

# Web ページ内容の理解を促進する画像の発見

伊原 啓晃<sup>†</sup> 山本 祐輔<sup>††</sup> 大島 裕明<sup>††</sup> 田中 克己<sup>††</sup>

<sup>†</sup> 京都大学 工学部 情報学科 〒 606-8501 京都府京都市左京区吉田本町

<sup>††</sup> 京都大学大学院 情報学研究科 社会情報学専攻 〒 606-8501 京都府京都市左京区吉田本町

E-mail: †{ihara,yamamoto,ohshima,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 手順や動作説明など、テキストのみでは理解しにくいような Web コンテンツが多く存在する。本稿では、与えられた Web テキストの理解を促進させるための補助的・補完的な画像・図形等を検索する方式を検討する。本稿では、化学実験や物理現象など、手法や現象を説明している理解が難しいテキストを対象とした。まず、与えられたテキストと画像周辺テキストの適合度、画像周辺テキストの可読性、専門性という 3 つの特徴量と画像の理解促進効果との関係分析を行った。その後、理解を促進させる画像を検索するためのランキング関数の構築を行った。傾向調査の結果、周辺テキストと対象テキストの適合性が高く、周辺テキストの可読性が中学レベルに近い画像ほどテキストに対する理解促進効果が高くなる傾向にあることが分かった。また、傾向調査の知見を利用して構築したランキング関数の評価結果から、理解促進効果と関係のある特徴は補完対象によって異なることが分かった。

キーワード 画像補完, Web 閲覧, Web マイニング, 情報検索, 理解容易性

## 1. はじめに

今や Web なしでは生活できない。ちょっとした調べ物から生活に密接に関連する情報に至るまで、多くの人が Web から情報を取得するようになった。検索エンジンやポータルサイトの発展によって、ユーザの情報要求に合致した情報に効率よくアクセスできるようになりつつある。しかし、欲しい情報が掲載された Web ページに出会ったとしても、Web ページに記述されている内容を理解することができないケースがしばしば起こりえる。

そのようなケースとして、専門的な内容が書かれた Web ページを閲覧するケースが挙げられる。例えば、ある理科系の知識が乏しいユーザがレポート課題のために「ティティウス・ボデーの法則」の解説ページを読んでいたとしよう。このページには「ティティウス・ボデーの法則は、太陽系の惑星の軌道長半径（太陽からの平均的な距離）は簡単な数列で表せるという法則」と書かれている。記載された内容は確かにティティウス・ボデーの法則を的確に説明している。しかし、知識の乏しいユーザにとっては、説明に用いられている単語が難解なため内容を理解することが難しい。

また、専門的な単語が使われていないものの、テキストを読んでも内容をイメージすることができないケースも存在する。例えば、「水切りできた豆腐を崩しながら入れる」と書かれたレシピサイトを考えよう。このテキストは豆腐や水切りといった比較的平易な単語で構成されているものの、料理に慣れていないユーザにとっては、一連の作業をテキストのみからイメージすることは難しい。

情報検索・情報閲覧システムに関する研究はこれまで盛んに行われてきた。しかし、既存の情報検索システムにおけるランキングアルゴリズムの多くは適合性や人気度に焦点を当ててい

るため、分かりやすい Web ページが上位にランキングされるとは限らない。また、効率よく内容が難しい Web ページをより簡単な内容に置き換えたり、理解が促進するコンテンツを補完したりといった、内容の理解を支援するシステムに関する研究もあまり行われていない。

現状では、情報検索・閲覧時にユーザの情報要求に適合する Web 情報を発見できたとしても、その内容をユーザが理解できるとは限らない。誰しもが検索・閲覧中の Web ページの内容を理解し、最終的に必要な情報を取得するためには、コンテンツの理解容易性を考慮した情報検索・閲覧支援システムが必要となる。

「百聞は一見にしかず」という諺にもあるように、画像コンテンツはテキストコンテンツの内容理解に非常に効果的である。また教育・学習分野で指摘されているように、画像コンテンツを学習に用いることで、注意喚起だけでなく理解促進の効果が期待されるということが報告されている [5]。本稿では、画像補完による Web テキストの理解促進を目指し、テキストコンテンツの理解を促進させる画像の特徴の分析を行う。

以後、2 章では関連研究について述べる。3 章では、画像の周辺テキストの可読性と専門性、理解しにくいテキストと画像の周辺テキストの適合性といった要素に着目し、理解を促進させる画像の特徴量の調査を行う。4 章では、3 章の調査結果をもとに理解を促進する画像のランキング関数について検討する。5 章では、3、4 章での結果をもとに、画像補完による Web ページの理解促進支援のためのシステムに関する展望をまとめる。

## 2. 関連研究

### 2.1 文書の理解容易性評価

文書の理解容易性を評価する研究の多くは、可読性に焦点を当てて行われてきた。これらの研究では教育分野での応用を目

指し、様々な可読性の評価手法が提案されてきた [3]。しかし、テキストの体裁のみに着目した可読性評価では、専門的なテキストの理解容易性を正確に評価することは難しい。Yan らは、医療用語のシソーラスを利用することで、医療分野のテキストの専門性を評価する手法を提案している [8]。Nakatani らは Wikipedia を利用し、特定のドメインに関係なくテキストの専門性を評価し、理解容易性の高い Web テキストを検索する手法を提案している [7]。Akamatsu らは、Web ページ間のリンク解析を行うことで、Web ページの理解容易性を評価する手法を提案している [2]。

