様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　　５ 年 １０ 月 ２３ 日現在

機関番号：１５１０１
研究種目：基盤研究(C)（一般）
研究期間：2020～2022
課題番号：２０Ｋ１１８６４

研究課題名（和文）人間の視覚特性に基づく人物印象属性の自動推定に関する研究

研究課題名（英文）A Study on Automatic Estimation of Impression Attributes based on Human Visual Function

研究代表者

西山　正志（NISHIYAMA, Masashi）

鳥取大学・工学研究科・教授

研究者番号：２０７５６４４９

交付決定額（研究期間全体）：（直接経費）　　3,400,000円

研究成果の概要（和文）：美しい，上品である，清潔感があるなどの印象を人物画像から感じる際に，その画像のどこにいつ注目したかを表す視線位置の時系列変化を本研究で明らかにした．視線位置は初期時刻に被写体の鼻へ近づき，その後は印象単語に応じて視線位置に差異が生じていくことを定量的に確認する方法を開発した．さらに，視線位置の時系列変化を認識モデルに導入し人間の見方と近づけることで，深層学習における新たなアテンション機構を開発し，推定精度が高まることを示した．

研究成果の学術的意義や社会的意義
システムが自動で人間の印象を推定できるようになると，その印象を手掛かりとしたコーディネートサービスを新たに創造できると考える．冠婚葬祭など非日常の特別な場で印象は重要となる．例えば，知人の結婚式へ出席するため，相応しい格好が与える印象を考えている．開発した推定方法を導入したシステムを構築できれば，これまでにないユーザ体験の提供が期待できる．その状況に適した格好を利用者は安心して選ぶことができ，自分では気付いていない格好を知る機会を得ることができる．

研究成果の概要（英文）：We investigated the gaze time-series signals, which indicate where and when the observer pays attention to a region of an image when perceiving the subject as beautiful, elegant, or clean. We developed a method to show that the gaze location approaches the subject’s nose initially, and then the gaze location varies according to the impression words. Furthermore, we developed a new attention mechanism by introducing the gaze time-series signals into a deep-learning model and making it closer to human perception. We confirmed that the model improved the estimation accuracy.

研究分野：画像認識

キーワード：人物印象属性　自動推定　視覚特性

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

１．研究開始当初の背景

美しい，可愛い，上品である，素敵である，清潔感があるなど，人物の見た目から感じる印象を本研究では取り扱う．冠婚葬祭など非日常の特別な場で印象は重要となる．例えば，知人の結婚式へ出席するために相応しい格好を選ぶことを考える．印象を新たな手掛かりとして格好をコーディネートするシステムを構築できれば，これまでにないユーザ体験の提供が期待できると考える．その状況に適した格好を利用者は安心して選ぶことができ，自分では気付いていない格好を知る機会を得ることができると考える．

本研究では大勢の人間が社会的な通念を共有できる状況での人物印象属性を前提とする．例えば結婚式で撮影された図１の人物画像を観察した場合，美しいや上品や可愛いなど共通の印象を人間は感じる．状況に適した印象を自動で推定するタスクは，画像を用いた深層学習などデータドリブンの手法と相性がよいと言える．ただし，人物画像と印象単語のペアを訓練サンプルとするデータドリブンの推定手法のみでは，システムが人間と同じ様に感じることができるレベルまで到達できるとは言えない．

図１：画像から感じる人物印象属性の例

本研究では，人間と同じ様に印象を感じるシステムを構築する手掛かりを探ることを学術的な主な問いかけとする．システムのカメラで撮影された人物画像を実際の人間が観察する際に，画像中の人物の見た目そのものが印象を決める手掛かりとなる．本研究では，人間が印象を判断する際の観察過程に着目し，人物画像のどこにある部位をどの順に見たかの知見をシステムに取り入れることを考える．観察される身体部位の位置と順序が，印象を表す単語によって違うという仮説を設ける．同じ順番で身体部位を見たとしても，その部位を見ていた停留時間が変わる可能性がある．

