

場所に焦点を当てた複数旅行ブログの自動要約

前田 剛[†] 難波 英嗣[†] 竹澤 寿幸[†]

[†] 広島市立大学大学院 情報科学研究科 〒731-3194 広島県広島市安佐南区大塚東 3-4-1

E-mail: [†] {maeda, nanba, takezawa}@ls.info.hiroshima-cu.ac.jp

あらまし 我々は、ブログ集合から旅行ブログエントリを自動的に検出し、さらに各エントリを地図上に表示するシステムを構築している。このシステムでは、ある地点を訪れた複数の旅行者が、それぞれ旅行ブログエントリを書いた場合、その地点に旅行ブログエントリが重なって表示されてしまうため、その地点で何が出来るのかがユーザにとって分かりにくいという問題があった。そこで、本研究では、ある地点に関する複数の旅行ブログエントリからひとつの要約を作成する。近年、テキスト要約研究分野では、LexRank と呼ばれるグラフベースの要約作成技術が広く使われるようになってきている。本研究では、この手法を拡張し、テキストだけでなく画像も対象とした要約作成手法を提案する。

キーワード 複数テキスト要約, 画像, 旅行, ブログ

1. はじめに

インターネット上には、多くの観光情報が存在しており、その一つとして旅行ブログエントリがある。旅行ブログエントリは、土産物、観光名所、旅行の際に参考にした Web ページへのリンクなど様々な観光情報を含んでいる。そのため、旅行ブログエントリは観光情報を得るための有益な情報源であると考えられる。旅行ブログエントリを、観光情報を提供するための情報源として利用した研究として、藤井ら[1]の研究がある。藤井らは旅行ブログエントリの観光タイプの自動判定を行い、地図上にマッピングし、ユーザへ提示するシステムを構築している。このシステムは、広島 p2walker で公開されている“ぶらり広島電停散歩 MAP”¹ に使用されている。図 1 にこのシステムの動作例を示す。図 1 は、“見る” (図中①) を選択した場合の例であり、ピン (図中②) を選択するとそのピンの場所について記載された“見る”に関する旅行ブログエントリのタイトルが図中③に表示される。そして、そのタイトルをクリックすると旅行ブログエントリを閲覧することができる。このシステムを利用することで、地図上にマッピングされた旅行ブログエントリから、目的の地点に関する観光情報を容易に検索することができる。さらに図 1 のシステムを拡張した世界版の観光情報システムも構築し、現在公開されている。しかし上記のシステムでは、地点によっては様々なトピックを含んだ複数の旅行ブログエントリが混在してしまい、その地点の特徴を捉えにくいという問題点がある。例として、宮島の場合、“厳島神社”、“宮島水中花火大会”、“紅葉狩り”などの様々なトピックを含んだタイトルの旅行ブログエントリが 1 つの地点に集約してしまうため、時期毎の見所などを見落とす可能性がある。

この問題を解決するために本研究では、LexRank と



図 1: ぶらり広島電停散歩 MAP の動作例

呼ばれるグラフベースのテキスト要約技術を用いて複数旅行ブログエントリの要約を行う。複数の旅行ブログエントリに対して要約を行うことで、地点毎に混在したトピックの要点を簡潔にまとめることが可能である。さらにテキストだけでなく画像も要約の対象とし、その地点に関する旅行ブログエントリの要約を重要画像付きで表示するシステムの構築する。これにより、旅行者は、目的の地点の特徴を視覚的かつ迅速に捉えることができると考えられる。テキストと画像を対象とした要約作成を試みる先行研究として、前田ら[2]の研究がある。前田らは、ソフトクラスタリングツール“bayon”²を利用し、テキストと画像のクラスタリングを行い、旅行ブログエントリの要約を生成している。しかし、前田らは要約生成の際に要約技術を使用しておらず、要約の精度そのものが十分ではなく、また 1 文 1 画像の要約しか生成できない。本研究では、テキスト要約研究において標準的な手法である LexRank を用いて要約を生成する。さらに、地点毎に 5 つのテキストと画像の要約を表示する。