これらの手法は、テキストの理解容易性を評価し、分かりやすいテキストを検索することを目的としている。一方、本研究は分かりやすい画像を検索・補完することを目的としている。テキストと画像はコンテンツの理解促進を図るために補完的な関係にある。それゆえ、テキストの理解容易性を考慮した情報閲覧・検索技術と本研究の知見を組み合わせることが、Web ページの理解促進を図る上で重要であると考えられる。

## 2.2 画像補完による情報検索・閲覧支援

Agrawal らは、テキスト中心で書かれた教科書に画像を補完するための手法を提案している [1]。Agrawal らが提案した手法は、教科書の各段落に適合する画像を Web から検索し、教科書全体で画像が重複しないように画像を補完する。この手法は、テキストの内容に適合する画像を補完することで学習者の理解促進を図ることを目的としているが、内容に適合する画像が理解を促進させるとは限らない。本稿で提案する手法は、テキストに対する画像の適合性以外の要素も考慮し、閲覧中のテキストの理解を促進させる画像を Web から検索・補完することを目的としている。

画像補完による情報検索支援に関する研究もいくつか行われている。Li らは Web ページに含まれ、かつ検索クエリに適合する画像を検索結果中に表示させることによって、ユーザの適合性判定を支援するシステムを提案している [4]。Yeh らは、キーワードと画像を組み合わせたクエリを入力できる Web 検索システムを提案している [9]。

## 3. 予備調査

本章では、テキストの理解を促進させる画像の特徴量の調査とその分析結果について述べる。本調査の目的は、テキストに対して画像補完を行ったときの理解促進の度合いと画像の周辺テキストの特徴量との関係を明らかにすることである。

### 3.1 被験者および調査材料

本調査には大学生および大学院生の計 4 人が参加した。

調査対象のテキストとして、Wikipedia から化学反応、化学実験、物理法則を説明するテキストを 10 個取得した。取得したテキストを表 1 に記す。分析対象となる画像の収集には、Yahoo!画像検索 API<sup>(注1)</sup>を用いた。実際にトピック  $t$  に関する説明テキスト  $s$  に対して関連する画像を収集するために、以下の操作を行った:

(1) Yahoo!キーフレーズ抽出 API<sup>(注2)</sup>を用いて、テキスト  $s$  から重要語リスト  $W = \{w_1, w_2, \dots, w_l\}$  を抽出。

(2)  $W$  から 2 もしくは 3 個の語からなるクエリの組み合わせ  $Q$  を生成。

(3) トピック  $t$  と  $q \in Q$  からなる AND クエリを Yahoo!画像検索 API に入力し、画像集合を取得。

(4) 取得した画像集合の中から、ノイズになりうる画像を除去。残った画像の集合を  $I = \{i_1, i_2, \dots, i_m\}$  とする。

(5) トピック  $t$  に関するテキスト  $s$  の補完画像候補として、取得された画像からノイズが除去された画像集合  $I$  よりランダムに 30 個選択。

### 3.2 調査対象となる画像の特徴量

本調査では、理解促進の効果が高い画像を発見するための特徴として、画像の周辺テキストに着目した。テキストに対する画像の理解促進効果を評価するには、色やテクスチャといった画像自身の特徴に着目することも考えられる。しかし、画像自身の特徴量を抽出するにはコストがかかる。本調査では画像の周辺テキストのみに着目し、(1) 周辺テキストと対象テキストの適合性、周辺テキストの (2) 可読性および (3) 専門性を分析し、それら特徴量とユーザが実際に感じた画像の理解促進効果との関係を分析した。

なお、今回の調査では画像周辺のテキスト部分前後 200 字を画像の周辺テキストとした。

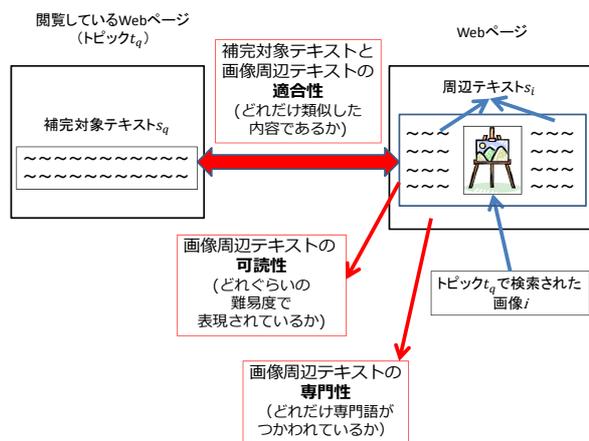


図 1 注目した特徴量

対象テキストに対する画像周辺テキストの適合性

対象テキストを  $s_q$ 、そのトピックを  $t_q$  とするとき、 $s_q$  と  $t_q$  によって検索された画像集合を  $I = \{i_1, i_2, \dots, i_n\}$ 、画像  $i \in I$  の周辺テキストを  $s_i$  と定義する。今、テキスト  $t$  の特徴ベクトルを  $v_t$  とするとき、テキスト  $s_q$  に対するテキスト  $s_i$  の適合性  $rel(s_q, s_i)$  を以下で定義する:

