

Webからの外観情報抽出とその画像特徴量変換に基づく特異画像の検索

服部 峻^{†,††} 田中 克己[†]

画像検索の研究分野では一般に、名称が与えられた対象オブジェクトに対して、ノイズ画像をできる限り排除して適合画像だけを返すことが目標である。本論文で我々は、さらなる次の段階として、適合画像から、典型画像や特異画像を選別することを目指す。単純に画像内容に基づくクラスタリングを行ったとしても、ノイズ画像ではなく特異画像だけを含むようなクラスタを見付けることは困難である。そこで我々は、大量の Web 文書中から抽出して来た対象オブジェクトの特異な外観に関する記述(テキスト)を活用して画像検索クエリを構成し、特異画像を精度良く検索する手法を提案する。

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

SHUN HATTORI^{†,††} and KATSUMI TANAKA[†]

Most researches on image retrieval have aimed at clearing away noisy images and allowing users to retrieve only approvable images for a target object specified by its name. In this paper, we try to discriminate between typical images and “Peculiar Images” in the approvable images, as a next step of image retrieval. It is difficult to find clusters which consist of not noisy but peculiar images only by clustering based on image content features. We propose a method to retrieve peculiar images for a target object by extracting its peculiar appearance descriptions (text) from the Web and converting them 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。次の段階として、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することが可能なシステムも提案されている^{1)~3)}。

一方で、適合画像の中から、対象オブジェクトの外観的特徴を把握し易く誰もが納得するような「典型画像⁴⁾」や、このような定番とは違う一風変わった意外な「特異画像」を選別して欲しいという要求もある。

例えば「ヒマワリ」と言えば「黄色いヒマワリ」を思い浮かべる人が多いが、「赤いヒマワリ」や「黒いヒマワリ」も存在する。我々は前者を「典型画像」と呼び、後者を「特異画像」と呼んでいる。対象オブジェクトについて網羅的に調べたい場合に、典型的な特徴に関する情報だけでなく、特異な特徴に関する情報も重要である。また、画像認識による一般物体識別などの基本的知識としても有用であると考えられる。

Google 画像検索⁵⁾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、Flickr⁷⁾のように、キーワード条件に対する関連(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

[†] 京都大学大学院 情報学研究所 社会情報学専攻
Department of Social Informatics, Graduate School of Informatics, Kyoto University

^{††} 日本学術振興会 特別研究員 DC2
Research Fellow (DC2) of the Japan Society for the Promotion of Science

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件になってしまったり、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ような、全体としては不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⁸⁾。

我々は、対象オブジェクトの名称が与えられたとき、そのオブジェクトを単に含むというだけの多種多様な適合画像をそのまま返すのではなく、典型画像や特異画像だけを選別することを目指す。本論文では特に、Web から特異画像を精度良く検索する手法について検討する。対象オブジェクトの特異な外観の特徴を画像が確かに備えていることを「特異画像」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の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、特異な画像特徴量を求める方法が考えられるが、特異ではあるが確かに対象オブジェクトが有する場合がある画像特徴量とノイズとを正確に切り分けることは容易ではない。

そこで我々は、対象オブジェクトに対する画像検索の結果は依然として多くのノイズ画像も含んでいるのに比べ、対象オブジェクトに対する文書検索の結果中の外観に関する言語的記述が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特異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特異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特異な外観記述、それを射影して得られた対象オブジェクトの特異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特異な外観特徴量を含む度合いを評価することでランキングし、対象オブジェクトの特異画像を Web から検索する手法を提案する。

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先行研究を紹介し、提案手法との類似点・相違点について述べる。次に、3章で Web から「特異画像」を検索するための手法について提案し、提案手法の検索精度について検証実験も併せて行う。最後に、4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「特異画像」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、そのオブジェクトにとって特異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、対象オブジェクトにとって特異ではあるが有する場合もある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^{10)~12)}。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¹³⁾。

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特異な外観特徴を備える「特異画像」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特異な外観記述を 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^{14)~16)}。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¹⁷⁾。

我々の提案手法も、対象オブジェクトの特異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特異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特異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特異な外観記述を射影することで自動的に獲得している点異なる。

2.3 テキストと内容に基づく統合的な画像検索

WebSeek¹⁸⁾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 Rover¹⁹⁾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特異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、目的の特異画像(に近い画像)をユーザ自身で見付け出す必要があるが、典型画像とは違い、特異画像が検索結果の上位に含まれている可能性はその性質上高くない。

