

研究課題名：旅行情報の連想検索システムの研究

政策メディア研究科 修士課程 2 年

メール: dejavu@sfc.keio.ac.jp

金澤 裕也

1. 研究概要

近年、インターネットで旅行情報を検索する旅行者が増加している。しかし、既存の旅行情報検索システムでは、行きたい旅行地のイメージを検索結果に反映させることができないので、行きたい旅行地が抽象的な段階での検索には向かない。よって、インターネットと比較して、旅行代理店の方がツアーの購買に繋がりやすい傾向がある[1]。

この課題を解決するためには、旅行動機などのイメージによる旅行情報の検索を実現しなければならない。しかし、これらのイメージは、言語化が困難であることから、通常のメニュー検索やテキスト検索等の方式は不適當である。そこで、本研究では、行きたい旅行地を、写真の印象を元に連想検索するシステムを提案する。そして、本システムにより、電子市場における旅行商品の購買意欲向上が実現されることを目標とする。

2. システムの概要

本研究では、写真の印象を元にツアーを連想検索するシステムについて提案する。連想検索とは、コンテンツ間の類似度を元に、コンテンツを検索する方式である。本システムでは、コンテンツ間の類似度を旅行地の印象を元に定義する。そのために、各写真に、旅行地の印象を表す言葉（印象語タグ）を成分とする意味ベクトルを付与し、この類似度を元にコンテンツ間の関連性を表す。次に連想検索の流れについて記す。

はじめに、旅行条件を元に初期のクエリ候補を設定する。検索者はクエリ候補の写真の中より、気になる写真を選択する。その後、検索者が選択した写真のベクトル Q と検索対象の写真のベクトル R_i の類似度を式(1)により算出し、類似度が近い写真 4 枚を検索結果として表示する。そして、これらの検索結果を次のクエリ候補として検索を繰り返す。最終的に、気になる検索結果が見つかった際には、予約用のウェブサイトからツアーを注文する。

$$S = \frac{QR_i}{\|Q\| \|R_i\|} \quad (1)$$

3. 昨年度までの実施内容

昨年度は、印象語タグとして適切なタグを抽出する作業を行なった。本研究では、Mikaの研究を参考に、フォーカスグループインタビューにより印象語の抽出を行なった[2]。インタビューでは、写真のみで構成した仮想のパンフレットを41枚作成し、写真の旅行地に行った際に、感じる印象について被験者に尋ねた。インタビューは、昨年度の1月から2月に掛けて、20代の学生を中心に計4回実施した。その結果、旅行地の印象を表す語として適切な語を43語抽出した。

さらに、これらの語の一般性を検証するために、インターネット調査を行なった。調査では、旅行パンフレットの写真をKJ法により分類した結果から策定した、代表的な旅行地の写真30枚に対して、フォーカスグループインタビューで抽出された38語の中より、印象としてあてはまる語を選択してもらった。調査は、首都圏在住の被験者1000人（男性500人、女性500人）に対して行なわれた。その結果、多くの語が200人以上に使用され、各語の一般性が検証された。しかし、一部の語については、使用数が少なかったことから省き、その結果、図1に示す38語を、意味ベクトルに使用する印象語として採用する事とした。

活動的な	ゆっくりとした	郷愁漂う	健康的な	日本的な
特別感のある	風情がある	トロピカルな	文化的な	おいしい
探求心をそそる	ほのぼのとした	フレンドリーな	安心な	
上質な	季節感のある	おしゃれな	やすらぎ	
プライベートな	うきうきする	ドキドキする	本格的な	
静かな	風光明媚な	貴重な	厳粛な	
異国的な	家庭的な	幻想的な	都市的な	
爽快な	庶民的な	わいわいとした	達成感のある	
じっくりと	モダンな	威勢の良い	乙女な	

図 1 印象語タグ

4. 今年度の実施内容

昨年度は、印象語タグに使用するために適切なタグを抽出する作業を行なった。今年度は、昨年度の結果を元に、システムの実装と評価を行なった。また、インターネットの調査の結果を元に、性別・年齢毎の印象語空間の違いについて検証を行なった。

