

# 物語展開を考慮した小説データからの表紙の自動生成

川口 晴会<sup>†</sup> 鈴木 伸崇<sup>††</sup>

<sup>†</sup> 筑波大学 図書館情報メディア研究科 〒 305-8550 茨城県つくば市春日 1-2

<sup>††</sup> 筑波大学 図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1-2

E-mail: <sup>†</sup>hkawaguchi@klis.tsukuba.ac.jp, <sup>††</sup>nsuzuki@slis.tsukuba.ac.jp

**あらまし** 本研究では小説を選ぶ際の支援になると考え、小説データからその内容の印象に沿った表紙を自動的に生成するシステムを提案する。読者にとって最も印象に残る場面が小説全体の印象に影響を与えられられることから、本手法では物語展開を考慮して場面を抽出した後、その場面のみを用いて表紙の構成要素である色、フォント、象徴物を推薦する。青空文庫をコーパスとして極性辞書を構築し、この辞書を用いて小説を時系列データに変換して特異値分解によって物語展開の特徴を求めた。特徴に基づいて読者の印象に残ると思われる場面の抽出を行う。色については色と単語を対応させた色彩データベースを用い、フォントについては各フォントと小説本文の印象を表した感性ベクトルを使用して推薦を行う。象徴物については、tf-idfを用いて抽出場面に出現する特徴的な名詞を獲得する。  
**キーワード** 表紙, 場面抽出, 色推薦, フォント推薦, 象徴物推薦

## 1 はじめに

電子書籍と電子書籍リーダーの普及により、近年多くの電子書籍が利用可能となっている。それと同時に、テキストファイルを電子書籍フォーマットである EPUB [1] やその他のフォーマットに変換するオンラインサービスやツール<sup>1</sup>が登場し、任意のテキストから電子書籍を作成して読むことができるようになった。例えば、著作権の消滅した作品を提供する青空文庫<sup>2</sup>の作品や、オンライン小説をダウンロードして電子書籍に加工することが考えられる。このような電子書籍の作成における主な問題として、多くの場合に適切な表紙が与えられていないことがある。表紙は書籍の印象を読者に伝える重要な要素であり、読む前の選別で大きな指標となる [2]。しかし、前述のサービスやツールでは入力テキストの印象を反映した表紙を自動的に付与することはできず、人手でそのような表紙を作成するには手間やコストがかかる。特に未読の小説を電子書籍に加工する場合、内容を反映した表紙を作成することは極めて困難である。

そこで、本稿では小説データからその内容の印象に沿った表紙の構成要素を推薦するシステムを提案する。ここで、本稿では表紙の構成要素として、感性を伝達する色とフォント、小説の内容を象徴する象徴物の3点に着目する。推薦された色、フォント、象徴物を用いることで、容易に表紙を作成することが可能となる。小説から色とフォントを推薦した筆者らの研究 [3] の評価実験では、読者が小説の一場面の印象を強く持っていた場合に推薦の精度が低いという問題があった。また、小説全体の単語を同じ重みで扱っているため、バッドエンドの話であっても幸福な場面が多ければ明るい表紙になる可能性がある。そのため、小説全体の本文を解析し、読者の印象に残ると考えられる場面を抽出後、その場面のみを用いて3点の推薦を行う手

法を提案する。色については色と単語を対応させた色彩データベースを用い、フォントについては各フォントと書籍本文の印象を表した感性ベクトルを使用して推薦を行う。象徴物については、tf-idfを用いて抽出場面に出現する特徴的な名詞を獲得する。推薦された色は表紙の背景色、フォントはタイトルや著者名の描画、象徴物はその物体の画像として表紙に使用できる。

小説の印象を伝えるために色に着目した研究では、書籍の本文とレビューの形容詞から抽出した色を背景色として表紙画像を生成した研究 [2] や、文献 [2] を発展させて小説から象徴物も合わせて抽出し、無料のイラストダウンロードサイトを用いて表紙の中央に描画する研究 [4] がある。これらの研究では、色 (RGB 値) と形容詞、色と形容詞の結びつきがどれだけ強いかを表すスコアの3つを格納した色彩データベースを使用して、辞書一致で色ごとのスコアを算出している。形容詞のみを対象として色を抽出しているが、その他の単語も色と結びつく可能性があり小説全体の印象とは言い難い。また、入力に書籍本文だけでなくレビューも用いた方が精度が高いとしているが、その選択基準については考察されていない。しかし、実際に購入サイトなどで描かれているレビューは一個人によって書かれているため内容の偏った意見や感想が書かれている可能性があり、そもそもレビューがない書籍も少なからず存在する。

フォント推薦に関する先行研究としては、入力されたテキストに合うフォントを推薦するもの [5] がある。この研究ではアンケート調査によって色とフォントに感性ベクトルを付与し、テキスト中の単語から推薦された色の感性ベクトルと類似度が高いフォントを推薦するという流れになっている。推薦の対象となるフォントは、文献 [5] の中でアンケートで感性ベクトルを得たフォントに限られる。また、入力テキストに対して色からフォントを推薦するという流れのため、色とフォントの組み合わせが固定化してしまう。

本研究では、色とフォントの組み合わせが固定化せずに多様な推薦が行われることを期待して、別のアプローチで推薦を行

1 : <https://ebook.online-convert.com/>

2 : <https://www.aozora.gr.jp/>

う。色抽出については、[2]を参考にして色彩データベースを構築し、辞書一致でスコアを算出する。書籍本文のみを入力として形容詞に加え名詞と動詞からも印象を抽出する。それに加え、小説全体ではなく印象に残ると思われる場面のみを用いることで、より具体的な象徴物を抽出できると考えられる。フォント推薦に関しては[5]を参考にして感性ベクトルを用いるが、未知のフォントも推薦の対象となるように深層学習を用いて感性ベクトルの推定を行えるようにする。