２．研究の目的

本研究の目的は，人物印象属性の自動推定を対象とし，人間が情報システムへ教師ラベルを与える作業を通して，視覚特性を計測しアルゴリズムに組み込む方法を開発することにある．人物印象属性に対する視覚特性を，ラベル付け作業時に高精細で計測する方法を探っていく．また本研究では深層学習のアテンション機構を発展させ，印象に対する視覚特性を組み込んだネットワークモデルを設計していく．本モデルを学習するために，人間が感じた印象と，モデルから出力された印象との相違を，視覚特性と共にフィードバックしていく．

３．研究の方法

項目１：印象を表す単語を設定し人物画像に付与することでデータセットを構築

印象属性を扱う上で，個人の感じ方の違いを起因とする難しさが想定される．その理由として，その人の性別，年代，置かれている状況などが考えらえる．このような難しさに対応するために，本研究では想定する応用先を絞り込む．応用の一例としてパーティーや式典など，非日常的な特別な場での衣装選びの状況を取り扱い，フォーマルな場に相応しい人物印象属性を対象とし主観評価を行う．実験協力者として日本人大学生の男女を対象とする．同年代同士での印象に関する共通性を明らかにしつつ，印象単語と人物画像のペアからなるデータセットを構築する．

項目２：人物画像の印象を人間が見分ける時の視覚特性を計測

印象を見分ける時の観察者の視線を高精細に計測するために，本研究では，人物画像を表示するディスプレイと視線計測装置を用いて実験協力者に刺激を与えていく．印象に対する視線の位置と順序を正確に捉えるために必要な表示装置と計測装置の時空間分解能を検証していく．さらに本研究では，刺激画像中の人物姿勢の自由度を高めるために，画像から身体部位を検出することで視線が集まる領域の位置合わせを行う．これにより据置型だけでなく装着型の視線計測装置も使用できる．視線計測を行う際，実験協力者に印象を判断させる指示を与えた上で，人物画像を一枚ずつ表示し，その印象に当てはまるかどうかを回答させる．

項目３：視覚特性に基づくフィードバックを用いて人物印象属性を判別するモデルを学習

本研究では，人間の視覚特性を深層学習モデルにアテンション機構として取り込んでいく．画像認識におけるアテンション機構とは，判別精度を高めるために画像中のどの位置の特徴を重視するかを制御することである．近年，アテンション機構を視線位置分布で更新する手法が登場したが，視線の順序関係は未だ導入に至っていない．本研究ではアテンション機構を発展させ，印象に対する視覚特性を組み込んだ深層学習モデルを設計していく．本モデルの特徴は，位置アテンションと順序アテンションをもつことにある．位置アテンションでは，人間が観察する位置と同じ位置を重視するように特徴抽出を制御する．順序アテンションでは，人間と同じ順序で観察するようデコーダを制御する．前段層では視線が集まる身体部位へ重みを詳細に与え，後段層では全体に対して重みを大まかに与える．本モデルを学習するために，教師ラベル付与時に人間が感じた印象とモデル出力の相違を各損失の中で考慮する．判別誤りの損失に加え，位置と順序の推定誤りの損失を導入し，それらの損失への相違フィードバックを組み込む．このために，判別タスクの印象で観察すべき視覚特性の重みを大きくし，誤って推定された印象で観察していた視覚特性の重みを小さくする．学習されたモデルのアテンション機構を可視化することで，人間の視覚特性に近づける本研究の考え方が有効であるかどうかを，推定精度と照らし合わせ説明していく．

４．研究成果

項目１の成果（印象単語が付与された人物画像データセット）

印象単語として，大勢の人間が社会的な通念を共有できるフォーマルな場面に当てはまるものを対象とした．具体的には，プレゼンテーションやパーティーなどの場面で感じると考えられる印象単語を対象とした．なお，フォーマルな場では相手にポジティブな印象を与えることを望むことが多いと考えられるため，ネガティブな印象を表す単語を除外した．想定する場面に合う印象単語を決めるためフリーディスカッションを行った．その結果，実験で用いる印象単語を以下の６個に決定した．

