

# 読みやすさと視覚的バランスを両立した改行・改頁位置の推定

大峠 和基<sup>†</sup> 佐藤 哲司<sup>††</sup>

<sup>†</sup> 筑波大学大学院 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>{otao,satoh}@ce.slis.tsukuba.ac.jp

**あらまし** 雑誌・広告・動画などのビジュアルデザインにおいてテキスト要素を配置する際、表示できる領域が限られるためにテキスト要素に対して適切な位置に改行・改頁を挿入する必要がある。この際、読者に負担なくテキスト要素を読ませる可読性と、デザイン全体のバランスを崩さない視覚的バランスとの2点を満たす必要があり、改行・改頁を挿入する工程は属人性が伴う。本研究ではこれら2つの要件を満たす改行・改頁位置の推定手法を提案する。具体的には、文字数や行数を制約として与え、可読性と視覚的バランスを考慮した評価関数を最適化問題として定式化する。のべ約400名の実験参加者による評価実験を通じて、ベースライン手法と比較して提案手法が可読性と視覚的バランスを向上できることを明らかにした。

**キーワード** ビジュアルデザイン、自動改行、可読性、視覚的バランス、テキストレイアウト

## 1 はじめに

ビジュアルデザインとは、イラストや画像、CGなどの視覚的な表現を使って顧客や大衆にメッセージを伝達することを目的としたクリエイティブの総称であり、雑誌・広告・ポスターなどが代表的な例である。その定義に基づけば企業のWebサイトや、対外発表と内部資料として用いられるプレゼンテーション用のスライドなどもビジュアルデザインに含まれる。特にデジタルデバイスの普及した現代においては、ビジュアルデザイン分野におけるクリエイティブの重要性は日々増加しており、自動でクリエイティブを生成する技術が求められている。ビジュアルデザインは、主題となる被写体を示すグラフィクス要素と、キャッチコピーや訴求文を示すテキスト要素から成り立っていると捉えることができ、グラフィクス要素とテキスト要素のレイアウトを推定する手法[1]、テキスト要素の色を推定する手法[2]、テキスト要素の大きさを最適化する手法[3]などが提案されてきた。

ここで、ビジュアルデザインにおけるテキスト要素はキャッチコピーなどのように顧客や大衆にメッセージを伝達する目的があるため、グラフィクス要素とテキスト要素の配置関係を考慮する際、表示する文字列をどのように分割・改行するかが重要である。例として、図1に「今日は動物園に行ってきました」というテキスト要素を異なる改行パターンで配置したときの様子を示す。図1(a)は文字列のうち最も大きな意味合いの区切りで改行を加えているため、可読性は高いといえるがグラフィクス要素として目立たせたい被写体(キリン)に被っている。一方で図1(b)はある程度可読性を担保しつつ行ごとの文字数が均等に近づくよう改行が加わっているため、テキスト要素の横幅が長くなりすぎず被写体に被っていない。また、テキスト要素がコンパクトな領域に収まっていることから、デザイン全体として見たときにバランスが良いと言える。このようにテキ

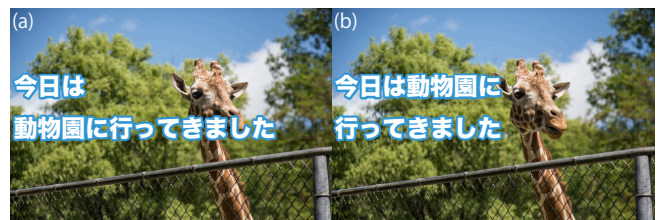


図1 「今日は動物園に行ってきました」という文字列に対して、異なる改行パターンをグラフィクス要素の上にオーバーレイして表示したときの様子。

スト要素に改行を付与する際には、可読性が高くなる位置に加えるという共通の認識に加えて、目立たせたい被写体に被らないことやテキスト要素自体の視覚的バランスが崩れないようにする必要がある。

改行位置の推定研究は、発話の音声波形を文字列に変換する音声認識の分野において多くの研究がなされてきた[4][5][6]。これらの研究では、改行アノテーション付きの文章を学習した推定器を用いて、与えられた文章に対して改行が加わる確率が高い位置に改行を挿入している。しかしながら、改行による可読性に焦点が当てられており改行が及ぼす視覚的印象については議論されていない。上述したようにビジュアルデザインにおいてはテキスト要素そのものがデザインを構成する一要素であるため、可読性に加えて視覚的バランスを保つことが重要であり、従来の改行位置推定の手法とは考慮すべき項目が異なる。

以上述べたように、可読性と視覚的バランスの双方を考慮したビジュアルデザインのためのテキスト要素への改行・改頁の自動挿入手法が求められる。本研究では、限られた領域内に文章を表示する際に、可動性と視覚的バランスを考慮した改行・改頁位置の推定手法を提案する。具体的には、可読性と視覚的バランスを考慮した評価関数を設計し、組み合わせ最適化として定式化することで改行・改頁パターンを選択する。その際、文

字数・行数の制約を与えることができ、これはシステムやユーザが自由に変更することができる。

従来手法に対する提案手法の大きな差分は以下の2点である。

- (1) 可読性と視覚的バランスを両立した改行・改頁位置の推定。
- (2) 行あたりの文字数や行数の制約をシステムやユーザが自由に変更可能。

提案手法により、従来人手で行われていたテキスト要素への改行・改頁挿入を自動化することを目指す。提案手法は雑誌や広告などのテキスト要素の自動レイアウトに応用できるほか、ブラウザに表示される文字の自動改行、ゲームやチャットアプリ上のテキストウィンドウ、表示できる領域が限られた電光掲示板など様々な範囲に応用ができる。

