

知的価値を有する観光対象を訪問する個人観光客のための 価値観計量による観光地推薦方式

川崎 彩子[†] 清木 康[†]

[†]慶應義塾大学環境情報学部 〒252-0882 神奈川県藤沢市遠藤 5322

E-mail: [†] {t14253sk, kiyoki}@sfc.keio.ac.jp

あらまし 本稿は、知的価値を有する観光対象を訪問する観光客を対象に、各々の価値観に合う価値を有する観光地を抽出・推薦する方式を提案する。観光対象が有する価値は、文化的価値、歴史的価値、美術的価値など多様であり、観光客が計画段階で正しく価値を理解し観光地を選別することが期待されている。本方式は、意味の数学モデルを利用することで観光対象に関する多次元価値計量空間を構築し、この空間上で観光地とユーザーの価値観を表すコンテキストとの相関計算を行うことにより、それぞれのユーザーに合った観光地を抽出し、推薦する。本方式は、ユーザーの興味に応じた価値を計量し、新たな興味への連想を可能にすることにより、ユーザーの観光対象や付随する学問に関する知識獲得と観光消費推進の実現を目指す。

キーワード データベース、多次元データベース、意味の数学モデル、情報推薦、感性計量、価値観、観光

1. はじめに

近年、景気回復や生活の質の向上に伴い、全世界的に観光客数が増加傾向にある[8]。観光立国という言葉もある通り、観光業は今や一大産業である。観光業は、消費者が継続的に観光行動を行うことで持続できる。そのためには、より良い観光体験を提供できる環境を整備することが必須である。

昨今の観光の主体は団体から個人へと変化しており[8]、2個人観光客が自ら観光行動を企画し、観光対象を選定する機会が増えている。観光対象の選定は観光体験の質を本質的に左右するプロセスであるため、個人観光客は良い観光体験を得るために、企画段階である程度情報を収集・解釈し観光対象の選定を行う傾向にある。

しかし、ここで必要とされる情報収集の程度は観光対象の種別によって大きく異なる。テーマパークなどの産業的な観光対象は運営会社の情報や口コミなど個人でも解しやすい情報が主流である一方、知的価値を有する観光対象は、人文科学、社会学、美術学など様々な分野でその価値について学問的に議論されてきた歴史を持つ。そのため一般的な個人観光客がこのような知的価値を有する観光対象から一つを選定する場合、膨大な情報・知見の存在により情報収集・解釈プロセスを厭うことが問題となり得る。

例えば寺社巡礼や世界遺産巡りに代表される遺産観光（ヘリテージツーリズム）において、各遺産の価値や特色を正しい意味で知り選定するためには、観光対象となる遺産やそれにまつわる文化や歴史などの背景を一つ一つ理解することが必要である。その際情報解釈の手間を惜しみ短絡的に観光対象を選定してしまうと、観光体験自体の質の低下を免れない。

そこで本稿では、このような知的価値を有する観光対象について長年議論されてきた知見を知識ベースとして構築する。観光における対象の価値やその形容方法を知識ベース化することで、ユーザーとして想定される個人観光客それぞれ価値観に相関のある情報を推薦し、有用な観光対象の選定に導く。

2. 基本方式

人間にはそれぞれの価値観、すなわちどのような価値を評価するか・どのような価値を評価しないかという異なる評価基準を持っており、人によってあらゆる事象・もの・ことに対して感じる価値は異なる。言い換えれば、対象がもつ価値の大きさは、どのような価値観の人間に見られるかの文脈によって異なる。この構造は、観光行動においても同じ様相を呈する。一般に価値ある観光対象とされているものでも人の価値観によって感じられる価値の大きさは異なるのである。

本方式は、これらユーザーの価値観を、「価値の大きさを決める文脈（コンテキスト）」ととらえ、意味的価値計量空間を構築することによって個々のユーザーの価値観に合致する観光対象を得るものである。

2.1. 意味的価値計量空間の生成

本方式では、観光対象のデータに意味の数学モデル[1][2][3][4]を適用しその価値を意味的に計量する。意味の数学モデルを用いると、特徴量付ベクトルとして定義したマルチメディアデータを、正規直交系を成す空間上でその時々々の文脈で選択された部分空間に射影し計量することで、文脈ごとに相関の強いデータを抽出することができる。