¹ <http://p2walker.jp/peace/ja/blog/>

² <https://code.google.com/p/bayon/>



図 2：システム動作例
(地点：宮島，観光タイプ：食べる)



図 3：システム動作例
(地点：宮島，観光タイプ：見る)



図 4：システム動作例
(地点：宮島，観光タイプ：見る，時期：7月～9月)

本論文の構成は以下の通りである。2 節では、システムの概要および動作例，3 節では、関連研究，4 節では、旅行プログエントリの要約，5 節では、実験，6 節では、考察について述べ，7 節で本稿をまとめる。

2. 旅行プログエントリ要約システムの概要および動作例

本節では、本研究で構築した旅行プログエントリ要約システム(以下、システム)の概要および動作例について説明する。まずシステムの概要について述べる。本研究で構築したシステムでは、“地点”，“観光タイプ”，“時期”を入力すると、該当する旅行プログエントリを 10 件抽出し、重要文、重要画像の選出を行い、その後、重要文と重要画像の対応付け、要約を生成する。

次に、システムの動作例について説明する。本システムでは、藤井ら[1]が提案する 6 つの観光タイプの内、“見る”，“食べる”，“買う”，“体験する”，“泊まる”の 5 つの観光タイプを考慮し、異なる要約結果を出力することが可能である。図 2 と図 3 は、宮島における観光タイプ“食べる”，および“見る”に関連する旅行プログエントリを対象とした場合の動作例である。図 2 では、“牡蠣”に関するテキストや画像が表示され、

一目で宮島では、牡蠣が有名であると判断することができる。また同様に図 3 は、“厳島神社”や“大鳥居”に関する要約が表示され、宮島の見所を迅速に捉えることができる。さらに本研究で構築したシステムでは、図 4 のように、時期を設定することで同じ地点、同じ観光タイプであっても異なる要約を得ることができる。図 4 では、宮島において 7 月～9 月に行われる行事として有名な“宮島水中花火大会”や“管弦祭”に関する要約が得られている。以上のように、同じ地点の入力においても、観光タイプや時期によって異なる要約結果を得ることができる。

3. 関連研究

本研究の関連研究として、3.1 節では、旅行プログエントリや観光支援サービスに関連する研究，3.2 節では、テキスト要約技術について述べる。

3.1. 旅行プログエントリや観光支援サービスに関連する研究

旅行者の過去の経験を収集する場としては、楽天トラベル³のような宿泊施設の予約サイトや、食ベログ⁴のような飲食店の口コミサイトがある。このようなサイトでは、一件の宿泊施設や飲食店に対し、その個別の施設に関する口コミしか得ることができない。そのため、本研究では、利用する情報源として、旅行プログエントリを採用した。

旅行プログエントリからは、複数の観光名所や宿泊施設に関連する情報の他に、以下のような情報を得ることができると考えられる。

- 季節や、天候に応じたお勧めの観光情報
- イベント等の観光情報

そこでプログデータベースから旅行プログエントリを自動的に収集する手法として、Nanba ら[3]や Ishino ら[4]の手法がある。Nanba らは、一般ブログから、機械学習を使用して旅行プログエントリを自動的に検出

³ <http://travel.rakuten.co.jp/>

⁴ <http://tabelog.com/>

する手法を提案している。機械学習の手法には、CRFを採用し、精度 0.867 と高い精度で旅行ブログエントリの検出に成功している。Ishino らは、広島県の特徴の一つである、広島電鉄の電車(広電)を使用した観光を支援するための枠組みの一つとして、広電の電停に関する旅行ブログエントリを収集する手法を提案している。Ishino らは、0.824 と高い精度で電停に関する旅行ブログエントリの検出に成功している。本研究では、Nanba らおよび Ishino らの手法により収集した旅行ブログエントリを使用する。