WebSeer²⁰⁾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可能である。従って、対象オブジェクトの名称をキーワード条件として、かつ、その特異な色を画像特徴量に関する条件として指定して検索すれば、提案手法と同様の検索結果を得ることができる。しかしながら、前述と同様に、特異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の上位を総覧しても、そのオブジェクトの典型的な色は把握できるかもしれないが、

必ずしも特異な色を把握できるとは限らない。ユーザ自身が、対象オブジェクトの特異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特異な色特徴量を自動抽出し、目的の特異画像を検索結果として返してくれる点異なる。

K-DIME¹⁾などの感性語や印象語に基づく画像検索^{2),3)}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、外観記述(色名)とその外観特徴量(色特徴量)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外観記述は、大量の Web 文書を解析して自動抽出する点異なる。

柳井⁹⁾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(クラスタリング)すること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、我々の案手法と同様である。しかしながら、「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」という条件を満たす A 群画像は、その大部分は対象オブジェクトの単なる適合画像であり、典型画像や特異画像が含まれる割合は非常に小さいため、典型画像や特異画像だけを選別することはできない。

3. 提案手法

本章では、対象オブジェクトの「特異画像」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。

3.1 概要

対象オブジェクトの名称が言語的に入力されると、そのオブジェクトの「特異画像」を検索結果として出力する。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特異な外観記述（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特異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特異な外観記述、それを射影して得られた対象オブジェクトの特異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特異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特異画像を検索結果として返す。

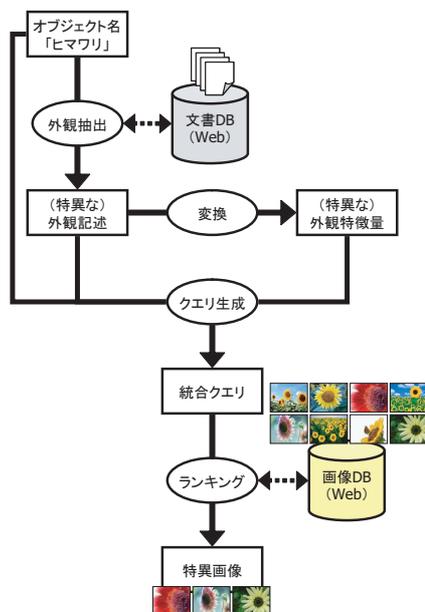


図 1 「特異画像」の検索手法の概要

Fig. 1 Overview of our search for “Peculiar Images”

3.2 オブジェクトの特異な色名の抽出

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 から抽出する手法について研究して来た^{21)~24)}。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特異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。

- 「(色名)の/い(オブジェクト名)」
- 「(オブジェクト名)の/は(色名)」

特異画像の検索に必要なのは、対象オブジェクトの典型的な色名ではなく特異な色名であり、不適合な色名だけでなく典型的な色名も排除できる評価尺度が必要である。そこで我々は、典型的な色名は後者の構文パターンによる出現頻度が大きく、一方、特異な色名は前者の構文パターンによる出現頻度が大きいという経験的観測に基づき、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、対象オブジェクト o_t の特異な色名としての相応しさの度合いを以下のように定義する。

$$\text{weight}(c_t, o_t) := \frac{\text{df}("c_t\text{の}/い o_t")}{\text{df}("o_t\text{の}/は c_t) + 1}$$

但し、 $\text{df}(q)$ は、検索質問 q を文書検索エンジンとしての Google ウェブ検索⁶⁾ で処理した検索結果の件数を表している。

対象オブジェクトの名称として「ヒマワリ」および「東京タワー」を与えると、特異な色名の抽出結果は表 1 および表 2 のようになる。「ヒマワリ」は典型的な品種は「黄色」であり、「赤」「白」「茶色」である品種も存在する。実験の結果を見ると、「黄色」「黄金色」「レモン色」などの典型的な色名は下位に、「赤」「白」「茶色」などの特異な色名は上位に来ていることが分かる。しかしながら、「虹色」という不適合な色名も上位に来てしまっており、単なる比による重み付けでは限界がある。後者の構文パターンによる出現頻度が非常に小さい色名は排除するなどの処理も必要であると考える。一方「東京タワー」は典型的な色は「赤」「オレンジ色」であり、特別なイベントの場合に「ピンク色」「青」「緑」などにライトアップされることがある。実験の結果を見ると、「赤」「オレンジ色」といった典型的な色名は下位に、「緑」「青」「紫色」などの特異な色名は上位に来ていることが分かる。しかしながら、特異な色名であると予測した「ピンク色」は上位に来ていない。これは「ピンク色」の「東京タワー」は他の「青」や「緑」に比べると既に一般的になりつつあるためであると考えられる。

