

文書中の地物画像を言語的記述で代替するための 地物の外観情報の Web からの抽出

服部 峻[†] 手塚 太郎[†] 田中 克己[†]

誰もが、どんな計算機環境からでも、Web 文書などのデジタル情報へユニバーサルアクセス可能にする技術の 1 つとして、画像メディアを出力できない非視覚的ブラウザのために、マルチメディアな Web 文書を文字情報だけからなる画像を含まないテキスト版に変換する技術が必要とされている。一般オブジェクトが撮影された画像に対して、その内容を説明する言語的記述を生成するという研究課題の小問題として、本論文では、文書に含まれ、固有名称を持つ地物（建築物）が撮影された写真に対して、その内容を説明する言語メディア表現に変換するシステムについて概説し、そのうえで、本システムを構成する要素技術である地物の外観情報を Web から抽出する手法について提案する。

Extracting Visual Descriptions of Geographic Features from the Web as the Linguistic Alternatives to Their Images in Digital Documents

SHUN HATTORI,[†] TARO TEZUKA[†] and KATSUMI TANAKA[†]

“Ubiquitous Societies” require technologies to enable us to access information universally. One of the most important technologies allows us to access not only multimedia Web documents via graphical browsers in common use but also their possibly equivalent text-only versions via non-graphical browsers such as text or audio browsers. As a conditioned solution to the media conversion from images of general objects into their linguistic explanations, this paper proposes a novel method to extract visual descriptions of geographic features by mining the Web, and to replace each photograph of geographic features in digital documents with their visual descriptions as its linguistic alternative. The method can convert multimedia Web documents into their text-only versions more equivalently than ever before.

1. はじめに

近年、ユビキタス社会を実現するための大きな課題として、誰もが、どんな計算機環境からでも、デジタル情報にユニバーサルアクセス可能にする技術が求められている。その 1 つとして、画像メディアを出力できないテキストブラウザ や音声ブラウザ といった非視覚的ブラウザからでも Web 文書にアクセス可能にするために、画像メディアも混合したマルチメディアな Web 文書を、言語的記述のみからなるテキスト版に変換する技術が必要とされている。

従来の非視覚的ブラウザでは、Web 画像に対して、その IMG 要素の ALT 属性の代替テキスト、SRC 属性の画像ファイル名、その画像がリンクアンカであればリンク先のタイトルや URL といった文字情報で代替することで、対象の Web 文書を文字情報だけからなる画像を含まないテキスト版に変換する。文字入りで

メニューリンクとして使われている画像の多くは、リンク先のタイトルや URL で置換しさえすれば、元々のマルチメディアな Web 文書とほぼ等価なテキスト版が得られ、また、単なる装飾として使われている画像の場合、単純に除去するだけでも特に問題はない。しかしながら、旅行記や観光ガイド中の地物画像など、Web 文書において画像が著者の伝えたい主要なコンテンツである場合、従来の単純な置換方法では、元々の Web 文書から視覚的に得られる情報とは程遠い価値のテキスト版しか得られない。一般の視覚的ブラウザからと非視覚的ブラウザからとでは、同一の Web 文書にアクセスしても得られる情報に差があり、依然としてデジタルデバイドの問題を払拭できていない。

Lynx: <http://lynx.browser.org/> (2007).

SimpleWeb Demo: <http://demo.simpleweb.jp/> (2007).

IBM Home Page Reader: http://www-06.ibm.com/jp/accessibility/solution_offerings/hpr/ (2007).

JAWS: http://www.freedomscientific.com/fs_products/software_jaws.asp (2007).

[†] 京都大学大学院情報学専攻

Department of Social Informatics, Graduate School of Informatics, Kyoto University

画像メディアと言語メディアは表現能力などの特長が異なるため、Web 画像を何らかの言語的記述で置換することで、マルチメディアな Web 文書を完全に等価なテキスト版に変換することは非常に困難である。しかしながら、Web 画像の内容を代替する言語的な表現として、その画像中のオブジェクトの名称やオブジェクト間の位置関係といった高次の意味的な情報とともに、そのオブジェクトの色名、大きさ、形状名などの高次の視覚的な情報を対応付けることができれば、Web 文書のいっそうのマルチメディア化にともなうデジタルデバイドの問題を緩和できると我々は考える。

画像の内容を説明する言語的な記述を求める研究は、画像認識や画像理解の分野で古くからさかに行われてはいるが、光学的文字認識 (OCR) や特定の対象領域に限定したオブジェクト認識などを除き、画像中の一般オブジェクトを高精度に認識する実用的な手法ははまだ考案されていない。一方、画像検索の分野では、Web 文書中の画像に対して、その画像に関連するキーワードや説明文などの文字情報を自動的に索引付けする様々な手法が提案されてはいるが、非視覚的ブラウザからのユニバーサルアクセスを実現するために、画像の代替テキストとして十分適当な文字情報を抽出できている研究例は見当たらない。

一般オブジェクトが撮影された画像に対して、その内容を説明する言語的記述を自動生成するという研究課題の小問題として、本論文ではまず、デジタル文書中に含まれ、固有名称を持つ地物 (建築物) が撮影された写真に対して、その内容を説明する言語メディア表現に変換するシステムについて概説する。我々が目標とするシステムは、対象の画像中の地物名の特定、その地物の外観情報の Web からの抽出、および、特定した地物名とその外観情報を用いた言語メディア表現の自動生成から構成される。そのうえで、本システムを構成する要素技術である地物の外観情報を Web から抽出する手法について提案し、評価実験を行う。

本論文の構成を以下に示す。まず、2 章では、本論文に関連する研究について紹介し、提案手法との比較を行う。3 章では、我々が目指す最終的な目標である「文書中の画像を代替する言語的記述の自動生成」システムの実現方法について概説し、以降で提案・評価する「地物の外観情報の Web からの抽出」手法の位置付けについて述べる。4 章では、地物の外観情報を Web から抽出する手法について具体的な提案を行う。5 章では、4 章で提案した手法に対する評価実験、および、その考察を行う。最後に、6 章で本論文をまとめ、今後の研究課題についても述べる。

2. 関連研究

本章では、本論文に関連する研究として、画像検索のために索引語や説明文を抽出する既存手法、および、地物の様々な特徴を抽出する既存手法について紹介し、提案手法との比較も行う。

2.1 画像検索のための索引語・説明文の抽出

画像検索は、画像に対して索引付けされたメタな文字情報に基づいて意味的なキーワードで検索する手法 (TBIR: Text-Based Image Retrieval) と、画像自身の内容 (視覚的な特徴量) に基づいて類似画像を検索する手法 (CBIR: Content-Based Image Retrieval)¹⁾ とに大別できる。本論文では、地物画像に対して、撮影されている地物の名称やそれを構成する部分要素の名称といった高次の意味的な情報を特定し、それらの外観情報として、色名、大きさ、形状名などの高次の視覚的な表現記述を Web から抽出しており、前者の TBIR 手法と非常に関連が強い。後者の CBIR 手法では、色、テクスチャ、形状などの低次の視覚的な特徴量に基づいて、類似度計算が行われる。

近年、デジタルカメラやカメラ付き携帯電話の普及により、Web 上には大量の画像が存在しているが、これらをキーワードで画像検索するためには、各画像に対してあらかじめ文字情報で索引付けしておく必要がある。画像認識や画像理解といったコンピュータ・ビジョンの分野では、一般的な画像に対して、写っているオブジェクトの名称、オブジェクト間の位置関係など、その画像の内容に関する高次の意味的な情報を抽出する一般物体認識の研究がさかに行われてはいるが、Web 上の多種多様な画像に対して適用することは現時点では容易ではない²⁾。また、画像の低次の視覚的な特徴量と印象語とを対応付けることにより、ユーザの印象語に基づく画像検索の研究も行われている³⁾。

一方、テキストマイニングや Web マイニングの分野では、Web 画像に対して、それ自身 (IMG 要素) の ALT 属性に記述された代替テキストや SRC 属性に記述されたファイル名、その画像を含む Web ページのタイトル、その周辺テキストや HTML タグの構造などを解析することにより、索引語や説明文といった文字情報を対応付ける様々な手法が提案されている⁴⁾。以下、いくつかの関連研究について詳述していく。

是津らは、Web 画像に対して、その前後にある周辺テキストや画像、ハイパーリンクによって関連付けられた周辺コンテンツから抜き出したリンクアンカや Web 文書のタイトルなどを画像の Web 文脈として抽出し、画像検索の結果とともに呈示することにより、

ユーザの検索要求と画像との関連性を視覚的にとらえることが可能なシステムを提案している⁵⁾。

相良らは、Web 画像を含むページ中で強く関連するテキスト領域をまず抽出し、そのテキスト領域内の重要文に基づく評価尺度と、画像からの距離やテキスト領域中での出現頻度に基づく従来の評価尺度とを組み合わせることで、人の感覚により近く、より高精度な索引付けを実現できることを示している⁶⁾。

竹内らは、Web 画像の ALT 属性を利用し、その代替テキスト中の語と同一 Web 文書中の語との共起関係を求めることで、対象の Web 画像により深く関連する複数の文を同一 Web 文書中の文章の中から抽出し、その Web 画像の説明文として対応付ける手法を提案している⁷⁾。ALT 属性が記述されていない場合には適用できず、索引付けできる Web 画像が少ないという問題があったが、類似した画像には、類似したキーワードが与えられるという経験則に着目し、類似画像検索による画像的な特徴と、語の共起関係に基づくキーワードの拡張による言語的な特徴との両方を併用することで、Web 画像の説明文を、その画像を含む同一 Web 文書中のテキスト部分から抽出する手法を提案している⁸⁾。しかしながら、Web 画像の説明文として、同一 Web 文書中のテキスト部分だけから抽出しているため、その文書中に画像の説明として適当な文が元々含まれていない場合、うまく抽出することはできない。成功例としてあげられている画像の説明文には、その画像の見た目を説明するような記述は含まれていない。竹内らの手法で抽出しうるのは画像中のオブジェクト(地物など)に関する多種多様な説明文であり、そのオブジェクトの外観に関する視覚的な情報が含まれるとは必ずしも限らない。また、非視覚的ブラウザによるアクセスにおいて、同一 Web 文書中のテキスト部分から抽出した文章で画像を代替したとしても、冗長なだけであり意味がない。一方、我々は、画像中の地物の名称をその周辺テキストを基に特定した後、その画像を含む文書中に限らず Web 全体から、その地物の外観に関する言語的な記述をマイニングしている。対象の地物画像を含む文書内にその地物の外観に関する言語的な記述が含まれていない場合でも抽出することが可能であり、非視覚的ブラウザのための画像の代替テキストの自動生成を実現しうる。

