価であると説明されている。こうした対象の性質は美的概念 (aesthetic concepts) と呼ばれる [6]~[8]. 審美性の予測精度を高めるためには、低次の情報である美的概念の把握が不可欠である。CNN を用いて美的概念 (aesthetic concepts) をモデル化した研究はまだ少数である。本研究では、美的概念を印象と定義し、スーツスタイルの視覚的印象と物理的特徴の関係を CNN 特徴量を用いてモデル化し、形状と質感で構成されるスーツスタイルから喚起される印象を自動で推定する手法を提案する。

2. 先行研究

プロダクトが持つ物理的特徴と、そのプロダクトから喚起される印象をモデル化する研究が行われている。これには、物理的特徴として色や光沢感や表面の粗さ、形状などを使用した技術 [9] や、2 次元画像を対象とした特徴量と印象の関係性を用いた検索技術 [10] があげられる。飛谷らは、車、花瓶、椅子を対象に 3 次元物体の印象を推定する手法を提案した [11]. 推定結果と実際に人が付与した評価値との相関係数を算出したところ、正の相関関係を示し、提案手法の実用上の有効性が確認された。しかし、この研究ではプロダクトの 3 次元形状のみを対象とし、プロダクトの質感は対象としていない。

質感を対象とした研究は盛んに行われている [12]~[14]. 寸田らは、衣服の柄画像を表現する物理的特徴としてスタイル特徴を用い、視覚的印象を自動で推定するモデルの構築を行った [15]. その後、構築したモデルに基づきテストデータの画像に対する視覚的印象を推定したところ、人が実際に感じる印象と強い正の相関があることが確認された。柄画像の印象推定においてスタイル特徴が有用であることが示された。

スーツは生地を体の形に作るプロダクトであるため、その印象推定では形状と質感の両方を考慮する必要がある。Gatys らは、画風変換アルゴリズムの提案において、一般物体認識に用いられる VGG-19 [16] から抽出される画像のコンテンツ特徴とスタイル特徴の両方に着目している [17]. コンテンツ特徴は VGG-19 の中間層から出力される特徴マップであり、スタイル特徴は特徴マップを Gram matrix 化したものである。Gatys らは、コンテンツ特徴は一般物体認識に必要な形状の情報を多く保持している特徴量であり、スタイル特徴は画像中の色やパターンの情報を多く保持している特徴量であることを示唆した。これをスーツに当てはめると、コンテンツ特徴はシルエットと関

連が強く、スタイル特徴は生地と関連が強いと考えられる。そこで、本研究ではプロダクトの形状と質感を考慮するためにコンテンツ特徴とスタイル特徴に着目し、視覚的印象と物理的特徴との関係を明らかにする。

3. 提案手法

本研究では、スーツスタイルから喚起される視覚的印象を自動的に推定する手法を提案する。その概要を図 1 に示す。(1) まず、印象評価実験を行い、視覚的印象を定量化する。(2) 次に、VGG-19 を用いて、CNN 特徴量であるコンテンツ特徴とスタイル特徴を物理的特徴として抽出する。(3) その後、視覚的印象と物理的特徴を Random Forest を用いてモデル化し、印象推定モデルを構築する。最後に、構築した印象推定モデルを用いて、テストデータの推定を行い、提案手法の有効性を確認する。

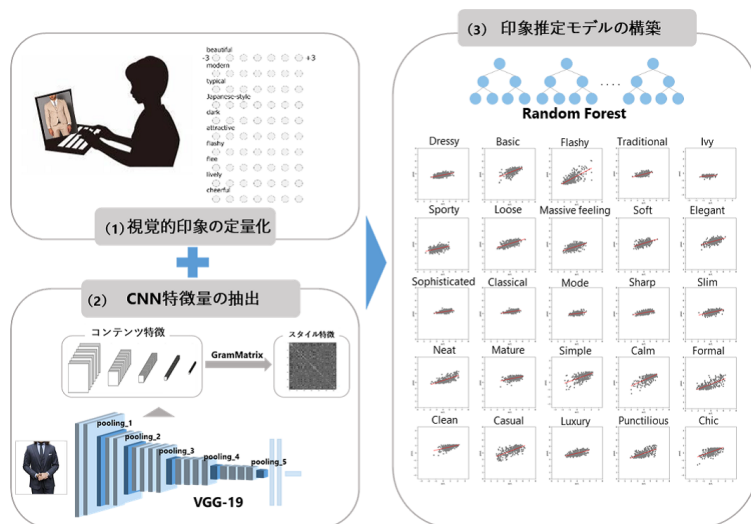


図 1 提案手法の概要

Fig. 1 Overview of our proposed method.