表 1 オブジェクト名「ヒマワリ」の特異な色名の抽出例
Table 1 Peculiar Color-Names for $o_t = \text{"sunflower"}$

色名 c_t	df($"c_t * o_t"$)	df($"o_t * c_t"$)	weight(c_t, o_t)
赤	497 (2nd)	22 (2nd)	21.6 (1st)
虹色	19	0	19.0 (2nd)
黒	160 (3rd)	9 (5th)	16.0 (3rd)
白	94 (5th)	6	13.4 (4th)
ココア色	24	1	12.0 (5th)
茶色	133 (4th)	12 (4th)	10.2
黒色	10	0	10.0
赤褐色	8	0	8.0
紫色	8	0	8.0
チョコレート色	55	6	7.9
レモン色	31	4	6.2
月色	6	0	6.0
白色	5	0	5.0
こげ茶色	5	0	5.0
オレンジ色	24	5	4.8
黄金色	20	4	4.0
夕日色	4	0	4.0
えんじ色	4	0	4.0
青	61	16 (3rd)	3.6
黄色	524 (1st)	391 (1st)	1.3

表 2 オブジェクト名「東京タワー」の特異な色名の抽出例
Table 2 Peculiar Color-Names for $o_t = \text{"tokyo tower"}$

色名 c_t	df($"c_t * o_t"$)	df($"o_t * c_t"$)	weight(c_t, o_t)
緑	117 (3rd)	13	8.4 (1st)
青	278 (1st)	37 (3rd)	7.3 (2nd)
紫色	23	4	4.6 (3rd)
セピア色	7	1	3.5 (4th)
白	108 (4th)	34 (5th)	3.1 (5th)
夏色	30	9	3.0
クリスマス色	3	0	3.0
黒	19	6	2.7
ピンク色	93 (5th)	36 (4th)	2.5
黄色	14	5	2.3
エメラルド色	2	0	2.0
オレンジ色	84	66 (2nd)	1.3
冬色	7	6	1.0
ミドリ色	3	2	1.0
紅白二色	1	0	1.0
ムラサキ色	1	0	1.0
雪景色	5	7	0.6
赤	160 (2nd)	285 (1st)	0.6
灰色	2	3	0.5
茶色	1	3	0.3

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。しかし、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²⁵⁾で規定されており、他種の外観記述から外観特徴量への変換と比べて非常に利用し易い。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。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を得ている。前述の特異な色名の抽出結果中の「虹色」「月色」「夏色」「冬色」などは日本工業規格で規定されていないため、画像的な外観特徴量に射影することはできない。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、日本工業規格で規定されている色名だけでなく、他の印象語へ拡張することは難しくないと考えられる。対象オブジェクトの典型画像と特異画像を選別するのに有効な外観特徴量が必ずしも色であるとは限らず、形やテクスチャである場合も考えられるため、色以外の外観特徴量への対応も今後必要である。

3.4 特異画像のクエリ生成

以上より、対象オブジェクトの名称、及び、その特異な色名が言語的な手掛かりとして、また、特異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「特異画像」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「特異な」「変わった」「珍しい」などのキーワードとをテキストに基づく条件とする画像クエリ
2. 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特異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とその特異な色名とをテキストに基づく条件とする画像クエリ
4. 対象オブジェクトの名称とその特異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特異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

特異画像を検索したいユーザの意図を最も反映しているのは第 1 の画像クエリであるが、目的の画像を含む Web 文書の作成者が「特異な」「変わった」といったキーワードを記述しているケースはあまり見当たらないため上手く検索できない。残りの中では、第 4 の画像クエリが指定する条件が最も強いいため、適合率が最も良くなると期待される。一方、第 2 の画像クエリを用いると、テキストに基づく条件が他よりも弱いいため、再現率は最も良くなると予測される。

3.5 統合クエリに基づく重み付け

「特異画像」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 (オブジェクト名および特異な色名) と画像内容に基づく条件 c_c (特異な色特徴量) が与えられた場合に、画像データベース中の各画像 i に対して、対象オブジェクトの「特異画像」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画像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²⁶⁾ 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、画像内容に基づく重み付けだけで対象オブジェクトの「特異画像」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⁵⁾ 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、画像内容に基づく条件である対象オブジェクトの特異な色特徴量 c_c による重み付けの評価値を次式で定義する。

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

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²⁷⁾ 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。

対象オブジェクトの名称として「ヒマワリ」および「東京タワー」を与え、3.2 節により特異な色名を Web から抽出し、3.3 節により色特徴量に変換したものを併せて構成した統合クエリを用いて特異画像を検索すると、図 2 および図 3 のようになる。