- W1 ： 温厚 （Gentle）
- W2 ： 意欲的 （Ambitious）
- W3 ： 個性的 （Unique）
- W4 ： 裕福 （Rich）
- W5 ： 知性的 （Intellectual）
- W6 ： スタイリッシュ （Stylish）

印象単語 W1 から W6 に対して感じる主観が，どの程度似ているかを確認するために実験を行った．この実験に参加した各人物に対して印象単語に関する多段階評価アンケートを実施した．具体的には，印象単語のペアを提示し，それらの単語から感じる印象がどれだけ似ているかについて尋ねた．各人物に以下から１つを選ぶよう回答させた．

- 似ている （1 点）
- やや似ている （2 点）
- やや似ていない （4 点）
- 似ていない （5 点）

なお括弧内の数字は主観スコアを表す．印象単語のペアとし W1 から W6 までの組合せの全通りを対象とし，尋ねる順序はランダムとした．印象単語の各ペアに対して，アンケートに参加した全人物から得た主観スコアの合計を求めた．これを距離行列とし，非計量の多次元尺度構成法（Multi Dimensional Scaling: MDS）を適用した．印象単語の主観的な近さを，非計量 MDS を用いて平面に配置した結果を図２に示す．意欲的の W2 と裕福の W4 とは，他の印象単語 W1，W3，W5，W6 と比較すると，平面上で近くに配置されていた．ただし，似ている，もしくは，やや似ているを全員が選んだわけではなかった．実験で用いた印象単語 W1 から W6 は，似ているかどうかの感じ方という観点で，重なりは無く，ある程度のばらつきを持っていると考えられる．

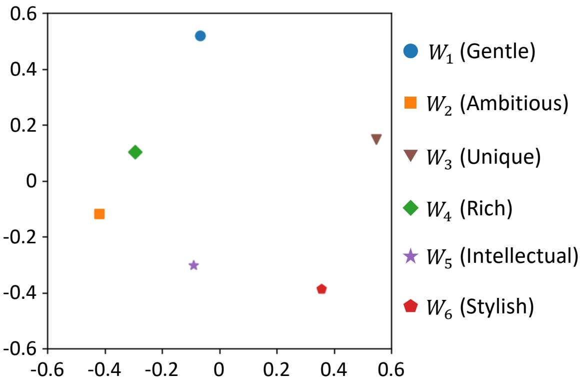

図２：印象単語の近さの可視化結果

次に，実験協力者が観察する刺激画像について詳細を設計した．その刺激画像の一部を図３に示す．画像中の被写体の姿勢について，ポスター会場や立食懇親会など，フォーマルな場面で多く遭遇する立位状態を対象とした．手や足の位置について，両手や踵をそろえるなどの条件は設けずに，自由とした．また，被写体の身体領域について頭からつま先まで全て見えるとし，かつ，被写体の顔が正面向きに近いとした．なお座位状態や臥位状態を除いた．刺激画像１枚に含まれ

る被写体の数を1名のみとした．被写体の男女比を1:1とした．被写体の服装として，プレゼンテーションやパーティーなどフォーマルな場面に合うものとした．フォーマルな場面では社交のため，笑顔やニュートラルの表情になると想定される．よって被写体の表情を，笑顔，もしくは，ニュートラルとした．刺激画像中の背景として，被写体より目立つものは存在しないとした．また，被写体の手前には他の物体が存在しないとした．

項目２の成果（視覚特性の計測結果）

　まず，視覚特性（具体的には視線位置）を計測するための実験環境を構築した．そのセッティングを図 4 に示す．センターバイアスを避けるため，ディスプレイ上のランダムな位置に刺激画像を表示した．視線位置を計測する際のサンプリングレートについて，装置の仕様上，視線位置の計測時刻の間隔にばらつきが生じるため，バイリニア補間を用いて等間隔となるようリサンプリングした．刺激画像中の被写体と実験協力者とは，直接の面識はなく親しくない間柄とした．視線位置計測時において，実験協力者は W1 から W6 の全ての印象単語に回答することとした．