本方式では、意味の数学モデル[1][2][3][4]を用いて

観光対象の価値と、ユーザーの重視する価値を意味的に解釈し、相関の強い観光対象を抽出する。

まず、意味の数学モデル[1][2][3][4]におけるイメージ空間の構成に則り、観光対象の意味的価値計量空間を生成する。観光対象の有する価値は、「美的価値」、「歴史的価値」など意味によって分類することができる。本稿では観光対象の有する価値を意味の種類を意味的に重複しない最小単位(以下「観光価値の意味素」という) n 個に分類できると仮定する。

分類した n 個の意味素を互いに正規直交な特徴量 f として定義し、 f の集合 $f_n = \{f_1, f_2, \dots, f_n\}$ を基底とする n 次元空間を意味的価値計量空間 I とする。

観光対象である文化財や自然の価値の分類について人文科学や観光社会学で諸説検討されており、これらの知見による分類を観光価値の意味素として用いることで本方式の実現が可能である。

2.2. 観光価値用語知識ベースの生成

ユーザーの価値観を観光対象の評価文脈として与えるため、自然言語で用いられるワードを「観光価値用語」として定義しなおす知識ベースを生成する。ユーザーに選択されたワードを観光価値の意味素値で表現することで、観光価値を形容する際に用いられる意味に変換する。

前提として、前節で定義した観光価値の意味素について記述したドキュメントが与えられていることとする。なお、このドキュメント上で、各観光価値の意味素が端的に自然言語で形容されていることは必要ではない。

観光価値用語知識ベースの生成には、観光価値用語抽出、意味素値の算出の三つのステップを要する。

Step-1: 観光価値用語抽出

各観光価値の意味素について記述されたドキュメントを形態素解析し、単語自体に意味を持つ品詞を抽出する。本稿では、助詞や接続詞は文章に意味をもたらすがそれ自体は観光価値的意味をもたないとし、名詞(固有名詞含む)、形容詞、動詞のみを観光価値用語として抽出するものとする。

Step-2: 意味素ごとの用語頻度の取得

各観光価値の意味素ごとに、抽出された観光価値用語が用いられた頻度をそれぞれ取る。この時点の各値を表に示す。

Step-3: 意味素値の算出

Step-2 の形から転置し、それぞれの観光価値用語に対し、意味素ごとの頻度を持つマトリクスを生成する。

ここで、各用語の持つ意味量は用語によらず一定であると仮定し、各行に対して2ノルム正規化を行い、正規化後の値を意味素値とする。

2.3. データマトリクスの定義

この節では、前節で示した知識ベースのデータと観光対象のデータを、意味の数学モデルで扱うための定義を行う。

任意の観光対象 o は n 個のうちいくつかの価値を有する。このとき、各観光価値の意味素での価値は o に関してそれぞれ存在が認められるか否かの2値であり、大きさは考えないものとする。つまり、 a 個の各観光対象について n 個の特徴 (f_1, f_2, \dots, f_n) を持つ特徴付ベクトル o_i ($i=1, \dots, m$) を考えた時、 o_i は0か1を各成分に持つベクトルである。

このベクトル o_i を a 個並べた $a \times n$ 行列 A を考える。図1左に示すのが観光対象 (A) のデータ行列である。

次に、観光価値用語知識ベースにより、 b 個の観光価値用語 w が n 個の観光価値の意味素 (f_1, f_2, \dots, f_n) に対して特徴値を定義されたとする。各観光価値用語に対して特徴付ベクトル w_i ($i=1, \dots, m$) を考えた時、 w_i は0以上1以下の正の数値を各成分に持つベクトルであると言える。

そのベクトルを並べた $b \times n$ 行列 B を考え、観光価値の形容語 (B) のデータ行列を図1右に示す。

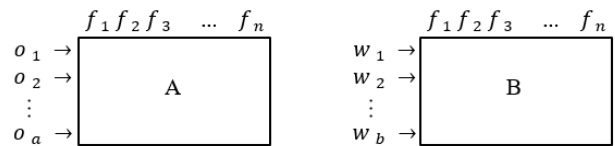


図1 データ行列 A , B によるメタデータの表現

2.4. 意味的価値計量と推薦データ抽出方式

この任意の観光価値用語ベクトル w に対し、成分が閾値 F 以上の観光価値の意味素にあたる軸のみで生成される部分空間 I_w を考える。 I 内に存在する観光対象ベクトル o を I_w 空間に写像した V' の大きさが w の意味での o の価値量となる。