$$rel(s_q, s_i) = \frac{v_{s_q} \cdot v_{s_i}}{|v_{s_q}| |v_{s_i}|} \quad (1)$$

(注1): Yahoo!画像検索 API: <http://developer.yahoo.co.jp/Webapi/search/jlp/>

(注2): Yahoo!キーフレーズ抽出 API: <http://developer.yahoo.co.jp/Webapi/>

表 1 調査に用いたテキスト集合

トピック	説明テキスト
ティティウス・ボーデの法則	太陽系の惑星の軌道長半径（太陽からの平均的な距離）は簡単な数列で表せるという法則．
右ねじの法則	電流と磁場（磁束）の向きに関する法則．電流を右ねじが進む方向に直進させると、磁場が右ねじの回転方向に生じること．
運動量保存の法則	ある系に外部から力が加わらないかぎり、その系の運動量の総和は不変であるという物理法則． 外部からの力が働かない問題の例としては、物体の衝突問題がある．
レンツの法則	電磁誘導に関する法則．何らかの原因によって誘導電流が発生する場合、電流の流れる方向は誘導電流の原因を妨げる方向と一致するというもの．
万有引力	「地上において質点（物体）が地球に引き寄せられるだけではなく、この宇宙においてはどこでも全ての質点（物体）は互いに gravitation（＝引き寄せる作用、引力、重力）を及ぼしあっている」とする考え方、概念、法則のことである．
フックの法則	力学や物理学において、ばねの伸びと弾性限度以下の荷重は正比例するという近似的な法則である．
地殻変動	地殻に応力が加わることで、長期間にわたり地殻の位置が年間数 mm から数 cm 程度移動する現象である． 地殻を構成するプレート運動や断層運動と密接に関係している．
ハーバー・ボッシュ法	鉄を主体とした触媒上で水素と窒素を 400 度から 600 度、200 から 1000 気圧の超臨界流体状態で直接反応させ、アンモニアを生産する方法である． 現代の工業化学では、メタンから不均一系触媒を使って分離された水素と大気中の窒素とを反応させてアンモニアを合成している．
蒸留	混合物を一度蒸発させ、後で再び凝縮させることで、沸点の異なる成分を分離・濃縮する操作をいう． 通常、目的成分が常温で液体であるか、融点が高々 100 程度の固体の場合に用いられる．
水上置換法	気体の捕集法の 1 つである．主に水に溶けにくい気体の捕集に用いられる．水槽などに水を入れ、気体の発生口をその中に入れる． はじめに発生口から出る気体は、気体を発生させる装置内の空気なので、しばらく待ってから気体を容器に集める． このとき容器の中は水で満たしておく．この方法では空気の混入も少なく、気体の捕集量の計測にも使われる．

特徴ベクトルの要素には、形態素解析エンジン MeCab<sup>(注3)</sup> で得られる名詞を用いた．また、用意したストップワードリストにより不要語を取り除いた．各要素の重み付けには tf-idf 値を用いた．

あるテキスト  $s$  において語  $w$  が出現する回数を  $c(w, s)$ 、テキスト  $s$  に含まれる語の総数を  $n_s$ 、語  $w$  が出現するテキストの数を  $df(w)$ 、特徴ベクトルを生成する全テキストの数を  $N (= |I| + 1)$  とするとき、特徴ベクトル  $v_s$  の語  $w$  に関する重みは以下の式で算出した：

$$v_s(w) = tf(w, s)idf(w) = \frac{c(w, s)}{n_s} \log \frac{N}{df(w)} \quad (2)$$

### 画像周辺テキストの可読性

画像周辺テキストの可読性を評価する手法として、本研究では可読性評価アルゴリズム obi<sup>(注4)</sup>を用いた．

obi は小中高大の教科書 127 冊から可読性レベルに応じて構成された 13 種類の言語モデルを用い、対象テキストの各モデルに対する尤度からテキストの可読性を評価するアルゴリズムである．obi はテキストを入力として与えると、1 から 13 の値を出力する（値が大きいほど難易度が高い）．本研究では、obi への入力テキストを  $s$ 、obi の出力する値を  $x$  としたとき、テキスト  $s$  の可読性  $read(s)$  は以下の式で算出した：

$$read(s) = \frac{14 - x}{13} \quad (3)$$

### 画像周辺テキストの専門性

本研究では、画像の周辺テキストの中にトピック  $t$  に関する専門的な語がどれだけ含まれているかで画像周辺テキストの専門性を評価した．

専門語の抽出と語の専門度の算出には、Nakatani らが提案

した手法を利用した [6]．Nakatani らの手法では、Web 検索エンジンの索引データからトピック  $t$  に関連する頻出語リスト  $W = \{w_1, w_2, \dots, w_n\}$  を抽出する．その後、語  $w \in W$  を見出し語とする Wikipedia 記事が特定の分野の Wikipedia 記事集合からどの程度偏ってリンクされているかを分析することで、語  $w$  の専門度  $s(w)$  を算出する．

今、Nakatani らの手法で得られたトピック  $t$  に関連する専門語のリスト  $W_t$ 、画像周辺テキスト  $s$  に含まれている語のリストを  $W_s$  とする．このとき、トピック  $t$  に関する画像周辺テキスト  $s$  の専門度  $spec(t, s)$  は以下の式を用いて算出した：

$$spec(t, s) = \sum_{w \in W_t \cap W_s} s(w) \quad (4)$$

### 3.3 調査方法

本調査では、4 人の被験者に合計 10 種類のタスクを依頼した．被験者は筆者らが事前に準備した実験用ウェブサイトを通じてタスクを実行した．各タスクの流れは以下の通りである：

(1) あるトピックに関する説明テキストを提示した．

(2) 30 枚の画像を提示し、事前に見せた説明テキストの内容を理解するために各画像がどの程度有用であったかを評価してもらった．評価には 5 段階のリッカート尺度を用いた（1: 全く役に立たない、5: 非常役に立つ）．