　次に，視覚特性を計測し分析するための方法を設計した．本研究における視覚特性の分析では，計測された視線位置から身体部位までの距離を用いた．ここでは図 5 の身体部位を用いた．印象単語に関する主観評価タスクを実験協力者へ与えた際，刺激画像を観察した時刻で計測された視線位置から身体部位までの距離を分析の対象にすることにした．具体例として，左肩に対する距離を計測する場合を図 6 に示す．次に，全ての実験協力者を含む集合と，全ての刺激画像を含む集合とを用いて，各時刻における身体部位の位置までの平均距離を求めた．本研究では，各時刻における平均距離の値が時間方向に並んだ信号を，視線位置の時系列変化と呼ぶことにした．

図 3：刺激画像の例

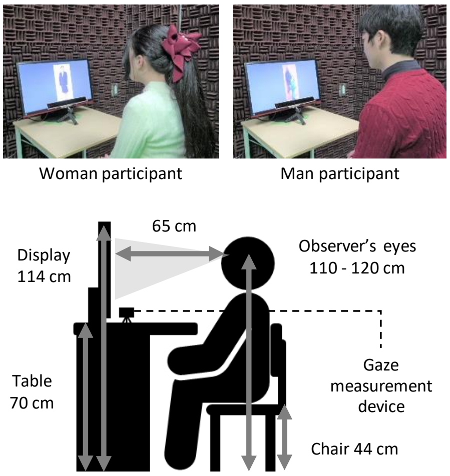

図 4：刺激画像を観察する時の視線位置を計測するための環境

図 5：視線分析に用いる身体部位（ここでは代表的な関節を利用）

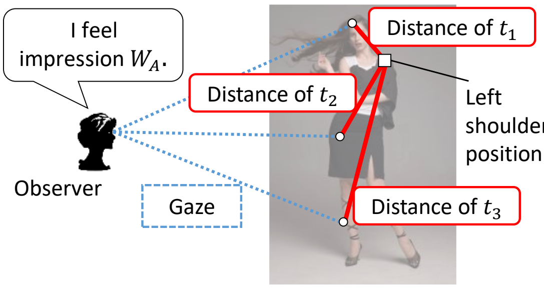

図 6：視線位置の時系列変化における各時刻の距離の算出例

視線位置の時系列変化に，印象単語の間でどのような差異が表れるかについて可視化し分析を行った．各印象単語 W1 から W6 における各時刻の視線位置について，身体部位までの距離の時系列変化を図 6 に示す．なお本実験では，身体部位として鼻に注目し考察を進めた．図中より，初期時刻 0 秒から 0.6 秒までは，印象単語間でほぼ共通の動きをしていることが分かる．時刻 0.6 秒から 6 秒までは，時間経過と共に距離が一旦大きくなり，その後は小さくなるという共通の傾向が見られるものの，距離の増減のタイミングに，印象単語の間で違いが見られた．以上の結果より，全印象単語において，実験協力者の視線位置は最初に鼻へ近づくことは共通であることを明らかにした．また，視線位置の移動の仕方には，鼻を見た後，印象単語の間で差異が生じる傾向があることが分かった．