藤井ら[1]は、旅行者が必要とする情報へのアクセスを容易にするために、6種類の観光タイプ毎に旅行ブログエントリを提示するシステムである、“ぶらり広島電停散歩 MAP”を構築している。このシステムでは、収集した旅行ブログエントリを地図上にマッピングすることで、どこに関する旅行ブログなのかが視覚的に分かるようになっていく。しかし、1点に複数の旅行ブログエントリがマッピングされると、必要とする情報を得ることが非常に困難になる。そこで本研究では、藤井らの提案する観光タイプの内、“見る”、“食べる”、“買う”、“体験する”、“泊まる”の5つの観光タイプを考慮し、地点毎に重要文とそれに対応する重要画像を提示する要約の生成を試みる。これにより、旅行者の必要とする情報をより迅速に得ることが可能になる。

3.2. テキスト要約技術

自動要約技術としては、多くの手法がこれまでに提案されてきている一方で、文間の類似度をグラフ表現したものを用いる手法が高い精度でテキスト要約を行えることが知られている[5, 6, 7]。Erkan ら[5]および Mihalcea ら[6]は、対象となるテキストの概要をまとめた要約生成を行っており、前者では複数テキストを、後者では単一テキストを対象としている。加えて、Otterbacher ら[7]は、クエリに特化した要約生成を行っている。テキスト要約技術において PageRank アルゴリズムを適用している代表的かつ初期の研究としては、LexRank がある。LexRank は Erkan らにより提案された、PageRank[8]の概念に基づいた複数テキスト要約手法である。LexRank は、対象テキスト内の文のグラフ表現における固有ベクトル中心性の概念に基づいて文の重要度を計算する手法である。これは、次数の多いノードを評価するだけでなく、次数の多いノードと隣接しているノードの重要度についても考慮し、その文に比例して対象ノードを評価することができる。この手法では、文間のコサイン類似度に基づいた連結性行列が文のグラフ表現の隣接行列として使われており、その隣接行列の第一固有ベクトルの成分を各ノードの中心性を表すスコアと考える。LexRank は、実際には上述の処理のみで要約文を生成するのではなく Radev

ら[9]の提案した要約システムである MEAD⁵の内部に組み込み、冗長性削減のための指標などと組み合わせることで要約文を生成することを前提としている。本研究では、Erkan らの提案する LexRank を用いて複数テキスト要約を行い、画像に対しても PageRank アルゴリズムの適用する。さらに、LexRank に従って文を抽出していくと、冗長性のある要約文が生成される可能性がある。これに対し、MMR(Maximal Marginal Relevance)[10]を適用することで、類似文及び類似画像の抽出を防ぐ。

4. 旅行ブログエントリの要約

4.1. システムの構成

現在、図 1 で示した“ぶらり広島電停散歩 MAP”を使用する際、地点によっては様々なトピックを含んだ複数の旅行ブログエントリが混在してしまい、その地点の特徴を捉えにくいという問題点がある。例として、宮島の場合、“厳島神社”、“宮島水中花火大会”、そして“紅葉狩り”などの様々なトピックを含んだ複数の旅行ブログエントリが1つの地点に集約してしまうため、時期毎の見所などを見落とす可能性がある。この問題を解決するために、本研究では、グラフベースのテキスト要約技術を用いて、複数旅行ブログエントリの要約を行うことで、これらの問題点の改善を試みる。図 5 に本研究で構築した、システムの構成を示す。図 5 における提案システムでは、「複数旅行ブログエントリからの重要文抽出」、「複数旅行ブログエントリからの重要画像抽出」、そして「要約生成」から構成されている。「複数旅行ブログエントリからの重要文抽出」は、4.2 節で、「複数旅行ブログエントリからの重要画像抽出」は 4.3 節で、「要約生成」は 4.4 節で説明する。