## 2 提案手法

提案手法は、小説の本文テキストから読者の印象に残ると思われる場面の抽出と、その場面を用いて行う色・フォント・象徴物の推薦からなる。それぞれの手法について詳しく説明する。

### 2.1 場面抽出

読者の印象に残ると思われる場面を特定するために、何らかの観点で場面ごとのスコア付けをする必要がある。本手法では、幸福度の観点に基づいてスコア付けされた単語の辞書を用いて小説を時系列データに変換し、分析を行っている研究[6]を参考にする。この時系列データを物語展開とみなし、青空文庫コーパスの小説を時系列データに変換して、特異値分解を用いて物語展開に多く見られる複数の特徴を得る。それぞれの特徴がハッピーエンドとバッドエンドのどちらであるかを、あらかじめ人手で付与する。入力された小説データを時系列データに変換し、どの特徴との類似度が高いかによって、その小説がハッピーエンドかバッドエンドであるかを判断する。ハッピーエンドであれば小説全体でもっとも幸福である場面、バッドエンドであれば不幸である場面を抽出する。

#### 2.1.1 物語展開の分類

良い印象を持つ（ポジティブな）単語が多く出現する場面は幸福で明るい場面、逆に悪い印象を持つ（ネガティブな）単語が多く出現する場面は不幸で暗い場面であると仮定する。文献[6]では、クラウドソーシングを用いて約10,000の英単語について幸福度を付与した辞書[7]を利用している。日本語の単語について、新たにクラウドソーシングなどを使用して辞書を作成するのは高コストであり好ましくない。そこで、単語がどれほどポジティブもしくはネガティブであるかという極性の強さを数値として表した極性辞書を、青空文庫コーパスを用いて自動構築する。本手法の極性辞書の自動構築においては変更可能なパラメータを3点設定した。実験参加者に複数の小説について場面ごとにポジティブ・ネガティブの観点でスコア付けしてもらい、このスコアを正解データとして、自動構築された極性辞書を用いた時系列データと類似度が高くなるパラメータを選択する。選択されたパラメータの辞書を用いて、青空文庫コーパスのそれぞれの小説を時系列データに変換し、特異値分解によって物語展開の特徴を得る。

#### 実験参加者による正解データ

自動構築の辞書を用いて読者の印象に沿った時系列データを作成できるかを確認するため、実験参加者によってスコアが付

与された正解データを準備した。以降、この実験を予備実験と呼ぶ。予備実験に用いた小説を表1に示す。

タイトル	単語数	場面数
押絵と旅する男	11,657	11
フランダースの犬	20,001	14
銀河鉄道の夜(角川文庫版)	24,992	17

実験参加者がスコア付けしやすいと考え、著名な作品を選択した。また、自動構築の辞書を用いた時系列データが単語数に依存しないことを確認するため、それぞれの小説の単語数が異なるように選択している。

予備実験の小説は章立てがないものがほとんどであるので、筆者が小説内の時間が変わる、あるいは主題が変わることを意識して場面を分けている。

実験参加者は1小説あたり大学生・大学院生の5名である。以下に実験の流れを示す。はじめに場面分けのない小説を一通り読んでもらい、次に場面分けを確認してもらう。実験参加者に、作品全体の大きな感情の流れを考慮して各場面について感情的にポジティブ（幸せ）かネガティブ（不幸せ）と感じるかを主観的に評価してもらう。各場面の評価は-5（とてもネガティブ）から5（とてもポジティブ）までの11段階でスコアをつけてもらった。

各場面について、実験参加者5人の評価の平均値を正解データとする。

#### 極性辞書の作成

極性辞書の作成においては、新聞記事を対象に複数の軸について印象辞書を構築した研究[8]を参考にする。文献[8]では、ある印象を持つ単語はその印象を持つ単語と共起しやすく、逆の印象を持つ単語とは共起しにくいという仮定に基づき、辞書を構築する。

極性辞書の構築では、ある単語がポジティブな単語あるいはネガティブな単語のどちらの周辺に出現しやすいかによってスコア付けを行う。新聞においては記事というまとまりがあるのに対し、小説では必ずしも章立てがあるとは限らない。それに加え、章立てがある場合でも作者によって基準が異なると考えられる。そこで、ある一定の単語数を一つの場面とみなし、その場面中にポジティブな単語とネガティブな単語のどちらがより多く出現するかで周辺単語のスコアを求める。どれほどの長さの文章を場面とみなすのが適しているかは不明であるため、場面として取り出す単語数を変更可能な一つのパラメータとする。その他、ポジティブな単語が一つだけ出現する場面より、ポジティブな単語が多数出現する場面の単語がよりポジティブな印象を持つと考えられる。また、出現回数が少ない単語を切り捨てることで確度の高い単語を残すことができると考えられる。

まとめると、極性辞書の自動構築において調整可能なパラメータは以下の3つである。

- 場面とみなす単語数（前後*i*単語）

• 1の場面においてポジティブあるいはネガティブな単語が何個出現するか ( $f$  個以上)

• ポジティブあるいはネガティブな単語の周辺単語として何回出現すれば辞書に取り入れるか ( $n$  回以上)

辞書の構築で用いるポジティブ・ネガティブな単語の集合として、語彙ネットワークを利用してポジティブ・ネガティブについての極性値を付与した単語感情極性対応表 [9] の一部を用いる。この極性対応表は-1 から+1 の実数値が単語に付与されている。ネガティブな (-1 に近い) 単語が多く含まれるため、極性値が-0.99 以下の単語をネガティブな単語として、このネガティブな単語と同数になるように極性値が高い順に単語を抜き出す。ポジティブな単語は極性値が1 から 0.969398 となる。それぞれ 1,723 語をポジティブ・ネガティブな単語として用いる。以降、これらの単語をポジティブ初期単語集合  $S_L$  とネガティブ初期単語集合  $S_R$  とする。