4.1 連想検索システムの実装

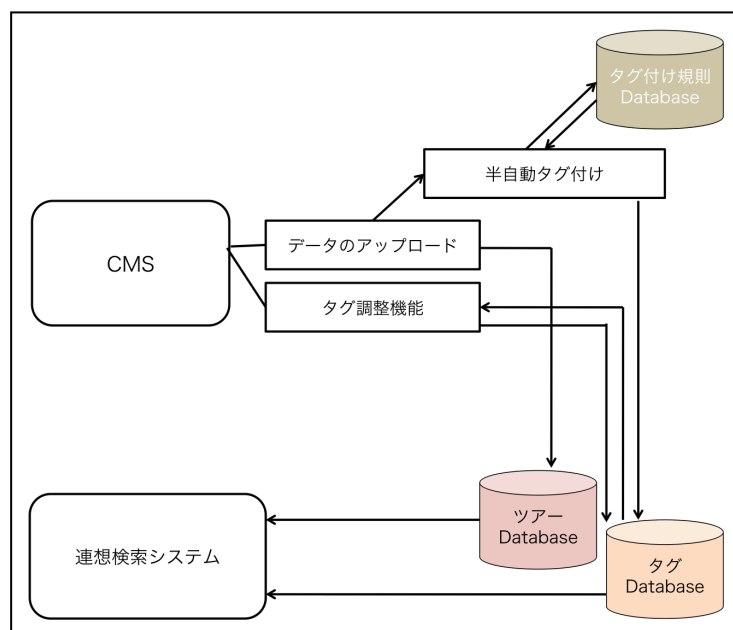


図 2 システムの構成図

2章で述べた方式で、ツアーの連想検索システムを構築した。システムはウェブシステムの形でPHP言語を用いて作成した。検索対象のツアーには、各ツアーの個性が明確である事、写真が豊富である事から、JR東日本の旅市を採用した。システムの構成を図2に示した。システムは連想検索システムと、タグ付けや情報の編集を行なうCMSに分かれている。

システムの実装に当たっては、各ツアーの写真にフォーカスグループインタビューにて抽出した印象語タグ38個を5段階で付与する必要がある。しかし、このタグ付け作業に、莫大な工数が掛かるという問題点がある。そこで、工数の削減を図るために、一部のタグをシステムで自動的に付与する半自動タグ付けシステムを実装する事にした。システムでは、写真と印象語タグの対応関係を元にタグを自動的に付与する。そのため、写真と印象語との対応関係を、印象評価実験の結果を元にして、主成分分析から求めた。

分析の結果、5つの主成分が抽出され、各主成分スコアが高い写真から、温泉観光地、歴史観光地、都市観光地、自然観光地、食の旅行地の写真の種類を表すと解釈された。その後、各写真のジャンルに対応する語を、因子負荷量を元に求めた。5段階相対評価で因子負荷量の値を5段階に変換した際に、4と5の値を取る語を各写真の種類と対応する語とした。表1に、各主成分と写真、語の対応を記した。

次に、この対応関係の正しさを検証するために、簡単な検証実験を行なった。実験では、旅行者4人に、表1の対応関係を元に、旅市のツアーの写真にタグ付けしたデータを修正してもらった。その結果、修正率が2割程度であり、ある程度、精度が高い事が確認された。これより対

応関係の正しさが検証されたため、表 1 の対応関係を元に、自動的にタグを付与する半自動タグ付けシステムを構築した。

システムでは、写真をアップロードする際に、上記の写真の種類を選択すると、該当するタグが自動的に付与される。その後は、タグ調整画面で、手動でタグの調整や付与を行なう。

シミュレーションから、半自動タグ付けシステムによって、タグ付け作業のコストを 2～3 割削減する事が可能であると、推察された。

表 1 主成分と印象語・写真の対応

タグ付けのジャンル	関連する語	代表的な写真
温泉観光地	プライベートな/やすらぎ/ゆっくりとした/上質な/特別感のある/じっくりと/貴重な/日本的な/風情のある/文化的な	
歴史観光地	貴重な/文化的な/本格的な/探求心をそそる/威勢の良い/上質な/特別感のある/活動的な/厳粛な/日本的な	
自然観光地	健康的な/やすらぎ/ゆっくりとした/ドキドキする/活動的な/季節感のある/静かな/探求心をそそる/風光明媚な	
都市観光地	おしゃれな/都市的な/モダンな/プライベートな/上質な/異国的な/うきうきする	
食	おいしい/季節感のある/郷愁漂う/威勢の良い/じっくりと/本格的な/うきうきする/日本的な/風情がある	

