

「見た目」によるバイアスを軽減するための Web ページデザインの自動変換

澤浦 司[†] 莊司 慶行[†] Martin J. Dürst[†]

[†] 青山学院大学 理工学部 情報テクノロジー学科 〒252-5258 神奈川県 相模原市 中央区 淵野辺

E-mail: [†]sawaura@sw.it.aoyama.ac.jp, ^{††}shoji@it.aoyama.ac.jp, ^{†††}duerst@it.aoyama.ac.jp

あらまし 本論文では、任意のページのデザインを変換し、別のページ集合のもつ「見た目」に近づける手法を提案する。Web ページにアクセスした際に、同一の内容でも、視覚的な違いによって意思決定に影響が出る場合がある。本研究ではこのような違いを「見た目」と定義し、敵対的学習に似たアルゴリズムを用いて Web デザインを自動変換する。この手法では、従来の情報抽出によるルールベースのデザイン変換と異なり、スクリーンショットを用いた CNN（畳み込みニューラルネットワーク）と遺伝的アルゴリズムを組み合わせる。はじめに、与えられた 1 つのページがあるページ集合に属するか判定する CNN 分類器を学習する。次に、分類器を騙せるまで、遺伝的アルゴリズムでページデザインの変更を繰り返す。これにより、任意のページに、別のページ集合のもつ「見た目」を付与できる。手法の有用性を確認するため、人為的に制作した Web ページからなるデータセットを用いて、被験者実験で評価した。4 種類の異なるカテゴリに属すページ集合間で変換を行い、変換後のページの属すカテゴリを被験者に推定させた。実験結果より、多くの場合、正しく「見た目」の変換が行えていることが分かった。

キーワード Web デザイン, CNN, 見た目, 印象, ページ変換

1 はじめに

Web ページの「見た目」は、時として人を面食らわせる。誰しも一度は、けげんばしい Web サイトや胡散臭いデザインの Web サイトを訪問し、内容を読むことなくブラウザを閉じたことがあるだろう。Web 全体の規模が増加し、用途やデザインが多様化したことで、ユーザが多くの Web サイトを比較し、情報を選別する必要に迫られる機会が増えてきている。例えば、あるユーザが「新しくテレビを買いたい」と思った場合を考える。ユーザはメーカーの公式サイトや通販サイトでスペックを調べたり、他の購入者の感想を知るためにレビューサイトやブログを閲覧したりと、様々な種類、様々なデザインのページを訪れる必要がある。この時、「デザインが格式高いページだったため、内容を吟味することなく、何となく閲覧を敬遠する」といった行動が一般的に見られる。このように、内容が同じであったとしても、ページの見た目が意思決定に影響を及ぼす場合がある。なかでも、視覚的な要素を原因とする第一印象がもたらす影響は、閲覧者の公平な情報比較の機会を減少させる [1]。

本研究ではこのような視覚的な要素の違いを「見た目」と定義し、ページ集合から「見た目」を抽出し、任意のページに付与することで、情報閲覧時のバイアスを軽減させる。例えば、コンピュータを扱うページが共通して持つ「見た目」を抽出し、それを化粧品ページのページに付与する、といったようなことを可能にする。このような変換が可能になれば、普段コンピュータ関係のページを多く閲覧しているユーザが「化粧品」に関する情報収集をする際などに、見慣れたデザインでの情報比較が可能になると考えられる。

そのために、CNN（畳み込みニューラルネットワーク）を用いた分類器と、遺伝的アルゴリズムに基づく CSS コードの変換機を用いた、Web ページの「見た目」を変換するアルゴリズムを提案する。このアルゴリズムでは、ある似通った「見た目」を持つページ集合と、任意の 1 つのページを入力とする。ページ集合から抽出した「見た目」を入力された 1 つのページに付与することで、「見た目」の差異を減少させる。このような潜在的特徴を見つけ出し、その特徴を有するデータを生成する課題には、敵対的生成ネットワーク (Generative Adversarial Network, GAN) を用いることが多い。しかし、今回のような Web ページに対し、通常の GAN を用いることは困難である。1 つ目の理由として、例えば「青か紫は正解クラスに属す『見た目』であるが、青紫はそうではない」というように、HTML や CSS が持つデザインに対して数値を連続的に変化させたときの判定結果が離散的である可能性が挙げられる。2 つ目の理由として、コンテンツやページの機能を保ったまま、文法的に正しいコードを GAN で生成することの困難性が挙げられる。そのため本手法では、GAN に近い発想のアプローチをとるが、分類と生成を個別に行った。

本研究の学術的貢献として、Web 一般の「見た目」を、ページの種別やページの構造に関わらず変換可能にしたことが挙げられる。情報比較を行う際に「見た目」が印象に占める影響は非常に大きいことが指摘されている [2]。「見た目」の差を減少させることで、複数の Web ページ上の情報を比較する際のバイアスの軽減を見込めるが、「見た目」に影響する要素を特定し、変換することは、困難である。「見た目」に関するコードは HTML や CSS の中に分割して記述されており、なおかつページの見た目は人間の認知に基づくため直接測定できないため