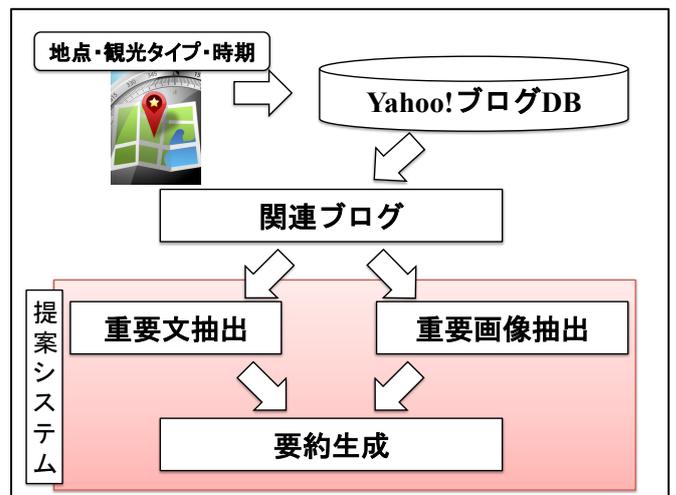


図 5: システムの概要

⁵ <http://www.summarization.com>

4.2. 複数旅行ブログエントリからの重要文抽出

本節では、旅行ブログエントリからの重要文抽出について説明する。本研究では、Erkan ら[5]の提案する LexRank を用いて重要文の抽出を行う。LexRank は、現在のテキスト要約研究においてベースライン手法として使用されている標準的な手法である。重要文抽出の手順を以下に示す。

(1) 文間の類似度

文間の類似度の計算には、旅行ブログエントリを一文に分割した文に対して MeCab⁶により形態素解析し、品詞が名詞、動詞、形容詞である単語のみを使用する。本研究では、類似性尺度としてコサイン類似度を用いる。また単語の重みには、tf*idf を用いる。表 1 は、文間の類似度を要素として持つ接続行列の一例である。表 1 の例では、総文数は 7 であり、S₀ は対象テキスト群に含まれる文のうち 0 番目の文を表す。

表 1: 文間の類似度

	S ₀	S ₁	S ₂	S ₃	S ₄	S ₅	S ₆
S ₀	1.00	0.27	0.24	0.34	0.30	0.30	0.68
S ₁	0.27	1.00	0.27	0.01	0.05	0.01	0.38
S ₂	0.24	0.27	1.00	0.09	0.12	0.18	0.21
S ₃	0.34	0.01	0.09	1.00	0.14	0.03	0.19
S ₄	0.30	0.05	0.12	0.14	1.00	0.16	0.11
S ₅	0.30	0.01	0.18	0.03	0.16	1.00	0.27
S ₆	0.68	0.38	0.21	0.19	0.11	0.27	1.00

類似度グラフは文間の類似度を要素とした接続行列に基づいて生成される。例えば、図 6 は表 1 の類似度グラフを示しており、各ノードは文を表し、各エッジの重みは文間の類似度を表している。ここで、類似度グラフは文間の類似度を重みとした重みつきグラフとして示されている。

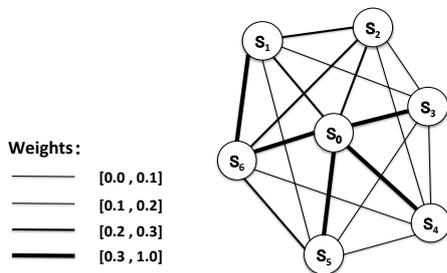


図 6: 類似度グラフ

(2) 文の重要度

次に、生成された類似度グラフに対して、固有ベクトル中心性に基づいた各文の重要度を計算する。文 u の重要度は、Erkan らの手法を参考にして、式(1)で求

められる。

$$p(u) = \frac{d}{N} + (1-d) \sum_{v \in \text{adj}[u]} \frac{\text{sim}(u,v)}{\sum_{z \in \text{adj}[v]} \text{sim}(z,v)} p(v) \quad (1)$$

- u, v, z : 文
- $p(u)$: 文 u の重要度
- N : 対象としているテキスト群の総文数
- $\text{adj}[u]$: 文 u の隣接ノード集合
- d : 非隣接ノードとの類似度を考慮するための制動係数 ($d = 0.85$)
- $\text{sim}(u, v)$: 文 u と文 v 間の類似度

次に、重要度を要素とした行列に対してべき乗法を用いて第一固有ベクトルを計算する。これにより、中心性の高い文と類似している文の重要度を求めることで重要文を抽出する。