## 2 関連研究

### 2.1 ビジュアルデザインにおける自動レイアウト

ビジュアルデザインの分野では、テキスト要素とグラフィクス要素の配置関係を推定する研究がなされてきた。特に雑誌やポスターは、ある程度その配置関係に決まった型があるという前提のもとで、テキスト要素とグラフィクス要素の配置関係を予め定めたテンプレートを用いる手法が提案されている [1] [2] [3]。Jahani らは顕著性マップを基にテキスト要素が被らないテンプレートを選択することで、テキスト要素とグラフィクス要素の配置関係を推定している [1]。また、推定された描画領域内にテキスト要素が収まるよう、テキスト要素の大きさの最適化にも取り組んでいる。Yang らはテキスト要素を雑誌名・ヘッドライン・見出しなどの細かい分類に分けた上でそれぞれ配置を決定している [2]。更に、グラフィクス要素の特徴を重要視してテキスト要素を配置するべく、顕著性マップに加えて顔検出や顔方向検出などを組み合わせることでテンプレートの推定に活用している。また、グラフィクス要素の色彩に調和するように、ハーモニックカラー [7] の考えを用いたテキスト要素の色推定を行っている。Donovan らはテンプレートを使わずビジュアルデザインの作例から直接レイアウトを学習する手法として、既存の作例に対して人手でレイアウトを表現するアノテーションデータを作成し、機械学習を使って配置関係を推定するモデルを提案した [3]。

### 2.2 改行位置の推定

自動で読点・改行を挿入する手法は、主に音声認識の分野において音声認識の結果として得られたテキストを整形するために発展してきた [8]。英語などのスペースで区切られる言語と異なり、日本語は読点の位置で読みやすさが変化する。古典的な手法として、鈴木らは既存の文章から句読点が置かれやすい品詞を分析し、形態素解析と独自に作ったルールで句読点を挿入した [9]。しかしながら、読点位置を決定するための共通した一意の規則は存在しないため、ルールベースで行う手法はすべてのパターンを網羅することができない。そこで近年では、村田らにより日本語コーパスから学習した確率モデルで読点を挿入

する機械学習的アプローチが提案されている [10]。彼らは推定のために、形態素・係り受け・節境界などの特徴量を用いた。ここで、同一のテキストであったとしても、アノータタによって読点を挿入する位置は異なりアノテーションデータに属人性が混入することを避けるのは難しい。秋田らは読点挿入の属人性の影響を排除するべく、複数人のアノテーションデータを組み合わせることで推定精度を向上させている [5]。秋田らの研究では、学習したモデルに対して 78.4% の適合率で読点を推定できることが報告されている。

音声認識結果を使った応用として、講演での発表者の発話をリアルタイムにスクリーンに表示することが考えられてきた [11]。この場合、得られた日本語の文章に対して句読点の代わりに改行を挿入することによって読みやすさの向上を狙う。文章に対して自動で改行を挿入する際も読点と同様に一意の規則は存在しないため、アノテーションデータを用意して機械学習モデルを学習させる。村田らは係り受けや節境界、音声波形のうち無音部分の時間を表すポーズなどの特徴量をもとに、統計的モデルを使って 80.2% の適合率で改行位置を推定できるとしている。近年では、Recurrent neural network (RNN) を使って、アノテーションデータに対する推定精度を高める試みがなされている [12]。

大野らはリアルタイムに改行位置を推定する際の手法の違いによる、手法のリアルタイム性と改行結果の質のトレードオフを比較している [6]。彼らの実験によると、文節ごとに改行を挿入する手法が最も定性的評価が高くなることを明らかにしている。句読点・改行推定のためには形態素解析や係り受け解析の結果などの様々な特徴量を入力とするが、それらの特徴量を計算するための実行時間とその特徴量が改行結果の質に対してどの程度影響しているかという点は今後考慮すべき項目であると考えられる。

## 3 提案手法

### 3.1 ビジュアルデザインのための改行・改頁

1 章で述べたように、ビジュアルデザインのための改行・改頁では、可読性と視覚的バランスを両立させる必要がある。本研究では可読性と視覚的バランスを以下のように定義する。

- 可読性が高い: 意味合いの塊で区切ったとき、できるだけ大きな塊の境界に改行が位置する。
- 視覚的バランスが良い: 行やページの間で文字数の差が小さい。

例えば「今日は動物園に行ってきました」という文字列を行数 2 行、1 行あたりの文字数は問わないという制約の下で改行を加えることを考える。このとき、総文字数は 12 文字であるので考えられる改行パターンは 11 通りである。最も可読性が高くなる例は「今日は/動物園に行ってきました」である。この場合、可読性は高いが、文字数を考慮にいけないため視覚的バランスは低いと言える。最もバランスが良いのは「今日は動物園/に行ってきました」である。これを実現するには、1 行目

表 1 形態素解析の結果.			
$M_i$	形態素	品詞	品詞細分類
0	今日	名詞	時相名詞
1	は	助詞	副助詞
2	動物	名詞	普通名詞
3	園	名詞	普通名詞
4	に	助詞	格助詞
5	行って	動詞	
6	き	接尾辞	動詞性接尾辞
7	ました	接尾辞	動詞性接尾辞

