

Webからの外観情報抽出とその画像特徴量変換に基づく「典型的な画像」の検索

服部 峻[†] 田中 克己[†]

近年, Web 上には文書データだけでなく画像データも大量に存在するようになって来ており, これらの情報を有効に活用できるように, Web 文書検索だけでなく Web 画像検索に対しても多種多様な要求が生まれて来ている. 対象オブジェクトが単に含まれているだけの適合画像ではなく, 対象オブジェクトの外観的特徴を把握し易い「典型的な画像」を検索したい場合も多い. 従来一般的な Web 画像検索エンジンに対して, 対象オブジェクトの名称をキーワード条件として入力することで, そのオブジェクトを確かに含むような適合画像を検索できる精度は高くなって来てはいるが, 「典型的な画像」を精度良く検索することは必ずしもできない. そこで, 我々は, テキストマイニング技術により Web から抽出して来た対象オブジェクトの典型的な外観記述 (本論文では色名のみ) を画像特徴量に射影した上で, 検索対象の各画像に対して, 対象オブジェクトの典型的な画像特徴量を含む度合いを評価することによって, 対象オブジェクトの典型的な画像を Web から検索する手法を提案する.

Search the Web for “Typical Images” based on Conversion from Web-extracted Appearance Descriptions to Image Features

SHUN HATTORI[†] and KATSUMI TANAKA[†]

As the number of Web images as well as Web documents available on the Web has grown exponentially in recent years, Web image searches as well as Web document searches are being pressed more various demands for. We often want to search the Web for not only acceptable images of a target object that simply contain it but also its “typical images” that assuredly represent its typical appearance features. If we submit its name as a keyword-based query to a conventional Web image search engine, we can acquire its acceptable images enough precisely to some extent but we can not always acquire its “typical images”. In this paper, we propose a method to search the Web for typical images of a target object based on the conversion from its Web-extracted appearance descriptions, especially color-names, to image features.

1. はじめに

近年, Web 上には文書データだけでなく画像データも大量に存在するようになって来ており, これらの情報を有効に活用できるように, Web 文書検索だけでなく Web 画像検索に対しても多種多様な要求が生まれて来ている. これまでの Web 画像検索は主に, 対象オブジェクトの名称をユーザが指定すると, そのオブジェクトが単に含まれている適合画像を返すといったものであったが, 対象オブジェクトの外観的特徴を把握し易い「典型的な画像」や, 印象語も指定して, その印象に相応しい特別な対象オブジェクトの画像を検索したい場合もある. もちろん, 多くの場合に対象オブジェクトの画像は大量に存在し, かつ, 文書デー

タとは異なり画像データは総覧し易いため, 対象オブジェクトの適合画像を検索結果として精度良く返してくれさえすれば, その中から多種多様なユーザ要求に合致した画像をユーザ自身の眼で探すことができるかもしれない. しかしながら, 対象オブジェクトを単に含まむというだけの適合画像ではなく, より多種多様なユーザ要求に確かに合致する画像だけを検索結果として返して欲しいには違いない.

Google 画像検索¹⁾ などの一般的な Web 画像検索エンジンは, 基本的に, 画像のファイル名や ALT 属性, 周辺テキストを手掛かりに, テキストに基づく画像検索 (TBIR: Text-Based Image Retrieval) を行う. ユーザがキーワード条件を指定すると, そのキーワード条件に適合する画像を検索結果として返す. また, 画像の投稿者が, 画像に関連すると思う語をタグとして自由に付与することができる写真共有コミュニティサイトの中には, Flickr³⁾ のように, キーワード

[†] 京都市立大学大学院 情報学専攻 社会情報学専攻
Department of Social Informatics, Graduate School of Informatics, Kyoto University

条件に対する関連（適合）性だけでなく、投稿日時に基づく新鮮さや、面白さなどの尺度でランキングすることができるものもある。

これらの Web 画像検索エンジンや写真共有サービスを利用すると、単一のキーワードで条件指定した場合には十分に精度良い検索結果が得られる場合も多くなって来てはいる。しかしながら、単一でも長いフレーズを条件指定した場合には検索結果が 0 件になってしまうり、複数のキーワードを条件指定した場合には全てのキーワードに関連する適合画像ではなくキーワード個々にだけ関連する非適合な画像しか得られなかったりと、多種多様なユーザ要求に対応するには未だ課題も多く残っている⁴⁾。