4. 印象の定量化

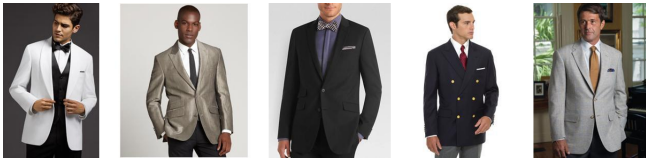
4.1 画像データの作成

はじめに、ファッション画像のデータベースである Deep Fashion [18] から、スーツの画像を 10,448 枚収集した。収集した画像のなかから、本研究で扱うデータを選定した。選定条件は以下の 3 つである。

- (1) 上半身が正面から写っている
- (2) 十分にスーツの形状を確認できる
- (3) 人がスーツを着用している

選定条件をすべて満たす画像は 3,398 枚であった。選定条件を満たす画像と満たさない画像の一部を図 2 に示す。

次に、印象評価実験や深層学習に適したデータセットを作成するため、スーツの領域のみの画像を作成した。Deep Image



(a) 選定条件を満たす画像

(a) Images that meet the selection criteria.



(b) 選定条件を満たさない画像

(b) Images that do not meet the selection criteria.

図2 収集したスーツの画像

Fig. 2 Collected suit images.

Matting [19] を用いて背景を除去し、画像編集ツールを用いて顔の領域を除去した。作成した画像の一部を図3に示す。



図3 スーツ領域のみの画像

Fig. 3 Image of suit region only.

4.2 評価語の収集選定

適切にスーツの印象を評価するには、網羅的かつ代表的な評価語を収集・選定する必要がある。また、スーツスタイルに対する印象は、エキスパートと非エキスパートで、異なる言葉が喚起される可能性がある。そこで、スーツの販売員歴20年のエキスパートと、大学生と大学院生である非エキスパートの両方から評価語を収集した。

まず、エキスパートを対象に、スーツの印象を表現する言葉

を調査し、26個の評価語を収集した。エキスパートから収集した言葉を表1に示す。次に、非エキスパートを対象に、飛谷

表1 エキスパートから収集した26語

Table 1 Twenty-six words collected from experts.

シュッとした	Iラインな	Aラインな	ドレッシーな	ベーシックな
派手な	トラディショナルな	アイビーな	スポーティーな	今風な
こなれた	ゆったりとした	かっぶくある	重厚感ある	やわらかい
上品な	洗練された	クラシカルな	モードな	豪華な
シャープな	スリムな	スラっとした	きちんとした	ユニセックスな
きれいな				

ら[20]を参考とした自由記述実験を行った。実験参加者は大学生及び大学院生10名(男性5名,女性5名,平均年齢22.5歳,SD=0.92)であった。手続きとして、実験参加者には、画面に呈示されたスーツの画像を観察してもらい、感じた印象をそれぞれ自由に書き出してもらった。その結果、304語のスーツスタイルに対する印象を表現する言葉を得た。意味の重複している言葉とエキスパートから収集した26語と重複している言葉を除いた150語を収集した。その後、自由記述実験により得られた言葉が視覚的印象を評価するのに適しているかどうかを検証する適合度実験を行った。実験参加者は大学生及び大学院生10名(男性5名,女性5名,平均年齢22.5歳,SD=1.02)であった。手続きとして、実験参加者には、画面に呈示されたスーツ画像を観察してもらい、自由記述実験により収集された150語が、スーツスタイルから感じる印象を評価するのに適しているのか「非常に適していない」、「適していない」、「やや適していない」、「どちらともいえない」、「やや適している」、「適している」、「非常に適している」の評価尺度からなる7件法で回答を求めた。「非常に適していない」を1点、「非常に適している」を7点として各評価尺度に対して1点刻みの点数化を行った。その結果、平均値が5.0以上の評価を得た26語を非エキスパートから収集した言葉として採用した。非エキスパートから収集した言葉を表2に示す。

表2 非エキスパートから収集した26語

Table 2 Twenty-six words collected from non-experts.