ある。しかし、あるページの「見た目」が他のどのようなページ集合の「見た目」に近いかどうかは、レンダリング後のページのスクリーンショットを用いれば、画像的に推定可能である。そこで本研究では、共通する「見た目」をもつページ集合（例として、同一カテゴリの商品を扱うページなど）のスクリーンショットから、「あるページがそのページ集合に属するか」を推定する CNN 分類器を学習する。次に、その集合に属さない任意のページについて、学習した分類器が「属す」と誤判定するまで、遺伝的アルゴリズムでページデザインを書き換える。このようなスクリーンショットによる画像分類によるアプローチをとることで、従来の情報抽出による手法や単なる CSS の切り替えでは対応できなかった「漠然と似た雰囲気を持つが異なる機能を持つページ群」などでも、「見た目」の変換が可能になる。

提案したアルゴリズムの実現可能性と有用性を議論するため、実際にページの「見た目」を変換して、そのページがどのカテゴリに属すと思うかラベル付けする被験者実験を行った。実験の際には、カテゴリごとに似たような見た目を持つページを自動的に多数生成するプログラムを作成し、4 種類の異なる「見た目」のページ集合を作成した。それぞれに含まれるページについて、カテゴリ間で相互に「見た目」を変換し、変換したページがどの集合に属するかを人手でラベル付けし、手法を評価した。

以下に、本論文の構成を述べる。第 2 章で本研究の対象とする問題を改めて定義する。第 3 章では本研究で提案する手法を述べる。第 4 章では評価のための実装と、実際の実験とその結果について述べる。第 5 章で実験を通して得られた結果に関して考察を行う。第 6 章では、関連研究について論じる。第 7 章でまとめと今後の展望について述べる。

2 問題定義

本研究で実際に解決する課題を整理し、問題を実際に定義する。そのために、まず Web ページ閲覧時の第一印象に影響を与える「見た目」を定義する。その後、問題解決のために考えられるいくつかの方針と、提案手法である「見た目」を抽出し付与するアプローチについて述べる。

2.1 「見た目」の定義

本研究では、情報の内容そのものは変わらないが、人の受ける印象に影響を与える視覚的特徴を「見た目」と定義する。見た目の単純な例を、図 1 に挙げる。この例では、4 つの異なるデザインでそれぞれ「500 円」という同じ情報を伝えている。この際、フォントや色使いなど、様々な外見特徴により、人がそれを一目見た時に、高そうあるいは安そうと感じるか、信頼感を感じるかなど、第一印象に大きな差が生じる。

このように、Web ページを閲覧する際、閲覧者は Web ページの「見た目」から、様々な第一印象を受ける。例えば「楽しそう」、「見慣れていて安心感がある」といった好印象もあれば、「このサイトは怪しい」、「何となく格式高すぎる」といった悪感情を覚える場合もある。第一印象は閲覧時の感情に影響を及ぼ

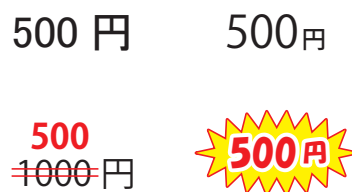


図 1 「500 円」という情報を表す複数の見た目を伴う表示。この場合、伝えたい情報の内容（500 円という価格）は変わらないが、デザインの違いで受ける印象は大きく異なる。

し、意思決定にも影響を与えることが知られている [3, 4]。Web 閲覧者による情報比較の際、第一印象が異なることは公平な比較機会の減少につながる。視覚的な印象が認知に与える影響は大きく、「見た目」の差異を減少させることは情報比較の機会を保つために重要である。

2.2 ページ集合に共通する「見た目」

Web ページの外見の細部はページ毎に異なるが、「見た目」に近いページ集合は存在すると考えられる。例えば、ある特定のジャンルに含まれる商品群に関する、複数のショッピングサイトの販売画面を収集したとする。収集したページに「手書き文字フォント」、「ピンクを基調としたパステルカラー」、「曲線が多く、角が丸い」という特徴をもつページが多く含まれた場合、直感的に、これらのページのトピックは、「コンピュータ」よりも「化粧品」である方が、可能性が高い。これは、先述の特徴が「化粧品」を扱うデザインに含まれていること、あるいは「コンピュータ」を扱うデザインに含まれていないことを意味する。こういったように、同じ読者層を想定したページ集合であったり、同じカテゴリに属する情報を扱うページ集合は、同一ではないにしろ何かしらの「見た目」上の共通点を有していると仮定する。

コンテンツはお互いに異なるが「見た目」に統一性がある実際のウェブページの例として、楽天市場の「アウトドア」カテゴリと、「弦楽器」カテゴリに属す、それぞれ異なる販売店のページのスクリーンショットを図 2 に示す。この例では、別の販売店でありながら、「アウトドア」カテゴリの各販売ページは赤系統の色を使いがちで、広い面積を強い色で塗りつぶす傾向があることが現れている。一方で、弦楽器カテゴリのそれぞれのページは、オレンジや茶色などの色を使い、余白を広くとっている。これらの各ページのデザインは、各販売店がそのカテゴリの顧客に合わせて独自に設定したものと考えられるが、カテゴリごとのデザイン傾向の違いが現れている。

2.3 問題設定と達成目標

本研究で対象とする問題として、Web ページの「見た目」を自動で変換する。図 3 に表すように、具体的な入力は、「見た目」上の共通点を含むページ集合と任意の 1 ページである。ページ集合から「見た目」の共通点を抽出し、入力として与えられた 1 つのページのデザインを変換したものを出力する。ページ集