例えば、「東京駅」の外観的特徴を把握し易い「典型的な」画像を検索しようと、Google イメージ検索に対して、「典型的な東京駅」という長いフレーズを検索クエリとして与えると検索結果は 0 件であった。そこで、やや緩和して、「東京駅 AND “典型的な”」という複数のキーワードを連言で検索クエリとして与えると検索結果の件数は十分に多かったが、赤レンガ造りが顕著な外観である「東京駅」の典型的な画像は容易には見当たらなかった。また、同様に「東京駅」の「ロマンチックな」画像を検索しようとしても上手く行かなかった。

このような画像検索への多種多様なユーザ要求のうち、印象語に基づく画像検索に関しては、予め学習された感性語と画像特徴量との対応関係を利用して感性語による画像フィルタリングを行う K-DIME⁵⁾ など、既に多くの研究がなされている^{6),7)}。一方で、「典型的な画像」の検索に関しては、あまり言及されていない。そこで、本論文で我々は、対象オブジェクトの名称がユーザから与えられた場合に、そのオブジェクトの外観的特徴を把握し易い「典型的な画像」を精度良く検索するための手法について取り組む。

まず「典型的な画像」とはどんな画像であるか、その要件を定義し、その上で、典型的な画像を検索するための手法について議論する。柳井⁸⁾は、類似している画像が多く存在している画像として、典型的な画像という言葉を用いている。そして、対象オブジェクトの名称をキーワード条件として検索された画像集合をクラスタリングし、大クラスタに分類された画像を対象オブジェクトの典型的な画像であると見なしている。対象オブジェクトの名称をキーワード条件として検索された画像集合の大部分が対象オブジェクトの典型的な画像である場合ならば、画像の内容に基づいてクラスタリングすることで典型的な画像を精度良く収

集できるであろう。しかしながら、対象オブジェクトを確かに含む適合画像ではあっても、それが生成（撮影）されるアングルや状況は多種多様である。また、対象オブジェクトの名称をキーワード条件として検索された（上位の）画像集合の大部分が必ずしも対象オブジェクトの適合画像ではない場合も未だ多い。

一方、本論文で我々は、対象オブジェクトと言えどのような外観をしているのかを把握することが容易である画像を「典型的な画像」と呼び、対象オブジェクトの典型的な外観的特徴を画像が確かに備えていることをその要件とする。検索対象の各画像に対して、この要件を満たすか否か、その度合いを評価するためには、対象オブジェクトの典型的な外観的特徴を求める必要がある。対象オブジェクトの名称をキーワード条件として検索された画像集合を解析することによって、典型的な画像特徴量を求める方法が考えられるが、上述のクラスタリング手法と同様に、対象オブジェクトの適合画像の多様性の問題から、各画像において、対象オブジェクトと多種多様な背景や周辺オブジェクトとを正確に切り分ける必要があり容易ではない。そこで、我々は、対象オブジェクトの適合画像が非常に多様であることに比べて、対象オブジェクトの外観に関する言語的表現がやや非多様であることに注目する。まず、テキストマイニング技術により、対象オブジェクトに関する大量の Web 文書を解析し、対象オブジェクトの典型的な外観記述を抽出する。次に、言語的な外観記述を画像特徴量に射影することで、対象オブジェクトの典型的な外観特徴量を得る。以上により、ユーザによって与えられた対象オブジェクトの名称、及び、Web から抽出して来た対象オブジェクトの典型的な外観記述、それを射影して得られた対象オブジェクトの典型的な外観特徴量に基づいて、テキストと内容の両方に基づく画像クエリを生成する。最後に、統合クエリのうちテキストに基づく条件に合致した各画像に対して、対象オブジェクトの典型的な外観特徴量を含む度合いを評価することによってランキングし、対象オブジェクトの典型的な画像を Web から検索する手法を提案する。但し、本論文の実験では、対象オブジェクトの典型的な外観記述として色名のみを Web から抽出し、それを色特徴量に射影している。

本論文の以下の構成を示す。まず、2 章で本論文に関連する研究をいくつか紹介し、提案手法との類似点・相違点についても述べる。次に、3 章で「典型的な画像」を検索するための手法について提案し、その手法の検索精度について 4 章で検証実験を行う。最後に、5 章で本論文をまとめ、今後の課題についても述べる。