大人っぽい	シンプルな	スタイリッシュな	落ち着いた	フォーマルな
スマートな	清潔感のある	おしゃれな	カジュアルな	高級感のある
びしっとした	紳士的な	クラシカルな	かっこいい	かっちりとした
モダンな	男らしい	真面目な	クールな	すっきりした
しっかりした	シックな	かしまった	知的な	個性的な
典型的な				

最後に、エキスパートと非エキスパートから収集した52個の評価語から印象評価実験に用いる評価語を選定するために評価語間の意味空間上での置換可能性を評価する距離測定実験を行った。実験参加者は大学生及び大学院生10名(男性5名,女性5名,平均年齢22.5歳,SD=0.90)であった。手続きとして、実験参加者には、スーツスタイルに対する印象を評価する

言葉 A が他の評価語 B で置き換え可能か全 1326 通り評価してもらった。その結果から、評価語間の距離行列を算出した。得られた距離行列に対して、多次元尺度構成法と ward 法を用いた階層的クラスタ分析を行うことで、評価語間の関係を可視化した。可視化した結果を図 4 に示す。この結果からエキスパートが選出した 25 語を、印象評価実験に用いる評価語とした。選定した 25 個の評価語を表 3 に示す。

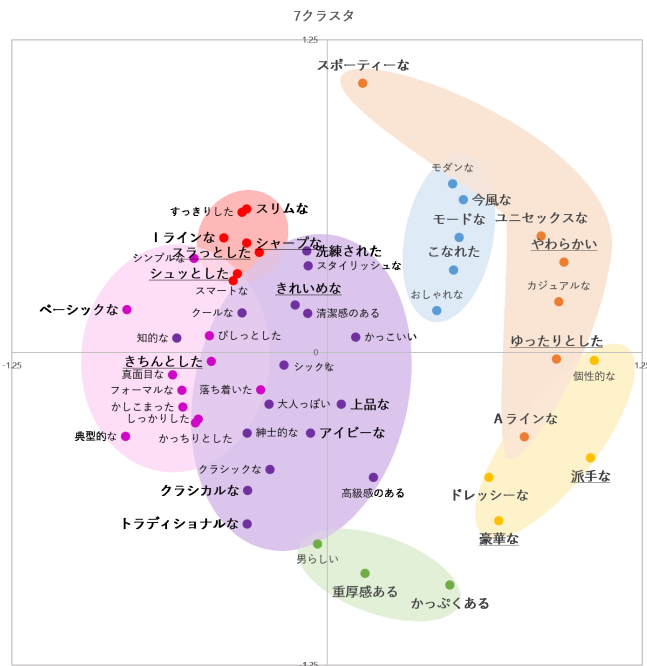


図 4 評価語間の関係を可視化した結果

Fig. 4 Result of visualizing the relationship between the evaluation words.

表 3 印象評価実験に用いる 25 評価語

Table 3 Twenty-five evaluation words used in the impression evaluation experiment.

ドレッシーな	ベーシックな	派手な	トラディショナルな	アイビーな
スポーティーな	ゆったりとした	重厚感ある	やわらかい	上品な
洗練された	クラシカルな	モードな	シャープな	スリームな
きちんとした	大人っぽい	シンプルな	落ち着いた	フォーマルな
清潔感のある	カジュアルな	高級感のある	かっちりとした	シックな