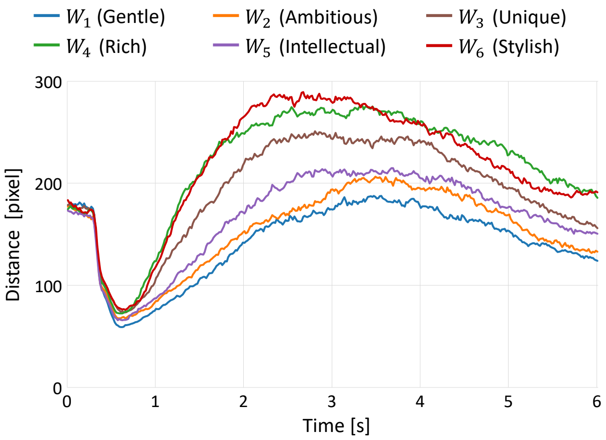

図 6：印象単語 W1 から W6 における各時刻の視線位置の時系列変化

項目３の成果（視覚特性に基づく判別モデル学習の性能評価）

ここでは視線位置の時系列変化を組み込んだ判別モデルを開発した．まず，本研究において開発した提案手法のベースとなるセルフアテンションの考え方について説明する．一般的なセルフアテンション層では，入力された特徴マップ上の一つのセルに与える重みを計算する際に，そのセルの値と近い値をもつ複数のセルを利用する．これにより，近傍に配置されているセルだけでなく，遠い位置に配置されているセルとの関係を，重み計算に取り入れることができる強みをもつ．しかし，既存手法のセルフアテンションをそのまま用いた場合，特徴マップの全てのセルをアテンション算出に等しく用いるため，人間の視覚特性とは異なる構造になっていると言える．そこで提案手法（GSA: Gaze-guided self-attention）では，人間から計測された視線位置分布を用いた重み付けをセルフアテンション層へ明示的に組み込む．これにより，人間が画像を見ている場合と近い形で，特徴マップ上で該当するセルの重みを大きくする重みづけを判別モデル内で行うことができる．提案手法の推定精度を表 1 に示す．一般的な深層学習の手法（B1, B2）や視線位置を用いる既存手法（GP）やセルフアテンションの既存手法（SA, CBAM, EA）と比較して推定精度が改善されていることを明らかにした．ただし，提案期間で構築した人物印象データセットでは画像枚数が少なかったため，この実験ではある程度の画像枚数を確保できる人物画像の性別認識で精度を評価した．今後はデータセット構築を継続し，印象認識の推定精度を確かめていく予定である．また，提案手法の GSA は位置アテンションであり，順序アテンションの導入について，この問題設定では効果が得られるところまで評価できていなかった．今後の課題として，静止画でなく動画像を対象として，位置と順序に関する人間の視覚特性を学習モデルに導入する方法を開発していく予定である．

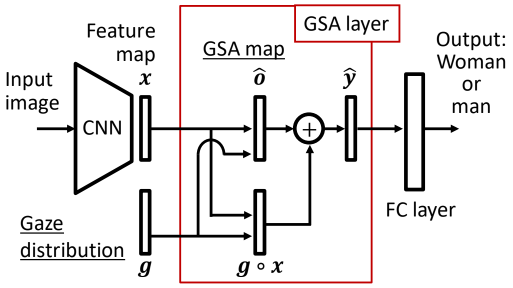

図 7：人間の視覚特性を導入したセルフアテンション（GSA）

略称	手法	推定精度
B1	MiniCNN	28.8±11.6
B2	ResNet50	35.8±4.4
GP	Gaze pooling	56.0±1.9
SA	Self-attention	29.8±10.6
CBAM	Convolutional block attention	33.8±5.6
EA	Efficient attention	24.2±7.2
GSA	**Gaze-guided self-attention (ours)**	**64.2±4.6**

表 1：推定精度の比較

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　0件／うちオープンアクセス　0件）

１．著者名 NISHIYAMA Masashi, INOUE Michiko, IWAI Yoshio	４．巻 E105.D
２．論文標題 Gender Recognition Using a Gaze-Guided Self-Attention Mechanism Robust Against Background Bias in Training Samples	５．発行年 2022年
３．雑誌名 IEICE Transactions on Information and Systems	６．最初と最後の頁 415～426
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2021EDP7117	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 河口 野乃加，木下 顕，加藤 直人，井上 路子，西山 正志，岩井 儀雄	４．巻 87
２．論文標題 観察者が人物画像の印象を判断する時に計測された視線位置の時系列変化に関する分析	５．発行年 2021年
３．雑誌名 精密工学会誌	６．最初と最後の頁 965～974
掲載論文のDOI（デジタルオブジェクト識別子） 10.2493/jjspe.87.965	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