2. 関連研究

本論文で提案する「典型的な画像」の検索手法は、対象オブジェクトの名称が言語的に与えられた場合に、そのオブジェクトの典型的な画像を返すシステムであり、ユーザから見れば、テキストに基づく画像検索 (TBIR: Text-Based Image Retrieval) である。しかしながら、内部的には、対象オブジェクトの典型的な外観特徴量を用いてランキングしており、内容に基づく画像検索 (CBIR: Content-Based Image Retrieval) でもある。そこで、本章では、関連研究として、テキストに基づく画像検索、内容に基づく画像検索、及び、これらを組み合わせた画像検索について紹介し、提案システムとの類似点・相違点についても述べる。

2.1 テキストに基づく画像検索

テキストに基づく画像検索では、検索対象である画像自体以外に何らかのテキストデータ (内容を表すキーワード) が付与されている必要がある。市販されていた Corel 社の画像集のように予め整備された画像データベースでは、画像の内容を表すキーワードが、その管理者の人手で付与されている。また、Flickr のような写真共有コミュニティサイトでは、その管理者ではなく、画像の投稿者個々の人手によってキーワードを付与する場合もある。

一方、一般的な Web 画像では、その内容を表すキーワードとして、画像ファイル名や代替テキストである ALT 属性が本来は利用可能ではあるが、画像ファイル名は機械的に付与されているだけであったり、ALT 属性は (適切に) 記述されていなかったりする場合が非常に多い。そこで、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テキストデータを各画像に対して対応付ける様々な手法が提案されている^{9)~11)}。また、類似する画像間でテキストデータを伝播させ、類似したテキストデータを対応付ける手法も提案されている¹²⁾。

我々の提案手法も、対象オブジェクトの名称というキーワード条件によってテキストに基づく画像検索を行うことで、初期の候補画像を収集しており、この点だけを見れば類似している。しかしながら、提案手法では、対象オブジェクトの典型的な外観特徴を備える「典型的な画像」に特化して検索するために、対象オブジェクトの名称から典型的な外観記述を Web から抽出して来ることで、テキストに基づく画像クエリを自動的に質問拡張している点が大きく異なる。

2.2 内容に基づく画像検索

テキストに基づく画像検索では、検索対象である画像自体以外に何らかのテキストデータが付与されている必要があるが、一方、内容に基づく画像検索では、検索対象である画像自体だけから画像特徴量を抽出し、ユーザが指定した画像に類似した画像を検索したり、ユーザが指定した画像特徴量を備える画像を検索したりすることができる^{13)~15)}。また、内容に基づく画像のクラスタリングをユーザへの画像提示に応用した研究も行われている¹⁶⁾。

我々の提案手法も、対象オブジェクトの典型的な外観特徴量という画像特徴量に関する条件によって内容に基づく画像検索を行うことで、各候補画像に対して、対象オブジェクトの典型的な画像としての相応しさの度合いを評価している点は類似している。しかしながら、提案手法では、対象オブジェクトの典型的な外観特徴量をユーザ自身が指定する必要はなく、ユーザが指定した対象オブジェクトの名称に基づいて大量の Web 文書から抽出して来た対象オブジェクトの典型的な外観記述を射影することで自動的に獲得している点が大きく異なる。

2.3 テキストと内容に基づく統合的な画像検索

WebSeek¹⁷⁾ では、最初にキーワードのみによって検索し、その検索結果の中からユーザの望む画像に近いものを選択することによって類似する画像が提示され、インタラクティブに目的の画像に到達することができる。また、Image Rover¹⁸⁾ も基本的には同様のシステムであるが、類似画像を検索する際に、画像の内容に基づく特徴ベクトルだけでなく、画像の周辺テキストに基づく特徴ベクトルも統合して行っている。これらの画像検索システムを用いて、対象オブジェクトの典型的な画像を検索するためには、提案手法とは異なり、対象オブジェクトの名称をキーワード条件として検索した結果の中から、目的の典型的な画像 (に近い画像) をユーザ自身で見つけ出す必要がある。