4.3 印象評価実験

スーツスタイルから喚起される視覚的印象を定量化するため、収集した 3,398 枚の画像に対し、クラウドソーシングを利用した印象評価実験を行った。実験にはエキスパートと非エキスパートを問わない 2,240 名が参加した。実験参加者には、画面に呈示されたスーツの画像を観察してもらい、各評価語につきどの程度あてはまるかについて、「非常にあてはまらない」、「かなりあてはまらない」、「ややあてはまらない」、「どちらでもない」、「ややあてはまる」、「かなりあてはまる」、「非常にあてはまる」の評価尺度からなる 7 件法で回答してもらった。なお

順序効果による影響をなくすため、刺激の提示順と評価語の並び順については試行ごとにランダムとした。実験画面を図 5 に示す。これにより各画像に対して、20 人分の 25 種類の評価語



図 5 印象評価実験の実験画面

Fig. 5 Experimental screen of the impression evaluation experiment.

における 7 段階の評価データを取得した。

また、得られた評価データに対し、ダミー画像の評価内容や回答時間からクリーニングを行った。クリーニングの結果、総回答者 2,240 人の内、有効な回答者は 2,055 人であった。各評価尺度を 1 から 7 までの 1 点刻みで得点化した。各評価語の平均点を算出し、それを各画像の印象に対する評価値とした。

5. 視覚的印象と物理的特性の関係性のモデル化

5.1 CNN 特徴の抽出

物理的特徴として、Gatys ら [17] が提案した、コンテンツ特徴とスタイル特徴を用いた。本研究では、ImageNet [21] で学習した VGG-19 [16] の pooling 層 2, 3, 4, 5 で出力されるコンテンツ特徴とスタイル特徴を使用した。それぞれのコンテンツ特徴の次元数は $56 \times 56 \times 128$, $28 \times 28 \times 256$, $14 \times 14 \times 512$, $7 \times 7 \times 512$, スタイル特徴の次元数は 128×128 , 256×256 , 512×512 , 512×512 である。

また、これらの特徴量は、学習するには次元数が大きすぎるため、縮小する必要がある。コンテンツ特徴は PCA を適用し、累積寄与率が 80% となるように縮小した。スタイル特徴は対称行列であるため重複部分を除き、さらに PCA を適用し累積寄与率が 80% となるように縮小した。PCA を適用後の CNN 特徴量の次元数を表 4 に示す。

表 4 PCA 適用後の CNN 特徴量の次元数

Table 4 Number of dimensions of CNN features after applying PCA.

抽出層	PCA 後のコンテンツ特徴の次元数	PCA 後のスタイル特徴の次元数
block2_pool	1118	7
block3_pool	662	9
block4_pool	966	400
block5_pool	359	130

5.2 Random Forest

Random Forest は、分類や回帰を行う機械学習手法の一つで、

決定木を用いたアルゴリズムである。Random Forest では、決定木を複数個構築し、それらから得た結果を組み合わせて予測を行う。個々の決定木は高い識別性能を持たないが、複数の決定木を組み合わせることで、高い精度のモデルを構築することができる。しかし、Random Forest の構造は決定木からなるため、モデルの精度は木の深さに依存し、深くしすぎると過学習を起こしやすい。そこで、最適な木の深さを設定するため、グリッドサーチを用いたパラメータチューニングを行う。本研究では、目的変数を評価値、説明変数を CNN 特徴量とする Random Forest 回帰を行った。また、グリッドサーチの際は、木の深さの探索範囲を 3, 5, 8, 10, 15 とした。各形容詞について pooling 層ごとに回帰モデルを構築し、4 つのモデルのうち相関係数が最も大きいモデルを印象推定モデルとして採用した。

表 5 各特徴量における評価値と推定値の相関係数

Table 5 Correlation coefficients between the evaluation score and the estimated score for each feature.