(3) このようなタスクを表 1 に記した説明テキストに関して合計 10 種類行ってもらった．

なお、各タスクでは時間制限を設けなかった．また、必要ならば拡大画像も閲覧できるようにした．

### 3.4 結果

#### 分析 1

調査終了後、説明テキストに対する画像周辺テキストの適合性、画像周辺テキストの可読性および専門性を計算し、画像による説明テキストの理解促進度との関係を分析した．本分析では、各画像による説明テキストの理解促進度を、4 人の被験者が付けたスコアの平均値とした．

図 2 は、説明テキストに対する画像周辺テキストの適合性がある値を持つ画像のうち、理解促進効果が高い画像/低い画像

(注3): 形態素解析エンジン MeCab: <http://mecab.sourceforge.net/>

(注4): obi: <http://kotoba.nuee.nagoya-u.ac.jp/sc/readability/obi.html>

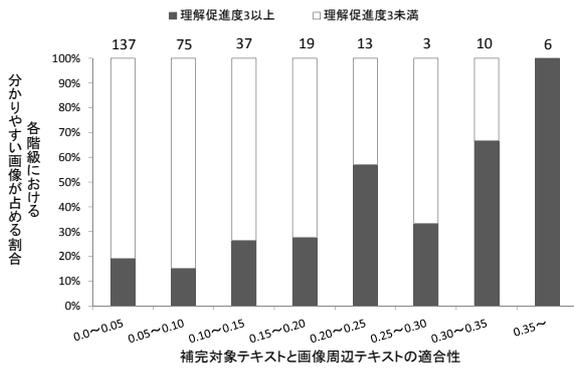


図 2 説明テキストに対する画像周辺テキストの適合性とテキストの理解促進度との関係。グラフ上の数字は、各階級に属する画像の個数の総和を表わす

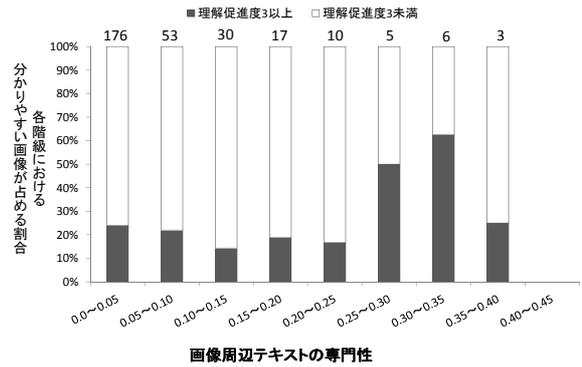


図 4 画像周辺テキストの専門性とテキストの理解促進度との関係。グラフ上の数字は、各階級に属する画像の個数の総和を表わす

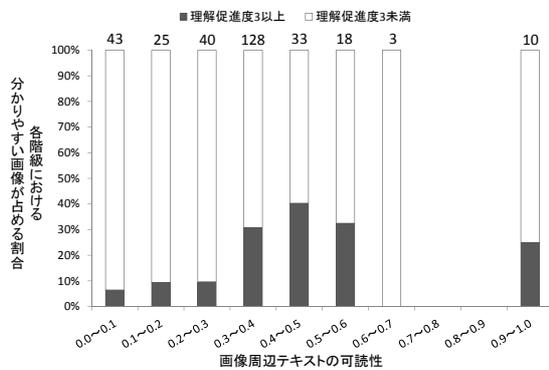


図 3 画像周辺テキストの可読性とテキストの理解促進度との関係。グラフ上の数字は、各階級に属する画像の個数の総和を表わす

の割合を示したグラフである。

図 2 によると、0.05~0.10 の部分と 0.25~0.30 の部分は例外として、適合性が上昇するにしたがって、理解促進効果が高いと評価された画像の割合は概ね大きくなっていることが分かった。これは、周辺テキストの適合性が低い画像は説明テキストと関係のない画像であるため、理解促進効果が得られないからであると考えられる。この結果から、説明テキストに対する画像周辺テキストの適合性が高い画像を提示することで、説明テキストに対する理解をある程度促進できる可能性があると考えられる。

図 3 は、画像周辺テキストの可読性がある値を持つ画像のうち、理解促進効果が高い画像/低い画像の割合を示したグラフである。

図 3 によると、理解促進効果が高いと評価された画像の割合は、概ね 0.4~0.5 の部分を頂点とした山なり分布になっていることがわかる。これは、簡単すぎるテキスト中では、テキストのレベルに合わせてあまりにも単純すぎる画像が用いられ、逆

に難しすぎるテキスト中では、その内容に合わせて画像が示す内容も難解になることが考えられる。この結果より、周辺テキストの可読性が 0.4~0.5 に近い値である画像ほど、説明テキストの内容の理解を促進させる効果が高いと考えられる。

テキストの可読性が 0.4~0.5 という値は、可読性分析アルゴリズム obi のシステムの出力でいうと 7~9 の難易度である。すなわち、このスコアを持つテキストは、中学レベルの表現で書かれたものであることを意味している。このことから、中学程度の文章レベルで書かれたテキスト中で用いられている画像が、テキストの内容を理解する上では最も効果が高いということが考えられる。

図 4 は、画像周辺テキストの専門性がある値を持つ画像のうち、理解促進効果が高い画像/低い画像の割合を示したグラフである。

図 4 によると、理解促進効果が高いと評価された画像の割合は、概ね 0.30~0.35 の部分を頂点とした山なり分布になっていることがわかる。この結果より、周辺テキストの専門性が 0.30~0.35 に近い値である画像ほど、説明テキストの内容の理解を促進させる効果が高いと考えられる。