岡田らは、画像データに対する感性検索を行うために、名詞だけでなく形容詞も索引語の候補として抽出している⁹⁾。抽出された形容詞の中には、オブジェクトの外観情報として、高次の視覚的な表現記述や印象語を含んでいる場合もあるが、あくまで、対象画像を

含む Web 文書中のテキストだけから抽出しているため、画像の内容に関する視覚的な説明が同一文書中に元々含まれていなければ抽出することができないという問題が、前述の竹内らの手法と同様に残る。

2.2 地物の特徴抽出やオブジェクトの評判抽出

地物の特徴として、本論文で抽出を試みている外観情報のほかにも、地物のランドマーク性や、その地物で典型的に行われている体験を抽出する研究などがある。

ランドマークとは、空間を移動する際に目的地を見つげるための目印となる地物のことである。Tezuka らは、人々が頻繁に目的地にし広く知られている地物であるオーソリティ型のランドマークを文書頻度に基づいて、空間的な位置関係を把握するために重要となる地理オブジェクトであり道案内の説明に頻繁に用いられるハブ型のランドマークを周辺の地名との共起度に基づいて、Web から抽出する手法を提案している¹⁰⁾。さらに、店舗などの開店している時間帯を抽出し、与えられた時間に活動している場所を可視化するシステム「ChronoSearch¹¹⁾」なども実装している。

一方、Kurashima らは、ある場所で実際に行った体験に関する文書を多く含み、かつ、日時とともに記録されるという Blog 文書の特性に着目し、特定の場所について記述された Blog 文書のテキストから、その場所の訪問者による体験を、時間、空間、動作、対象属性の間の相関ルールマイニングによって抽出する手法を提案している¹²⁾。この手法により、ある場所の上で行われた典型的な人々の活動を把握することが可能となる。

地物に限らず、何らかの一般の対象物が与えられたときに、その評判情報を Web や Blog などの文書から抽出する研究もさかに行われている。小林らは、評価対象表現、属性表現、評価表現の共起パターンを利用することで、これら領域依存の表現セットを効率的に収集する手法を提案している¹³⁾。藤村らは、Web 文書から評判情報を抽出し、肯定・否定・非評価を判定する分類器を構築している¹⁴⁾。また、赤木らは、ある対象に対して、どのような観点から評価されているかを表す評価属性の名称を抽出している¹⁵⁾。本論文では、「八坂神社」の「西楼門」は「朱塗り」というように、ある地物に対して、その地物の構成要素名、および、その外観的な評価を表す語の組を抽出しており、一般オブジェクトに対する評価情報抽出の研究と非常に関連しているが、実世界上の地物の外観は、いつ、どこから、どんな状況下で見るとにより大きく変化する場合がある点が特徴的である。ただし、本論文では、時空間や状況に依存した外観情報の抽出までは行えておらず、今後の研究課題の 1 つとして取り組む必要がある。

3. 目標システムと提案手法の位置付け

本章では、我々が目指している最終的な目標である「文書中の画像を代替する言語的記述の自動生成」システムを実現する要素技術について概説し、4章で提案する「地物の外観情報の Web からの抽出」手法の位置付けを明確にする。

3.1 画像を代替する言語的記述とは

対象の画像を代替する言語的な記述とは、その画像を晴眼者が視覚的に閲覧して容易に獲得することが可能な情報を言語メディアで表現した文字情報のことであり、その画像の内容を表す説明文に相当する。したがって、ある写真を晴眼者が見たとしても、過度な事前知識がない限り獲得することが不可能な情報に関しては、その写真の内容を説明する言語的記述としては不適であると見なす。ただし、多くのユーザが共有しているような一般的なオブジェクトの名称や色概念などの最低限の事前知識はあるものとする。

まったく同じ写真を閲覧したとしても、その閲覧者が持つ事前知識の量に依存して、より具体的な情報を獲得できるか否かには差が生じる。図1の写真に対して「八坂神社」の「西楼門」が既知である者もいれば、未知で「門」としか特定できない者もいる。過度な事前知識がない限り一般に固有名称は分からないが、画像中に名札や表札があれば分かることもある。また、対象の画像を代替する言語的記述として、非視覚的ブラウザの利用者に呈示する場合にも、具体的な情報になればなるほど、その利用者の事前知識の量に依存して、正しく理解（共感）できるか否かに差が生じる。しかしながら、具体的な情報を呈示された方が、ユーザ個々の事前知識にはなかったとしても、別途調べることで正しく理解しやすい。一方、抽象的な情報である色名を呈示したとしても、それから受ける色のイメージは人それぞれであり、共感し難い「赤」よりも「朱」の方がより具体的であり、イメージが限定される。画像を代替する言語的記述として、イメージを共有するためには具体的な情報の方が良いが、事前知識にない可能性も高い。一方、抽象的な情報の方が過度な事前知識は必要ないが、イメージの幅が広く共有し難い。

このように、各々長短はあるものの、本論文では、抽象的な情報、および、具体的な情報を、画像を代替する言語的記述としての適合解とする。一方、対象の画像を晴眼者が見たとしても、過度な事前知識がない限り対応付けることが不可能な情報は、その画像の内容を説明する言語的記述としては不適合とする。以下では、八坂神社の西楼門の写真に対する例を付す。



図1 京都祇園の八坂神社の西楼門の写真の一例

Fig.1 Photograph of Nishi-Romon of Yasaka Shrine.

- 抽象的な文字情報：
 - － 画像中のオブジェクトの一般名称
ex) 門, 楼門, 階段, 森, 道路
 - － オブジェクトの見た目(色や形状, 材質など)
ex) 朱色の, 2階建の門, 石の階段, 緑色の森
 - － 周りのオブジェクトとの相対的な大きさ
ex) 大きな門, 大きな表札
 - － オブジェクトの画像中での絶対的な配置
ex) 中央に門, 上部に森, 下部に道路
 - － オブジェクト間の位置関係
ex) 階段の上に門, その後ろに森
- 具体的な文字情報：
 - － 画像中のオブジェクトの固有名称
ex) 八坂神社・西楼門, 四条通, 祇園交差点
 - － オブジェクトの構造様式
ex) 三間一戸, 屋根が切妻造の
- 不適合な文字情報：
 - － 高さや横幅を表すメートルなどの数値データ
 - － 地理的な所在地(経緯度や住所など)
ex) 京都府京都市東山区祇園町にある
 - － 製作者(会社), 製作時期
ex) 室町時代, 1497年に建てられた
 - － その他, 歴史, 沿革, エピソードなど
ex) 重要文化財の, シンボル(主門)である

3.2 目標システムの実現方法

人間の目をシミュレートするコンピュータ・ビジョンの分野では、メタな文字情報なしに画像だけが与えられた場合に、その画像中に含まれているオブジェクトの名称や外見、また、互いの位置関係などを認識することが最終的な目標の1つである。この一般画像認識技術が実現すれば、画像を代替する言語的記述を自動生成することも可能となるであろう。

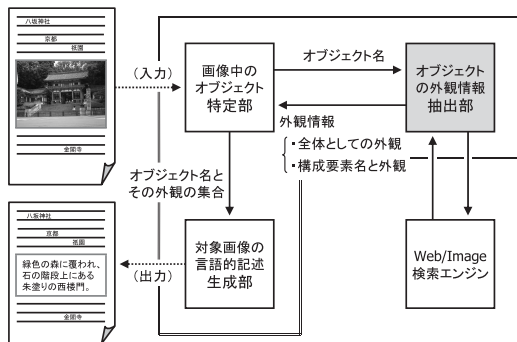


図 2 目標システムの概観

Fig. 2 Overview of our goal system.

一方、我々が目指している最終的な目標は、画像のみが独立して与えられた場合ではなく、文書中にコンテンツとして存在する画像が指定された場合に、その画像の内容を表す説明文を自動生成することである。つまり、言語メディア表現へ変換したい対象の画像自身の内容の視覚的な特徴量しか解析する資料がないという条件下ではなく、対象の画像は Web ページや Blog エントリなどのデジタル文書中に存在し、画像自身だけでなく同一文書中の他のコンテンツも解析する資料として活用できるという条件下を想定する。

最終目標のシステム構成と本論文の位置付けは図 2 のようになる。ただし、本論文の提案手法は、一般オブジェクトではなく、地物限定の外観情報の抽出である。また、以下の 4 つのステップを経ることにより、文書中の対象画像の内容を説明する言語記述を生成する。

Step 1. 画像中のオブジェクト名の初期候補の取得：

「特定部」は、対象の画像を含む文書中のコンテンツのテキスト部から形態素解析器を用いて名詞句を切り出し、対象の画像中のオブジェクトの名称を特定するための初期候補とする。

Step 2. 画像中のオブジェクト名の候補の細粒化：

「抽出部」は、「特定部」から送られた画像中のオブジェクト名の初期候補に対して、全体としての外観、および、構成要素名と外観のペア集合を Web から抽出するため、「Web/Image 検索エンジン」に対して適切に検索クエリを発行し、その検索結果を解析する。そして、各オブジェクト名に対して抽出した外観情報を「特定部」に返す。