前述のパラメータと初期単語集合を用いた極性辞書の構築の流れは以下になる。

(1) 青空文庫コーパスを分かち書きする

(2) (1) の先頭の単語から見て、 $S_L$  が出現すれば、前後  $i$  単語を取得する

(3) (2) のうち、 $S_L$  の単語の数が  $S_R$  の単語の数より多い、かつ  $S_L$  の単語が  $f$  個以上出現すれば、(2) の単語群から名詞・動詞・形容詞を取り出す ( $N_L$  を 1 増やす)

(4) (3) で取り出した単語  $w$  それぞれについて  $S_L$  の周辺単語として数える ( $N_L(w)$  を 1 増やす)

(5) (2) に戻り単語を進める

前後  $i$  単語を取得する際に、小説をまたがることはないことに注意する。ここで、 $S_L$  が多く出現する場面の数、つまりポジティブな場面の数が  $N_L$  となり、ある単語  $w$  がポジティブな場面に含まれる数が  $N_L(w)$  となる。上記の 2 から 5 の処理を  $S_R$  についても行い、ネガティブな場面の数  $N_R$  と、ある単語  $w$  がネガティブな場面に含まれる数  $N_R(w)$  を求める。

$n$  回より少ない出現回数が少ない単語の切り捨てについては、 $N_L(w) + N_R(w) < n$  となる単語を除去を行う。

このとき、単語  $w$  のポジティブ場面・ネガティブ場面における条件付き確率は以下の通りになる。

$$P_L(w) = \frac{N_L(w)}{N_L}, P_R(w) = \frac{N_R(w)}{N_R} \quad (1)$$

これらを用いて、単語  $w$  の極性値  $v(w)$  を次の式で求める。

$$v(w) = \frac{P_L(w) * weight_L}{P_L(w) * weight_L + P_R(w) * weight_R} \quad (2)$$

$$weight_L = \log_{10} N_L, weight_R = \log_{10} N_R \quad (3)$$

単語  $w$  がポジティブな場面で多く登場すれば高い極性値に、逆にネガティブな場面で多く登場すれば低い極性値となる。

辞書の作成における 3 つのパラメータは、 $i$  は 10 から 70 まで 5 ずつ増やした値、 $f$  は 1 から 10 までの整数値、 $n$  は 0 から 20 まで 5 ずつ増やした値についてすべての組み合わせを試し、650 個の辞書を用意した。

#### 時系列データの作成

作成した極性辞書を用いて、小説から時系列データを求める。小説の頭から一定の単語数を取り出して辞書一致でスコアを算出し、小説の末尾に達するまでこの単語群をずらしながらスコアを得ることで、時系列データに変換する。以降、時系列データを求める際のスコアを算出する一定の単語数を window と呼ぶ。

window ごとのスコアの算出では、window 中の単語と一致する極性辞書の単語の極性値の平均を、その window のスコアとする。window の単語数は、すべての小説に対して同じ単語数で算出すると小説に占める window の割合が大幅に変わってしまうことを考え、(小説の単語数/5) としている。スライドさせる回数は、特異値分解のために揃える必要があり、今回は 100 回としている。つまり、時系列データの長さは 100 となる。window をスライドさせる幅は、(小説の単語数-window-1)/99 の計算式で求める。

#### 作成した極性辞書の評価

パラメータを変えて作成した極性辞書のうち、どの辞書がもっとも読者の印象を表すことができるのかを評価する。正解データとともにも類似度が高い時系列データを算出できる極性辞書を探るとともに、自動構築した辞書で読者の印象に合う時系列データを導出できることを示す。

類似度を求める前処理として、正解データは一つの小説につき場面の数のみスコアが付与されており、時系列データと近い長さに変換する必要がある。ここで、小説の進行度を  $x$  軸、スコアを  $y$  軸としてプロットすることを考える。辞書から計算された時系列データは  $x$  軸の 1 から 100 まで、それぞれ window のスコアをプロットすることでグラフを描画できる。正解データをこの時系列データに近い形に変換する。正解データは異なる長さの場面に対して付与されているため、小説全体の単語数に対する場面の単語数の割合に応じて正解データのスコアをプロットする位置を決定する。例えば、1 つめの場面の単語数が小説全体の 20% である場合、 $x = 10$  のときの  $y$  を、その場面の正解データのスコアとする。正解データのスコアをプロットする位置、つまり  $x$  の値を求めたあとに場面間に直線を引き、その他の  $x$  におけるスコアを補完する。

もう一つの前処理として、場面間のスコアを補完した正解データと辞書から求めた時系列データの正規化を行う。正解データの最大値と最小値に収まるように、時系列データのスケールを変更する。

以上の前処理を行った正解データと時系列データの類似度を測定する。予備実験では、実験参加者は分けられた場面に対してスコアを付与しており、その場面は小説全体の割合として数%から 10%程度である。それに対し時系列データは小説の 20%の単語からスコアを算出しているため、もし実験参加者が強くポジティブに感じる場面に対して時系列データが全く同じ位置で強くポジティブなスコアになるとは限らず、多少位置がずれる可能性がある。そこで、平均二乗誤差など位置がずれると類似度が低くなるような尺度は好ましくないと考え、動的時間伸縮法 (Dynamic Time Warping, 以降 DTW) [10] を用いて類似度を測る。DTW は 2 つの時系列データにおいて同じ時

点だけでなく他の時点とも距離を測り、すべての組み合わせを比較してもっとも短くなる距離を求める。この DTW を利用して、正解データとの距離がもっとも短い（類似度が高い）時系列データを算出できる極性辞書を求めた。すべての極性辞書から 3 作品の時系列データを計算し、正解データと時系列データの DTW の距離を求める。極性辞書ごとに 3 作品の距離を足しあわせ、もっとも値が小さい辞書を今回採用する辞書とする。