画像周辺テキストの専門性スコアが高いということは周辺テキストに専門記述が多いことを意味している。そういう専門的なテキストを周辺に持つ画像は専門的な説明に使用されるような画像であるため、理解が困難である場合が多い。よって、画像周辺テキストの専門性スコアが高い画像は理解促進効果が小さい画像であると考えられる。

## 分析 2

上記の分析から以下の傾向があることがわかった。

- 説明テキストに対する画像周辺テキストの適合性スコアが高い画像ほど理解促進効果が高い。
- 画像周辺テキストの可読性スコアが 0.4~0.5 に近い値を持つ画像ほど理解促進効果が高い。
- 画像周辺テキストの専門性スコアが 0.30~0.35 に近い値を持つ画像ほど理解促進効果が高い。

これらをより詳細に検証するために、理解促進効果と各特微量

表 2 各画像の説明テキストの理解促進度と  $rel, 1 - |read - 0.4|, 1 - |spec - 0.30|$  それぞれとの相関

設問 ID	rel	$1 -  read - 0.4 $	$1 -  spec - 0.30 $
0	0.584	0.123	0.111
1	0.222	0.118	-0.145
2	0.666	0.321	0.350
3	0.444	0.202	0.201
4	0.305	-0.425	-0.104
5	0.319	-0.181	-0.0150
6	0.411	0.320	0.0936
7	0.511	0.281	0.359
8	0.156	0.416	-0.380
9	0.125	0.230	-0.161
絶対値平均	0.374	0.262	0.192

との相関の分析を行った。

以下では、説明テキストに対する画像周辺テキストの適合性スコアを  $rel$ 、画像周辺テキストの可読性スコアを  $read$ 、画像周辺テキストの専門性スコアを  $spec$  とし、各画像の説明テキストの理解促進度と  $rel, 1 - |read - 0.4|, 1 - |spec - 0.30|$  それぞれとの相関を分析した。

表 2 は、各画像による説明テキストの理解促進度と  $rel, 1 - |read - 0.4|, 1 - |spec - 0.30|$  それぞれとの相関値の値を設問ごとにまとめたものである。表 2 から、画像による説明テキストの理解促進度と  $rel$  との相関の全体的な傾向を分析すると、設問 ID が 8, 9 のを例外として、概ね画像による説明テキストの理解促進度と画像周辺文章の適合性スコアに相関関係があることが読み取れた。この結果より「説明テキストに対する画像周辺テキストの適合性スコアが高い画像ほど理解促進効果が高い」傾向は存在すると考えられる。

画像による説明テキストの理解促進度と  $1 - |read - 0.4|$  との相関の全体的な傾向を分析すると、設問 ID が 0, 1, 5 のときは例外として画像による説明テキストの理解促進度と画像周辺文章の  $1 - |read - 0.4|$  スコアとの間には弱い相関がみられた。この結果より「画像周辺テキストの可読性スコアが 0.4 ~ 0.5 に近い値を持つ画像ほど理解促進効果が高い」傾向が少しはあると考えられる。

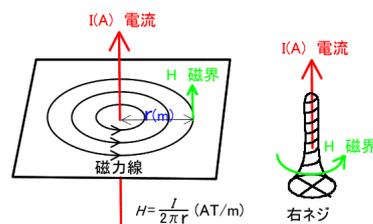
画像による説明テキストの理解促進度と  $1 - |spec - 0.30|$  スコアとの相関の全体的な傾向を分析すると、設問 ID が 2, 7, 8 のときは例外として画像による説明テキストの理解促進度と画像周辺文章の  $1 - |spec - 0.30|$  の間には相関は見られなかった。この結果より「画像周辺テキストの専門性スコアが 0.30 ~ 0.35 に近い値を持つ画像ほど理解促進効果が高い」傾向は見られないと考えられる。

さらに詳細に検証を行うために設問別の相関値に注目した。

設問 ID1 は「右ねじの法則」についての説明テキストである。表 2 から、この設問は、説明テキストの理解促進度と  $rel$  には弱い相関、説明テキストの理解促進度と  $1 - |read - 0.4|, 1 - |spec - 0.30|$  には相関がないことがわかった。この説明テキストの補完にふさわしいと判断された画像はその画像だけで右ねじの法則を説明できてしまうようなものが多く、そのような画像の周辺テキストでは説明記述として書くべきことがほとんどない状況が見受けられた。あまりにも画像だけで説明が成

アンペア右ねじの法則とはなにか [top map](#)  
map2 電気の基本法則

電流が流れるときには磁界が生じ、その磁界の方向は右回りのネジを電流の方向に捻じ込む時のネジの回転する方向である。



発見者  
アンペール  
アンドレ・マリー・アンペール (Andre Marie Ampere  
1775/1/20-1836/6/10)  
フランスの物理学者で、アンペールの法則を発見した電磁気学の創始者の一人です。電流の単位アンペア(A ampere)はアンペールの名にちなんでいます。

図 5 右ねじの法則の例

り立ってしまうような事例に対しては、周辺テキストの説明記述の必要性が薄いため周辺テキストの記述の内容が異なってしまうことが考えられる。例えば図 5<sup>(注5)</sup> は、画像周辺文章として「電流が流れるときには磁界が生じ、その磁界の方向は右回りのネジを電流の方向に捻じ込む時のネジの回転する方向である。」と書かれた一例である。この画像は「右ねじの法則は下の図のことである。」のように一切テキストによる説明をなくとも画像の説明情報だけで右ねじの法則を説明出来てしまう。よって、周辺テキストに書くべき記述がほとんどないため、理解促進効果が高い画像同士でも周辺テキストに書かれている内容がテキストごとに異なった傾向になってしまっているのではないかと考えられる。