Step 3. 画像中のオブジェクト名の特定：

「特定部」は「抽出部」から返されたオブジェクト名と外観記述のペア集合を用いて、対象の画像中に含まれているオブジェクト名を候補集合から選定する。実際には、以下のような複数の尺度を組み合わせることで相応度を評価する。

● 従来の画像インデキシング手法：

文書中の画像に対して、周辺テキスト中のキーワードによってインデキシングする従来の重み付け手法を用いる。ただし、候補のオブジェクト名のうち、初期候補に含まれるものに対してしか評価できない。対象画像との位置関係や構造的な関係などに基づく。

● 画像の文脈と候補名・外観ペアとの照合：

対象画像の周辺テキストの文脈と、オブジェクト名の候補、および、その外観記述のペアとをマッチングすることにより、各オブジェクト名の候補の相応度を評価する。元々の文書中には含まれない初期候補以外のオブジェクト名に対しても評価できる。たとえば、画像の前後一定数までの文からなる周辺テキスト中のキーワード集合との共起頻度により算出する。

● 画像の内容と候補名・外観ペアとの照合：

対象画像の内容の特徴量と、オブジェクト名の候補、および、その外観記述のペアとをマッチングすることにより、各オブジェクト名の候補の相応度を評価する。画像の内容からは、色、形状、テクスチャなどの特徴量を解析できるが、抽出精度を考慮すると、色情報のマッチングが現実的である。対象の画像中の占有面積が大きい色などを代表色とする。一方、オブジェクト名の候補に対して抽出された色名は、日本工業規格 (JIS Z 8102:2001) で規定されている物体色の慣用色名、および、そのマンセル値に基づいて、RGB 色空間上に射影する。HSV 色空間に変換したうえで、両者の色の類似度を計算する¹⁶⁾。精度を向上させるには、前景と背景を分離し、オブジェクト領域を切り出したうえで代表色を求めるといった、より詳細な画像解析を行う必要がある。

Step 4. 画像の内容を説明する言語的記述の生成：

「特定部」は、特定したオブジェクト名に対する外観記述の中から、対象画像の内容や文脈にふさわしいものだけにフィルタリングした後、「生成部」に送る。「生成部」は、あらかじめ定義された自然言語文の生成規則に従い、対象画像を代替する言語的記述を生成し、元文書のテキスト版として出力する。

上述の実現方法がうまく機能するためには、対象の画像中に撮影されたオブジェクトの名称が、その画像を含む文書中のタイトルやボディなどに、適切な限定性を持って言語的に記述されている必要がある。

たとえば、図1のように京都祇園の八坂神社の西楼門という名称の建造物が撮影された写真が存在する文書中に、限定性が粗い「京都」という地名しか記述されていなければ、京都には寺社仏閣などの地物が非常に多数存在し、「京都」の構成要素名として様々な地名が抽出されてしまうため、画像の内容と候補名・外観ペアとの照合だけを考慮してオブジェクト特定を行った場合、候補名の中から「八坂神社」と特定することはできないかもしれない。なぜならば、正解の「八坂神社」と同様に「赤い」「鳥居」を外観情報として持ち、かつ「京都」の構成要素名である「平安神宮」と誤って特定してしまうかもしれないからである。

一方「祇園」という地名しか記述されていなくても、「京都」より限定性が強く、祇園で典型的に写真を撮影される地物としては「八坂神社」「知恩院」「南座」程度であり、この中で「赤い」外観が特徴的な地物は「八坂神社」であるため特定を誤り難い。「八坂神社」と特定できれば、さらに構成要素名を展開することで、「西楼門」まで細かく特定できるかもしれない。

八坂神社の西楼門の写真は多くの場合、建物の朱色、周りの森の緑色、空や道路の灰色などが、代表色として求められる。5章での実験結果を参照すると、表2では「八坂神社」の構成要素名として「本殿」「鳥居」「西楼門」などが抽出できており、表4より、色情報としていずれも「朱」を持つが、「本殿」「西楼門」は「緑」を持たず、一方「鳥居」は「緑」を持っている。したがって、画像の内容と候補名・外観ペアとの照合によってのみ相応度を評価した場合、システムは正解の「西楼門」ではなく「鳥居」と特定してしまう。

最後に「八坂神社」の「鳥居」に対して抽出された外観記述である「大きな」「石の」「朱塗りの」などを用い、「八坂神社の大きな朱塗りの石鳥居」といった言語的記述を対象の画像の説明文として生成する。八坂神社には石鳥居もあり、「石の」は「鳥居」の外観としてはふさわしいが、西楼門は石造ではないため、対象の画像に対してはふさわしくない。このような問題を回避するためには、オブジェクト名を特定する際、画像の内容との照合によってのみ相応度を評価するのではなく、画像の文脈との照合なども考慮する必要がある。

3.3 本論文の位置付け

我々が目指している最終的な目標システムである「文書中の画像を代替する言語的記述の生成」を実現するための方法について述べてきたが、本論文で具体的に提案し評価する手法は、最終的な目標システム構成のうち「抽出部」に相当し、かつ、一般的なオブジェクトに対する外観情報の Web から抽出する手法ではなく、固有名称を持つ地物という特定ドメインに限定したオブジェクトに対する外観情報を Web から抽出する手法について提案する。

地物に限定したのは、地名だけでなく外観情報も付属した道案内文を生成するために、地物の外観情報を抽出する手法について研究してきたことに起因し¹⁷⁾、また、旅行記や観光ガイドから地物画像を単純に除去してテキスト版を生成してしまうと、コンテンツの価値が大きく損なわれるため、地物画像を言語的記述で代替する意義が大きいと考えたためでもある。

本論文で提案する外観情報の抽出手法のうち、対象のオブジェクトの構成要素名を抽出する手法、および、「(オブジェクト名)の(外観候補)(構成要素名)」という型自由な表現パターンに合う記述を Web から収集し、「(外観候補)」の部分を集約して対象のオブジェクトの構成要素ごとの外観記述とする手法に関しては、地物に特有の項を含んではいないため、一般オブジェクトに対してもただちに適用することは可能である。しかしながら、同手法をそのまま適用したとしても、地物に対して適用した結果の精度よりは格段に劣る結果が得られてしまったため、本論文では地物に限定している。地物、特に寺社仏閣などは、それについて記述された Web 文書の中で、外観に関する話題の占める割合が他のドメインに比べて大きいいため、抽出精度が良かったと考えられる。もちろん、地物であっても精度は異なり、たとえば「東京タワー」のように文学作品のタイトルでもある地物は、より劣る結果となる。

地物以外のドメインのオブジェクトに適用した結果を考察すると、Web において外観に関する話題がやや重要視される動植物では、地物よりは劣るが利用できる可能性が感じられた¹⁸⁾。しかし、工業製品では、ファッション要素の強い物を除くと、その機能や性能、価格の方が外観よりも関心事である場合が多く、うまく抽出できなかった。また、著名な人物に対しては、その人の活動(行動)に関する話題が占める割合が高く、「顔」「腕」「脚」などの構成要素名を抽出することさえ困難であった。もちろん、与えられたオブジェクトの名称が人物名であることが分かれば、構成要素名をあらかじめ辞書として保持する方法もとりうる。

本論文の提案手法を用いて Web から抽出してこなくても、地理情報システム (GIS) から取得する方が容易である。

4. 地物の外観情報の Web からの抽出

本章では、地物、特に固有名称を持つ建築物が典型的にどのように見られているかを表す外観情報を、Webマイニングにより抽出する手法について述べる。まず、地物の外観情報について、本論文での定義と適合解の判断基準を明確にする。次に、本章で提案する手法の概要を述べた後、対象の地物の構成要素名を Web から抽出する手法、および、その構成要素ごとの外観情報を Web から抽出する手法について提案を行う。

4.1 地物の外観情報とは

地物の外観情報としては、建築物全体としての色、大きさ、高さ（階数や数値データ）、形状、および、建築部を構成する部分要素である外壁や屋根などの色、材質、大きさ、形状などが考えられる。地物が撮影された写真の内容を説明する言語的記述として適当な情報は、その地物の外観情報のサブセットである。2種類の情報の差は、地物の高さや幅などを表すメートル単位などの数値データ情報を、地物の外観情報としては適合解とするが、一方、地物画像の内容を説明する言語的記述としては、晴眼者が見たとしても、その地物個々について過度な事前知識がない限りは得ることが不可能な情報であるため、不適合とする点である。

4.2 提案手法の概要

文書中の地物画像の周辺テキスト中には、その画像に写っている地物の名称に関しては記述されている場合が多い。しかしながら、その地物の特にどの構成要素について撮られた写真であるかまでは記述されていない場合もある。このような画像と周辺の言語的記述とのギャップは、著者にとって、地物名を読者に伝えることは重要であっても、その地物の特定の構成要素名を伝えることは重要でないと考えている場合に起こりうるが、一方で、その地物の画像を文書中に掲載しているということは、その地物の外観情報は伝えたいと考えているはずである。文書中の画像に写っている地物は、多数の構成要素を持ち、かつ、その構成要素ごとに外観は異なる場合も非常に多い。むしろ、その建築物の全体が一様な外観を持つことの方が珍しいであろう。したがって、対象の地物画像に対して、その地物名しか特定せず、その地物の全体的な外観情報しか抽出しない我々の従来手法では不十分である¹⁷⁾。一方、本論文では、対象の地物を構成する部分要素の名称までを特定し、その地物の特定の構成要素ごとの外観情報を抽出することができれば、その地物画像を代替する言語的記述として、より正確かつ詳細な情報となると考える。もちろん、対象の画像を含む文書中

地物名がまったく記述されていない場合には、その画像自体の内容を画像解析しない限り、その画像に写っている地物の名称や外観情報を抽出することは不可能である。本論文では、地物画像の周辺テキスト中にその地物名は少なくとも記述されており特定できるという前提で、その周辺テキストだけではなく Web 全体から、対象の地物の全体としての外観情報、その地物を構成する部分要素の名称、および、その地物の特定の構成要素ごとの外観情報を抽出する手法を提案する。

4.3 地物の構成要素名の抽出

対象の地物に対して、それを構成する部分要素の名称を Web から抽出する手法について述べる。対象の地物の名称が与えられた場合に、その地物名の直後に「の」を連結した「(地物名)の」という検索クエリを Web 検索エンジンに入力し、その検索結果の文書集合から「(地物名)の」に続く名詞句を形態素解析で切り出して対象の地物の構成要素名の候補語とし、各候補に対して「(地物名)の(構成要素名の候補)」を含む文書数に基づいて相応度を評価するという手法がまず考えられる。しかしながら、日本語の文法における「の」という助詞は様々な目的で使用されるため、単純に、Google などの Web 検索エンジンで検索した結果を解析したとしても、地物の構成要素名とは何ら関係のないキーワードも多数混合してしまう。たとえば「八坂神社」という地物名に対して、この手法を適用すると表 1 のような結果となる。地物の構成要素名として適合解と判定した候補語の頭には「*」印を付し、太字で表す。上位 10 件の適合率は 10.0%、上位 20 件では 15.0%、上位 30 件では 26.6%、上位 40 件では 27.5%である。対象の地物に関する Web 文書の典型的な話題を把握する一助にはなるが、構成要素名を抽出する手法としての適合率は十分ではない。

表 1 Web 検索エンジンによる「八坂神社」の構成要素名の抽出
Table 1 Component extraction for geographic feature
“Yasaka Shrine” by using Web search engine.