WebSeer¹⁹⁾ では、言語的なキーワード条件だけでなく、画像特徴量として主要な色などを検索条件として最初から指定することができる。典型的な画像の検索対象であるオブジェクトの名称をキーワード条件として指定して検索した結果を総覧して、そのオブジェクトの典型的な色を把握した上で、対象オブジェクトの名称をキーワード条件として、かつ、その典型的な色を画像特徴量に関する条件として指定して再検索すれば、提案手法と同様の検索結果を得ることができるであろう。しかしながら、ユーザ自身が、対象オブジェクトの典型的な色特徴量を把握する必要があり、

また、検索条件を何度も指定して再検索する必要もある。一方、提案手法では、ユーザは対象オブジェクトの名称をキーワード条件として指定して一度検索するだけで、システムが内部的に、そのオブジェクトの典型的な色特徴量を自動抽出し、目的の典型的な画像を検索結果として返してくれる点が異なる。

K-DIME⁵⁾などの感性語や印象語に基づく画像検索^{6),7)}では、印象語とその画像特徴量との予め対応付けておくことによって、ユーザは対象オブジェクトを表す語と特化したい印象を表す印象語とをキーワード条件として指定するだけで、印象語を画像特徴量に自動的に変換し、対象オブジェクトの名称をテキストに基づく条件として、印象語から変換された画像特徴量を内容に基づく条件として検索することができる。我々の提案手法でも、色名(外観記述)とその色特徴量(外観特徴量)との対応付けを予め行っておくことで、対象オブジェクトの名称をテキストに基づく条件として、その典型的な色特徴量を内容に基づく条件として検索している点は非常に類似している。しかしながら、提案手法では、ユーザ自身が指定する必要がある条件は対象オブジェクトの名称だけであり、これらのシステムではユーザ自身が指定する必要がある印象語に相当する対象オブジェクトの典型的な色名は、大量のWeb文書を解析して自動抽出する点が異なる。

柳井⁸⁾は、Webを画像データベースとして活用するため、テキストに基づく画像検索と内容に基づく画像検索とを組み合わせることで、ユーザが指定したキーワードを表す画像を大量に収集するシステムを提案している。画像のファイル名や代替テキストであるALT属性にキーワードが含まれているという厳密な条件によって選別された少数のA群画像を正例として画像特徴量を抽出し、画像の周辺テキストにキーワードが含まれるという緩い条件によって選別されたより多数のB群画像に対して、内容に基づく画像検索によって取捨選択を行う。ユーザは最初にキーワードを指定するだけで、処理の途中でユーザによる画像や画像特徴量の指定を不要としている点は、我々の提案手法と同様である。もし、厳密な条件によって選別された少数の画像の大部分が対象オブジェクトの典型的な画像として相応しければ、典型的な画像を大量に収集できるかもしれない。しかし、「画像のファイル名や代替テキストであるALT属性に対象オブジェクトの名称が含まれている」ならば「その画像は対象オブジェクトの適合画像である」と仮定することはできたとしても、「その画像は対象オブジェクトに対して典型的な画像である」と仮定することは安易にはできない。

3. 提案手法

本章では、対象オブジェクトの典型的な外観特徴を備えている「典型的な画像」をWebから検索するための手法について提案する。まず、提案手法の基本的な概要について述べた後、提案手法を構成する各機構について詳細に述べて行くことにする。

3.1 概要

対象オブジェクトの名称が言語的に与えられると、そのオブジェクトの「典型的な画像」を検索結果として返す。提案手法の内部の基本的な概要を示すと図1のようになる。まず、ユーザから与えられた対象オブジェクトの名称に基づいて、テキストマイニング技術により、対象オブジェクトに関する大量のWeb文書を解析し、対象オブジェクトの典型的な外観記述(本論文では色名のみ)を抽出する。次に、言語的な外観記述を画像特徴量に射影することで、対象オブジェクトの典型的な外観特徴量を得る。以上より、ユーザによって与えられた対象オブジェクトの名称、及び、Webから抽出して来た対象オブジェクトの典型的な外観記述、それを射影して得られた対象オブジェクトの典型的な外観特徴量に基づいて、テキストと内容の両方に基づく画像クエリを統合的に生成する。最後に、統合クエリのうちテキストに基づく条件に合致した各画像に対して、対象オブジェクトの典型的な外観特徴量を含む度合いを評価することによってランキングし、対象オブジェクトの典型的な画像を検索結果として返す。