もっとも類似度が高くなったのは、パラメータが  $i = 60$ ,  $f = 8$ ,  $n = 20$  の極性辞書であった。この極性辞書で求めた時系列データと正解データのグラフを図 1 に示す。スコアが大きく上がる時点や下がる時点が多少ずれているものの、概ね時系列データが正解データに近い形になっていることがわかる。



図 1 もっとも類似度が高い時系列データと正解データのグラフ

今後、特異値分解などで使用する極性辞書は前述したパラメータの辞書である。

### 特異値分解による分析

時系列データから物語展開の特徴を求めるために、特異値分解を用いる。特異値分解の前処理として、それぞれの時系列データの平均が 0 になるよう正規化する。極性辞書の評価の際に読者の印象に沿った時系列データを作成できることが示された。10,000 以上の単語数である小説を特異値分解に用いる。青空文庫コーパスのうち 10,000 単語以上である 1,382 小説の時系列データ  $t_i$  を行とした行列  $A$  を用意し、特異値分解を適用して以下の式を求める。

$$A = U\Sigma V^T = WV^T \quad (4)$$

今回は行列  $A$  の時系列データに対する正規直交基底である  $V^T$  に着目し、 $V^T$  の行である右特異ベクトルを mode と呼ぶ。この mode が物語展開の特徴となる。また、 $U$  と  $\Sigma$  の積を  $W$  とし、mode に対する係数とみなす。行列  $W$  における  $i$  行は時系列データ  $t_i$  の係数となっており、行列  $W$  の  $i$  行と  $V^T$  の積をとると時系列データ  $t_i$  が得られる。さらに、行列  $W$  の  $i$  行の  $j$  番目の要素は、 $j$  番目の mode の重みとなっている。

行列  $\Sigma$  における対角成分は行列  $A$  の特異値であり、降順に並んでいる。そのため、mode のうち行列  $V^T$  において 1 行目にあたる mode の特徴が基本的には強く影響し、2 行目、3 行目と徐々に影響が弱くなる。作品によっては、1 番目より 2 番目以降の mode に対する係数が大きく、そのような mode の影響がもっとも強いこともある。ここで、行列  $V^T$  の  $i$  行目の mode を mode  $i$  と表す。

特異値分解によって得られた mode を 1 から 6 まで図 2 に示す。

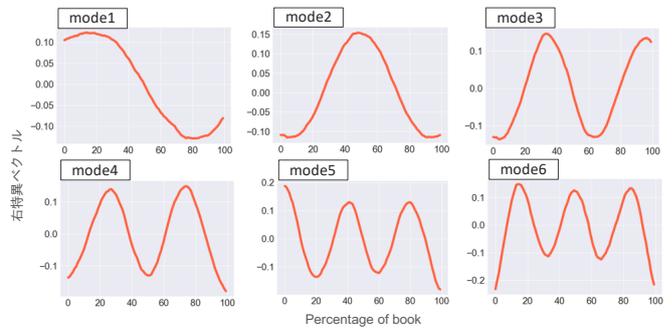


図 2 mode の例

mode 係数が負であれば、その mode は符号が反転し、グラフの上下が反転した形になる。mode1 の係数が負であれば、実際の特徴は  $-\text{mode1}$  であるとする。

行列  $A$  の  $i$  行目にあたる小説がどの mode に近いかが判別するためには、行列  $W$  の  $i$  行において絶対値が最大である要素を求める。その要素が  $j$  番目であれば mode  $j$  がその小説の特徴を表した mode である。もっともよく見られる特徴は mode1 であり、その次に  $-\text{mode2}$ ,  $-\text{mode1}$  と続き、最大の mode は 12 となった。

物語展開の特徴を、以上の方法で求めた mode とする。実際に場面を抽出する際には、特異値分解に用いた作品であれば絶対値が最大である mode 係数を求め、そうでない作品は各 mode との類似度をコサイン類似度などを使用して求めることで、どの特徴に当てはまるのかを見る。

#### 2.1.2 場面抽出の基準

抽出する場面は、ハッピーエンドの物語であればもっともポジティブな場面、反対にバッドエンドの物語であればネガティブな場面という方針で行う。特異値分解で求めた mode に対して、ラストのスコアが大きく上昇しているか下降しているかによってハッピーエンドかバッドエンドかを付与する。

入力された小説を極性辞書を用いて時系列データに変換し、どの mode に近いかを判別することでハッピーエンドとバッドエンドかを判断する。もしハッピーエンドの場合は時系列データのうち、スコアが最大値である箇所を中心として周囲の単語を収集する。バッドエンドであれば、スコアが最小値である箇所の周囲の単語を収集する。単語を抜き出す範囲は、小説全体の単語数の 20% の単語とする。

### 2.2 色の推薦手法

本文テキストからの色抽出では、word2vec により拡張した色彩データベースを用いて色を想起させる単語を特定する。単語の出現頻度を考慮して色のスコアを算出し、色相ごとにスコアを合計する。そして、合計したスコアが最上位となった色相において最もスコアの大きい色を抽出色とする。

#### 2.2.1 色彩データベース

色彩データベースは、日本カラーデザイン研究所により作成されたカラーイメージスケール [11] を利用して構築する。カラーイメージスケールでは、この代表色 130 色に対し、人々の感性を捉えるために有効な形容詞 180 語を定義している。ここ