地物名	地物の構成要素名の候補語と検索件数(降順)
八坂神社	祭礼 (772), *境内 (761), 近く (661), 祇園祭 (595), お祭り (391), 祭神 (363), 南 (313), 夏 (294), 夏祭り (247), 公式サイト (238) ¹⁰⁾ , 奥 (191), 祭 (189), *神輿 (180), 中 (177), 氏子 (173), *本殿 (169), 西 (168), おけら詣り (163), リソース (155), *鳥居 (148) ²⁰⁾ , *西楼門 (141), 例祭 (133), *石段 (129), 例大祭 (116), *能舞台 (113), 神事 (108), おけら参り (106), 節分 (101), 大晦日 (99), *桜 (99) ³⁰⁾ , *狛犬 (93), *宮司 (92), そば (87), 裏 (82), 写真 (81), 御祭神 (79), *石段下 (77), 東 (76), 裏手 (74), をけら詣り (72) ⁴⁰⁾ , ...
$P_{10} = 1/10$	
$P_{20} = 3/20$	
$P_{30} = 8/30$	
$P_{40} = 11/40$	

一方で、我々は、「(地物名)の」という記述が画像の周辺テキストに現れる場合、単に文書中に現れるという条件と比べて、「(地物名)の」に続く名詞句がその地物を構成する部分要素の名称を表すことが多いことを発見した。一般には、対象のオブジェクトを撮影する際の典型的なテーマを表す語句が、そのオブジェクトの写真の周辺テキストで「(オブジェクト名)の」に続いてよく現れて抽出されることになるが、地物という特定ドメインにおいては、撮影のテーマとなる語句の多くが地物の特徴的な外観を持った構成要素名である場合が多いためであると考えられる。

そこで、我々が発見した以上の経験則を活用し、次のステップを実行することによって、対象の地物の構成要素名を Web から抽出する手法を提案する。

Step 1. 地物の構成要素名の候補語の取得：

対象の地物名の直後に「の」を連結した検索クエリ [(地物名)の] を Web 検索エンジンではなく画像検索エンジンに入力し、その検索結果のスニペットから「(地物名)の」に続く名詞句を日本語形態素解析により切り出し、その地物の構成要素名の候補語とする。

Step 2. 地物の構成要素名の候補語の重み付け：

各候補語 c に対して、対象の地物名と候補語を「の」で接続した [(地物名)の(構成要素名)] という検索クエリ、および、対象の地物名と候補語とを直接連結した [(地物名)(構成要素名)] という検索クエリを画像検索エンジンで処理した結果の検索件数に基づく次式により、対象地物 o の構成要素名としての相応度 $sc_o(c)$ を評価する。

$$sc_o(c) = df^{img}([o\ の\ c]) + df^{img}([oc]) \quad (1)$$

ここで、 $df^{img}([q])$ は、画像検索エンジンで検索クエリ q を処理した結果の検索件数を表す。

4.4 地物の構成要素ごとの外観情報の抽出

対象の地物の名称とそれを構成する部分要素の名称が与えられた場合に、その地物の構成要素ごとの外観情報を抽出する 2 種類の手法について述べる。第 1 の手法は、「(地物名)の(外観修飾句)(構成要素名)」という型自由な表現パターンに記述に注目して抽出する。一方、第 2 の手法は、抽出したい外観属性として、色名、建築様式、階数などを固定的に想定し、対応する「～色」「～造り」「～階建て」といった型依存な表現パターンの記述と、対象の地物名、および、その構成要素名との文書共起頻度に基づく手法である。

4.4.1 型自由な表現パターンに基づく抽出手法

地物の構成要素ごとの外観情報を抽出する第 1 の手法は、「(地物名)の(外観修飾句)(構成要素名)」という外観属性の型に依存しない表現パターンに合致する記述に注目し、「(地物名)の」と「(構成要素名)」に挟まれた文字列中に、対象の地物の構成要素ごとの外観を説明する修飾句が現れることが多いという経験則に基づいて、次のステップで構成される。

Step 1. 対象の外観を表す修飾句の候補語の取得：

対象の地物名の直後に助詞「の」を連結した後に、Google などがサポートするワイルドカード検索を表す「*」を、さらに、その後に構成要素名を連結した [“(地物名)の*(構成要素名)”] という検索クエリを Web 検索エンジンに入力し、その検索結果のスニペットから「(地物名)の」と「(構成要素名)」との間のテキストを切り出す。その際、「(構成要素名)」に直接係る語句だけを日本語係り受け解析により選定し、候補語とする。直接係らないが「(構成要素名)」の直前にある名詞句については、対象地物の構成要素ごとの外観修飾句の候補語としては採用しないが、その名詞句の直後に「(構成要素名)」を連結し、対象地物の構成要素名の候補語としてフィードバックする。

Step 2. 対象の外観修飾句の候補語の重み付け：

各候補語 v に対して、対象の地物名と候補語を助詞「の」で接続した直後に構成要素名を連結した [“(地物名)の(外観候補)(構成要素名)”] という検索クエリを Web 検索エンジンで処理した検索件数に基づく次式を用いて、対象の地物 o の構成要素 c ごとの外観情報(外観修飾句)としての相応度 $sv_{o,c}(v)$ を評価する。

$$sv_{o,c}(v) = df^{web}([“o\ の\ cv”]) \quad (2)$$

ここで、 $df^{web}([q])$ は、Web 検索エンジンで検索クエリ q を処理した結果の検索件数を表す。

4.4.2 型依存な表現パターンに基づく抽出手法

地物の構成要素ごとの外観情報を Web から抽出するための第 2 の手法は、抽出したい地物の外観属性の種類として、色名、素材または建築様式、階層数、屋根の葺き方、形状などをあらかじめ固定的に想定し、地物という特定ドメインで対応する「～色」「～造り」「～層」「～階建て」「～葺き」「～型」といった外観属性の型に依存した表現パターンに合致する記述と、対象の地物名、および、その構成要素名との文書共起頻度に基づき、いずれの外観属性の値抽出についても次のステップで構成される。

Step 1. 外観属性の値の候補語の取得：

対象の地物名をタイトル中に含み、かつ、構成要素名と抽出したい外観属性の型に依存した典型的な表現パターン(「色」「造り」など)とをどこかに含む文書を Web から収集し、表現パターンの直前の語句を日本語形態素解析により切り出し、対象の地物の構成要素の外観属性の値の候補語として採用する。また、個々の地物によらず、頻出する語句をあらかじめ辞書化しておき、それらの語句も候補語として用いる。ただし、構成要素ごとではなく、地物全体としての外観属性の値の候補語を求める場合には、対象の地物名をタイトル中に含み、かつ、抽出したい外観属性の型に依存した典型的な表現パターンとをどこかに含む文書を Web から収集して、以上と同様に解析する。

Step 2. 外観属性の値の候補語の重み付け：

外観属性の値の各候補語 v に対して、外観属性の種類ごとに定義された評価式に基づいて、対象の地物 o の構成要素 c ごとの外観属性 x の値としての相応度 $sv_{o,c}^x(v)$ を計算する。そのうえで、一定閾値以上の相応度を持つ候補語を、対象の地物の構成要素ごとを外観属性の値として採択する。ただし、本論文の実験では、閾値として 0.25 を採用している。

以下、本論文で抽出を試みる地物の外観属性の種類ごとに、その外観属性の値の候補を取得する際に用いる典型パターン、および、相応度の評価式を順に定義する。

● 色名の抽出：

- 候補語の取得パターン：「色」「塗り」
- 候補語の相応度の評価式：

$$sv_{o,c}^c(v) = \frac{v_{o,c}^c(v)}{v_{o,c}^c} \quad (3)$$

$$v_{o,c}^c(v) = df^{title}(o, c, "v 色") \\ + df^{title}(o, c, "v い") \\ + df^{title}(o, c, "v の") \\ + df^{title}(o, c, "v 塗り")$$

$$v_{o,c}^c = df^{title}(o, c, "色") \\ + df^{title}(o, c, "塗り")$$

ここで、 $df^{title}(o, c, q)$ は Web 文書コーパス中で、地物名 o をタイトルに含み、かつ、構成要素名 c と典型パターンに基づく語句 q を含む文書の総数を表す。本論文では、Google の Web 検索エンジンで [intitle:" o " " c " q] という検索クエリを処理した結果の検索件数を用いる。ただし、構成要素名 c についてはオプション引数である。