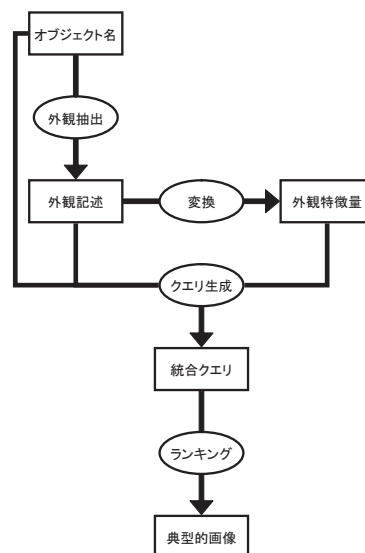


図1 「典型的な画像」の検索手法の概要

Fig. 1 Overview of our search for "typical images"

3.2 オブジェクトの典型的な色名の抽出

我々はこれまで、色名だけでなく、形状や質感など、オブジェクトの様々な外観記述を Web から精度良く抽出する手法について研究して来た^{20),21)}。これらの知見を参考にして、本論文では、ユーザによって与えられた対象オブジェクトの名称に対して、そのオブジェクトの典型的な色名を構文パターンに基づいて抽出する手法を採用し、以下に挙げる構文パターンを用いる。

- 「～色の(オブジェクト名)」
- 「(オブジェクト名)の～色」

構文パターンを用いて大量の Web 文書から抽出された各候補 c_t に対して、次式を用いることで、対象オブジェクト o_t の典型的な色名としての相応しさの度合いを評価する。

$$\text{weight}(c_t, o_t) := (\text{df}(\text{"c}_t\text{色の } o_t\text{"}) + 1) \cdot (\text{df}(\text{"o}_t\text{の } c_t\text{色"})) + 1)$$

但し、 $\text{df}(q)$ は、検索質問 q を文書検索エンジンで処理した検索結果の件数を表している。本論文の実験では、文書検索エンジンとして、Google ウェブ検索²⁾を用いている。

そして、その評価値が他に比べて十分に大きな候補一つを、対象オブジェクトの典型的な色名として採択することにする。もちろん、他に比べて評価値が十分に大きな候補だけでなく、全ての各候補と評価値とのペアをベクトルとして、以降のステップに受け渡して利用することも可能ではあるが、対象オブジェクトの典型的な色名としてノイズである候補の影響をできる限り除去するためである。

3.3 色名から色特徴量への変換

基本的に必要な機能は、言語的な外観記述から、画像的な外観特徴量へ変換することである。外観記述には、色名だけでなく、形状や表面の材質・質感に関する記述なども含まれる。色名から色特徴量への変換は、JIS 慣用色名とそのマンセル値との対応関係が日本工業規格²²⁾で規定されており、他種の外観記述から特徴量への変換と比べて利用し易い。また、色特徴量は対象とする画像のドメインを限定せず、単純な処理で画像の特徴を表現できる汎用的な手法であるため、外観特徴量として第一に採用することにした。もちろん、印象語とその画像特徴量との対応関係に関する研究は既に多数行われており、この知識を利用することで、色名だけでなく、印象語への拡張は難しくはないであろう。本論文では、日本工業規格で規定されている各色名のマンセル値に基づいて、機械的に RGB 色空間や HSV 色空間上に射影することによって、各色名を表現する画像的な色特徴量として用いる。

3.4 典型的画像のクエリ生成

以上より、対象オブジェクトの名称、及び、その典型的な色名が言語的な手掛かりとして、また、典型的な色特徴量が画像的な手掛かりとして得られている。対象オブジェクトの「典型的な画像」を検索するためのクエリとして、以下のような様々な候補が考えられる。

1. 対象オブジェクトの名称と「典型的な」というキーワードをテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ
 2. 対象オブジェクトの名称とその典型的な色名とをテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ
 3. 対象オブジェクトの名称だけをテキストに基づく条件として、対象オブジェクトの典型的な色特徴量を内容に基づく条件として統合した画像クエリ
- しかしながら、いずれのクエリ候補においても、内容に基づく条件として用いているのは対象オブジェクトの典型的な色特徴量である。従って、各クエリ候補のうちのテキストに基づく条件だけを用いて検索した場合の精度が最も良くなるようなクエリ候補を採用すべきである。テキストに基づく画像検索である一般的な Web 画像検索エンジンでは、キーワード条件が複数の語から成ると、特に、外観語や印象語を含むと精度が著しく悪化することが観察されるため、本論文では、対象オブジェクトの「典型的な画像」を検索するためのクエリとして 3 番目のクエリ候補を採用する。