これを踏まえ、価値評価文脈を決定する k 個の単語の列 $S_k = \{w_1, w_2, \dots, w_k\}$ が与えられた時の、文脈による部分空間選択と各観光対象の価値の導き方を示す。

S_k の各要素は、前章で定義したマトリクス A 上に観光価値の意味素値を持つため、意味的価値計量空間 I 上にベクトルとして定義可能である。 S_k の重心ベクトル G は、 S_k の全要素ベクトルが持つ意味の重心、すなわち与えられた文脈の意味重心ととらえられる。したがって、 G によって選択される部分空間 I_G に射影した観光対象ベクトル o の大きさが与えられた文脈にお

いて o の持つ価値となる。

本方式では前章マトリクス B 上の各要素 o を全て I_0 上に射影し、その射影ベクトルの大きさをはかることで、文脈に応じた観光対象の意味的価値を計量する。また、この価値の大きい順に推薦結果を返す。

3. 実現方式

本章では、前章で述べた基本方式を実現するためのシステム概要、使用データとその構造、さらに実現環境と構成を示す。

3.1. システム概要

想定される使用の流れは以下の図に示す通りである。

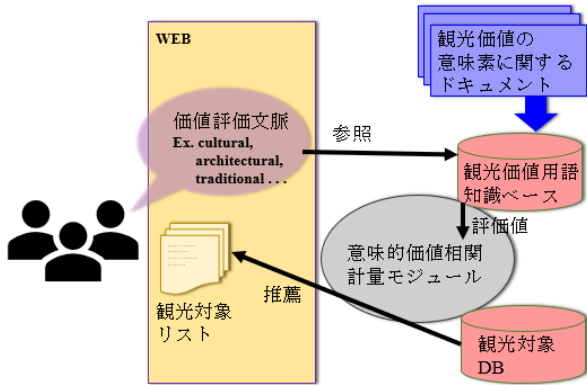


図2 基本方式を実現するシステムの使用想定フロー

3.2. 使用データ

本方式の実現にあたっては、価値計量空間構築のため、まず観光対象の分類例が必要である。また、観光対象についてその価値の種類を特徴量とした知識ベースを構築する性質上、その価値について十分議論がされている観光資源を対象とすることが望ましい。そこで本稿では、実現のための対象データとして世界遺産とその登録基準を用いる。本節では、世界遺産とその分類・ドキュメントを本方式の実現に応用する方式を説明する。

3.2.1. 世界遺産とは

世界遺産とは、「顕著な普遍的価値」を有することが認められ、国際連合教育科学文化機関（United Nations Educational, Scientific and Cultural Organization U.N.E.S.C.O., 以下ユネスコという）の世界遺産委員会により世界遺産一覧表に記載されたもののことを言う。「世界遺産条約履行のための作業指針(Operational Guidelines for the Implementation of the World Heritage Convention)」[10][11]49 項によれば、顕著な普遍的価値とは「国家間の境界を超越し、人類全体にとって現代及び将来世代に共通した重要性をもつような、傑出した文化的な意義及び/又は自然的な価値を意味する。」さらに同項には「委員会は、世界遺産一覧表に資産を登録するための基準の定義を行う。」とあることから、世界遺産は以下の表に示す 10 項目のいずれかの意味で顕著な普遍的価値を有するものとして定義されていると解することができる。

表1 世界遺産登録基準

| | |
|------|---|
| i | 人間の創造的才能を表す傑作である。 |
| ii | 建築、科学技術、記念碑、都市計画、景観設計の発展に重要な影響を与えた、ある期間にわたる価値観の交流又はある文化圏内の価値観の交流を示すものである。 |
| iii | 現存するか消滅しているかにかかわらず、ある文化的伝統又は文明の存在を伝承する物証として無二の存在（少なくとも希有な存在）である。 |
| iv | 歴史上の重要な段階を物語る建築物、その集合体、科学技術の集合体、あるいは景観を代表する顕著な見本である。 |
| v | あるひとつの文化（または複数の文化）を特徴づけるような伝統的居住形態若しくは陸上・海上の土地利用形態を代表する顕著な見本である。又は、人類と環境とのふれあいを代表する顕著な見本である（特に不可逆的な変化によりその存続が危ぶまれているもの） |
| vi | 顕著な普遍的価値を有する出来事（行事）、生きた伝統、思想、信仰、芸術的作品、あるいは文学的作品と直接または実質的関連がある（この基準は他の基準とあわせて用いられることが望ましい）。 |
| vii | 上級の自然現象、又は、類まれな自然美・美的価値を有する地域を包含する。 |
| viii | 生命進化の記録や、地形形成における重要な進行中の地質学的過程、あるいは重要な地形学的又は自然地理学的特徴といった、地球の歴史の主要な段階を代表する顕著な見本である。 |
| ix | 陸上・淡水域・沿岸・海洋の生態系や動植物群集の進化、発展において、重要な進行中の生態学的過程又は生物学的過程を代表する顕著な見本である。 |
| x | 学術上又は保全上顕著な普遍的価値を有する絶滅のおそれのある種の生息地など、生物多様性の生息域内保全にとって最も重要な自然の生息地を包含する。 |