● 素材または建築様式の抽出：

- 候補語の取得パターン：「造り」「造」「製」
- 候補語の相応度の評価式：

$$sv_{o,c}^m(v) = \frac{v_{o,c}^m(v)}{v_{o,c}^m} \quad (4)$$

$$v_{o,c}^m(v) = df^{title}(o, c, "v 造り") \\ + df^{title}(o, c, "v 造") \\ + df^{title}(o, c, "v 製")$$

$$v_{o,c}^m = df^{title}(o, c, "造り") \\ + df^{title}(o, c, "造") \\ + df^{title}(o, c, "製")$$

● 建物の層数の抽出：

- 候補語の取得パターン：「層」
- 候補語の相応度の評価式：

$$sv_{o,c}^l(v) = \frac{df^{title}(o, c, "v 層")}{df^{title}(o, c, "層")} \quad (5)$$

● 建物の階数の抽出：

- 候補語の取得パターン：「階建て」「階建」
- 候補語の相応度の評価式：

$$sv_{o,c}^f(v) = \frac{v_{o,c}^f(v)}{v_{o,c}^f} \quad (6)$$

$$v_{o,c}^f(v) = df^{title}(o, c, "v 階建て") \\ + df^{title}(o, c, "v 階建")$$

$$v_{o,c}^f = df^{title}(o, c, "階建て") \\ + df^{title}(o, c, "階建")$$

● 屋根の葺き方の抽出：

- 候補語の取得パターン：「葺き」「葺」
- 候補語の相応度の評価式：

$$sv_{o,c}^r(v) = \frac{v_{o,c}^r(v)}{v_{o,c}^r} \quad (7)$$

$$v_{o,c}^r(v) = df^{title}(o, c, "v 葺き") \\ + df^{title}(o, c, "v 葺")$$

$$v_{o,c}^r = df^{title}(o, c, "葺き") \\ + df^{title}(o, c, "葺")$$

● 建物の形状の抽出：

- 候補語の取得パターン：「型」
- 候補語の相応度の評価式：

$$sv_{o,c}^t(v) = \frac{df^{title}(o, c, "v 型")}{df^{title}(o, c, "型")} \quad (8)$$

5. 評価実験と考察

本論文では、文書中の地物画像を代替する言語的記述を自動生成する目標システムを構築するのに必要な要素技術である「地物の外観情報の Web からの抽出」手法として、対象の地物を構成する部分要素の名称の抽出、および、地物の構成要素ごとの外観情報の抽出について提案した。本章では、4章で提案した各々の抽出手法について評価実験と考察を順に行う。

5.1 地物の構成要素名の抽出実験

平均適合率を算出する際に地物の種類に依存しないように、「神社」「神宮」「寺」「城」「タワー」「駅」といった種類からそれぞれ選択した14個の著名な地物の名称に対して、その構成要素名を Web から抽出すると、表2のような結果となる。ただし、画像検索エンジンとして、Google イメージ検索を採用している。「写真」「画像」「イメージ」といった語は、Google イメージ検索エンジンが検索クエリから自動的に除去してしまい、かつ、地物の構成要素名としてふさわしくないため、あらかじめ候補から除外している。

まず、「(地物名)の」という検索クエリを Web 検索エンジンで処理した検索結果から「(地物名)の」に続く語句を対象地物の構成要素名の候補語とし、そのうえで、「(地物名)の(構成要素名の候補語)」という検索クエリを Web 検索エンジンで処理した結果の検索件数に基づいて、各候補語の相応度を評価する手法では、「八坂神社」という地物の構成要素名の抽出精度は、上位10件の適合率が10.0%、上位20件では15.0%、上位30件では26.6%、上位40件では27.5%であったが、同様のプロセスを Web 検索エンジンではなく、画像検索エンジンを用いて行う本論文の提案手法では、上位10件の適合率が30.0%、上位20件では50.0%、上位30件では53.3%、上位40件では55.0%となっており、上位20件まででは約3倍、上位40件まででは約2倍、精度が向上している。

次に、14個の著名な地物名に対する構成要素名の抽出精度に関しては、上位10件の平均適合率が60.7%、上位20件では58.6%であった。さらに、個々の地物に依存せず共通にノイズとして出現している位置関係を表す語、方角を表す語、時間帯を表す語、季節を表す語、および「歴史」という語句をあらかじめストップワードとして除外したとすると、上位10件の平均適合率が73.6%に、上位20件では67.5%にまで向上するため、提案手法は十分な抽出精度を実現できているといえる。

地物の種類ごとに考察すると、6個の神社の名称に対する上位10件の平均適合率は61.6%、上位20件では62.5%、2個の寺名に対する上位10件の平均適合率は60.0%、上位20件では57.5%、2個の城名に対する上位10件の平均適合率は90.0%、上位20件では77.5%、2個のタワーの名称に対する上位10件の平均適合率は40.0%、上位20件でも40.0%、1個の駅名に対する上位10件の平均適合率は70.0%、上位20件では55.0%であった。各種類から選択している数が異なり、また、各々の数が非常に少ないため、種類ごとの正確な平均適合率を表しているとはいえないが、顕著な傾向も見受けられる。たとえば、城名は平均適合率が他に比べて非常に高く、一方、タワーの名称は他に比べて低い。タワーや塔といった種類に属する地物の多くは、典型的な構成要素名として「展望台」「階段」「エレベータ」程度しか持たず、また、タワーの写真撮影する場合、特定の構成要素に焦点を当てて撮影することも珍しいため、当然、その写真を Web 文書中に掲載する際、周辺テキストに記述する撮影テーマとして、構成要素名が使われることが少ない。実験の結果では、これらの「展望台」という構成要素名は確実に上位で抽出できているため、再現率が頭打ちに到達しているのではないかと考えられる。その他の顕著な傾向としては、神社の名称と寺名の平均適合率が非常に似ていることが見受けられる。

5.2 地物の構成要素ごとの外観情報の抽出

まず、前節で用いた14個の著名な地物名に対して、「(地物名)の(外観修飾句)(構成要素名)」という外観属性の型に依存しない表現パターンに注目した第1の手法により、その地物の外観情報を Web から抽出すると、表3のような結果となる。ただし、Web 文書検索エンジンとして、Google ウェブ検索を採用している。14個の著名な地物名に対する外観情報としての抽出精度は、上位10件の平均適合率が51.4%であった。一方、各地物を撮った写真の内容を説明する言語的記述としての抽出精度は、上位10件の平均適合率が50.7%であった。ただし、表3では、後者の意味で適合であると判定した外観修飾句と構成要素名のペアの頭に「*」印を付し、太字で表す。たとえば、各々の平均適合率を算出するのに用いた上位10件中では「出雲大社」に対する「23メートルの/大鳥居」のみ、上位10件より以降の順位では「東京タワー」に対する「250mの/特別展望台」や「150mの/大展望台」などを、後者の意味においては不適合と見なしている。

表 2 画像検索エンジンを用いた地物の構成要素名の抽出例

Table 2 Component extraction for each geographical feature by using Image search engine.