3.5 統合クエリに基づく重み付け

「典型的な画像」を検索するための統合クエリとして、テキストに基づく条件 o_t と内容に基づく条件 c_c が与えられた場合に、検索対象の各画像 i に対して、対象オブジェクトの「典型的な画像」としての相応しさの度合いを評価する必要がある。基本的には、テキストに基づく重み付けと内容に基づく重み付けの評価値とを何らかの合成関数 f で結合することになる。

$$\text{weight}(i, o_t, c_c) := f(\text{weight}(i, o_t), \text{weight}(i, c_c))$$

両者を線形結合し、その結合パラメータを学習するという手法も Jing ら²³⁾によって提案されているが、本論文の提案手法では、最終的に検索結果として返される全ての画像はテキストに基づく条件に合致し、かつ、重みに関しては差を付けず、その上で、内容に基づく重み付けだけで対象オブジェクトの「典型的な画像」としての相応しさの度合いを評価している。つまり、以下の式で定義する。

$$\text{weight}(i, o_t, c_c) := \text{weight}(i, o_t) \cdot \text{weight}(i, c_c)$$

$$\text{weight}(i, o_t) := \begin{cases} 1 & \text{if 画像 } i \text{ が条件 } o_t \text{ を満たす} \\ 0 & \text{otherwise.} \end{cases}$$

但し、画像 i が条件 o_t を満たすとは、画像のファイル名や ALT 属性、周辺テキスト中にキーワード条件 o_t が含まれている場合を表す。本論文の実験では、画像検索エンジンとして Google イメージ検索¹⁾を用いており、キーワード条件 o_t による検索結果に含まれる全ての画像に対して、テキストに基づく重み付けの評価値 $\text{weight}(i, o_t)$ として 1 を等しく与えている。

最後に、検索対象の各画像 i に対して、内容に基づく条件である対象オブジェクトの典型的な色特徴量 c_c による重み付けの評価値を次式で定義する。

$$\text{weight}(i, c_c) := \sum_{\forall c} \text{sim}(c, c_c) \cdot \text{prop}(c, i)$$

但し、 $\text{sim}(c, c_c)$ は、何らかの色空間における色特徴量 c と c_c との間の類似度を表す。本論文の実験では、HSV 色空間における色の類似度²⁴⁾を用いて算出し、0.8 以下である場合には無視 (0.0 に) している。また、 $\text{prop}(c, i)$ とは、検索対象の各画像 i において、色特徴量 c が占有する画素面積の割合を表す。

4. 実験

本章では、Web からの外観情報抽出に基づく「典型的な画像」の検索手法の精度を検証する。まず、提案手法を構成する最初の機構である、オブジェクトの典型的な色名の抽出の精度を検証する。次に、抽出された色名から変換した色特徴量に基づいてリランキングすることによって、キーワード条件として対象オブジェクトの名称を用いて検索した場合に比べて、典型的な画像を検索する精度が改善されることを確認する。

4.1 オブジェクトの典型的な色名の抽出

オブジェクトの名称が与えられた場合に、その典型的な色名を構文パターンに基づいて Web から抽出する手法を、地物名の「東京駅」、植物名の「ひまわり」、動物名の「カメレオン」に適用した結果が表 1 であり、精度良く抽出できている。

一方で、「福田康夫」などの人物名に対しては多くの場合、候補さえ得ることができないため、提案手法によって典型的な画像を検索する精度を改善することはできない。しかしながら、これは、人物の外観において色名が重要でないことを意味しており、色名以外の外観記述や印象語を抽出することができれば、人物名に対しても典型的な画像を検索できる可能性はある。また、「PS3」などの工業製品名においては、実際の色

と消費者が希望するカラーリングとが混ざってしまう場合も多く、非適な色名 (ノイズ) の方が採択されてしまい、提案手法によって典型的な画像を検索する精度を悪化させてしまうこともある。対象ドメインをより広げるためには、対象オブジェクトの名称で検索した結果の画像群を内容解析することによって得られる画像特徴量も組み合わせることによって、オブジェクトの典型的な色名 (色特徴量) の抽出のロバスト性を向上させる必要があると考える。