「アウトドア」カテゴリの楽天ページ



「弦楽器」カテゴリの楽天ページ



図 2 楽天市場の 2 つのカテゴリに属す、それぞれ異なる販売店のページのスクリーンショット（実際の商品情報を含むため、ぼかし加工を施した）。

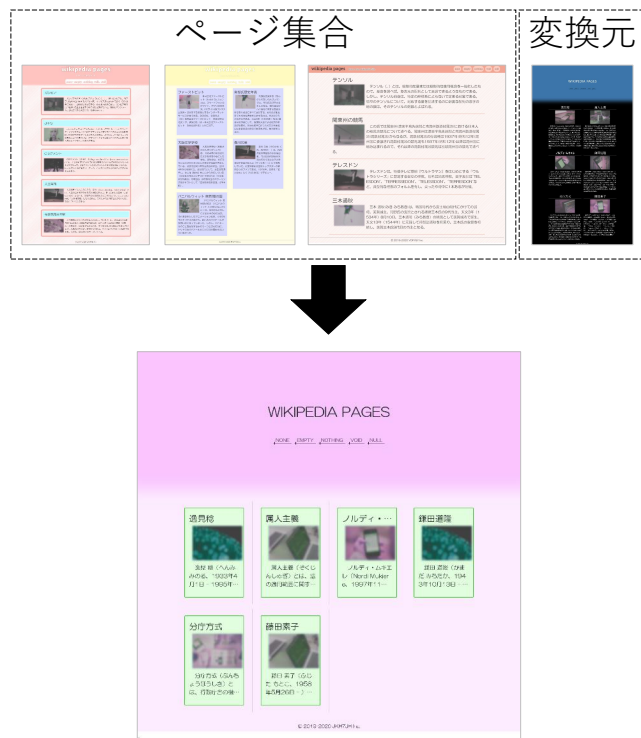


図 3 設定した問題の入出力例。図中の各画像は、実際に実験で用いたスクリーンショットである。

合から抽出した「見た目」を 1 つのページに付与することで、そのページがあたかもページ集合に属すかのような「見た目」に変更されれば、目標を達成できたとみなす。

2.4 本研究における解決方針

このようなページ集合から「見た目」を抽出し、それを与え

られた任意の 1 つのページに付与する方法はいくつか考えられる。ひとつの解決方法として、古典的な情報抽出手法によるページデザインの変換が挙げられる。情報抽出では、あるページのどの部分に何の情報が書いてあるかを特定することを可能にしている。これらの技術を使えば、ページ内で同じ意味を持つ内容を入れ替えることで、ページデザインの相互変換が可能である。一方で、こういった手法では、同じ情報を扱っているサイト同士しか変換できない。すなわち、同じ粒度のショッピングサイト同士では変換できるが、パソコンのレビューページと販売ページといったような、似たデザインを持ちながらコンテンツの異なるページには対応できない。

他の解決へのアプローチとして、敵対的生成ネットワーク (Generative Adversarial Network, GAN) が、画像や文体のスタイル変換で用いられている。本研究で取り扱うような「見た目」は、スタイルの一種であると考えられる。しかしながら本研究では、「見た目」の判定には画像と CNN 分類器を用いるが、生成には GAN をはじめとしたニューラルネットワークを使用しない。GAN によってスクリーンショットを直接変更した場合、Web ページの機能が失われる (すなわち、画像化した時点で、リンクやボタンをクリックできなくなる)。そのため、画像の中身を変更するために、HTML や CSS 中に散発的に記述された、文法的制約の強いコードを変更する必要がある。このような文法規則の強いコンピュータ言語を GAN で扱うのは困難である。

そこで、本研究では、スクリーンショットを分類する CNN 分類器と、ページデザインを遺伝的アルゴリズム (Genetic Algorithm, GA) に則って変更する変換機を、それぞれ個別に作成する。

3 提案手法

本章では、Web ページ集合が持つ「見た目」を、任意のページに付与する変換手法について述べる。提案するアルゴリズムの概要を図 4 に示す。提案手法は、入力としてページ集合と任意のひとつのページを与えられた際に、出力として任意のページをあたかもページ集合に属すかのように変換したものを返すアルゴリズムである。アルゴリズムを構成するモジュールの概要を節 3.1 で述べる。また、アルゴリズムの実際の計算手順について、前処理に関して 3.2 節で、分類器に関して 3.3 節で、変換器の作成に関して 3.4 節でそれぞれ述べる。

3.1 手法の概要

提案手法は大きく分けて、2 つのモジュールから形成される。ひとつめのモジュールは、ページのスクリーンショットから、あるページの「見た目」が、与えられたページ集合に属すかを判定する分類器からなる。デザインが完全に同一でなくても、人間が見た際に、何となく「これらは同じようなジャンルのページだろう」と推測できるような、漠然とした共通性を持つページ集合がある。図 4 中の例では、与えられたページ集合 P_{target} は、全て IT 機器に関するページである。例として、これらの