表 2 文節生成の結果.	
$C_j$	文節
0	今日は
1	動物
2	園に
3	行ってきました

と 2 行目の文字数が近くなる位置に改行を加えれば良い。この場合、視覚的バランスは良いが、意味の塊とは無関係に文字列が区切られてしまうため可読性は低くなる。可読性と視覚的バランスを両立する改行パターンとして、1 文字ずらした位置の「今日は動物園に/行ってきました」が挙げられる。本研究では、このようなある程度可読性を担保しつつも視覚的バランスの取れた改行位置を推定することを目指す。

また、表示するテキスト要素の領域はデザインによって変化する。そこで、自動改行の際にテキスト要素の文字数や行数を外部からインタラクティブに設定できるように設計する。このように行数や行内文字数の上限を外部から設定可能とすることで、背景となるグラフィクス要素の重要な被写体に重なることなくテキスト要素を配置できるようにする。

### 3.2 改行可能性推定

改行位置を推定するためには与えられた文字列に対して 1 文字ずつ改行可能性  $p$  を推定する。従来研究では、入力として文字列の形態素や係り受け情報、出力として人手でアノテーションされた改行位置を与えることで改行位置を学習・推定している。しかしながら、行数や文字数の制限が動的に変化し、その中で視覚的バランスを考慮する必要のある本研究ではアノテーションは困難である。そこで本研究では簡易に改行可能性  $p$  を推定するルールベース手法を使う。

日本語における改行の候補となる位置を考える際に、最初の手がかりとなるのが形態素である。形態素とは、文字の塊として意味を持つ最小限の単位である。「今日は動物園に行ってきました」という文字列が与えられたとき、形態素解析の結果として表 1 が得られる。これらの形態素を  $M_i (0 \leq i < M)$  と表現する。ここで、 $M_0$  の形態素は今日、品詞は名詞、品詞細分類は時相名詞である。本研究では形態素解析ツールとして JUMAN++ [13] を用いた。

次に、より大きな粒度でグルーピングする際概念として文節を使うことができる。文節とは、言葉を細かく区切った際に不自然にならない最小の単位で、形態素より大きな粒度となる。文節を構成するルールとして、例外は多く存在するものの、基本的には 1 個以上の自立語に 0 個以上の付属語もしくは接辞が接続されたものと定義される<sup>1</sup>。付属語とは、助動詞と助詞の

ことで「は」「に」「が」「です」などが該当する。自立語とは、名詞や動詞、形容詞などの付属語以外の品詞で、「今日」「動物」「行って」などの例が該当する。そこで、形態素を分節にまとめるために、各形態素  $M_i$  に対して以下の条件を満たす位置を分節の区切りとする。

- $M_i$  が接続詞ではない
- かつ、 $M_{i+1}$  が付属語もしくは接尾辞ではない

これを表 1 に適用した結果として、表 2 が得られる。これらの文節を  $C_j (0 \leq j < C)$  で表すこととする。

ここで、この手法により求めた文節は正確な文節ではないことに注意する。「動物園に」という文字列に着目すると「動物」「園に」という 2 つの文節で表現されている。主に哺乳類を表現する名詞である「動物」とある一定の目的のために区分けされた領域を表現する名詞である「園」は別々の形態素であり、「動物園」で 1 つの複合名詞として扱うことが一般的である。正しく分節を判定するために、Wikipedia などの外部辞書を参照するアプローチも考えられるが、このような固有名詞は文字列として長くなりがちが多く、改行パターンを列挙する際に 1 行あたりの文字数の制限を超えてしまうケースがあるため、本研究では 1 つの文節で扱わず「動物」「園に」と複数の文節として扱うことにする。

最後に改行可能性  $p$  を求める。例えば「～です」「～ます」「～ました」のように、日本語として文章が終了することを示す助詞の直後では強く改行を加えることが望ましい。また、「～に」「～が」のように、現在の境界の直前が付属語で終わっている場合は積極的に改行を挿入したい。一方で、名詞などで終わっている場合は区切りとして弱いことから、改行は避けたい。これらを踏まえて各文字に対して以下のルールを適用して改行可能性  $p$  を決定する。

- 対象の文字が  $C_j$  の境界でない場合、 $p = 0.0$  とする。
- 対象の文字が  $C_j$  の境界で、その  $C_j$  の最後の形態素が付属語と接尾辞ではない場合、 $p = 0.2$  とする。
- 対象の文字が  $C_j$  の境界で、その  $C_j$  の最後の形態素が付属語もしくは接尾辞の場合、 $p = 0.5$  とする
- 対象の文字が  $C_j$  の境界で、その  $C_j$  の最後の形態素が動詞性接尾辞か終助詞の場合、 $p = 1.0$  とする

動詞性接尾辞には「れる」「させる」などが該当する。終助詞とは、文の終わりに位置する命令・疑問・願望などを意味する助詞のことで、「かしら」「ぞ」などが該当する。

このルールを「今日は動物園に行ってきました」に適用すると、入力した文章の長さと同じ大きさを持つ配列  $[0.0, 0.0, 0.5,$

1 : <https://ja.wikipedia.org/wiki/%E6%96%87%E7%AF%80> (最終閲覧: 2022 年 2 月)