表 1 オブジェクトの典型的な色名の抽出例
Table 1 Examples of Web-extracted typical color-names

オブジェクト名 o_t	色名 c_t	$\text{weight}(c_t, o_t)$
東京駅	レンガ色	27 (= 3 · 9)
	オレンジ色	4 (= 2 · 2)
	セピア色	2 (= 2 · 1)
ひまわり	オレンジ色	460 (= 92 · 5)
	レモン色	144 (= 36 · 4)
	黄金色	88 (= 22 · 4)
カメレオン	緑色	28 (= 28 · 1)
	オレンジ色	5 (= 5 · 1)
	グレー色	2 (= 2 · 1)

4.2 オブジェクトの典型的な画像の検索

オブジェクトの名称が与えられた場合に、その外観的特徴を把握し易い「典型的な画像」を Web から検索する提案手法を、前節で用いた 3 種類のオブジェクト名に適用した結果について考察する。

対象オブジェクトの名称である「東京駅」だけで画像検索した結果の上位 100 件と、その上位 100 件を「東京駅」の典型的な色名として抽出された「レンガ色」から変換された色特徴量に基づいてリランキングした結果との各々について、上位 k 件までにおける典型的な画像の再現数グラフを描くと図 2 のようになり、大きく改善されていることが確認できる。

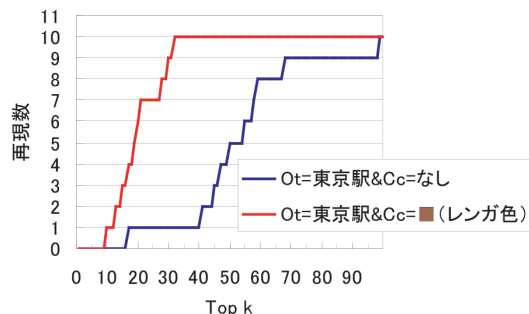


図 2 「東京駅」の典型的な画像の検索結果上位 k 件の再現数
Fig. 2 Top k recall-number of Web-searched typical images for "tokyo station"

3種類のオブジェクトに対して、各々の名称だけで画像検索した結果と、Web抽出した典型的な色名から変換した色特徴量に基づいて重み付けした結果との上位20件の具体的な画像を示すと図3の通りである。「ひまわり」に対しては、典型的な画像の検索精度に大差は見受けられない。また「カメレオン」に対しては、「オレンジ色」など「緑色」以外の非典型的な体色を持つ画像は除去できているが、「緑色」を外観として持つ「カメレオン」以外のオブジェクトである「基板」や背景の「草木」が上位にノイズとして現れている。

5. おわりに

本論文で我々は、対象オブジェクトの典型的な外観特徴量を確かに備えることを「典型的な画像」の要件とし、対象オブジェクトの名称がユーザから与えられた場合に、テキストマイニング技術によりWebから抽出して来た対象オブジェクトの典型的な外観情報を画像特徴量に変換した上で、検索対象の各画像に対して、対象オブジェクトの典型的な外観特徴量を含む度合いを評価することによって、対象オブジェクトが単に含まれているだけの適合画像ではなく、対象オブジェクトの外観の特徴を把握し易い「典型的な画像」をWebから検索する手法を提案した。

人物名などのように、対象オブジェクトの名称に対して、典型的な色名がWebから抽出できない(典型的な色名が元々存在しない)場合には、典型的な色特徴量による重み付けを行うことはできないが、色名以外の外観記述や印象語もWebから抽出することによって「典型的な画像」を精度良く検索できる対象ドメインを広げることができると考えている。

また、対象オブジェクトの名称(キーワード条件)として、「清水寺」などの単一のフレーズだけでなく、「清水寺 AND “夜”」や「清水寺 AND “京都”」などの複数のフレーズにも対応できるように手法を改良し、その検索精度の検証実験を行っていく予定である。

謝辞 本研究の一部は、京都大学グローバルCOEプログラム「知識循環社会のための情報学教育研究拠点」、及び、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(研究代表者:田中克己)、及び、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」における計画研究「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」(研究代表者:田中克己, A01-00-02,