以上のような、半自動タグ付けシステムを利用して、写真に対してタグを付与し、タグデータベースを作成した。データベースの構築後、システムのインターフェースを作成した。システムのインターフェースと検索の流れについて以下に示した。



図 3 検索の初期画面

はじめは、通常の旅行情報検索システムと同様に、旅行条件を入力する（図3）。ただし、この旅行条件は、検索の初期候補を抽出するための旅行条件であり、これらの条件による検索結果のフィルタリングは行なわない。



図 4 クエリ候補の提示

次に検索条件に合うツアーの中から、印象が異なるように、半自動タグ付けで指定した旅行地の種類別にツアーの写真をも4枚表示する。検索者は、表示された候補の中より行きたいと感じるツアーの写真を選択する。その後、選択された写真をクエリとして、連想検索が行なわれる。検索結果として、選択したツアーの詳細と印象が類似するツアーの写真が、4枚表示される(図5)。



図 5 連想結果

その後、連想結果の中から、興味のある写真を選択し、探索的にイメージを膨らませながら、検索を進めて行く。その過程で、自身の気に入ったコンテンツが見つかった際には、旅市のサイトに進み、ツアーを注文する。

以上、連想検索システムを構築した。次に、このシステムがツアーの選択行動に与える影響、検索アルゴリズムの検証を目的として評価実験を行った。

4.2 連想検索システムの評価

実装した検索システムの有効性を検証するために、2つの評価実験を行なった。

評価1 検索アルゴリズムの評価

はじめに、本システムの検索アルゴリズムを検証するための評価実験を行なった。この評価実験では、連想検索の結果の適切さ、クエリが抽象的である際に、システムの検索精度が通常の検索システムと比較して優位であるかについて検証を行なった。

実験は7月に9人の被験者(20代 男性5名、女性4名)を対象にして行なわれた。

結果の適切さを検証する実験では、写真5枚と対応する7つの連想結果間の、印象の類似度を被験者に0-3の4段階で評価してもらい、これより第何位の結果までが、適切であるかを検証した。

検索精度を測定する実験では、地名検索（通常の検索システム）と連想検索で、再現率、適合率の変化を比較し、精度の検証を行なった。被験者には、「同性の友達3人と旅行に行く」という前提条件のもと、各システムで、「癒し」などの抽象的なクエリによって、ツアーの検索を行なってもらった。この検索結果を元にして、適合率、再現率を算出した。その際に、正解セットについては、紙媒体で同様のコンテンツを見てもらい、各クエリに適するコンテンツに印を付けてもらった結果を元に作成した。

図6に各評価実験の結果を記した。

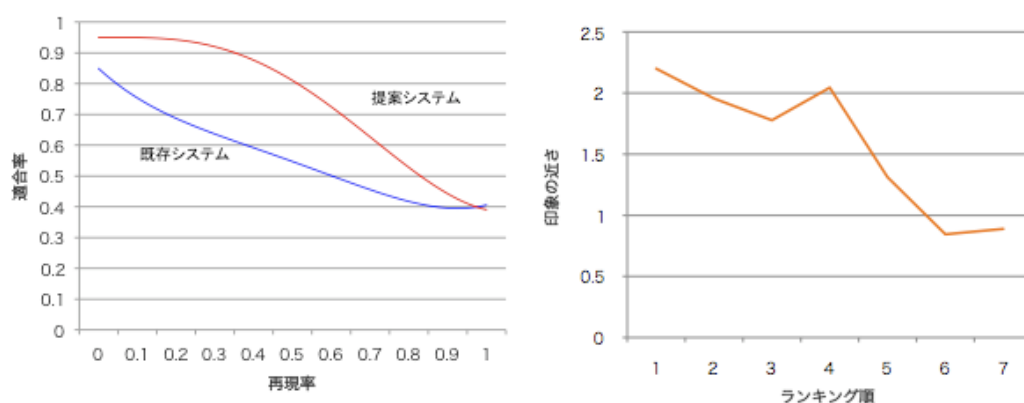


図6 評価実験の結果

結果の適切さについては、第4位までの結果は、類似度が高いと判定され、連想結果として適切である事が確認された。検索精度については、連想検索システムが既存のシステム(地名検索)と比較して、高い適合率を有する事がわかり、抽象的なクエリによる検索では、連想検索システムの検索精度が高い事が確認できた。