では仮に、180語の形容詞を色彩形容詞と呼ぶ。また、色彩形容詞に対して代表色が使用される頻度を1から5の整数値で表してある。

代表色のRGB値、色彩形容詞の語幹、使用頻度の3属性を格納したものを色彩データベースとする。

### 2.2.2 色彩データベースの拡張

従来の色彩データベースは限られた形容詞しか色と対応づけられていないが、実際は形容詞以外の単語も色を連想させる可能性がある。形容詞に加え動詞と名詞も色を連想させ、意味が類似している単語は同一の色を想起させると仮定して、提案手法では2つの段階で色彩データベースの単語を拡張する。まず、日本語 WordNet [12] を用いて色彩形容詞の類義語を収集する。次に、色彩形容詞と類義語に類似した単語を、word2vec においてコサイン類似度が高い単語を抽出することで収集する。

word2vec とは、Mikolov *et al.* によって提唱されたニューラルネットワーク [13] を実装したツールである。テキストコーパスを入力すると、コーパスの共起関係に基づいて任意の次元のベクトルに圧縮する。これにより、単語同士の類似度をコサイン類似度によって表すことができる。学習において使用するモデルは Skip-gram を用いて、青空文庫コーパスを入力として word2vec のモデルの学習を行った。この際のコーパスは、形容詞、動詞、名詞のリストとなっている。Skip-gram において入力単語の前後5単語を予測するように学習を行い、100次元のベクトルを作成した。また、コーパス中で出現回数が20回に満たない単語は切り捨てて学習を行っている。

極性辞書には小説の印象に関わる単語が多く含まれると考え、色彩形容詞と類義語と意味が近い単語を極性辞書から収集した。具体的には、色彩形容詞と類義語に対してコサイン類似度が0.75以上の単語を類似度が高い単語として抽出を行った。

使用頻度の算出について述べる。日本語 WordNet で収集した単語には、収集する際に使用した色彩形容詞と等しいRGB値と使用頻度を与える。word2vec で収集した単語には、色彩形容詞と類語と等しいRGB値を与え、使用頻度にはその単語のコサイン類似度と使用頻度との積を与える。複数の色彩形容詞と類義語に対して辞書の1単語の類似度が高い場合には、重複した色ごとにそれらのコサイン類似度と使用頻度の積の平均をとる。

以上の処理によって、約23,300の単語、RGB値、使用頻度の対応づけが得られた。

### 2.2.3 色のスコアの算出方法及び色の抽出方法

色抽出においては本文テキストを形態素解析し、色彩データベースと一致した単語の使用頻度の値をスコアとしてRGB値ごとに足し合わせる。より精度の高い色抽出を行うために、否定後に続く単語を除去し、常用対数を用いた頻出単語の正規化を行う。ある色  $C_i$  に対する色のスコアを  $S_i$  とし、本文テキスト中に出現した色彩データベースと一致する単語集合を  $W_i$ 、そのうち否定語に続く語の集合を  $N_i$  とする。このとき、 $S_i$  は次の式によって得られる。

$$S_i = \sum_{w \in W_i \setminus N_i} (w_f \times \log_{10}(w_k + 1)) \quad (5)$$

算出した色のスコアを用いて、全130色のスコアを考慮して色の推薦を行う。まず、130色を色相を基準にして11のグループに分ける。有彩色(赤、青、黄など)の場合は1つのグループが12の色を持ち、無彩色(黒、白など)の10色を1つのグループにまとめる。グループごとに色のスコア  $S_i$  を合計し、その合計スコアを昇順に並び替え、上位グループの中で最大のスコアを持つ色を抽出色とする。

## 2.3 フォントの推薦手法

フォント推薦には感性イメージをベクトルとして表した感性ベクトルを用いる。PCにあるフォントファイルの各フォントと本文テキストの感性ベクトルを推定する。フォントについては、アンケート調査によって得たフォントと感性ベクトルの対を教師データとして学習を行い、作成された学習器を用いて感性ベクトルを推定する。本文テキストについては、極性辞書の作成方法で感性ベクトルの印象について辞書を作成し、テキスト中に出現する単語によって感性ベクトルを算出する。本文テキストの感性ベクトルとコサイン類似度が最も高い感性ベクトルを持ったフォントを推薦し、表紙に使用する。

### 2.3.1 感性ベクトル

感性ベクトルとは、感性イメージを表すベクトルであり、「暗い」「明るい」などの形容詞対の尺度を要素とする。形容詞対の尺度を感性値とすると感性ベクトルの次元数は感性値、つまり形容詞対の数になる。文献[5]を参考に選択した、表2に示す5つの形容詞対を感性ベクトルに用いる。

表2 次元削減後の感性ベクトルの形容詞対

暗い	明るい
穏やかな	激しい
下品な	上品な
つまらない	楽しい
冷たい	暖かい

### 2.3.2 フォントの感性ベクトルの推定方法

フォントの感性ベクトルを推定するために、本研究では感性ベクトルを推定する推定器を構成する。推定器は、アンケートにより収集した感性値をCNNで学習することにより構成する。実験に使用したフォントはPCにあらかじめインストールされているフォントや、フリーフォントなどから収集した日本語フォント計171個である。

実験参加者に171のフォントで表記された無意味文字列を提示し、フォントの印象をそれぞれの形容詞対について7段階SD法で評価してもらった。

実験参加者は大学生・大学院生の合計32名であり、171のフォントを4グループに分け1フォント当たり8名の回答を得た。各フォントの感性値の平均を算出して、フォントの感性ベクトルとして用いる。

アンケートに用いたフォントは数が限られており、それ以外の日本語フォントも無数に存在する。そこで、未知の日本語フォントであっても感性ベクトルの推定が行えるように、アンケートで得たデータを教師データとしてCNNによって学習を

行う。フォントについては、アンケートに用いたフォントで描画した文字の画像を用いる。フォントの特徴が現れやすいと考え「永」「あ」という2文字を使用して、90ポイントで115ピクセル×210ピクセルの白背景の画像に描画して作成した。この画像を1ピクセルずつグレースケールに応じて0から255の数値に変換した数値列を学習の際に入力する。