地物名	地物の構成要素名の候補語, および, 評価値 (降順)
八坂神社 $P_{10} = 3/10$ $P_{20} = 10/20$ $P_{30} = 16/30$ $P_{40} = 22/40$	祭礼 (215), *境内 (213), 前 (165), 祇園祭 (150), 天王祭 (107), *神輿 (100), *本殿 (88), 例大祭 (85), 節分祭 (59), 近く (57) ₁₀ , *鳥居 (52), *狛犬 (51), *桜 (39), *舞殿 (37), *西楼門 (33), *夏祭 (32), *サツキ (32), 祭 (31), *地車 (29), 祭り (25) ₂₀ , *提灯 (16), *御旅所 (16), 祭神 (15), *正門 (14), 奥 (13), *枝垂れ桜 (11), *ケヤキ (10), おけら参り (9), 舞妓 (9), *摂社 (8) ₃₀ , *紅葉 (8), 裏手 (8), 下 (7), *南楼門 (7), *森 (6), おけらまいり (6), *末社 (5), *シイ (5), 御田祭 (5), *十二冠女 (5) ₄₀ , ドライブ (5), ...
厳島神社 $P_{10} = 9/10$ $P_{20} = 12/20$	*大鳥居 (346), *鳥居 (274), *回廊 (80), *境内 (61), *国宝 (56), *社殿 (51), *神輿 (41), *宝物 (32), 歴史 (29), *入り口 (27) ₁₀ , 中 (24), *鹿 (18), *狛犬 (17), 拳式 (17), 舞楽 (9), 祭神 (9), 朱 (9), 由来 (6), 海中 (6), *外宮 (5) ₂₀ , 朱丹 (4), 位置 (4), 東側 (3), *建築物 (2), 花火 (2), シンボル (2), ...
生田神社 $P_{10} = 5/10$ $P_{20} = 11/20$ $P_{30} = 13/30$	*会馆 (51), 前店 (46), 神戸 (39), 西側 (27), *境内 (11), *楼門 (11), *参道 (10), 初詣 (9), 獅子頭保存会 (9), *生田の森 (7) ₁₀ , *提灯 (5), *桜 (5), 参拝 (5), お正月 (5), *本殿 (4), *社頭の鳥居 (4), *参道の鳥居 (4), 観桜祭 (4), 夏祭り (4), *里桜 (3) ₂₀ , 生田祭 (3), *桜門 (3), *記念碑 (3), 西 (3), 裏 (2), 葉っぱ (2), 地域性 (2), 妙 (2), 中 (2), 豆まき (2) ₃₀ , 下 (2), おわたり (2), ...
平安神宮 $P_{10} = 7/10$ $P_{20} = 16/20$	*神苑 (506), *桜 (284), *鳥居 (164), *大鳥居 (151), *花菖蒲 (132), *庭園 (110), 近く (105), *睡蓮 (76), 新能 (59), 初詣 (44) ₁₀ , *紅枝垂桜 (34), *枝垂れ桜 (30), *しだれ桜 (23), *菖蒲 (22), *庭 (15), *夜桜 (15), *お庭 (14), *門 (13), *カキツバタ (13), 正面 (12) ₂₀ , *入り口 (11), 中 (7), 創建 (7), ...
伊勢神宮 $P_{10} = 6/10$ $P_{20} = 12/20$	*内宮 (1251), *外宮 (703), 式年遷宮 (319), *別宮 (77), お木曳 (48), 近く (46), *御神木 (45), *参道 (42), *宇治橋 (40), 旅 (32) ₁₀ , *勾玉池 (31), *森 (15), *錦鯉 (14), *木 (13), 巻 (11), 門前 (8), 祭事 (6), 中 (6), *伊弉宮 (5), 歴史 (5) ₂₀ , *神馬 (5), *杉 (5), *遷拝殿 (4), 宮大工 (4), 神事 (4), ...
出雲大社 $P_{10} = 7/10$ $P_{20} = 14/20$	*本殿 (207), 前 (186), *境内 (119), *神楽殿 (75), 近く (39), すべて (35), *正門 (32), *御本殿 (32), *おみくじ (32), *鳥居 (27) ₁₀ , 初詣 (22), 復元 (20), *しめ縄 (19), *神殿 (17), *大鳥居 (17), *駐車場 (15), *御札 (9), 東 (9), *巫女 (8), *松並木 (7) ₂₀ , *分社 (7), 模型 (7), *注連縄 (6), 例大祭 (5), ...
金閣寺 $P_{10} = 6/10$ $P_{20} = 13/20$	前 (135), *庭園 (102), *紅葉 (96), *杜若 (62), 近く (49), *参道 (40), 雪景色 (21), *鳳凰 (20), 金色 (18), *駐車場 (15) ₁₀ , *茶店 (14), 製作 (14), 撮影 (11), *入り口 (11), *一枚天井 (10), *屋根 (8), *庭 (8), 様子 (8), *茶室 (7), *もみじ (6) ₂₀ , 茶席 (5), *お守り (5), 巻 (5), 横 (4), *裏山 (4), ...
銀閣寺 $P_{10} = 6/10$ $P_{20} = 10/20$	前 (172), *参道 (124), *庭園 (123), 近く (116), *紅葉 (105), *庭 (36), *入り口 (31), 横 (25), *苔 (23), 方 (22) ₁₀ , 観光 (20), 全景 (19), *総門 (15), *家 (15), *ススキ (13), 後 (12), 雪 (11), そば (11), 死線 (7), *門 (6) ₂₀ , 中 (5), *駐車場 (5), 上 (5), 様子 (3), *竹 (3), 味 (3), *石庭 (2), *敷地 (2), ...
大阪城 $P_{10} = 10/10$ $P_{20} = 17/20$ $P_{30} = 23/30$ $P_{40} = 29/40$	*梅林 (1508), *天守閣 (941), *桜 (354), *梅 (253), *西の丸 (157), *西の丸庭園 (144), *桃園 (123), *さくら (123), *花 (94), *石垣 (92) ₁₀ , *大手門 (78), 近く (62), 前 (60), *堀 (58), *お堀 (48), *桃 (45), 歴史 (34), *桜門 (33), *日本庭園 (31), *梅林公園 (30) ₂₀ , お花見 (17), 四季 (15), *紅葉 (14), *梅花 (12), *広場 (11), 地図 (11), *黄葉 (11), *外濠 (10), *山里 (9), 住所 (9) ₃₀ , 菊花展 (9), *石 (9), *櫓 (7), 隣 (7), 謎 (6), *紅梅 (6), *銀杏 (6), *玄關口 (6), 今 (5), *六番桶 (4) ₄₀ , ミニちょうちん (4), 研究 (4), ...
名古屋城 $P_{10} = 8/10$ $P_{20} = 14/20$	*天守閣 (245), *桜 (135), *お堀 (63), *シャチホコ (56), *金鯱 (52), *本丸 (47), 近く (37), 四季 (37), *金のシャチホコ (36), *石垣 (32) ₁₀ , *金の鯱 (30), *天守 (27), 歴史 (24), 西 (22), *鹿 (21), *梅 (21), *外堀 (21), 東 (19), ライトアップ (16), *堀 (13) ₂₀ , シンボル (12), そば (10), 秋 (10), 敷地内 (7), *藤 (5), ...
東京タワー $P_{10} = 4/10$ $P_{20} = 8/20$	DVD(241), 下 (223), 夜景 (170), *展望台 (169), *大展望台 (156), クリスマス (135), 真下 (107), *階段 (83), *特別展望台 (56), 足元 (40) ₁₀ , *キティー (38), 高さ (36), *クリスマスツリー (34), 試写会 (16), 消灯 (14), 足下 (11), 人気 (11), 歩き方 (11), *影 (10), *展望室 (9) ₂₀ , 電気 (9), 都市伝説 (7), 夕暮れ (6), ミニ提灯 (6), *外階段 (5), *ボタン (4), 頂上 (4), *トイレ (3), ドライブ (3), ...
通天閣 $P_{10} = 4/10$ $P_{20} = 8/20$	*ピリケン (105), *展望台 (105), 下 (72), 近く (69), 真下 (55), *ピリケンさん (36), *ネオン (29), 上 (25), そば (14), 歌姫 (13) ₁₀ , *入り口 (12), *エレベータ (10), 夜 (9), 中 (9), 守り神 (8), *展望室 (7), 足元 (7), モデル (6), *ピリケン像 (6), 艶姿 (5) ₂₀ , *足 (5), *二階 (4), 全景 (4), 勇姿 (3), 所 (3), たもと (3), ...
東京駅 $P_{10} = 7/10$ $P_{20} = 11/20$	前 (1108), *コンコース (344), *丸の内口 (284), 新幹線 (246), 近く (240), *名店街 (150), *ホーム (149), *赤レンガ (131), *日本橋口 (128), *駅弁屋 (121) ₁₀ , モデル (93), 東北 (63), 夜景 (47), 定番 (47), *居酒屋 (46), 発車 (39), *赤レンガ駅舎 (33), *駅舎 (31), EF66(30), *ドーム (23) ₂₀ , ミレナリオ (20), *ホテル (18), *大丸デパート (16), *店舗 (14), *天井 (14), ご案内 (11), *トイレ (10), *クリスマスツリー (9), ...
国会議事堂 $P_{10} = 3/10$ $P_{20} = 8/20$	前 (841), 正面 (71), 周辺 (56), 中 (45), 裏 (31), *中庭 (30), *正門 (25), 中央 (23), 横 (21), *前庭 (21) ₁₀ , *正門前 (19), 全景 (18), *中央広間 (12), *建物 (12), 裏手 (11), 裏側 (10), 向かい (7), *正面玄関 (6), 夕暮れ (6), *入り口 (6) ₂₀ , *時計台 (5), 北側 (5), 設計 (4), *食堂 (4), *入口 (4), 石 (4), 街 (3), ...

(注 1) 「写真」「画像」「イメージ」は, Google イメージ検索がクエリから自動的に除去するため, 候補からあらかじめ除外している.

(注 2) 各地物の構成要素名として適合解と判定した候補語の頭に「*」印を付し, 太字で表示. ただし, 「八坂神社」「国会議事堂」のように, 同一名称の地物が複数存在する場合には, その名称を持つ地物の何れかの構成要素名であれば適合解と判定している.

(注 3) 共通してノイズとして出現しており「近く」「周辺」「傍」「前」「後」「奥」「裏」「横」「隣」「中」「上」「真上」「下」「真下」といった対象の地物を基点とした相対的な位置関係を表す語「東」「西」「南」「北」といった方角を表す語「春」「夏」「秋」「冬」「四季」といった四季を表す語, および「歴史」をストップワードとして除外したとすると, 14 個の地物名に対する構成要素名の抽出結果の上位 10 件の平均適合率 \overline{P}_{10} は 6.1/10 から 7.4/10 に約 13% (相対的に約 1.21 倍に), 上位 20 件の平均適合率 \overline{P}_{20} は 11.7/20 から 13.5/20 に約 9% (相対的に約 1.15 倍に) 改善される.

表 3 型自由な表現パターンに基づく地物の構成要素ごとの外観抽出の例

Table 3 Type-free visual description extraction for each component of geographic feature.