0.0, 0.2, 0.0, 0.5, 0.0, 0.0, 0.0, 0.0, 0.0, 1.0] が改行可能性  $p$  として得られる。

### 3.3 評価関数

改行可能性  $p$  を用いて改行位置を決定するために、単なるしきい値処理ではなく視覚的バランスを盛り込んだ評価関数を設計し、評価値が最も良い改行パターンを採用する。また、動画の字幕やチャットアプリのテキストウィンドウなどのアプリケーションでは改行だけでなく改ページも必要であるため考慮に入れる。

例えば、「美しい写真」という文字列が与えられたとする。ここで改行可能性  $p$  の配列として、[0.0, 0.0, 0.5, 0.0, 0.0, 0.2] が得られる。全ての文字の間に改行・改ページが挿入されると改行・改ページのパターンは 81 通りとなる。しかし、 $p = 0.0$  の位置は改行・改ページを挿入されることがないことから、この場合の可能な改行・改ページパターンは、「美しい」と「写真」を単に結合しただけの「美しい写真」、改行記号を挿入した「美しい / 写真」、改ページを加えた「美しい」「写真」の 3 通りとなる。

すなわち、文節として得られた  $C_j$  を使って、以下のようなパターンを考慮すれば良い。

- (1)  $C_j + C_{j+1}$
- (2)  $C_j + \text{linefeed} + C_{j+1}$
- (3)  $C_j + \text{pagefeed} + C_{j+1}$

ここで+記号は文字列の結合を示し、*linefeed* は改行記号の挿入、*pagefeed* は改ページ記号の挿入処理に置き換えられる。文節の数  $C$  に対して、改行・改ページパターンは  $3^{(C-1)}$  通りとなる。

続いて、それぞれの改行・改ページパターンに対して評価値を求める。可読性については、改行・改ページを挿入した直前の文字の改行可能性  $p$  が高いほどよい。そこで、改行・改ページが挿入される直前の文字の改行可能性  $p$  の合計値を  $S_{sum}$  とする。視覚的バランスについては、可読性が高い位置にひたすら改行を入れることでページが分散しないよう可能な限り少ないページ数に収めつつ、各行と各ページにおいて文字数の偏りが無いことが求められる。そこで、 $P_{sum}$  をページ数の総和、 $EP_{textmin}$  を各ページの総文字数の最小値、 $EL_{textmin}$  を各行の総文字数の最小値とする。また、ビジュアルデザインに応用する際に、表示できる領域が限られる、コンテンツによって行数を変えたい、などの要求がある。そこで、 $EL_{textmax}$  を各行の総文字数の最大値、 $EL_{linemax}$  を各頁の総行数の最大値とする。

これらの  $S_{sum}$ ,  $P_{sum}$ ,  $EP_{textmin}$ ,  $EL_{textmin}$ ,  $EL_{textmax}$ ,  $EL_{linemax}$  を用いて設計した評価関数を以下に示す。

$$Eval = -(\alpha \cdot S_{sum} - \beta \cdot P_{sum} + \gamma \cdot EP_{textmin} + \delta \cdot EL_{textmin}), \\ (EL_{textmax} \leq A, EP_{linemax} \leq B)$$

$\alpha, \beta, \gamma, \delta$  は係数であり、括弧内の条件を満たすとする。A,

B は制約であり、ユーザあるいは提案手法を組み込むシステムから与えられるアプリケーション用の変数である。

「今日は動物園に行ってきました」に対して制約  $A \cdot B$  をそれぞれ  $8 \cdot 2$  としたとき、制約を満たす組み合わせは 5 通りであり、この評価関数による最も評価値の良かったパターンは「今日は動物園に / 行ってきました」であった。これは 3.1 節で述べた提案手法が目指す改行の要件を満たしている。

## 4 評価実験

### 4.1 短文における視覚的バランスの検証

#### 4.1.1 デザイン

この実験では、提案手法による文章への改行付与がベースライン手法と比較して視覚的バランスを向上させることを検証する。比較対象として、視覚的バランスを考慮せず改行の可能性がある位置でとにかく区切る手法を再現するために、3.2 節で述べた改行可能性  $p$  の推定の結果を基に、しきい値 0.5 を越えた位置で改行を加えるベースライン手法を用意した。

実験に使用する文章として、公益財団法人日本国際教育支援協会と独立行政法人国際交流基金が主催する日本語能力試験<sup>2</sup>より、4 段階あるうち上から二番目の難易度に位置する N2 の問題を用いた。中文問題に該当する文章の中から 30 文字前後となる文字列を抜き出し、文章が不自然にならないように接続詞等を修正した上で 28 文字～32 文字の範囲に収まるように整形し、これを 10 セット作成した。読点は読みやすさに関する手がかりを与えてしまうため整形の際に削除した。

実験参加者は評価実験用の Web サービスにアクセスし、実験開始ボタンを押す。10 セットの文章の中からランダムに 1 つが選ばれ、行あたりの文字数制限を 8～20 文字の範囲にランダムに決めた上で、ベースライン手法と提案手法によって自動で改行を加えられた結果がランダムな順序で左右に並べて提示される。実験参加者は読みやすさに関する項目として「読みやすさ (Readability)」、文章を見た際に受ける印象に関する項目として「視覚的バランス (Visual Balance)」「自然さ (Naturality)」「親しみやすさ (Familiarity)」の項目について、2 つのパターンのうちどちらがより良いかを「左の方がとても良い」「左の方が良い」「左の方がやや良い」「右の方がやや良い」「右の方が良い」「右の方がとても良い」の 6 段階の中からそれぞれ選択するよう求められた。10 セットの文章がランダムな順序で全て出題され、全ての項目について回答を行った後、実験終了ボタンを押すことで回答をサーバに送信した。

#### 4.1.2 結果

クラウドソーシングサービス「ランサーズ<sup>3</sup>」を用いて実験参加者を集めた。実験参加者は実験に関する注意事項やプライバシーポリシーについて同意したのち手順に従って実験に参加した<sup>4</sup>。そのうち 207 件の回答が得られ、これらを有効な結果として扱った。

2 : <https://www.jlpt.jp/index.html> (最終閲覧: 2022 年 2 月)

3 : <https://www.lancers.jp/> (最終閲覧: 2022 年 2 月)

4 : 筑波大学図書館情報メディア系 研究倫理審査 承認第 21-71 号

表 3 6 段階評価の平均値.