世界遺産条約履行のための作業指針(文化庁仮訳)[11]

本方式実現にあたり、世界的に認められているこの価値種別を観光価値の意味素として採用し、2018年1月現在認められている1073件の世界遺産を観光対象データとする。

なお各世界遺産の有するそれぞれの観光価値の意味素についてはユネスコの国際記念物遺跡会議

(ICOMOS/ International Council on Monuments and Sites) によるドキュメンテーション (以下遺産価値評価文書という) 上で詳しく述べられている。この遺産価値評価文書は作業指針に則ってユネスコで保存され、公開されている。

3.2.2. ドキュメント分析による世界遺産用語知識ベースの自動生成

遺産価値評価文書を用い、世界遺産用語知識ベースを構築する。本稿で用いる遺産価値評価文書は世界遺産ごとにその世界遺産が有する意味の価値を記述しているため、実現には基本方式で述べた3つのステップに加え、準備段階として解析する文章を選出するステップが必要となる。

Step-0: 文章選出

遺産価値評価文書のうち、各観光価値の意味素について評価が述べられている部分を意味素ごとに収集する。

Step-1: 世界遺産用語抽出

収集した文章群を形態素解析し、名詞、動詞、形容詞を抽出しマトリクスをつくる。

形態素解析には Tree Tagger[12]を用いた。

Step-2: 意味素ごとの用語頻度の取得

マトリクスに列を追加し、各観光価値の意味素ごとに、抽出された観光価値用語が用いられた頻度を記す。

Step-3: 意味素値の算出

Step-2 の形から転置し、それぞれの観光価値用語に対し、意味素ごとの頻度を持つマトリクスを生成する(表)。ここで、各用語の持つ意味量は用語によらず一定であると仮定し、各行に対して2ノルム正規化を行い、正規化後の値を意味素値とする。

3.3. システムの実現環境と構成

3.3.1. システム構成図

本システムでは、ユーザーの入力取得と出力表示、またその制御のためにHTML形式のユーザーインターフェースを実装した。全体のシステム構成図を以下に示す。

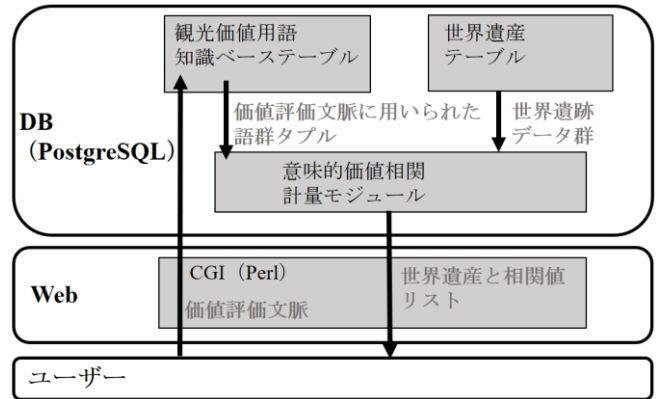


図3 本方式実現のためのシステム構成図

3.3.2. データ格納と計算の実現

PostgreSQL を用いて世界遺産データベースを構築し、CLI 上での推薦を実現した。PostgreSQL 上では、データベースの他に、意味的価値相関計量モジュールの役割として、任意の2データの意味素値について内積を計算する関数を定義した。

3.3.2.1. 世界遺産データ

世界遺産データは PostgreSQL 上に記述されている。データ内容はユネスコ公式サイトの World Heritage List[9]を基に1073個のデータを収納している。