地物名	外観修飾句と地物の構成要素名のペア, および, 評価値(降順)
八坂神社 $P_{10} = 6/10$ $P_{20} = 7/20$	*赤い/鳥居 (291), *三基の/神輿 (269), *大きな/鳥居 (198), *石の/鳥居 (126), *朱塗りの/鳥居 (91), 境内の/桜 (88), 有名な/祇園祭 (74), 祭礼の/祇園祭 (73), 夏の/祭礼 (39), *大きな/ケヤキ (38) ₁₀ , 入り口の/鳥居 (33), 夏の/例大祭 (29), 正面の/鳥居 (28), 入口の/鳥居 (22), 南の/鳥居 (20), 枝垂桜の/桜 (18), 南側の/鳥居 (17), 正門である/西楼門 (15), *朱塗りの/西楼門 (14), 東の/鳥居 (14) ₂₀ , 境内の/舞殿 (9), 祇園祭の/祭礼 (9), *さして広くない/境内 (6), 合同の/祭礼 (6), 角のある/狛犬 (4), 主門である/西楼門 (4), ...
厳島神社 $P_{10} = 8/10$ $P_{20} = 15/20$	*赤い/鳥居 (691), *大きな/鳥居 (382), 現在の/社殿 (231), *朱の/大鳥居 (179), かわいい/鹿 (149), *朱色の/大鳥居 (129), *朱丹の/大鳥居 (101), *赤い/大鳥居 (98), *でっかい/鳥居 (89), *壮麗な/社殿 (89) ₁₀ , *石の/鳥居 (85), *朱塗りの/社殿 (76), *小さな/鳥居 (66), 有名な/大鳥居 (52), *ライトアップされた/鳥居 (36), *真っ赤な/鳥居 (34), *朱の/回廊 (29), *華麗な/社殿 (26), *朱色の/回廊 (25), 奉納した/国宝 (22) ₂₀ , *海中の/大鳥居 (18), 数多くの/国宝 (17), *海中の/鳥居 (16), *壮大な/社殿 (15), シンボルの/大鳥居 (13), ...
生田神社 $P_{10} = 4/10$	*朱色の/鳥居 (74), *赤い/鳥居 (65), 三が日の/初詣 (62), *広い/本殿 (61), 初詣の/参拝 (61), 三が日の/参拝 (22), *折れた/鳥居 (11), やすらぎの/森 (5), 裏手にある/神戸 (4), 年始の/初詣 (3) ₁₀ , ...
平安神宮 $P_{10} = 8/10$ $P_{20} = 16/20$	*大きな/鳥居 (772), *満開の/桜 (581), *赤い/鳥居 (499), 左近の/桜 (362), *広い/庭園 (183), 神苑の/桜 (176), *広大な/庭園 (144), *でっかい/鳥居 (133), *大きな/門 (131), *朱色の/鳥居 (129) ₁₀ , 回遊式の/庭園 (102), 正面の/門 (89), *巨大な/鳥居 (67), *ライトアップされた/桜 (66), *大きい/鳥居 (60), *朱色の/門 (58), *赤い/門 (57), *朱の/鳥居 (56), *日本の/庭園 (55), *朱塗りの/大鳥居 (49) ₂₀ , *赤い/大鳥居 (48), *朱色の/大鳥居 (35), 元旦の/初詣 (30), 神苑の/花菖蒲 (28), *朱塗りの/門 (27), 庭園である/神苑 (26), ...
伊勢神宮 $P_{10} = 2/10$	*深い/森 (604), *長い/参道 (401), 内宮の/別宮 (331), 神の/森 (302), 五十鈴川にかかる/宇治橋 (202), 内宮の/宇治橋 (154), 内宮の/参道 (127), 外宮の/別宮 (90), 静かな/森 (86), 次の/式年遷宮 (82) ₁₀ , *杉の/森 (56), 外宮の/勾玉池 (51), 素晴らしい/森 (46), 式年遷宮の/お木曳 (42), 外宮の/参道 (39), 次回の/式年遷宮 (37), ...
出雲大社 $P_{10} = 6/10$ $P_{20} = 11/20$	*大きな/しめ縄 (571), *大きな/鳥居 (403), *広い/境内 (337), 現在の/本殿 (270), 近くの/駐車場 (151), *巨大な/しめ縄 (103), *太い/しめ縄 (102), 23メートルの/大鳥居 (95), *大きな/駐車場 (78), 日本一の/大鳥居 (75) ₁₀ , *巨大な/神殿 (74), *大きい/しめ縄 (54), *石の/鳥居 (52), 日本一の/しめ縄 (46), *銅の/鳥居 (42), *大きな/神殿 (42), 正面の/鳥居 (41), 参道の/鳥居 (33), 拝殿の/しめ縄 (33), 現在の/御本殿 (28) ₂₀ , 神楽殿の/大しめ縄 (23), 古代の/神殿 (22), 無料の/駐車場 (21), 古代の/御本殿 (19), *巨大な/本殿 (17), *白い/大鳥居 (14), *大きな/本殿 (12), 婚儀殿の/大しめ縄 (12), *碧銅の/鳥居 (7), 正門の/鳥居 (6), ...
金閣寺 $P_{10} = 5/10$	*素晴らしい/紅葉 (283), *日本の/庭園 (136), 参道の/紅葉 (132), 現在の/様子 (99), 境内の/紅葉 (92), 今の/様子 (54), *美しい/雪景色 (45), 入り口の/紅葉 (42), *古い/庭園 (35), *こけら葺き/屋根 (26) ₁₀ , *青い/紅葉 (26), *てっぺんの/鳳凰 (25), *綺麗な/雪景色 (22), 見学の/様子 (21), 雪の/様子 (20), ...
銀閣寺 $P_{10} = 7/10$	*美しい/庭園 (902), *邪魔な/苔 (406), 大切な/苔 (305), *とても邪魔な/苔 (287), *素晴らしい/紅葉 (257), 友人の家 (222), *日本の/庭園 (114), *見事な/庭園 (97), *苔の/庭 (88), 周りの/紅葉 (56) ₁₀ , お庭の/紅葉 (38), 大事な/苔 (36), 裏の/庭園 (33), *見事な/苔 (32), 近くの/家 (32), 参道の/入り口 (28), *苔の/庭園 (26), ...
大阪城 $P_{10} = 5/10$ $P_{20} = 10/20$	梅の花 (20400), 桜の花 (16300), *満開の/桜 (952), 現在の/天守閣 (316), 梅林の/梅 (130), *高い/石垣 (127), 早咲きの/梅 (111), *満開の/梅 (106), *大きな/石垣 (66), *巨大な/石垣 (63) ₁₀ , *ピンクの/梅 (65), 堀の/石垣 (43), 西の丸庭園の/桜 (42), 天守閣の/石垣 (42), 周りの/桜 (29), *大きな/天守閣 (28), 桃園の/桃 (27), *満開の/梅林 (24), *深い/堀 (23), *埋められた/堀 (19) ₂₀ , 外堀の/石垣 (16), 梅林の/花 (15), 本丸の/天守閣 (12), *黒い/天守閣 (10), ...
名古屋城 $P_{10} = 4/10$ $P_{20} = 10/20$	*金の/シャチホコ (21200), *満開の/桜 (381), シンボルである/金のシャチホコ (311), *立派な/石垣 (115), 天守台の/石垣 (53), *綺麗な/桜 (50), 天守閣の/石垣 (47), 有名な/金の鯨 (36), 堀の/桜 (34), お堀の/桜 (33) ₁₀ , *大きな/天守閣 (27), *紅白の/梅 (21), 天守閣の/金鯨 (20), 天守閣の/金の鯨 (18), *大きな/堀 (13), お堀の/鹿 (11), 鯨の/歴史 (11), *埋め立てられた/外堀 (10), *大小の/天守閣 (9), *小さな/天守閣 (9) ₂₀ , 本丸の/天守閣 (9), ...
東京タワー $P_{10} = 2/10$ $P_{20} = 6/20$	今年の/クリスマス (47600), 美しい/夜景 (10500), 見える/夜景 (993), 綺麗な/夜景 (968), *大きな/クリスマスツリー (815), 前の/クリスマス (655), *長い/階段 (545), きれいな/夜景 (543), 素晴らしい/夜景 (511), すばらしい/夜景 (500) ₁₀ , 倍の高さ (473), *きらめく/夜景 (429), 巨大な/クリスマス (399), *巨大な/クリスマスツリー (344), 250mの/特別展望台 (289), 150mの/大展望台 (265), *輝く/夜景 (256), 特典の/DVD (164), *大きな/階段 (148), 地上 250mの/特別展望台 (141) ₂₀ , 地上 150mの/大展望台 (140), ...
通天閣 $P_{10} = 1/10$	ビリケンさんの/足 (561), 幸運の/神様ビリケン (239), 有名な/ビリケン (94), 商店街の/入り口 (86), 新しい/ネオン (78), 上の/展望台 (73), 5階の/展望台 (65), 夜の/ネオン (61), *赤い/ネオン (57), 最上階の/展望台 (37) ₁₀ , *丸い/ネオン (36), *派手な/ネオン (36), 日立の/ネオン (32), ビリケン像の/足 (32), てっぺんの/ネオン (22), 展望台の/ビリケン (20), 展望台にある幸運の/神様ビリケン (16), 光る/ネオン (14), まるい/ネオン (13), ...
東京駅 $P_{10} = 6/10$ $P_{20} = 11/20$	周辺の/ホテル (20700), 近くの/居酒屋 (14200), *高い/天井 (11800), 素晴らしい/夜景 (298), *古い/駅舎 (231), *立派な/駅舎 (189), *新しい/駅舎 (175), *煉瓦の/駅舎 (105), *大きな/ドーム (105), ドームの/天井 (104) ₁₀ , *丸い/ドーム (92), 丸の内の/東京ミナライオ (81), *煌く/夜景 (70), *レンガ造りの/駅舎 (69), そばの/居酒屋 (69), *ドーム型の/天井 (66), 前の/赤レンガ (57), 丸の内口の/赤レンガ (42), 地下の/名店街 (40), *美しい/赤レンガ (36) ₂₀ , *煉瓦造りの/駅舎 (35), 地下街の/居酒屋 (29), 天井の/ドーム (27), *長い/コンコース (22), ...
国会議事堂 $P_{10} = 7/10$	*高い/建物 (623), *大きな/建物 (605), *白い/建物 (523), *立派な/建物 (502), *重厚な/建物 (162), 隣の/建物 (129), *二つの/塔 (113), 建物の/正面 (90), *荘厳な/建物 (55), 中央の/塔 (51) ₁₀ , 中の/食堂 (41), *大きな/塔 (35), 有名な/時計台 (27), 横の/建物 (26), 門の/入り口 (22), 真ん中の/建物 (17), ...

表 4 型依存な表現パターンに基づく地物の構成要素ごとの外観抽出の例

Table 4 Type-dependent visual description extraction for each component of geographic feature.