「ヒマワリ」というオブジェクト名そのまま Google イメージ検索した結果の上位 100 件を見ると、特異な色としては「赤」と「えんじ色」が 1 件ずつ確認でき、他の約 9 割は一般的な「黄色」で、残り約 1 割はノイズ画像であった。「ヒマワリ」でキーワード検索した結果の上位 100 件中にもそもそも含まれていないため、「黒」「茶色」「白」の色特徴量でリランキングしても上位に現れることはない。オブジェクト名だけでなく特異な色名も加えてキーワード検索した場合が最も良い結果であるように思われる。さらに特異な色名から変換した色特徴量でリランキングすると、より相応しい画像を上位に引き上げられている一方で、全く不適当な画像も現れてしまっている。

	1st	2nd	3rd	4th	5th
$o_t = \text{ヒマワリ}$ $c_c = \text{なし}$					
$o_t = \text{ヒマワリ}$ $c_c = \text{(赤)}$	 (0.67978)	 (0.15051)	 (0.09130)	 (0.05273)	 (0.01654)
$o_t = \text{赤いヒマワリ}$ $c_c = \text{なし}$					
$o_t = \text{赤いヒマワリ}$ $c_c = \text{(赤)}$	 (0.67978)	 (0.46467)	 (0.43346)	 (0.40223)	 (0.27214)
$o_t = \text{ヒマワリ}$ $c_c = \text{(黒)}$	 (0.70461)	 (0.48403)	 (0.40225)	 (0.34994)	 (0.30787)
$o_t = \text{黒いヒマワリ}$ $c_c = \text{なし}$					
$o_t = \text{黒いヒマワリ}$ $c_c = \text{(黒)}$	 (0.84271)	 (0.56097)	 (0.55464)	 (0.52331)	 (0.50347)
$o_t = \text{ヒマワリ}$ $c_c = \text{(茶色)}$	 (0.57311)	 (0.53778)	 (0.44692)	 (0.40929)	 (0.36935)
$o_t = \text{茶色いヒマワリ}$ $c_c = \text{なし}$					
$o_t = \text{茶色いヒマワリ}$ $c_c = \text{(茶色)}$	 (0.68377)	 (0.68092)	 (0.59003)	 (0.57918)	 (0.55231)
$o_t = \text{ヒマワリ}$ $c_c = \text{(白)}$	 (0.78630)	 (0.67564)	 (0.66968)	 (0.63002)	 (0.62101)
$o_t = \text{白いヒマワリ}$ $c_c = \text{なし}$					
$o_t = \text{白いヒマワリ}$ $c_c = \text{(白)}$	 (0.73676)	 (0.71736)	 (0.70039)	 (0.66011)	 (0.35824)

図 2 「ヒマワリ」の特異画像の検索結果
Fig. 2 Web-searched Peculiar Images for “sunflower”

「東京タワー」というオブジェクト名そのまま Google イメージ検索した結果の上位 100 件を見ると、特異な色としては「ピンク色」が 3 件、「青」が 1 件確認でき、他の約 7 割は一般的な「赤」で、残り約 3 割はノイズ画像であった。「東京タワー」でキーワード検索した結果の上位 100 件中にも含まれていないため、「緑」や「紫色」の色特徴量でランキングしても上位に現れることはない。オブジェクト名だけでなく特異な色名も加えてキーワード検索した場合が最もバランスが良いように思われる。さらに特異な色特徴量でランキングすると、「ピンク色」のようにノイズ画像で占められてしまう場合もあり問題が残る。

	1st	2nd	3rd	4th	5th
$o_t =$ 東京タワー $c_c =$ なし					
$o_t =$ 東京タワー $c_c =$ (緑)	 (0.07814)	 (0.03779)	 (0.01790)	 (0.01713)	 (0.01166)
$o_t =$ 緑の東京タワー $c_c =$ なし					
$o_t =$ 緑の東京タワー $c_c =$ (緑)	 (0.23871)	 (0.22918)	 (0.15307)	 (0.15206)	 (0.13857)
$o_t =$ 東京タワー $c_c =$ (青)	 (0.28349)	 (0.10937)	 (0.06056)	 (0.03498)	 (0.02039)
$o_t =$ 青い東京タワー $c_c =$ なし					
$o_t =$ 青い東京タワー $c_c =$ (青)	 (0.33134)	 (0.21984)	 (0.21313)	 (0.20359)	 (0.13177)
$o_t =$ 東京タワー $c_c =$ (紫色)	 (0.43249)	 (0.25317)	 (0.14402)	 (0.10626)	 (0.03427)
$o_t =$ 紫色の東京タワー $c_c =$ なし					
$o_t =$ 紫色の東京タワー $c_c =$ (紫色)	 (0.09933)	 (0.09933)	 (0.09933)	 (0.07188)	 (0.07188)
$o_t =$ 東京タワー $c_c =$ (ピンク色)	 (0.78601)	 (0.74187)	 (0.72034)	 (0.67454)	 (0.57592)
$o_t =$ ピンク色の東京タワー $c_c =$ なし					
$o_t =$ ピンク色の東京タワー $c_c =$ (ピンク色)	 (0.89570)	 (0.86740)	 (0.86133)	 (0.81559)	 (0.80578)