(3) 冗長性の削減

重要文は、重要度の高い文が抽出されているため、冗長性のある重要文が抽出される可能性がある。この問題を避けるために、本研究では、MMR(Maximal Marginal Relevance)[10]を適用することで、冗長性の削減を行う。この指標は、新しく抽出された文と既に抽出された文との間の類似度に対応するペナルティ値を与えることにより、類似した文の抽出を防ぐことができる。以上により得られた重要文を本研究における重要文として用いる。なお、重要文抽出の段階で、10 語以下の文は省いている。

4.2. 複数旅行ブログエントリからの重要画像抽出

本節では、旅行ブログエントリからの重要画像抽出手法について述べる。本研究では、重要文抽出と同様に、Erkan ら[5]の提案した LexRank を適用することで旅行ブログエントリ中の各画像の重要度を測る。画像間の類似度として Color Histogram のヒストグラム値を要素とし、コサイン類似度を適用する。さらに、重要文抽出の際と同様に、MMR[10]を適用することで、類似画像の抽出を防ぎ、画像のランク付けを行う。ここで、本研究で用いる Color Histogram には、HSV の値を使用する。HSV は、H(Hue/色相), S(Saturation/彩度), V(Value/明度)をそれぞれ表している。色相はスペクトル上での色の位置を表しており、0 度から 360 度の角度により表される。彩度は色の鮮やかさを表しており、明度は色の明るさを示している。本研究では、H, S, V の値をそれぞれ 10, 4, 4 分割することで、160 色に減色させ、ヒストグラム値を計算する。

⁶ <http://mecab.sourceforge.net/>

表 2：旅行プログエントリの地点リスト

地点	観光タイプ(件)				
	見る	食べる	買う	体験する	泊まる
北海道(札幌市)	5	3	4		
北海道(函館市)	6	3	2	1	
東京(葛飾柴又)	4	4	3		
東京(墨田区)	10				
東京(上野)	3	7			
栃木	9	2			1
神奈川(横浜市)	3	7		1	
神奈川(小田原市)	9	2	1	2	
長野	9			1	
愛知	8	2			
富山	9	1			
京都	9	1	1		
奈良	9		1		
和歌山	10				
大阪	6	4			
兵庫	9	1	2		
広島(広島駅)		10			
広島(平和記念公園)	9			1	
広島(宮島)	2	9	1	7	
下関	3	7	1		
福岡	8	3			
長崎	6	3		1	2
計	146	69	16	13	3

4.3. 複数旅行プログエントリの要約生成

本節では、4.1 節と 4.2 節で得られた、重要文と重要画像の対応付け、および要約生成について説明する。まず、重要文と重要画像の対応付けについて述べる。旅行プログエントリに含まれる文と画像の関係として、画像と距離の近い文はその画像を説明している可能性が高い。そこで、4.2 節でランク付けを行った画像と画像が含まれる旅行プログエントリにおいて、画像の前後に存在する一文の対応付けを行う。そのために、文に対して形態素解析を行い、名詞句を抽出する。その後、抽出した名詞句と画像を対応付ける。次に、4.1 節により抽出された重要文に対しても形態素解析を行い、名詞句を抽出する。この時、重要文に含まれる名詞句と画像と対応付けた名詞句を上位から比較し、一致した場合、重要文と重要画像の対応付けを行う。しかし、重要文に含まれる名詞句と重要画像に対応付いた名詞句が一致しない場合、対応付けることができない重要文は省く。したがって、4.1 節で抽出した重要文と画像と対応付けることで抽出した重要文は異なる可能性がある。以上の要約生成までの流れを図式化したものを図 7 に示す。