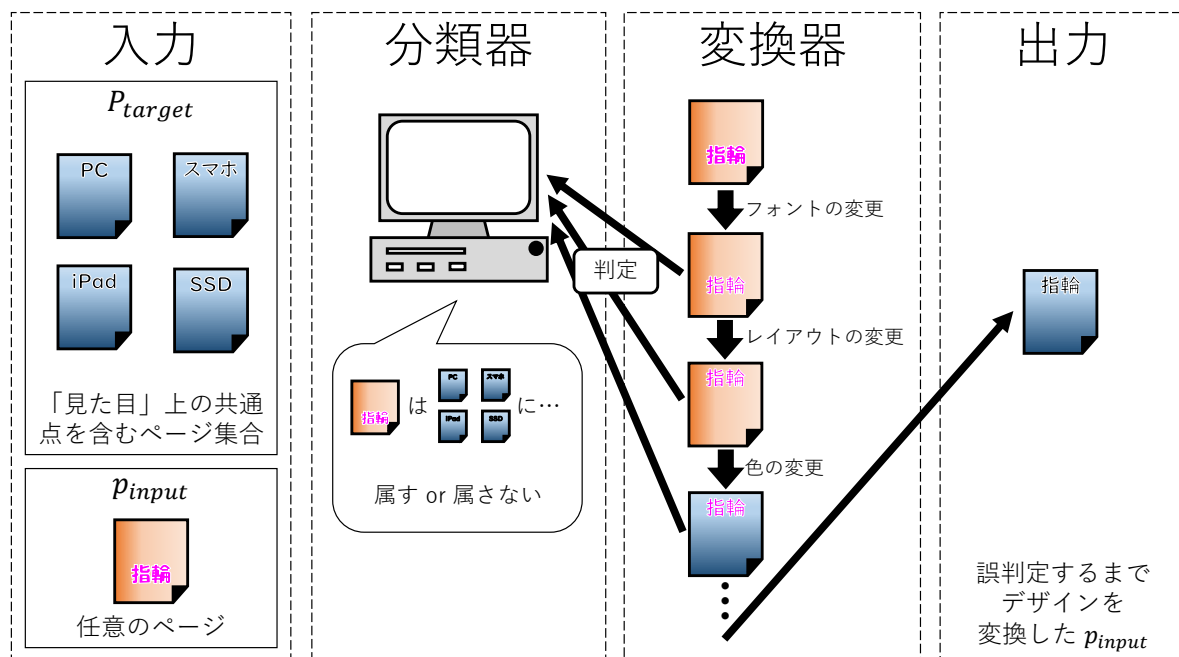


図 4 提案手法の概要. 入力されたページがあるページ集合に属するかを判定する分類器を用い、変換器がページデザインの変更と判定を繰り返す. この例では、IT 機器のページ集合が持つ「見た目」を、指輪のページに与えている.

ページは、ソリッドで寒色系のデザインを持つ場合が多い. そこで、任意のページが、その集合に属すかどうかを、ページのスクリーンショットのみから、CNN 分類器によって分類する.

二つ目のモジュールは、変換器である. この変換器は、任意のページ p_{input} に対して、CSS や HTML の編集を行い、「見た目」を変更する. 変換器は、もともと異なる「見た目」を持った入力ページ p_{input} を、「このページは P_{target} に属す」と分類器に誤判定されるまで、段階的に変更を繰り返す. この変換器は、進化的アルゴリズムの 1 つである遺伝的アルゴリズムを用い、見た目に関連する要素、すなわち使われているフォントや、画面のレイアウト、色調などを変更する.

最終的に分類器が「このページは P_{target} に属す」と誤判定するまで p_{input} に対して変換器でデザイン変更を行ったページを、アルゴリズムの最終出力とする.

3.2 前処理

入力されたページ集合 P_{target} と任意のページ p_{input} について、「見た目」を判定するために前処理を行う. HTML と CSS からなるコードは、そのままだと人間が第一印象を覚えるような、「見た目」を伴った形式をとっていない. そこで、一般的なウェブブラウザでレンダリングし、スクリーンショットを撮影し、画像化する. 今回用いる物を含め、一般的な CNN 分類器は入力が固定長のベクトルであり、事前に入力する画像の大きさを揃えておく必要がある. そこで、前処理として画像のランダムクロッピングを行う.

スクリーンショットの撮影について、近年見られる多くのウェブページは、一般的に横長の PC モニタに対して、スクロールを伴う縦長のページを有している場合が多い. そこで、はじめ

にページの全体を 1 枚の縦長の画像に変換する. スクロール機能を有するヘッドレスブラウザを用いることで、HTML や CSS, JavaScript による動的生成によってレンダリングされる Web ページを観たままの画像に変換できる.

次に、ランダムクロッピングの処理では、ウェブページのスクリーンショットから共通した大きさを持つ矩形（ここでは分類器の入力サイズに合わせ、224px 角とした）を複数回ランダムに切り取り、小さい画像の集合に変換する.

クロッピング処理の目的として、画像の大域的な特徴と、画像の細部に表れる特徴を、両方反映させることが挙げられる. 例えば、使用しているフォントが持つ特徴などは、画像全体の拡大縮小によって、もしくは学習の過程で潰れてしまうことが考えられる. そのため、小さく分割した画像でより多様な特徴の学習を図る.

また、Web ページの「見た目」の特徴は一様に分布しているわけではない. 例えば、コンテンツが無く一色の背景だけである部分に関して、その部分をページの特徴と見なすことは困難である. そこで、クロッピング後の画像について、全てのピクセルが同色であるようなものは事前に除去した.

3.3 CNN 分類器の学習と判定

画像認識の分野では、各ピクセル間の関係をニューラルネットワークに学習するため、あらかじめ画像セットで学習したモデルから、転移学習を用いて実際のデータに適応する場合が多い. 本研究においても、同様に事前学習済みのモデルから転移学習する.

実際に分類に用いる CNN 分類器の詳細について説明する. 本研究では、モデルの基盤として ResNet-50 モデルを用いた.