評価2 ツアーの検索過程に与える影響の検証

次に、システムがツアーの選択行動に与える影響について検証を行なった。実験では、既存の旅市の検索システム（地名による検索が可能）と、開発した連想検索システム間で、ツアーの検索行動の比較を行なった。実験は、12月に、学生14名（男性6名、女性8名）に対して行われた。被験者はシステム毎に7人ずつに分割された。被験者には、「12月に同性の友人3人と旅行に行く」という前提条件で、各検索システムで10分間、自由にツアーを検索してもらい、その中で、気に入ったツアーを3つ選択してもらった。

ツアーの選択後、検索の過程、選択したツアーの満足度について被験者にインタビューを行なった。その結果から、3つのツアーの選択行動に与える効果が明らかになった。

① 写真の選択から、検索欲求の自覚へ

1つ目に、行きたい旅行地が不明瞭である際に、写真の選択を通じて、検索欲求が具体化される点が挙げられる。

本システムでは、検索を進めるために、検索のクエリを4枚の連想結果の写真の中から選択しなくてはならない。その過程において、各写真を比較する中で、自身の行きたい旅行地の特徴が言語化され、検索欲求の自覚に繋がる事が観察された。

通常の地名による検索でも、検索欲求は言語化されるが、言語化までに、複数のコンテンツを閲覧しなければ成らないため、効率が良くない。連想検索では、1～2回の検索でクエリが言語化される事がほとんどであった。

② ツアーの比較から、納得の行く選択へ

2つ目に、検索要求が具体化した後に、同種のツアー間での比較が、容易にできる点が挙げられる。

本システムでは、印象語により写真の関係性を定義するため、一部の写真が旅行体験の種類により纏まる傾向がある。これにより、同種のツアーの写真が連想結果に表示され、上記のような内容の比較を可能としている。

この際に、複数のツアーを比較した上で、最もイメージに合うツアーを選択するため、選択されたツアーの満足度が高くなる傾向にあった。

③ 意外性のある結果から、新たな視点へ

3つ目に、検索者が予期せぬ結果を表示し、検索者に、新たな視点を与える点が挙げられる。

検索要求が具体化した後、検索者は同種の旅行地の写真を辿って行く傾向にある。しかし、システムでは印象語により写真間の繋がりを定義しているため、印象は類似しているが、種類の異なる写真が結果として表示される場合がある。この際、ツアーにより満たされる内的要因が同様であれば、検索者に、新たな検索の視点を提供する。

この新しい視点により、検索の判断軸が増加し、満足度の高い結果の提供へと繋がった。また、内容の新規性から、新たな視点を与えたツアーを結果として選択する被験者も多かった。

以上のような3つの効果が、実験中に確認された。連想検索システムには、行きたい旅行地の特徴が具体化されるという利点がある。しかし、連想検索では、検索要求が具体化したあとに、

具体的な旅行地を検索することができないため、テキスト検索と条件検索を組み合わせ、検索のストーリーを作り、購買に繋げることが重要であるとする。

4.3 人口学的属性による印象語空間の違いの検証

上記の評価実験によって、連想検索システムが抽象的なクエリによる検索の際に有効である事がわかった。しかし、学生を中心に実験を行なったため、システムの一般性については、検証の余地がある。検索プロセスについては、大差がないと感じるが、年齢や性別によって、旅行地に対して感じる印象は異なるため、連想の精度については差が生じると予想される。そこで、先年度行なった、印象評価実験の結果を元に、分散分析により、人口学的属性による印象語空間の違いについて検証した。