データ構造は、世界遺産名称を主キーとし、観光価値の意味素10項目を属性に持つ。また、推薦には用いないが推薦結果としてユーザーに表示するための情報として、所属国名とより詳細なエリア名、緯度経度と簡単な説明文をメタデータとして持っている。観光価値の意味素については、10項目それぞれに認められているか否かで0か1かの値を持つ。

以下の表2にデータ例を示す。ここでは、推薦に用いない情報を省略する。

表2 世界遺産データの例

| Name | ... | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 | f_{10} |
|--------------------|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| Burgos Cathedral | ... | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| Yakushima | ... | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |
| Old City of Sana'a | ... | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |

3.3.2.2. 世界遺産価値用語データ

3.2.2 にて構築した知識ベースもまた、PostgreSQL に格納することで検索・計算を可能にした。データ構造は、ワード名を主キーとし、観光価値の意味素10項目を属性に持つ形である。前述通り、各観光価値の意味素の値は2ノルムで正規化され、0以上1以下の値である。以下に、表3にデータ例を示す。

表 3 世界遺産価値用語データの例

| Word | f_1 | f_2 | f_3 | f_4 | f_5 | f_6 | f_7 | f_8 | f_9 | f_{10} |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| spiritual | 0 | 0.162 | 0.135 | 0.081 | 0.081 | 0.486 | 0.054 | 0 | 0 | 0 |
| permanent | 0 | 0.2 | 0.2 | 0 | 0 | 0 | 0.1 | 0.2 | 0.2 | 0.1 |
| ideal | 0.375 | 0.25 | 0.125 | 0.062 | 0.062 | 0.125 | 0 | 0 | 0 | 0 |

3.3.2.3. ユーザーデータ

ユーザーコンテキストは、本システムの動的推薦を実現するため、システムの入力として単語列を都度入手する形式を採った。データベースとしては保持せず、問い合わせの要素として入力する。

例としては、“cultural”, “spiritual, ritual, traditional” などである。このように、任意の単語を一つ以上指定する。

3.3.2.4. PostgreSQL 上での本方式実現

上記本節で定義したマトリクスに対し、PostgreSQL 上で以下に説明するクエリを送ることで、本方式の CLI 上での実現が可能となる。

まず、入力となる単語列（ユーザーデータ、検索したい単語の組）が含む各単語について、観光価値の意味素値を世界遺産価値用語データマトリクスに参照する。次に、各単語の観光価値の意味素値を合計したものを求め、結果を TEMP VIEW として保存するクエリを発行する。最後に、各世界遺産のデータと TEMPVIEW の意味素値を意味的価値相関計量モジュール（関数）に送り、その計算結果の値（スコア）でソートし表示させる。

以上で得られた表示が、本方式による推薦結果である。

Recommended World Heritage List

RANK:1

Score:1.9065420560747662

| | | | | | | | | | | | | | | | |
|---|---|----------------------------|----------------------------------|-----------------------|------------------------------|--|-------------------|--|---------------------------|----------------------------|----------------------------------|---------------------|------------------|--|----------------------|
| (1) Pergamon and its Multi-Layered Cultural Landscape | (2) Cultural Landscape and Archaeological Remains of the Bamiyan Valley | (3) The Persian Garden | (4) Buddhist Monuments at Sanchi | (5) Acropolis, Athens | (6) Ancient City of Damascus | (7) Assisi, the Basilica of San Francesco and Other Franciscan Sites | (8) Mount Taishan | (9) Mahabodhi Temple Complex at Bodhi Gaya | (10) Villa d'Este, Tivoli | (11) Takht-e Soleyman | (12) Historic Centre of Florence | (13) The Great Wall | (14) Mogao Caves | (15) Sanctuary of Asklepios at Epidaurus | (16) Arc Site of Oly |
| Turkey | Afghanistan | Iran (Islamic Republic of) | India | Greece | Syrian Arab Republic | Italy | China | India | Italy | Iran (Islamic Republic of) | Italy | China | China | Greece | Gre |
| | | | | | | | | | | | | | | | |

図 5 GUI による推薦結果出力の様子

3.3.3. 入力取得と出力表示制御の方式

本節では本方式の GUI 上での実現の例とこれによる推薦出力の例を示す。

本稿では、HTML 形式のユーザーインターフェースへの情報の入出力、データベースとの通信、および問い合わせの実行にプログラミング言語 Perl を用い CGI を実装した。