最後に、要約生成として上記の手順で対応付いた文と画像の上位 5 件を旅行プログエントリ要約システムの入力結果として表示する。

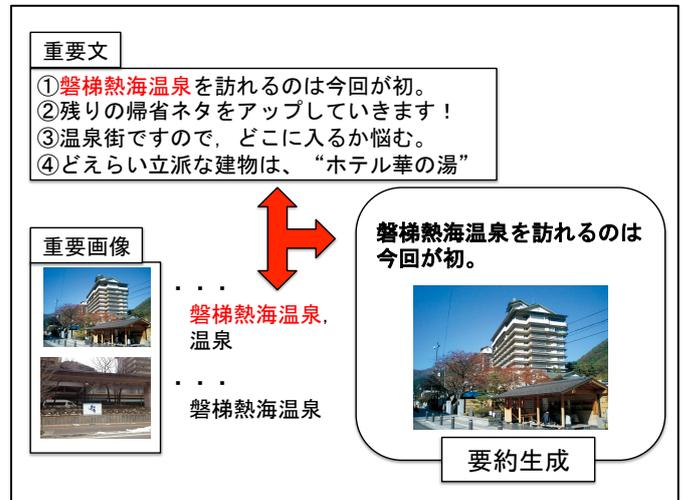


図 7：重要文と重要画像の対応付けおよび要約生成手順

5. 実験

5.1. 実験設定

実験データとして、Nanba ら[3]および Ishino ら[4]の手法により収集した旅行プログエントリの中から、22 地点各 10 件の旅行プログエントリ 220 件を手手で選定し、実験に用いた。本研究で用いた旅行プログエントリの地点リストを表 2 に示す。また本実験では、旅行プログエントリに対して LexRank[5]を適用した結

果をベースライン手法(LexRank)とし、重要文に対して画像を対応付けることで抽出した結果を提案手法(LexRank+IMG)として比較実験を行った。評価指標としては、それぞれによって抽出された要約に対して、ROUGE[11]を適用し、ROUGE-1とROUGE-2のスコアを比較した。

5.2. 実験結果

表3に実験結果を示す。表3より、ROUGE-1では0.022、ROUGE-2では0.018の向上が見られ、提案手法の有効性を示すことができた。

表3：重要文抽出の実験結果

	ROUGE-1	ROUGE-2
LexRank	0.167	0.043
LexRank+IMG	0.189	0.061

6. 考察

本節では、重要文抽出の実験結果に対して考察を行う。本研究において重要文抽出を行う際、LexRankを利用した手法とLexRankにより得られた重要文に対して重要画像を対応付けることで、再度抽出した重要文で比較を行っている。図8に、上記の2つの例を示す。後方で重要文を抽出する場合、4.3節でも述べたように、重要文から名詞句を抽出した際、そこに含まれる名詞句が重要画像と対応付けたものと一致しない場合がある。本研究では、このような対応付かない文は要約生成の段階で省いている。表3において得られた結果から、重要文と重要画像の対応付けを行うことで、不要な文が排除されたため、結果が向上したと考えられる。

また更なる結果の向上のために、情報利得を用いた手法が挙げられる。情報利得とは、「ある単語の出現の有無」の情報が、クラスに関する曖昧さ(エントロピー)をどれくらい減少させるかを示す値である。曖昧さを減少させる単語を選択することにより、クラスの特徴が容易になり、情報利得を用いて収集した単語は、クラス分類において、有効な単語であると言える。そこで、本研究で使用する5つの観光タイプ、“見る”、“食べる”、“買う”、“体験する”、“泊まる”のクラス毎に情報利得により収集した手掛かり語を重要文抽出の際に組み合わせることで、結果を向上させることができると考えられる。

7. おわりに

本研究では、グラフベースのテキスト要約技術を利用し、テキストだけでなく画像も対象としたマルチメディア要約を行った。さらに、その地点における重要画像付きの要約を提示するシステムを構築した。旅行ブログエントリの要約を行う手順として、以下の4つのステップで行った。旅行ブログエントリの要約を生成する手順として、複数の旅行ブログエントリからの