評価語	コンテンツ特徴	スタイル特徴	連結した特徴量
ドレッシーな	0.554	0.565	0.581
ベーシックな	0.583	0.626	0.648
派手な	0.573	0.669	0.655
トラディショナルな	0.633	0.628	0.642
アイビーな	0.256	0.396	0.306
スポーティーな	0.679	0.680	0.694
ゆったりとした	0.679	0.677	0.688
重厚感ある	0.641	0.673	0.658
やわらかい	0.621	0.663	0.661
上品な	0.619	0.659	0.648
洗練された	0.522	0.510	0.500
クラシカルな	0.668	0.665	0.664
モードな	0.414	0.459	0.427
シャープな	0.553	0.568	0.575
スリムな	0.570	0.580	0.592
きちんとした	0.717	0.701	0.728
大人っぽい	0.549	0.585	0.569
シンプルな	0.609	0.633	0.628
落ち着いた	0.593	0.663	0.651
フォーマルな	0.777	0.765	0.790
清潔感のある	0.553	0.595	0.610
カジュアルな	0.766	0.748	0.783
高級感のある	0.637	0.610	0.630
かっちりとした	0.745	0.733	0.752
シックな	0.585	0.625	0.616
平均	0.604	0.627	0.628

5.3 CNN 特徴量の比較

提案手法の有効性を検証するために、9-fold 交差検証を行った。CNN 特徴量として、コンテンツ特徴、スタイル特徴、コンテンツ特徴とスタイル特徴を連結した特徴量の 3 つを比較した。評価値と推定値との相関係数を推定精度の指標とする。

各特徴量の相関係数を表 5 に示す。全評価語の印象推定モデルの相関係数の平均は、コンテンツ特徴を使用した場合は 0.604、スタイル特徴を使用した場合は 0.627、コンテンツ特徴とスタイル特徴を連結した特徴量の場合は 0.628 であり、全て

中程度以上の正の値を示した。これにより、本手法の有効性が示された。

Gatys らは、コンテンツ特徴は形状の情報を多く保持している特徴量であり、スタイル特徴は色やパタンの情報を多く保持している特徴量であることを示唆した。このことから、生地（質感）とシルエット（形状）から構成されるスーツスタイルの印象推定において、コンテンツ特徴とスタイル特徴を連結した特徴量を用いることで精度が向上すると考えられる。しかし、スタイル特徴の相関係数の平均値は、コンテンツ特徴とスタイル特徴を連結した場合の相関係数と大きな差はなく、コンテンツ特徴の相関係数よりもわずかに高い値となった。この結果から、スタイル特徴はコンテンツ特徴の情報を保持していることが可能性の 1 つとして考えられる。すなわち、スタイル特徴は質感だけでなく、形状の印象推定にも有効であることが示唆された。今後、コンテンツ特徴とスタイル特徴の関係についてより詳細な解析を行う予定である。

6. おわりに

本研究では、スーツスタイルから喚起される視覚的印象を自動で推定する手法を提案し、以下の手順を行った。(1) 視覚的印象を印象評価実験によって定量化し、(2) CNN 特徴量であるコンテンツ特徴とスタイル特徴を物理的特徴として抽出し、(3) 視覚的印象と物理的特徴の関係性を Random Forest によってモデル化した。その後、得られたモデルをもとに、テストデータに対する印象を推定した。評価値と推定値の相関係数が中程度以上の正の値を示したことから、提案手法が有効であることを確認した。コンテンツ特徴とスタイル特徴を比較すると、スタイル特徴の相関係数の平均は、コンテンツ特徴よりもわずかに高く、両特徴量を連結した場合とも大きな違いは示さなかった。仮説では、コンテンツ特徴はシルエットと関連が強く、スタイル特徴は生地と関連が強いと考えられたが、この結果ではむしろスタイル特徴が生地（質感）だけでなくシルエット（形状）の印象推定にも有効である可能性が示唆された。今後は、特徴量や、回帰モデルの比較を行い、高精度なモデルを構築する予定である。

謝辞 本研究に関し、ご支援、ご議論いただいた株式会社センチュリーエール 西岡稔晃氏に感謝致します。また本研究の一部は、国立研究開発法人科学技術振興機構 (JST) の研究成果展開事業センター・オブ・イノベーション (COI) プログラム「感性とデジタル製造を直結し、生活者の創造性を拡張するファブ地球社会創造拠点 (JPMJCE1314)」の支援を受けたものである。

文 献

- [1] 加藤哲夫, “インダストリー 4.0 に想起される新たなイノベーション分類の提言,” 国際 P2M 学会誌, vol.12, no.2, pp.129-144, 2018.
- [2] F. Ross, “Refashioning london’s bespoke and demi-bespoke tailors: New textiles, technology and design in contemporary menswear,”