フォントの画像から感性ベクトルを推定する学習には、多層のニューラルネットワークを用いたDeep Learningの一種であり、主に画像認識で用いられているCNN(Convolutional Neural Network)を使用した。具体的な学習のモデルは、CNNを用いてフォントの自動識別を行っている[14]を参考にして作成した。畳み込み層とプーリング層のペアを2回繰り返して、3回畳み込んだ後に全結合させ、シグモイド関数を通して1,000のユニットに出力し、線形出力関数を通して感性ベクトルの次元数のユニットに出力する。フィルタの数を30、畳み込み層とプーリング層の大きさは共に3×3ピクセルとする。学習時の損失関数は平均二乗誤差、最適化関数はRMSpropを用い、教師データの感性ベクトルとフォントの画像から予測した感性ベクトルの平均二乗誤差が少なくなるように学習を進める。このモデルを深層学習ライブラリであるKerasで実装し、全171のデータのうち7割を訓練データ、3割を評価データとして50回(50 epoch)の学習を行った。50 epoch目の訓練データのlossは0.42244、評価データのlossは0.50658であった。学習済みのモデルの重みを用いることで、フォントの画像を入力するとフォントの特徴に応じた感性ベクトルを得ることができる。

本文の感性ベクトルの感性値は形容詞対についてどちらでもない場合に0となるので、1から7の値をとるフォントの感性値 $v$ を、 $(v-4)/6$ の式で変換する。

## 2.4 本文の感性ベクトルの推定方法

本文テキストの感性ベクトルについては、極性辞書と同様に作成した辞書を用いて、抽出場面に出現する単語から感性値を求める。この辞書を、以降は感性辞書と呼ぶ。

感性辞書の作成方法について説明する。形容詞対それぞれの類義語を日本語WordNetを用いて収集し、初期単語集合とする。感性辞書の作成において、一つの場面とみなす前後の単語数と、出現回数による単語の切り捨ては採用された極性辞書のパラメータに合わせる。ただし、初期単語集合の大きさが極性辞書のものとは大きく異なるため、一つの場面に2個以上の初期単語集合に属する単語が出現する場面を取り出す。それぞれの形容詞対について感性辞書を作成する。

感性値は、抽出場面の単語と一致する感性辞書のスコアの平均値を求め、その値から感性辞書全体のスコアの平均を引くことで求める。

## 2.5 象徴物の推薦手法

象徴物の推薦では、抽出場面と青空文庫コーパスの小説のtf-idfを用いて特徴語となる名詞を抽出する。

tf-idfとは、Term Frequency(単語の出現頻度)とInverse Document Frequency(逆文書頻度)の2つの指標の積をとつ

たものである。tf-idfを用いることで、ある文書には多く出現するが、その他の文書にはあまり出現しないような単語、つまり文書の特徴となるような単語を求めることができる。

抽出場面と青空文庫コーパスの4,031小説をそれぞれ文書とみなして、tf-idfを計算する。小説にしばしば登場するような名詞や1小説にしか登場しない名詞を取り除くため、全小説のうち10%以上に含まれる名詞と全体の出現頻度が20回未満である名詞は削除している。抽出場面においてtf-idfが高い名詞を、象徴物として推薦する。

## 3 評価実験

本章では、提案手法により推薦された色とフォント、象徴物が、小説の印象を反映しているかを検証する評価実験について述べる。評価方法としては、提案手法で推薦された5つの色・フォント・象徴物を含む選択肢を提示して、推薦されたものが選択されるかを検証した。

本実験は、クラウドソーシングを提供するLancers<sup>3</sup>にて行い、青空文庫の2つの小説について、1小説あたり8人の実験参加者の評価をしてもらった。実験参加者の平均年齢は43.3歳である。

実験の流れは以下の通りである。

(1) 青空文庫の小説を読んでもらう。青空文庫の作品リンクだけでなく、Kindle小説や縦書きで読めるサービスも提示して、好きな環境で読んでもらう。

(2) 実際に読んでいるのか確認するために、小説の内容に関する質問に答えてもらう。

(3) 小説の中で特に印象に残った場面について印象に合う色・フォント・象徴物を選択してもらい、選択肢のうち小説の印象に近いものを5位まで選択してもらい、それぞれ簡単に理由を記入してもらう。

以降では、評価実験で用いる小説と、色・フォント・象徴物の選択肢の作成方法、実験結果、考察についてそれぞれ説明する。

### 3.1 評価実験に用いる小説

評価実験に用いる小説は、青空文庫で公開されているアクセスランキング<sup>4</sup>の中から2作品を選択した。作品名とその単語数を表3に示す。

表3 評価実験に使用する小説

タイトル	著者	単語数
アッシャー家の崩壊	エドガー・アラン・ポー	11,152
戯作三昧	芥川竜之介	13,271

評価実験に使用する小説の時系列データと、そのデータと類似度をもっとも高いmodeのグラフを図3を示す。それぞれの小説の物語展開の特徴となるのがこのmodeであり、『アッシャー家の崩壊』とはバッドエンド、『戯作三昧』はハッピーエンドと判断する。前者は時系列データのスコアがもっとも低い

3: <https://www.lancers.jp/>

4: [https://www.aozora.gr.jp/access\\_ranking/2017\\_xhtml.html](https://www.aozora.gr.jp/access_ranking/2017_xhtml.html)

時点を中心として小説のうち 20%の単語数を取り出し、後者はスコアがもっとも高い時点の周辺単語を取り出す。その結果、『戯作三昧』は小説の末尾の 20%、『アッシャー家の崩壊』は小説の 54%から 74%にあたる部分が抜き出される。