このモデルは 50 層からなる 1,000 値分類用のモデルであるので、「あるページがページ集合に属するか」という 2 値分類を行うため、1,000 値分類用の全結合層を 2 値分類用のものに変更して用いる。

分類器を用いて p_{input} が P_{target} に属するか予測する方法について述べる。分類器の入力に合わせて p_{input} をクロッピングし、それを入力とすることで、ページが P_{target} に属するかを 2 値で出力する。この時、クロップされた場所が局所的に異なるデザインを持っている場合、その他の大多数の部分クロップしたときと比較して、意図しない外れ値が出る可能性が考えられる。

この問題に対処するため、 p_{input} に対して複数回、ランダムな位置に対してクロッピングを行い、 P_{target} に属すると判定された回数の割合を、 p_{input} が P_{target} に属する確率 $\text{Sim}(p_{\text{input}}, P_{\text{target}})$ とする。これにより、あるページ p_{input} が P_{target} に属するかを、確率値で表現可能になる。

3.4 遺伝的アルゴリズムによる変換器の作成

変換器は、異なる「見た目」を持つページ p_{input} を P_{target} が持つ「見た目」に近づけるために、遺伝的アルゴリズムで学習と変換を繰り返す。変換の際は、前節で作成した分類器が与える確率 $\text{Sim}(p_{\text{input}}, P_{\text{target}})$ を 1 に近づけることを目標とする。ページの変換対象となるのは、使用する色の傾向とコントラスト、各ブロックの角のデザインおよび枠線の有無、使用フォントの種類と大きさ、画像および文章の表示位置、余白の大きさである。これらの変更を CSS のコードで表記し、変換前のページに対して適用する。これにより $\text{Sim}(p_{\text{input}}, P_{\text{target}})$ が上昇すれば、要素の変更の有意性が認められる。

以上の変換を、遺伝的アルゴリズムによって繰り返し実行する。各個体の適応度を $\text{Sim}(p_{\text{input}}, P_{\text{target}})$ として与える。1 世代内の個体の適応度が十分大きい値になった時、 p_{input} は P_{target} に属するように変換されたと判定し、変換を終了する。

4 評価実験

本章では、提案手法の評価のための、実装の詳細および評価実験とその結果について述べる。実装について 4.1 節で説明し、実際に使用したデータセットの詳細について 4.2 節、評価実験で設定したタスクについて 4.3 節でそれぞれ述べる。実験結果については 4.4 節で述べる。

4.1 実装

提案手法の実現のため、CNN 分類器と GA (Genetic Algorithm, 遺伝的アルゴリズム) による変換器を実装した。本節では、それぞれの実装における詳細を述べる。

分類器として、本研究ではモデルの基盤として ResNet-50 モデルを用いた。このモデルは 50 層からなる 1,000 値分類用の汎用画像認識モデルで、ImageNet が提供する 1,000,000 枚を超える画像セットで訓練済みのものである。また、本研究では 2 値分類を行うため、1,000 値分類のための全結合層を 2 値分類用のものに変更した。

表 1 GA 学習時のパラメータおよび各遺伝子座と対立遺伝子の一覧

種別	項目	値
パラメータ	学習世代数	100 世代
	集団サイズ	25 体
	世代交代モデル	Simple GA
	選択方式	Roulette Wheel Selection
	交叉方式	Two Point Crossover
	個体突然変異確率	10%
遺伝子座	遺伝子突然変異確率	10%
	基本色	16,777,216 色
	拡大率	61 段階 (0.90 - 1.50)
	余白	31 段階 (-2.0 - 1.0)
	段組み	4 段階 (1 - 4)
	背景スタイル	7 種類
	字体スタイル	3 種類
	見出しスタイル	9 種類
	タイトルスタイル	8 種類
	本文スタイル	8 種類
	画像スタイル	3 種類

表 2 あらかじめ人手で設定したクラスごとのデザインの特徴

	キュート	クール	うさん臭い	ナチュラル
使う色の傾向	パステルカラー	黒, 寒色	黄, 赤	薄緑
色のコントラスト	弱い	強い	弱い	弱い
項目ごとの角のデザイン	丸い	鋭い	無し	鋭い
使用するフォント	丸ゴシック系	明朝系	ゴシック系	ゴシック系
余白の広さ	普通	広い	広い	普通
画像の量	多い	普通	少ない	普通
文章の量	少ない	普通	普通	少ない

ページデザインの変換器は、HTML と CSS、および CNN 分類器を入力として、デザインを変換した HTML を出力する。本実装では、遺伝的アルゴリズムを用いてデザイン変換を行う。データセットに対して、設定した学習時のパラメータおよび、各遺伝子座とその対立遺伝子を表 1 に示す。

4.2 データセット

本実験では、4 つのクラス (キュート, クール, うさん臭い, ナチュラル) を設定し、実際に Web サイトを閲覧しながら人手でそれぞれのサイトの持つ特徴を列挙した。次に、列挙した特徴を人手でグループ分けし、「このクラスのサイト群はこういうデザインを持ちがちである」という要素にまとめた。それぞれの持つデザイン要素を表 2 に示す。

次に、各クラスが持つ特徴を、それぞれ HTML と CSS からなる、複数種類のテンプレートとして記述した。それらのテンプレートをランダムに組み合わせることで、各クラスに属するページを生成した。ページのコンテンツは wikipedia¹ から無作為に抽出した本文、画像は PlaceIMG² のものを使用した。こうして作成されたデータセットを用い、CNN 分類器の学習を行った。学習には、各クラス 400 枚ずつの画像を用いた。

4.3 評価タスク

提案した問題が実際に解決可能かどうか、提案手法が実際に働くかどうかを、人手で作成したデータセットを用いた実験に

1 : <https://ja.wikipedia.org/wiki>

2 : <http://placeimg.com/>

表 3 人手で作成したデータセットにおける実験結果. 回答者が変換後のクラスにより近いと回答した回数を表す. 各質問は計 20 回ずつ行われた.