- The Journal of The Textile Institute, vol.98, no.3, pp.281-288, 2007.
- [3] H. Talebi, and P. Milanfar, "NIMA: Neural image assessment," IEEE Transactions on Image Processing, vol.27, no.8, pp.3398-4011, 2018.
 - [4] L. H. Chew, J. Teo and J. Mountstephens, "Aesthetic preference recognition of 3D shapes using EEG," vol.10, no.2, pp.165-173, 2016.
 - [5] 片平建史, 武藤和仁, 橋本 翔, 飛谷謙介, 長田典子, "SD 法を用いた感性の測定における評価の階層性," 日本感性工学会論文誌, 2018.
 - [6] F. Sibley, "Aesthetic Concepts," The Philosophical Review, vol.68, no.4, pp.421-450, 1959.
 - [7] J. A. Black Jr, K. Kahol, P. Tripathi, P. Kuchi, and S. Panchanathan, "Indexing Natural Images for Retrieval Based on Kansei Factors," In Human Vision and Electronic Imaging IX, vol.5292, pp.363-375, 2004.
 - [8] Y. W. Chen, D. Chen, X. H. Han, and X. Huang, "Generic and Specific Impression Estimation of Clothing Fabric Images Based on Machine Learning," 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp.1753-1757, 2015.
 - [9] 丹羽志門, 青山祥貴, 数藤恭子, 谷口行信, 加藤俊一, "Modeling relationship between visual impression of commodities and their graphical features," IPSJ SIG Technical Reports, vol.2013-HCI-152, no.24, pp.1-4, 2013.
 - [10] Y. W. Chen, X. Huang, D. Chen, and X. H. Han, "Generic and specific impressions estimation and their application to KANSEI-based clothing fabric image retrieval," Int. J. Pattern Recognit. and Artif. Intell, vol.32, no.10, pp1854024, 2018.
 - [11] 飛谷謙介, 田口皓一, 橋本学, 阪下啓祐, 谷伊織, 橋本翔, 片平建史, 長田典子, "多視点画像群を用いた DNN による 3 次元物体の印象推定," 電子情報通信学会論文誌 D, vol.103, no.11, pp.844-848, 2020.
 - [12] B. Julesz, "Textons, the elements of texture perception, and their interactions," Nature, vol.290, no.5802, pp.91-97, 1981.
 - [13] J. Portilla, and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," International Journal of Computer Vision, vol.40, no.1, pp.49-70, 2000.
 - [14] 飛谷謙介, 白岩 史, 片平建史, 長田典子, 荷方邦夫, 荒川 薫, "Modeling of "high-class feeling" on a cosmetic package design," Journal of the Japan Society of Precision Engineering, vol.87, no.1, pp.134-139, 2021.
 - [15] N. Sunda, K. Tobitani, I. Tani, Y. Tani, N. Nagata, and N. Morita, "Impression estimation model for clothing patterns using neural style features," HCI International 2020, pp.689-697, 2020.
 - [16] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," The 3rd International Conference on Learning Representations, 2015.
 - [17] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," The IEEE Conference on Computer Vision and Pattern Recognition 2016, pp.2414-2423, 2016.
 - [18] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "Deepfashion: Powering robust clothes recognition and retrieval with rich annotations," The IEEE Conference on Computer Vision and Pattern Recognition 2016, 2016.
 - [19] N. Xu, B. L. Price, S. Cohen, and T. S. Huang, "Deep image matting. The IEEE Conference on Computer Vision and Pattern Recognition 2017," pp.2970-2979, 2017.
 - [20] 飛谷謙介, 松本達也, 谿 雄祐, 藤井宏樹, 長田典子, "素肌の質感表現における印象と物理特性の関係性のモデル化," The Journal of Image Information and Television Engineers, vol.71, no.11, pp.J259-J268, 2017.
 - [21] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, L. FeiFei, "ImageNet large scale visual recognition challenge," International Journal of Computer Vision, vol.115, pp.211-252, 2015.