LexRank	
①	そんな大事故があつて間もない時期だったので、ここだけでなく近隣の花火大会は軒並み中止となりました。
②	終了後の大混雑で駅まで徒歩で約1時間、そして帰りの電車内から外の土砂降りに2度目の「ホッ!」と云う事で、相変わらずの絵ですがUP出来ましたのでどうぞ。
③	大阪のホテルに荷物を置き、早速ルミナリエに行くべく元町駅へ向かいました。
④	夕焼けの空って綺麗けど寂しいねえ。
⑤	三連休は初日からエンジョイしてるま〜ぶ〜は・・・笑昨日は母とバイキング
LexRank+IMG	
①	そんな大事故があつて間もない時期だったので、ここだけでなく近隣の花火大会は軒並み中止となりました。
②	終了後の大混雑で駅まで徒歩で約1時間、そして帰りの電車内から外の土砂降りに2度目の「ホッ!」と云う事で、相変わらずの絵ですがUP出来ましたのでどうぞ。
③	大阪のホテルに荷物を置き、早速ルミナリエに行くべく元町駅へ向かいました。
④	三連休は初日からエンジョイしてるま〜ぶ〜は...笑昨日は母とバイキング。
⑤	神戸市の紋章なんです。

図8：2パターンで抽出した重要文比較
(地点：兵庫，観光タイプ：見る)

重要文抽出、複数の旅行ブログエントリからの重要画像抽出、重要文と重要画像の対応付け、旅行ブログの要約生成の4つのステップで行った。有効性を確認するための実験の結果、画像を対応付けることで抽出した重要文において結果が向上し、提案手法の有効性を示した。

最後に、本研究では、図1のような従来システムにおいて、旅行ブログエントリのタイトルのみを表示するのにに対して、旅行ブログエントリ要約システムを構築した。本システムにより、テキストや画像を利用することで、視覚的かつ迅速に旅行ブログエントリの内容を理解することが可能になった。

参考文献

[1] 藤井一輝, 石野亜耶, 藤原泰士, 前田剛, 難波英嗣, 竹澤寿幸, “多言語旅行ブログエントリを用いた観光情報提示システム”, 第6回データ工学と情報マネジメントに関するフォーラム(DEIM 2014), 2014.

[2] 前田剛, 河野有希, 石野亜耶, 難波英嗣, 竹澤寿幸, “場所に焦点を当てた複数ブログの自動要約”, 第6回データ工学と情報マネジメントに関するフォーラム(DEIM 2014), 2014.

- [3] H. Nanba, H. Taguma, T. Ozaki, D. Kobayashi, A. Ishino and T. Takezawa, "Automatic Compilation of Travel Information from Automatically Identified Travel Blogs", Proc. of the Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Proceeding, Short Paper, pp. 205-208, 2009.
- [4] A. Ishino, H. Nanba, and T. Takezawa, "Construction of a System for Providing Travel Information along Hiroden Streetcar Lines", Proc. of the 3rd IIAI International Conference on e-Services and Knowledge Management (IIAI ESKM 2012), 2012.
- [5] G. Erkan and D. R. Radev, "LexRank: Graph-based Lexical Centrality as Saliency in Text Summarization", Journal of Artificial Intelligence Research, pp. 457-479, 2004.
- [6] R. Mihalcea and P. Tarau, "TextRank: Bringing Order into Texts", Proc. of the 2004 Conference on Empirical Methods in Natural Language Processing, pp. 401-411, 2004.
- [7] J. Otterbacher, G. Erkan and D. R. Radev, "Using Random Walks for Question-focused Sentence Retrieval", Proc. of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, pp. 915-922, 2005.
- [8] S. Brin and L. Page. "The Anatomy of a Large-scale Hypertextual Web Search Engine", Computer Networks and ISDN Systems, pp. 107-117, 1998.
- [9] D. R. Radev, S. B. Goldensohn and Z. Zhang, "Experiments in Single and Multi-document Summarization using MEAD", First Document Understanding Conference New Orleans, 2001.
- [10] J. Goldstein, V. Mittal, J. Carbonell and M. Kantrowitz, "Multi-document Summarization by Sentence Extraction", Proc. of the 2000 NAALP-ANLP Workshop on Automatic Summarization, pp. 40-48, 2000.
- [11] C. Y. Lin, "ROUGE: a Package for Automatic Evaluation of Summaries", Proc. of Workshop on Text Summarization Branches Out, pp. 74-81, 2004.