評価項目	平均値	ニュートラル値からの差
読みやすさ	3.33	+0.83
視覚的バランス	3.30	+0.80
自然さ	3.36	+0.86
親しみやすさ	3.33	+0.83

「ベースライン手法の方がとても良い」を 0 点, 「提案手法の方がとても良い」を 5 点として, 6 段階評価を平均で取ったものを表 3 に示す. 読みやすさ, 視覚的バランス, 自然さ, 親しみやすさの全ての項目で提案手法の方が良いと評価された. また, ベースライン手法と提案手法の中間を示すニュートラル値は 2.5 であるが, ニュートラル値からの差は 0.8 程度であった. 最も大きく差が出たのは自然さに関する項目であった.

## 4.2 長文における可読性と理解度の検証

### 4.2.1 デザイン

この実験では, 500 文字程度の長文を文字数が限られた領域に分割して表示し, ユーザーがページ送りをしていくことを想定したときに, 機械的に改行・改頁を挿入するベースライン手法と, 行やページのバランスを考慮する提案手法を比較する.

実験参加者が実験用の Web サービスにアクセスし, 実験開始ボタンを押すと文章が表示されるテキストウィンドウが提示される. テキストウィンドウは行あたりの文字数を 20 文字以内, ページあたりの行数を 3 行までに制限されており, この制限を超えてテキスト要素を表示することはできない. 実験参加者はテキストウィンドウ上に表示される文章をできるだけ素早く正確に読むように指示される. この際, 実験参加者はテキストウィンドウを操作しページ送りとページ戻しを行うことができる. テキストウィンドウ上に表示される課題文は, ベースライン手法または提案手法によって自動的に改行・改頁が付与される.

機械的に改行・改頁を挿入するベースライン手法として,

- (1) ベースライン手法 A: 規定の文字数・行数を超えたときに改行・改頁を挿入する.
- (2) ベースライン手法 B: 形態素解析で文章を形態素に分解し, 規定の文字数・行数の範囲であれば形態素を追加, そうでなければ改行・改頁を挿入する.

の 2 つの実装を準備した. これらは両方ともナイーブな手法であるが, ベースライン手法 A は意味合いとしてランダムに改行・改頁を挿入したものとみなせるのに対して, ベースライン手法 B はある程度文章の意味合いを反映していることとなる.

実験に使う文章として, 実験 1 と同様に日本語能力試験 N2 に記載された読解問題から中文問題に該当する問題を引用した. 実験参加者が問題の全てのページを閲覧したとき, 文章を読み終えるボタンが出現し, これ以上読む必要がなければボタンを押すことでクイズ・アンケート画面に遷移する. クイズ・アンケート画面では, 課題文に対応する 3 つのクイズと 2 つのアンケートが出題される. クイズの内容は日本語能力試験の公式問

表 4 出題した問題の結果.

	WPM	正答数	読解時間
問題 A	730	2.71	55.48
問題 B	466	2.39	55.64
問題 C	456	1.96	63.46

題集より引用した課題文に対応する設問であり, 4 択のうちから最も回答にふさわしいものを選択する. アンケートとして, 「文章は読みやすいと感じましたか?」「文章は理解しやすいと感じましたか?」の 2 つが提示され, 「とても読みにくかった/理解しにくかった」「読みにくかった/理解しにくかった」「少し読みにくかった/少し理解しにくかった」「少し読みやすかった/少し理解しやすかった」「読みやすかった/理解しやすかった」「とても読みやすかった/とても理解しやすかった」の 6 択から 1 つを選択する.

3 問の問題がランダムな順序で出題され, それぞれ課題文の読解とクイズ・アンケートの解答を行う. ここでは, 実験参加者の能力によって特定の改行手法の読解結果にバイアスが生じることを防ぐため, 同じ実験参加者に対して出す 3 問の問題のうち, 適用する改行手法はベースライン手法 A・B・提案手法がランダムな順序で全て適用されるようにプログラムした. 全ての問題について回答を行った後, 実験終了ボタンを押すことで回答をサーバに送信し実験を終了する.

### 4.2.2 結果

クラウドソーシングサービス「ランサーズ」を用いて実験参加者を集めた. 実験参加者は実験に関する注意事項やプライバシーポリシーについて同意したのち, 手順に従って実験に参加した. 実験を通じて 235 件の回答が得られた.