地物名と構成要素名	色/い/の/塗り	造り/造/製	層	階建て/階建	葺き/葺	型(形状)
o: 八坂神社	朱, 赤, 白	祇園	閾値以上の候補なし	2, 3, 4	桧皮, 檜皮	閾値以上の候補なし
c: null (全体)	(33, 30, 27/79)	(34/113)		(1, 1, 1/4)	(7, 5/16)	
o: 八坂神社	朱, 赤	祇園	閾値以上の候補なし	2, 4	桧皮, 檜皮	閾値以上の候補なし
c: 本殿	(23, 11/27)	(33/68)		(1, 1/2)	(5, 5/13)	
o: 八坂神社	朱, 赤, 白, 緑	石	3, 木	2, 3	銅板	明神
c: 鳥居	(8, 6, 5, 5/13)	(9/30)	(2, 1/3)	(1, 1/2)	(2/5)	(2/6)
o: 八坂神社	朱, 丹	祇園	重	2	柿	明神, 瓢箪
c: 西楼門	(15, 6/17)	(20/29)	(2/3)	(1/1)	(1/1)	(1, 1/2)
o: 大阪城	赤, 白, 青, 金	閾値以上の候補なし	2	8	銅	閾値以上の候補なし
c: null (全体)	(208, 188, 188,.../366)		(44/163)	(13/51)	(3/6)	
o: 大阪城	金, 白, 赤, 黒	鉄筋コンクリート	5	8	銅	閾値以上の候補なし
c: 天守閣	(48, 35, 26, 16/73)	(33/108)	(25/43)	(10/23)	(2/3)	
o: 大阪城	緑, 黒, 赤	閾値以上の候補なし	5	8	銅	閾値以上の候補なし
c: 天守閣, 屋根	(9, 3, 3/6)		(5/8)	(2/5)	(2/3)	
o: 大阪城	緑, 金, 白, 黒	閾値以上の候補なし	5	閾値以上の候補なし	候補なし	閾値以上の候補なし
c: 天守閣, 壁	(10, 6, 4, 3/11)		(5/11)			
o: 東京タワー	白, 赤, 黒, 青, オレンジ	鉄骨	2	閾値以上の候補なし	閾値以上の候補なし	タワー
c: null (全体)	(1605, 1566,.../1130)	(107/316)	(435/494)			(270/412)
o: 東京タワー	オレンジ, 赤, ピンク, 白	閾値以上の候補なし	2	4, 2	銅板	閾値以上の候補なし
c: 大展望台	(50, 43, 41, 40/135)		(8/13)	(13, 5/18)	(1/1)	
o: 東京タワー	オレンジ, 白, 赤	閾値以上の候補なし	2	4	候補なし	閾値以上の候補なし
c: 特別展望台	(33, 30, 26/98)		(10/15)	(10/13)		
o: 東京タワー	白, 黒, 赤	閾値以上の候補なし	閾値以上の候補なし	閾値以上の候補なし	候補なし	閾値以上の候補なし
c: 外階段	(117, 115, 112/370)					

次に、外観属性の型に依存した表現パターンと、対象の地物名、および、その構成要素名との文書共起頻度に基づく第2の手法を用いて、「八坂神社」「大阪城」「東京タワー」という地物名に対して、各々の地物の全体としての外観情報、および、その地物の3種類の構成要素ごとの外観情報をWebから抽出すると、表4のような結果となる。第1の手法では抽出できない、地物の全体としての外観情報を抽出できている。また、地物の構成要素ごとの外観情報に関しても、第1の手法では抽出に失敗したが、第2の手法では成功したのものもある。たとえば、「八坂神社」の「西楼門」の外観情報として、第2の手法でのみ、「2階建ての」という外観修飾句を抽出できており、「八坂神社」の「西楼門」の写真の内容を代替する言語的記述としても有用である。一方で、単純な文書共起頻度に基づく手法であるためにノイズも多くなっている。たとえば、「八坂神社」の「西楼門」の建築様式として「祇園造り」が抽出されているが、正しくは「切妻造り」であり不適である。「西楼門」に対して誤って抽出された「祇園造り」を外観として持つのは「本殿」であり、正しく抽出できている。第1の手法では抽出できない適切な外観修飾句を第2の手法では抽出できる場合もあり、この長所を残しつつ、不適切な外観修飾句が対応付けられてしまうのを回避する手法について、今後、検討していく。

6. おわりに

本論文では、文書中の画像、特に地物(建築物)が撮影された写真に対して、その内容を説明する言語的記述を生成することを目標に、実現に必要な要素技術である「地物の外観情報のWebからの抽出」手法について提案した。14個の著名な地物名に対する評価実験の結果、地物の構成要素名の抽出手法としては上位10件の平均適合率が61.6%、地物の外観情報の抽出手法としては上位10件の平均適合率が51.4%であった。

今後は、提案手法をより多くの地物名に対して適用し評価したうえで、地物の外観情報としての相応度を評価する尺度の改良を行う。また、最終的な目標である「文書中の画像を代替する言語的記述の自動生成」システムの実装にも取り組んでいく。

謝辞 本研究の一部は、文部科学省研究委託事業「知的資産の電子的な保存・活用を支援するソフトウェア技術基盤の構築」、異メディア・アーカイブの横断的検索・統合ソフトウェア開発(代表者: 田中克己)、および、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しいIT基盤技術の研究」、計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」(代表者: 田中克己, A01-00-02, 課題番号: 18049041)により、ここに記して謝意を表す。

参 考 文 献

- 1) Gudivada, V. and Raghavan, V.: Content-Based Image Retrieval-Systems, *IEEE Computer*, Vol.28, No.9, pp.18-22 (1995).
- 2) 彌富 仁, 萩原将文: ファジー推論ニューラルネットワークを用いた風景画像からの知識抽出と認識, 電子情報通信学会論文誌, Vol.82-D2, No.4, pp.685-693 (1999).
- 3) 栗田多喜夫, 加藤俊一, 福田郁美, 坂倉あゆみ: 印象語による絵画データベースの検索, 情報処理学会論文誌, Vol.33, No.11, pp.1373-1383 (1992).
- 4) 出原 博, 藤本典幸, 竹野 浩, 萩原兼一: WWW 画像検索における画像周辺の HTML 構文構造を考慮した画像説明文の抽出手法, 信学会技術研究報告, DE2005-136, pp.19-24 (2005).
- 5) 是津耕司, 田中克己: Web からの画像の文脈情報の抽出と提示, 第 14 回データ工学ワークショップ (DEWS'03) 論文集, 6-P-05 (2003).
- 6) 相良直樹, 砂山 渡, 谷内田正彦: HTML テキストの重要文を用いた画像ラベリング手法, 信学会論文誌, Vol.J87-D1, No.2, pp.145-153 (2004).
- 7) 竹内謹治, 黄瀬浩一, 松本啓之亮: 語の共起の統計情報を用いた画像と説明文の対応付け, 信学会総合大会講演論文集 2005, D-5-1 (2005).
- 8) 竹内謹治, 黄瀬浩一: 類似画像とキーワードを利用した Web 画像の説明文抽出, 情報処理学会研究報告「自然言語処理」, Vol.2006, No.1, pp.7-12 (2006).
- 9) 岡田 真, 浜田浩史, 宝珍輝尚: マルチメディアデータの効率的検索のためのキーワード自動抽出手法, 情報処理学会研究報告「自然言語処理」, Vol.2005, No.94, pp.73-78 (2005).
- 10) Tezuka, T. and Tanaka, K.: Landmark Extraction: A Web Mining Approach, *COSIT'05, LNCS Vol.3693*, pp.379-396 (2005).
- 11) Tezuka, T. and Tanaka, K.: Temporal and Spatial Attribute Extraction from Web Documents and Time-Specific Regional Web Search System, *W2GIS'04, LNCS Vol.3428*, pp.14-25 (2004).
- 12) Kurashima, T., Tezuka, T. and Tanaka, K.: Blog Map of Experience: Extracting and Geographically Mapping Visitor Experiences from Urban Blogs, *WISE'05, LNCS, Vol.3806*, pp.496-503 (2005).
- 13) 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: テキストマイニングによる評価表現の収集, 情報処理学会研究報告「自然言語処理」, Vol.2003, No.23, pp.77-84 (2003).
- 14) 藤村 滋, 豊田正史, 喜連川優: 文の構造を考慮した評判抽出手法, 第 16 回データ工学ワークショップ (DEWS'05) 論文集, 6C-i8 (2005).
- 15) 赤木法生, 大島裕明, 小山 聡, 田島敬史, 田中克己: レビューページ例からの属性抽出に基づくレビューページ検索, 第 17 回データ工学ワークショップ (DEWS'06) 論文集, 2C-i10 (2006).
- 16) Smith, J.R. and Chang, S.-F.: VisualSEEK: A Fully Automated Content-Based Image Query System, *ACM/Multimedia'96*, pp.87-98 (1996).
- 17) 服部 峻, 手塚太郎, 田中克己: サーチエンジンのメタデータを用いた道案内補助のための外観情報の抽出, 情報処理学会研究報告「データベースシステム」, Vol.2006, No.78, pp.169-176 (2006).
- 18) 服部 峻, 手塚太郎, 田中克己: オブジェクトの外観情報の Web マイニング, 電子情報通信学会第 18 回データ工学ワークショップ (DEWS'07) 論文集, L4-6 (2007).

(平成 18 年 9 月 15 日受付)

(平成 19 年 2 月 27 日採録)

(担当編集委員 石川 博, 有次正義, 片山 薫, 木俣 豊, 中島伸介)



服部 峻 (学生会員)

2004 年京都大学工学部情報学科卒業。2006 年同大学大学院情報学研究科社会情報学専攻修士課程修了。同年同博士後期課程入学, 現在に至る。主にユビキタス社会の情報アクセス技術の研究に従事。IEICE, DBSJ 各学生会員。



手塚 太郎 (正会員)

2005 年京都大学大学院情報学研究科社会情報学専攻博士後期課程修了。同年京都大学大学院情報学研究科社会情報学専攻助手。2007 年同助教, 現在に至る。博士 (情報学)。主に地域情報検索システム, Web からの知識発見の研究に従事。DBSJ 会員。



田中 克己 (正会員)

1974 年京都大学工学部情報工学科卒業。1976 年同大学大学院修士課程修了。1979 年神戸大学教養部助手。1986 年同大学工学部助教授。1994 年同大学工学部教授 (情報知能工学専攻)。1995 年同大学大学院自然科学研究科情報メディア科学専攻専任教授。2001 年京都大学大学院情報学研究科社会情報学専攻教授, 現在に至る。工学博士。主にデータベースとマルチメディア情報システムの研究に従事。人工知能学会, 日本ソフトウェア科学会, IEEE Computer Society, ACM 等各会員。