設問 ID8 は「蒸留」についての説明テキストである。表 2 を見ると、ほかの設問と比べてこれらは非常に適合性スコアとの相関が弱いことがわかった。一方、強い相関とは言えないもののこれらは説明テキストの理解促進度と  $1 - |read - 0.4|$  に相関があることがわかった。正解画像について分析するとイラスト販売のページの画像が半分くらいを占めていた。このため、先のような相関関係が表れていたと考えられる。

設問 ID9 は「水上置換」についての説明テキストである。表 2 から、説明テキストの理解促進度と  $1 - |read - 0.4|$  には弱い相関、説明テキストの理解促進度と  $rel, 1 - |spec - 0.30|$  には相関がないことがわかった。この設問も設問 ID1 と同様に、 $rel, 1 - |read - 0.4|, 1 - |spec - 0.30|$  どれもあまり相関がなかった。正解画像について分析すると、物質の捕集方法に水上置換を使ったということによって水上置換の図が入っていたものが少なからず存在した。水上置換という手法を使って別の用語を説明しているページのような水上置換の手法そのものについて説明していないページでも水上置換の内容理解を促進させるような画像が掲載されていたために、理解促進効果が高い画像同士でも周辺テキストに書かれている内容が違った内容になってしまっているから適合性の相関が低くなっているのではないかと

(注 5): <http://www.geocities.jp/hiroyuki0620785/k4housesoku/ampeare.htm>

と考えられる。

#### 4. 理解促進を支援する画像の検索

この章では、予備調査の知見を用いて、ランキング関数を構築し、そのランキング関数によって得られた結果から考察を行う。

##### 4.1 ランキング関数

本章では、予備調査で得られた調査データに対して線形回帰分析を行うことで得られた回帰式をランキング関数として用いた。

今、説明テキストに対する画像周辺テキストの適合性スコアを  $rel$ 、画像周辺テキストの可読性スコアを  $read$  と呼ぶことにする。先の調査結果から、ランキング関数に使用する説明変数には  $rel, 1 - |read - 0.4|$  を選択した。その上で、訓練データを入力として上記で決めた説明変数を使った線形回帰分析を行い回帰式を求めた。回帰分析には、最小二乗法による線形回帰分析の方法を選択した。テストデータの画像  $i$  の理解促進度を  $c(i)$ 、適合性スコアを  $rel(i)$ 、可読性スコアを  $read(i)$ 、訓練データを  $I = \{i_1, i_2, \dots, i_n\}$  とし、 $E(I) = \sum_{i \in I} (c(i) - y_i)^2$  の値が最小になるような  $y_i = \alpha \cdot rel(i) + \beta \cdot (1 - |read(i) - 0.4|) + \gamma$  の係数  $\alpha, \beta, \gamma$  を用いた関数

$$y(i) = \alpha \cdot rel(i) + \beta \cdot (1 - |read(i) - 0.4|) + \gamma \quad (5)$$

が訓練データ  $i$  の回帰式である。この回帰式をランキング関数として用いた。

##### 4.2 評価実験

$rel$  と  $1 - |read - 0.4|$  を説明変数に用いたランキング関数の決定とその評価を行うために以下の操作を行った：

(1) テストデータを5つのグループに分ける。今回は、

- 設問 ID0, 1 であるデータをグループ 1
- 設問 ID2, 3 であるデータをグループ 2
- 設問 ID4, 5 であるデータをグループ 3
- 設問 ID6, 7 であるデータをグループ 4
- 設問 ID8, 9 であるデータをグループ 5

というように設定した。

(2) グループ 1~5 のうちグループ  $i$  ( $i = 1, 2, 3, 4, 5$ ) を除いたデータを使用して線形回帰分析を行い一次回帰式を得た。

(3) 得られた回帰式をランキング関数としてグループ  $i$  の2個の各設問に適用し、ランキングを行った。

(4) グループ  $i$  を構成する各設問の Average Precision(AP) 計算し、グループ  $i$  の Mean Average Precision(MAP) を求めた。

(5) 各グループに対して 2. ~ 4. の操作を行ったという方法を用いた。

またベースラインは、以下の手法を用いた。

- 適合性スコア降順でランキングした手法 (rel)
- 可読性スコア降順でランキングした手法 (read)
- 専門性スコア昇順でランキングした手法 (spec)

表 3 各グループ毎の MAP

グループ	rel	read	spec	提案手法
グループ 1	0.613	0.493	0.565	0.631
グループ 2	0.545	0.235	0.272	0.615
グループ 3	0.327	0.639	0.708	0.325
グループ 4	0.466	0.365	0.227	0.466
グループ 5	0.715	0.782	0.680	0.702
グループ平均	0.533	0.503	0.490	0.548

表 4 各グループ内の設問ごとの AP

グループ	設問 ID	rel	read	spec	提案手法
グループ 1	0	0.443	0.279	0.372	0.483
	1	0.782	0.707	0.757	0.778
グループ 2	2	0.608	0.297	0.144	0.740
	3	0.481	0.173	0.399	0.490
グループ 3	4	0.248	0.722	0.708	0.227
	5	0.407	0.556	0.707	0.423
グループ 4	6	0.424	0.447	0.357	0.448
	7	0.508	0.282	0.097	0.483
グループ 5	8	0.753	0.755	0.658	0.738
	9	0.677	0.810	0.702	0.666