図 2 に横軸に読解時間に  $\log$  の 10 底を取った値, 縦軸に正答数を取ったときの実験参加者の結果を青色の丸でプロットした結果を示す. 読解時間は, 実験参加者が課題文を読むのにかかった時間の合計値であり, 説明画面やクイズ・アンケート画面, 終了画面の閲覧時間は含まれていない. 多くの実験参加者が合計で 300 秒以内に課題文を読んでいるのに対して, 最長で 752 秒かかった実験参加者の存在が確認できた. 正答数は課題文に対応した 3 つのクイズの正答数の合計であり, 全問不正解で 0, 全問正解で 9 の値を取る. このとき, 読解時間と正答数の相関係数は  $r = 0.55$  であり, 正の相関にあった. また, 上位の正答数を記録している実験参加者は読解時間が 4 から 7 の間に位置することが分かる. 一方で, 1 から 2 に位置する実験参加者は低い正答数となっており, 加えてこの時間の範囲で課題文を読むことは非常に困難であるため, 報酬目的の不真面目な実験参加者であると言える. そこで, 不真面目な実験参加者が結果に及ぼす影響を考慮するため, 合計読解時間が 3 以下の実験参加者を集計から除外した.

問題ごとの Word per minute (WPM), 正答数, 読解時間の平均を表 4 に示す.

WPM とは, 以下の式で表される項目であり, 短い時間で正答数が多くなると高くなる.



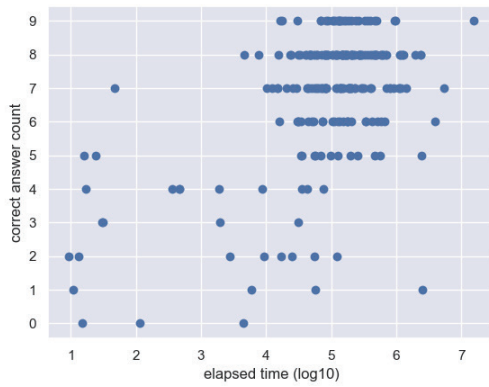


図2 読解時間と正答数の関係。

$$WPM = words\_count \times \frac{60}{elapsed\_time} \times \frac{correct\_answer\_count}{3} \quad (1)$$

*words\_count* は課題文に含まれる単語の数、*elapsed\_time* は読解時間、*correct\_answer\_count* は正答数を示す。ただし、*words\_count* は主に英語においてスペースで区切られた範囲を1と数えることで表現する指標であり、日本語ではその判定が使えないことから、文字の数とした。WPMは問題A、B、Cの順に高かった。問題A、問題Bの正答数の平均は全問正解した場合の3に近い2.75と2.41であり読解時間も60秒を切っていたのに対して、問題Cでは、正答数の平均が2を下回っており読解時間も60秒を越えた。このことから、問題A・Bは比較的簡単な問題であるのに対して、問題Cは比較的難しい問題であると言える。

全ての問題をまとめた際の、ベースライン手法A・Bおよび提案手法による違いを表5に示す。全体としては、正答数、読解時間、読みやすさ、理解しやすさの項目において提案手法が最も良かった。しかしながら、これらの結果に統計的有意差は見受けられなかった。

問題Aについての、ベースライン手法A・Bおよび提案手法による違いを表6に示す。WPM、正答数、読解時間、読みやすさ、理解しやすさの項目について提案手法が最も良い値だった。WPM以外の項目で2番目に良い値となったのはベースライン手法Bで、最下位はベースライン手法Aであった。

問題Bについての、ベースライン手法A・Bおよび提案手法による違いを図7に示す。正答数、読解時間、読みやすさ、理解しやすさの項目について提案手法が最も良い値だった。ベースライン手法の間では、WPMと読みやすさがベースライン手法Bの方が良かったものの、正答数、読解時間、理解しやすさではベースライン手法Aの方が良い値となり、問題Aと逆転した。

問題Cについての、ベースライン手法A・Bおよび提案手法による違いを表8に示す。WPMは問題Aと同様の傾向となった。正答数ではベースライン手法Bと提案手法が同程度であったものの、読解時間、読みやすさはベースライン手法Bが最も良かった。そして、読解時間、読みやすさ、理解しやすさの項目では提案手法が最も悪かった。これらは問題A・Bの傾向と

異なっており、また定量的な指標と定性的な指標の傾向が一致しなかった。

### 4.3 考察

1つ目の実験では、30文字程度の短文において提案手法の視覚的印象を評価した。結果として、読みやすさ、視覚的バランス、自然さ、親しみやすさの全ての項目でベースライン手法と比較して指標が改善したことを確認した。評価関数として考慮に含まれている読みやすさと視覚的バランスよりも、文章の自然さや親しみやすさの方が上昇幅が大きかった。これは、視覚的バランスを考慮することによって、普段の生活でよく見る媒体上での表現に近くなり、結果的に文章が自然であると判断されたと想定できる。また、6段階評価で表現したときに、ニュートラル値からの差は+0.8程度であった。これは、「提案手法の方がやや良い」と「提案手法の方が良い」の間のうち、やや「提案手法の方が良い」に近い位置にプロットされるものである。このことから、提案手法による読みやすさ、視覚的バランス、自然さ、親しみやすさの向上は劇的なものではなく限定されたものであることに注意する。また、6段階評価を二値化した場合は、20%程度の実験参加者がベースライン手法の方が良いと評価しており、この結果には属人性が存在することにも留意する。

2つ目の実験では、500文字程度の長文において提案手法の読解速度と理解度への影響を調べた。図2で示したように読解時間と正答数には正の相関があり、文章を読むのに長い時間をかけるほど正答数は上がる傾向が確かめられた。一方で、表5に示したように、提案手法による改行付与は読解時間を短くしたまま正答数を上げた。これは全体のトレードオフと反しており、限られた領域に表示されるページ送りを伴う長文において、行やページの視覚的バランスを考慮することで読みやすさと理解しやすさを上げることができるといえる。これらの間には統計的有意差は認められなかったものの、読解時間や正答数などの定量的指標と読みやすさと理解しやすさの定性的な評価の傾向は一致していた。