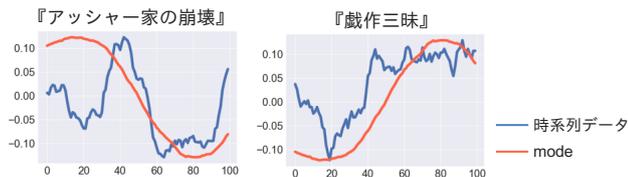


図 3 評価実験の小説の時系列データと mode のグラフ

### 3.2 色・フォント・象徴物の選択肢の作成

ここでは、評価実験における色・フォント・象徴物の選択肢の作成方法についてそれぞれ説明する。なお、上位に推薦されたもの以外の選択肢をダミーの選択肢と呼ぶ。

#### 3.2.1 色の選択肢の作成

カラーイメージスケール [11] によると、有彩色（赤，青，黄など）は色相とトーンの違いを持ち、それぞれの色相には明暗，派手，地味というように共通した色の調子（トーン）がある。トーンは明暗の違いを生む明度と派手・地味の違いに関わる彩度との相互の関係のもとに成り立つ。実験では色相と色調をもとに色の選択を行う。提案手法で色のスコアを算出し色相ごとのスコアを昇順して上位 5 位までの色相グループを代表グループとし、それぞれのグループで最大スコアを持つ色を推薦色とする。5 つの抽出色の属する色相と異なる色相の色で、色抽出の際に色ごとのスコアを昇順にし中央より下位になる色をダミーの選択肢とする。ダミーの選択肢については抽出色と異なる色調になるよう選ぶ。

#### 3.2.2 フォントの選択肢の作成

フォントについては、アンケートに用いた 171 のフォントから書体が偏らないように選んだ 23 フォントと新しく収集した 9 フォントの合わせて 32 フォントを推薦の対象とした。提案手法でのフォント推薦の結果、コサイン類似度を昇順にした際の上位 5 フォントを推薦フォントとする。明朝体やゴシック体のような書体を基準にして、コサイン類似度のランキングが下位のフォントからダミーの 5 つの選択肢が似ないように書体が異なるものを選ぶ。実験参加者には 10 のフォントを用いて、それぞれテキストのタイトルを描画して提示する。

#### 3.2.3 象徴物の選択肢の作成

抽出場面で tf-idf の値が高い順に 5 位までの名詞を提案手法で推薦された象徴物とする。ダミーの選択肢として、抽出場面の代わりに小説全体を文書と見なした場合に tf-idf の値が高い 5 位までの名詞を用いる。小説全体を用いて推薦された名詞が提案手法で推薦された名詞と重複した場合には、象徴物の選択肢の数が 10 個未満となる。

本実験では、象徴物を写真として表紙に使用する場合に小説の印象を表現できるかを評価する。写真の収集では、Google 画像検索を用いる。検索オプションで、画像の種類を写真、ライ

センスを「変更後の再使用が許可された画像」に指定して、名詞を検索して象徴物の写真を得る。検索結果で 1 位の写真から象徴物について読み取りにくい場合を考え、上位 3 位までの写真を一つのグループとして実験参加者に提示した。評価実験では、名詞の記載はせずに写真のみを提示している。

### 3.3 実験結果

評価実験に用いた 2 つの小説に対する提案手法による推薦色・フォント・象徴物を表 4 に示す。

表 4 提案手法で推薦された色・フォント・象徴物

小説	順位	色	フォント	象徴物
アッシャー家の崩壊	1	0,0,0 (黒)	あおぞら明朝 Heavy	死者
	2	255,217,0 (黄)	HOKKORI フォント	ゴシック
	3	176,143,119 (茶ねずみ)	ロゴたいぶゴシックコンデンスド	夢魔
	4	0,38,102 (紺色)	Kirieji (切絵字)	息吹
	5	217,253,255 (白群)	正調祥南行書体 EX	ランプ
戯作三昧	1	255,255,255 (白)	ロゴたいぶゴシックコンデンスド	蟋蟀
	1	255,242,191 (象牙色)	あおぞら明朝 Heavy	戯作
	3	206,181,159 (とのこ色)	IPAex ゴシック	祖父
	4	0,38,102 (紺色)	うずらフォント	原稿
	5	213,255,236 (うすあさぎ)	クレー	茶の間

評価実験において、実験参加者が選んだ選択肢の中で推薦色・フォント・象徴物が含まれた個数の平均を回答平均値として表 5 に示す。また、実験参加者に提示した選択肢の数を ( ) の中に示す。

表 5 評価実験の結果

タイトル	色の回答平均値	フォントの回答平均値	象徴物の回答平均値
アッシャー家の崩壊	3.375 (10)	3.875 (10)	2.75 (10)
戯作三昧	2.625 (10)	3 (10)	3.625 (9)

### 3.4 考察

評価実験の結果から、提案手法ではフォントについては小説の印象に沿う場合が多いが、色については読者の印象に沿う場合と沿わない場合に分かれることが示された。また、象徴物については小説によっては抽出場面に限った方が小説全体よりも印象に合ったものを推薦できることがわかった。

『アッシャー家の崩壊』については、ほとんどの実験参加者は小説に対して不気味さや怖さ、悲しさ、冷たさを感じ、暗い夜の場が印象に残ったと回答している。そのため、フォントにおいては文字自体が崩れているように見える切絵字フォント