変換前 \ 変換後	キュート	クール	うさん臭い	ナチュラル
キュート	——	18	9	12
クール	17	——	12	20
うさん臭い	11	13	——	10
ナチュラル	14	17	6	——

よって評価した. 実験では, 実際にあるページのデザインを別のページ集合のデザインに変換し, 変換後のページを 5 秒を目安に被験者に提示し, 実際にデザインが変換されているかをラベル付けさせた.

データセットは複数の Web ページのスクリーンショットからなる. 各ページは「可愛い」, 「クール」などの印象クラスに属すページが含まれる.

実験は複数の質問から構成される. 1 つの質問では, 2 種類のクラス A , B のページ群と, A から B へと変換したページを, それぞれスクリーンショットで提示する. この際, 各クラスのページ群のどちらが変換前(変換後)であるかは提示せず, 回答者はより近いと考えたクラスを選択する. この質問で, 回答者がクラス B を選択した場合, このページに対する変換が成功しているとみなす.

実際の評価システムのスクリーンショットを図 5 に示す. 本実験では, 異なる 2 クラスの順序対からなる全 12 組の変換について, それぞれ 2 回ずつ, 計 24 問の設問を提示した. 回答者は 10 人である.

4.4 結 果

それぞれの設問に対して, 変換前に属していたクラスよりも変換後のクラスに近いと回答された回数を表 3 に示す. 全 240 回の評価クエリのうち, 159 回で変換後のクラスのほうが近いと回答され, 変換の成功率は 66.3%となった.

また, 各列および各行の最低値を見ると, 全てが「うさん臭い」クラスに関わる変換であることが分かる. 実際に, 「うさん臭い」クラスに関わらない 120 回の変換では, 81.7%にあたる 98 回の変換が成功している. 「うさん臭い」クラスに関わる変換は, 変換前後に関わらず有効な変換が行えていないことを示している.

回答者間の意見の一致度を測るために, Cohen の重み付き κ 係数を用いる. これは通常の κ 係数に, 差に比例する重みを加えて計算する指標である. 各回答者対ごとの, 重み付き κ 係数による回答の一致度を平均すると 0.189 となり, 被験者間にわずかな相関 (slight agreement) があることが示された [5]. そのため, 被験者間の Web ページ閲覧時の第一印象には, 少なくとも意味のある一致が存在する.

5 考 察

評価実験での結果が導かれた理由を考察し論じる. 実験にお



図 5 被験者実験システムのスクリーンショット. 2 つのページ集合と, 1 つのページが提示される. 被験者はそのページが左右どちらのクラスに属するかを入力する.

いて, 異なるデザインを持つページ集合に対して, 実際に提案手法を適用し, 被験者評価を行った. 評価実験の結果から, ほとんどのクラス間において, 変換は正しく行われたと被験者に判断された.

実験結果全体から, スクリーンショットを CNN 分類器によって分類すると, 人の認知と近く「見た目」を判別できることが明らかになった. また, GA を用いることで, 「見た目」に合わせてデザインを変換できることが示された.

一方で, 上手くいかなかった例として, 「うさん臭い」クラスの存在が変換成功率へ与える影響について考察する. 「うさん臭い」クラスが関係する変換では, このクラスが変換の前後のどちらであっても関係なく成功率が低かった. これは, ページ閲覧時に目に入る画像の量の多寡が「見た目」に占める影響が大きかったためと考えられる.

表 2 にあるように, 「うさん臭い」クラスのジェネレータが生成するページには画像の量が少ない. 具体的には他クラスのページが 3 枚から 7 枚の画像を持ちうるのに対し, 「うさん臭い」クラスのページには 1 枚しか画像が含まれない.

Web ページ内に含まれる画像は, デザインの一部であると同時に, それ自体が情報量を持っている. 本研究はデザインのみを変換することで「見た目」がもたらす第一印象の変化を評価した. ページが持つ情報の量の差が第一印象にもたらす影響を考慮せず, 「見た目」のみを論じるために, GA による変換の際には画像および文章の量は変化させていない. そのため, 2 クラスのページ集合をそれぞれ提示した際, 「画像が多いほうと少ないほう」という差が明確に現れ, 容易に識別可能となったと考えられる.

次に, 「うさん臭い」クラスを除く 3 クラス間のクラス変換について述べる. こちらでは, (キュート, ナチュラル) および (ナチュラル, キュート) の変換において, 他クラス間の変換と比べ, 被験者の正答率が相対的に低かった. これらの 2 クラスは, 他のクラス対と比較して生成されるページの持つデザインの色調が近かったことが考えられる. 「ナチュラル」が薄緑を基調とするページなのに対して, 「キュート」はパステルな色調を持つページである. このように, ページ全体の色が近いクラス対では, どちらのクラスの分類器も「自クラスに属する」ペー

ジであると判定するようなページが多く生成された。

逆に、他クラスと比較してまったく異なる背景色を持つ「クール」クラスに関しては、多くの変換で被験者に正しく変換できていると見なされた。これらの事実より、ページが持つ色の傾向は第一印象に大きな影響を与える「見た目」の一つであることを示している。