ユーザーは表示された HTML 形式の入力欄に価値評価文脈として単語列を入力することができ、「Submit」ボタンをクリックすることでシステムへの入力を行う。

CGI は、前節の手順に従い PostgreSQL 上で 3 つのクエリを発行し、結果を推薦取得する。

続けて、CGI はこの結果をリストとして HTML 形式でユーザーに返すことで情報の推薦を実現する。

入力ページと表示した結果の様子を以下の図 4、5 に示す。

Send some keywords

Ex1: spiritual, ritual, traditional

Ex2: change, history

change,history

送信

図 4 GUI によるユーザーコンテキスト入力の様子

4. 実験と評価

本章では、前章で示した実現方式で実験環境を構築し、実際に情報推薦を行うことにより、本方式の有効性を示す。

4.1. 推薦の有効性検証実験

4.1.1. 推薦の有効性検証実験の概要

本実験は、前章で示した実現方式に則って構築した実験環境を用い、実際の推薦結果を収集・分析することで、推薦の有効性を示すものである。すなわち、世界遺産 1073 か所のデータに対し、ユーザーコンテキストとして入力される単語列との相関計算を行い、相関量が一定閾値以上のものを推薦し、その結果の有効性を検証する。

本実験では使用する入力単語列に対して、あらかじめ正解リストが存在することが必要である。ここでは、多次元空間の軸に用いた世界遺産の登録基準を用い、任意の軸の概念を表すような単語列を入力、その基準を見たす世界遺産のリストを正解リストとして採用することにする。

具体的には、以下のような実験を行う。

実験(i) : criteria(i)"to represent a masterpiece of human creative genius"を代入する実験

- ・ 入力 – 上記概念を特徴づける単語の列「represent, masterpiece, human, creative, genius」
- ・ 正解リスト – criteria(i)を満たすことが認定されている世界遺産のリスト

同様に、世界遺産登録基準(i)~(x)を用い実験(i)から実験(x)まで、10 種の実験を行った。本方式は、相関量が 0 以上の対象をランキング形式で推薦するものである。しかしより詳細な精度検証のために、ここでは、各実験で 1.0 から相関量の上限值まで、1.0 刻みで閾値ごとに複数回の検証を行った。

4.1.2. 推薦の有効性検証実験の結果

実験(i)から実験(x)のそれぞれについて、正解数、検索結果数、検索結果中正解数、再現率、適合率、F 値を記録した。

10 の実験のうち、実験(i)と実験(x)では F 値が最大の 1 になるような閾値が存在し、実験(ii), (iii), (iv), (ix)では F 値が 0.9 を超える閾値が存在した。また、実験(i)から実験(x)において、実験ごとの F 値の最大値の平均は 0.881032 であった。

以下、最も精度が低かった実験(viii)の結果を表 4 に、最も精度が高かった実験(i)の結果を表 5 に示す。

表 4 : 実験(viii)の推薦結果の分析

| 閾値 | 検索結果数 | 検索結果中 | | 再現率 | 適合率 | F値 |
|-----|-------|-------|-----|----------|----------|----------|
| | | 正解数 | 正解数 | | | |
| 7.0 | 1 | 1 | 90 | 0.011111 | 1 | 0.021978 |
| 6.0 | 2 | 2 | 90 | 0.022222 | 1 | 0.043478 |
| 5.0 | 7 | 5 | 90 | 0.055556 | 0.714286 | 0.103093 |
| 4.0 | 36 | 29 | 90 | 0.322222 | 0.805556 | 0.460317 |
| 3.0 | 135 | 46 | 90 | 0.511111 | 0.340741 | 0.408889 |
| 2.0 | 655 | 90 | 90 | 1 | 0.137405 | 0.241611 |
| 1.0 | 942 | 90 | 90 | 1 | 0.095541 | 0.174419 |

表 5 : 実験(i)の推薦結果の分析

| 閾値 | 検索結果数 | 検索結果中 | | 再現率 | 適合率 | F値 |
|-----|-------|-------|-----|----------|--------|----------|
| | | 正解数 | 正解数 | | | |
| 4.0 | 69 | 69 | 251 | 0.2749 | 1 | 0.43125 |
| 3.0 | 248 | 248 | 251 | 0.988048 | 1 | 0.993988 |
| 2.0 | 251 | 251 | 251 | 1 | 1 | 1 |
| 1.0 | 400 | 251 | 251 | 1 | 0.6275 | 0.771121 |