課題番号:18049041)、及び、計画研究「情報爆発に対応する新IT基盤研究支援プラットフォームの構築」(研究代表者:安達淳, Y00-01, 課題番号:18049073)による。ここに記して謝意を表します。



図3 「東京駅」「ひまわり」「カメレオン」の典型的画像の検索結果
Fig. 3 Top 20 examples of Web-searched typical images for “tokyo station”, “sunflower”, and “cameroon”

参 考 文 献

- 1) Google イメージ検索,
<http://images.google.co.jp/> (2007).
- 2) Google ウェブ検索,
<http://www.google.co.jp/> (2007).
- 3) Flickr, <http://www.flickr.com/> (2007).
- 4) Ulug, F., Emirzade, E., and Bitirim, Y.: The Impact of Number of Query Words on Image Search Engines, Proceedings of the Second International Conference on Internet and Web Applications and Services (ICIW'07), pp.50 (2007).
- 5) Inder, R., Bianchi-Berthouze, N., and Kato, T.: K-DIME: A Software Framework for Kansei Filtering of Internet Material, Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC'99), Vol.6, pp.241-246 (1999).
- 6) 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).
- 7) 木本晴夫: 感性語による画像検索とその精度評価, 情報処理学会論文誌, Vol.40, No.3, pp.886-898 (1999).
- 8) 柳井啓司: キーワードと画像特徴を利用した WWW からの画像収集システム, 情報処理学会論文誌 (トランザクション) データベース, Vol.42, No.SIG10(TOD11), pp.79-91 (2001).
- 9) 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, 電子情報通信学会論文誌, Vol.J87-D1, No.2, pp.145-153 (2004).
- 10) 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, 電子情報通信学会技術研究報告, DE2005-136, pp.19-24 (2005).
- 11) 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).
- 12) 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).
- 13) Gudivada, V. N. and Raghavan, V. V.: Content-Based Image Retrieval-Systems, IEEE Computer, Vol.28, No.9, pp.18-22 (1995).
- 14) 串間和彦, 赤間浩樹, 紺谷精一, 山室雅司: 色や形状等の表層的特徴量にもとづく画像内容検索技術, 情報処理学会論文誌 (トランザクション) データベース, Vol.40, No.SIG03(TOD1), pp.171-184 (1999).
- 15) Veltkamp, R. C. and Tanase M.: Content-Based Image Retrieval Systems: A Survey, Technical Report, UU-CS-2000-34 (2000).
- 16) 串間和彦, 佐藤路恵, 赤間浩樹, 山室雅司: 大量画像の閲覧を目的とする階層的分類支援機能 - 画像目録の実装と評価 -, 情報処理学会論文誌 (トランザクション) データベース, Vol.41, No.SIG01(TOD5), pp.54-63 (2000).
- 17) Smith, J. R. and Chang, S.-F.: Visually Searching the Web for Content, IEEE Multimedia, Vol.4, No.3, pp.12-20 (1997).
- 18) Sclaroff, S., La Cascia, M., Sethi, S., and Taycher, L.: Unifying Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web, Computer Vision and Image Understanding, Vol.75, No.1/2, pp.86-98 (1999).
- 19) Frankel, C., Swain, M. J., and Athitsos, V.: WebSeer: An Image Search Engine for the World Wide Web, Technical Report, TR-96-14, University of Chicago (1996).
- 20) 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, 第 18 回データ工学ワークショップ (DEWS'07) 論文集, L4-6 (2007).
- 21) 服部 峻, 手塚太郎, 田中克己: 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, 情報処理学会論文誌 (トランザクション) データベース, Vol.48, No.SIG11(TOD34), pp.69-82 (2007).
- 22) 日本工業規格: 物体色の色名, JIS Z 8102:2001 (2001).
- 23) Jing, F., Li, M., Zhang, H.-J., and Zhang, B.: A Unified Framework for Image Retrieval Using Keyword and Visual Features, IEEE Transactions on Image Processing, Vol.14, No.7, pp.979-989 (2005).
- 24) Smith, J. R. and Chang, S.-F.: VisualSEEK: A Fully Automated Content-Based Image Query System, Proceedings of the 4th ACM International Conference on Multimedia (ACM Multimedia'96), pp.87-98 (1996).