最後に、変換手法の実用性について論じる。本評価実験で行った変換手法は、実際のすべての Web ページの変換には、そのまま適用することができない。変換の候補となる集団を表 1 に示したようにごく小さい範囲に定めているためである。今回、変換の候補は 10 種のパラメータに対するテンプレートとして定めた。しかし実際の Web ページを構成する要素は膨大であり、テンプレートを用意することは現実的でない。ページに含まれる画像の数および大きさ、使用される色の傾向や強調の方法など、商品カテゴリ内で「見た目」上の共通点は存在する。これらの HTML および CSS から共通するコード断片を抽出できればデザインのテンプレートが構成可能である。しかし、こういったページ集合は、HTML および CSS のソースコードとして見た場合には、共通したコードは含まれていない。そのため、「見た目」上の共通点を持つ実際の Web ページに対して、その共通点がソースコード上で各要素およびタグにどう表れるかを区別し、デザイン特徴ごとに構築し、表現する必要がある。

本稿における評価の際は、テンプレートの組み合わせで生成され得る全てのページを表現できた。これは生成されるページのデザインの組み合わせが高々4,600 兆通り程度であり GA による学習が現実的であったためである。実際の Web ページを変換する場合、ページデザインとコードによる表現は無数に存在し、探索空間は無限に近くなる。そのため本手法と同様に変換の候補を現実的な範囲まで絞る必要があると考えられる。

6 関連研究

本研究は、「見た目」がもたらす Web 情報閲覧時のバイアス軽減を目標に、Web デザインを自動変換するものである。そこで本章では、Web ページ閲覧時の視覚がもたらす印象、バイアス軽減のための異なるアプローチである情報抽出、および Web デザインの自動生成に関する研究について述べ、本研究の立ち位置を明らかにする。

6.1 Web ページの視覚による印象

第一印象を構成する要素の中でも、内容に影響しない視覚的な印象が閲覧者の認知・第一印象に占める割合は大きく、研究も盛んである。Tractinsky ら [2] は、視覚的な印象が最初に検出され閲覧者の経験に影響を及ぼすことを示している。また、Lindgaard ら [6] は、ページを閲覧してから僅か 50 ミリ秒の間に視覚が受ける情報が印象の好悪を決定づけるとしている。これまでに議論されてきた視覚的な印象を構成する要素は、画像・テキスト・アニメーション・リンクの量、トップページの長さ [7] や、左上隅の要素、使用単語がもつ感情の傾向 [8] など、多岐にわたる。

6.2 情報抽出によるページ変換

非構造化データから構造化された情報を抽出する、情報抽出分野の研究が盛んである。

情報抽出の研究では、Web ページの主コンテンツを認識し、抽出する研究もおこなわれている。Labsky ら [9] は、HTML で作成された企業の製品カタログからの情報抽出手法を提案している。この研究では、サイトまたは企業ごとに異なる製品の記載フォーマットに対して、隠れマルコフモデルと多層パーセプトロンを独立に用いることでルールベースでない情報抽出を可能にしている。Gogar ら [10] は、Web ページに対して CNN を適用し、通販サイトのスクリーンショットから商品のメイン画像と現在の価格の位置を予測している。

これらの技術でも、例えば見た目と無関係にページのコンテンツを抽出して提示したり、抽出したコンテンツに一樣な見た目を付与することは可能である。Web ページに対する情報抽出がコンテンツの抽出・付与によって目標となる情報の獲得を目指すのに対し、本研究ではデザインの抽出・付与によって、コンテンツを欠落させることなく、公平な情報比較を促す。

6.3 Web デザインの自動生成

Web 閲覧時のユーザ体験の向上などのため、Web デザインの自動生成に関する研究が行われている。自然言語をクエリとして Web デザインテンプレートを生成する研究 [11] や、Web デザインの大規模な機械学習のためのプラットフォームの構築 [12] が行われている。

本研究と同様に、目標となる「見た目」を指定することでデザインを生成する研究も行われている。Kumar ら [13] は、2つの HTML 間で視覚・意味的に類似したページ要素をマッピングする手法を提案している。この研究では作成したマッピング結果を用いることで、実際に 2つのページからそれぞれデザインとコンテンツを抽出し合成することが可能である。Beltramelli ら [14] は、Web サイトやアプリケーションの GUI 部分のイラストから、実際のコードを生成する手法を提案している。この際、コードとスクリーンショットを独立に機械学習し、得られた結果を再度学習することで、解釈可能なコードを生成している。これらの研究では、目標となる「見た目」を HTML や画像で直接指定しているが、本研究では、目標の「見た目」を入力されたページ集合から抽出している。

また、生成したページに対して得られる判定結果を利用して目標の Web ページを構築する研究も行われている。Sorn ら [15] は、対話型遺伝的アルゴリズムを用いて Web ページデザインのテンプレートの自動生成を実現している。この研究では、HTML や CSS の複雑性・構造化から学習器に対する報酬の決定が難しいという問題を、ユーザとの対話インタフェースを提供し評価を逐次入力とすることで解決している。

本研究では、ページデザインの変換という目的に応じて、遺伝的アルゴリズムによる簡易な CSS の変換を行った。今後、より高度なデザイン変換や、高効率化のためには、これらの手法で使われているようなコード生成的アプローチが必要になると考えられる。