4.1.3. 推薦の有効性検証実験の考察

全体を通して再現率、適合率、F 値は高い水準の推薦結果を得た。特に、実験(i)では、閾値 2.0~4.0 で適合率が、閾値 4.0 で再現率、適合率、F 値全てが最高値 1 を記録し、正解リストを完全に抽出できてきた。実験(viii)においても、最低限の水準が満たされている。

また、本方式はランキング形式での推薦を行うが、F 値の最高値の平均が 0.88 を超えていることから、一定閾値以上の上位に高い割合で正解を抽出できていることもわかる。

以上より、情報推薦の有効性について、優位な結果が得られたと言える。

4.2. 既存推薦方式との比較実験

4.2.1. 実験概要

本実験では、パターンマッチによる情報推薦結果との比較により、本方式の特性を記述する。

前提として、本方式による情報検索・推薦とパターンマッチ検索による情報検索との相違を述べる。両者とも、単語の集合を入力として、関連する情報をリストアップする点では同様である。異なる点は、パターンマッチによる検索が、ユーザーが欲す情報が明確に存在し、それに関連する具体的なインプットによって情報を抽出する点にある。

これに対し本方式は、情報が指す観光対象の持つ価値の意味を解釈・計算し抽出するという点で本質的に異なる。すなわち、実際的には異なる現象・事象であって、パターンマッチでは検索されない情報も、それらを繋ぐコンテキストが存在すれば同時に出現し得る。

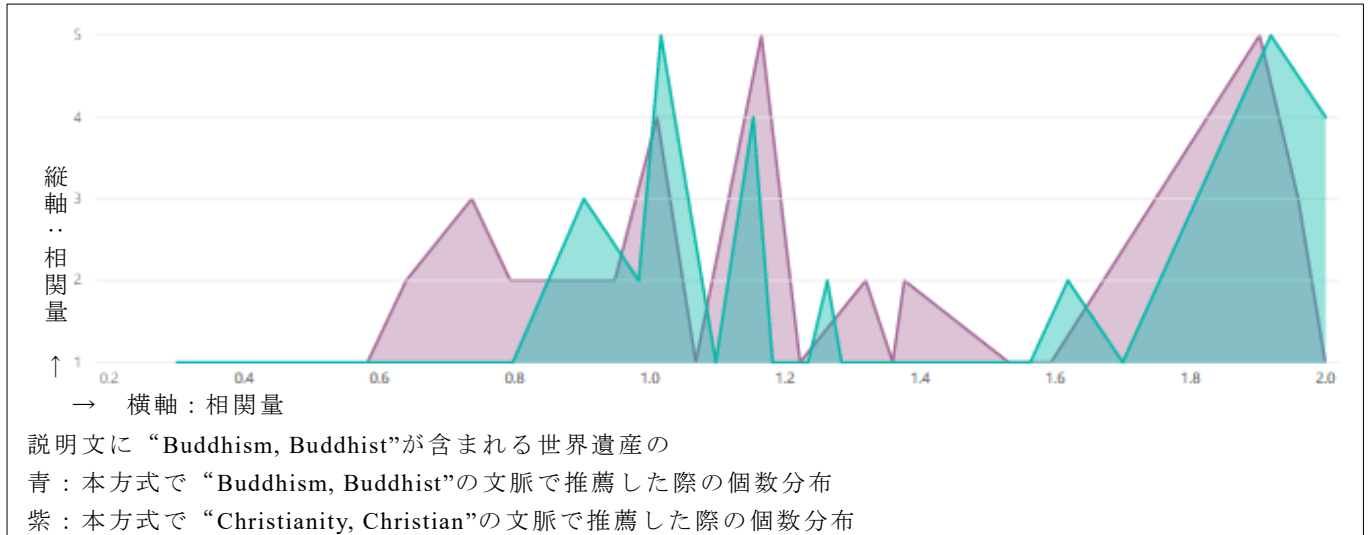


図 6 単語列 AB による本方式分析結果における、パターンマッチ検索結果リストの分布

また、本方式の推薦結果への出現によって、対象に関する隠れたコンテキストの推論も可能である。

以上に述べた特性を示すため、観光対象データにパターンマッチ検索結果によるタグを付加し、情報推薦を行った。

具体的には、まず (A) 「Buddhism, Buddhist」、(B) 「Christianity, Christian」の2つの単語列について、各世界遺産の説明文[9]とのパターンマッチ検索を行った。