#### Mean Average Precision(MAP)

MAP は、求めた各グループの平均適合率 (AP) の平均値であり、グループ  $Q = \{q_1, q_2, \dots, q_n\}$  の  $MAP(Q)$  は、設問  $q$  の平均適合率  $AP(q)$  を用いて以下の式で算出した。

$$MAP(Q) = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \quad (6)$$

設問  $q$  の平均適合率  $AP(q)$  は、理解促進度 3 以上の画像を正解画像であるとみなして、設問  $q$  の全正解画像集合  $I_q = \{i_1, i_2, \dots, i_n\}$ 、正解画像  $i \in I_q$  の時点での適合率  $R(i)$  を用いて以下の式で算出した。

$$AP(q) = \frac{1}{|I_q|} \sum_{i \in I_q} R(i) \quad (7)$$

#### 4.3 結果

表 3 は、 $rel$  と  $1 - |read - 0.4|$  を説明変数とした回帰式でランキングしたものとベースラインの結果の各グループごとの MAP を表した表である表 3 の MAP 平均を比べてみると、提案手法の結果は、0.548 であり、適合性 (0.533)、可読性 (0.503)、専門性 (0.490) をベースとする手法よりも優れていた。

また、設問ごとのランキング評価結果である表 4 から、提案手法が有効に働いた例として設問 ID2、提案手法が有効に働かなかった例として設問 ID1, 9 がある。それらの提案手法によるランキング上位 3 件の画像をそれぞれ図 6~8 に示す。

#### 4.4 考察

提案手法をより詳細に分析するために、設問ごとのランキング結果と表 2 の相関分析表に注目した。

設問 ID9 のように理解促進度と適合性の間に相関がない時、この提案手法は有効に働かなかった。このケースでは、理解が促進したと判断された画像の周辺テキストではその手法の具体

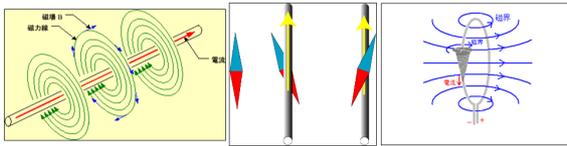


図 6 設問 ID1「右ねじの法則」における提案手法のランキング結果 上位 3 件の画像（左から順に 1 位，2 位，3 位の画像）

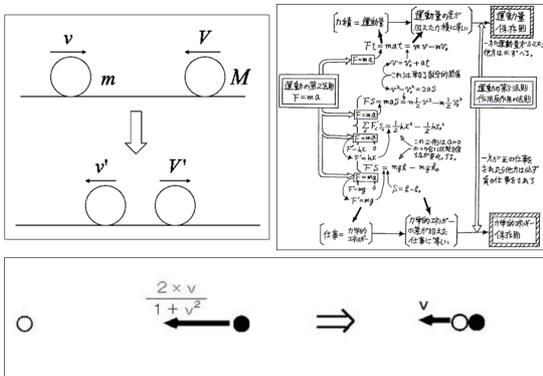


図 7 設問 ID2「運動量保存の法則」における提案手法のランキング結果 上位 3 件の画像（左上が 1 位，右上が 2 位，下が 3 位の画像）

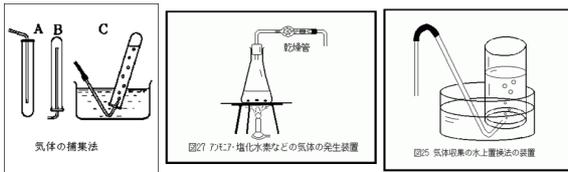


図 8 設問 ID9「水上置換」における提案手法のランキング結果上位 3 件の画像（左から順に 1 位，2 位，3 位の画像）

例について説明されていた．具体例について説明しているテキストは表層的に補充対象テキストと文章が異なっている．そういった画像は周辺テキストの適合性が低い画像であり，そのような理解が促進したと判断された画像を上位にランキングできなかったため提案手法は有効に働かなかったと考えられる．

また，設問 ID1 のように，今回注目したどの特徴量とも相関がほとんどない時も提案手法は有効に働かなかった．このケースでは理解が促進したと判断された画像は，対象について説明しているテキスト情報が画像中に書き込まれていたため周辺テキストに対象について本質的に説明しているテキスト情報の記述がほとんど存在していなかった．よって周辺テキストは本質的に説明していないテキスト情報が書かれており，本質的でない記述は各周辺テキストごとに異なっていたため，理解が促進した画像の周辺テキストの傾向の一貫性が見られなかった．理解が促進したと判断された画像に傾向が一貫しておらず，そのような理解が促進したと判断された画像をまとめて上位にランキングできなかったため提案手法は有効に働かなかったと考えられる．

一方，設問 ID2 のように提案手法が有効に働いた場合も存在

した．このケースでは，理解が促進したと判断された画像では周辺テキストで対象の内容を簡単すぎない程度に簡潔に説明された画像，つまり，周辺テキストの適合性が高く，可読性が中くらいの画像であった．よって，提案手法は理解が促進したと判断された周辺テキストの適合性が高く，可読性が中くらいの画像を上位にランキングできたため有効に働いたと考えられる．

## 5. ま と め

本稿では，理解を促進させる効果が高い画像を発見するための特徴として，画像の周辺テキストのみに着目し，周辺テキストと対象テキストの適合性，周辺テキストの可読性および専門性という 3 つの特徴量と，テキストに対する画像の理解促進効果との関係を分析した．また，分析結果を基に，理解が難しいテキストの理解を促進する画像の検索ランキング手法を提案した．