印象評価実験では、年齢、性別、家族構成、職業、収入等の人口学的属性について尋ねた。最初にこれらの人口学的属性の中から、代表的な属性を取り出すために、主成分分析を行なった。その結果、性別と家族構成に、各属性が集約される事がわかった。しかし、年齢は、インターネット調査のモニターに依存したため、分散が低くなったと考えられ、相関から、家族構成が年齢を示していると考えられる。

表 2 性別により有意差が生じる印象語

言葉	P 値	頻度(高)	言葉	P 値	頻度(高)
トロピカルな	0.000	男性	ほのぼのとした	0.000	女性
風光明媚な	0.013	男性	うきうきとした	0.001	女性
異国的な	0.014	男性	わいわいとした	0.001	女性
活動的な	0.033	女性	静かな	0.006	女性
やすらぎ	0.017	女性	庶民的な	0.000	女性
ゆっくりとした	0.000	女性	家庭的な	0.000	女性
プライベートな	0.025	女性			

表 2 は、性別によって、印象語が異なる語を示している。この結果より、男性は「風光明媚な」「異国的な」のような、旅行地の様態を表す語を、女性は「ほのぼのとした」「ゆっくりとした」のような自身の内面を表す語を使用する傾向がある事が確認された。また、実際に、写真間の関係を表すネットワークを構築すると、これらの使用語の違いによって、ネットワークの構造が異なる事が確認された。

表 3 家族構成により有意差が生じる印象語

言葉	P 値	頻度(高)
活動的な	0.008	単身、夫婦、拡大家族

わいわいとした	0.037	単身、夫婦、拡大家族
文化的な	0.014	拡大家族
幻想的な	0.033	夫婦と子、拡大家族
異国的な	0.017	夫婦と子、拡大家族

表3は、家族構成により異なる語を示している。性別程の差は生じなかったが、単身世帯や夫婦世帯など、若い世帯は「活動的な」「わいわいとした」といった、活発さを表す印象語を使用し、反対に、夫婦と子世帯や拡大家族など年齢層の高い世帯は、「幻想的な」「異国的な」といった非日常的な感動を表す語を使用する傾向が見られた。

このように、性別、年齢別に印象語空間が異なる事が明らかになった。今後はこれらの結果を元に、印象語タグの重みを変化させる事が連想検索の精度向上のために適切であると考えられる。

5. 今年度の対外発表

・金澤裕也, 天笠邦一, 日高洋祐, 小川克彦 “写真の印象を元にした旅行地連想検索システムの構築” ヒューマンインターフェースシンポジウム論文集 2012, 1505D, HI学会

・ORF2012 “観光メディア”

・Yuya Kanazawa , Yosuke Hidaka, Katsuhiko Ogawa “Destination Retrieval System Using an Association Retrieval Method” International Journal of Future Computer and Communication, Vol. 2, No. 3, June 2013

・3月発表予定：金澤裕也,日高洋祐,小川克彦 “旅行地のイメージに基づく旅行検索システムの開発と評価” 画像工学研究会

6. 今後の課題

(1) 人口学的属性別の印象語の重み付け

今回は、実装できなかったが、写真の印象評価実験の結果から、年齢、性別により印象語空間が異なる事が観察されたため、年齢、性別により印象語の重みを変化させることによって、検索の精度が高まる事が期待される。

さらに、印象語のセット自体も、年齢・性別によって変化させることが適切であると考ええる。旅行業者に対するインタビューの中でも、年齢によって適切な語が異なるとの指摘があった。そ

のため、今後は、旅行地のさまざまな人口学的属性毎に言葉を抽出して印象語の空間について再度検証を行なう必要がある。

(2) 他の検索方式との併用

評価実験の結果より、検索の時点で行きたい場所が具体化している場合は、連想検索システムは有効でない事がわかった。そのため、多様な検索ニーズを網羅するためには、テキスト検索や条件検索と本検索方式を併用することが望ましい。この併用方式の策定についての考察は、今後の課題となる。

■文献

[1]日本旅行業業界 “数字が語る旅行業 2012”, 2012.

[2] K. Mackay, D. Fesenmaier: Pictorial Element Of Destination In Image Formation, Annals of Tourism Research, Vol.21, No. 3, pp537-565, 1997.