次に、パターンマッチ検索で該当した世界遺産にメタデータをつけ、本方式によって情報推薦し、メタデータの個数分布を相関量について検証した。

4.2.2. 実験結果と考察

結果を図 6 に示す。図 6 に示すのは “Buddhism, Buddhist” のメタデータが付与された世界遺産（すなわち寺や仏教に関わる建物、街等）を、本方式により “Buddhism, Buddhist” “Christianity, Christian” の2つの文脈で推薦した際の相関量ごとの検索個数分布の比較図である。右に行くほど推薦度（相関量）が高い。単語列 (A) (B) で検索を行った場合、パターンマッチ検索ではそれぞれ異なる結果を得るが、それらの異なる結果群は、本方式に推薦された際そのスコアについて非常に似た分布を示すことがわかる。

文脈 (A), (B) はそれぞれ異なる二つの概念を表す単語列であるが、“宗教” という共通概念を共有するものであり、それらの文脈に関連する共通の世界遺産が、本方式により発見されている。すなわち図 6 は、本方式が、意味を計算し推薦結果を導出するメタデータを自動付与し、個別の文脈が「宗教」という意味で類似するものを共通的に推薦できることを示すものである。

以上の結果により、検索対象についてタグ情報でな

い隠れたコンテキストを発見する本方式の特性が示された。

5. おわりに

本稿は、知的価値を有する観光対象を訪問する観光客を対象にそれぞれの価値観に合う価値を有する観光地を抽出・推薦する、「価値観計量による観光地推薦方式」を提案した。

本方式は観光対象となる文化、美、歴史などに関する知の成熟に伴い進化していく。ユーザー個人に即した情報推薦で新たな興味への連想を可能とし、学問の発展へのキャッチアップを手助けすることで、ユーザーに知識獲得を促す。

今後、計算モデルの改善やデータの選定により、精度の向上を図るとともに、本方式の展開の可能性を探っていく。また、実験で示した特性を用いて解析を進めることにより、観光対象同士の関連や新たな知的文脈の発見をめざす。

参考文献

- [1] KITAGAWA T. and Kiyoki.Y, "A mathematical model of meaning and its application to multidatabase systems," Proc. IEEE International Workshop on Research Issues on Data Engineering: Interperability in Multidatabase Systems. 1993, p. 130-135.
- [2] 北川高嗣, 清木 康, 人見 洋一, “意味の数学モデルとその実現方式について” 信学技報. 1993, vol. 93, p. 25-31.
- [3] 清木 康, 金子 昌史, 北川 高嗣, “意味の数学モデルによる画像データベース探索方式とその学習機構” 電子情報通信学会論文誌.D-2, 情報・システム 2-情報処理. 1996, vol. 79, no. 4, p. 509-519.
- [4] 宮川 祥子, 清木 康, “特定分野ドキュメントを対象とした意味的連想検索のためのメタデータ空間生成方式” 情報処理学会論文誌データベース (TOD) . 1999, vol. 40, no. 5, p. 15-28.

- [5] 前田 勇,“新現代観光総論“ , 学文社, 2015, x, 215 p.p.
- [6] 北川 宗忠,“観光文化論“ , ミネルヴァ書房, 2004, vii, 234 p.p.
- [7] 垣内 恵美子, 岩本 博幸,“文化財の価値を評価する : 景観・観光・まちづくり“ , 水曜社, 2011, 203 p.p.
- [8] 総理府, 国土交通省, 観光庁, "観光白書" 2017, 平成 29 年版.
- [9] United Nations Educational, Scientific and Cultural Organization U.N.E.S.C.O. “World Heritage List”, <http://whc.unesco.org/en/list> , (3 月 28 日参照).
- [10] United Nations Educational, Scientific and Cultural Organization U.N.E.S.C.O. “Operational Guidelines for the Implementation of the World Heritage Convention” <http://whc.unesco.org/document/163852> , (3 月 28 日参照).
- [11] United Nations Educational, Scientific and Cultural Organization U.N.E.S.C.O. “世界遺産条約履行のための作業指針 (文化庁仮訳)” 平成 17 年度版, http://bunka.nii.ac.jp/special_content/hlink13 , (3 月 28 日参照).
- [12] Tree Tagger
<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/> (3 月 28 日参照).