予備調査では，理解促進効果に関係のあるパラメータの発見を行うために理解促進度と，周辺テキストと対象テキストの適合性，周辺テキストの可読性および専門性という 3 つの特徴量の相関分析を行った．予備調査の結果，周辺テキストと対象テキストの適合性が高く，周辺テキストの可読性が中学校レベルに近い画像ほど，テキストに対する理解促進効果が高くなる傾向にあることが分かった．一方で，画像の周辺テキストの専門性と理解促進効果には強い関係は見られなかった．

これら予備調査の結果の知見から，理解促進を支援する画像のランキング関数に用いる説明変数を決定し，その説明変数を使った線形回帰分析を行いランキング関数を構築した．本稿では，適合性，可読性，専門性のみに着目したランキング関数をベースラインとし，提案したランキング関数の評価を評価を行った．実験の結果，提案手法は合計 10 個のテストセット平均では他のベースライン結果よりも優れた検索精度を示した．しかし，設問によっては提案手法が他のベースラインよりも精度が劣ることがあった．

提案手法は画像の周辺テキストの適合性が高く，可読性がある値に近づくほど，画像の理解促進効果が高いと評価するものであった．各設問ごとの実験結果と予備調査結果を照合したところ，設問によっては適合性の値とは無関係に可読性が高いほど理解促進効果が高いケース，可読性の値とは無関係に適合性が高いほど理解促進効果が高いケースがあることが分かった．テキスト中で画像がどのように使用されるかはケースによって異なる．例えば，ティティウスボーデの法則のように説明に使われている語句が難しくなくても，テキストを読んでも内容が想像しにくいテキストには，内容を具象化した画像が通常補充されることが考えられる．もし，そのような説明テキストに画像を補充したい場合は，説明テキストの適合性のみを考慮すれば良いと考えられる．一方で，水上置換法の説明テキストのように，テキストだけ読んでもある程度内容を理解できるが，追加情報を与えるために画像が補充されるケースもある．そのような場合，可読性がある値に近づくほど理解促進効果が高い傾向があることが分かった．また，説明テキストによっては適合性，可読性，専門性のいずれとも相関関係がないケースも見受

けられた。これらから、説明テキストの内容、ジャンルによって補完すべき画像の特徴が異なることが考えられる。

説明テキストによって理解促進度と相関がある説明変数は異なるため、Web ページを閲覧中に分かりにくいテキストの理解を促進させる画像の補完を行うブラウザはある程度様々なカテゴリのテキストに対して適用出来なければならない。このため、対象によってランキング関数に使う説明変数や重みを変更できるようにシステムを作る必要がある。そのために、まずは対象によってランキング関数に使う説明変数を判別できるようなランキング関数を作るために説明テキストを傾向ごとにカテゴリ分けして個々のカテゴリごとに特徴を分析することを行う予定である。そして、カテゴリごとの分析が十分にできたのちにランキング関数に使う説明変数や重みを説明テキストによって動的に決定するシステムを構築する予定である。また、画像自身の特徴量など理解促進効果に関係する新しい特徴の発見も検討していく予定である。

謝辞 本研究の一部は、グローバル COE 拠点形成プログラム「知識循環社会のための情報学教育研究拠点」(拠点リーダー: 田中克己) によるものです。ここに記して謝意を表します。

## 6. 参考文献

### 文 献

- [1] R. Agrawal, S. Gollapudi, A. Kannan, and K. Kenthapadi. Enriching Textbooks with Images. In *Proceedings of the 20th ACM international conference on Information and knowledge management (CIKM 2011)*, pages 1847–1856. ACM, 2011.
- [2] K. Akamatsu, N. Pattanasri, A. Jatowt, and K. Tanaka. Measuring Comprehensibility of Web Pages Based on Link Analysis. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI 2011)*, pages 40–46, 2011.
- [3] J. Chall. *Readability revisited: The new Dale-Chall readability formula*. Brookline Books Cambridge, MA, 1995.
- [4] Z. Li, S. Shi, and L. Zhang. Improving Relevance Judgment of Web Search Results with Image Excerpts. In *Proceedings of the 17th international conference on World Wide Web (WWW 2008)*, pages 21–30. ACM, 2008.
- [5] R. Mayer. *Multimedia Learning*. Cambridge University Press, 2001.
- [6] M. Nakatani, A. Jatowt, H. Ohshima, and K. Tanaka. Quality Evaluation of Search Results by Typicality and Speciality of Terms Extracted from Wikipedia. In *Proceedings of the 14th International Conference on Database Systems for Advanced Applications (DASFAA 2009)*, pages 570–584, 2009.
- [7] M. Nakatani, A. Jatowt, and K. Tanaka. Easiest-first Search: Towards Comprehension-based Web Search. In *Proceedings of the 18th ACM conference on Information and knowledge management (CIKM 2009)*, pages 2057–2060. ACM, 2009.
- [8] X. Yan, D. Song, and X. Li. Concept-based Document Readability in Domain Specific Information Retrieval. In *Proceedings of the 15th ACM international conference on Information and knowledge management (CIKM 2006)*, pages 540–549. ACM, 2006.
- [9] T. Yeh, B. White, J. San Pedro, B. Katz, and L. S. Davis. A Case for Query by Image and Text Content: Searching Computer Help using Screenshots and Keywords. In *Pro-*