また、難易度が簡単な問題A・Bと、比較的難問である問題Cで傾向が変わることも注意したい。問題が簡単な問題A・Bであれば視覚的バランスを考慮することで結果的に読みやすくなったが、難しい問題である問題Cでは視覚的バランスを考慮せず意味合いのみを考慮したベースライン手法Bが優れた結果となった。これは、問題Cは難解な単語が多く文章の接続の関係も理解しづらい構成になっていたことから、下手に見た目のバランスを整えるよりもできるだけテキストウィンドウに表示する文字数が多くなるようにした方が理解しやすかったのだと推測される。これらの傾向より、視覚的バランスを考慮した改行は簡単な文章であれば理解を助け可読性を向上させる要素になるが、問題が難しければその効果は薄れる傾向にあり、かえって逆効果になってしまうことがいえる。

表 5 実験全体の結果.

	WPM	正答数	読解時間	読みやすさ	理解しやすさ
ベースライン手法 A	557	2.68	59.16	4.42	4.58
ベースライン手法 B	543	2.70	54.52	4.55	4.60
提案手法	551	2.74	52.15	4.63	4.82

表 6 問題 A の結果.

	WPM	正答数	読解時間	読みやすさ	理解しやすさ
ベースライン手法 A	723	2.68	59.16	4.42	4.58
ベースライン手法 B	707	2.70	54.52	4.55	4.60
提案手法	765	2.74	52.15	4.63	4.82

表 7 問題 B の結果.

	WPM	正答数	読解時間	読みやすさ	理解しやすさ
ベースライン手法 A	464	2.35	52.53	3.77	3.87
ベースライン手法 B	470	2.33	59.29	3.83	3.82
提案手法	437	2.50	55.25	3.99	3.96

表 8 問題 C の結果.

	WPM	正答数	読解時間	読みやすさ	理解しやすさ
ベースライン手法 A	466	1.91	63.40	3.70	3.66
ベースライン手法 B	442	1.99	61.51	3.72	3.50
提案手法	490	1.98	65.23	3.68	3.46

## 5 議論と今後の展望

### 5.1 境界に対する改行可能性の推定精度

提案手法では、改行可能性  $p$  を推定するために形態素解析の結果を使用した。具体的には、形態素解析から文節を推定し、文節の境界に対して  $p = 0.2$ ,  $p = 0.5$ ,  $p = 1.0$  を付与した。すなわち、この改行可能性は 3 段階の強度で表されており可読性の表現に欠ける。本研究の目的は可読性と視覚的バランスの両立であり、この改行可能性  $p$  の推定の精度を上げることが結果に大きく寄与しないと判断したが、従来手法のように係り受け情報などを使うことが改行の質がより向上する可能性がある。例えば、改行可能性  $p$  の推定を他の従来研究や Google が提供する budoux<sup>5</sup> に差し替えることが可能である。一方で、これらの手法はデータセットに依存するため、用途に合わせて正しい改行可能性の推定手法を選択することが求められる。

### 5.2 多言語への適用

本研究では、日本語に焦点を当てて改行手法を論じている。提案手法は改行可能性  $p$  の推定と組合せ最適化の 2 つのフェーズから構成されるが、英語などの他の言語に提案手法を応用する際は、言語ごとに改行可能性  $p$  を推定する部分を差し替える必要がある。ここで、英語に焦点を当てると、英語圏の文化では読みやすい位置に改行を加えるという文化は見受けられず複数の行にまたがる単語はハイフネーション (-) で接続される。そのため、既存のデータセットから学習することは難しく、人手によるアノテーションは労力がかかる。一方で、英文を読ん

でいく際に、ある程度まとまった塊に区切ることによって読みやすくする「スラッシュリーディング」と呼ばれるテクニックもある。そこで、本研究でも用いた形態素解析とルールベースによる改行可能性  $p$  の推定に、このスラッシュリーディングのルールを取り込むことができる。スラッシュリーディングには一意のルールは存在せず、人によってやり方はまちまちであるが、一定のルールとして「前置詞」「副詞」「準動詞」「関係詞」「接続詞」などの直前で区切ることが知られている。このように、3.2 節で述べた手順で形態素解析および境界推定を行うことで、英語への応用は可能である。

### 5.3 ビジュアルデザインのためのデータドリブンな改行位置推定

本研究では、視覚的バランスを考慮した評価関数を設計した。しかし、偏った文字数の行を持つテキスト要素を意図的に配置する場合があることに注意する。すなわち通常は視覚的バランスを考慮してテキスト要素に改行を加えるが、デザイナーがあえてそれをずらすことによってひと目を引くデザインを実現するテクニックである。提案手法では、このような意図的にバランスを崩した改行は再現できない。これは、グラフィクス要素の構成やテキスト要素のコンテキストに依存するため、これらの事情を反映するのが難しいためである。グラフィクス要素やテキスト要素の持つコンテキストを反映した改行を実現するために、ビジュアルデザインの作例から改行アノテーションを作成し機械学習アプローチによって改行位置を推定する手法が考えられる。テキスト要素のみを考慮に入れた従来の改行位置推定と異なり、グラフィクス要素の顕著性マップなどを入力に組み合わせる必要があり、データセットの準備やマルチモーダルな推定器の構成には更なる工夫が必要となる。