7 まとめと今後の課題

本研究では、Web ページの閲覧時に受ける第一印象に影響を与える視覚的特徴を「見た目」と定義し、ある任意のページとページ集合を入力として与えると、あたかもそのページがページ集合に属しているかのように「見た目」を変換する手法を提案した。提案手法は、あるページがページ集合に属するか判定する CNN 分類器と、ページデザインの変更を繰り返す Web ページの変換器を遺伝的アルゴリズムの上で組み合わせるアルゴリズムである。提案手法に対して、人手で作成したデータセットを用いた実験を行った。実験結果の分析から、提案手法が多くの場合正しく「見た目」を変換可能であること、人間の「見た目」の認識において画像の量と色が重要であることがわかった。

今後の課題として、実際の Web サイトへの適用に向けた学習および変換の精度上昇があげられる。人為的に作成したデータセットと比べて、実際のウェブサイトはデザインのパターンが多い。またカテゴリ間でのデザイン要素の重複が多く、不可分な場合もある。より高度な CNN 分類器を使ったり、データのクレンジングや、スクリーンショットの拡張やマスクといった処理により、改善できる可能性がある。

評価実験では、ページ全体のスクリーンショットを提示した。しかしブラウザを用いた Web ページの閲覧時には、ページの全てが一度に視認されるわけではない。Web ページが持つ「見た目」は、ページの一定部分に集約している可能性の考慮についても、今後の課題である。

また、現時点では提案手法の適用によって、本研究の目的である、実際に意思決定への影響を減少させられるかを評価する段階までは到達していない。本研究の貢献は、「見た目」の構成に寄与する要素の提示および分類器と変換器による手法の有用性を示すに留まっている。本研究を礎とし、手法の各段階が持つ課題を解決することで、実際の Web ページに対する適応および意思決定への影響の測定という、今回提起した問題の解決が可能になると考えられる。

謝 辞

本研究は JSPS 科研費 18K18161 (代表: 莊司慶行), 18H03243 (代表: 田中克己) の助成を受けたものです。ここに記して謝意を表します。

文 献

- [1] Bassam Hasan. Perceived irritation in online shopping: The impact of website design characteristics. *Computers in Human Behavior*, Vol. 54, pp. 224–230, 2016.
- [2] Noam Tractinsky, Adi S Katz, and Dror Ikar. What is beautiful is usable. *Interacting with computers*, Vol. 13, No. 2, pp. 127–145, 2000.
- [3] Bo N Schenkman and Fredrik U Jönsson. Aesthetics and preferences of web pages. *Behaviour & Information Technology*, Vol. 19, No. 5, pp. 367–377, 2000.
- [4] Andrea Basso, David Goldberg, Steven Greenspan, and David Weimer. First impressions: Emotional and cognitive factors underlying judgments of trust e-commerce. In *Proceedings of the 3rd ACM conference on Electronic Commerce*, pp. 137–143. ACM, 2001.
- [5] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, pp. 159–174, 1977.
- [6] Gitte Lindgaard, Gary Fernandes, Cathy Dudek, and Judith Brown. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & information technology*, Vol. 25, No. 2, pp. 115–126, 2006.
- [7] Gary Geissler, George Zinkhan, and Richard T Watson. Web home page complexity and communication effectiveness. *Journal of the Association for Information Systems*, Vol. 2, No. 1, p. 2, 2001.
- [8] Eleni Michailidou, Simon Harper, and Sean Bechhofer. Visual complexity and aesthetic perception of web pages. In *Proceedings of the 26th annual ACM international conference on Design of communication*, pp. 215–224. ACM, 2008.
- [9] Martin Labsky, Vojtech Svatek, Ondrej Svab, Pavel Praks, Michal Kratky, and Vaclav Snasel. Information extraction from html product catalogues: from source code and images to rdf. In *The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI'05)*, pp. 401–404. IEEE, 2005.
- [10] Tomas Gogar, Ondrej Hubacek, and Jan Sedivy. Deep neural networks for web page information extraction. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pp. 154–163. Springer, 2016.
- [11] Nandini Sethi, Abhishek Kumar, and Rohit Swami. Automated web development: theme detection and code generation using mix-nlp. In *Proceedings of the Third International Conference on Advanced Informatics for Computing Research*, pp. 1–6, 2019.
- [12] Arvind Satyanarayan, Maxine Lim, and Scott Klemmer. A platform for large-scale machine learning on web design. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pp. 1697–1702. ACM, 2012.
- [13] Ranjitha Kumar, Jerry O Talton, Salman Ahmad, and Scott R Klemmer. Bricolage: example-based retargeting for web design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2197–2206. ACM, 2011.
- [14] Tony Beltramelli. pix2code: Generating code from a graphical user interface screenshot. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, p. 3. ACM, 2018.
- [15] Davy Sorn and Sunisa Rimcharoen. Web page template design using interactive genetic algorithm. In *2013 international computer science and engineering conference (IC-SEC)*, pp. 201–206. IEEE, 2013.

- [1] Bassam Hasan. Perceived irritation in online shopping: The impact of website design characteristics. *Computers in Human Behavior*, Vol. 54, pp. 224–230, 2016.
- [2] Noam Tractinsky, Adi S Katz, and Dror Ikar. What is beautiful is usable. *Interacting with computers*, Vol. 13, No. 2, pp. 127–145, 2000.
- [3] Bo N Schenkman and Fredrik U Jönsson. Aesthetics and preferences of web pages. *Behaviour & Information Technology*, Vol. 19, No. 5, pp. 367–377, 2000.
- [4] Andrea Basso, David Goldberg, Steven Greenspan, and David Weimer. First impressions: Emotional and cogni-