図 3 「東京タワー」の特異画像の検索結果
Fig. 3 Web-searched Peculiar Images for “tokyo tower”

4. まとめと今後の課題

ノイズ画像を排除し適合画像だけを得るという画像検索の研究分野のさらなる次の段階として、適合画像から特異画像（あるいは典型画像）だけを選別することを我々は目指した。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特異な外観記述（色名）を画像特徴量（色特徴量）に変換した上で、画像データベース中の各画像に対して、テキストに基づく条件として対象オブジェクトの名称だけでなく特異な外観記述も追加し、画像内容に基づく条件として対象オブジェクトの特異な外観特徴量を含む度合いを評価しランキングすることによって、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定番とは違う「特異画像」をより精度良く Web から検索できる手法を提案した。

「クリスマス色」「冬色」のように日本工業規格で規定されていない非慣用色名は色特徴量へ直接的に変換することはできないが、「クリスマス色」と言えば「赤」「緑」「白」のペアというように、非慣用色名から慣用色名への想起関係を推定できれば、色特徴量へ間接的に変換することができる。このように複合された色が暗示されている場合、どのように色特徴量（ベクトル）を構成するべきか、今後議論を深めて行く。

人物名などのように、対象オブジェクトの名称に対して、特異な色名が Web から抽出できない（元々存在しない）場合もあるが、色名以外の形状、表面の材質といった外観記述や印象語も Web から抽出することで、対象ドメインを広げることができると考える。

また、従来の画像内容に基づくクラスタリング技術を用いた特異画像の検索との比較実験を行った上で、提案手法との併用によって精度向上を図ることができないかという可能性についても検討して行く。

謝辞 本研究は、科学研究費補助金特別研究員奨励費「モバイル・ユビキタス環境における空間情報アクセスに関する研究」（研究代表者：服部峻，課題番号：1955301，平成 19～20 年度），及び、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」（研究代表者：田中克己，平成 19～23 年度），及び、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）の助成を受けたものである。ここに記して謝意を表す。

参 考 文 献

- 1) 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).
- 2) 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).
- 3) 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).
- 4) 服部 峻, 田中克己: 色名抽出と色特徴量変換に基づく典型的画像の Web 検索, 日本データベース学会 Letters, Vol.6, No.4, pp.9-12 (2008).
- 5) Google イメージ検索, <http://images.google.co.jp/> (2008).
- 6) Google ウェブ検索, <http://www.google.co.jp/> (2008).
- 7) Flickr, <http://www.flickr.com/> (2008).
- 8) Uluc, 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).
- 9) 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).
- 10) 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).
- 11) 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).
- 12) 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).
- 13) 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).
- 14) Gudivada, V. N. and Raghavan, V. V.: Content-Based Image Retrieval-Systems, IEEE Computer, Vol.28, No.9, pp.18-22 (1995).
- 15) 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).
- 16) Veltkamp, R. C. and Tanase M.: Content-Based Image Retrieval Systems: A Survey, Technical Report, UU-CS-2000-34 (2000).
- 17) 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).
- 18) Smith, J. R. and Chang, S.-F.: Visually Searching the Web for Content, IEEE Multimedia, Vol.4, No.3, pp.12-20 (1997).
- 19) 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).
- 20) 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).
- 21) 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).
- 22) 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).
- 23) Hattori, S., Tezuka, T, and Tanaka, K.: Mining the Web for Appearance Description, Proceedings of the 18th International Conference on Database and Expert Systems Applications (DEXA'07), LNCS Vol.4653, pp.790-800 (2007).
- 24) 服部 峻, 田中克己: コンテキストに依存する外観情報の Web からの抽出, 電子情報通信学会 第 19 回データ工学ワークショップ (DEWS'08) 論文集, A2-1 (2008).
- 25) 日本工業規格: 物体色の色名, JIS Z 8102:2001 (2001).
- 26) 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).
- 27) 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).