１．著者名 木下 顕，井上 路子，西山 正志，岩井 儀雄	４．巻 86
２．論文標題 身体への印象単語を判別する際に計測された視線位置分布の確率表現	５．発行年 2020年
３．雑誌名 精密工学会誌	６．最初と最後の頁 989～996
掲載論文のDOI（デジタルオブジェクト識別子） 10.2493/jjspe.86.989	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 －

〔学会発表〕　計11件（うち招待講演　1件／うち国際学会　3件）

１．発表者名 西山 正志
２．発表標題 人間の視覚に基づく画像認識の技術開発
３．学会等名 産業技術連携推進会議 情報通信・エレクトロニクス部会 情報技術分科会分科会 情報通信研究会（招待講演）
４．発表年 2022年

１．発表者名 Michiko Inoue, Masashi Nishiyama, Yoshio Iwai
２．発表標題 Subjective Scores and Gaze Distribution in Personality Evaluations: Effect of Subjects' Clothing on Observers' Impressions of Them
３．学会等名 Proceedings of 24th International Conference on Human-Computer Interaction (HCII)（国際学会）
４．発表年 2022年

１．発表者名 Fuyuko Iwasaki, Masashi Nishiyama, Yoshio Iwai
２．発表標題 Investigation of temporal changes of gaze locations during characteristic evaluation when viewing whole-body photos
３．学会等名 Proceedings of 24th International Conference on Human-Computer Interaction (HCII)（国際学会）
４．発表年 2022年

１．発表者名 岩崎 芙由子，西山 正志，岩井 儀雄
２．発表標題 動画中の被写体を観察する際に計測された視線分布を三次元人体モデルで可視化する手法の検討
３．学会等名 画像の認識・理解シンポジウム(MIRU)
４．発表年 2022年

１．発表者名 Ken Kinoshita, Michiko Inoue, Masashi Nishiyama, Yoshio Iwai
２．発表標題 Body-part Attention Probability for Measuring Gaze During Impression Word Evaluation
３．学会等名 Proceedings of 23rd International Conference on Human-Computer Interaction（国際学会）
４．発表年 2021年

１．発表者名
木下 顕，西山 正志，岩井 儀雄

２．発表標題
観察者の視線が集まる身体部位の三次元可視化の提案

３．学会等名
HCGシンポジウム

４．発表年
2021年

１．発表者名
岩崎 芙由子，西山 正志，岩井 儀雄

２．発表標題
全身画像や頭隠れ画像から受ける衣服印象に関する視線の時系列変化の調査

３．学会等名
HCGシンポジウム

４．発表年
2021年

１．発表者名
井上 路子，西山 正志，岩井 儀雄

２．発表標題
被写体の服装カテゴリが及ぼす性格印象の主観スコア及び視線分布の調査

３．学会等名
ヒューマンインタフェースシンポジウム

４．発表年
2021年

１．発表者名
井上 路子，西山 正志，岩井 儀雄

２．発表標題
被写体の性格印象への主観スコアと視線位置の評価 ～刺激画像が頭部または全身の場合の比較～

３．学会等名
情報処理学会シンポジウム インタラクション

４．発表年
2021年

１．発表者名 河口 野乃加，木下 顕，井上 路子，西山 正志，岩井 儀雄
２．発表標題 人物画像の印象判断時における視線位置の時間変化の解析
３．学会等名 ビジョン技術の実利用ワークショップ（ViEW）
４．発表年 2020年

１．発表者名 西山 正志，山口 紗也加，岩井 儀雄
２．発表標題 視線位置分布を組み込んだセルフアテンションを用いた性別認識
３．学会等名 画像センシングシンポジウム（SSII）
４．発表年 2020年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

－

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究協力者	岩崎　芙由子 (IWASAKI Fuyuko)		
研究協力者	岩井　儀雄 (IWAI Yoshio) (70294163)	鳥取大学・工学研究科・教授 (15101)	
研究協力者	井上　路子 (INOUE Michiko) (10980724)	鳥取大学・工学研究科・助教 (15101)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関