や HOKKORI フォントが選択され、推薦フォントが選択されやすい結果となった。色の評価では、推薦順位が1位の黒をすべての実験参加者が選択しており、3位の茶ねずみ色を7人、4位の紺色を6人が選択していた。推薦色のうち、実験参加者にほとんど選択されなかったのは白群色（うすい水色）である。この色は、抽出場面のうち「静か」や「明らか」、「確実」などの単語と一致して抽出されたが、実験参加者は暗い印象を感じる明度が低い色を選択しており、明度の高い白群色は選択されなかったと考えられる。色彩データベースで使用したカラーイメージスケールでは、色に紐付いた形容詞が基本的にポジティブな単語が多く、ネガティブな場面でも明るい色が推薦されることがあることがわかった。象徴物については、小説全体から推薦した象徴物が多く選択されており回答平均値が低い結果となった。理由として、提案手法で1位に推薦された「死者」の写真がすべて「死者の日」という祭りに関する写真となっており鮮やかに着色された骸骨の人形などの写真が含まれるため、作品に暗いイメージを持った実験参加者は選択しなかったことが考えられる。また、推薦順位が4位の「息吹」についても、Web ページでコメントにこの単語が含まれた植物の写真となっているため選択されなかった。ただし、推薦2位の「ゴシック」は6人の実験参加者が1位に選択しており、印象に近いものが推薦できたと考えられる。

『戯作三昧』では、主人公が悩み苦しんでいる場面や、最後にその悩みが解消される場面など実験参加者によって印象に残った場面が異なっていた。このことが、色とフォントの選択において実験参加者による回答のばらつきが大きいという結果に影響していると考えられる。色において、主人公が老人であることから渋い色として明度の低い色を選択する実験参加者がいたが、推薦色以外でもそのような色が含まれていたため回答平均値が低くなっている。推薦色の中でもっとも選択されていたのは推薦順位が4位の紺色であり、理由としては主人公が苦悩する場面や戯作を執筆する暗い部屋の印象が残っていることが挙げられていた。ダミーの色は、それぞれ3、4人程度に選択されていた。フォントについても色と同じく、古風に感じることを理由にフォントを選択する実験参加者もいたが、実験参加者によって古風に感じるフォントが異なっていた。また、全体としてフォントのシンプルさなどが理由として多く挙げられており、実験参加者が小説からフォントのイメージを掴みにくかった可能性がある。象徴物については高い回答平均値となったが、提案手法で1位に推薦される「蟋蟀（こおろぎ）」を評価実験で1位に選択した実験参加者はおらず、5人の実験参加者が推薦4位の「原稿」を1位に選択する結果となった。「蟋蟀」は抽出場面である小説のラストで出現する名詞であり、参加者はその場面の印象が薄かったと考えられる。

## 4 む す び

本研究では、書籍内容の印象を反映した表紙を生成することを目的として、小説データから読者の印象に残ると思われる場面を抽出して、その場面から色・フォント・象徴物を推薦す

る手法を提案した。評価実験の結果、抽出場面が読者の印象に残った場面と合致すれば、概ね印象に合う推薦が行えることがわかった。色彩データベース上にポジティブな単語が多いため印象に合う色の推薦が難しい場合があること、読者による小説の印象が共通している場合は印象に沿うフォントを概ね推薦できるがそうでない場合は難しいことが示された。また、象徴物については小説によっては抽出場面に限った方が印象に合ったものを推薦できることがわかった。

今後の課題として、色の推薦において色彩データベースにより幅広い色と単語の紐付けを導入すること、推薦された象徴物を正しく伝えられる提示方法を考案すること、青空文庫以外の小説を用いて評価実験を行うことが考えられる。

## 文 献

- [1] International Digital Publishing Forum, “EPUB,” <http://idpf.org/epub/>, (参照 2018-12-28).
- [2] 飯田拓也, 梶山朋子, 大内紀知, 越前功, “読者の印象を反映させた書籍表紙画像生成のための色抽出手法,” 電子情報通信学会論文誌, vol. J97-D, no.1, pp.75-84, 2014.
- [3] H. Kawaguchi, and N. Suzuki, “Recommending Colors and Fonts for Cover Page of EPUB Book,” the 18th ACM Symposium on Document Engineering (DocEng 2018), 4p.
- [4] 蒔野充裕, 梶山朋子, 大内紀知, 越前功, “読者の印象を反映させた書籍表紙画像生成のための色彩と象徴物の抽出,” 電子情報通信学会技術研究報告 = IEICE technical report : 信学技報, vol.114, no.68, pp.89-94, 2014.
- [5] 飯場咲紀, 土斐崎龍一, 坂本真樹, “テキストの感性イメージを反映した色彩・フォント推薦,” 日本バーチャルリアリティ学会論文誌, vol.18, no.3, pp.217-226, 2013.
- [6] A. J. Reagan, L. Mitchell, D. Kiley, C. M. Danforth, and P. S. Dodds, “The emotional arcs of stories are dominated by six basic shapes,” EPJ Data Science, vol.5, no.1, 31p, 2016.
- [7] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss, and C. M. Danforth, “Temporal Patterns of Happiness and Information in a Global Social Network: Hedonometrics and Twitter,” PLOS ONE, vol.6, no.12, 2011.
- [8] T. Kumamoto, and K. Tanaka, “Proposal of Impression Mining from News Articles,” Knowledge-Based Intelligent Information and Engineering Systems, pp.901-910, 2005.
- [9] 高村大也, 乾孝司, 奥村学, “スピンモデルによる単語の感情極性抽出,” 情報処理学会論文誌ジャーナル, vol.47, no.2, pp.627-637, 2006.
- [10] H. Sakoe, and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.26, no.1, pp.43-49, 1978.
- [11] 小林重順, カラーイメージスケール改訂版, 講談社, 2001.
- [12] H. Isahara, F. Bond, K. Uchimoto, M. Utiyama, and K. Kanzaki, “Development of the Japanese WordNet,” In LREC-2008.
- [13] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” In Advances in neural information processing systems, pp.3111-3119, 2013.
- [14] Z. Wang, J. Yang, H. Jin, E. Shechtman, A. Agarwala, J. Brandt, and T. S. Huang, “DeepFont: Identify Your Font from An Image,” ACM International Conference on Multimedia (ACM MM), pp.451-459, 2015.