<sup>5</sup> : <https://github.com/google/budoux> (最終閲覧: 2022 年 2 月)

## 6 結 論

本研究では、ビジュアルデザインにおけるテキスト要素の自動レイアウトの一要素として、文字数と行数に制約を与えた下でテキスト要素に対して自動で改行・改頁を付与する手法を提案した。従来よりビジュアルデザインの分野ではテキスト要素とグラフィクス要素の自動レイアウト技術が探求されてきたが、特に日本語においては単にテキスト要素を配置するだけでなく、テキスト要素をどのように分割するかで視覚的印象が大きく変化する。また、改行位置を推定する従来研究は、文章の可読性を最適化することに焦点が当てられており、視覚的バランスは考慮されていなかった。本研究では、ビジュアルデザインにおける改行では可読性のみならず視覚的バランスが重要であるとし、視覚的バランスを考慮した改行位置推定を実現するために、(1) 改行可能性  $p$  を推定する簡易なルールベース手法、(2) 可読性と視覚的バランスの両立を目指す評価関数による最適化、の2つのフェーズで構成される組合せ最適化手法を提案した。

提案手法を評価するためにクラウドソーシングを使って2つの実験を実施した。1つ目の実験では、30文字程度の短文に提案手法とベースライン手法を適用し、ベースライン手法と比較して提案手法による改行が文章の「読みやすさ」「視覚的バランス」「自然さ」「親しみやすさ」を向上させることを明らかにした。この改行による印象の改善は劇的ではないものの、70%の実験参加者が提案手法に対してポジティブな反応をみせた。2つ目の実験では、500文字程度の長文に提案手法とベースライン手法を適用し、視覚的バランスを考慮することによって文章の「理解度」と「読解速度」が定量的・定性的の両面で向上することを明らかにした。これは、人間が文章を読むとき、記号としてシーケンシャルに追うのではなく、ある程度文字列を画像として認識しており、視覚的バランスの考慮がその認識プロセスを補助したのだと推測される。また、文章が簡潔なものほど理解度と読解速度の改善効果が大きくなることがわかった。

今後の展望として、グラフィクス要素やテキスト要素の持つコンテキストを反映した改行を実現するために、ビジュアルデザインの作例から改行アノテーションを作成し、機械学習アプローチによって改行位置を推定するデータドリブンな手法の発展に繋げていきたい。

## 文 献

- [1] Ali Jahanian, Jerry Liu, Qian Lin, Daniel Tretter, Eamonn O'Brien-Strain, Seungyon Claire Lee, Nic Lyons, and Jan Allebach. Recommendation system for automatic design of magazine covers. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, pp. 95–106, 2013.
- [2] Xuyong Yang, Tao Mei, Ying-Qing Xu, Yong Rui, and Shipeng Li. Automatic generation of visual-textual presentation layout. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol. 12, No. 2, pp. 1–22, 2016.
- [3] Peter O'Donovan, Aseem Agarwala, and Aaron Hertzmann. Learning layouts for single-page graphic designs. *IEEE transactions on visualization and computer graphics*, Vol. 20, No. 8, pp. 1200–1213, 2014.
- [4] Masaki Murata, Tomohiro Ohno, and Shigeki Matsubara. Construction of linefeed insertion rules for lecture transcript and their evaluation. *International Journal of Knowledge and Web Intelligence*, Vol. 1, No. 3-4, pp. 227–242, 2010.
- [5] 秋田祐哉, 河原達也. 講演に対する読点の複数アノテーションに基づく自動挿入. *情報処理学会論文誌*, Vol. 54, No. 2, pp. 463–470, 2013.
- [6] 大野誠寛, 村田匡輝, 松原茂樹. 講演のリアルタイム字幕生成のための逐次的な改行挿入. *電気学会論文誌 C (電子・情報・システム部門誌)*, Vol. 133, No. 2, pp. 418–426, 2013.
- [7] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. Color harmonization. In *ACM SIGGRAPH 2006 Papers*, pp. 624–630, 2006.
- [8] Ji-Hwan Kim and Philip C Woodland. The use of prosody in a combined system for punctuation generation and speech recognition. In *Seventh European conference on speech communication and technology*, 2001.
- [9] 鈴木英二, 島田静雄, 近藤邦雄, 佐藤尚. 日本語文章における句読点自動最適配置. *全国大会講演論文集, データ処理*, pp. 185–186, 1995.
- [10] 村田匡輝, 大野誠寛, 松原茂樹. 読点の用法的分類に基づく自動読点挿入. *研究報告音声言語情報処理 (SLP)*, Vol. 2010, No. 8, pp. 1–8, 2010.
- [11] 村田匡輝, 大野誠寛, 松原茂樹. 読みやすい字幕生成のための講演テキストへの改行挿入. *電子情報通信学会論文誌 D*, Vol. 92, No. 9, pp. 1621–1631, 2009.
- [12] 飯泉智朗, 大野誠寛, 松原茂樹. 読みやすい字幕生成のための rnn を用いた講演テキストへの改行挿入. 第 82 回全国大会講演論文集, Vol. 2020, No. 1, pp. 451–452, 2020.
- [13] Hajime Morita, Daisuke Kawahara, and Sadao Kurohashi. Morphological analysis for unsegmented languages using recurrent neural network language model. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 2292–2297, 2015.